Title
Compositional Strategies in Spectral Spatialization

Author
Jaroszewicz, Martin

Publication Date
2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Compositional Strategies in Spectral Spatialization

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

Doctor of Philosophy

in

Music

by

Martin Jaroszewicz

March 2015

Dissertation Committee:

Dr. Paulo C. Chagas, Chairperson
Dr. Miller Puckette
Dr. Ian Dicke

The Dissertation of Martin Jaroszewicz is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

I am deeply grateful to my supervisor, Dr. Paulo C. Chagas, for his teaching and the guidance over the past four years. Without his incredible knowledge, this thesis would not have been possible.

I would also like to thank Wouter Snoei from the Game of Life Foundation and Andreas Piesiewicz from TU Berlin for their time and support.

Immense gratitude to Hans-Ulrich Werner for giving me the opportunity to be in residence at Hochschule Offenburg and connecting me with all the people and resources pertaining to this thesis.

Finally, I am very grateful to my wife for all the sacrifices and infinite support.

To my beloved wife.

ABSTRACT OF THE DISSERTATION

Compositional Strategies in Spectral Spatialization

by

Martin Jaroszewicz

Doctor of Philosophy, Graduate Program in Music
University of California, Riverside, March 2015
Dr. Paulo C. Chagas, Chairperson

The use of space as a structural part of a composition is a complex issue that involves rethinking the idea of the sound object – as described by Pierre Schaeffer in his *Treatise on Musical Objects* (1966) – and involves considering movement as a means to manipulate spectrum in the frequency domain. In addition, the contemporary composer – especially when writing acousmatic music – needs to consider the materiality of the studio where music is created. These interrelated factors and others, such as the means of reproduction, the acoustics of the venue, the choice of loudspeakers and eventually the software that executes the algorithms for complex calculations of movement and other parameters of sound spatialization, are part of the *techné* involved in the creative process.[1] Thus, the choice of sound material and the aesthetics of movement should be considered to be elements intrinsically related to the technical means of the performance as spatialization produces an impact on the timbral footprint of sound.

One of the motivations behind the research conducted at EARS was to answer Agostino Di Scipio's question: "How can I design the tools that are necessary to realize my own idea of composition ?" (Scipio 1995a). The other one, was to realize some ideas inspired by Borges' story *The Library of Babel*, which became the starting point behind

[1]For a discussion about the significance of *techné* see (Manning 2006).

the *Laberintos* series of electroacoustic études, each one related to one aspect of working with space and using the electronic music techniques described in this thesis.

There is a discussion involving the meaning of working with space from the point of view of a composer of contemporary electroacoustic music. This dissertation deals with strategies for working with sounds objects whose trajectories are predetermined by the composer during the composition process, as opposed to using space as a resonant body to "enhance" music. A new definition of the sound object that considers its "materiality" and its relationship with space is discussed. It views sound–objects from a different perspective than Pierre Schaeffer, considering them as volumetric objects that occupy a space like any object or "thing" – a view in tandem with speculative realism (Harman 2011). These sound objects are real; they can travel in space and acquire a tangible property.

I describe a unique approach for dealing with sound trajectories in spatialized music using 3D modeling software. Although, none of these applications were designed for music or sound and lack any synthesis capabilities – besides the creation of basic waveforms – it is possible to extract the data containing the coordinates of a 3D virtual space in the form of a simple text file. In addition, Blender[2] features a game engine that can be used to send Open Sound Control (OSC) messages in real time using the Python scripting language. As an alternative, I have created several externals written in C for the Pure Data language for the creation of sound trajectories using ideas taken from parametric design. These objects offer the possibility to work using algorithmic and stochastic approaches to spatialization simplifying the compositional process. The technical part of this dissertation deals with the tools I have developed at EARS and how can be implemented using the techniques exemplified with my own work.

[2]Open source 3D graphics and animation software.

Another aspect of this thesis is the description of the systems available for spatialization in music and how they differ from commercial systems. At EARS, I built – and worked with – tools that were specifically designed for composition and spatialization thus contributing to the *techné* and the aesthetics that influenced my musical ideas. The result of the research conducted at EARS and the experienced gained working with Wave Field Synthesis (WFS) systems – the Game of Life in Netherlands and the systems at the Technische Universitt Berlin – not only generated new apparatus but suggested a new rethinking of composition in space.

From a technical point of view, among the programming languages for music and sound, I found that Pure Data,[3] Supercollider,[4] Faust, and Chuck are suitable open source tools for the composer working with electronic music as they are portable and available for most platforms including Mac, Linux and Windows.[5] In addition, Pure Data, with its visual approach, is a great tool for quickly sketching musical ideas and for demonstrating the theory and technique of electronic music without writing lines of code.

[3]Pure Data PD is an object based graphical environment for sound synthesis developed by Miller Puckette, professor at the University of California San Diego. In PD, like in Supercollider, it is possible to create custom synthesizers, effects, musical patterns, and sonic and musical machines by connecting on-screen patch cords, but most importantly, PD is a great tool for sound research, analysis and re-synthesis.

[4]Supercollider is an object-oriented programming language designed specifically for describing sound processes in real time. SuperCollider was written by James McCartney and is now an open source (GPL) project maintained and developed by various people. It is used by musicians, scientists, and artists working with sound.

[5]Software portability is the usability of the software in different platforms. For example, a Pure Data patch can be run in a Macintosh, Linux or Windows computer without any modification of the code or need to recompile.

Contents

List of Figures

Part I

Theoretical Aspects

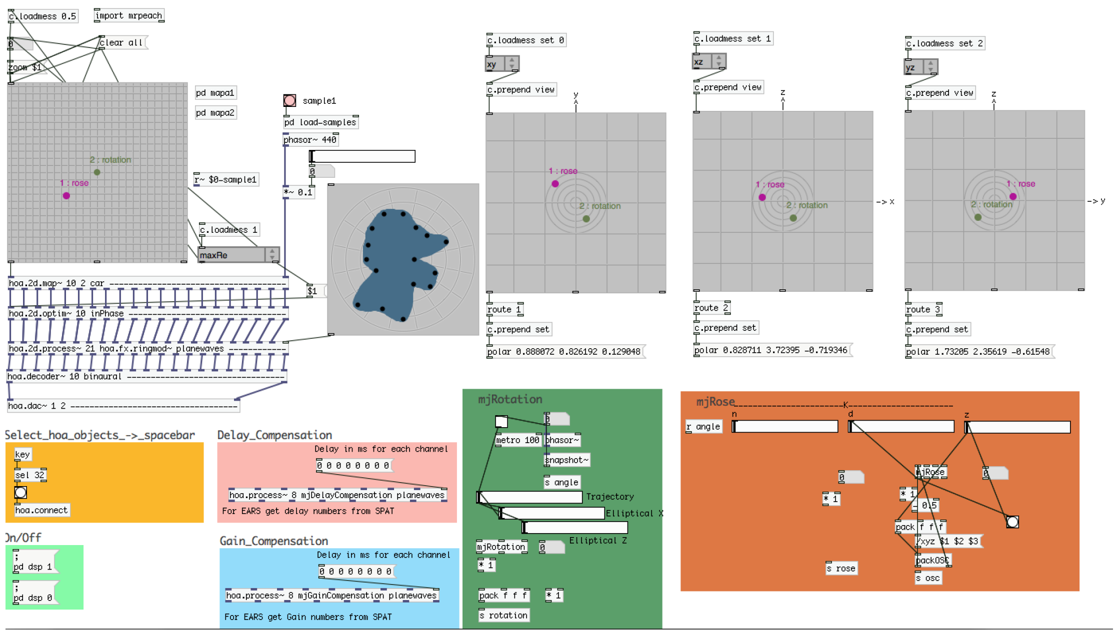


Figure 0.1: Spatialization with the mj library for Pure Data and HOA

Chapter 1

Introduction

El universo (que otros llaman la
biblioteca) se compone de un número
indefinido, y tal vez infinito, de
galerías hexagonales...

La biblioteca de Babel

JORGE LUIS BORGES

1.1 Space and Spatialization

The use of technology for creating listening spaces to **enhance** music goes back to ancient Greece and Rome where engineers built special theaters for music and speech. Aristotle – 4th century B.C.– knew that the law of reflections could be applied to sound as well and refers to the analogy between light and sound in his explanation of echoes and resonances, demonstrating how timbre changes due to distance and absorption.[1]

[1]See Aristotle's *Problemata*. Book XI (Aristotle 2001).

A different use of space concerned with music is found in antiphonal practice – Medieval Gregorian Chant and Responsorial Psalmody – performed during religious services in the 4th century.[2] The difference between antiphonal and responsorial is the manner of performance. Antiphonal features alternating choirs while responsorial, a choir responding to a soloist. There were also secular chants based on non-biblical texts, prose and poetry.

During the 16th. and 17th. centuries composers took advantage of the spatial properties of cathedrals, such as the St. Mark's Basilica of Venice, to "color" their music by exploring timbre possibilities thus expanding their musical ideas. This practice can be seen in the works of composers from the Venetian region who often wrote for double chorus. The style grew with the polychoral music of Giovanni Gabrieli (ca. 1553–1612) who explored the sonic capabilities of mixing up to seven choruses combining different timbres: low or high voices mixed with diverse instruments.

The practice in San Marco in Venice raises questions about music and space. Did the composer selected the performance site and how musicians should be placed? How these decisions affected the work of the composer? Any orchestral work requires a specific layout, for example, in a standard orchestra we find the violin section to the left of the conductor – and the audience – and double basses to the right. If halls did not have reverberation the listener would have a completely different experience as he would be perceiving great spatial separation.[3] In addition, higher frequencies – from violins – and lower frequencies from the bass section would definitely be perceived as a sort of spectral separation.

[2]In this manner of performing, a soloist or reader sings the first part of the psalm, typically two halves of a verse, and a choir or congregation responds singing a second part.

[3]Note that the layout of the orchestra has nothing to do with the range of the instruments but instead with volume and visibility. For an example see: Di Grazia, Donna.19th-Century Music Vol. 22, No. 2 (Autumn, 1998), pp. 190-209.

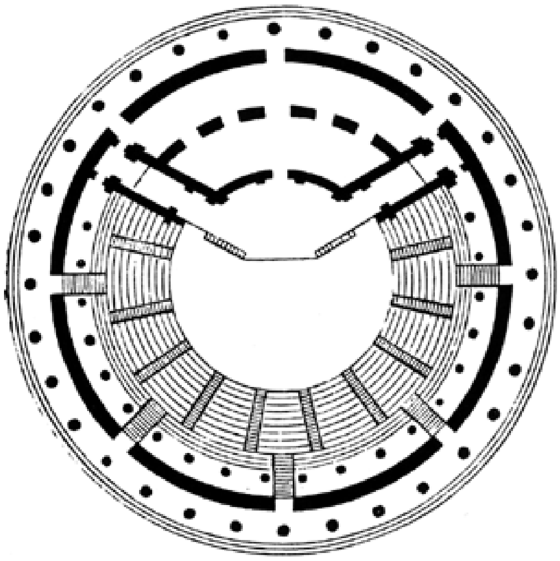


Figure 1.1: [4]

The law of reflections – produced by sound bouncing from a surface – was well known by the Greeks. The work of the Roman Marcus Vitruvius Pollio *De Architectura libri decem* (c. 30 BC.) provides an interesting account of early thought in room acoustics and the use of different materials for acoustic conditioning applied to music and speech. Vitruvius' work explains how spaces with no reverberations – such as the open air – are more suitable for speech intelligibility and reverberant spaces are adequate for the performance of music (Pollio 2008, 149). For music performance, there were vessels under the seats of the theaters that in theory would resonate with the frequencies produced by the instruments:

> The Greeks knew that the clear but dry acoustics of the open-air *theatron*(in Greek "a place to see") were excellent for speaking but inadequate for music. Therefore, the *odeion* (Greek "a place to hear") was developed for song performances during musical competitions. A hall of limited size with excellent acoustics erected on a rectangular ground plan, the *odeion* featured steeply raking rows of seats and a flat horizontal ceiling. The walls and the ceiling with structured surfaces produced well supported *reflexiones consonantes*, the open window absorption (Baumann 2011).

[4]Odeion. McGraw-Hill Dictionary of Architecture and Construction. Copyright ©2003 by McGraw-Hill Companies, Inc.

In many cases, large reverberant spaces such as the domes of basilicas of the roman tradition can produce reverberations times of eight seconds of duration[5] due to the angles of incidence and non-absorbent materials of their walls.

Throughout the 17th and 18th centuries many architectural techniques were used to reinforce or "enhance" sound including the use of water under the orchestra [6] and/or the audience. Some of these ideas had unpredictable or no results at all and some did work. Only during the mid 19th century, the acoustics of the halls were greatly improved with efficient architecture. According to Dorothea Baumann in her book *Music and Space: A systematic and historical investigation into the impact of architectural acoustics on performance practice followed by a study of Handel's Messiah*: "this period deserves to be called the "golden age" in the history of opera houses and concert halls" (Baumann 2011, 41).

In the mid 19th, Berlioz writes about the importance of room reverberations and the effect on the listener: "And one must vibrate with the instruments and voices and because of them in order to have a genuine musical sensation" (Berlioz 1994).

There is evidence of a clear separation between acoustic spaces designed for speech and those designed for music. Moreover, the inclination to use reverberant spaces for music shows that music was meant to be perceived without being localized. In other words, music needed to "fill" the space as one amalgamated form. On the other hand, theater pieces took advantage of dry spaces in order for the audience to be able to localize the voice of the actor/s on stage, for example, to easily follow a plot.[7]

[5]Reverberation – Reverb – time is calculated by measuring the time a reverberated sound takes to fall to 1/1000 of its peak amplitude. This is referred as the RT60 value because it represents a 60 decibels (dB) drop from the intensity of the first reflection.

[6]The word *orchestra* originally meant the semi-circular space in front of a stage in a Greek theater.

[7]For a brief history of the development of architectural acoustics see Shankland (Shankland 1972) and for an example of a scientific analysis of the acoustics in a Greek theater see Declercq (Declercq 2007).

The technology for sound reproduction – including the invention of the gramophone in the late 19th century and the use of filters in the 20th century – made it possible to recreate acoustic spaces with artificial reverberation. The technique is often used in pop and rock music as recording studios are "dry", that is, they are designed to not produce any reverberations . In "classical" music, reverb may be not necessary as most performances are recorded in "live" halls as opposed to studios. Moreover, the musician often needs to adjust his playing to the acoustics of the concert halls and playing in a dry space requires different techniques.

The more realistic the artificial space is reproduced, the more important is for the conductor to adjust the balance of the orchestra. Then, it is the task of the architects, engineers, conductors and performers to adjust those parameters to create and aesthetics such as in the music of the cathedral of San Marco. In addition, it is possible to virtually recreate any concert hall or space by using convolution reverb, a type of reverb that is obtained by "capturing" the space with microphones and convolving the result with the music.

A different aspect of the use of space is the concept of *spatialization* which is, as I see it, part of the compositional process; it can be incorporated on the "score" and in the material supporting the work. That is, as a parameter that can be "written" in the music by the composer whereas as part of the score or by the use of techniques and practices – algorithmic or "composed" –Moreover, in this thesis, I use the term spatialization to refer to the techniques for moving acousmatic sounds from a virtual source in a real space using computer algorithms and parametric implementations that can be written in the score thus becoming a structural part of the work.[8] Spatialization

[8]As are dynamics, pitch collections, tempo markings, etc.

as a parameter of composition has strong links to serialism and *Elektronische Musik* is especially suitable for working with timbre composition using parametric techniques:

> The methods of timbre composition concentrated on parametric and combinatorial thinking link *elektronische Musik* to the polyphony of the Middle Ages and the Renaissance and to the experimental music of the beginning of the 20th century, particularly the aesthetics of atonalism, twelve-tone music, and serialism (Chagas 2014).

Although spatialization deals with localization and movement and not with the use of natural or artificial reverb, it is possible to combine these two different ideas as composer Pierre Boulez did in the 20th Century with his work *Répons* .[9]

While spatialization does not concern with the acoustics of the space, the manner of playing or the conductor's decisions, the concepts are interconnected. Paulo C. Chagas, in his book *Unsayable Music* clearly explains the connection between space, reverb, technology and spatialization:

> The acoustics properties of the performance hall play an important role in determining how the listener perceives the music. Reverberation for example, can significantly impact both the performance and the listening experiences. It affects the clarity and liveliness of the sound and the localization of the sound source. In the case of electroacoustic music, the listening experience also depends on the quality of the sound system: the audio equipment, the characteristics of the speakers, and their distribution throughout the performance space, Sound spatialization, a crucial aspect of electroacoustic music performance, is in fact a virtual construction of sound space; the virtual space can be generated by the composition itself[..](Chagas 2014).

It is in the 20th. century when composers could and began to explore other aspects of the use of space other than the one imposed by the architecture of the hall. This was possible by the invention of the tape recorder and the loudspeaker. In the last fifty years, with the growth of the movie industry, new developments in audio spatialization and multichannel systems such as "surround", that offer an immersive experience to

[9]Pierre Boulez recreates the massive sonorous climaxes of antiphonal chant with his work *Répons* for ensemble, soloists and spatialized electroacoustic sounds. Boulez's work, in addition to a virtual sound space created by the circular motion of acousmatic and electronic sounds around the listener, presents a composer who is aware of the spectral changes imposed by the acoustics of the performance space. Moreover, a closer look at the orchestration and electronic transformations in *Répons*, reveals how the composer methodically works with the timbral possibilities of spatialization (Jaroszewicz 2013).

the movie aficionado, had been commercialized and standardized. Unfortunately, commercial surround techniques have little to offer to contemporary music; their surround panning techniques do not work in tandem with the aesthetics of contemporary music adding little to nothing to the development of spatial techniques and aesthetics.

A very special case in the history of electronic music and the development of spatial techniques is Karlheinz Stockhausen's *Gesang der Jünglinge* [Song of the Youths] (1955-56), which was the first piece to use multiple loudspeakers creating the psychoacoustic illusion of circular motion around the listener: an aesthetics that has influenced composers of electroacoustic music since the 1950s (Decoupret 1998). But the most salient aspect of the work is undoubtedly the linkage of movement and spectral transformations:

> The revolutionary accomplishment of *Gesang der Jünglinge* is to provide the first example of the integration of the composition of timbres with spatial articulation (Chagas 2014, 111).

Being the first electroacoustic work to integrate space and timbre as part of the composition process, *Gesang der Jünglinge* opened the path to new possibilities for the composer of electronic music. Moreover, it changed the way electroacoustic music should be performed including its technological requirements. Spatialization needs "dry" spaces as opposed to large reverberant halls – such as those of churches – detaching music from the tradition of the concert hall or the church .[10]

To further clarify the concept of spatialization, we can infer that there are three different aspects of music and space according to several parameters; the last being the topic of this dissertation:

1. **Architectural Space**: The use of technology and architecture to "enhance" music by modifying physical spaces.

[10]For a discussion on the role of the concert hall see Boulez' *Orientations: collected writings*(Boulez 1986, 448).

Table 1.1: Different interpretations of the use of space in music.

	Architectural Space	Performance Space	Spatialization
Example	Greek Theater	Concert Hall	Wave Field Synthesis
Function	To enhance or suppress reverberations	To enhance music	To add movement to sound
Controlled by	Shape and Materials	Conductor / Performers	Composer
Sound Directionality	Omni/Direct	Omni	Direct
Listener's attention	Non-Focused/Focused	Non-Focused	Focused
Structural in Composition	No	No	Yes

2. **Performance Space**: The use of orchestration as a "special effect" with the intention to modify timbre.

3. C**omposed Space or Spatialization**: The use of space as a container for creating movement as a structural part of a composition.[11]

In order to incorporate spatialization to my own work, I had to approach the task of composing music from a different perspective. Working with space adds another level of complexity and I thought spatialization had to be "composed" with the music. The analogy is the use of dynamics which some composers incorporate last as if it was a parameter of less hierarchically value. But spatialization, as any aspect of composition can be approached from different angles and using different techniques. One of them is algorithmic composition using stochastic methods for moving sounds in space. Another one is the use of simple trajectories such as circular and elliptical paths. Lastly, the

[11]Table 1.1 shows how correlations between music and space result in different ideas that are clearly defined by their function and the intention of the composer to alter – or not – timbre. It also takes into consideration the point of view of the listener who has the ability to focus on different aspects of music such as timbre, harmony, melody, etc. When spatialization is part of the structure of a work, the intention of the composer is to draw the attention of the audience to movement and the spectral changes linked to the movement of the sound object as it travels through predefined – composed or algorithmically generated – trajectories.

"composition" – written in the score or somehow notated – of paths that work in tandem with the music.

Moreover, there are two aspects of composing spatialization: the creation of trajectories and the spectral transformations sound undergoes as it moves. I found that a good composition practice was to work with these two elements at the same time. For example, as a sound moves from point A to point B, it undergoes a timbral transformation using cross-synthesis techniques in the spectral domain, which I found to be the best way to link movement and spectra. At EARS,[12] I had the opportunity to developed different strategies by using its 8 channel movable system.

At EARS, I experimented with different speaker layouts beginning with a circle – for Ambisonics – and finally placing speakers accordingly to the shape of the room and adjusting for distance using delay and gain compensation techniques. The software tools at EARS allowed me to try different cross-synthesis techniques. I was also able to experiment with different approaches for the creation of trajectories. There were several steps in the creation of a system that can be used to incorporate spatialization as a structural part of composition and the following items were considered:

- Speaker layout: circular / square / room-adapted
- Spatialization techniques: SPAT / Ambisonics / WFS
- Cross-synthesis techniques : Mixing, source-filter / convolution / cross-modulation / square-root convolution / cross-product
- Trajectories: Algorithmic / Parametric / Stochastic

The following chapters introduce the work I developed at EARS, including different aspects of spatialization techniques, new technological developments, a reflec-

[12]The Experimental Acoustic Research Studio at the University of California Riverside is a satellite facility founded by Paulo C. Chagas and expanded into a studio for spatialization in collaboration with the author.

tion on the aesthetics of sound trajectories, and a brief historical background for the purpose of connecting ideas about timbre and space, always from the point of view of composition. The reader will be able to learn about recent practices in sound spatialization and get acquainted with the tools and technologies available to the composer of contemporary music.

1.1.1 Definitions

Some definitions are provided to clarify and introduce the reader to spatialization terminology:

1. **Monoaural**: Listening through one ear.
2. **Binaural**: Listening through two ears. This is the natural way of listening to sounds. The term is also used for “binaural” recordings , which are made with an anatomically correct dummy head that captures sounds binaurally with two microphones attached to its “ears”. There are also binaural microphones that are designed as earplugs – like headphones – and are able to record binaurally using a human head instead. The reproduction of binaural recordings is through headphones as this type of recording technique produces excellent separation of the sounds.
3. **Stereophonic**: The word “stereo” comes from the Greeks meaning “solid” referring to three-dimensional. Originally stereo recordings were intended to be reproduced through two loudspeakers creating a psychoacoustic three-dimensional image which in fact is not “stereo” in the strict meaning of the word. Stereo systems can have more than two channels.
4. **Multichannel**: A system than can reproduce more than one channel. At the present time, commercial systems offer up to 7.2 channels, that is seven surround channels and two subwoofers. There are 10.2 and 22.2 systems not commercially available yet. The definition includes all the systems listed below.
5. **Surround**: I use this term referring to multichannel commercial systems such as 5.1 and 7.1 surround. These include an array of speakers surrounding the listener. The most common layout is the following[13]: Front Left(L), Front Center(C),Front

[13]Clockwise.

Right(R), Surround Right(SR) and Surround Left(SR). the sub-woofers can be placed anywhere as low frequencies cannot be spatialized.

6. **High Order Ambisonics** A technique for creating a sound field requiring a full-sphere array where speakers are distributed in a dome surrounding the listener.
7. **Wave Field Synthesis** A technique for spatialization of sounds in a 2D field with an array of speakers completely surrounding the listener.

It is convenient for composers to write acousmatic music for commercial surround formats as systems are widely available and most Digital Audio Workstations, such as ProTools and Logic, offer plugins or built-in tools for bouncing surround files and for creating surround automation. Moreover, if writing for multimedia, it is possible to work with pictures and sounds in the same software environment. Multimedia works – such as my Laberintos V – can be "performed" in a variety of venues including a movie theater setting; they can be encapsulated in a single file and stream or store in a DVD. Although surround is convenient, it is limited to positioning and moving sounds from speaker to speaker using standard panning techniques. The focus of attention is not the sound itself and movement usually adds little to nothing to the composition. For this reason, alternative systems have been developed exclusively for music with emphasis on contemporary practices and acousmatic music.

1.2 Immersion and Fields

The idea of immersion underlies the development of sound spatialization .[14] Multichannel systems position the listener inside a space surrounded by loudspeakers which in addition act as physical boundaries. Surround techniques, which were originally

[14]Immersion – without interactivity – creates a sense of depth integrating the listener and the space. Sounds can be perceived as moving from point to point and in the case of surround systems, they can only be perceived as moving from speaker to speaker, and only left to right and the opposite.

developed for the movie theater, are limited to the movement of sounds which is the product of a psychoacoustic effect. Moreover, there is no correlation between movement and timbre.

The use of multichannel systems by composers of contemporary music such as Karlheinz Stockhausen and Paulo C. Chagas influenced the development of immersive sonic environments where the listener can experience timbral changes and movement interrelated as parameters of composition. This idea is further explored by the use of devices such as the gmebaphone (1973), which is capable of doing spectral spatialization by splitting frequencies into multiple channels and using cross-over techniques.[15] It was a combination of instruments specially designed for the diffusion of electroacoustic works. These instruments were built under the premise of considering diffusion of electroacoustic music inseparable from the process of composition. Christian Clozier argues that music is a complex ensemble of parts and not single units that move in time:

> Music is the only system of symbolic communication and exchange that unfolds along the irreversible line of time, it is no mere information, nor is it the simple flow of sequence of single sound units, but rather a complex ensemble of parts related to each other not only in the present but as fragments of processing time and future time (Clozier 2001).

Spatialization should be part of that complex ensemble and not be regarded as a technological "trick".

The gmebaphone is an immersive system that creates a sound field surrounding the listener. Modern surround systems are designed to be "equal" in all directions and are used for the reproduction of works where sound trajectories are not relevant and not suitable for music reproduction.[16] On the other hand, "geometric systems" are

[15] Later versions of the gmebaphone moved away from the cross-over filtering of signals into a software based instrument with fader motion capture and scene automation with crossfade. It was renamed cybernephone in 1997 (Emmerson 2007).

[16] In the 1990, the European standards organization ITU (International Telecommunications Union) conducted research for optimal speaker placement. The recommendations are part of the document

considered when space is a structural part of the composition as in, for example, Stockhausen's *Oktophonie* and Paulo C. Chagas' *Migration*, the latter – for twelve channels – "was the first electroacoustic piece that systematically explored the circular configuration of 12 loudspeakers in the WDR studio" (Chagas 2008). Chagas also designed the spatialization system for Stockhausen's *Oktophonie.* The intention to create a more accurate immersive experience is evident by the increasing number of loudspeakers:

- 4 channels. *Kontakte.* 1958. (Stockhausen)
- 8 channels. *Oktophonie.* 1991. (Stockhausen)
- 12 channels. *Migration.* 1995. (Chagas)
- 12 channels. *Projektion.* 2000. (Chagas)

Systems such as WFS and ambisonics allow for a more flexible approach to spatialization without the need to work or compose with the loudspeaker in mind. One of the reasons behind the creation of the mj library was to detach the loudspeaker from the sound–object allowing the composer to focus on the composition and the spectral changes instead.

ITU-R BS.775-1. The research was conducted using classical music material and left the surround channels for "effects" or "ambience". The document was created well before the development of new surround methods which give equal importance to all five main speakers which may not be ideal for music applications. See: The Recording Academy's Producers and Engineers Wing. *Recommendations for Surround Sound Production.* The National Academy of Recording Arts & Sciences, Inc. 2004.

1.3 From Stereo to Wave Field Synthesis

First attempts to reproduce sound were made on a single monophonic channel with the invention of the phonograph in 1877 by Thomas Edison. The ability to record sounds – and music – paved the way to the development of *musique concrète* which began with recordings of natural sounds. Although, it "engaged the persistent myth of listening to the sound of the world as a source of music creativity" (Chagas 2014, 107), the first phonographs had already inspired musicians to think of it as a musical instrument.[17]

The earliest known use of a stereo system was Clement Ader's experiment with distant telephone listening in 1881. He demonstrated his invention at the Paris Electrical Exhibition where the attendees where able to listen to music broadcast live via telephone from the Grand Opera at Paris. The system consisted of two microphones on the stage at the opera and stereo pairs of headphones connected in series at the exhibition. Using the same principle as the stereoscope, which enables a person, by means of the superposed visual impressions of the two eyes, to see the stereoscopic images with their natural relief, Adler adds "relief" to sounds.

> Suppose two microphonic transmitters, placed on the stage at T and T. Let one of them be connected by wires with the receiver R, and the other with the receiver R, these receivers being applied to the ears of the auditor, and suppose an actor to stand first at A and then at A. In the former position, as he is nearer to the transmitter T than to T, his singing will be heard loudest with the left ear; but when he is at A he is nearer to T than to T, and the right ear will receive the strongest impression. Thus as he goes from A to A the definite sensation which the auditor will receive will be that of a diminution of loudness in the left and an increase of loudness in the right ear, which is the same as he sensation which we experience when a person who is speaking walks from our left to our right (Nature, 1881).

[17] Philip G. Hubert, Jr. "What the Phonograph will do for Music and Music-Lovers," Scribner's Monthly 46 (1893:May/Oct.), pp.152-4."Open Letters" section. Can be accessed at: http://www.phonozoic.net/n0128.htm.

Aders experiment is a demonstration of stereo sound spatialization as a psychoacoustic phenomenon or "sensation" and of course, he was aware of the changes in volume on each headphone: "diminution of loudness in the left and an increase of loudness in the right ear". Today we use the terms *constant linear panning* or *linear crossfade*. Usually we can change the position of a mono-source signal by feeding each channel with the same signal and adjusting the relative amplitudes of the channels. Although an interesting way to create the illusion of moving sounds from left to right and the opposite, linear crossfading does not preserve the loudness and creates a void in the middle of the stereo front.

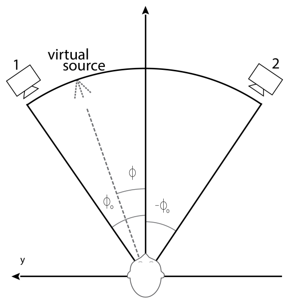


Figure 1.2: The simplest stereo panorama produces a void in the center

In the United States Harvey Fletcher, a physicist most famous for his contributions to the creation of the equal-loudness contours, invented a "Wall of Sound" or "Acoustic Curtain" at the Bell Labs in 1932. For him, the ideal stereophonic system could work with up to 80 microphones and speakers in a hall. This is perhaps the closest idea to the Wave Field Synthesis systems we have today. His research included the creation of a system capable of recording and reproducing the entire frequency range and to be able to synchronize sound with motion pictures because silent film was the prevalent format for movies. Sound was introduced by using Bell Lab systems but there was no interest from the film industry to incorporate spatialization to movies. Fletcher wanted to add a spatial dimension to movies by using binaural techniques which he initially called "auditory perspective" and later became known as stereophonic sound or stereo (Fletcher 1992, 182).

Fletcher demonstrated a binaural system in 1932 at the Bell Labs World Fair in Chicago using a dummy head with microphones attached to each ear. As someone walking around the head would speak, the public listening through headphones connected to the microphones would get the impression of being surrounded by someone talking around them. Fletcher later contacted different orchestras and conductors to see if they would be interested in a spatialization system for music. No one seemed interested except Leopold Stokowski, at the moment conductor of the Philadelphia Orchestra and who later collaborated with Disney in several projects including the first stereophonic film *Fantasia*.

> It was about 1931 when I first met Stokowski and we made tests of stereophonic sound down at the Academy of Music in Philadelphia where the Philadelphia Orchestra held its concerts. There was a spare room in the Academy building which was large enough to house the orchestra so that we could have them play on the stage and listen to it up in this large room. In this way we tried experiments until we felt we had developed a stereophonic system. Originally, the theory of this system was that it should have an infinite number of loud speakers at one end and the same number of microphones at the other end. However, we found that in stage productions, three microphones, three transmitting lines and three loud speakers were sufficient. I'll not go into the details of the development work that was necessary to produce this. There are several printed papers on it. However, we did make nine loud speakers expecting that we might have to use three across and three up and down. However, we found that most of the action was horizontal and consequently, three loud speakers were sufficient . . . so realistic was the effect that to the audience the act seemed to be taking place on the stage before them. Not only were the sounds of sawing, hammering, and talking faith- fully reproduced correctly, but the auditory perspective enabled the listeners to place each sound in its proper position, and to follow the movements of the actors by their footsteps and voices [...](Fletcher 1992).

Fletcher understood the importance of spatialization for commercial purposes such as in film, telematic music and theater. He focused on two aspects: fidelity and localization. During the 1980's, fifty years after Fletcher's experiments, A.J. Berkhout and Diemer de Vries from the Delft University of Technology in the Netherlands developed the first Wave Field Synthesis system that was based on seismic data analysis – elastic wave field extrapolation – and research in acoustics (Berkhout 1993).The WFS technique changes the psychoacoustic stereo paradigm to a model where sound sources are physically created in an acoustic field. The speaker is not the source of the sound but part of a system that creates the acoustic field. WFS is costly to implement and there are few places where it can be experienced. As of today, the most important centers for research in WFS and music are at the Technische Universität Berlin, Germany and Institute of Sonology in Den Haag, Netherlands.

1.4 Acoustic spatialization

Although spatial music is usually associated with acousmatic or electroacoustic music of the 20th Century, the practice of placing a sound in space with the purpose of using the space as a musical parameter was explored by composers during the early Christian tradition in the form of antiphonal choral music of *cori spezzati*.[18]

The interest in spatial writing was gradually lost[19] but by the end of the 18th Century composers occasionally placed group of musicians away from the orchestra only to create a "dramatic" effect. Wolfgang A. Mozart (1756–1791) wrote for four identical chamber ensembles of strings and horns in his *Notturno* in D Major, K 286 (1776).[20] In

[18]The principle is also used in large polychoral compositions – for two or more choirs–. The term *cori spezzati* or split choirs was used to describe polychoral singing in Venice in the later 16th century. qPolychoral singing. Encyclopedia Britannica, 2013.

[19]Many factors affected the way music was performed including performance practices, performances in secular and private venues, the invention of the pianoforte, etc.

[20]A nocturne is a piece intended for an evening party.

Figure 1.3: The dome of San Marco basilica in Venice

the score, Mozart notated the four entrances of the orchestras using the word "echo"[21] – *Erstes Echo, Zweites Echo and Drittes Echo* – corresponding to the entrances of the second, third and fourth orchestras respectively.

In addition to expressively indicate the "echoes" in the score, the music does evoke the acoustical phenomenon of echo; during the first ten measures of the piece, the repeated entrances of the opening phrase become shorter, spanning from sixteen eight notes to six eight notes at the last entrance. There is no polyphonic writing between the ensembles until the last bars of the third movement – *menuetto* – before the "Trio" where the four orchestras, at one point, sound simultaneously. Another interesting aspect of the piece is that the trio is played by only one of the ensembles. In Mozart's *Notturno* the ensembles always play sequentially following the same order throughout the piece. Mozart did not specify where the orchestras should be placed with respect to

[21]From a scientific point of view, echoes are delays between approximately 30 milliseconds and about a second in duration that usually do not change the "shape" of melodies or phrases (Roads 2004). For more on time shifts and delays, see the corresponding chapter in Miller Puckette's *Theory and Techniques of Electronic Music.*

Figure 1.4: Beginning of Mozart's *Notturno* for 4 orchestras

the audience. Many different combinations are possible and, by analyzing the way the piece was written, some are more relevant than others.

