ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

INTUITIVE CONTROL: NAVIGATING THE DEPTHS OF SOUND WITH A MODULAR SYSTEM

FINAL PROJECT

Fatih ACIKGOZ

The Department of Music

Music Programme (Without Thesis)

MAY 2023

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

INTUITIVE CONTROL: NAVIGATING THE DEPTHS OF SOUND WITH A MODULAR SYSTEM

FINAL PROJECT

Fatih ACIKGOZ

(409211013)

The Department of Music

Music Programme (Without Thesis)

Project Advisor: Dr. Manolis Ekmektsoglou

MAY 2023

Fatih ACIKGOZ, a graduate student with student ID number 409211013 at Istanbul Technical University's Graduate School of Education, has successfully presented his thesis titled "An Extensive Live Performance Tool to Create an Alternative for Traditional DAW and Meter Systems" before a jury with the signatures below, after fulfilling all necessary requirements set forth by relevant regulations.

Project Advisor:	**Dr. Manolis Ekmektsoglou** Istanbul Technical University	**....……………….**
Jury Members:	**Dr. Manolis Ekmektsoglou** Istanbul Technical University	**....……………….**
	Assoc. Prof. Dr. Cevdet Erek Istanbul Technical University	**....……………….**
	Dr. Sair Sinan Kestelli Istanbul Technical University	**....……………….**

Date of Submission: 26 May 2023
Date of Defence : 8 June 2023

FOREWORD

I would like to thank all the members of the MIAM community, who have had a profound effect on my perspectives on my life, and all my dear friends who touched my mental thinking.

May 2023 Fatih ACIKGOZ

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

BPM	Beats Per Minute
CD	Compact Disc
CMI	Computer Musical Instrument
DAW	Digital Audio Workstation
DSP	Digital Signal Processing
EDO	Equal Division of Octave
FluCoMa	Fluid Corpus Manipulation
FM	Frequency Modulation
Hz	Hertz
LP	Long Playing
Ms	Millisecond
MLP	Multi-Layer Perceptron
MC	Multi Channel
PRNGs	Pseudo Number Generators

LIST OF FIGURES

SUMMARY

This project explores the realm of sound synthesis and control within a modular interface, aiming to create generative music systems that offer flexibility, creativity, and efficient parameter control. The study focuses on two main sound sources: multichannel-based additive synthesis and granular synthesis. Additive synthesis, utilizing the sum of sine waves, provides a comprehensive sound source, while granular synthesis involves dividing complex waveforms into small grains for sonic manipulation. The modular interface employs optimized and fast multichannel (MC) objects, enabling efficient additive processes and the generation of harmonic series. To maintain smooth transitions between grains and prevent clipping, a buffer~ approach is employed, dividing waveforms into grains and applying Gaussian envelopes to ensure optimal playback. Additionally, a hybrid control mechanism is adopted, allowing for both collective and individual parameter control. The system's parameters are designed as "modes," representing the simultaneous movement of all parameters. To facilitate mode transitions, the concept of interpolation is introduced, with MPLRegressor~ from the Fluid Corpus Manipulation Project serving as a neural network-based tool for predicting parameter values based on assigned coordinates. The study highlights the importance of mapping parameters and controlling the interface through a pitchslider and 24 parameter values. Pinpointing specific modes within the system enables optimized interpolation between data sets, resulting in smooth transitions between modes. Through this research, the modular interface offers a powerful platform for sound synthesis and control, empowering musicians and artists to create dynamic and immersive sonic atmospheres. Overall, this project contributes to the field of sound synthesis and control by presenting a hybrid approach that combines additive and granular synthesis techniques within a modular interface. The integration of optimized multichannel objects, buffer~ approaches, and interpolation mechanisms enhances the creative potential and efficiency of the system, providing a flexible and intuitive platform for generative music production.