Although the details of the first performance of *Notturno* in D are not known, and there is no indications of a layout in the score, four orchestras allow for the following combinations, including clockwise and counterclockwise rotation:

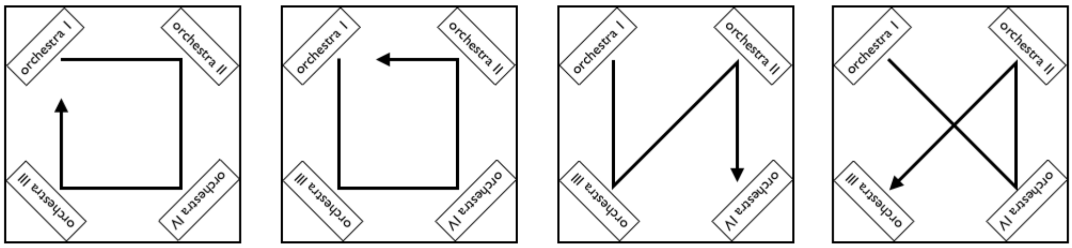


Figure 1.5: Some of the possible spatialization layouts in Mozart's *Notturno* for 4 orchestras

Hector Berlioz's *Requiem* (1837) is another example of the use of space in music.[22] The piece was written for four antiphonal brass groups placed on each corner of the stage and a massive orchestra of singers, woodwinds, horns, strings and percussion. Composed for Saint-Louis des Invalides in Paris, Berlioz exploited the acoustic characteristics of its gigantic dome.

The four brass choirs enter at unison at the *Tuba mirum* section of the mass playing an Eb chord then splitting into grandiose antiphons, polyphonic passages and unisons; the same orchestration idea is repeated throughout the work.

Although Berlioz was aware of the use of space in music and its implications, the composer creates an impact on the listener as the majestic choir of brass first appears in the *Dies Irae – Tuba mirum* – section of the mass[23] (Berlioz 1994).

[22]Hector Berlioz. French. (1803–1869).

[23]"Day of Wrath". The *Dies Irae* was used in the Roman liturgy as the sequence for the Requiem Mass for centuries being the most famous Christian doomsday chant, painting an apocalyptic vision of the world's dissolution into ashes.(Slonimsky 1998, 39).

Figure 1.6: Berlioz's *Rex Tremendae*

Another example is the work of American composer Charles Ives (1874–1954), who experimented with polyrhythms and polytonality incorporating a spatial element in many of his works. His *Symphony No. 4* often requires two conductors as the composer divides the gigantic orchestra into smaller groups and incorporates additional off-stage ensembles. Ives was strongly influenced by the sounds of simultaneous marching bands in the town of Danbury, Connecticut, where he was born. This influence can be clearly seen in his innovate way of writing for overlapping keys and meters and the combination of European, American and church music.

Ives is an early postmodernist; as he juxtaposes multiple elements and experiments with microtonality, polytonality, polyrhythm an aleatoric elements such as in his work *The Unanswered Question*, his intention is to separate different layers of independent material. The work features three different layers of disparate material that is tempo and key independent. In addition, these layers are further separated by the spatial placement of the ensembles: the string orchestra off-stage, the woodwind ensemble on stage, and the solo trumpet off-stage at a distant position.

> The string quarter or the string orchestra (con sordini), if possible, should be "off -stage", or away from the trumpet and flutes. The trumpet should use a mute unless playing in very large room, or with a larger string orchestra (Ives 1908).

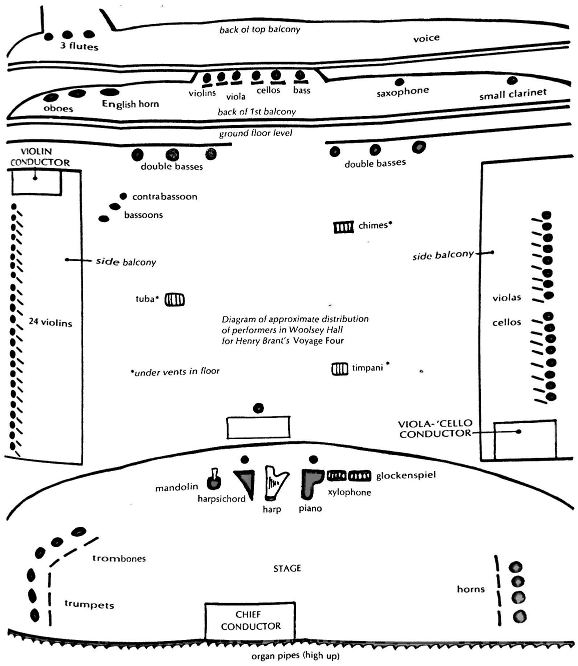
Ives uses a spatial layout to create a sense of atmospheric perspective and distance, this is written in the foreword of *The Unanswered Question* where he asks the trumpet to use a mute if too close to the string ensemble or playing in a small room. The composer continued exploring the use of space with his *Symphony No. 4* where he takes the technique to a an unprecedented level applying the same ideas to a very large orchestra.

From the point of view of Auditory Scene Analysis, the spatialization in Ives' music aids the listener in the segregation of audio streams preventing the computing of dissonances. In other words, dissonant musical lines played by instruments placed in different parts of the hall are not perceived as one dissonant passage but as different melodies. There is no intention by Ives to create new timbres. The juxtaposed materials remain unrelated throughout the work whose layers are to be discretely perceived by the listener.[24]

Henry Brant (1913–2008) continued and expanded the experiments initiated by Ives. A pioneer in the development of spatial music, he was greatly overshadowed by European contemporary music composers, especially Karlheinz Stockhausen. Most of Brant's works are spatial. His first one being *Antiphony* (1953) where the orchestra is divided into five groups and each section play in different keys. The result is a level of complexity created by dissonant relationships. In 1967, Brant summarizes his ideas on space and music in relation to timbre (Brant 1967):

- Spatial separation changes harmonic content of a texture – if the music contains several layers within the same range of an octave, by-product unisons sound confusing if they come from the same location, for example: performers that are close together sharing a stage. If the same layers occur from widely separate locations in the hall, the unisons are no longer perceived as unisons but as distinct tone qualities created by diffusion and distance. In other words, same pitches on same

[24] To break a complex sound apart our brain uses simultaneous strategies, the first one being *segregation*, itself composed of different strategies that attempt to identify individuals objects or events with a composite stream of information. Another primitive feature is the regulatory of harmonic structure or *harmonicity*. We understand the harmonics series as one event corresponding to one timbre and not as separates sounds. For example, a note from a violin is one identifiable sound and not a series of simultaneous streams with different pitches. For more information on the topic see the chapter on sound cognition in Andy Farnell's book *Designing Sound* (Farnell 2010).



[h]

Figure 1.7: Distribution plan for *Voyage Four* by Henry Brant

instruments that are widely separated in a hall will have different spectral content that is added by the room's acoustics.

- Distance hinders coordination – simultaneities in the music, e.g. unisons –, are difficult to coordinate between two spatial separated groups
- Spatial separation allows for expanded complexity – if writing the music within a close pitch range e.g octave –,creating no collisions or crossing of textures it is possible to have clarity within a group that is close together. On the other hand, if groups are widely separated, the writing can be freed of restrictions hence adding textural complexity.
- Placement of performers cannot be optional – the layout of groups of performers in the hall must be carefully planned as part of the work in order to create a controlled musical result. Brant originally wrote his ideas about the use of space in the process of composition in 1953 publishing them a year after.

For many years composers focused on three elements of music: melody, harmony and rhythm. By considering spatialization a structural element of a work, Brant suggests that composers need to be aware of texture, spectral content and dynamics. All these elements add complexity to the music and should not be arbitrarily considered.

1.5 Electroacoustic Spatialization

It is during the early 20th century when composers started to think differently about music and space. With the invention of the telephone and the gramophone, composers – and the general public – began to develop different ideas about time and space including timbre. For them, the earliest telephone systems were capable of transporting a distorted but intelligible human voice – timbre – miles away from one space to another.

German composer Karlheinz Stockhausen may have been aware of Brant's writings since they predate Stockhausen's "spatial music manifesto" *Musik im Raum* (1959) written after his acousmatic piece *Gesang der Jünglinge* (1955–56) and Gruppen (1955–57) for 3 orchestras.

> In my *Gesang der Jünglinge* I attempted to form the direction and movement of sound in space, and to make them accessible as a new dimension in musical experience. The work was composed for 5 groups of loudspeakers, which should be placed around the listeners in the hall. From which side, by how many loudspeakers at a time, whether rotation to left or right, whether motionless or moving, how sounds and sound groups are projected into space: all this is decisive for a comprehension of the work (Stockhausen 1958).

Stockhausen opposes Brant's ideas about the use of space. In his "manifesto", the German composer dismisses Gabrieli's use of partial techniques. He argues that Berlioz and Mahler used spatialization on their works only to add drama accusing them of being "too theatrical". He argues that direction is the only spatial feature of sound worthy of compositional attentions because it could be serialized. The composer claims that perception of distance and timbral changes due the acoustics of a hall are parameters that cannot be serialized hence cannot be a structural part of a composition. In contrast, when a sound is placed in a circle around the audience, parameters such as localization and speed of movement can be operated with exact proportions. It is possible to create "the scale of localities corresponding to the scales of pitch, duration, timbre and loudness" (Miller 2009).

Stockhausen's music and ideas were novel at the time. Today, a sound orbiting a space in circular motion can be parameterized as a function of time obtaining phase for a given radius in a two dimensional polar coordinate system, allowing the sound to be placed at a precise location, for example, using Wave Field Synthesis or High Order Ambisonics. The perception of distance and timbral changes are a consequence of the room's acoustics and cannot be "composed" but highly controlled using artificial reverb algorithms and software tools.

1.6 Gesang der Jünglinge

Gesang der Jünglinge is Stockhausen's fourth acousmatic piece, being, Etude (1952, *musique concrète*), Studie I (1953, *musik elektronische*) and Studie II (1954, *musik elektronische*) his previous works for tape. His first three pieces are monophonic works; Studie I and II are completely serialized compositions. In contrast, *Gesang der Jünglinge* was written for five channels and incorporates the idea of fusion between sounds, in this case: the sound of a human voice and electronic generated sounds. This is the first time a composer applies the idea of morphing acoustic sounds with artificially generated spectra to create music. According to Paulo C. Chagas, *Gesang der Jünglinge*:

> [...] develops an aesthetics of hybridism that integrates different kinds of sound material. Inspired by Meyer-Eppler's research in phonetics, Stockhausen explores the living quality of voice to create electronics imitations

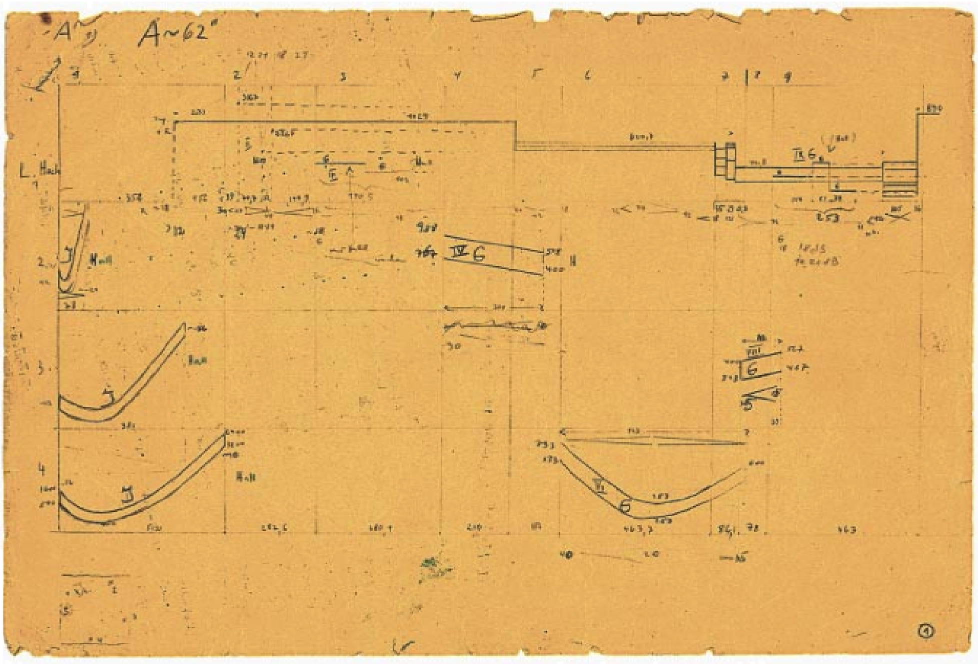


Figure 1.8: Excerpt from the manuscript of *Gesang der Jünglinge*

> of phonemes and other elements of spoken speech and language such as formants, articulation and intonation[...] Stockhausen applied serial principles of composition to the generalization of this material as well to other sound dimensions and to the organization of the formal levels of the composition (Chagas 2014, 110).

Gesang der Jünglinge is serialized work that employs the idea of functions of functions – transposition of the series onto its own elements – developed by Boulez (Decroupet 1998, 100). The text from the prerecorded material is transformed by permutations that obscure its meaning. Although continuing the German tradition of serialism, Stockhausen gives a historically important step forward from early experimental stages of electronic music and *musique concrète* showing a more open minded approach to composition departing from the aesthetical constrains imposed by the studios in Cologne and Paris.

Gesang is historically important for another reason. It includes five channels of audio: four speakers surrounding the audience and one hanging from the ceiling at the center of the hall, and it is the first electronic piece to serialize the projection of the music in space. For its first performance at its 1956 premiere, the fifth loudspeaker was set up on stage. The first four channels where played by a four-track tape machine. the fifth by a separate machine. After the performance, Stockhausen mixed the fifth track onto the first track (Smalley 2000, 11).

Gesang's spatialization seems to aid to clarify the serial texture not as a discrete parameter of the composition. This aspect of *Gesang* shows strong evidence that the composer was well aware of Brant's ideas with regards to spectral content of textures in spatial music. Although the logical step would have been to go from experimenting with monophonic playback to stereo, the idea from working with multiple discrete sound sources can be attributed to Henry Brant who wrote for five orchestras before Stockhausen's *Gruppen* (1955).[25]

1.6.1 Circular Sound Space

The aesthetics of circular motion, emerged after the creation of a device Stockhausen designed himself: the *Rotationslautsprecher* or *Rotationstisch* which allowed sound images to be rotated 360 degrees within a two dimensional plane. The device consisted of a table capable of rotating a loudspeaker that sit on top. Rotation was first controlled manually, later motorized. There were four microphones in a square array capturing the sounds emanating by the rotating speaker. The sounds were recorded into a newly acquired four-track tape recorder, all sounds to be mixed later with mono tape recorders (Manning 2006). This approach to the process of composition was not simple or straightforward requiring different stages of recording and mixing. One important aspect of the process that is usually not considered by many scholars is the fact that

[25] *Antiphony I*(1953) was written for an orchestra divided into 5 sections situated in different parts of the hall to create an effect the composer called *spatial-polyphony* (Brant 1977).

the *Rotationslautsprecher* made noise and generated wind as the table rotated. On top of air and noise there was the addition of the Doppler effect as the sound from the loudspeaker move from one speaker to the next (Decroupet 1998).

Although Stockhausen aimed for great precision and control over the serialized material, the various monophonic tracks could never be synchronized perfectly, the recorders would work at different speeds, the oscillator would never play back the exact frequency of a sine tone the composer meticulously calculated (Manning 2006). All these elements added to the aesthetics of the composition. The result of the parameterization of electronic sounds in early electronic music could never be as "precise" as the ones created by today's digital computers but they incorporate to the aesthetics of the works.

The rotating loudspeaker mechanism was used for *Kontakte* (1959–1960), for four-channel tape, piano and percussion. By using this device, Stockhausen created a new aesthetics of sound spatialization that will be explored and expanded by other composers such as Pierre Boulez notably in his work *Répons*. The aesthetics of circular motion is still today the most popular approach to sound spatialization.

The idea of a moving sounds in space created an aesthetics that cannot be dissociated with the importance of the materiality on the resultant sound space. The circular motion was not randomly chosen by Stockhausen. It is the consequence of a real physical room where microphones could be placed in a quadrant with a rotating table in the middle. The recorded sounds were to be reproduced in a concert hall of rectangular shape with the speakers oriented to the audience at the same angles. The microphones are replaced by speakers, the *Rotationstisch* by an audience.

For Stockhausen, moving sounds in a circular fashion was not enough. Considering that sounds are "non-stationary objects": movement is related to space as it is related to time. The spatial perception of a sound moving in space is hard to achieve with only two speakers. As the perception of a sound moving left to right on a stereo setup is a psychoacoustic phenomena, it only works when the listener is at an equidistant distance from the two speakers. We localize sounds by their variations on their timbres and the different times of arrival of sound waves to our left and right ears. In order to make the circular sound space a structural part of a work, sounds need to change as they move. Stockhausen applied morphing techniques[26] from one sound object to another as sounds rotate. This technique originates the concept of *spectral spatialization*. As the sounds morph while moving in circular motion, they make transitions clear. No matter where the listener is, sounds belong to a space which in return shapes the spatiality of the physical space.

1.6.1.1 *Répons* by Pierre Boulez

Since the beginnings of the microcomputer revolution during the mid 1970s and through the 1980s, processing of large amount of data became faster and cheaper. Research institutions such as Ircam in France applied the technology to computer music, with special emphasis on real-time signal processing. Composer and Ircam's director Pierre Boulez, and a group of collaborators designed and developed a machine that was capable of synthesis, spatialization and transformation of sounds in real time: the 4x. Boulez used the computer for his composition *Répons* which he wrote as a showcase for Ircam. The machine was capable of doing the following real-time transformations:

- The modulation of a solo instrument by a synthetic sound.
- Time shifts systems, mostly delays and echoes.

[26] Mixing is not morphing. Mixing is a change in volume – fade-in and fade-out –. One can transition from one sound to another by fading them accordingly but morphing requires combining two spectra. There are several techniques, for example one can take the amplitude (from its frequency bins in the spectral domain) of one sound and apply the gains of another sound.

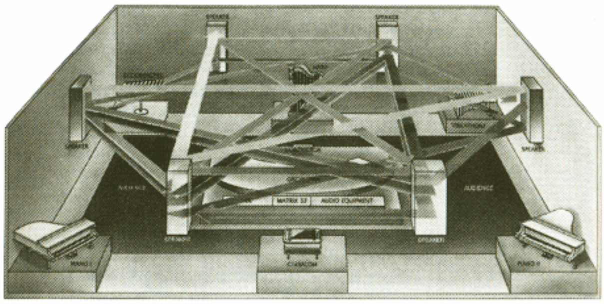


Diagram of the sound projection in *Répons*

Figure 1.9: Répons' sound projection

- The spatialization of sound at different speeds controlled by dynamics (of the soloists).

The spatialization was implemented using an envelope follower as follows:

- A microphone captures the sound of the performance
- The sound is stored in a buffer that contains an envelope
- The envelope (buffer) is divided into four sections
- Each section is transferred to the loudspeakers using a flip-flop unit.[27]

Spatialization in *Répons* is performed in the time-domain. The portion of the sound that is first captured by the microphone is the first one to be transferred to a loudspeaker as in a FIFO system.[28] It was not possible with the 4x to perform a FFT, that is to say, to apply transformations in the spectral-domain.

Boulez wanted to get rid of the "tyranny of the tape" as in Mario Davidovsky's *Synchronisms* where performers need to precisely synchronize with a fixed media leaving no room for personal interpretation and the freedom of internal time as opposed to metronomic time (Jaroszewicz 2013). Boulez positioned six instrumental soloists and

[27]In electronics, a flip-flop unit is a circuit that has two states (0 or 1) and can be used for storing state information. These circuits are the building blocks of digital computers.

[28]FIFO: First In, First Out. An abstraction related to ways of organizing and manipulation of data relative to time and prioritization.

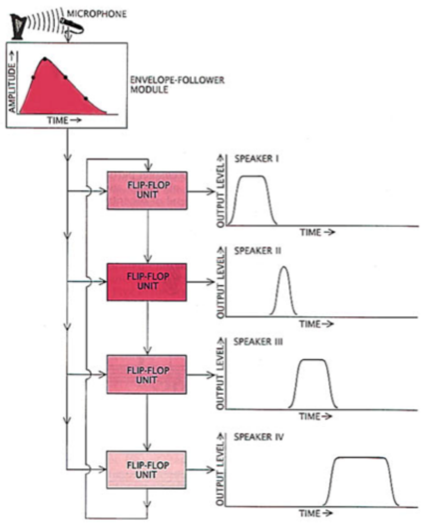


Figure 1.10: Répons' spatial switch

six loudspeakers surrounding the audience, leaving a small orchestra at the center of the performance hall. The "surround" instruments are highly resonant: piano, glockenspiel, harp, vibraphone and cimbalon. The work was inspired by medieval responsorial chants (*Répons*) and the spatialization is used to provide multiple "responses" to a soloist's question. After the opening movement at the entrance of the electronics and the introduction of spatialization the 4x captures a series of arpeggios and creates a new sound field by opening a circular space.

> The attention of the audience is thereby suddenly turned away from the center of the hall to the perimeter, where the soloists and the speakers are. The audience hears the soloist's sounds traveling around the hall without being able to distinguish the paths followed by the individual sounds. The overall effect highlights the antiphonal relation between the central group and the soloists by making the audience aware of the spatial dimensions that separate the ensemble from the soloists and that separate the individual soloists as well (Boulez 1988).

In this case, the trajectories are not intended to be perceived as such. They are a timbral extension of the soloists. Velocities and volume envelopes are map to speed of rotation with the implementation of a switching mechanism. An arpeggio is stored and divided into four pieces which are later spatialized by the matrix. Although spatialization in *Répons* seems to be executed in the time-domain, the four different "dissected" sections of each arpeggio correspond to different parts of the spectrum as sounds change their spectral content over time. A different approach – and technology – than Stockhausen, in *Répons*, there is no morphing or cross-synthesis between sounds.

1.7 The Sound Source

In his article "Composition in Circular Sound Space", Paulo C. Chagas argues that the materiality of the studio where music is created influences the aesthetics of sound spatialization. According to Chagas, space as an embodiment of composition has a special role in the history of electroacoustic music – *elektronische Musik* – and how it relates to serial music. Most importantly, Chagas describes how sound can acquire a tangible property thus becoming material:"When space becomes a parameter of composition, sound develops a "tactile" dimension. Similar to a body, it occupies a unique position in the space from which it can exclude other spaces."

At EARS, I had the opportunity to experiment with systems for spatialization using different layouts and software including Ircam's *Spat*, ICST's *Ambisonics Tools* and the *HOA Library*. I worked with different speaker layouts and different sound sources – ranging from acoustic instruments to computer generated noise – in order to test sounds that had contrasting spectra and repetition rate. The results of the research were applied to a series of spatialization etudes. The apparatus at EARS and the physical space influenced the compositions of the etudes resulting in an aesthetic imposed by the materials. Moreover, upon listening and working with different Wave Field Synthesis systems in which sound sources become "physical", a different way of thinking was needed in order to understand what it meant to compose with sounds that occupy a real space and exclude other spaces.

The choice of sound sources – synthesized or sampled – and their transformations in space change the way sounds are perceived: they occupy a space therefore acquire a "tactile" dimension. In order to understand this transformation, the topic necessitates a different thinking style, one in which sound is not seen as what it appears – perceived – but as what it really is. Thus, a real world perspective is required that defines sound objects as *things* or "concrete" objects.

A thing has the quality of being something that occupies a space in time. In other words, a space-temporal concrete object that in global 21st century culture we associate with material things, things we can observe and are perceptible by touch, which for most people, is the ultimate proof of existence. We can perceive a flower by its fragrance but it becomes a flower when we are able to see it and define its shape. We reaffirm its quality as object by using a combination of senses; a rock becomes a rock when we are able to see it, touch it,[29] move it, measure its weight and feel its shape.

A rock and a flower have different space-temporal qualities, as sound does. A rock occupies a space and might not change its shape in a human perceivable time lapse. We know it is changing by looking at the erosion produced by the elements. For us, the rock is the same rock. On the other hand, a flower grows from a seed and decays in a time span that is shorter than our own. Rocks and flowers are both concrete objects and can be physically displaced while retaining its identity. Men, like the flower, perceivable change over time but we still retain our identity even if those transformations are not part of the natural aging process. This identity is conscious and experiential. Sounds

[29] It is possible to physically touch sounds as demonstrated by Miha Ciglar's tactile controller which uses 97 ultrasonic transducers to focus acoustic radiation. "The result is a relatively strong and spatio-temporally precise tactile reproduction of the projected audio signal". http://www.ciglar.mur.at/sonicbeams.html. Last accessed 01/20/2015.

morph as they travel in space following a trajectory and their identities become part of the musical form.

For the physicalist and the materialist, the world is comprised of small particles or atoms that create things that posses unity. Iannis Xenakis (1922-2001) – composer and architect – applied these ideas as a metaphor for composing with microsounds. In his book *Formalized Music: Thought and Mathematics in Composition*, Xenakis uses the concepts of "granular sounds" and "sound-points" defined by an audible area called "screen". This screen defines the boundaries of the thing, in this case, a sound-thing or a sound object (Xenakis 1992, 51).

A "thing", to be considered as the same thing as it moves in space and time, should preserve those boundaries that define it. For example, given that we cannot perceive ultrasounds with our human ears,[30] a sound object is limited by its frequency components but the combinations of those components are what defines and gives a sound its identity, which we call timbre. For example, the sound of a locomotive changes over time depending on parameters such as the speed of the train, but it preservers its identity; the frequency components undergo small variations but preserve their relationship. Then timbre is analogous to shape and the contour of the shape is the screen. This idea corresponds with our modern conception of matter. We think every object is composed of smaller particles, from elements to atoms to point-like particles. These conceptions thus remain local belonging to some spatial scale. When does a sound object begin and when does it end when we only look at its spatial dimension? If the frequency components keep their relationship over time then the object preserves its unity but then what is the space the sound occupies or is there an occupied space?

[30]Sounds above 20,000 Hz, varying from person to person.

It is difficult to define boundaries for some objects and especially human created things. Considering a sound–object that preservers its unity as time advances, it may occupy a space in our consciousness and does occupy a physical space as sound waves traverse an elastic medium such as the air. A perturbation of the pressure of the medium can be measured with a sound level meter and a sound field can be established. The boundaries of the sound field cannot be established by the listener and the sound object is in reality a *hyperobject*. It goes beyond the limits of human perception. An attribute of a hyperobject is non–locality; these objects are massively distributed in time and space.

Examples of hyperobjects are icebergs, the universe, radiation and music (Morton 2013). Imagine a helicopter flying in a city like N.Y. with a dense distribution of skyscrapers. For the listener on the street, the sound from the helicopter only means "helicopter". It does not carry other information as the reverberations makes it impossible for the human ear to know where the sound is coming from, and for the listener to point to its source. Then, another attribute of a sound object is to be able to be localized. Then the sound–hyperobject becomes a sound–object establishing its spatio–temporal unity.

For everyday material objects, space unity takes precedence over temporal unity (Garcia 2014, 35); an object that can be moved retaining its shape has more permanence than and object that changes as it moves, thus creating a hierarchy. Spatial consistency is an attribute of a thing. In music, a motif, has spatial consistency as it reappears unchanged or slightly varied as the music progresses. This idea was explored by Wagner who used shorts musical ideas – the *leitmotiv* – associated with physical objects or characters in his cycle *Der Ring des Nibelungen*. In Wagner's music, these objects do not undergo a temporal transformation by themselves or are affected by external forces. Leitmotivs are sound objects that are associate to a greater object and from a different perspective, sound objects in Wagner's music acquire another dimension as they carry a musical idea attached to them. For example, as with the fragrance of a flower signifying the flower, the "ring" leitmotiv signifies the ring.

If an object changes and loses its unity as it is being spatially displaced, it becomes "unthingly", bounded with the environment and of an inferior hierarchy. Spatial consistency is an attribute of objects that allow us to re-identify them as time progresses. In music, this consistency creates form and coherence. If a sound–object lacks the attribute of localization remaining a hyperobject, there is an impossibility to reduce it to a spatio–temporal object. But as material things change and our perception of them also changes over time, their spatial limits are different at different scales, for example, a sound particle or microsund can be part of a stream of sound particles, the latter having its own timbral qualities. As things are constantly changing at different temporal scales an insect might live one day and a rock could take millions of years to dissolve. According to Tristan Garcia in his book *Form and Object*, for the formalist or metaphysical abstract things are objects that can exists regardless their spatio-temporal attributes: abstract things exist outside those boundaries. "Many things exist, and we cannot do without the concept of "thing". In the absence of things, the world becomes undifferentiated: the world is a self-saturated whole in itself which knows no differentiation" (Garcia 2014, 37).

In acousmatic music, the sound world of a composition could be a mass of non-localized sound-objects. For example, in an ambience or soundscape composition, where does the sound object begin and where does it end?

According to Pierre Shaeffer, the sound object is not the instrument that has been played (Schaeffer 2003, 58). For example, if we listen to a recording of a band whose instruments are unknown to us, what is what we listen to? For him, that is the sound object. The sound object is neither the instrument nor the tape which is a support for the sound. This tape can contain many different sound objects and when the composer manipulates sound by splicing and pasting its pieces, he is not modifying the sound but creating new sound objects. For Schaeffer, the sound object is not an emotional state because the object does not change from one listener to another or between our different moods or attention spans. The object does exceed our individual experiences: the visual, auditory, tactile impressions and the way we interact with it through perception, memory, imagination, etc.

Shaeffer defines the object as an independent thing that is perceived from a level that is deeper than the acousmatic reduction (Schaeffer 2003, 50). There is no need to interact or find its significance. When we listen to a speech we focus on the concepts that are transmitted by the speaker, the sound is a significant. On the other hand, the sound objects appears when we do this different level of listening that is more rigorous which Schaeffer calls *l'écoute réduite*. It is the sound itself that takes our attention without any meaning attached to; we disregard its significance. A deeper level of listening is achieved by detaching the information about the sign. If we listen to the sound of a helicopter in a recording, it will be hard not to think about a helicopter at the first listening, which, according to Schaeffer, is listening by reference. At the moment when we do not think of the helicopter anymore and instead our attention focuses on the sound object, we are doing a deeper level of listening and the sound object reveals itself.

With Schaeffer's ideas, it is very difficult to see how a sound object can become a thing or how the definition of object applies to an acoustic sound. For us, it is missing some elemental features in our definition of things. As mentioned above, an attribute of a hyperobject is that it is non-local and it is massively distributed in time and space. Music is a hyperobject, but in a micro scale we could consider Pierre Schaeffer's sound–object to be a hyperobject as it is a non-localized object. We could also put in this category Ambient music and ambisonics fields for background support. These non-localized sounds occupy the whole space and they are meant to have no clear beginning or end, they usually give the listener a background layer for other sounds. Brian Eno, describes immersion being the point of Ambient music; "music to swim in, to float in, to get lost inside" (Eno 2004, 95).

Soundscape composition is different "as it tends to involve material offering a connection with a listener's life" (Landy 2007, 45). In order to find the sound object in a soundscape it is necessary to apply a deeper level of listening; but *l'écoute réduite* is not what the composer of a soundscape wants as he intends to communicate something and tries to offer and interpretation of the material that was transformed. On the one hand, Ambience music offers the listener a space to explore without having to search for meanings, on the other hand, spatialization, gives the listener the opportunity to attach to, follow and hear the transformations of the sound object as it morphs into something else, all in a clearly defined sound space by composed or algorithmically generated trajectories.

1.8 The Aesthetics of Circular Motion

We shall focus now on one of the aspects that define the sound object: the property of localization. According to Chagas, the aesthetics of circular motion works very well for the ears. Extended experimentation with circular panning techniques at EARS and at the Wave Field Synthesis system in TU Berlin resulted in excellent perception of circular motion regardless of the system (Chagas 2008). This perception is enhanced by the the visual layout of the speakers, which are usually positioned surrounding the listener. Complex trajectories are more difficult to follow with the WFS system rendering the best results.

But an object that moves elliptically – loops – around the listener creates more than movement, it becomes part of the structure of the composition by adding to its form. The use of concentric circles in a WFS plane can create form as objects move towards the listener at different time intervals. If sound objects preserve their amplitudes throughout their orbits, then a special event occurs when it reaches the closest point to the listener. This varies of course, depending if the listener is static or moving inside the space. For static listeners, there is something that comes back musically. For example, if an orbit takes one minute to complete, and the object completes an orbit during the piece, this could be perceive as A (object's position closest to listener X) and A again when object arrives at the starting location. As a part of a piece, this could be perceived as A - B - A if at some point the object disappears from the listener's ears. We have a similar experience when listening a theme that returns during the recapitulation of a sonata. The sound object becomes a building element of the musical form like a recurring motive.

Another approach is to use orbits that reach listeners at different time intervals. This creates a new challenge for the composer as he needs to think about form from a different perspective. In some cases, the composer could be dealing with as multiple – and complex – forms as there are listeners in the space, even if he wrote the piece with a particular form in mind.

This also raises the question about form and the perception of the work by different listeners at different points in the listening space. A moving audience is very

Figure 1.11: “Wall” of Speakers at the Zuiderstrandtheater in Den Haag

likely when concerts are held in large warehouses or works are part of sound installations. To give an example, the Zuiderstrandtheater in Den Haag held in 2014 a series of concerts where sound was projected from a gigantic array of speakers and the audience was free to move, walk or lay around the listening space.

How many objects can be continuously followed? It varies. It is possible to follow the orbital paths of four[31] objects sounding at the same time. I experimented with sounds including repetitive patterns (drum loops, clicks, noise, short samples) and those provided no problems for the listener but is it possible to follow the location of many objects that transform in time? Usually when there is a great timbral transformation, for example by means of cross-synthesis, where sound A gradually acquires some properties of sound B eventually morphing into B, and the paths are not elliptical, there is chance to lose the connection with the sound. But the question is difficult to answer because if the listener is not static then, what does it mean to follow a sound in space? With a

[31]Stockhausen talks about spatial depth and the ability to hear up to 6 layers in *Four Criteria of Electronic Music. Lecture IV, Kontakte* given at the Oxford Union on May 6Th 1972.

WFS system it is possible to physically "follow" the sound by walking around the space especially if the source is created as a plane wave instead of a point source .[32]

Extensive experimentation and listening with a spatialization system is required in order to find out what works and what does not for the composer. As spatialization becomes a structural part of a musical work, aesthetic choices are left to the composer who can "preview" his work using a binaural tool. Nonetheless, one must spend hours listening and fine tuning paths, dynamics, velocities, etc. The process itself is like any other aspect of a composition. I had the opportunity to listen to many works written for the Game of Life system and found unique approaches when working with moving objects. Some composers see spatialization as a way to create interesting paths that are more visually and conceptually appealing than their musical outcomes. But a great example of a very effective use of circular motion is the piece "Iron Age" by Robert Henke. The composer uses static concentric circles distributed around the field creating a form that has different timing for listeners located at different points inside the field.