ÖZET

Bu proje, modüler bir arayüz içinde ses sentezi ve kontrolün alanını keşfederek, esneklik, yaratıcılık ve etkili parametre kontrolü sunmayı amaçlamaktadır. Çalışma, iki ana ses kaynağı üzerinde odaklanmaktadır: çok kanallı temelli toplamsal sentez ve taneli sentez. Toplamsal sentez, sinüs dalgalarının toplamını kullanarak kapsamlı bir ses kaynağı sağlarken, taneli sentez karmaşık dalga formlarını küçük tanelere bölmeyi ve bunları ses manipülasyonu için kullanmayı içerir. Modüler arayüz, optimize edilmiş ve hızlı çok kanallı (MC) nesneleri kullanarak etkili toplama işlemleri ve harmonik serilerin üretilmesini sağlar. Taneler arasındaki geçişleri düzgün bir şekilde sağlamak ve klipleme olasılığını önlemek için bir buffer~ yaklaşımı kullanılır. Dalga formları tanelere bölünerek, her biri Gauss zarfıyla oynatma için optimize edilmiş hale getirilir. Ek olarak, hibrit bir kontrol mekanizması benimsenir ve hem kolektif hem de bireysel parametre kontrolüne izin verilir. Sistemin parametreleri "modlar" olarak tasarlanır ve tüm parametrelerin aynı anda hareket etmesini temsil eder. Mod geçişlerini kolaylaştırmak için, ara değerlere dayanarak parametre değerlerini tahmin eden Fluid Corpus Manipulation Projesi'nden MPLRegressor~ kullanılır. Çalışma, parametreleri eşlemek ve arayüzü bir pitchslider ve 24 parametre değeri üzerinden kontrol etmenin öneminin altını çizer. Sistem içinde belirli modları belirleyerek, veri setleri arasında optimize edilmiş bir ara değerlendirme sağlanır ve modlar arasında düzgün geçişler elde edilir. Bu araştırma sayesinde modüler arayüz, müzisyenlere ve sanatçılara dinamik ve etkileyici ses atmosferleri yaratma imkanı sağlayan güçlü bir platform sunmaktadır. Genel olarak, bu proje, toplamsal ve taneli sentez tekniklerini modüler bir arayüz içinde birleştiren bir hibrit yaklaşım sunarak ses sentezi ve kontrol alanına katkıda bulunur. Optimize edilmiş çok kanallı nesnelerin, buffer~ yaklaşımlarının ve ara değerlendirme mekanizmalarının entegrasyonu, sistemin yaratıcı potansiyelini ve verimliliğini artırır.

1 INTRODUCTION

Although the expression of musical thought is similar from past to present, it has experienced critical breakdowns in many ways. At its core, before the use of tools for musical expression, the result was a highly personal and unpredictable one. This is because only individual experiences play a role in the development of musical expression in the mental process. However, the introduction of tools has led to the establishment of an inevitable organic link between this production process and tools. This new depth of musical expression is, in a sense, limited by the inherent limits of the instruments used. It should not be concluded here that instruments limit musical expression. Because tools have provided the opportunity to expand the places where humanoid production is limited. In this context, all acoustic instruments, which have been used as a means of expression of musical thought for a long time, have been an example of this. When it comes to virtuosity, the limits of instruments have been exceeded.

The tools mentioned so far have illuminated many untouched areas on the map of the musical expression process. However, the fact that digital technologies have become a tool of musical expression has turned this historical chart upside down. In the tools mentioned before, the practice of personal thinking and the time for the tools to reflect this are in a natural balance. However, the information processing capability of digital technology has encouraged musical expression to think far beyond the possibilities. The audacity of information processing power can be cited as a sign that the winds of experimental movement will slowly begin to blow in personal music approaches. At this point, musical expression has reached the creative power that not a single tool can achieve, but a multitude of tools at the same time. In this sense, the myriad of interfaces developed on computers and hardware have increasingly made musical expression more unpredictable than ever before. Increasingly and popularly, these digital interfaces have now become more standard and self-forming mechanisms. At this point, we can bring up the concept of Digital Audio Workstation (DAW)[1] for the first time.