Complex patterns with irregular shapes are difficult to follow, they depend on the speed of movement and the timbral characteristics of the objects. The interpolation of continuous movement between two points may not be perceived by the listener but this is hard to achieve even with sounds in the real world.

Using circular or elliptical orbits and straight paths with a WFS system is an excellent combination to provide the listener with a sense of motion and form, as the WFS produces a good sense of depth; sounds can be positioned close to the listener's head and can gradually disappear moving away from the listener.

[32] See the Wave Field Synthesis. Section 3.1

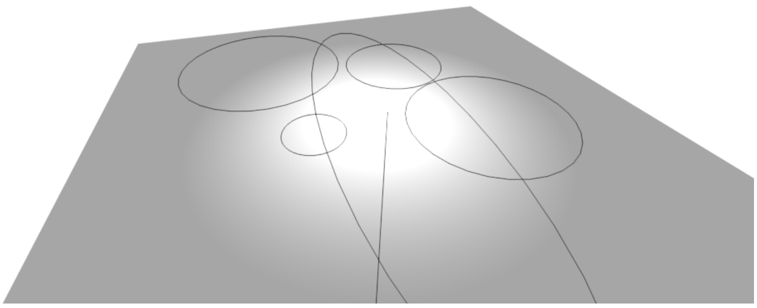


Figure 1.12: Elliptical and straight paths in a WFS system

Part II

Technical Aspects

Chapter 2

Sound Fields

2.1 Stereo and Panning

A *panorama* is a wide angle representation to conform to a flat or curved background, which surrounds or is unrolled before the viewer.[1] In audio, panning is the technique to place a monaural –one channel – sound in the stereo sound field between the left and right speakers to create a sense of space. For example, a performer can be virtually positioned in a semicircle (0°–180°) in front of the listener.

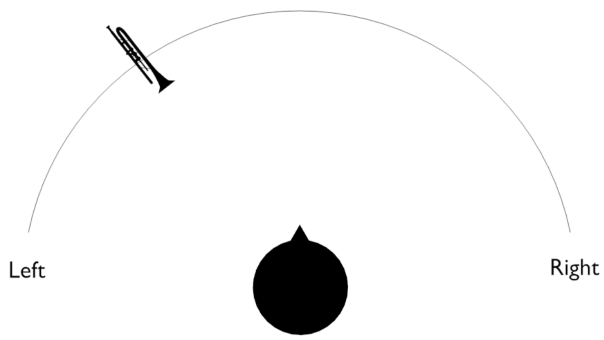


Figure 2.1: Theoretical width of an audio image.

[1]Encyclopædia Britannica Online, s.v. "panorama", accessed January 21, 2015. http://www.britannica.com/EBchecked/topic/441452/panorama.

The field of an audio *panorama* is called the *image*. Although it is delimited by a semicircle of 180°, a width of 90°or 60°is used (Farnell 2010, 221).

There are several techniques for panning including the following:

- Simple Linear Panner
- Square Root Panner
- Cosine Panner

2.1.0.2 Simple Linear Panner

The simplest form of panning by equally decreasing and increasing the power of both left and right signals. A control signal (x: 0.0, 1.0) is used to calculate $(1 - x)$.

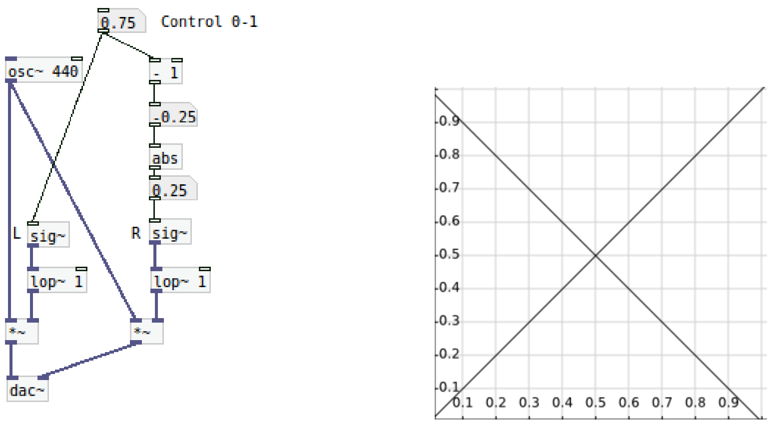


Figure 2.2: Simple Linear Panner in PD with Low Pass Filters for smoothing the signals.

As shown in Figure 2.2, the signal loses power at the center. The square root law panner and the sine-cosine law panner compensate for that loss. See Figure 2.3. The sine-cosine panner is smoother at the edges as it approaches them at 45°. The latter is also good for creating the impression of circular motion in front of the listener, for example, it can be used for placing the musicians of a large orchestra like in a concert hall. For small ensembles –rarely panned hard left or right – a square root panning

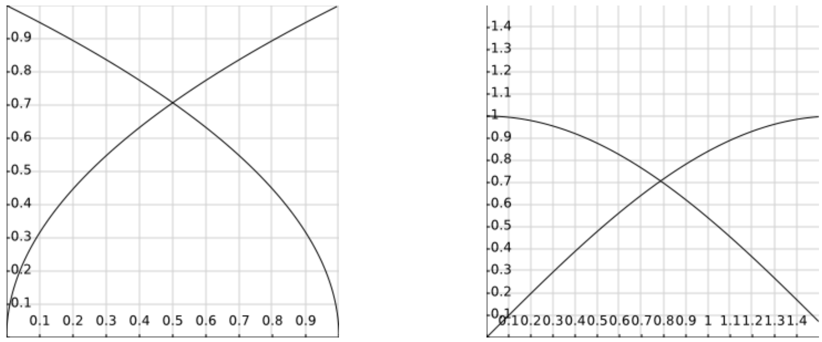


Figure 2.3: Square root panner (left) and Sine-Cosine panner (right)

technique is more desirable as it does have better response at the center (Farnell 2010, 222).

In a stereophonic setting, the output of the speakers is adjusted to a point towards the center of the listening space – a concert hall, studio, etc –. This point between the speakers is known as the *sweet spot*. Panning, due to several psychoacoustic mechanisms, gives the listener the impression that sound moves. This is perhaps the oldest and most common technique that has been implemented for the reproduction of music. Surround sound such as 5.1 is based on the same technique and also requires – like stereo – that the speakers are laid out using always the same setup. This is a *channel based* technique because the mixing is done working on each channel. For example, in Apple's DAW *Logic Pro X* one can automate how sound moves from left to right individually on each track (Figure 2.4).

Although it seems fairly easy to use the pencil tool in a DAW to add movement to sound, it only works with stereo or surround panning. For complex spatialization techniques it is not possible to use a commercial DAW such as Logic or ProTools, instead, it is necessary a dedicated programming language for audio such as Pure Data or Supercollider, especially when working in real time. Alternatively, for the creation of trajectories, 3D modeling software can be used in combination with Python scripting.

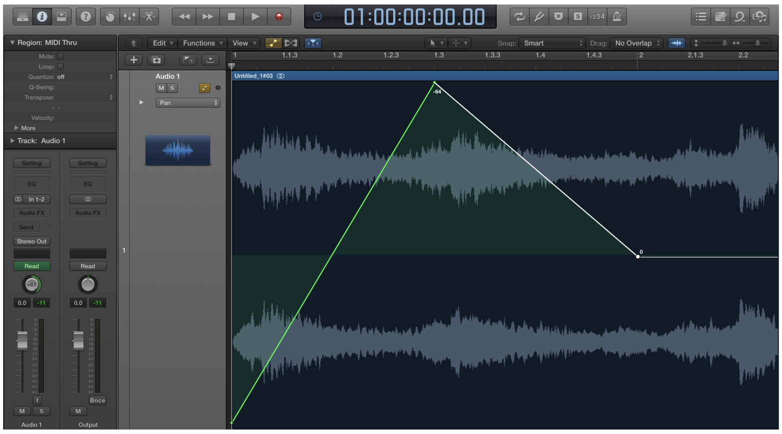


Figure 2.4: Logic Pro X. Panning using automation. From right to left to center in 2 seconds

The example below was written in Supercollider where a grain of sound is panned using a two channel equal power panner class.

```
var env=Env.new([-1, 1, 0], [1, 1]);
SynthDef("pan2",{ Out.ar(0, Pan2.ar(Dust.ar(200),
    EnvGen.kr(env,doneAction: 2),0.3)) }).play;
```

Another and better approach to panning is to use the *Equal Distance Crossfade* technique, where the curve of level differences needed for a specific angle, is approximated by the Blumlein Law:

$$\sin\phi = \frac{gainL - gainR}{gainL + gainR}\sin\omega \tag{2.1}$$

where gainL and gainR are the gains for each channel, ϕ is the angle of the virtual sound source and ω is the angle formed by the loudspeaker. Blumlein Law works only for frequencies lower than 600Hz and a listener's head pointing directly forward. To correct this we can apply the tangent law[2]:

$$\tan\phi = \frac{gainL - gainR}{gainL + gainR}\tan\omega \tag{2.2}$$

[2] These equations can be easily implemented using C-like expressions in Max/MSP or Pure Data.

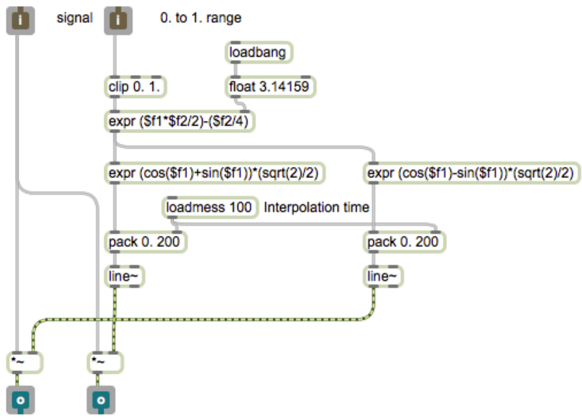


Figure 2.5: Stereo panning in Max/MSP

and compensate for when the source is in central position $\varphi = 45°$

$$gainL = \frac{\sqrt{2}}{2}\cos(\phi) + \sin(\phi) \tag{2.3}$$

$$gainR = \frac{\sqrt{2}}{2}\cos(\phi) - \sin(\phi)) \tag{2.4}$$

Blumlein Law was named after Alan Blumlein, an English electronics engineer who in 1931 developed what he called "binaural sound" when working at EMI in England. His technique used a two channel stereo system for recording and playback as opposed to Ader's stereophonic technique for broadcasting music. It is important to note that recording techniques begun in the middle 1800s with the invention of recording devices such as the phonoautograph, the paleophone and the later phonograph which was capable of both recording and reproduction of sounds.[3]

Stereophonic sound was a practical approach to spatialization limited by the number of speakers necessary to reproduce a wave front that could position a sound source moving from left to right on a theater's stage or a cinema screen. Stereo as the

[3]The phonoautograph, the paleophone and the phonograph were invented by Edouard-Leon Scott de Martinville in 1857, Charles Chros in 1877 and Thomas Edison in 1857 respectively.

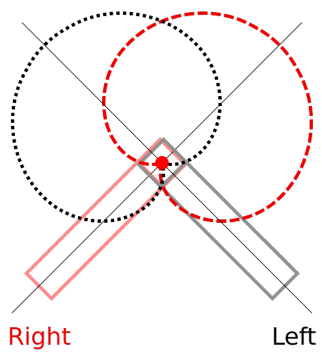


Figure 2.6: XY Stereo microphone placement technique

standard format for the music industry today serves the purpose of virtually reproducing an image – a phantom image – of the location of instruments in a concert hall or studio. Most recordings of live events are done using two microphones. Common microphone placements techniques are x-y, ORTF, NOS, Stereosonic, MS and parallel spaced. Many of these techniques only vary in their angles of maximum response.

2.1.0.3 Surround

Surround refers to an array of speakers *surrounding* the listener. Five-point-one (5.1) and Seven-point-one (7.1) are the most common surround formats adopted by the movie industry today. In surround, the .1 or 0.1 is the Low-Frequency channel or Low-Frequency Enhancement (LFE) that is connected to a subwoofer speaker.

Other formats such as quad – four speakers in a square – never became commercially popular due to a reluctancy from the public to add more speakers to their home-entertainment systems. In addition, there were different competing formats for LPs and tapes between the late 1960s and through the 1970s (Holman 2008, 7). Today, the general public has access to inexpensive surround systems for the home which include a surround decoder or receiver, woofers and a subwoofer. However, quadraphonic systems – sometimes called 4.0 surround – have been adopted by composers willing to add the element of spatialization to electroacoustic music. It works for placement of sounds on each channel as discrete direct radiators at 90°. Pans from the front to the surround speakers do not work. As an aesthetic choice it might be desirable to work with quadraphonic sound when the location of a loudspeaker wants to be perceived as the source of the sound.

The 0.1 is widely used in music to accurately reproduced the low end of an acousmatic piece and it should be used for frequencies lower than 120Hz. In addition to aesthetic reasons, a dedicated subwoofer decreases the intermodulation distortion of the main loudspeakers when handling large amounts of low bass [4] (Holman 2008, 59).

2.1.0.4 Ambisonics

Ambisonics – 1st order or B-format and Higher Order Ambisonics – is a technique developed by Michael Gerzon (1945-1996) in the 1970's as an alternative to quadraphonic systems which he considered to be incapable of good results:

> The fault lies partly in studio equipment incapable of giving good quadraphonic results, and partly in erroneously conceived "quadraphonic" systems which leave the apparent localization of sounds at the mercy of the listener's imagination (Gerzon 1974).

Ambisonics can spatialize sound fields using 2D and 3D multi-speaker systems. It decomposes the sound field at a point using spherical harmonics up to a certain order (1st-order ambisonics, 2nd order ambisonics, etc). In most implementations it only accounts for plane waves, that is, only the direction of the source is considered.

[4] Intermodulation distortion is a type of non-linear distortion that comes with a non-linear relationship between the input signal and the resulting sound. For example, a common distortion is created when a smaller sound is masked by a larger sound. According to Floyd Toole in "Sound Reproduction" this is not a problem with loudspeakers unless they are put into a limiting condition.

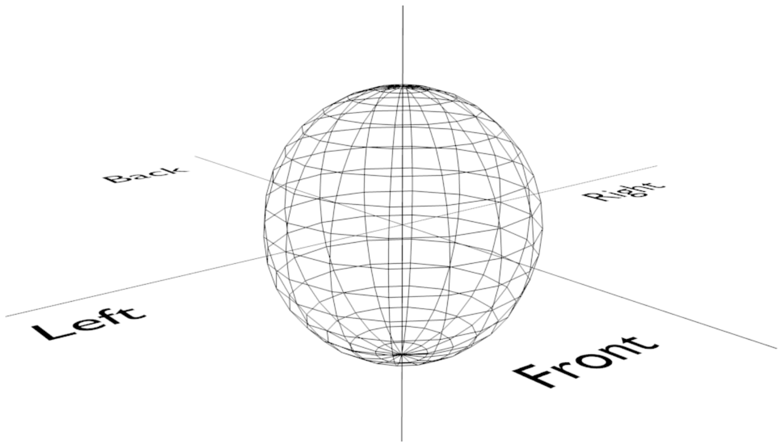


Figure 2.7

Ambisonics can be used for recording and playing back multichannel audio in a 2-dimensional or planar space (pantophonic system) or a 3 dimensional space or "full sphere" (periophonic system) (Figure 2.7). This system encodes the signal in three channels for pantophonic systems and requires an extra "height" channel for periophonic ambisonics encoding. The main purpose of using Ambisonics is the creation of a surround sound field. It is possible to render virtual acoustic spaces with sound sources that can be positioned and moved around a space in real-time.

The main advantage of using ambisonics for musical composition is *isotropy*. [5] Ambisonics is *isotropic* meaning that sounds from all directions are treated equally. Another advantage of using Ambisonics is that the position of the speakers can vary from square to circles to regular polygons. All speakers are used to localize sound creating a sound field that is stable regardless of the listener position. In order to understand how panning works is best to understand how 2D panning works first. We saw that panning is a technique for positioning a monophonic sound within a stereophonic image. Panning uses differences in gain that are fed into the loudspeakers nearest to the virtual sound

[5] Uniform in all orientations; it is derived from the Greek isos ("equal") and tropos ("way"). For example, a bell is an isotropic radiator of sound. The sun is an isotropic point source of light.

source. In a stereo system, using linear panning or the square of the cosine function, the gains or the squares of the gains add to 1.

$$\sqrt[p]{gainL^p + gainR^p} = 1 \tag{2.5}$$

or

$$gainL^p + gainR^p = 1 \tag{2.6}$$

In Ambisonics, all channels add to 1 at the same time thus creating a sound field instead. The panning functions are defined for 3 dimensions. In addition to the recording of the sound source, a spherical array of microphones will also record room information that can be reproduced (decoded) independently of the speaker setup.

The decoding of Ambisonics can vary from different degrees of accuracy depending on the desired directionality for each speaker that is used to create the virtual sound field. This directional accuracy is given by the so-called order of Ambisonics. The zeroth order corresponds to a mono signal of equal loudness for all speakers. An order of 7 with an octophonic speaker layout gives a good amount of directionality while a higher order localizes the sound at one speaker as shown in Figure 2.9.

Ambisonics is not speaker-dependent or listener-dependent. Although there is a *sweet-spot* and better images are created at the center of the field, an accurate representation depends on the order. As shown in Figures 2.10a, the ambisonics order defines the accuracy. As ambisonics does not provide information about distance, when working with synthesized sources, it is possible to simulate *distance encoding* with the use of artificial reverb and filters. In addition it is also possible to simulate sound sources with other shapes other than points (Pérez-López 2014). Ambisonics – 1$^{\text{st}}$ order

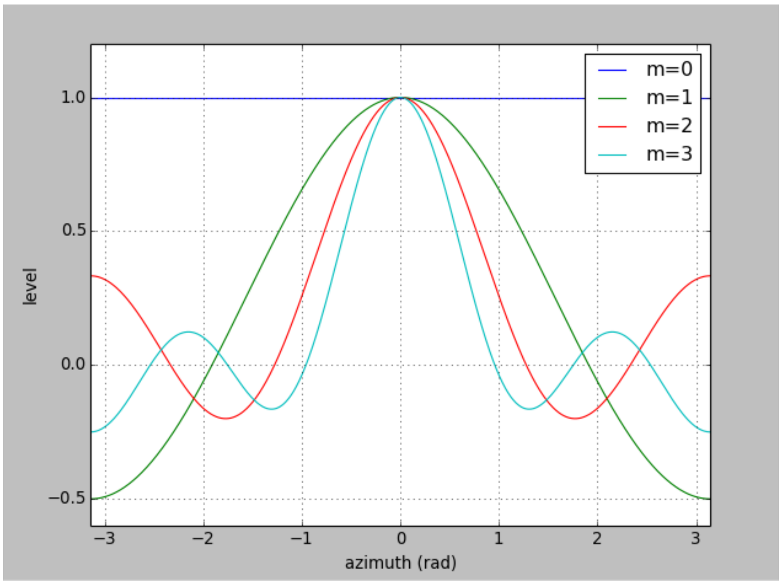
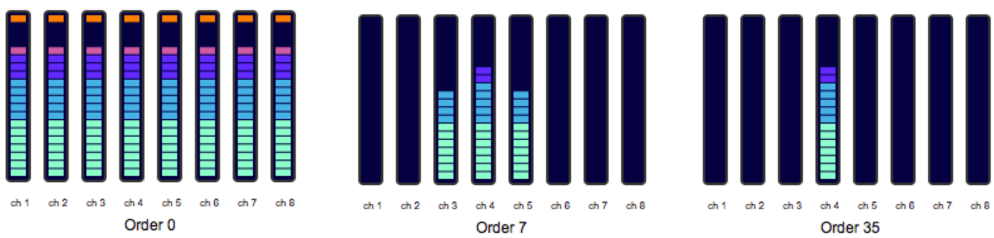


Figure 2.8

Figure 2.9: Directivity of a point source encoding with different Ambisonic Levels (Pérez-López 2014).



(a) Order zero (b) Order seven (c) Order 35

Figure 2.10: Different ambisonics orders and their directionality in Max. The patch was developed at EARS for the works *Laberintos*.

– is a good system for composing a sound field that surrounds the listener without moving sound objects. For example, it is possible to create a sound field with only four speakers, but for moving sources, higher orders are needed – and more speakers – as the order defines the level of localization accuracy. For example, one of the advantages of using Ambisonics is the ability to record using an array of microphones to capture a soundscape:

> Recorded 1^{st}. order sources, although of a lower spatial resolution, capture "reality information" from the original source, including proximity, image size, the real balance between events in the original scene and a unified 3D image of the fore- mid- and background – or in other words the sense of environment and perspective (Barrett 2010).

Ambisonics is a technique that requires two stages: encoding and decoding. A mono signal enters the system and it is encoded into a number of different channels – depending on the HOA format – and is sent to the a decoding algorithm of the same order to be distributed to N speakers. During the encoding process the sound waves are projected into the spherical harmonics using four channels with omnidirectional w, x, y and z-directional information. The decoding process reconstruct the spatial scene for a minimum number of speakers in a circular array specified by the formula:

$$2 * (order + 1) \tag{2.7}$$

Mathematically, the directional encoding of a sound source S on three orthogonal directions x,y,z of the unit vector u - or rather an incident plane wave u carrying a signal S results in the following equations:

$$\begin{cases} W = S \\ X = \sqrt{2}\vec{u}.\vec{x}S = \sqrt{2}\cos\theta\cos\delta S \\ Y = \sqrt{2}\vec{u}.\vec{x}S = \sqrt{2}\sin\theta\cos\delta S \\ Z = \sqrt{2}\vec{u}.\vec{x}S = \sqrt{2}\sin\delta S \end{cases} \tag{2.8}$$

where the vector $\vec{u}$ is described by the spherical harmonic (θ, δ) (Daniel 2001).

2.1.0.5 Wave Field Synthesis

Wave field synthesis is also a spatial sound field representation with the purpose of the reproduction of auditory scenes.[6]

WFS is among the newest techniques for sound spatialization and requires large number of speakers to create a virtual auditory scene. The system completely overcomes the limitations of stereophonic and ambisonics regarding the listener's position in the sound field such as the *sweet-spot*. Rather, there is a *sweet-area* delimited by the layout of the speakers. Given the number of speakers required to implement the technique, attempts have been made using a linear distribution thus creating a planar listening area as shown in Figure 2.11.

WFS is based on the principle of Huygens which states that when there is a wave front, it is possible to synthesize the next wave with an infinite number of small sources or *spherical waves*. Sound sources can be located outside or inside the front, though, the sound source should always remain before the listener. In other word, the listener is positioned inside an enclosing array of speakers (Figure 2.11). In

[6]Auditory scenes refer to all audible sounds belonging to a sound field from the perspective of a listener. *Auditory scene analysis* aims to study how the human brain reconstructs sound. There are practical applications in the field of artificial intelligence such as *machine listening*

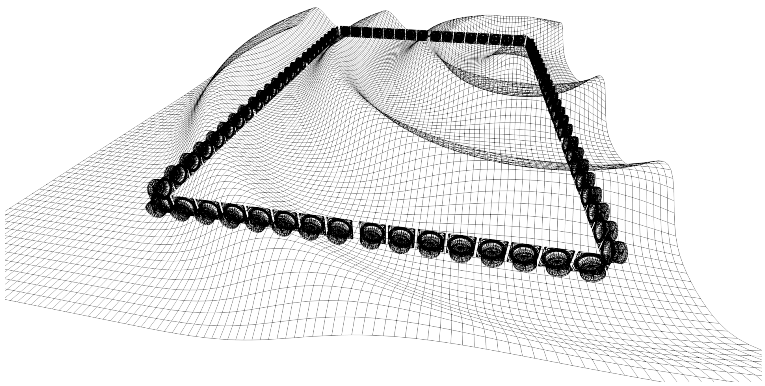


Figure 2.11: Speaker array, point source and wave generated by the system.

WFS, loudspeakers are used as **secondary sources**. This way, virtual sources can be synthesized in the listening area by making use of the Huygens principle.

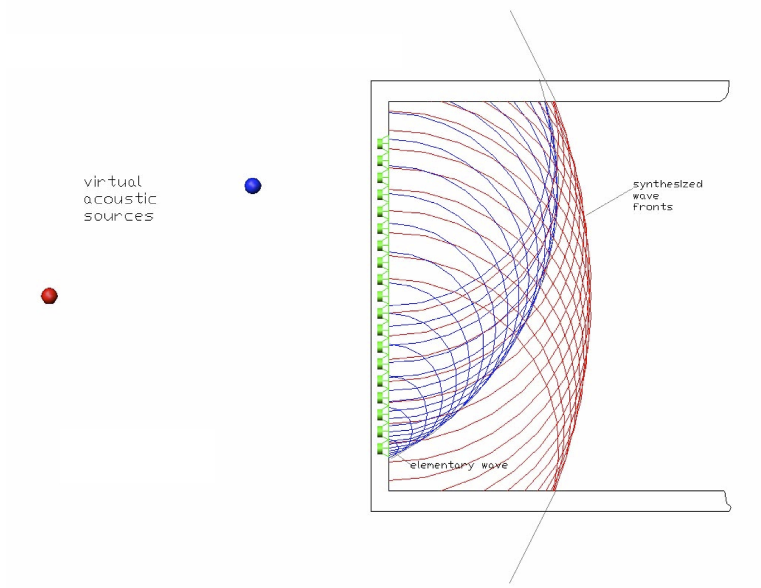


Figure 2.12: Wave Field Synthesis. Huygens' Principle (Snoei 2014)

From a mathematical[7] point of view this principle states that a wave field at time $t+\delta t$ can be synthesized by replacing the wave front at time t by an infinite number

[7]For an extensive discussion behind the physics and mathematics of WFS see (Sascha 2008).

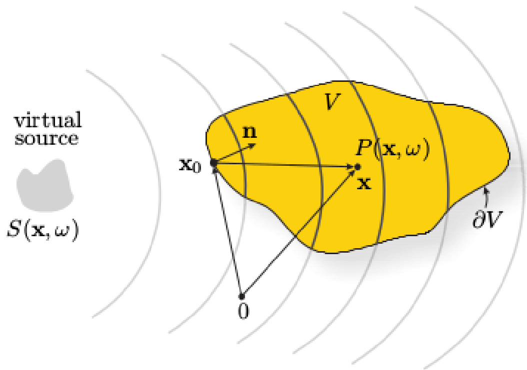


Figure 2.13: Kirchhoff-Helmutz. Illustration of the geometry (Sascha 2008).

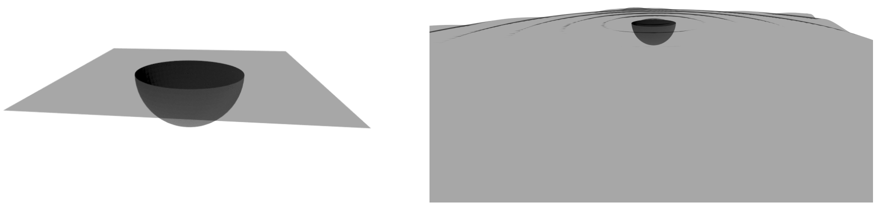


Figure 2.14: Wave Field 2D plane and listener's perspective

of secondary sources at an infinite small distance δx from each other; this is basically the Kirchhoff Integral (Van Dorp Schuitman 2005) which states that it is possible to reconstruct the amplitude of a source at a given point inside of an enclosing surface:

$$P_A(\mathbf{r},\omega) = \frac{1}{4\pi}\int_S \left[P \frac{\partial}{\partial \hat{\mathbf{n}}}\left(\frac{e^{-jkr}}{r}\right) - \frac{e^{-jkr}}{r}\frac{\partial P}{\partial \hat{\mathbf{n}}} \right] dS \tag{2.9}$$

A WFS system creates a point source at the desired location. The point source is where the wave will start to propagate in all directions like an inflating balloon (Snoei 2014).

but WFS is not a 3 dimensional system thus it only uses the plane that intersects with the wave field as shown in Figures 2.14. This simplifies the wave field to a 2

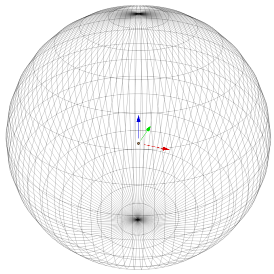


Figure 2.15: WFS Spherical Harmonics

dimensional system of coordinates: (x,y) and from the listener's point of view the first waves to arrive provide localization cues. [8]

As shown in Figure 2.12 the array of speakers recreates the waves originating in a virtual source as if they were passing through a "glass"[9] that transforms them in "real" waves, each speaker reproducing a portion of the source. Thus, for the system to work, all speakers must be working together and must be synchronized. Given that there is no *sweet spot*, all listeners hear the source radiating from the exact same location – the source has an **absolute position** – and a physical wave is generated .[10] The acoustics of the room responds accordingly and does not play an important role in localization.

The integral in 2.9 tells us that we need an infinitesimal small distance between the speakers[11]but this is impossible in practice and one of the physical limitations of the system, as loudspeakers are used to recreate a wavefront. This approximation of sound sources by loudspeakers results in a spatially discrete source distribution with

[8]Localization cues are ITD (Interaural Time Difference, IID (Interaural Intensity Difference) and Filtering provided by the shape of the head.) ITD is the difference in time sound arrives at the left and right ears and IID is the measure of the amplitudes arriving at each ear. Combined they are used in a model known as the *head transfer function*.

[9]The Huygens' principle was first applied to optics.

[10]As opposed to ambisonics or stereophonic systems that produce a psychoacoustic effect.

[11]r is the distance to the source.

speakers located next to each other as close as possible. The resulting discretization effects may be described in terms of *spatial sampling* . For the human range of hearing (about 20,000 Hz) loudspeakers have to be placed at a distance less than 1 centimeter apart, something not possible considering the size of available speakers and the quantity needed for the construction of the array (Rabenstein 2006). The system will distort when representing frequencies above the *spatial aliasing frequency* as frequencies start to overlap. Distorted[12] frequencies will not contribute to spatial cues and the listener perceives them as *coloration* (Van Dorp Schuitman 2005). Speakers that are approximately 17 centimeters apart will have an spatial aliasing at 1000 Hz. so to in order to avoid distortions in the human hearing range, secondary sources should be less than one[13] centimeter apart. Fortunately, higher frequencies do not contribute with spatialization cues as much as frequencies in the range 800-1600 Hz where the brain uses a combination of cues from both ITD and IID.[14]

For the system to calculate the wave field, each speaker has to process an algorithm that considers the distance from the speaker to the sound source in meters d, the speed of sound (344m/s) c, delay d/c and the scaling amplitude factor in a 2 dimensional system of coordinates.

Basically, calculations are based on the same theory used to calculate an ITD panner[15] but from the perspective of a speaker array in space with a system of coordinates that includes a virtual and a real space. Knowing the distance between each speaker, the attenuation – loss of intensity due to the air – and the position of the source accordingly to a system of coordinates we can calculate the distances using trigonometry:

[12] In this context, distortion refers to *spatial distortion* and the ability to properly localize sound

[13] 8.5mm

[14] ITD is most effective at frequencies below 700 Hz because of the wavelength, stopping at around 1500Hz where IID is used instead (Farnell 2010).

[15] A simple panner with the addition of a delay line.

$$d_L(r, d, \alpha) = \sqrt{(r \cdot sin(\alpha))^2 + (r \cdot cos(\alpha) - \frac{d}{2})^2}$$
$$d_R(r, d, \alpha) = \sqrt{(r \cdot sin(\alpha))^2 + (r \cdot cos(\alpha) + \frac{d}{2})^2}$$

$$delta(r, d, \alpha) = d_L(r, d, \alpha) - d_R(r, d, \alpha) = \sqrt{r^2 - d \cdot r \cdot cos(\alpha) + \frac{d^2}{4}} - \sqrt{r^2 + d \cdot r \cdot cos(\alpha) + \frac{d^2}{4}}$$

Figure 2.16 below shows a system of coordinates with the origin at the center of the leftmost speaker on the x axis. The y axis is oriented as the depth and the unit is 1 meter. A sound source has positive coordinates and the listener negative coordinates.

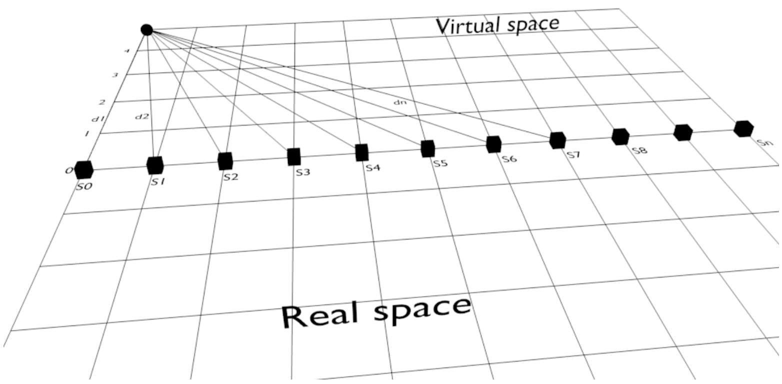


Figure 2.16

We can express speakers' positions as a function of their indexes, going from 0 to n:

$$S_i(d) = (d \cdot i, 0) \quad i = 0..n \tag{2.10}$$

If the speaker line has a length S_n and (x, y) are the speaker coordinates then:

$$D_i(d, x, y) = \sqrt{(x - d \cdot i)^2 + y^2} \tag{2.11}$$

Sound pressure p is a force per unit area, in N/m^2 and the peak sound pressure of a wave is inversely proportional to the distance decreasing $\frac{1}{r}$ for a distance r[16] so we have to use a level scaling factor L_i proportional to $\frac{1}{D_i}$ for each speaker (Bole 2008):

$$L_i(d, x, y) = \frac{1}{D_i(d, x, y)} \tag{2.12}$$

and delay:

$$Del_i(d, x, y) = \frac{D_i(d, x, y)}{343m/s} \tag{2.13}$$

An implementation only taking into account delay times and amplitudes for each speaker is shown by using the Faust programming language.[17] The hypothetical scenario has 8 speakers that are 10 centimeters apart – for simplification – with the sound source located at (0,5) – far left – as shown in Figure 2.16 s:

```
import("math.lib");
import("music.lib");
import("filter.lib");

d = 1; // speakers at 1 meter apart
x = 0; // in meters
y = 5; // in meters
nSpeakers = 8;

Quad(x) = x * x ; // x squared
D(d,i,x,y) = Quad(x - (i - 1) * d) + Quad(y) : sqrt ; // distance calculation for each speaker

// Amplitudes assignments:
```

[16]Sound pressure is an absolute measurement and applies to a point in space where the measurement is made without taking into consideration the direction of the wave. Sound Pressure Level (SPL) is a ration given in decibels (Farnell 2010).

[17]Faust is a **functional** programming language that offers high-performance signal processing at the sample level thus suitable for low level DSP operations. http://faust.grame.fr/. It can be used for building VST plugins, PD externals, Max and others.

```
Amp(d,i,x,y,sig) = sig / D(d,i,x,y) ;
OutA(d,1,x,y,sig) = Amp(d,1,x,y,sig) ;
OutA(d,i,x,y,sig) = OutA(d,i-1,x,y,sig), Amp(d,i,x,y,sig) ;

// Delay amounts assignments:
R(d,i,x,y) = fdelay1s(D(d,i,x,y) * SR / 343) ;
OutR(d,1,x,y) = R(d,1,x,y) ;
OutR(d,i,x,y) = OutR(d,i-1,x,y), R(d,i,x,y) ;

// sequence composition:
Out(d,n,x,y,sig) = OutA(d,n,x,y,sig) : OutR(d,n,x,y) ;
process = Out(d,nSpeakers,x,y) ;
```

The program generates output signals y_i for $i \in [1, 8]$

1. $y_1(t) = 0.2$
2. $y_2(t) = 0.196116135138184$
3. $y_3(t) = 0.185695338177052$
4. $y_4(t) = 0.171498585142509$
5. $y_5(t) = 0.156173761888606$
6. $y_6(t) = 0.14142135623731$
7. $y_7(t) = 0.12803687993289$
8. $y_8(t) = 0.116247638743819$

2.2 Characteristics of the WFS

There are three different types of sources:

1. Point Source
2. Focus Point Source
3. Plane Wave

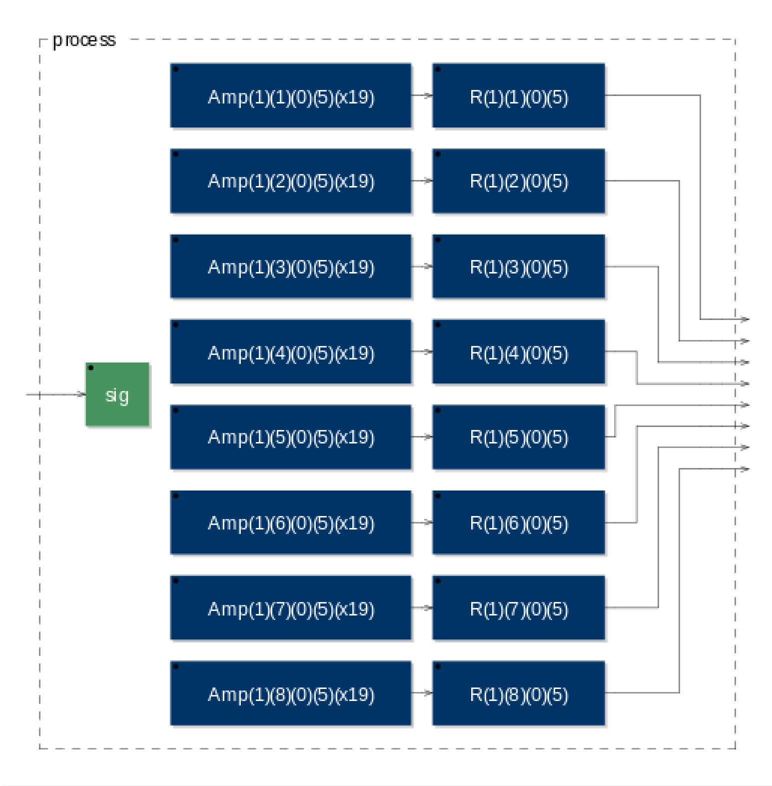


Figure 2.17: Block diagram of the process generated by Faust svg compiler

The *point source* gives an absolute position for every listener in the room, when the listener moves the source is still perceived as being in the same position. If there are multiple sources, there is a sense of perspective of the sound field. A *focus point source* is a source that is located in front of the speaker array.[18] The location of these sources will be perceived accurately only by listeners who are in front of the source. Finally, *plane waves* are sources that have the same angle for every listener. They “”follow the listener when he moves inside the space. The latter is mostly used for reverb/reflections or for emulating a conventional stereo reproduction[19] (Snoei 2014).

[18]Sources located in the *real* space as opposed to the *virtual* space have negative x coordinates.
[19]It is possible to emulate any stereophonic or surround setup when placing sources –virtual loudspeakers– accordingly.

2.2.0.6 Delays, Doppler and Distance Cues

In the example above delay lines are interpolated using a 1st-order Lagrange interpolation .[20] Given that there must be some sort of interpolation when sources are moving, there is a cost on CPU if applying higher orders in real time but they give better sound quality and sound localization. In addition, when sources move from the "virtual" space behind the speakers to the "real" space in front of them, delay times are inverted causing a "click". The Game of Life uses a crossfading technique to solve this issue, other systems may use interpolation techniques but according to Wouter Snoei these are not as effective.[21]

Doppler shifts are generated as a natural side effect on WFS systems and happen because of the distance between the source and the listener: *normal Doppler shift*, and the distance between the source and each speaker: *Speaker Doppler shift.* The former can be canceled out but the latter, caused by the discontinuity of the array, cannot be canceled and could be exacerbated by the cancellation of the *normal Doppler shift, producing a chorus effect* (Snoei 2014).

It is possible to improve distance cues by using filters and adding a global amplitude roll-off relative to a reference point or line.[22] As with any system, one can apply a low-pass filter to mimic distance cues as higher frequencies decay faster than lower ones.

[20]"Lagrange interpolation is a well known, classical technique for interpolation [194]. It is also called Waring-Lagrange interpolation, since Waring actually published it 16 years before Lagrange [312, p. 323]. More generically, the term polynomial interpolation normally refers to Lagrange interpolation. In the first-order case, it reduces to linear interpolation." (Smith 2010).

[21]Wouter Snoei is the lead developer behind The Game of Life software end of the system. An alumni of the Institute of Sonology in Den Haag, the composer still maintains, updates and add new features to the Supercollider driven system. WFSCollider can be installed from its Sourceforge repository at: http://sourceforge.net/projects/wfscollider/.

[22]The Game of Life system imposes a roll-of of 6dB per distance doubling according to the inverse square law (l/r) (Snoei 2014).

To conclude, WFS is more accurate than first-order-Ambisonics for moving sources as the distance cues are clearer. Given that there is no sweet spot, it can be used for installations and acousmatic performances where listeners can freely move around to experience a work from different points or perspectives. There is a visual discrepancy between the output of the system and the location of the speakers as sound sources can be positioned anywhere – up to 200 meters away in the Game of Life system – around an audience which is mainly accustomed to listen to stereophonic and surround setups. In concert situations, I have seen people looking at the speakers surrounding them "looking for" sounds as if they were moving from speaker to speaker. With a WFS system we see something but we hear something that does not corresponds with the visual cues we rely on for spatial processing. I do not think this is a disadvantage of the system, on the contrary, it opens the doors to a new way of listening that goes beyond the loudspeaker and beyond the constrains of the physical space.

2.3 The Design of The Portable Game of Life System

Game of Life is a portable WFS system that was located in room ANNA RW47, at Raamweg 47, The Hague, Netherlands in November 2014 when I visited. It is maintained by the Game of Life foundation which is subsidized by the Performing Arts Fund NL (Fonds Podiumkunsten NL). The foundation was founded in September of 1999 and its purpose is to promote electroacoustic music organizing and curating concerts in the field of spatial reproduction.[23] The GOL was built as a system that can be easily used by composers with good sound quality and the most important aspect: portability. The system can be pack by a few people and transport anywhere by land.

[23]The site of the Game of Life Foundation is active and contains documentation and how to place a request to work with the system. It can be accessed at: http://gameoflife.nl/en/.

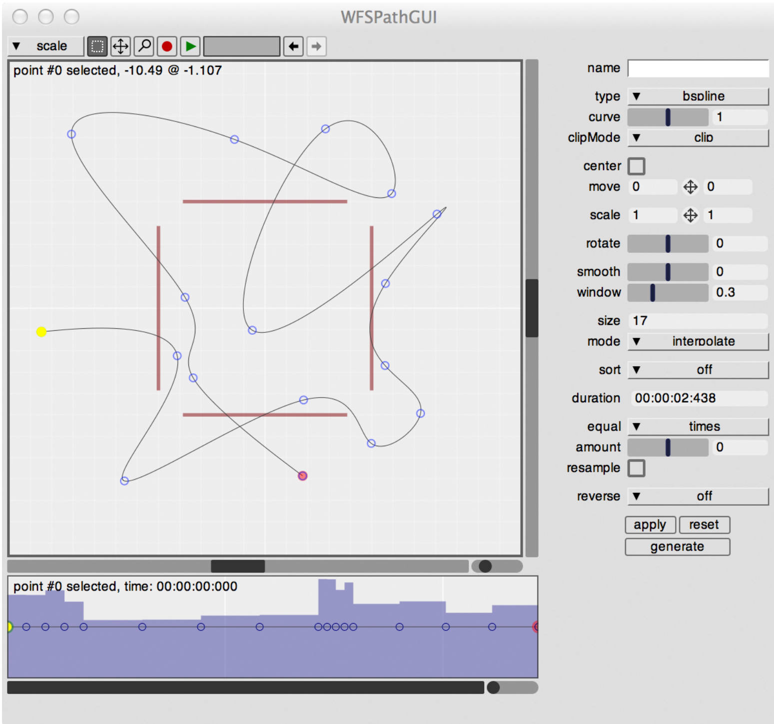


Figure 2.18: a complex path created with the WFSCollider software

One of the most remarkable aspects of the GOL is its user interface that allows the composer to interact with the system in an intuitive manner.

The hardware was designed by Raviv Ganchrow accordingly to the following specifications:

- 192 coaxial speakers (SEAS)
- 12 active subwoofers (Hypex)
- 24 8-channel Class D amps (Hypex)
- 24 Behringer ADA8000
- 8 Motu 2408mk3 / 2 PCle
- 2 Mac Pro 8-core 21.8Ghz
- MacBook with a Motu 828 sound card
- ± 300 meters of cables

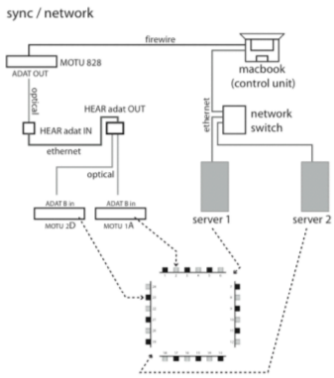


Figure 2.19

The system can be configured as a 10x10 meters array. There are 24 modules; for a square setup there are 6 modules on each side, each containing 8 speakers. It can also be configured in any shape allowed by the 24 modules including for example, 5x7 modules, circular setups, etc. There is a subwoofer every other module. Each module is driven by one 8-channel amplifier that receive an analog signal from a Behringer ADA8000.[24] The MacBook computer is connected to the Motu828 audio interface for synchronizing the digital signals with the 8 2408mk3 connected to 3 ADAs, each 2408mk3 controlling three modules. There is a network switch between the control unit (MacBook) and a HEAR ADAT ethernet extender between the Motu828 and the 2408mk3. The Mac Pro servers controls 4 of the motu2408s as shown in Figure 2.19 (Snoei 2014).

The open source software fully written in Supercollider features a solid interface with a live input and control of every sound source's spatial position. The system is being expanded to consider directionality[25] and the addition of more effects and synthesis techniques (Snoei 2014b).

[24]A Digital to Analog Converter.
[25]Aperture of the virtual speakers such as in IRCAM's SPAT (Spatialisateur 2012).

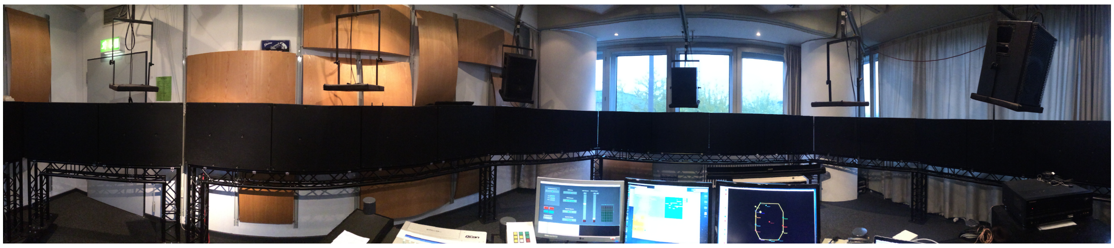


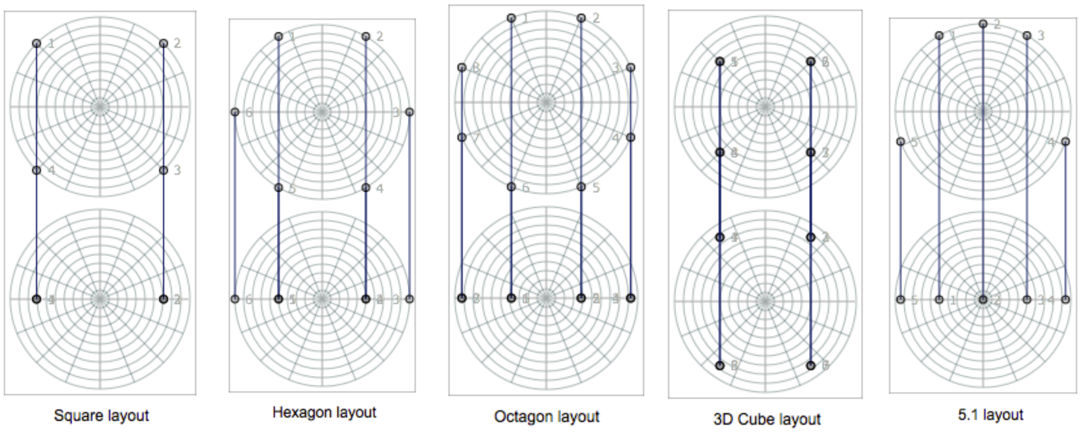
Figure 2.20: WFS studio at TU Berlin controlled by parametric spatialization tools written by the author for the Pure Data software.

2.4 The TU Berlin Systems

The Technischen Universität Berlin has two WFS systems. The lecture and event hall WellenFeld H104 features a system with more than 2000 speakers grouped in modules of 8 channels with 3 drivers for each channel and two subwoofers. Drivers are at a distance of 10 centimeters each. The speakers are controlled by a computer cluster on the left side of the stage with 832 audio channels. The hardware is controller by a Linux system running the sWonder software written at the TU Berlin.

The other WFS system is located at the 3rd floor in the EN - building of the Department of Audio and Communication (Fachgebiet Audiokommunikation) and is mainly used for research and composition. It is possible to compose in the studio and transfer the work to the larger system. Like the Game of Life, there is a Supercollider server that is located in a different room used to control the spatialization. The system features the following hardware and software:

- 24 panels with 8 channels each, 2 woofers connected to 4 channels. 24 tweeters.
- Ethernet/Network based Dante Controller Matrix by Audinate.
- 1 main computer with the matrix software (Mac Pro in different room)
- 2 nodes computers where the spatialization software/interface runs.



(a) 3D view (b) Side view (c) Top view (d) Top view (e) Top view

Figure 2.21: Speaker layouts.

- Xwonder. Custom software built at TU.
- Iannix software for creating algorithmic trajectories with OSC in/out.

The system can be controller by OSC messages from other applications such as Pure Data.

Chapter 3

Spectral Spatialization

This chapter deals with a thorough description of the research conducted at EARS using its eight channel system. It includes a description of sound morphologies (Landy 2007), spectral typologies and morphological archetypes used with different spatialization techniques. I emphasize spatialization techniques such as elliptical motion, paths and spirals that were used for a series of pieces under the name *Laberintos* inspired by Jorge Luis Borges' story *La biblioteca de Babel*.[1] These etudes explore the use of spatialization techniques using ambisonics, SPAT and discrete spectral spatialization, the latter a technique developed at EARS that combines the creation of an isotropic sound field with directionality and spectral panning.

The EARS studio is a satellite facility of the University of California Riverside. Designed and built by director Paulo C. Chagas for research in experimental electroacoustic music, it became a flexible spatialization studio that allows for different

[1] *The Library of Babel.* Borges' short story describes a universe that consists of a vast library containing an undefined number of hexagonal rooms that are identical. The order of books and their content seem to be random and extend to the infinity. The books contain every possible permutation of characters, from books with only one letter to meaningless books of random words. Given that all the combinations of the alphabet are possible, the library holds all possible books in the universe.

configurations of speakers which can be arranged very quickly. The following is the audio equipment used during the course of the research:

- 8 Genelec 8030 loudspeakers
- RMS firewire audio interface
- Mackie Pro FX16 Mixer
- iMac 2011
- ProTools, Max/MSP, Pure Data, Supercollider, Ircam Forum plugins

I have worked with several electronic music techniques in the time and spectral domains but special attention is given to spectral morphing, a technique for merging two sound spectra. Basically, frequencies that are common to both sounds are emphasized while the rest of the spectrum is diminished. The spectrum of a wave is a two-dimensional representation of its frequencies and their relative amounts at some point in time. By morphing two sounds new spectra with some characteristics of both sounds is created. Morphing between two sounds requires a process that starts with the Fast Fourier Transform and the technique is particularly interesting when combined with synchronized sound trajectories that add musical meaning to movement, the main goal of this research. Two basic procedures can be applied: sounds and musical gestures can be spatialized to their entire length or can be timed stretch to fit their trajectory's length using a one-to-one – injective – mapping technique, the most common approach to spatialization, or by spectral morphing two sounds – bijective – in a single trajectory which is a more interesting approach to spatialization.

The main tools used for the design and implementation of spectral spatialization were the following[2]:

[2]Software tools were developed using the Pure Data and Max/MSP programming environments and sound examples were also generated with Csound, Open Music and/or Apple Logic.

- The Fast Fourier Transform
- Cross-synthesis
- Envelopes
- Panning and Ambisonics

The FFT

3.1 Spectral Domain. The Fourier Transform and its musical applications

A periodic wave is a wave that repeats itself. Although natural phenomena seems to be governed by chaos, there is an element that makes it comprehensible for us: patterns of repetition. Seasons repeat every year, Comet Halley is visible from Earth every 75 years, water undergoes a cycle from precipitation to evaporation. If we look at the smaller scale of natural phenomena we find that mechanical waves, oscillations that travel through space and time, are an essential component of nature given their ability to transport energy from one place to another.

We have evolved to perceive the types of waves that manifest in the form of light and sound with our eyes and ears respectively. They differ in the way they propagate, their frequency content and length. Visible light is electromagnetic radiation with waves in the frequency range of 405 THz to 790 THz with a wavelength in a range from 380 nanometers (nm) to about 740 nm. If light travels through a prism, it decomposes into its constituents colors; it was Isaac Newton who discovered that light could be reconstituted back into its original form if passed through a prism again. We cannot use a prism to decompose sound into its sinusoidal components but there is a mathematical device: the Fourier Transform, which can deconstruct a sound wave and under certain conditions, the same wave can be mathematically reconstructed using the Inverse Fourier Transform.

There are several terms related to the Fourier Transform and its implementations: *Fourier series*, *Fourier Transform*, *Discrete Fourier Transform* (DFT), *Short Time Fourier Transform* (STFT) and *Fast Fourier Transform* (FFT). Although they refer to different things, the latter is the most common term used in audio applications.

Joseph Fourier (1768-1830) was a French mathematician and physicist who, after studying how heat propagates to an object, proposed that a periodic wave could be expressed as the sum of infinite simple waves. The Fourier series decomposes periodic functions or periodic signals [sig˜] into the sum of an infinite set of cosine and sine waves

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos nwt + b_n \sin nwt \tag{3.1}$$

or equivalent complex exponential functions of the form

$$e^{int} \tag{3.2}$$

where a_n and b_n are the coefficients that represent the amplitude of the waves and nwt represents their angular velocity.

To analyze music digitally stored as a recorded audio signal it becomes necessary to create software or hardware devices that emulate how our hearing works. The first step would be to create a tool that can detect when events happen in time and how fast they repeat. These two elements, time and repetition, appear to be fundamental parts of any music. For example, we can look at the rhythm cycles of the Indian *tala*, the key cycles of a classical sonata or rondo, a melody or motif that is repeated in its original form, intervallic relationships, etc. If we can count these repetitions and localize them in time, then we can extract a statistical analysis of the given work. In computer music, it is possible to modify the spectrum of the work and perform an Inverse Fourier Transform to create new sounds.

The standard Fast Fourier Transform (FFT) is an efficient[3] algorithm that is used to compute the Discrete Fourier Transform (DFT) which is based on the mathematical series of Fourier. Jean-Baptiste Fourier formulated a theory stating that periodic waveforms – for example sound – can be deconstructed into infinite combinations of sine waves of different amplitudes, frequencies and phases. In order for a computer to calculate the series it is necessary to discretize it into small segments. Given that the human hearing range has its limits around 22 KHz, digital sound is sampled at 44.1 kHz –Nyquist sampling theorem–. By only computing a small range of the real spectrum the infinite series is transformed into a discrete algorithm. Additional segmentation is done – *windowing* – depending on the application, and segment to be analyzed.

In order to extract information about time-allocation of frequencies, it is necessary to use the Short-Time Fourier Transform (STFT) which is essentially a DFT adapted to perform that task.[4] Although there are several algorithms to compute the FFT, all compute the DFT ; in this chapter the general term FFT will be used to describe the process of spectral domain analysis, that is the conversion of any periodic signal into the sum of its infinite sinusoidal (sine and cosine) components obtaining magnitude (energy) and phase for each bin. The FFT can be viewed as a transform that converts any finite, discrete signal into a finite, discrete sum of discretized sinusoidal components. With the FFT, it is possible to obtain magnitudes and amplitudes of the desired frequencies of an entire audio clip and look at differences and similarities across the spectrum. The output of the FFT is the spectrum. Therefore, by applying a FFT

[3]The FFT performs a fast computation of the Discrete Fourier Transform (DFT)

[4]The Short Time Fourier Transform splits the signal into equal overlapping blocks and calculates a DFT for each block which is windowed.

to an incoming audio signal we are taking the signal from the time domain to get useful statistical information in the spectral domain.

Given that any type of Fourier analysis assumes infinite periodic signals, periodicity is obtained by repeating the signal forcing it to become periodic. To obtain a small segment of the signal we could multiply the original audio by a signal of value 1 – a rectangular window – during the time period of interest. Moreover, by multiplying the original signal by a window that has a shape of a bell, we minimize spectral leakage that occurs at the beginning and end of the windowed signal that adds distortion or noise (Ramirez 1985).

Different types of windows are commonly used, for example: Hann (Hanning), Hamming, Blackman (Roads 2004). These series of windows are overlapped – shifted by n-samples and summed – for better time localization. The FFT is a function of both time and frequency. To summarize, the spectrum provides insightful information about the components of a sound which is commonly obtained in a computer by implementing a Fast Fourier Transform. With computer languages for real-time audio processing such as Max/MSP, Pure Data, Csound or Supercollider, it is possible to compute an FFT and its inverse iFFT with minimum coding. Finally, The FFT differs from its mathematical counterpart – the Fourier Transform – in three different aspects. Firstly, it applies to discrete-time sequences such as an audio recording in a digital format. This is useful for storing and manipulating a spectrum. Secondly, it is a sum rather than an integral making it easier to implement with software and hardware. Lastly, there is no need to define the function over time given that the function operates on a finite data record.

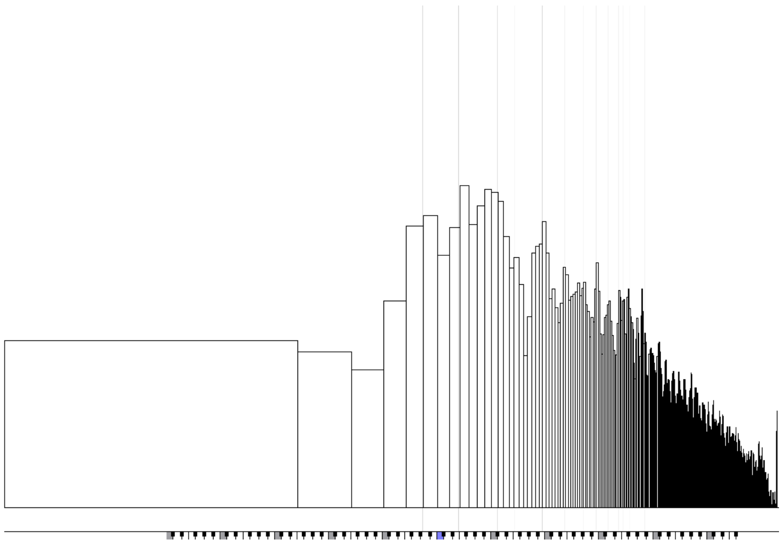


Figure 3.1: Spectrogram of an entire performance of Pierre Boulez's *Répons*

One of the uses of spectral domain analysis in contemporary music is for finding changes in texture or the gravitational center of a work in order to understand its form or key features.[5]

For the purpose of this research, a Hanning window was used to compute the FFT. In addition sound buffers containing this type of window were used in replacement of fading and panning algorithms. Different windows –not limited to the common ones– were generated using buffers in Pd and Max. Experimentation at EARS have shown that different windowing techniques affect the resultant spectrum by introducing sidebands and clutter that are translated into noise which may be desired – as part of the work's aesthetics – for some morphing techniques.

Given:

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi f t}\, df \tag{3.3}$$

we can evaluate $X(f)$from $x(t)$ with the Fourier Transform:

[5]See (Jaroszewicz 2013).

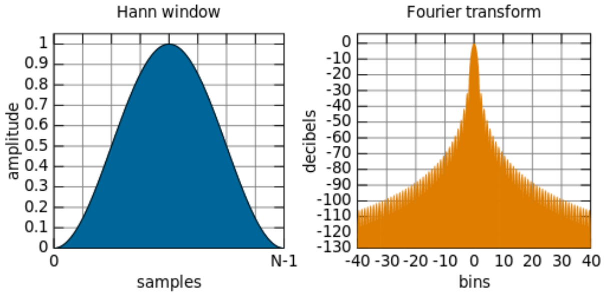


Figure 3.2: Hann Window and resultant spectrum

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{j2\pi ft}\, dt \tag{3.4}$$

The first equation $x(t)$ refers to a time-domain signal while the second one, $X(f)$, represents a magnitude of the Fourier Transform of the preview time-domain signal.

The Discrete Fourier Transform (DFT)

$$X[n] = \sum_{k=0}^{N-1} x[k]e^{-j2\pi nk/N} \tag{3.5}$$

$n = 0, 1, 2, 3..., N - 1$

returns information containing *magnitude*, *frequency* and *phase* of a discrete signal.

3.1.1 Windowing

The theory behind the FT allows us to analyze any signal or data. In order to do that, it would be necessary to apply a window the size of the work which in return will be converted to a spectrum containing information about all the frequencies present in the piece, which their magnitudes and phase providing statistical information without telling much about time or when those frequencies are located.

As oppose to real-time analyses using many overlapping windows, conserving memory and obtaining results quicker, the analysis of an entire piece using long windows applied to the desired sections could give us more information about the section itself and the opportunity to compare it to other sections of the same piece.

According to Curtis Roads:

> A spectrum analyzer measures not just the input signal but the product of the input signal and the window envelope. The law of convolution, states that the multiplication in the time-domain is equivalent to convolution in the frequency-domain. Thus the analyzed spectrum is the convolution of the spectra of the input and the window signals. In effect, the window modulates the input signal, and this introduces sidebands clutter into the analyzed spectrum.
>
> A smooth bell-shaped window minimizes the clutter (Roads 2004).

As mentioned above, windows such as Hann, Hamming, Gaussian, Blackman-Harris are commonly used for analysis and resynthesis in music. One can always try different type of windows and decide what is more convenient for the type of analysis to be done.

3.1.2 Window creation with Max/MSP

In Max/MSP it is possible to create a window by writing the values of a function into an object that stores audio samples $\boxed{\text{buffer\~{}}}$. In the following example a window of size 512 values is generated using the Hanning function:

$$\omega(n) = 0.5\left(1 - \cos\left(2\pi\frac{n}{N}\right)\right) \tag{3.6}$$

If the sample rate is 44100 and the FFT size is 1024, then a window that is half the FFT size (512) creates band sizes of approximately 86 Hz.

$$\frac{44100}{512} = 86.15 \tag{3.7}$$

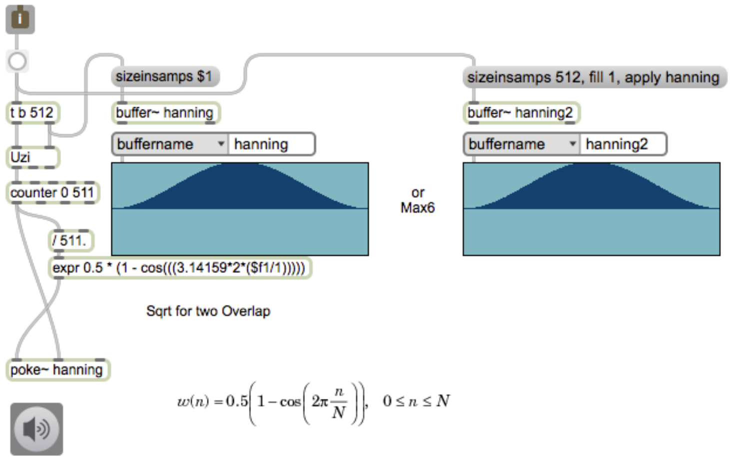


Figure 3.3: Hanning Window in Max

3.2 FFT in Max/MSP

If a window is created with a function as shown in Figure 3.3, a way to perform an FFT in Max is to use the object fft~ and multiply the incoming signal by the window. Another option is to use a patcher loaded by the pfft~subpatch object. Inside the patcher there should be an fftin~hanning object with an argument specifying the desired window type. If using the pair pfft $\longrightarrow$ fftin, three elements can be obtained from the signal: a real part, an imaginary part and the FFT bin index. These are the three outputs of the fftin~hanning object inside the pfft patch. With the real and imaginary signals then it is possible to convert to polar form and the opposite, from polar back to cartesian before the ifft~. In polar form the signal becomes magnitude and phase. The third outlet is very useful for analysis, storing or displaying a signal.

A form of spectral spatialization can be achieved by dividing the spectrum into many different channels.[6] By knowing the FFT bin index, it is possible to gate the signal to a specific buffer when a particular index has been reached. Moreover, the

[6] 8 channels at EARS.

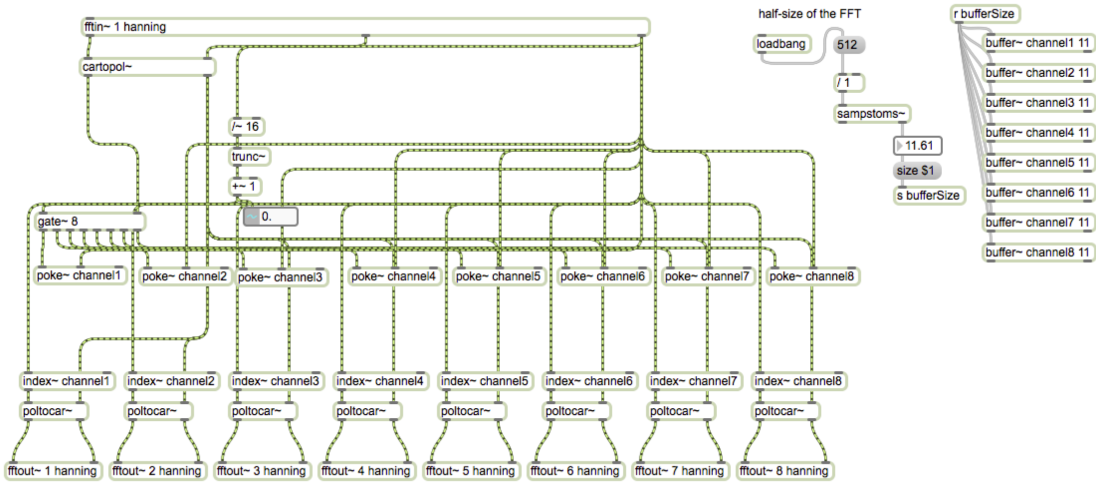


Figure 3.4: Dividing a spectrum into 8 different buffers

signal can be stored on different buffers that can be read and output to the time-domain multiplying them by their respective fftout˜hanning objects, as shown in Figure 3.4 . Each output of the pfft˜subpatch can be sent to an ambisonics encoder for further spatialization.

Inside the FFT: Cross-synthesis

Cross-synthesis describes a number of techniques that in some way combine the properties of two sounds into a single one. Convolution is a special case of cross-synthesis and serves as a bridge between time-domain and spectral-domain. In signal processing, Convolution is the multiplication of two spectra (Roads 2004, 213).

Several techniques were used and tested at EARS using a wide variety of sounds with contrasting spectral content.[7] For the purpose of spatialization, morphing between two sounds requires control over their lengths. In addition, expansion and compression (compand) in the spectral domain was implemented to match the amplitudes of signals

[7]see Appendix B.

with strong spectral energy.[8] Techniques included cross-synthesis by mixing, source-filter, convolution, cross-modulation, square-root convolution and cross-product.

The following implementations have been applied in the spectral domain in combination with cross-synthesis:

- Spectral shredding
- Spectral delays
- Spectral compand
- Spectral gates

Spectral shredding is a term applied to the technique of segmenting the spectrum into as many parts as channels are available in the system. For example, with an octagon array the spectrum of a sound is divided into eight equally[9] divided sections. This is achieved in the spectral domain after performing an FFT. Each segment is then sent to the time domain after an iFFT for further spatialization. I have found that spatialization at fast speeds creates an effect similar to applying granular synthesis, that is dividing a sound into particles that are less than 50ms in length. Spectral shredding in combination with ambisonics is a great technique for creating a virtual space where more directional sounds can be layered thus creating a complex and rich sound space.

Spectral gates aid in the control of the amount of noise that is filtered. Applying spectral noise gates before cross-synthesis creates a smoother transition between the two sounds, something that may be desired or not. Spectral delays can be used for aesthetical reasons but delaying across the spectrum with different times and feedback amounts can help perceive a smoother morphing process.

[8] Soundhack's *Spectral Shapers* including its +spectralcompand are great tools in the form of plugins for spectral processing. http://www.soundhack.com/spectral-shapers/.

[9] Depending on the nature of the sounds, one could emphasize different parts of the spectrum by dividing it into unequal segments.

Spatialization

- Circular and elliptical motion
- Paths, Spirals
- Point Clouds

Different techniques and software for spatialization had been tested at EARS. For transformations using two, three or four channels in a circular array, logarithmic panning in combination with Hanning windows for control of fading work well. In addition, a window and its inverse can be used to cross-fade between original sounds and morphed spectra. In this case windows are not used for convolution but as a control function.

Circular and elliptical motion was first tested using IRCAM's[10] Spat, ICST's ambisonics[11] objects for Max/MSP and CICM's HOA library for Pure Data. Working with ICTS's and CICM's libraries allow for a better control of specific output as opposed to the more user-friendly Spat. Most importantly, both libraries are open source and well documented.

Circular motion works very well especially when working with sounds with a short attack and fast repetition patterns and for applying morphing techniques.

Many of the sounds were previewed by using an ambisonics to binaural object.

Summary of Techniques Applied to Spatialization:

Cross-synthesis

1. Cross-synthesis by mixing
2. Source-filter
3. Convolution

[10] http://forumnet.ircam.fr/product/spat/?lang=en
[11] http://www.icst.net/research/projects/ambisonics-tools/.

4. Cross-modulation
5. Square-root convolution
6. cross-product

Spectral Techniques with Cross-Synthesis

1. Spectral shredding
2. Spectral delays
3. Spectral compand
4. Spectral gates

Spatialization

1. circular and elliptical motion
2. Free Paths / Spirals
3. Point Cloud

3.3 Cross-synthesis and Spectral Shredding

Cross-synthesis describes a number of techniques that in some way combine the properties of two sounds into a single one. Convolution is a special case of cross-synthesis and serves as a bridge between time–domain and spectral–domain. In signal processing, convolution is the multiplication of two spectra. The Law of Convolution states the following:

Convolution in the time domain is equal to multiplication in the frequency domain, and vice versa (Roads 2004).