When DAW was first discovered, it was a gigantic virtual studio that offered unlimited possibilities and challenged the minds with its results. However, the policy of standardization, which has progressed in parallel with the music industry, has turned the digital interfaces, which always involve the new and deepen, into standard and highly optimized tools focused on meeting the demand in the industry. Just like acoustic instruments, it provided users with a comfort zone in a structure that brought the concept of personal virtuosity back to the fore. In this article, I will explore the foundations of my search for a new musical interface that can expand the boundaries of DAWs without denying the existence of DAWs.

[1] Here, a comparison is made with DAWs specific to Ableton, which stands out with its live performance features.

2 TRADITIONAL DAW ENVIRONMENTS

2.1 Overview

The phenomenon of collective self-existence, which music carries by its nature, has undergone many changes from past to present. The self-creation of music emerged as a result of organic-based biological factors. The enlightenment of the early people regarding the use of instruments deeply affected the way of musical expression and heralded the journey of accompaniment mechanisms to the new chakras. After the phonograph, the fact that the phenomenon of music became open to experience again and increased its diffusion power would redefine musical existence. The change in musical performance has affected people's listening experiences. This was the beginning of "on demands" listening (Thompson, 2016). Music performance is now in a position where it can be recorded and listened to over and over by more people. The first recorded music on vinyl was made in 1898 by the Victor Talking Machine Company. This track, "Twinkle, Twinkle, Little Star," was the birthplace of the record label and the music industry (*A Brief History of Sound Recording*, 2022). The critical point here is that the musical experience now moves out of one's individual space and becomes open to interaction in wider areas. Here, there emerges the existence of different dynamics that the musical experience must take into account when realizing itself - both in terms of the music creator and the music listener. What is meant here is that the dynamics of producers and listeners involved in musical production are deeply affected and connected to each other as an industrial concern.

The revolutionary effect of the phonograph gradually faded after a while due to the poor quality of the recordings. After the Second World War, the introduction of new technologies such as LP (long-playing) records and compact discs (CD) into our lives enabled this process to reach a more detailed and quantitatively powerful level. In 1981, Tony Banks developed the first Digital Audio Workstation (DAW), called "Sound Maker", by making a digital version of the Fairlight CMI (Computer Musical Instruments), the popular synthesizer of the time (*A Brief History of Sound Recording*, 2022). This can be considered as the beginning of music production in a personal sense. Can it be said that this situation now provides space for the music producer to convey his personal musical experience? In one sense, the answer to this question is positive, on the other hand, it is negative. If we consider a record label, we can understand that the source of motivation for the music produced here can be the audience it reaches and the number of compact discs sold. This is the point at which we must consider the musical experience as a musical 'product'. For this reason, it would not be surprising if the personal musical production process provided by DAW evolved on similar concerns. Because the development, use and processing of recording equipment owned by record label as a company is possible with a sustainable commercial concern. This situation is optimized according to the need for the specified equipment to have the power to reach as many people as the musical product can, rather than the musical depth. For this reason, after the first DAW was produced, it was an inevitable consequence of optimizing it to the needs outlined here.

2.2 The Structure of DAWs

2.2.1 Timing Mechanism

In various fields such as music production, sound design, film, and game development, Digital Audio Workstations (DAWs) provide an organized interface for editing and automating a wide range of audio tasks. A critical element of this interface is time organization, which is essential to understanding the overall concept. DAWs use the term BPM (beats per minute) as a global time reference, with a global transport[2] value that serves as a basis for applying other time signatures.