There are several techniques used to combine aspects of two sounds. For example, IRCAM's *Audiosculpt*[12] offers three different types:

[12] *Audiosculpt* is a software for viewing, analysis and processing of sounds by IRCAM. Among other features, it offers cross–synthesis powered by *SuperVP*, IRCAM's phase vocoder analysis engine. More information here: http://anasynth.ircam.fr/home/english/software/Audiosculpt.

- cross–synthesis by mixing
- source–filter
- generalized cross-synthesis

Cross-synthesis S of two sounds S_1, S_2 with corresponding spectra M and phase θ can be calculated in several ways:

1. Summation: $S = S_1 + S_2$ Same as cross-synthesis by mixing, while the amplitude of S1 increases, the amplitude of S2 decreases.
2. Convolution: $M = M_1 M_2, \theta = \theta_1 + \theta_2$. Multiplication of magnitudes and summation of phases after FFT and conversion from cartesian to polar.
3. Cross-modulation: $M = M_1, \theta = \theta_2$. Magnitudes of S_1, phase of S_2
4. Square-root convolution: $M = \sqrt{M_1 M_2}, \theta = \frac{1}{2}(\theta_1 + \theta_2)$
5. Cross-product #1: $M = M_1 M_2, \theta = \theta_1$
6. Cross-product #2: $M = M_1 M_2, \theta = \theta_2$

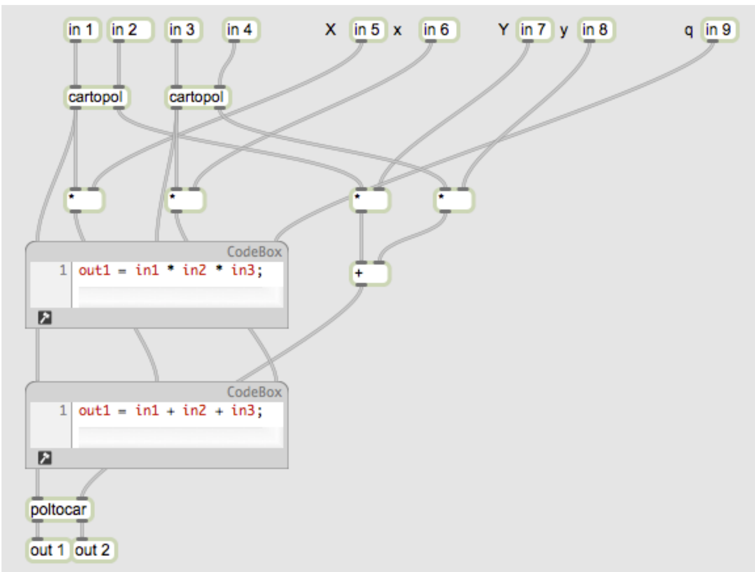


Figure 3.5: Generalized cross–synthesis with *gen*˜in Max/MSP

The first technique: summation or cross–synthesis by mixing, involves mixing of two sounds, that is cross-changing their gains in the time-domain. A spatialized summation would involve fading in and fading out the two sounds respectively. Although there is no spectral processing involved, it can give the listener the impression of sounds morphing into different sounds and space.

The rest of the techniques are performed in the spectral domain after analyzing the signal with an FFT and converting the sines and cosines to amplitudes and phases. In other words, converting from cartesian to polar coordinates. These techniques differ in the way the amplitudes and phases are combined. Operations between magnitudes are multiplicative while phases are subject to addition.

Source-filter cross synthesis requires an extra step taking the signal to the *cepstrum* domain. The word cepstrum is a rearrangement of the word spectrum.The following steps are required:

$$x(n) \longrightarrow \boxed{\text{FFT}} \longrightarrow \boxed{\text{abs()}} \longrightarrow \boxed{\text{log()}} \longrightarrow \boxed{\text{iFFT}} \longrightarrow c_x(n)$$

In Max, and especially with *gen~*,[13] it is possible to perform these operations at the sample-level in real-time. By synchronizing the cross cross–synthesis of two – or more – sounds as they move from one loudspeaker to another, it is possible to create "real-time spectral spatialization".

In addition, any of these techniques can be combined with "spectral shredding" creating spectral spatialization at different times between portions of the spectrum.

[13] Max/MSP's $\boxed{gen\tilde{}}$ compiles in real time any visually generated code. Code generated with $\boxed{gen\tilde{}}$ can be exported to C++.

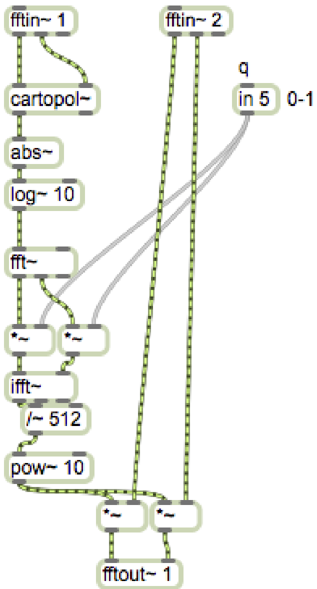


Figure 3.6: Source–filter cross–synthesis in cepstrum domain with Max

3.4 Tools and Software for Spatialization

Several aspects of spatialization and parameters can be modified in real-time using software tools.

- Speed of circular motion of a sound source including Doppler effects
- Spatial trajectories of sound sources
- Behavior of sound source e.g. swarm
- Directionality (ambisonics order)
- The creation of a broad sound image
- The creation of an acoustic environment using filters and delays.

Perceptual parameters include:

- Azimuth and elevation
- Source presence and distance attenuation
- Room presence

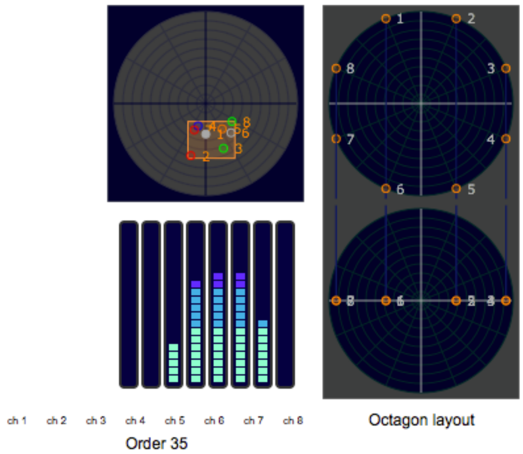


Figure 3.7: Swarm behavior of eight sources after spectral shredding

- Reverberation
- Envelopes

DSP parameters include:

- Equalization
- Doppler effect
- Air absorption
- Reverberation
- Directional distribution of various speaker layouts e.g. hexagon
- Delay

Many of these parameters can be set in IRCAM's *SPAT*, ICST's *Ambisonics Tools* and the *HOA Library*.

3.5 Some Remarks

The ability to localize sounds in space is affected by the spectral content of the sound itself, especially its attack. As lower frequency sounds tend to be harder to localize, sounds with a complex spectrum are preferred and work well for morphing. The techniques the author has implemented for his own work range from cross-synthesis of two sounds to spatialization techniques including basic rotation using an array of eight speakers, swarm behavior, trajectories and combined movements using a set of tools for parametric spatialization developed in Pure Data. See Figure 3.7.

3.6 Envelopes

From a psychoacoustic perspective, we perceive music as events and flows that can be continuous or a succession of discrete time intervals, for example: floats and integers, analog and digital. In music, the analogy corresponds to staccato versus legato or a scale versus a glissando. Dealing with events or streams can change the way the micro and macro structures are perceived. If there is a gradual transformation throughout a piece, changes may not be perceived by the listener but contrasting events belonging to the macrostructure clearly define sections.

A way to divide sections is to consider the change in direction of a signal. An envelope or *transient generator* makes a signal to rise and fall "smoothly" approaching its limits $[0, x], 0 >= x <= 1$ without any discontinuity that can cause unwanted clicks (Puckette 2007, 89). As envelopes can be applied to modulate a variety of musical parameters, when applying this concept to computer music, the main application is the modulation of the amplitude of a signal. A typical synthesizer implementation has four components: Attack, Decay, Sustain, Release (ADSR). Any continuous function can be used as an envelope and some are designed to emulate the changes in amplitude of acoustic instruments. For example, Supercollider has built-in envelopes that can be used for that purpose:

```
Env.perc.test.plot;
Env.triangle.test.plot;
Env.linen.test.plot;
Env.sine.test.plot;
Env.asr.test.plot;
Env.adsr.test.plot;
```

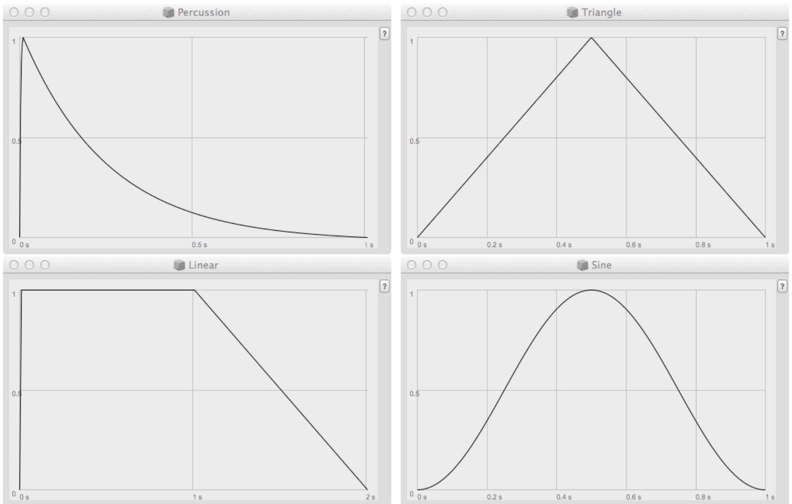


Figure 3.8: Percussion, triangle, linear and sine envelopes created in Supercollider

or one can generate a more complex envelop as shown in the code snippet and Figure 3.8.

```
({
var curve = \sine; //or \step \exp \sin \wel
var dur = 2;
var envgen, trig, output;
var env1 = Env([0,1,0,0.5,0.3,0.7,0],[dur/2,dur/2,dur/2,dur/2,dur/2,dur/2],curve);
    env1.plot;
envgen = EnvGen.ar(env1,doneAction:2);
    output = SinOsc.ar(mul:envgen) ;
}.play
)
```

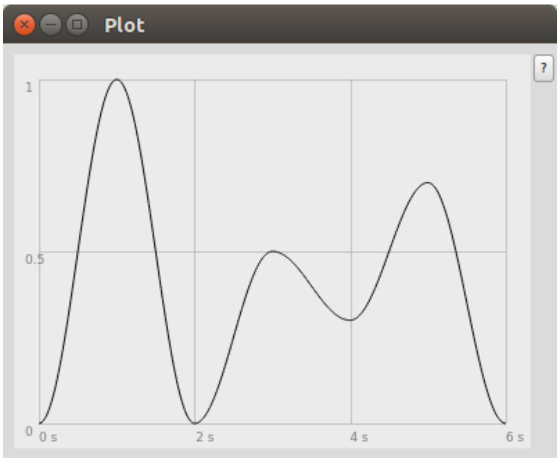


Figure 3.9: A more complex envelop applied to a sawtooth oscillator

3.6.1 The Sound Envelope as a metaphor for composition

Envelopes can be used for other purposes than modulating the amplitude of a sound in time. For example, an envelope could be used to modify parameters such as density or texture when working with acoustic instruments. The envelope becomes a composition tool. The following excerpt shows how to use a linear envelope to control orchestration. For example, measure 46 of Ligeti's *Atmosphères* (1961) can be interpreted as a Quasi-Gaussian envelope applied to density in the orchestration (see Figure 3.10). *Atmosphères* originated as an electronic piece (*Pièce électronique no. 3*) after writing two electronic works (*Glisandi*, 1957 and *Artikulation*, 1958) at the WDR studio in Cologne. The beginning of the piece gradually builds as a timbral changing cluster without any identifiable melody or rhythm: "The micropoliphonic web enters extremely softly, increasing in volume only in the last four bars. Each part has a step-wise pitch sequence except were octave shifts are necessary[..](Steinitz 2003)".

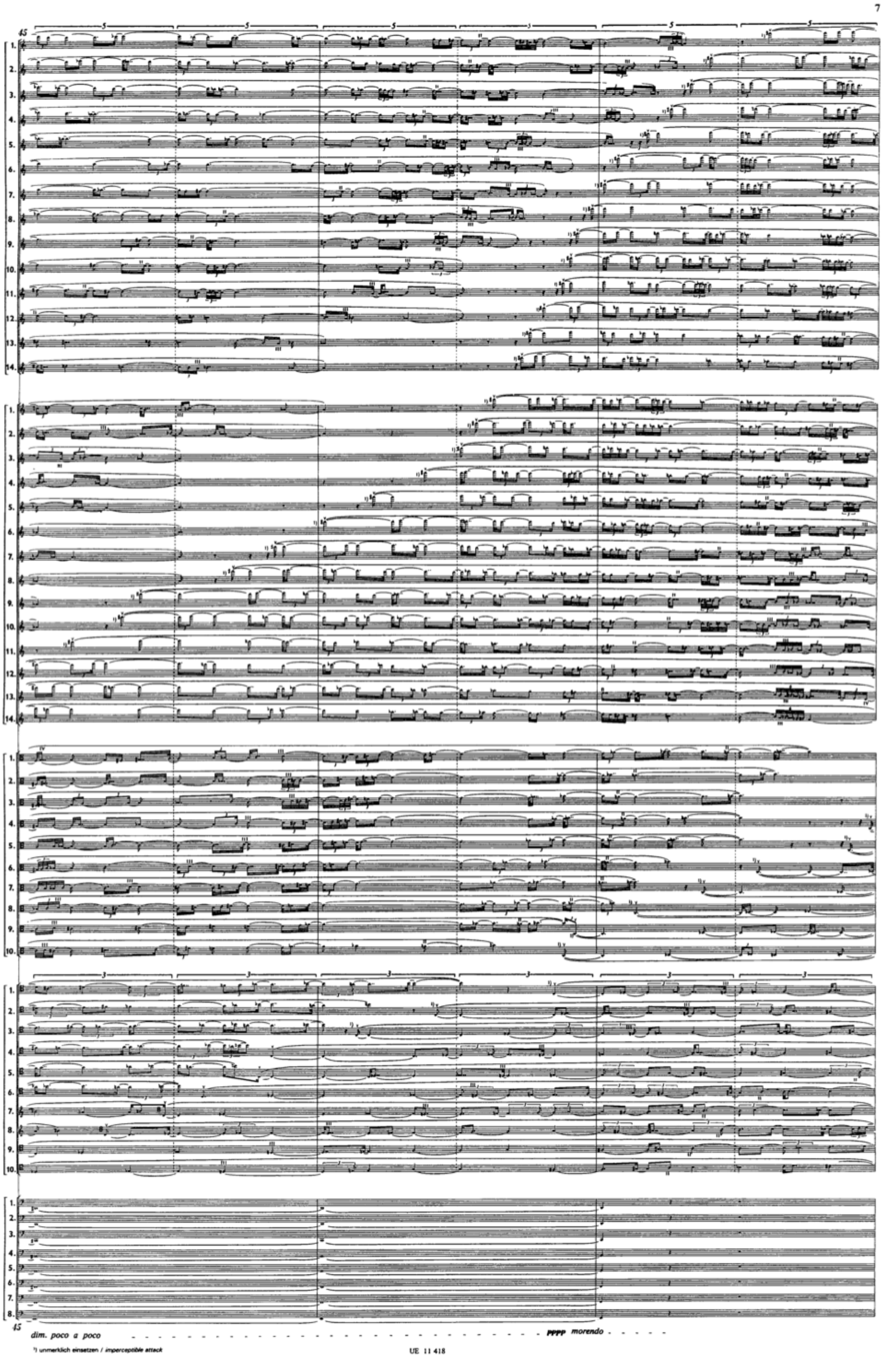


Figure 3.10: The shape of the increasing texture resembles a Gaussian envelope

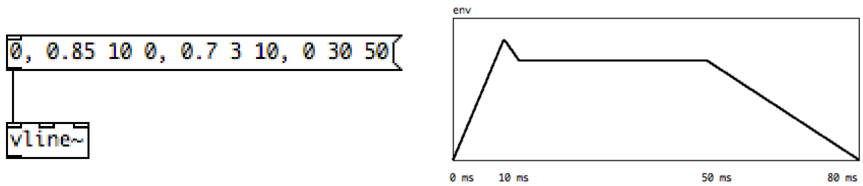


Figure 3.11: The PD patch showing the creation of an envelope (ADSR) applied to a table

In PD, the object *vline˜*, Figure 3.11 is used for signal rate control data. vline˜ is a high precision audio ramp generator that can be used to create an ADSR envelope. As in Supercollider, the object can be programmed with an array of numbers or list.

3.7 Parametric Tools for Spatialization

3.8 Parametric Spatialization

The idea of creating PD externals for spatialization came from the difficulties the composer faces when working with libraries such as Ircam's SPAT, ICST's Ambisonics Tools and CICM's Hoa Library. There is no standardized system of coordinates and/or spaces. For example, the HOA Library includes an object for the rotation of the ambisonics field. Ircam's SPAT includes examples on how to achieve different rotational figures using signals but the composition of paths has to be done with programming.

I created a series of externals for Pure Data to help the composer working with algorithms simplifying the use of trajectories by using parametric or procedural ideas. These externals specify trajectories of three different categories: circular motion, rectilinear motion and a stochastic point cloud. These can be applied to a two-dimensional or three-dimensional field such as WFS or ambisonics respectively.

The externals work well with the HOA library[14] and can be used in any environment that receives OSC data as they output coordinates (x,y,z)as floating point numbers. The system of coordinates is based on a cube of 2 units with (0,0,0) at the center with a Cartesian system of coordinates (-1..1)

[14]The HOA Library is an open source high order ambisonics spatialization tools collection programmed in C++ with externals for Max/MSP, Pure Data, included in Faust and also available as a VST plugin. It was developed by the CICM, Center for research in Computer Science and Musical Creation of the University Paris

3.8.1 HOA Library externals

3.8.1.1 $\boxed{\textbf{hoa.2d.decoder}\sim}$

Decodes an ambisonic sound field for several loudspeakers.

Arguments:

1. Ambisonics order
2. Decoding Mode:
    - 0 or ambisonics
    - 1 or binaural
    - 2 or irregular (see below)
3. Number of speakers (min= 2 * order +1)
4. Offset of speakers in degrees) e.g. [@angle] 0 30 60 90

Note that ambisonics decoding is used for a circle of equidistant speakers. For an irregular configuration there should be delay and gain compensation for each speaker. At EARS, I used IRCAM's spat.align∼ for calculating those parameters.[15] It is possible to do it manually knowing the distance to each speaker or measure with a microphone. Any configuration of speakers can be used if this method is applied. I created two abstractions that use the $\boxed{\text{hoa.process}\sim}$ external for working in PD with the HOA library:

1. $\boxed{\text{mjDelayCompensation}}$ (processes multiple delay lines in parallel)
2. $\boxed{\text{mjGainCompensation}}$ (processes the gains so all speakers are in phase according to a reference position which should be the middle of the ambisonics field)

[15]IRCAM SPAT calibration: Ircam's SPAT (Max/MSP) includes objects for calibration delay and gain: spat.gaincalibration∼ and spat.delay calibration∼.

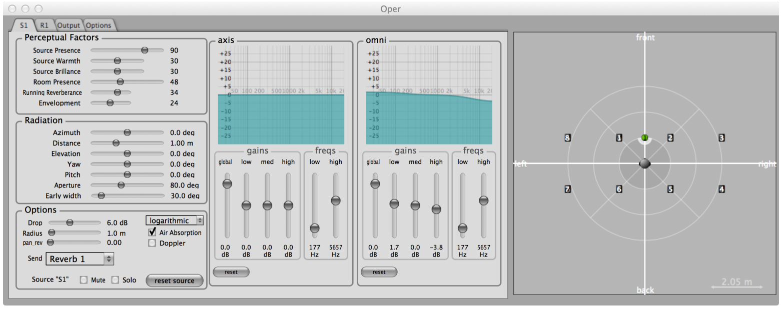


Figure 3.12: Speakers configuration at EARS

Both abstractions need a list with floats representing the values for each speaker. Delay is in milliseconds and gain is a multiplier (0–1). By applying delay and gain compensation we are creating a virtual circle regardless of the layout of the studio allowing to work with spatialization in different conditions.

mjDelayCompensation processes multiple delay lines in parallel Irregular mode is used for standard configurations:

1. mono
2. stereo
3. 3.1
4. quad
5. 5.1
6. 6.1
7. 7.1

The hoa.decoder needs phase optimization to avoid artifacts. This can be achieved by using the hoa.optim∼ external

3.8.1.2 hoa.optim∼

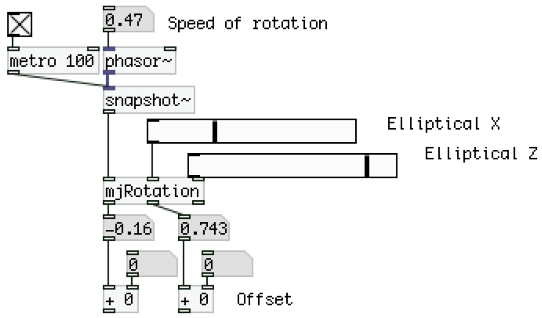
Weights the spherical harmonics for different configurations

Arguments:

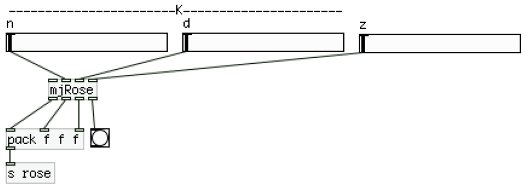
1. Ambisonic order
2. Optimization type:
   - basic (no effect)
   - maxRe (when the listener is at the center of the circle)
   - inPhase (for an audience that covers the entire circle)

3.9 Parametric Spatialization objects for PD

- mjRotation



An external that can be used for the creation of elliptical, circular or linear motion paths. The inlet parameters are *trajectory(0..1)*, *ellipticalX(0..1)* and *exllipticalY(0..1)* and the outlets output *(x,y)* coordinates in the form of floats. It is possible to create pendular rectilinear motion over the Y axis and vice versa by sending *ellipticalX* = **0** and *ellipticalY* = **1** . The trajectory parameter can be driven by a phasor∼ connected to a snapshot∼ thus sending floating point values from 0..1 to the *trajectory* inlet.




```
void mjRotation_float(t_mjRotation *x, t_float f){
    t_float a1 = x->f_x;
    t_float a2 = x->f_y;
    t_float a3 = x->f_z;
    ;
    x->f_x = sin(f * x->f_mult)*x->f_elipX;
    x->f_y = 0;
    x->f_z = cos(f * x->f_mult)*x->f_elipZ;
    outlet_float(x->l1_out, a1);
    outlet_float(x->l2_out, a3);
    outlet_float(x->l3_out, a2); // This works for HOA Library
}
```

- $\boxed{\text{mjRose}}$

A mathematical *rose curve*.[16] This is an expansion of the idea of circular motion for HOA or WFS. Given that this external outputs *(x,y,z)* coordinates, it is suitable for working with a spherical system but can be used with a planar system with $Z = 0$. the mathematical rose curve is best know by its polar equation:

$r = asin(n\theta)$,

or

$r = acos(n\theta)$.

- If n is odd, the rose is n-petalled. If n is even, the rose is $2n$-petalled.

[16] Also known as Rhodonea

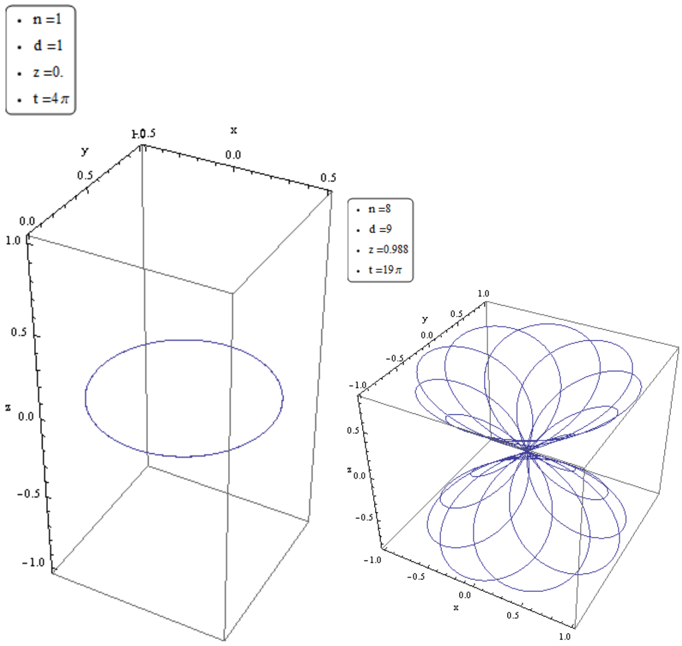


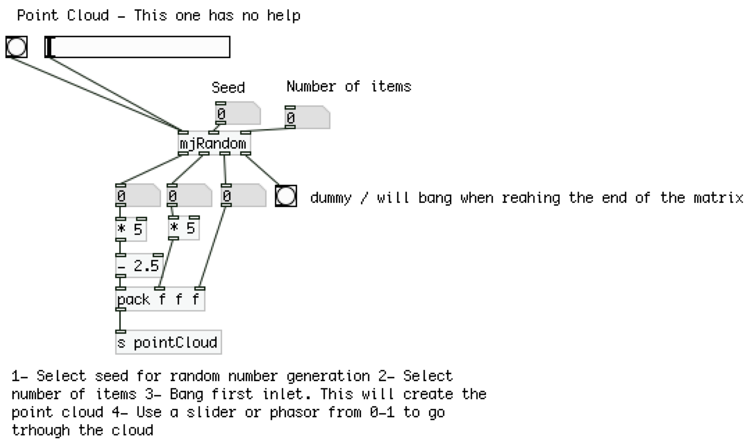
Figure 3.13: Simple and complex paths can be created with mjRose

- If $n = r/s$ is a rational number, then the curve closes at a polar angle of $\theta = \pi sp$, where $p = 1$ if rs is odd and $p = 2$ if rs is even.
- If n is irrational, then there are an infinite number of petals (Weisstein 2104).

```
void mjRose_float(t_mjRose *x, t_float f){

    double n;
    double d;
    modf(x->f_n,&n);
    modf(x->f_d,&d);
    int z = x->f_z;
    int in = n;
    int id = d;
    if(in%in != 0 && id%id !=0){
        x->f_mult = M_PI*2*d;
    } else {
        x->f_mult = M_PI*d;
    }
    x->f_position = f;
```

```
    float position = x->f_position*x->f_mult;
    float rose = sinf((n/d)*position);
    float out_x = rose*cosf(position);
    float out_y = rose*sinf(position);
    float out_z = rose*sinf(z);
    outlet_float(x->l1_out, out_x);
    outlet_float(x->l2_out, out_y);
    outlet_float(x->l3_out, out_z);
    if (out_z == 0.0 && out_y == 0.0 && out_z == 0.0){
        outlet_bang(x->l4_out);
    }
}
```

- mjRandom



Stochastic point cloud generator. The first inlet – as with the other externals – accepts a float from 0..1 to traverse the cloud. The second inlet takes a value for the seed of the random number, the third inlet selects the number of points. It generates a 3-dimensional matrix of unit 1 where $(x : 0..1)$ $(y : 0..1)$ $(z : 0..1)$. Random values are calculated using the C Macro **RAND_MAX** and the function **rand()**.[17]

The data provided by these objects is not limited to providing coordinates for sound sources. It can be used for other purposes. For example, to easily create tables without writing much code for modulating other parameters.[18]

[17] See code.

[18] See Appendix A for source code. The C files can be compile using the makefile provided with Pure Data's source, also included in Appendix A.

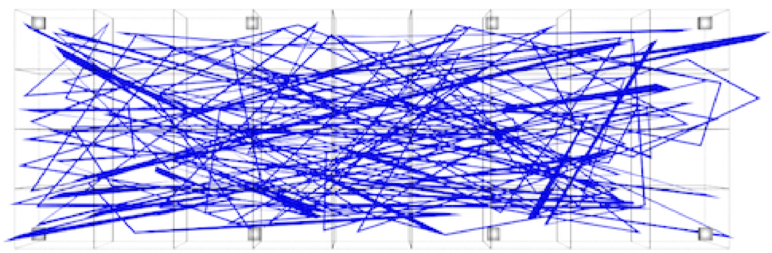


Figure 3.14: 3-D model of a point cloud at EARS

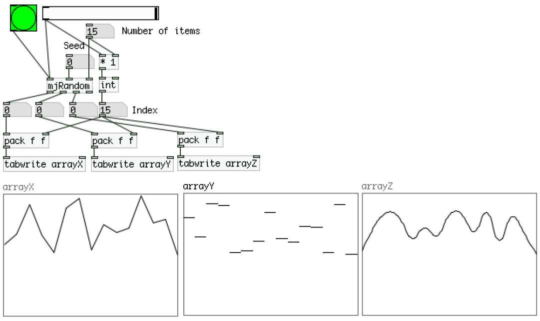


Figure 3.15

```
void mjRandom_float(t_mjRandom *x, t_float f){

    int index = f*x->i_num_points;
    post("Incoming float %f with index %d",f, index);
    outlet_float(x->l1_out, x->a_x[index]);
    outlet_float(x->l2_out, x->a_y[index]);
    outlet_float(x->l3_out, x->a_z[index]);
    if (f == 0.999){ // a phasor never gets to 1 but this does not work
        outlet_bang(x->l4_out);//change seed here
    }
}

void mjRandom_bang(t_mjRandom *x){

    float seed = x->f_seed;
    int points = x->i_num_points;
    if (points > 100){
        post("Max number of points is 100");
        points = 100;
    }
    if (points < 0){
        post("Please provide a value between 0 and 100");
        points = 0;
    }
    int index;
```

```
    float m1[points];
    float m2[points];
    float m3[points];
    for (index = 0; index < points; index++){
        m1[index] = (float)rand() / (float)RAND_MAX;
        m2[index] = (float)rand() / (float)RAND_MAX;
        m3[index] = (float)rand() / (float)RAND_MAX;
        post("x[%d] = %f",index,m1[index]);
        post("y[%d] = %f",index,m2[index]);
        post("z[%d] = %f",index,m3[index]);
        x->a_x[index] = m1[index];
        x->a_y[index] = m2[index];
        x->a_z[index] = m3[index];

    }
}
```

3.10 Parametric Spatialization using 3D Procedural Animation Software

Increasingly popular 3D procedural tools for animation are being added as plugins to animation and SFX applications such as Side Effects' *Houdini* and the Open Source *Blender*. Blender includes a game engine that can be used in real time to send – and receive – coordinates via the OSC protocol. Both softwares can be extended by using Python scripting, Blender being more flexible as it is possible to create a full application for spatialization using its engine. For algorithmic spatialization, there is an open source graphical sequencer *Iannix*,[19] which can be programmed and extended

[19] http://www.iannix.org/

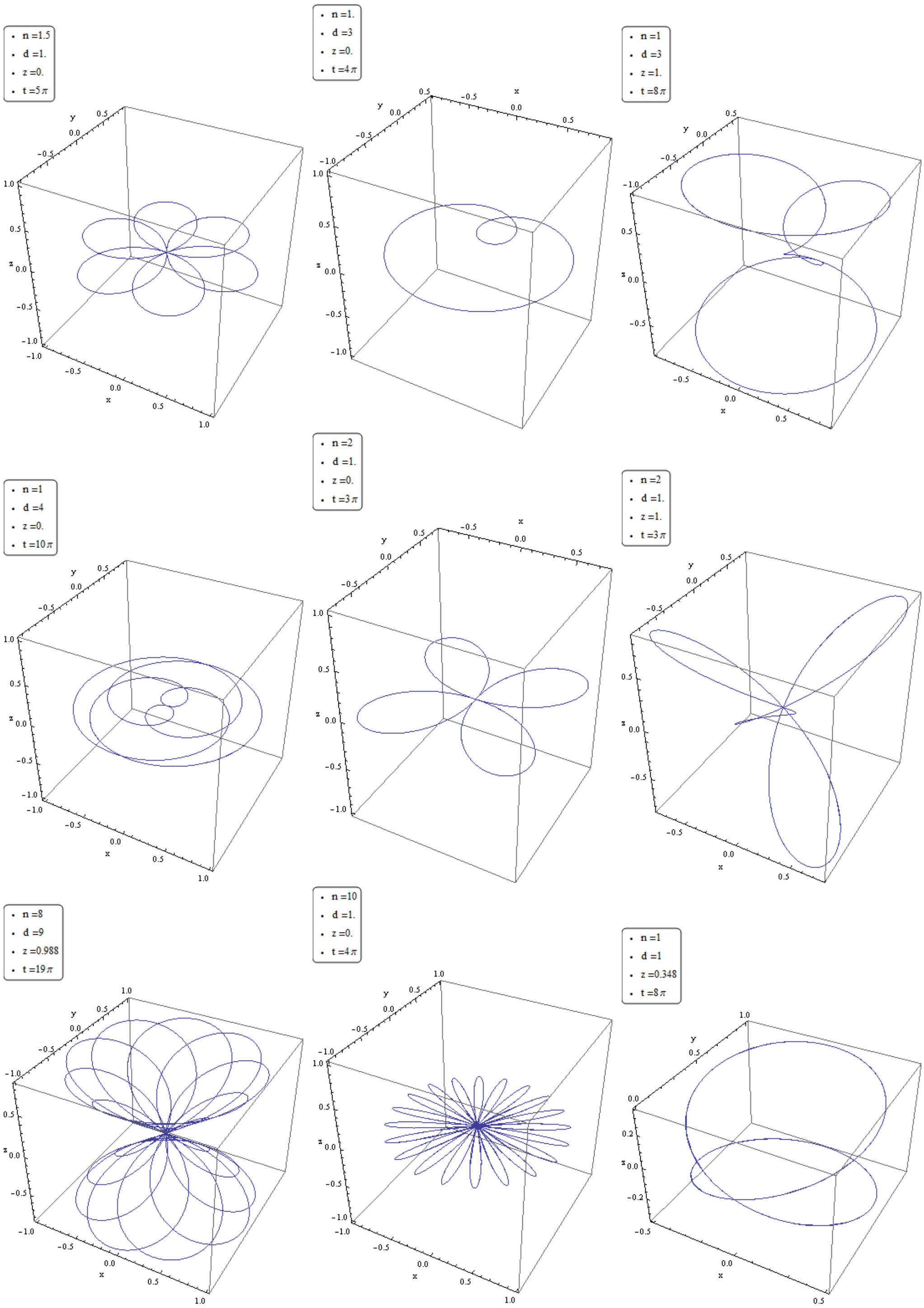


Figure 3.16: $\boxed{\text{mjRose}}$ with different values for n, d and z

with JavaScript. With Iannix it is possible to create Bezier and elliptical curves in 3D, interface with OSC and follow objects in real time.

By using Blender or Houdini, the composer can create or use any imaginable shape not limited to an algorithm. For example, one can import a 3D model of a building and use its vertices as (x, y, z) coordinates. In addition, the same algorithmic shapes can be easily created with some scripting. More complex planes and surfaces can be used as both softwares include *dynamics*, *simulations*, *deformations*, etc, as shown in Figure 3.17.

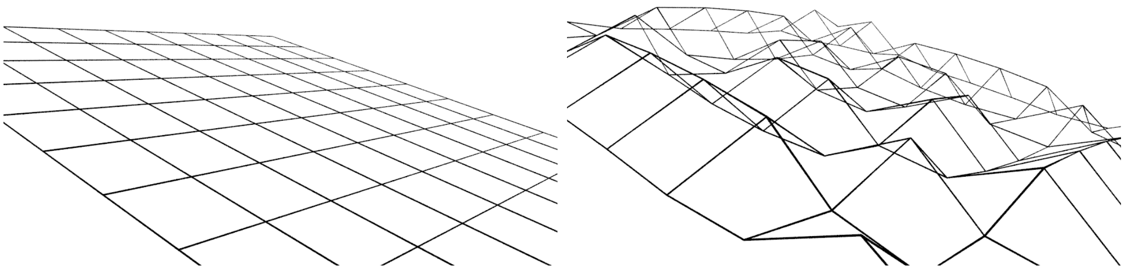
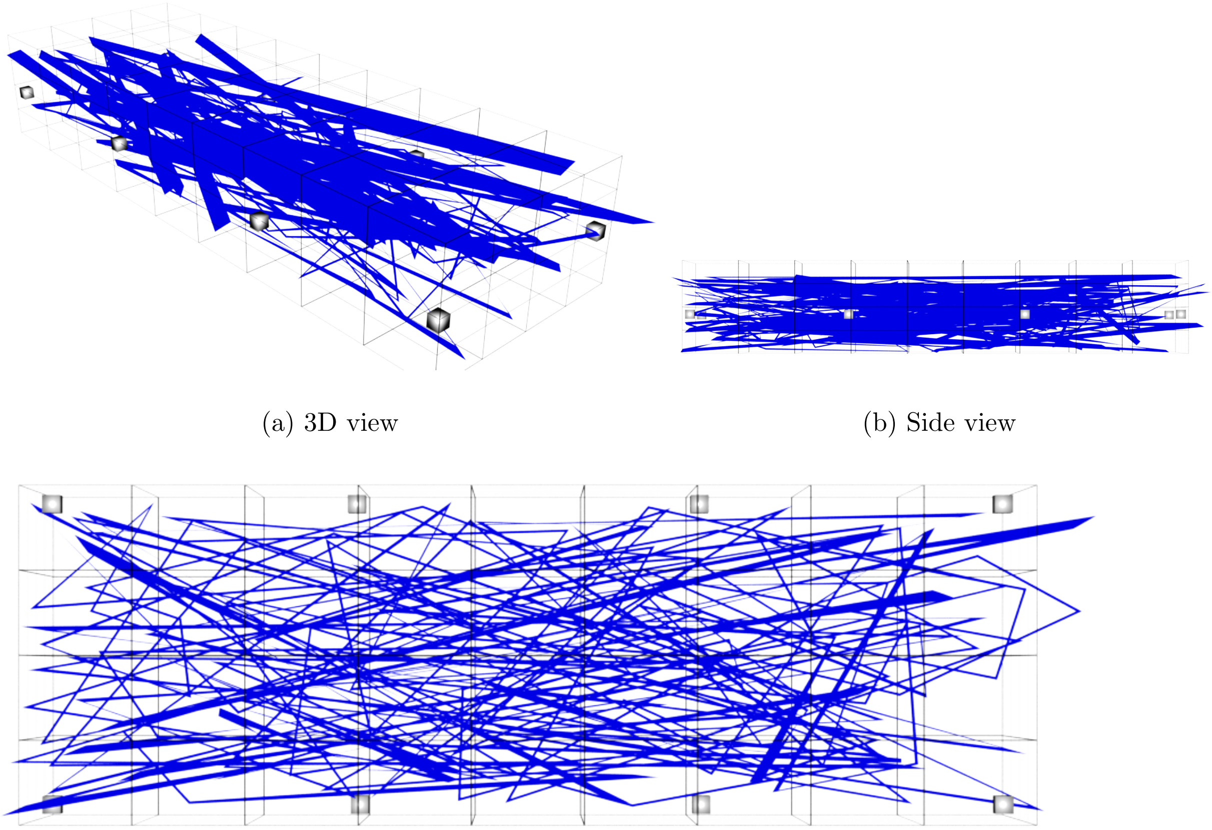


Figure 3.17: A 3D mesh (left) with a wave deformation (right) in Blender.

For non-real time work, vertices and its coordinates can be exported as Wavefront Object files (.OBJ) or other formats which can be used to populate a list or matrix in PD or Max (Murray 1996). As these type of files include more information than the necessary and the format is not readable by Max's $\boxed{\text{coll}}$ or PD's $\boxed{\text{textfile}}$ the format needs to be converted to something else. For that task, I created a command-line tool – polyConverter – for Mac that can transform Houdini *.poly* files to a text file that can be read with both Max and Pd.

The application **polyConverter** takes the following arguments *source(file) target(file) max(pd)*. For example, one can write *fileFromHoudini.poly polyForPd.txt pd*

```
if ok {
    // ------------------------------ CONVERT ------------------------------
    var isError : NSError?
    let stringFromFile: NSString = NSString(contentsOfFile: sourceFile, encoding:
         NSWindowsCP1250StringEncoding,
        error:&isError)
    var copy = stringFromFile.stringByReplacingOccurrencesOfString("POINTS", withString:"")
    copy = copy.stringByReplacingOccurrencesOfString("POLYS", withString:"")
    copy = copy.stringByReplacingOccurrencesOfString("END", withString:"")
    if conversionFormat == "max" {
        copy = copy.stringByReplacingOccurrencesOfString(":", withString:",")
    } else {
        copy = copy.stringByReplacingOccurrencesOfString(":", withString:"")
    }

    copy = copy.stringByTrimmingCharactersInSet(NSCharacterSet.whitespaceAndNewlineCharacterSet())
    copy = copy.stringByReplacingOccurrencesOfString("\n", withString:";\n")
    copy = copy.stringByAppendingString(";")
```

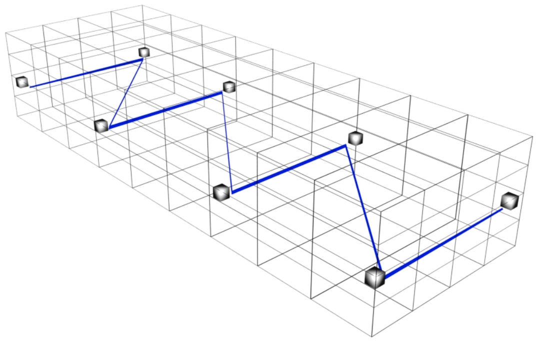


(a) 3D view

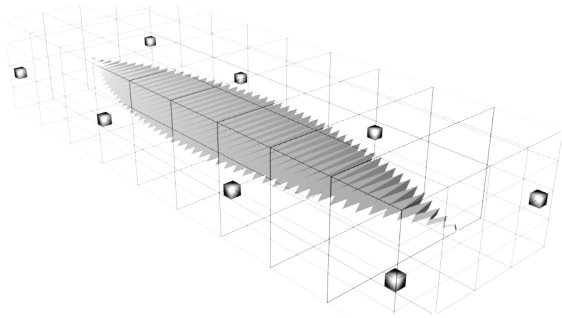
(b) Side view

(c) Top view

Figure 3.18: Different views of a point cloud. In addition to the creation of curves using parameters or manually drawing curves, 3D software allows the composer to visualize the curve in real time from any angle and modify any vertex or point as desired.



(a) Lines



(b) Helix

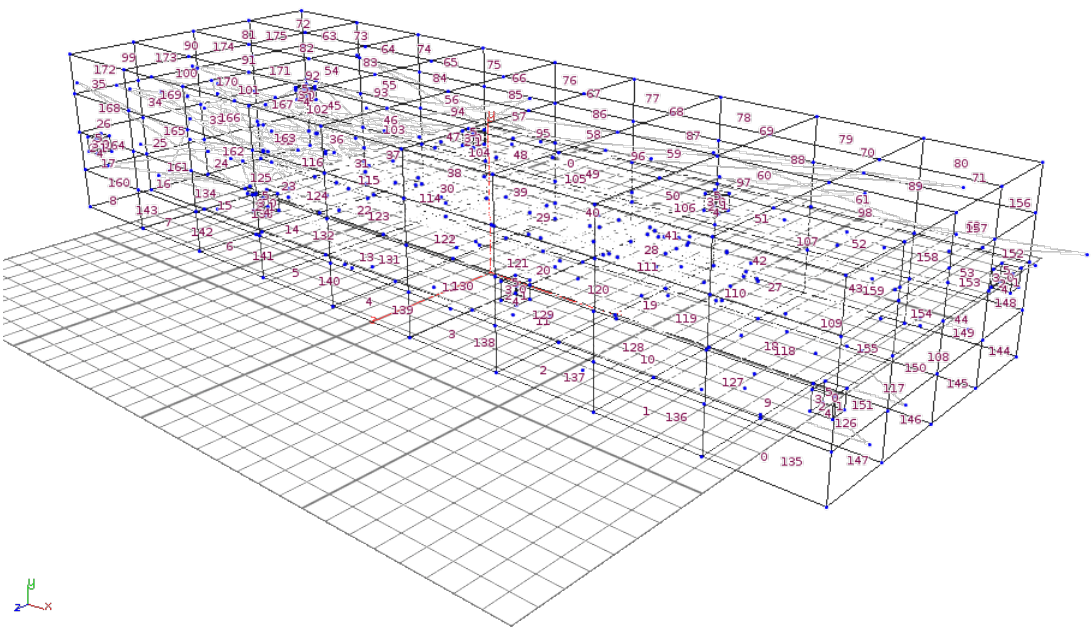


Figure 3.19: Point Cloud

Figure 3.20: Different paths created with Houdini and Python. See Appendix A for source code.

Part III

Laberintos

Chapter 4

Laberintos

This chapter presents s series of works inspired by Borges' short story *The Library of Babel*[1] and the book *The Unimaginable Mathematics of Borges' Library of Babel* by William Bloch (Bloch 2008). In his book, Bloch analyses the possible numerical outcomes of the story which he uses for the purpose of teaching some mathematical concepts such as Information Theory, Topology and Combinatorics. Most interestingly is the chapter *Geometry and Graph Theory* where he describes how the library in Borges' story could look like with its hexagonal and spiral shapes.

> The next observation is that it's conceivable the floor plan of a level of the Library may look like the preceding illustrations, in which the paths run straight through the hexagons. However, it is consistent with the text – and the atmosphere of the story – that the corridors weave and spiral around symmetrically or chaotically (Bloch 2008, 97).

As a composer, the infinite permutations and the paths were the main motivation behind the music and the research conducted at EARS. Beginning with the hexagonal shapes of the library, one can create a path that traverses each room as shown in the figures below

[1] *La biblioteca de Babel* (Borges 1997).

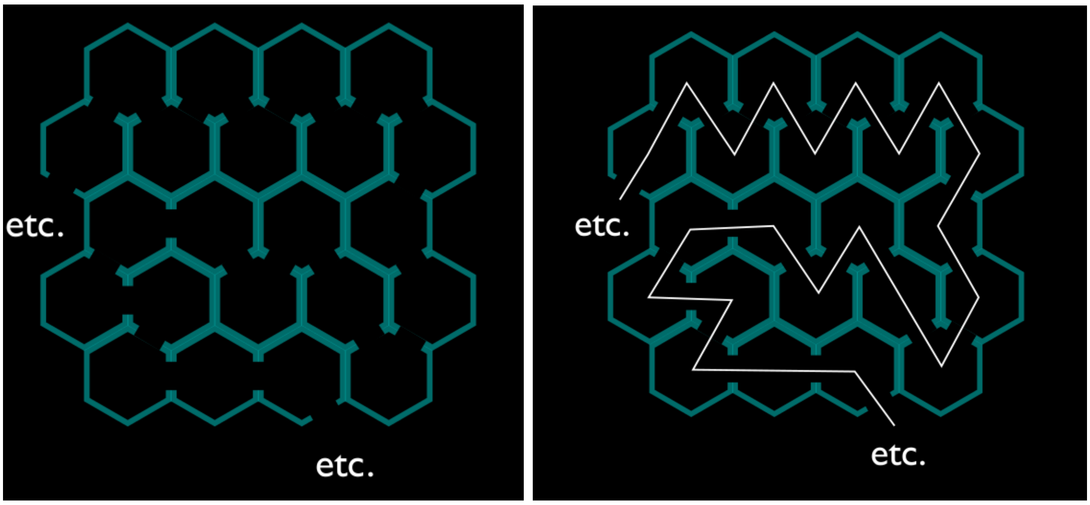
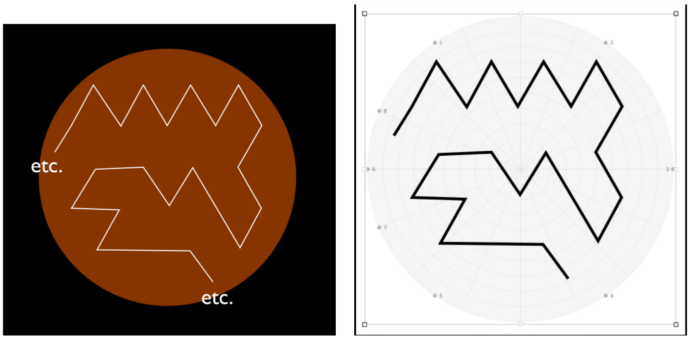


Figure 4.1: A contorted passage and path through the library.

This idea of an algorithmically designed path was the main inspiration that led me to pursue research in sound spatialization.[2] Firstly, and to be more consistent with the story, six speakers in an hexagonal array surrounding the listener were used. The space could be virtually rotated and transformed in real time. I think of parametric generated shapes and their permutations as a new form of working with space in music and a way to connect sound spatialization to form thus breaking away from the aesthetics of circular motion, at least from a conceptual point to view. The paths in Borges' story became sound trajectories in an ambisonics space and from there other curves were created and new ideas were generated.

All the etudes were originally conceived from Borges' stories and realized using the theory and technique of spatialization, envelopes, transformations in the spectral domain and other computer music techniques. *Laberinto 5* is an audiovisual algorithmic composition with the strongest connection to Borges' story; Laberinto I, for flute and live electronics uses point cloud spatialization; Laberinto 3, for Clarinet and live elec-

[2]The path is a metaphor for the trace left by a librarian in Borges' story. The laberynthic geometry of the hexagons and plausible paths are discussed by Bloch (Bloch 2008, 99).



(a) A path (b) Ambisonics trajectory

Figure 4.2: A path applied to an ambisonics field

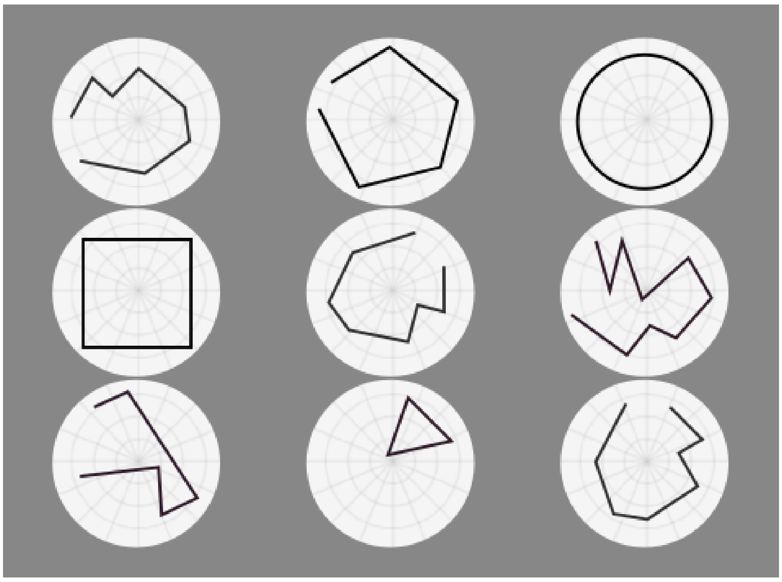


Figure 4.3: Some trajectories used in *Laberintos*' ambisonics field.

tronics, uses envelopes and circular motion with delay lines to create an infinite spiral of sounds; Laberinto 4, for piano an electronics, uses granular synthesis and algorithms as a metaphor for the Minotaur and the Labyrinth in Borges' *The House of Asterión* . [3] In *Laberintos*, labyrinths become paths and sound trajectories in which the listener can discover his own labyrinth or a meaningful connection between movement and the music.

[3] La casa de Asterión.

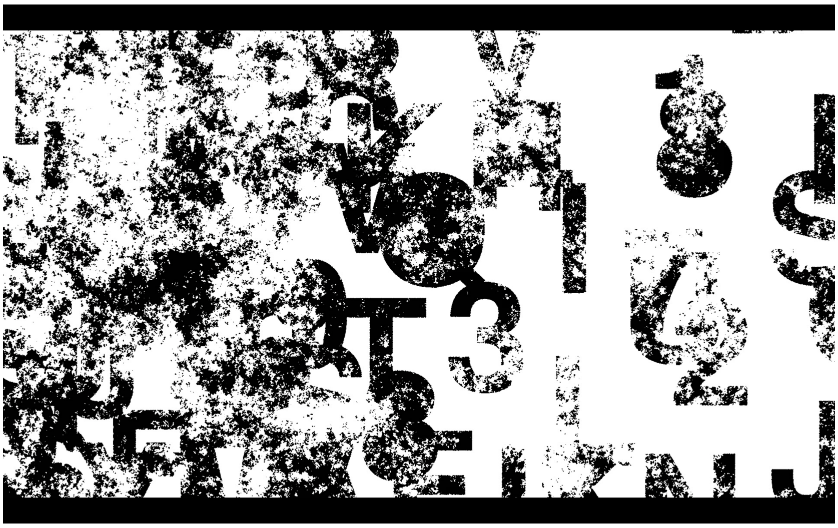


Figure 4.4: Laberinto V screenshot

4.1 Laberinto 5

Audiovisual composition. 5.1 Surround

Laberinto 5 (2014),[4] the last of the pieces in the series *Laberintos*, is an audiovisual algorithmic composition that include elements that reflect the beginning and the last stages of the research conducted at EARS. The beginning, because it incorporates the metaphor of the Labyrinth with its infinite permutations as imagined from Borges' story. It is also a work written for 5.1 surround which is somehow a primitive form of spatialization but developed with the algorithmic tools – computer music and 3D software – I investigated during the last years.

In the story, the Library of Babel is an endless universe made up of small hexagonal rooms. On the shelves of these rooms are books and all the books put together contain **every possible permutation** of single characters –all the letters of the alphabet and punctuation marks–. Most books are complete nonsense but some of the books contain true and accurate descriptions of the universe, including everything

[4]Laberinto 5. Stereo. http://vimeo.com/100868756.

that has ever happened and everything that will happen. Departing from the idea of infinite permutations of characters, I used a random algorithm in Houdini to access a list with Western characters. Those characters were also randomly put in a matrix for visualization in a 3D space. Visual effects were used to transform each character into particles representing the starts of the Universe.

The sounds in *Laberinto 5* were created using the technique of granular synthesis[5] with Csound, Max/MSP and mixed with ProTools. Firstly, phonemes were recorded and converted to samples cutting unnecessary silence to make them all the same length, and to optimize the files for granular synthesis. Using a random algorithm, sections of each sample were randomly selected using parameters such as the *sample index*, *size* and *duration*. The small "pre-grains" were then recorded and sent to Csound were granular synthesis was applied.

One of the interesting aspects of the piece is the camera movement done in Houdini. The letters only move forward and there is a slow motion rotation that is more perceivable in the second section. The reason behind such unusual way to move the camera without having objects moving from left to right –and vice versa–, as one might see in a film, is to focus the audience's attention on the movement of the sound as the spatialization was carefully composed.

The piece is divided in two sections. The first section begins with a short introduction where vertical stripes of black against a bright background move back and forward in pendulum motion. The lines dissolve into recognizable characters which move towards the viewer at a fast rate. The scene is saturated with indistinguishable sounds which occupy all 5 channels. Although the spatialization is minimum and the visuals are

[5] Granular synthesis is a technique that uses sounds that are between 10-100ms and windowed by a function. See Roads *Microsound* (Roads 2004).

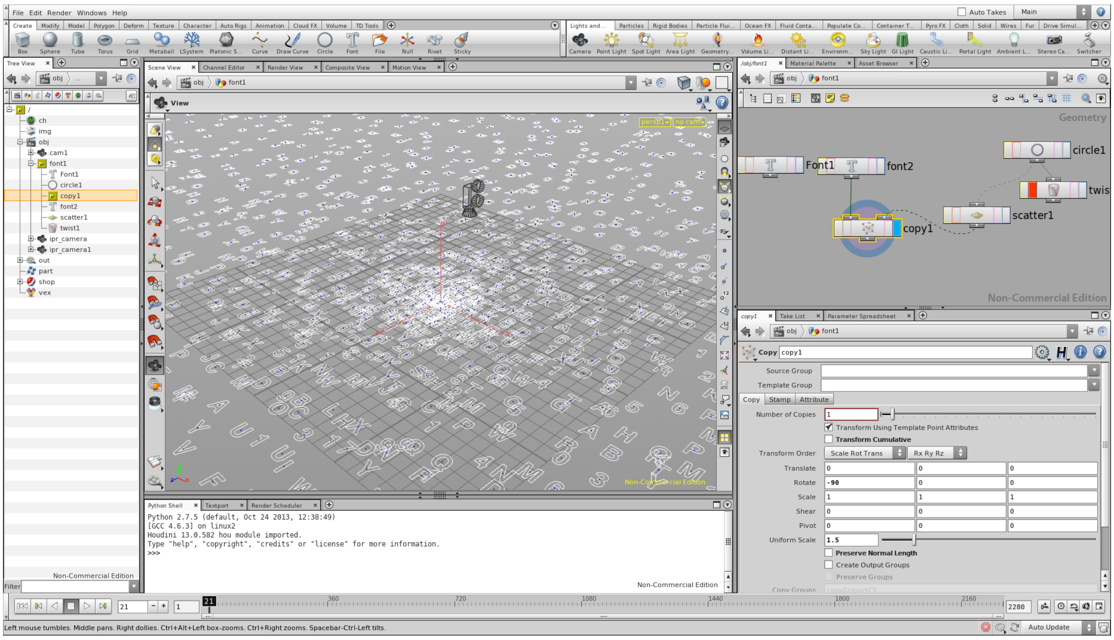


Figure 4.5: Algorithmically created letters in Houdini.

not synchronized with the sound, the granular synthesis creates a strong connection with the visuals as the stripes are composed of small particles. The contrasting second section is dark and slow. Small particles move counterclockwise at a slow rate slowly revealing as text characters. The movement is synchronized with the sounds and spatialization is noticeable and surrounds the listener. The visuals accompany the sound and not vice versa. Grains are of longer duration and although they change fast, one can understand they are phonemes. As time move forward layers are added and the particles on the screen reveal as letters as they become closer to the viewer. Towards the end, sounds and movement –like in the first part–saturate the medium. Finally, sounds abruptly stop and there is a short fading out of a single character that occupies the entire screen.

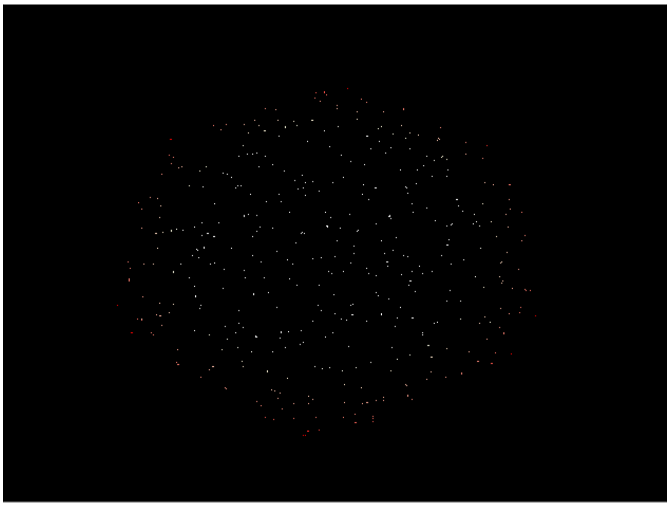


Figure 4.6: Laberinto V Frame.

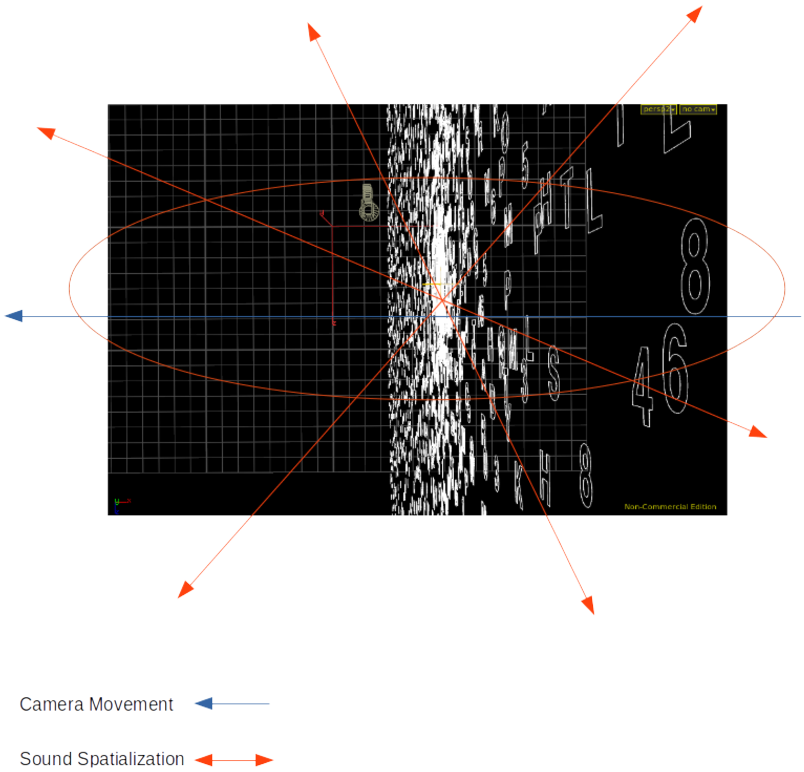


Figure 4.7: Images only move forward while sounds move in all directions.

Summary of techniques

- Sampling
- Granular Synthesis
- Stochastic algorithms for both images and sound using /Python
- 5.1 Surround

4.2 Laberinto I

for flute and live electronics. 8 channels The first piece of the series and a short spatialization étude for flute or piccolo flute. It features an assortment of extended techniques including multiphonics, whistle tones, key clicks, microtonal oscillations and different sounds produced by variations in the position of the instrument's embouchure. The processing and spatialization is in real time requiring a live input from a microphone to be connected to the sound interface. See Diagram 4.1.

$$\boxed{\text{Flute}} -> \boxed{\text{Mic}} -> \boxed{\text{ADC}} -> \boxed{\text{Sound Processing}} ->$$
$$-> \boxed{\text{Spatialization}} -> \boxed{\text{DAC}} -> \boxed{\text{Mixer}} -> \boxed{\text{Speaker Array}} \quad (4.1)$$

A good quality microphone with a cardioid pattern and a frequency range of 20Hz-20KHz is preferred[6] to avoid noise from the performance space and feedback from the speakers. Several patches were made using PD and Max/MSP to find which one could do better pitch tracking. Best results were obtained in PD with the $\boxed{\text{sigmund}\sim}$ object and Ircam's $\boxed{\text{yin}\sim}$ in Max/MSP.

$\boxed{\text{yin}\sim}$ uses the YIN algorithm which is a fundamental frequency estimator for speech and music based on the autocorrelation method[7] .

The pitch tracing object is used to split the signal in two according to a threshold. Detected clicks and high pitch sounds are sent to a granular process *(g)* and lower pitches to a harmonizer *(h)*. If using a flute, frequencies below 500Hz – first octave – are sent to the harmonizer. The idea is divide the spectrum and apply different synthesis and spatialization techniques *(p)*. An ambisonics field is created with the noise *(n)*

[6]Such as the DPA 4011A. Cardioid microphones have an angle of rejection of 180°.

[7]The autocorrelation method compares the signal to its shifted self. It applies a Fourier transform of the power spectrum and can be seen as measuring the regular spacing of zcs within that spectrum. (De Cheveigne 2002).

Figure 4.8: Two excerpts. The first shows micro-glissandi (long sounds); the second, short sounds. These gestures are spatialized using rotation and point cloud respectively.

generated by the extended techniques. See Diagram 4.2. Both processes are sent to a spectral delay before the spatialization.

$$ADC{-}>sigmund<(p,n):(p:h<:$$

$$(long{-}>\boxed{\text{mjRotation}},short{-}>\boxed{\text{mjRandom}}),$$

$$n{-}>g) \qquad (4.2)$$

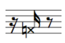
The spatialization can be done in real time using the $\boxed{\text{hoa.mapl}}$ object or algorithmically with the *mj* Library if working in PD. For this piece, I have chosen to use rotation for longer sounds with micro glissandi, and point cloud spatialization for short sounds.

Martin Jaroszewicz

Laberinto I

for piccolo ute and interactive electronics

LEGENDA - *LEGEND*



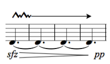
key clicks

gradual accelerando

random/approximate key clicks

pizzicato

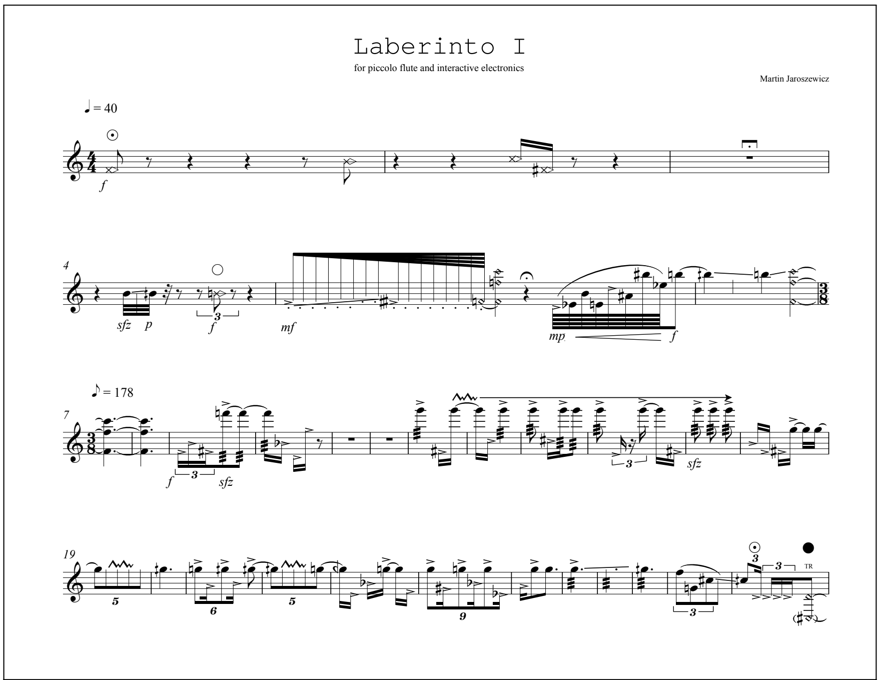
pizzicato with key click

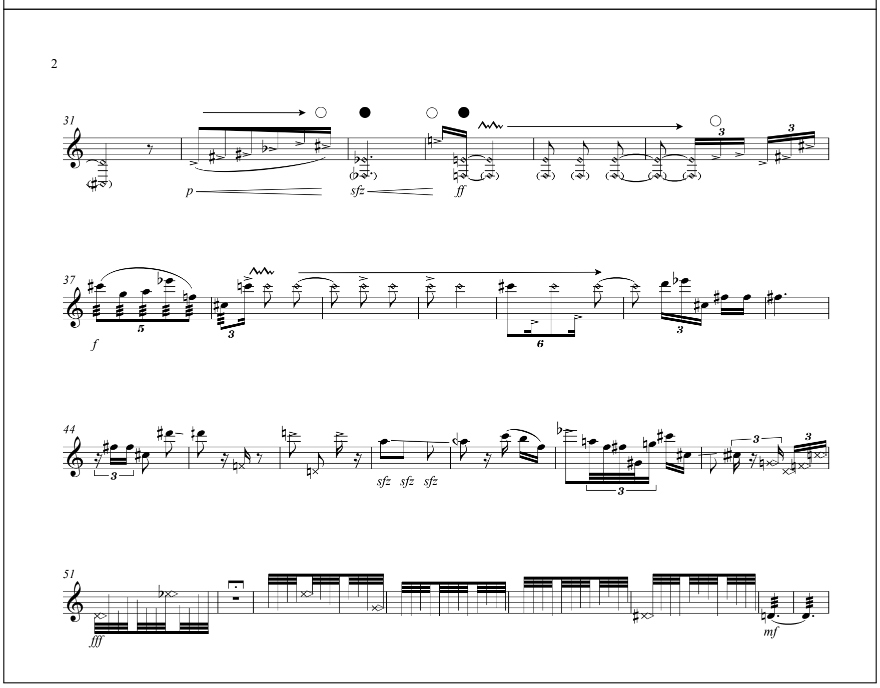


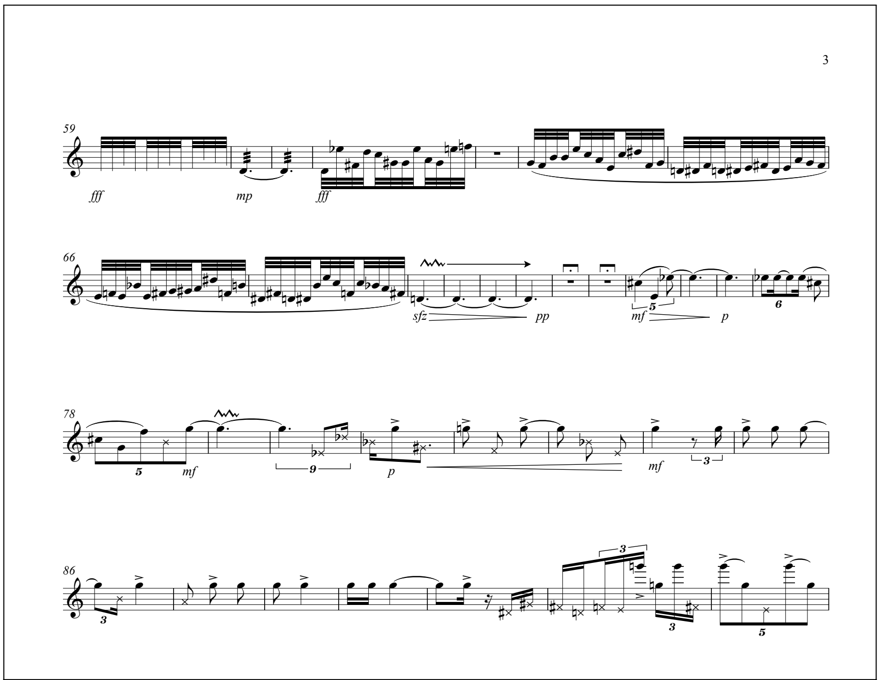
Microtonal oscillation (more or less fast) by embouchure rotation

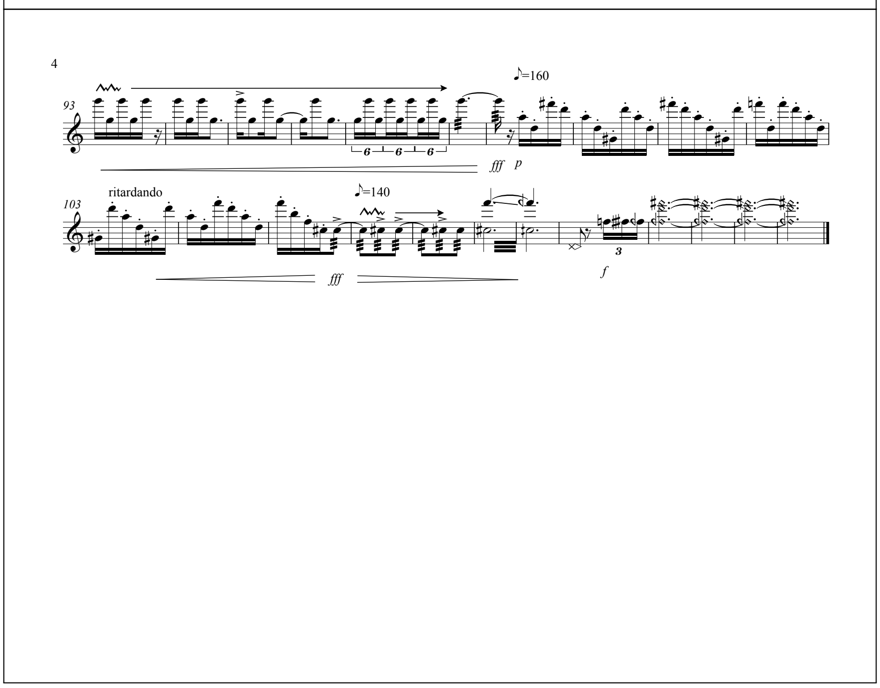
○ ⊙ ●

Embouchure open, close, and tongue in mouthpiece


Laberinto I
for piccolo flute and interactive electronics
Martin Jaroszewicz
♩ = 40
♪ = 178
sfz p f mf mp f sfz sfz TR


2
p sfz ff f sfz sfz sfz fff mf


3
59
fff
mp
fff
66
sfz
pp
mf
p
78
mf
p
mf
86


4
♪=160
93
fff
p
103
ritardando
♪=140
fff
f

4.3 Laberinto III

for Bb clarinet and live electronics. 8 channels The piece is divided into two sections: A, a meditation with long notes that resembles the playing of a *shakuhachi* flute[8]; B is a contrasting section with long chromatic lines of fast motion interlocked with multiphonics and *Flatterzunge*. The processes are simple, there is a slow rotation in the first section with each long note transposed at the micro level using envelopes connected to a real-time pitch transposition module that uses a vocoder. The fast section is a dense spiral of sound that moves upwards, an effect that is achieved by using spectral delays and the coordinates of a spiral curve exported from Houdini[9] and spatialized using ambisonics. The ideal format for this pieces is a high order ambisonics field with a hemispherical dome or multilevel rings.