[2] https://docs.cycling74.com/max5/refpages/max-ref/transport.html

> Fundamental to the definition of music itself is that music must move through time—it is not static. Hence, music is sound organized through time. This organization of music through time is managed in the Western music system through time signatures. (Aichele, 2020)

Time organization in DAWs is based on musical note values and time signatures, with all edited elements on the interface dependent on each other in regular or irregular patterns. While complex rhythmic structures and time designs can be created on these interfaces, the resulting irregularity cannot escape the character of being based on regularity. In this sense, it is necessary to speak of a structure that is not based on regular time signatures and meters, but rather on a "microtemporal" structure (Fell, 2013, p.93) - the subdivision of the irregular small parts of the global transport - which is fully integrated into the entire interface. As mentioned above, interfaces of this kind have developed in a certain direction due to the influence of industry standards in music production. There is a strong connection between the rhythmic structure and the listener's engagement with the music they are listening to. Moreover, in auditory perception[3], we tend to hear the elements we hear as templates that we have culturally developed. In this sense, when we hear materials that conform to the patterns we are used to, our brain's processing of this information will be more efficient. This is why regular metric divisions have been integrated into DAW interfaces as a standard, as we can understand the reason for this.

> Metrical sequences, containing regular accents, induce a “good” clock that leads to efficient encoding of serial time intervals, whereas non-metrical sequences, with irregular accents, do not. (Hallam et al., 2014, p.126)

In this context, we can assert that music produced through DAWs may have a certain compositional and organizational structure. In addition, we can say that there is a correlation between the motivation behind the development of these interfaces and the motivation of music producers to create music using these interfaces. For example, in

[3] https://www.oxfordreference.com/display/10.1093/oi/authority.20110803095433742;jsessionid=1E0A23972E57660D3CAA42238509964E

order to sell more coffee to a group of people who have a coffee culture, certain formulas have been created and the most preferred ones have been determined (latte, macchiato, etc.). While a barista who makes coffee may prefer to make one of these coffees freely, a coffee machine that does this job may also be preferred. In this context, if the coffee produced adheres to a certain standard, it is more likely to be accepted by consumers.

2.2.2 Interface

In the field of music production, there are many different DAWs with varying styles. This situation can be interpreted as each interface having its own strengths and weaknesses. In its most basic sense, a digital interface provides the producer with ease in various dynamics such as temporal integration, editing options, arrangement, and composition form. It must be acknowledged that this provides a significantly large area of time for creativity in the music production process. However, while each DAW has its own unique interface, such a pre-prepared production space poses a risk of creating a stereotype in the music produced. It should be noted here that such interfaces provide an area where an individual can potentially become a virtuoso. However, in this case, a specialization developed by many individuals using the same interface will be at issue. Taking a guitar virtuoso as an example, the musical result produced can reach a high level of creative satisfaction and generate unexpected elements. However, by the nature of thing, it should be noted that there are limits to the options that can be produced.

2.2.3 Restrictions

Interfaces designed for content editing are created with specific production concepts in mind. In these production concepts, the intention is to allow users to modify only certain quantitative data in order for these interfaces to function optimally. While this increases efficiency in extracting data from the interface, it may overshadow the creative process. However, this does not imply that these interfaces hinder creative music production. It should be noted that the optimized and powerful aspects of these interfaces can contribute to the production process as additional time.

3 MODULAR DAW INTERFACE FOR LIVE PERFORMANCE

3.1 Overview

We have emphasized that all musical interfaces exhibit a focused, optimized attitude. At this point, the idea of flexible coexistence of these modules may come to mind. I would call it, in my own words, “interDAW communication”. In this context, we can think of all the interfaces that are on the agenda in computer music as an exhibition curation[4]. We can describe the process of combining mechanisms that operate differently from each other with variations that can have flexible functions, as a flood of thoughts for the curator who brings together the works of art to be exhibited in an avant-garde exhibition. Here, the concept of “modularity” comes to light in the context of digital music. In general terms, we can perceive a modular structure as a combination of many parts to form a higher function. For example, if we take the bookcase as a modular furniture, the combination of many sub-pieces creates the functionality of the bookcase. However, it should be noted that the lower parts have no function other than to bring the whole together (I exclude the new generation smart furniture from this example). What is meant by the modular approach here is that it is a part that is flexible in itself and at the same time is not fixed in its entire musical composition.

> …reuse of modules in new contexts allows more of the material's potential to be explored through new configurations rather than limiting it with a fixed

[4] Information on the "curation" metaphor can be obtained from this article. https://mymodernmet.com/what-is-curating/

> relation to other material. This is both a creatively valid position and an efficient use of composing time. (Saunders, 2008, p.154)

The existence of flexible models that can produce different functions will make the results that the current composition can produce wider and less limited. This is because each of them has a set of parameters with different dynamics in itself. In this context, we can predict that the music and live performances produced with these interfaces can create an independent but inter-related atmosphere each time.

3.2 The reason of personalized the interface

We mentioned above that DAWs are highly optimized and fast musical arrangement applications. However, moving in this production space risks staying within the boundaries drawn by these interfaces. The limits here mean that the digital signal processing (DSP) networks provided by the interfaces are minimally customizable. In addition, in the context of time management and rhythmic structures, it is not always desirable to be connected to the classical meter system. As we mentioned above, these applications are one of the most powerful pillars that make up the fluid structure of the popular music industry, as they can make musical output as optimized and focused as possible.

Creating customizable versions of this interface might mean removing the stereotype costume here. However, DAWs have a fairly fast and optimized routing system, although they have certain limitations in the creative process. At this point, it would be appropriate to bring up the term “interDAW communication” that I mentioned above. For a modular compositional structure that can be created, is it necessary to reconstruct certain apparatuses? At this point, my answer would be “no”. A modular interface, such as a personalized DAWs, can communicate with these applications in many ways such as routing, mixing. An important detail that I would like to underline here is again "musical interface curation". All these applications are designed as platforms where you can communicate with each other and use as many of them as you want. In this context, the new interface to be created can use the interface of many DAWs and reassign certain tools (within the scope of functionality).

3.3 Overall Content

In this section, I will talk about the content of my personal modular composition interface, and my interDAW communication strategies. Here I can start by first defining the concept of "module".

> …one of several parts of a piece of computer software that does a particular job.[5]

In this context, we can see each module developed for this interface as an application with independent dynamics. However, it should be noted that most of the models used in this design are tuned to produce integrated data. To develop these models, I used Max/MSP, an object-oriented programming language developed by Cycling '74.[6] Here, one of the reasons for my preference for Max/MSP is that it has a user interface where I can see the DSP flow[7] between the models more clearly. The most obvious reason why this is important to me is that it provides clear and accurate information to properly route communication with other DAWs. Another reason for my preference is that it has an object called "bpatcher", which I need in the modular approach, where I can create more than one module with different functions and call each one separately.

> Abstracts the contents of a patcher or sub patcher for use in other patchers, displaying only those visual elements which are desired. The number of inlets and outlets in a bpatcher object is determined by the number of inlet and outlet objects contained in its sub patch window. (Bpatcher Reference - Max 8 Documentation, n.d.)

Thanks to the inlets and outlets provided by each module bpatcher, it has become easier for me to design each module to be integrated with each other. This also

[5] https://www.ldoceonline.com/dictionary/module
[6] https://cycling74.com/company
[7] https://docs.cycling74.com/max5/vignettes/core/dsp_status.html

provided the possibility to use a module multiple times in the same interface, with presets with different dynamics. The modules to be used in this modular composition system are listed below:

- Global Phasor
- Equal Division of Octave (EDO)
- MC based Additive Synthesizer
- Granular Synthesizer based on MC[8] (Multichannel) library, and two buffer~ approach and polybuffer~ as a sound source.
- FM base number modulator
- Parameter distributor based on Machine Learning tools of FluCoMa[9] library.

All these developed models can be thought of as flexible programs that run on the max/MSP interface and have specific functions. Here the question may arise whether the interDAW communication element exists. However, all these modules are designed to be split into different channels and routed to other DAWs. What we mean is that the sounds produced by the modules will be sent on separate channels to Ableton for mixing and arranging. But how will this process work? Here, the BlackHole[10] application, which is a virtual audio loopback driver that can routing between interfaces within the computer, will be integrated into the system.