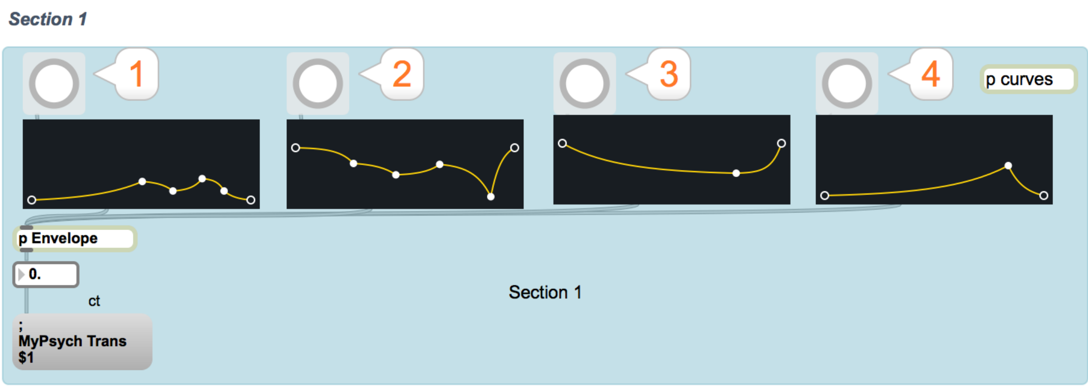


Figure 4.9: Envelopes used for micro-pitch modulation

[8]A *shakuhachi* is a Japanese flute made of bamboo introduced from China in the 8th Century. It was used by the Fuke monks of Zen Buddhism.

[9]This can also be achieved by using mjRotation with an offset on the y axis driven by a phasor and changing the radius of the circle as the piece progresses.

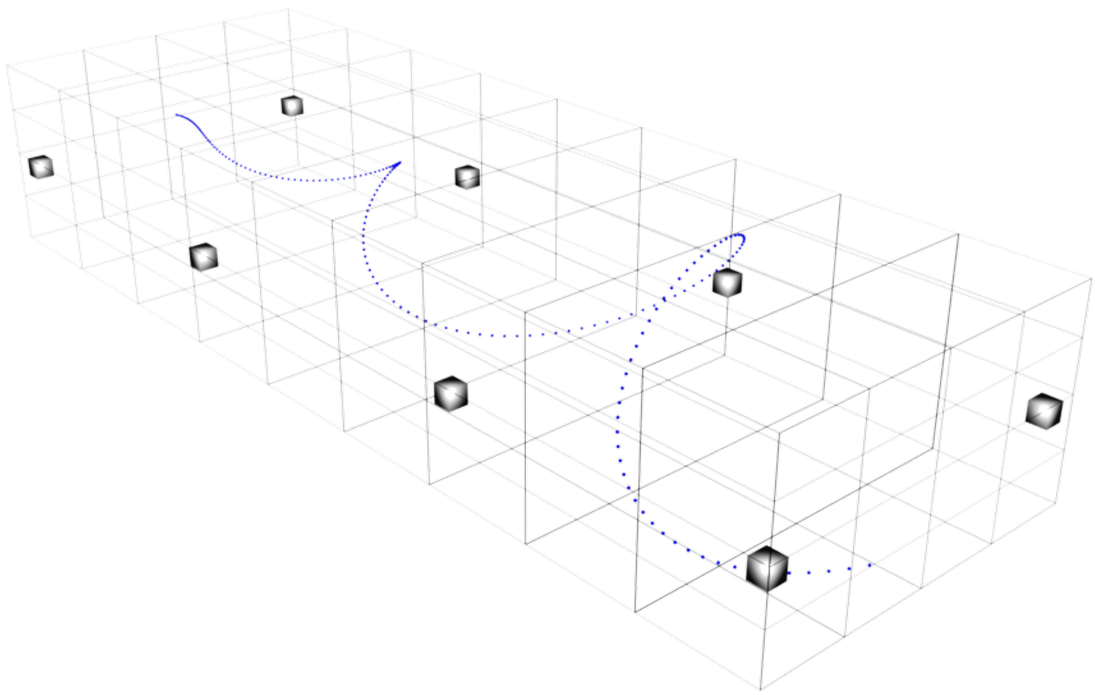


Figure 4.10: A spiral curve used in Laberinto III

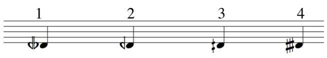
Martin Jaroszewicz

Laberinto III

for clarinet in Bb and interactive electronics

LEGENDA - *LEGEND*

n.v.	non vibrato
v.	vibrato
⟶	progressive transformation
s.g.	slow glissando



1. three quarter flat
2. quarter flat
3. quarter sharp
4. three quarter sharp

Multiphonics (in parenthesis) are notated as sounding.

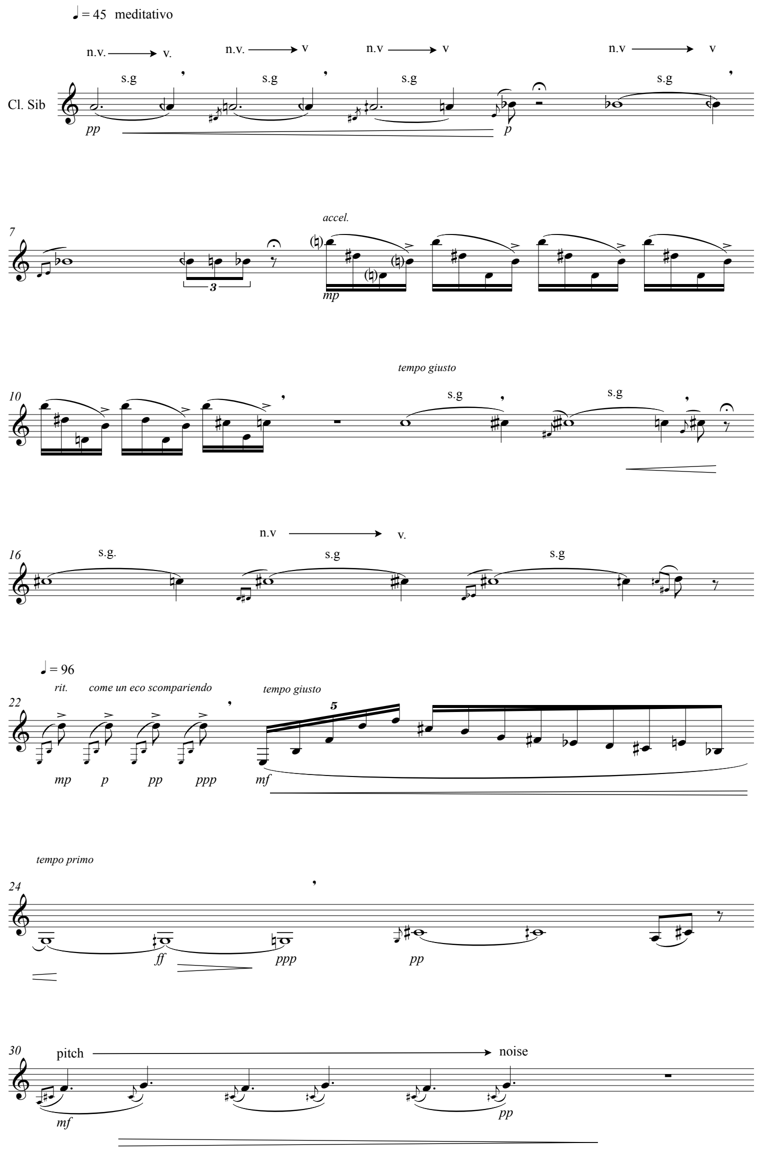
The piece starts with a mediation in the style of traditional Japanese Shakuhachi music. The durations are not strict and more or less related to the breathing of the performer.

All the chromatic passages should be executed *legatissimo.*

The gestures in the second part of the piece marked f*loatingly, ethereal*, are analyzed and transformed by the electronics creating a series of spectral delays that follow the performance. Gradual *accelerandi* and *decrescendi* and gradual changes in dynamics are recommended to create a sense of a timeless moving object.

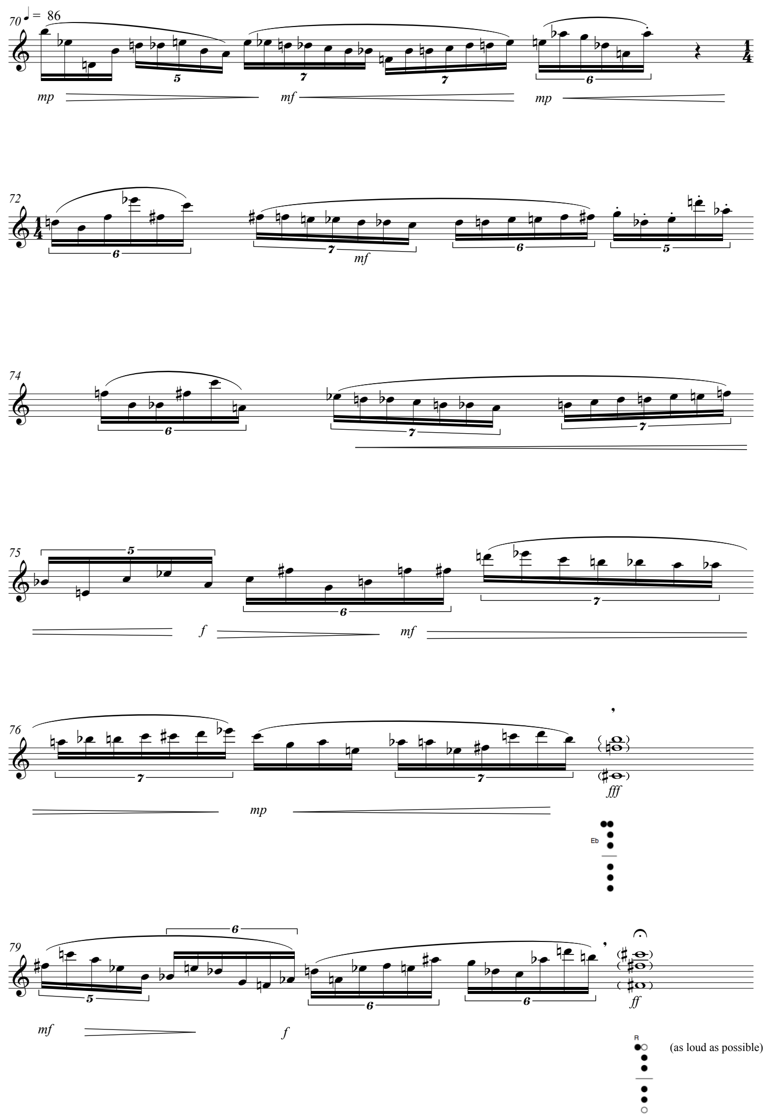
Laberinto III

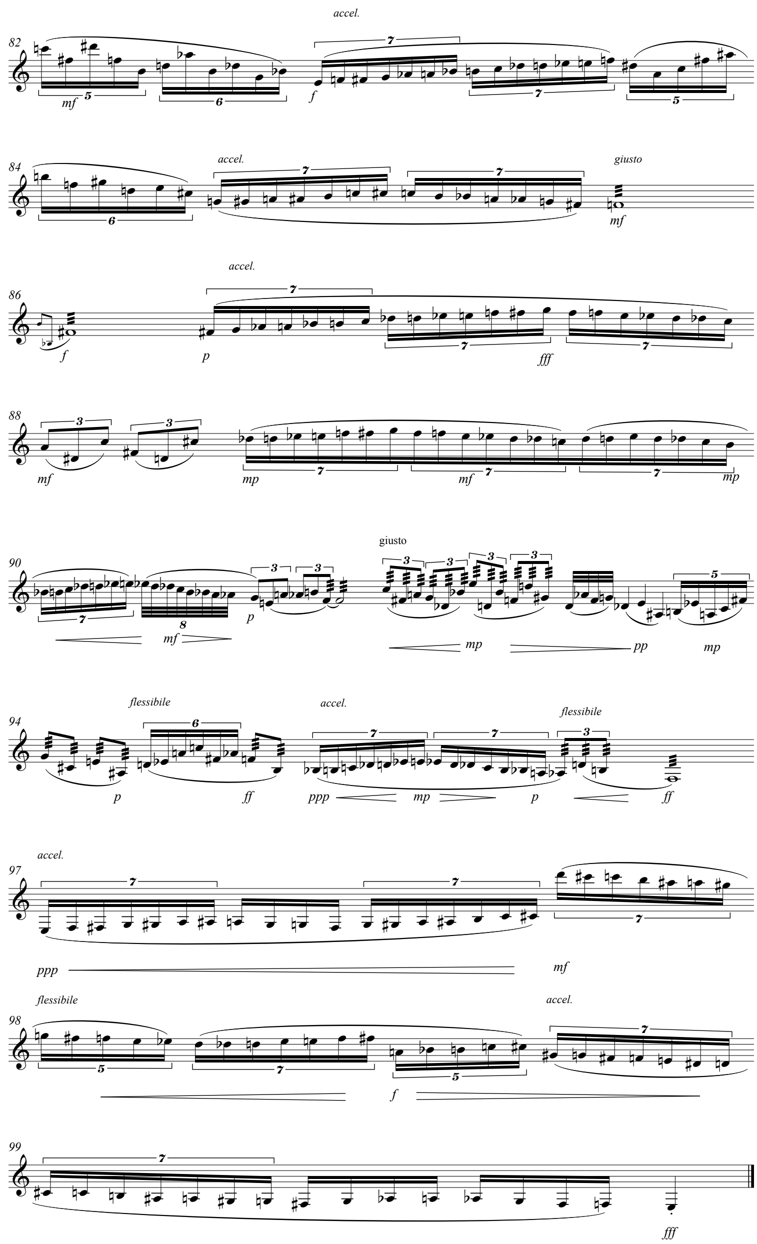
Martin Jaroszewicz




pitch
noise
(audible)
♩ = 96
tempo primo
♩ = 96
accel.
tempo giusto
rit.
tempo giusto

Floatingly, ethereal




82
mf
accel.
f
84
accel.
giusto
mf
86
f
accel.
p
fff
88
mf
mp
mf
mp
90
mf
p
giusto
mp
pp
mp
94
flessibile
p
ff
accel.
ppp
mp
p
flessibile
ff
97
accel.
ppp
mf
98
flessibile
f
accel.
99
fff

4.4 Laberinto IV

for piano and live electronics. Stereo *Laberinto IV* is based on the piece *Laberinto Borgiano* inspired by Borges short story La casa de Asterión[10] and the formation of globular clusters in the universe.*Laberinto Borgiano* is a metaphor for a unicursal labyrinth, a type of Greek labyrinth that hold the Minotaur and a metaphor for the creation of the universe and the composer's creative labyrinths. These two elements: the man and the bull, are developed and expanded. The first one is gradually transformed and leads into the B section of the work where the sense of a path is lost. The second short motif (the entrance to the labyrinth) becomes the labyrinth itself which reveals itself at the last section of the piece gradually moving towards the center of the labyrinth where the universe ends.

The piece was constructed through the concept of duality which is found at every level in its form. The first two measures contain the material that is developed in different sections of the work. At the immediate surface level, there are two motives that are expanded and transformed into sections respectively. There is an introduction, where the material is presented, and there are short transitions to the main sections of the piece. The concept of duality is furthered explored by trying to create independence of the hands and the feeling of more than one piano, or two hands, being heard.

Laberinto borgiano is a work that constantly juxtaposes two different ideas without trying to create a sense of linear counterpoint or polyphony. The idea of duality is present throughout the work and is presented at the beginning with two different motifs that become the pillars of the work.

[10] Asterión is the mythological creature that was half man and half bull. According to Greek mythology the name Asterión means "ruler of the stars"or "starry".

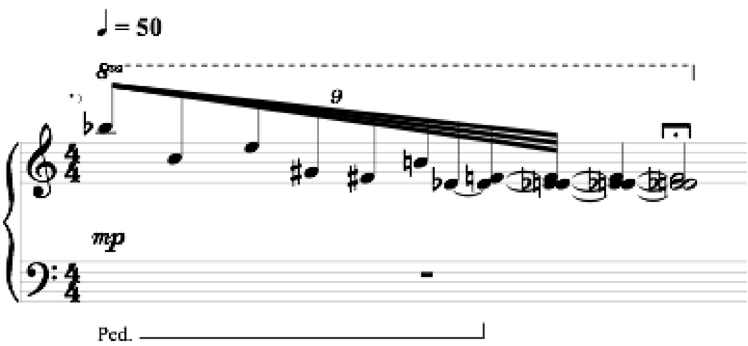


Figure 4.11: First measure of *Laberinto Borgiano*

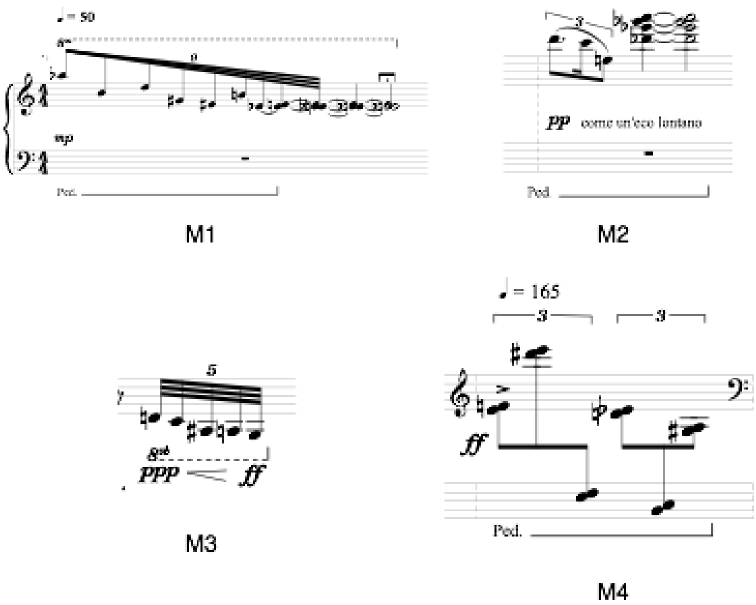


Figure 4.12: Main motifs in *Laberinto Borgiano*

There are four main motifs (Figure 4.12). The first two, M1 and M2 are developed in each section of the work respectively. M3 gradually becomes and important element in A at the end of the section. The addition of M4 in A creates the material for the transition between A and B. M1 is deconstructed at the end of the work intersecting with M2 chord *come un'eco lontano.*

In *Laberinto IV* another element was added to incorporate the electronics in a meaningful manner. A short introduction based on the beginning of Beethoven's piano

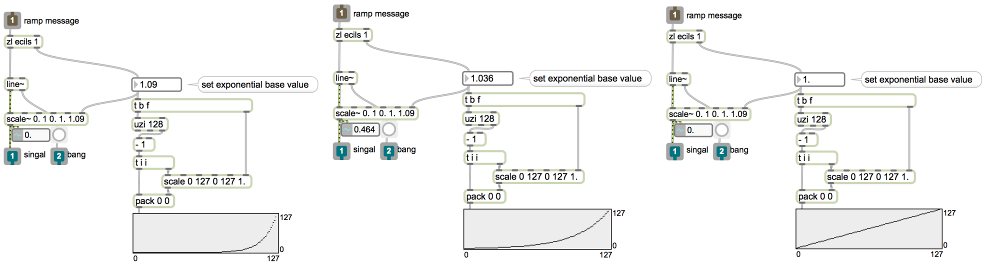


Figure 4.13: Time envelopes.

sonata Op. 100, III movement, alludes to the *bebung* technique .[11] The electronics extend the piano and play upon the notions of vibrato and the use of repeated notes and/or slurs.

For each section different time envelopes were used to mix the sound of the piano and the electronics. Combining pitch tracking and time envelopes, a cheap form of “score following” was implemented. The performer also controls time by manually advancing sections – and triggering the envelopes – with a ”user friendly” interface.

There are three basic envelopes that trigger eight sections of the piece. The first one has an exponential value of 1.09, the second one and exponential value of 1.036 and the last one, a linear envelope –with a value of 1.0–as shown in Figure 4.13. The sections trigger different settings for a granular synthesizer with a smooth overlap and[12] time stretching/freeze capabilities. The signal is divide into two different streams at middle C with a crossover filter at 261 Hz. This allows for processing the left and right hands separately, especially in the last section of the piece where the resonances of high pitches are extended by a “virtual *tremolo*”.

[11] A type of vibrato executed on the clavichord by applying force to a depressed key. Although not often written, performers apply the technique as ornamentation.

[12] IRCAM’s sogs~.

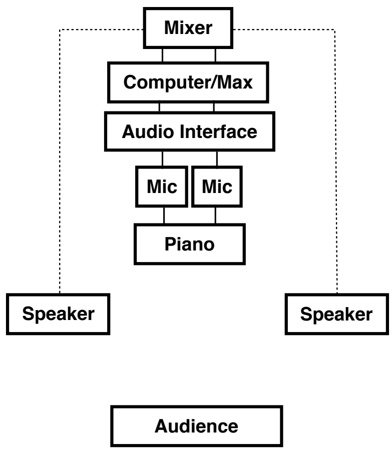
Martin Jaroszewicz

Laberinto IV

for piano and interactive electronics

Requirements and Setup

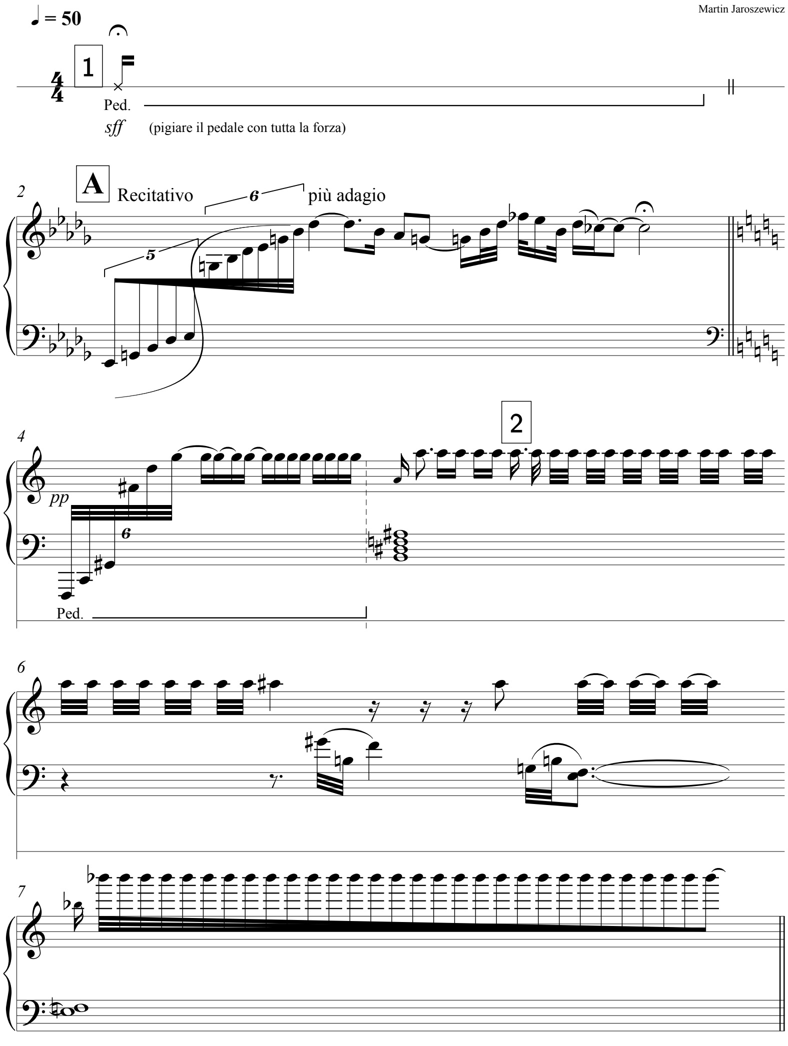
- Accompanying standalone Max/MSP patch. The patch requires the *sogs* granulator from IRCAM's SoundBox
- Two condenser microphones and cables
- An audio interface with at least two inputs/ outputs. Direct monitoring through outputs must be avoided.
- Mixer
- Two speakers / PA system
- The Max/MPS patch contains eight buttons the performer has to click as notated in the score. The buttons' numbers turn green when the mouse is over for better visibility and precision.
- Balanced should be achieved between the sound of the piano and the electronics.



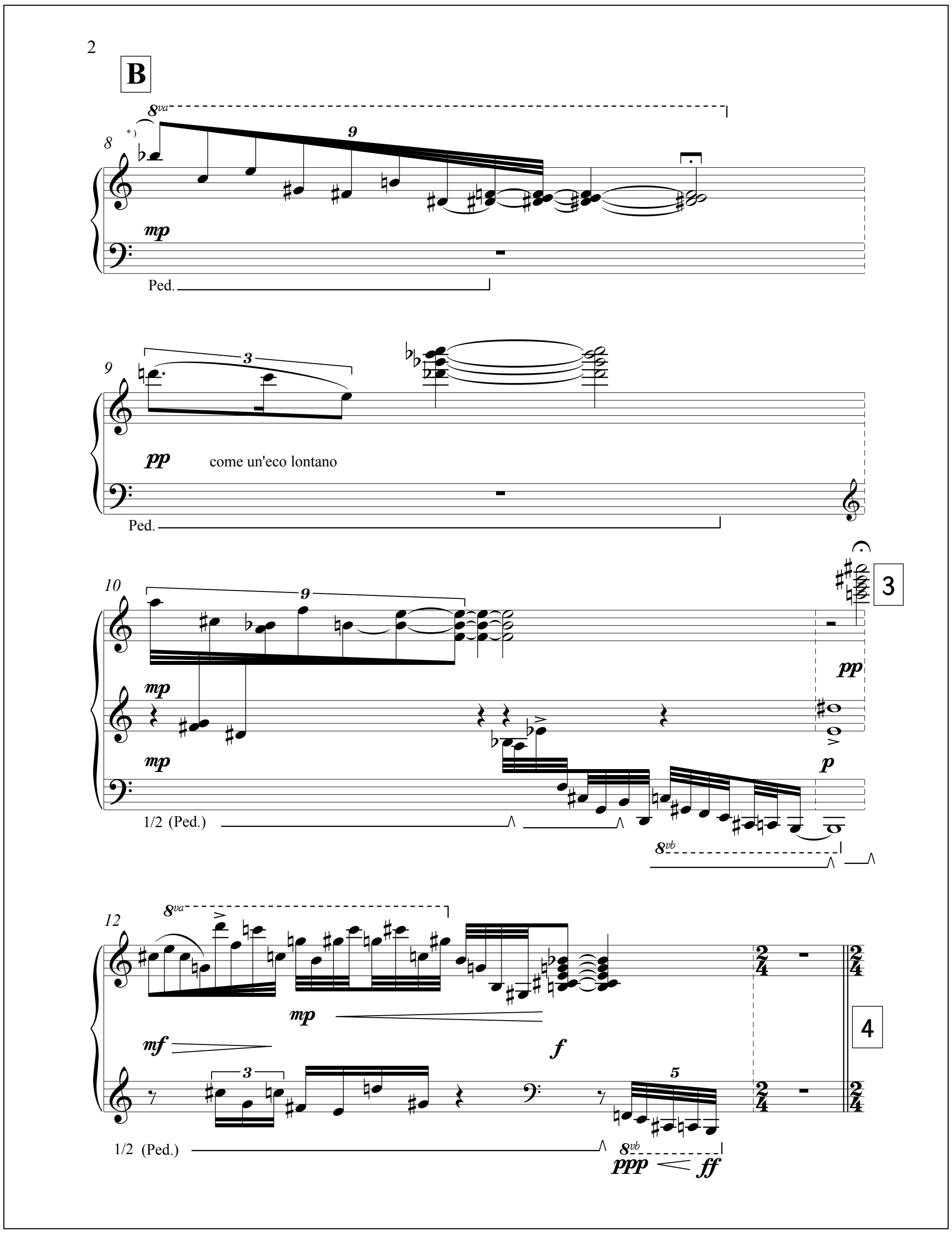
- Double click the *adc~* object to con gure inputs and outputs.

Laberinto IV

Martin Jaroszewicz

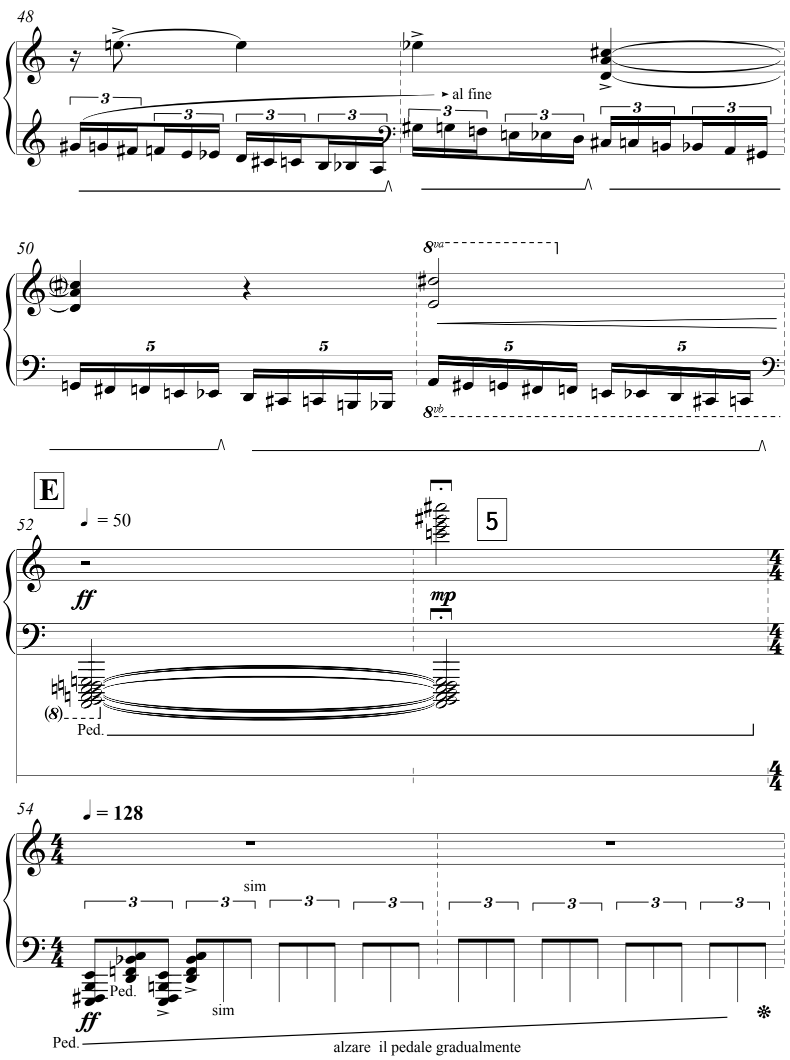


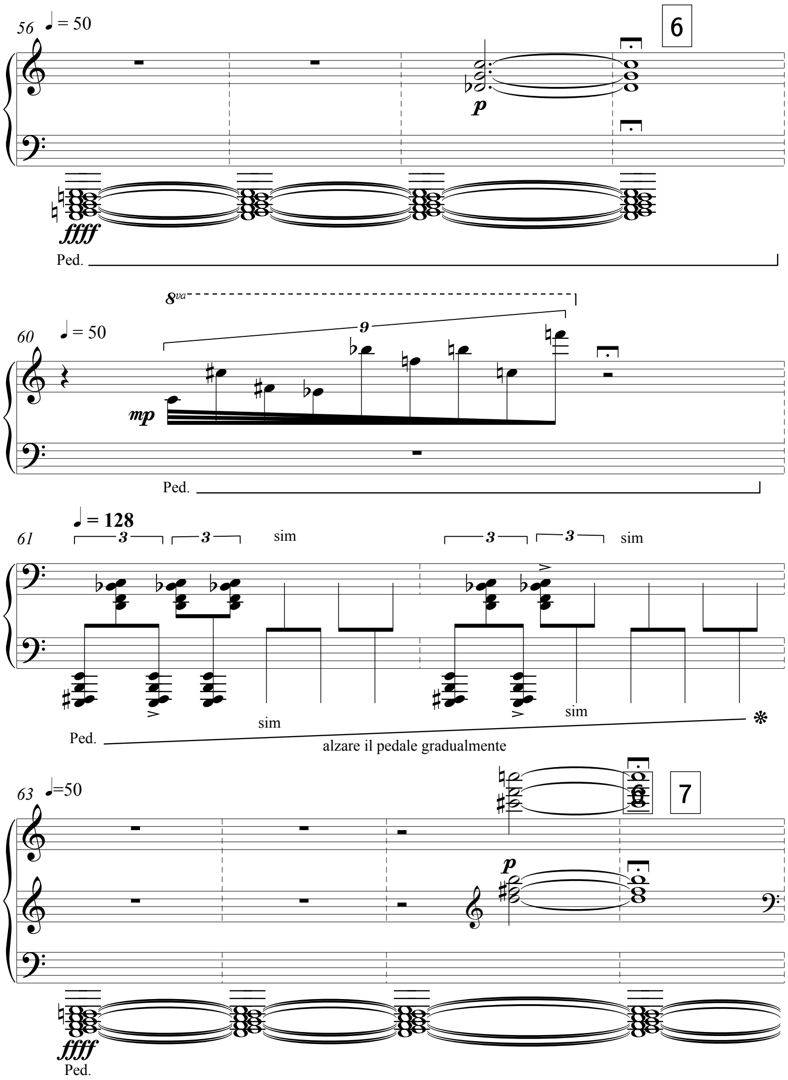
*) The use of dotted barlines is evident at rehearsal C. The performer can repeat and randomize this section ad libitum keeping the transition to section D connected. Pitches in measure 17 can be altered on every repetition.