[8] https://docs.cycling74.com/max8/vignettes/mc_signals_newobjects
[9] https://www.flucoma.org/
[10] https://existential.audio/blackhole/

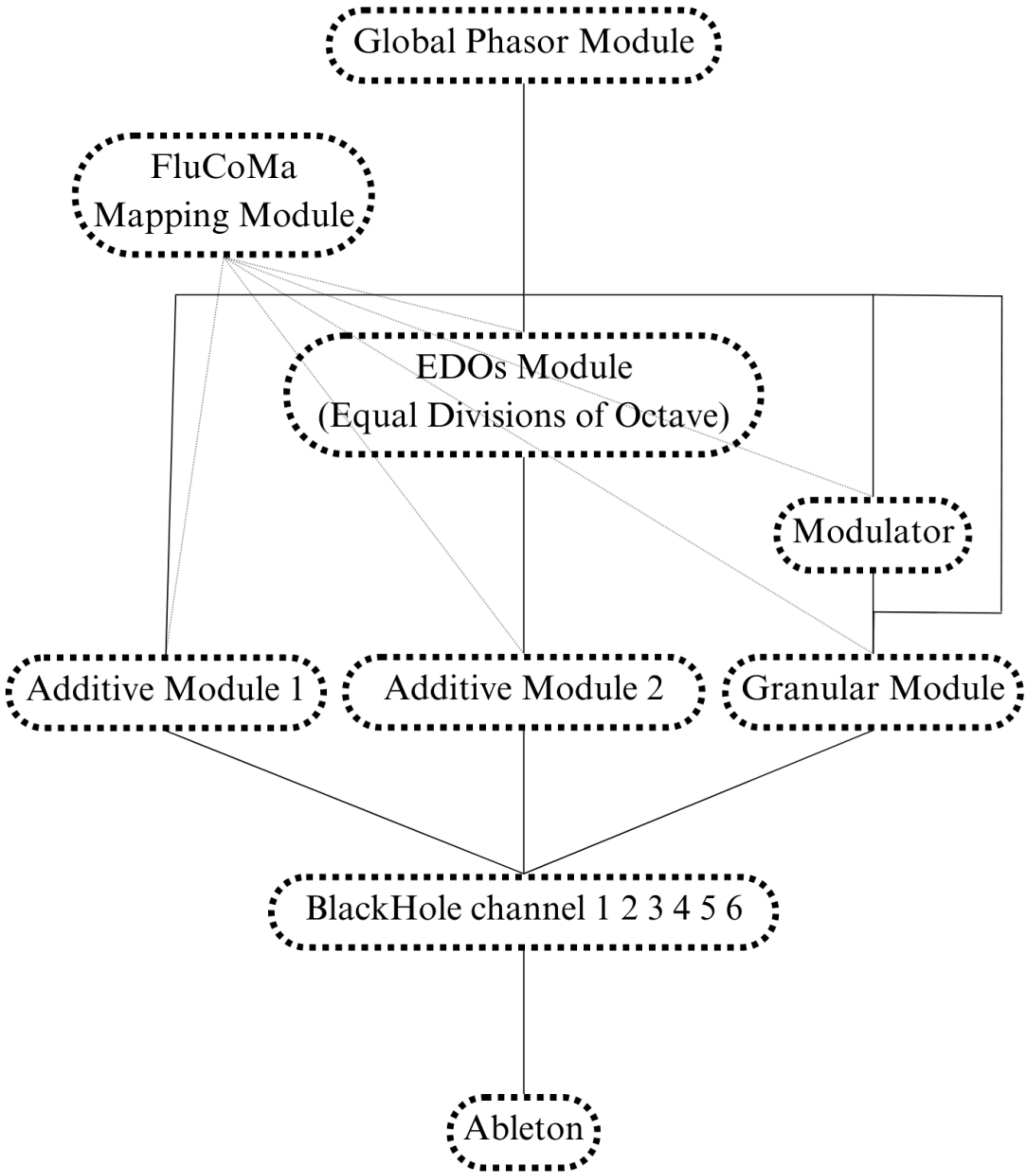

Figure 3.1: Basic demonstration of the modular system

4 THE CONTENT OF THE INTERFACE

4.1 Global Phasor

As we mentioned above to the time management in musical interfaces, we can remember that they are based on the classical meter approach. In this context, there are many time management options in the max/MSP interface created by the functional modules. The most basic of these options is the "metro" object. This object takes arguments in millisecond (ms) and if it is set to 1000, it triggers a bang (pulse) every 1000 ms.

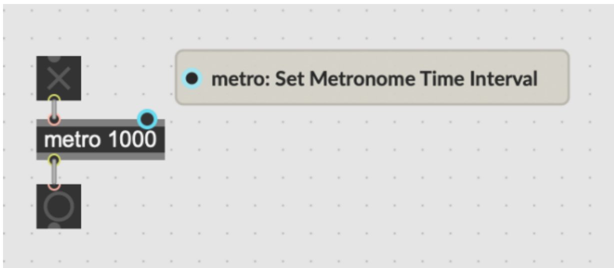

Figure 4.1: Basic metro construction

However, since this mechanism does not respond to the expected response in fairly fast configurations, I used the phasor~ object, which basically generates a sawtooth signal, for time management.

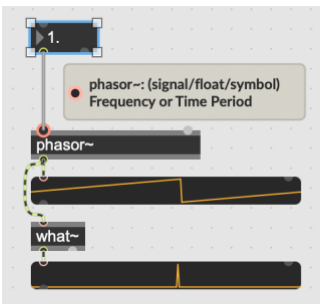

Figure 4.2: Basic phasor~ construction

4.1.1 Irregular Time Management