2
B
mp
Ped.
pp
come un'eco lontano
Ped.
1/2 (Ped.)
mf
f
ppp
ff

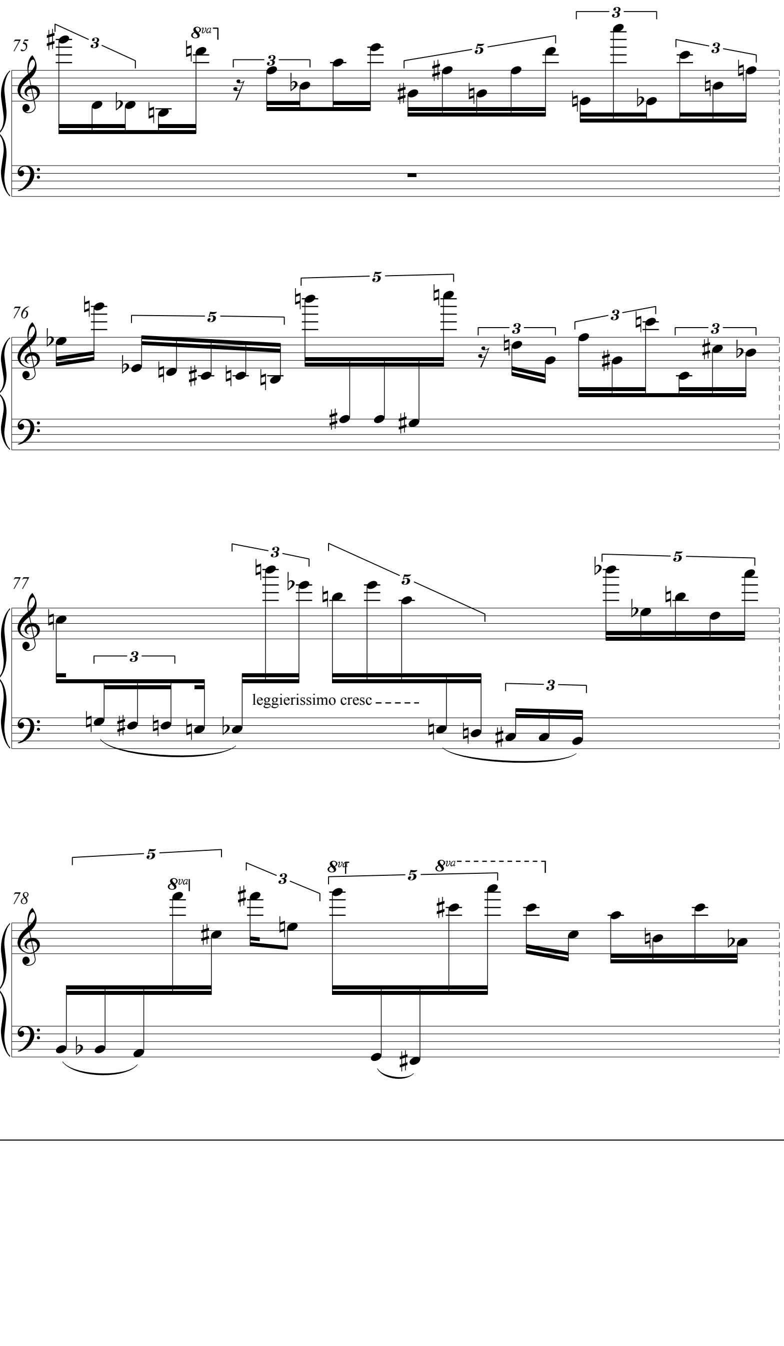

3
C
♩=120
Meccanico e secco
poco accel.
♩ = 165
martellato
senza pedale
salvo indicazione
Ped.
18 ♩=120


34
38
gradualmente legatissimo e leggiero
D
42
44
46


48
al fine
50
8va
8vb
E
52
= 50
ff
mp
5
(8)
Ped.
54
= 128
sim
Ped.
sim
ff
Ped.
alzare il pedale gradualmente


56 ♩ = 50
6
p
ffff
Ped.
8va
60 ♩ = 50
9
mp
Ped.
61 ♩ = 128
3
sim
sim
Ped.
alzare il pedale gradualmente
63 ♩=50
7
p
ffff
Ped.


F
(♩ = 50)
67
p
mp
69
8
mp sempre
8va
legatiss. sempre, quasi sempre Ped.
(8)
71
73
8va


75
76
77
leggierissimo cresc
78
8va


poco accel
79
3
5
8va
5
decresc
80
tempo
8va
8vb
8va
9
82
15ma
9
p
mf
p
Ped.
ritardando
come un'eco lontano
84
15ma
15ma
ppp

4.5 9

electroacoustic. Quadraphonic This work uses the technique of micro montage with pseudo random algorithms and composed spatialization. In the context of computer and electroacoustic music, a sound collage or montage is a technique where a work is made by combining a collection of sound objects in a pseudo random manner. Each object and morphology contributes to the whole resulting in a work with its own unique identity. The end result is intrinsically related to the time scale of each individual object. As the time scale gets smaller, the whole acquires a stronger morphological identity as perception in the micro sound[13] scale –less than 100ms–. This does not leave enough time for the listener to aurally discriminate sound morphologies and envelopes, especially sounds in which the attack has been masked by a window as in granular synthesis. If the time scale varies or fluctuates from the microtime domain to a bigger scale, microsounds can become a portion of an envelope of a sound that will be created and become an entity of the multiscale system escaping from the microstructure. For the work to be considered a montage or micro-montage, the composer must not create these macro morphologies intentionally.

My approach to creating a macro structure with microsounds is using a pseudo random sequence to preserve what we might expect from a random selection of objects. [14] The production of microsounds is, in my compositional work 9, based on pseudo random operations on a single sound object containing rich and varied material. This object is accessed randomly at different points in time providing the micro material for building the macrostructure. More than 5000 "particles" were generated with this

[13]For a discussion on time scales and granular synthesis see Roads (Roads 2004).

[14]A quasi random algorithm will be more appropriate for this purpose than a "random" or "pseudo-random" one, which has the tendency to clump objects together instead of being scattered somehow uniformly.

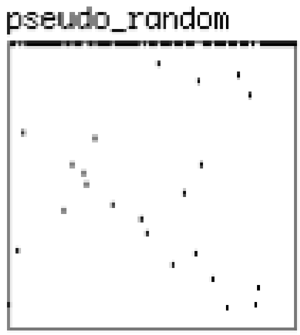


Figure 4.14: A random or "pseudo random" distribution in PD

technique for later construction of the montage. The use of particles in granulation differs from a montage in many respects, the most important of which is the process by which these particles become part of the work. In granulation, particles are generated by a computer as part of an automated process, while a micro-montage is the work of a composer "painting" as in the works of pointillists such as the French artist Seurat (Figure 4.15). The composer not only adds particles, he also adds silence, which could be compared to the non-intentional white space between the dots on the canvas which seems to be a by product of the painter's technique. The distance between sound particles is then carefully planned as an integral part of the montage.

The durations in 9 are very short, and silences of different lengths were randomly introduced to avoid saturation of density; other elements pertaining to a different time scale were also added. These layers that are not part of the random micro-structure add contrast to the work creating two independent macro-structures that are balanced with carefully composed quadraphonic spatialization. The addition of a multidimensional sound space to a somewhat linear composition, a characteristic of sound collages, creates a sense of direction that is perceivable by the listener as a formal structure cre-

Figure 4.15: Georges Seurat - "Circus Sideshow" (1889) - detail showing pointillism technique. Public domain image.

ating "something to hold on to" (Landy 2007). Silence creates the time scale for the microsounds to move in the quadraphonic space.

Spatialization at the microsound level has a powerful impact on the listener. As sounds move in a swarm-like pattern, similar to that of insects, the listener begins to perceive them in an almost tangible manner. Microsounds become dense creating a spatial effect in a process that can be thought of as forming a cloud in space. This manifests a vivid three-dimensional spatial morphology that can be also perceived in the stereo reduction of the work.[15]

[15]The stereo version of the piece was mixed creating a virtual quadraphonic stage with four speakers panned in front of the listener at 30,60,90,120 degrees

Chapter 5

Conclusion

In this thesis, I presented spectral spatialization, its background, applications and implementations. In addition, several musical examples that summarize many of the techniques discussed were analyzed. Most importantly, we looked at composition in space and the use of technology as a creative tool. Developed by the author to find solutions to his own compositional strategies, these parametric and algorithmic tools can help the composer of contemporary music to focus on the composition of sound spectra as it changes over time.

As part of the research conducted at EARS, and given the portability and flexibility of the studio, I took advantage of every possible configuration of loudspeakers allowing me to experiment with different spatialization techniques. This experimental approach was essential to the development of the collection of works *Laberintos* and their metaphors. As the apparatus becomes part of the composition in many different ways, a studio that is more flexible opens an opportunity for the composer to develop new musical ideas, overcoming the limitations of the past and finding new ones that in turn will be part of the *techné* that will shape future interactions. The mj Library for PD is a small contribution to the integration of spatialization in the compositional process, but it triggered new ideas for future work and the way the composer can interact with sound trajectories. The creation of a tool for spatialization that can work in tandem with synthesis techniques allowed me to reflect on what it means to work with sound–objects that can transform their spectral content as they move along a composed path. The research conducted at EARS suggested that a new understanding of the sound-object is needed, one that considers new philosophical ideas that are current and can contribute to the understanding of new technological developments as they shape the work of the composer.

5.1 Contribution

The contributions of the research conducted at EARS include the following:

- A series of objects for the Pure Data programming language to aid the composer with the design of parametric curves and random algorithmically driven paths.
- A command-line application that converts points from Houdini to coordinates that can be used in Max/MSP and PD.
- Different approaches to cross-synthesis – implemented in PD and Max/MSP – that can be synchronized with spatialization curves.
- A series of études that range from stereo to 8 channels using ambisonics and different computer music techniques including granular synthesis in the time domain and several techniques in the spectral domain.

Much of what I have discussed rests on the different compositional approaches to spatialization in music, sound imaging, sound fields and the aesthetics of circular motion. Most systems share a fundamental limitation: they are determined by the location of the listener, with some systems relying more on the *sweet spot*. There is a need of more tools to deal with movement in space as most libraries only provide encoding and decoding with few or no composition tools. The Game of Life is the only system that allows the composer to "compose" the spatialization along with the music but is still limited to basic curves and hand drawn paths.

I proposed a system that can be compatible with current 3D software such as Blender where any shape can be created and exported as coordinates. I also proposed the use of algorithmically created curves that can be synchronize to a cross-synthesis system for a spectral spatialization approach that is more meaningful in terms of composition and form. Spectral spatialization by changing spectra as sound moves is the correct approach to add meaning and clarity to electroacoustic music, whereas the creation of sound fields and diffusion creates an immersive experience that is more related to the space than the work. While spatialization focuses on the movement of sound objects, spectral spatialization focuses on the object as part of the work, that is, it focuses on form, thus movement becoming a structural part of it. There is a need for an integrated system that can allow the composer to work with spatialization in tandem with morphing.

The contorted passages and the hexagonal shapes in Borges' *La biblioteca de Babel* and the labyrinth in *La casa de Asterión* were the sources of inspiration to create a series of works where spatialization played an important role in defining form.

5.2 Future Work

At the time of this writing, a small contribution to the groundwork for specifying a framework for composers to work with spatialization in a meaningful manner was given. Tools for the – open source – Blender 3D software are being developed to be able to communicate in real time using the OSC protocol. Composing with the space in a intuitive manner or by using algorithms can benefit from a "user-frienldy" 3D interface where the composer can transform non-musical ideas into meaningful spatialization.

Bibliography

Aristotle. 2001. *Problems.* Translated by R. Mayhew. Cambridge, MA: Harvard University Press.

Barrett, Natasha. 2010. "Ambisonics Spatialisation and Spatial Ontology in an Acousmatic Context." Paper presented at the *Sixth Electroacoustic Music Studies Network Conference*. http://www.ems-network.org/ems10/ (accessed February 12, 2015).

Baumann, Dorothea. 2011. *Music and Space: A Systematic and Historical Investigation Into the Impact of Architectural Acoustics on Performance Practice Followed by a Study of Handel's Messiah.* Bern: Peter Lang, AG Internationaler Verlag der Wissenschaften.

Berkhout, A.J, D. de Vries, and P. Vogel. 1993. "Acoustic Control by Wave Field Synthesis." *Acoustic Society of America* 93: 2764-78.

Berlioz, Hector. 1994. *The Art of Music and Other Essays.* Edited by Elizabeth Csicsery-Ronay. Bloomington, IN: Indiana University Press.

Bloch, William. 2008. *The Unimaginable Mathematics of Borges' Library of Babel.* Oxford, NY: Oxford University Press.

Bole, Tiziano. 2008. "Faust Tutorial 2". Grame. http://faust.grame.fr/ (accessed February 12, 2015).

Borges, Jorge Luis. 1997. *Ficciones.* Madrid: Alianza Editorial.

Boulez, Pierre. 1986. *Orientations: Collected Writings.* London: Faber.

Boulez, Pierre and Andrew Gerzso. 1988. "Computers in Music." *Scientific American* 258. http://articles.ircam.fr/textes/Boulez88c/ (accessed February 12, 2015).

Brant, Henry. 1967. "Space as an Essential Aspect of Musical Composition". In *Contemporary Composers on Contemporary Music,* edited by Elliott Schwartz, Barney Childs, and Jim Fox, 221–42. New York, NY: Da Capo Press.

Brant, Henry. 1977. *Antiphony One: For Symphony Orchestra.* New York, NY: Fischer.

Chagas, Paulo. 2014. *Unsayable Music: Six Reflections on Musical Semiotics, Electroacoustic and Digital* Music. Leuven: Leuven University Press.

Chagas, Paulo. 2008. "Composition in Circular Sound Space: Migration 12-channel electronic music 1995-97". *Organised Sound* 13 (3): 189-98.

De Cheveigne, Alain, and Hideki Kawahara. 2002. "YIN, a Fundamental Frequency Estimator for Speech and Music." *The Journal of the Acoustical Society of America* 111: 1917-30.

Clozier, Christian. 2001. "The Gmebaphone Concept and the Cyberphone Instrument." *Computer Music Journal* 25: 81-90.

Daniel, Jérome. 2001. "Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia." PhD dissertation. Paris: Université Paris 6.

Declercq, Nico F. 2007. "Acoustic Diffraction Effects at the Hellenistic Amphitheater of Epidaurus: Seat Rows Responsible for the Marvelous Acoustics." *The Journal of the Acoustical Society of America* 121: 2011-22.

Decroupet, Pascal, Elena Ungeheuer and Jerome Kohl. 1998. "Through the Sensory Looking-Glass: The Aesthetic and Serial Foundations of Gesang der Jünglinge." *Perspectives of New Music* 36: 97-142.

Van Dorp Schuitman, J. 2005. "Wave Field Synthesis Using Multi Actuator Panel Loudspeakers: Design and Application of Various Digital Filtering Algorithms." Phd dissertation. Delft: Delft University of Technology.

Emmerson, Simon. 2007. *Living Electronic Music.* Hampshire: Ashgate Publishing Limited.

Eno, Brian. 2004. *Audio Culture: Readings in Modern Music.* Edited by Dan Warner. New York, NY: Continuum.

Farnell, Andy. 2010. *Designing Sound.* Cambridge, MA: MIT Press.

Fletcher, Stephen. 1992. *A Biographical Memory.* Washington, D.C.: National Academy of Sciences.

Garcia, Tristan. 2014. *Form and Object.* Edinburg: Edinburg University Press.

Gerzon, Michael. 1974. "What's Wrong With Quadraphonics?" *Studio Sound*. http://audiosignal.co.uk/Resources (accessed February 12, 2015).

Harman, Graham. 2011. "Realism Without Materialism." *SubStance* 40: 52-72.

Holman, Tomlinson. 2008. *Surround Sound: Up and Running.* Boston, MA: Focal Press.

Ives, Charles. 1908. *The Unanswered Question.* New York, NY: Southern Music Publishing Co. Inc.

Jacquemin, Guillaume, Thierry Coduys and Matthieu Ranc. 2012. "IANNIX 0.8". Paper presented at *des Journées d'Informatique Musicale (JIM 2012)*, Mons, Belgium, 9-11 May, 2012. http://jim.afim-asso.org/jim12/ (accessed February 12, 2015).

Jaroszewicz, Martin. 2013. "Pierre Boulez's Répons: A Virtual Cathedral". Paper presented for PhD qualifying exam. University of California Riverside.

Landy, Leigh. 2007. *Understanding the Art of Sound Organization.* Cambridge, MA: MIT Press.

Manning, Peter. 2006. "The Significance of Techné in Understanding the Art and Practice of Electroacoustic Composition." *Organised Sound* 11 (1): 81-90.

Miller, Paul. 2009. "Stockhausen and the Serial shaping of Space". PhD diss. Rochester, NY: Eastman School of Music, University of Rochester.

Morton, Timothy. 2013. *Hyperobjects: Philosophy and Ecology after the End of the World.* Minneapolis, MN: University of Minnesota Press.

Murray, James and William van Ryper. 1996. *Encyclopedia of Graphics File Formats.* Bonn: O'Reilly and Associates.

Pérez-López, Andrés. 2014. "Real-Time Audio Spatialization Tools for Interactive Performance." Master thesis, Universitat Pompeu Fabra, Barcelona. http://mtg.upf.edu/ (accessed February 12, 2015).

Pollio, Vitruvius. 2008. *Los diez libros de arquitectura.* Barcelona: Linkgua Ediciones.

Puckette, Miller. 2007. *The Theory and Technique of Electronic Music.* Hackensack, NJ: World Scientific Publishing Company.

Rabenstein, R., S. Spors and P. Steffen. 2006. "Wave Field Synthesis Techniques for Spatial Sound Reproduction." In *Topics in Acoustic Echo and Noise Control,* 517-45. Berlin: Springer.

Ramirez, Robert. 1985. *The FFT: Fundamentals and Concepts.* Englewood Cliffs, NJ: Prentice-Hall.

Roads, Curtis. 2004. *Microsound.* Cambridge, MA: MIT Press.

Sascha Spors, Rudolf Rabenstain and Jens Ahrens. 2008. "The Theory of Wave Field Synthesis Revisited". Presented at *the 124th Convention of Audio Engineering Society* http://www.aes.org/events/124/ (accessed February 12, 2015).

Schaeffer, Pierre. 2003. *Tratado de los objetos musicales.* Translated and edited by A. C. de Diego. Madrid: Alianza Editorial.

Scipio, Agostino Di. 1995a. "Inseparable Models of Materials and of Musical Design in Electroacoustic and Computer Music." *Journal of New Music Research* 24 (1): 34-50.

Sèdes Anne, Pierre Guillot and Eliott Paris. 2014. "The HOA Library, Reviews and Prospects." Paper presented at *ICMC|SMC|2014.* http://www.icmc14-smc14.net/ (accessed February 12, 2015).

Shankland, Robert S. 1972. "The Development of Architectural Acoustics: The Application of the Science of Acoustics to Architectural Designs Has Produced Greatly Improved Halls for Both Speech and Music." *American Scientist* 60: 201-09. http://www.jstor.org/stable/27843021 (accessed February 12, 2015).

Slonimmsky, Nicolas. 1998. *Webster's New World Dictionary of Music.* Edited by R. Kassel. New york, NY: Macmillan.

Smalley, John. 2000. "Gesang der Jünglinge: History and Analysis." Columbia University Department of Music. http://www.music.columbia.edu/masterpieces/notes/stockhausen (accessed February 12, 2015).

Smith III, Julius O. 2010. "Physical Audio Signal Processing: For Virtual Musical Instruments and Audio Effects. JOS On-Line Publications, Stanford University. https://ccrma.stanford.edu/˜jos/pubs.html (accessed February 12, 2015).

Snoei, Wouter. 2014. "Wave Field Synthesis". Personal Powerpoint.

Snoei, Wouter and Martin Jaroszewicz. 2014b. Personal conversation.

Spatialisateur. 2012. *User Manual.* Paris, France: IRCAM. http://forumnet.ircam.fr/product/spat/ (accessed February 12, 2015).

Steinitz, Richard. 2003. *Gyorgy Ligeti: Music of the Imagination.* Boston, MA: Northeastern University Press.

Stockhausen, Karlheinz. 1958. "Musik im Raum." Die Reihe 5: 152-175.

Weisstein, Eric. 2014. "Rose." *Wolfram Mathworld.* http://mathworld.wolfram.com/Rose.html (accessed February 12, 2015).

Xenakis, Iannis. 1992. *Formalized Music: Thought and Mathematics in Composition.* Hillsdale, NY: Pendragon Press.

.1 Appendix A - mj Library Source Code

- mjRotation

```
//
// mjRotation.c
// mjRotation
//
// Created by Martin Jaroszewicz on 7/20/14.
// Copyright (c) 2014 com.martinjaroszewicz. All rights reserved.
//

#include "m_pd.h"
#include <math.h>
#include <string.h>

static t_class *mjRotation_class;

typedef struct _mjRotation {
    t_object x_obj;
    //local variables
    t_float f_count;
    t_float f_x;
    t_float f_y;
    t_float f_z;
    t_float f_elipX;
    t_float f_elipZ;
    t_float f_input;
    t_float f_mult;
    t_outlet *l1_out, *l2_out, *l3_out;
} t_mjRotation;


void mjRotation_float(t_mjRotation *x, t_float f){
    t_float a1 = x->f_x;
    t_float a2 = x->f_y;
```

```c
    t_float a3 = x->f_z;
    ;
    x->f_x = sin(f * x->f_mult)*x->f_elipX;
    x->f_y = 0;
    x->f_z = cos(f * x->f_mult)*x->f_elipZ;
    outlet_float(x->l1_out, a1);
    outlet_float(x->l2_out, a3);
    outlet_float(x->l3_out, a2); // This works for HOA Library
}

void *mjRotation_new(t_floatarg f){

    t_mjRotation *x = (t_mjRotation *)pd_new(mjRotation_class);
    x->f_mult = 3.14159 * 2;
    x->l1_out = outlet_new(&x->x_obj, &s_float);
    x->l2_out = outlet_new(&x->x_obj, &s_float);
    x->l3_out = outlet_new(&x->x_obj, &s_float);
    floatinlet_new(&x->x_obj, &x->f_elipX);
    floatinlet_new(&x->x_obj, &x->f_elipZ);

    return (void *)x;
}

void mjRotation_setup(void) {
    mjRotation_class = class_new(gensym("mjRotation"),
    (t_newmethod)mjRotation_new,
            0, sizeof(t_mjRotation),
            CLASS_DEFAULT,
            A_DEFFLOAT, 0);
    class_addfloat(mjRotation_class, mjRotation_float);

}
```

- mjRose

```c
//
```

```
// mjRose.c
// mjRose
//
// Created by Martin Jaroszewicz on 7/20/14.
// Copyright (c) 2014 com.martinjaroszewicz. All rights reserved.
// // This works with HOA Library

#include "m_pd.h"
#include <math.h>
#include <string.h>

static t_class *mjRose_class;

typedef struct _mjRose {
    t_object x_obj;
    //local variables
    t_float f_position;
    t_float f_mult;
    t_float f_n;
    t_float f_d;
    t_float f_z;
    t_outlet *l1_out, *l2_out, *l3_out, *l4_out;
} t_mjRose;


void mjRose_float(t_mjRose *x, t_float f){

    double n;
    double d;
    modf(x->f_n,&n);
    modf(x->f_d,&d);
    int z = x->f_z;
    int in = n;
    int id = d;
    if(in%in != 0 && id%id !=0){
```

```
        x->f_mult = M_PI*2*d;
    } else {
        x->f_mult = M_PI*d;
    }
    x->f_position = f;
    float position = x->f_position*x->f_mult;
    float rose = sinf((n/d)*position);
    float out_x = rose*cosf(position);
    float out_y = rose*sinf(position);
    float out_z = rose*sinf(z);
    outlet_float(x->l1_out, out_x);
    outlet_float(x->l2_out, out_y);
    outlet_float(x->l3_out, out_z);
    if (out_z == 0.0 && out_y == 0.0 && out_z == 0.0){
        outlet_bang(x->l4_out);
    }
}

void *mjRose_new(t_floatarg f){

    t_mjRose *x = (t_mjRose *)pd_new(mjRose_class);
    x->f_n = 1;
    x->f_d = 1;
    x->f_z = 0;
    x->l1_out = outlet_new(&x->x_obj, &s_float);
    x->l2_out = outlet_new(&x->x_obj, &s_float);
    x->l3_out = outlet_new(&x->x_obj, &s_float);
    x->l4_out = outlet_new(&x->x_obj, &s_bang);
    floatinlet_new(&x->x_obj, &x->f_n);
    floatinlet_new(&x->x_obj, &x->f_d);
    floatinlet_new(&x->x_obj, &x->f_z);
    post("mj Library by Martin Jaroszewicz.");
    post("  2014 UCR Riverside");
    post("Version Beta 0.1 (August 9 2014) for PD-extended v.0.43.4");

```

```
    return (void *)x;
}

void mjRose_setup(void) {
    mjRose_class = class_new(gensym("mjRose"),
    (t_newmethod)mjRose_new,
            0, sizeof(t_mjRose),
            CLASS_DEFAULT,
            A_DEFFLOAT, 0);
    class_addfloat(mjRose_class, mjRose_float);

}
```

- mjRandom

```
//
// mjRandom.c
// mjRandom
//
// Created by Martin Jaroszewicz on 7/20/14.
// Copyright (c) 2014 com.martinjaroszewicz. All rights reserved.
//

#include "m_pd.h"
#include <math.h>
#include <string.h>
#include <stdlib.h>

static t_class *mjRandom_class;

typedef struct _mjRandom {
    t_object x_obj;
    //local variables
    t_float i_num_points;
    t_float a_x[100];
    t_float a_y[100];
```

```
    t_float a_z[100];
    t_float f_seed;
    t_float f_input;
    t_outlet *l1_out, *l2_out, *l3_out,*l4_out;
} t_mjRandom;


void mjRandom_float(t_mjRandom *x, t_float f){

    int index = f*x->i_num_points;
    post("Incoming float %f with index %d",f, index);
    outlet_float(x->l1_out, x->a_x[index]);
    outlet_float(x->l2_out, x->a_y[index]);
    outlet_float(x->l3_out, x->a_z[index]);
    if (f == 0.999){ // a phasor never gets to 1 but this does not work
        outlet_bang(x->l4_out);//change seed here
    }
}

void mjRandom_bang(t_mjRandom *x){

    float seed = x->f_seed;
    int points = x->i_num_points;
    if (points > 100){
        post("Max number of points is 100");
        points = 100;
    }
    if (points < 0){
        post("Please provide a value between 0 and 100");
        points = 0;
    }
    int index;
    float m1[points];
    float m2[points];
    float m3[points];
```

```
    for (index = 0; index < points; index++){
        m1[index] = (float)rand() / (float)RAND_MAX;
        m2[index] = (float)rand() / (float)RAND_MAX;
        m3[index] = (float)rand() / (float)RAND_MAX;
        post("x[%d] = %f",index,m1[index]);
        post("y[%d] = %f",index,m2[index]);
        post("z[%d] = %f",index,m3[index]);
        x->a_x[index] = m1[index];
        x->a_y[index] = m2[index];
        x->a_z[index] = m3[index];

    }
}

void *mjRandom_new(t_floatarg f){

    t_mjRandom *x = (t_mjRandom *)pd_new(mjRandom_class);
    x->l1_out = outlet_new(&x->x_obj, &s_float);
    x->l2_out = outlet_new(&x->x_obj, &s_float);
    x->l3_out = outlet_new(&x->x_obj, &s_float);
    x->l4_out = outlet_new(&x->x_obj, &s_bang);
    floatinlet_new(&x->x_obj, &x->f_seed);
    floatinlet_new(&x->x_obj, &x->i_num_points);

    return (void *)x;
}

void mjRandom_setup(void) {
    mjRandom_class = class_new(gensym("mjRandom"),
    (t_newmethod)mjRandom_new,
            0, sizeof(t_mjRandom),
            CLASS_DEFAULT,
            A_DEFFLOAT, 0);
    class_addbang(mjRandom_class, mjRandom_bang);
    class_addfloat(mjRandom_class, mjRandom_float);
```

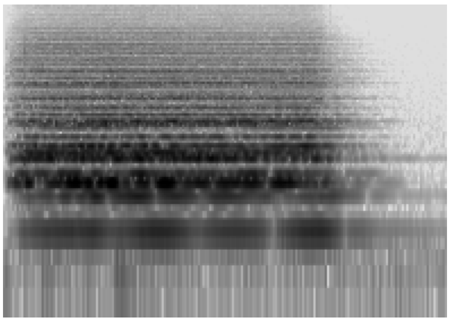
```

}
```

- Python script that generates a point cloud in Houdini.

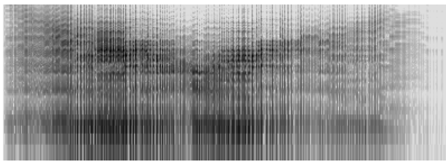
```
import hou
import math

# Creates a sine wave path and exports to file

def createGridGeo():
    # Creates main obj GEO
    aGeoNode = hou.node('obj').createNode('geo',run_init_scripts=False,node_name = "Random_points"
        )
    # Creates a grid node with 1 row and 180 columns
    aGridNode = aGeoNode.createNode('grid')
    aGridNode.parm('cols').set(180) # The number of points
    aGridNode.parm('rows').set(1)
    aGridNode.parm('sizex').set(2) #size 2 (from -1 to 1)
    aGridNode.parm('sizey').set(1)
    # Creates a point node and adds a formula to create a sine wave
    aPointNode = aGeoNode.createNode('point')
    aPointNode.setFirstInput(aGridNode)
    aPointNode.parm('tx').setExpression("$TY + rand ($PT) * 0.2") #c
    aPointNode.parm('ty').setExpression("$TZ + rand ($PT*3) * 0.2")
    aPointNode.parm('tz').setExpression("$TX + rand ($PT*4) * 0.2")
    # Creates a delete node to delete everything but the points
    aDeleteNode = aGeoNode.createNode('delete')
    aDeleteNode.setFirstInput(aPointNode)
    aGeoNode.layoutChildren()
    aDeleteNode.parm('pattern').set("0 $N")
    aDeleteNode.parm('keeppoints').set(1)
    aDeleteNode.setDisplayFlag(1)
    # Creates a file node to export
    aFileNode = aGeoNode.createNode('file')
    aFileNode.setFirstInput(aDeleteNode)
```

```
    aFileNode.parm('filemode').set(2) #write
    aFileNode.parm('file').set("test.poly")
    aFileNode.setDisplayFlag(1)


createGridGeo()
```

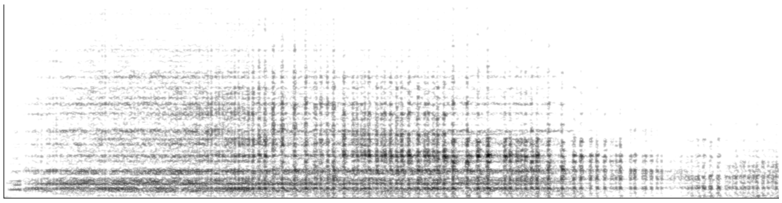
.1 Appendix B - Spectrograms



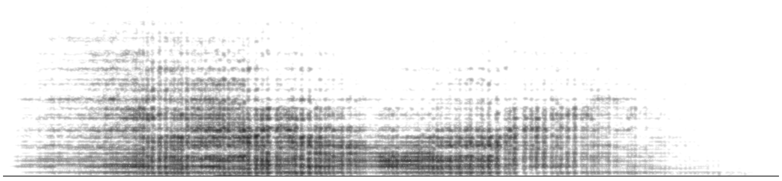
(a) A chord played by a string ensemble. Dur: 7 seconds.



(b) Granular texture. Dur: 14 seconds.

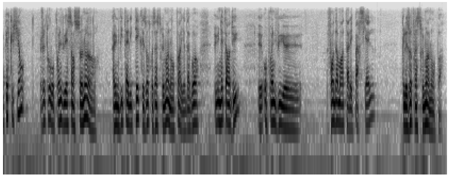


(c) General Cross-Synthesis between (a) and (b)

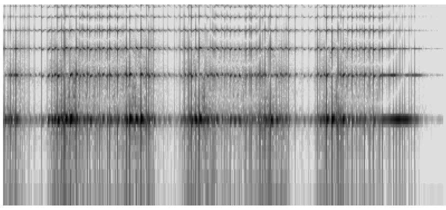


(d) Source Filter Cross-Synthesis between (a) and (b)

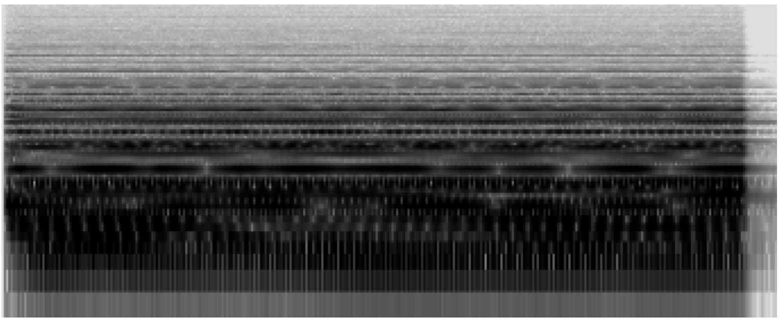
Figure .1: Different Cross-synthesis techniques.



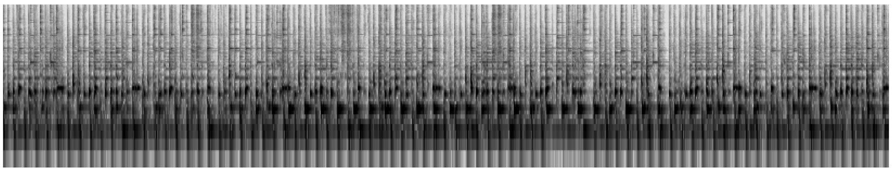
(a) Male voice speaking in French. Dur: 15 seconds.



(b) Granular glass sounds. Dur: 16 seconds.



(c) Organ C chord. Dur: 12 seconds.



(d) Latin Percussion. Dur: 29 seconds.

Figure .2: Sounds with contrasting spectrum.