It should be noted that the Phasor~ object has a stronger side that it provides besides its optimizing work. The generated sawtooth signal can be divided into desired subsections with the subdiv~ object. This reminds us of the classical metric system. For example, if the value of phasor~ is 1 Hz, if we give the value of 4 as an argument to the subdiv~, the ramp produced by the phasor gives us a measure, beats in 4 subsections (Basically 4/4-time signature). However, cooperation of the phasor~ object with the what~[11] object, which can work in more depth here, may lead to a more micro temporal approach. Here, the what~ object generates impulses if the values (an array) exceed the threshold level. So how does this object communicate with the phasor~? Basically, the phasor ~ generates a ramp between 0 and 1, and when it reaches a value of 1, it immediately goes back to 0. Here, the what~ object is assigned a set threshold list between 0 and 1 and generates an impulse every time it exceeds the value.

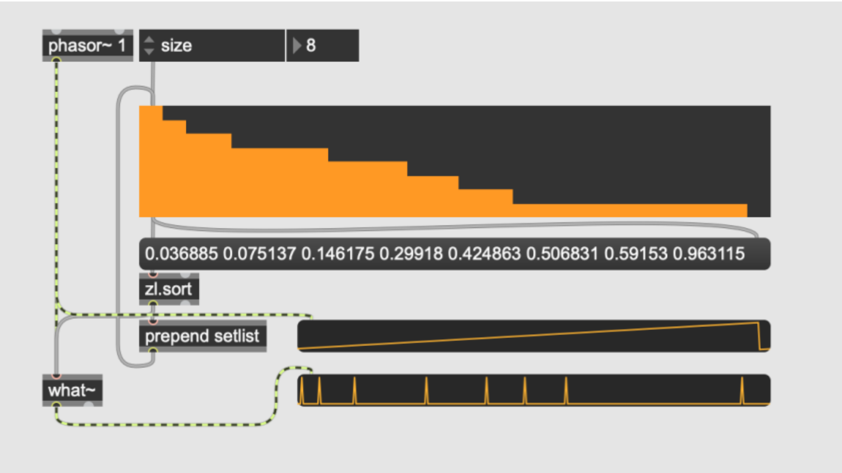

Figure 4.3: 8 elements array assigned to what~ object via multislider object

The phenomenon described in Figure 4.3 points to the basic time mechanism in the modular composition system. Basically, a mechanism can be observed here that can

[11] https://docs.cycling74.com/max8/refpages/what~?q=what~

split 1 measure into its very small particles. In addition, the multislider size is set to 8 in Figure 4.3. This 1 measure will divide it into 8 unequal parts. However, these divisions can be changed more "actively" according to the composition flow. Here, it promises more control than the grid system in DAWs. The main goal here is to control all modules irregularly with a common time mechanism. The irregular at first sense here can be perceived as random times. However, what is meant by irregularity here is to avoid the repetitions of sounds that move away from the classical meter system (grid structure) and move regularly on top of each other. Again, it should be emphasized that here we mean a *regular control over disorder.* As stated above in the time mechanism of DAWs, the musical contents produced by all producers by sharing a similar time structure carry a sub-signature of DAW in a sense. It can be said that a more personal time control mechanism is aimed here.

4.2 Equal Division of Octaves (EDO scales)

When we hear musical content, the first thing that comes to our mind is undoubtedly the sheet music. As we know, there are 7 pitch values for dominant music notation. The 8th note is known as the octave of the 1st note. In fact, if we take into account the half notes, 1 octave is divided into 12 equal parts.

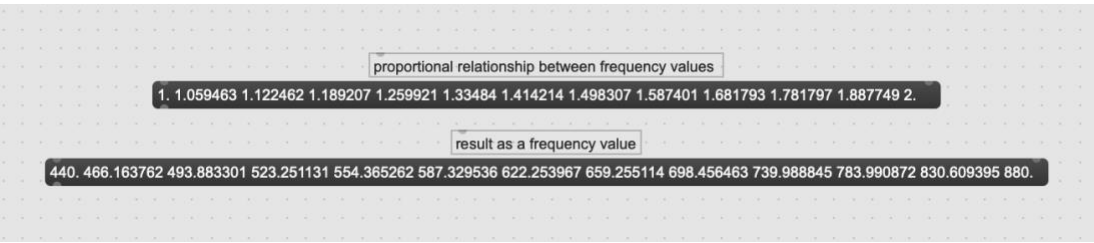

Figure 4.4: Frequency values of all 12 notes between A4 and A5

As we can see from Figure 4.4, the distance between A4 (440) and A5 (880) is divided into 12 equal parts. As we understand here, it is the proportional equalization of all interval values in the scale. Expressively, this is called the Equal Division of Octave[12]. Here, it is called “12edo”[13]. We can see that the approach here helps us to understand the relationship between the frequency values that come to the fore in musical harmony. Here, in a fundamental sense, it would not be wrong to bring up the relationship of musical scales with one's cultural backgrounds. 12edo, which is the dominant musical scale, has a great influence on the way people perceive the proportional relations of the notes in the musical content in many ways. What we mean here is that the proportional relationship between these frequency values plays a very important role in the auditory perception process, when we decide whether a music will sound good or bad.

> We tend to recognize patterns of pitches that form melodies. We do this presumably by recognizing the musical intervals between successive, and most of us seem relatively insensitive to the absolute pitch values of the individual note, so long as the pitch relationships between notes are correct. However, exactly how the pitch is extracted from each note and how it is represented in the auditory system remain unclear, despite many decades of intense research. (Oxenham, 2013, p.9)

4.2.1 Randomizing issues

In the world of sound synthesis, we know that randomly generated numbers have a wide range of uses. So, by the word "random", do we mean a number generator that is completely random? At this point, we need to introduce the concept of pseudo number generators (PRNGs[14]). Basically, computers implement algorithms based on PSNGs, rather than getting random data from a random outside source. Therefore, no matter how many random numbers generated, after a while it will start to repeat over its original weave. So why is this pattern important? Because, in this sense, the “seed” value of the random numbers produced is important because of its power to reproduce

[12] https://en.xen.wiki/w/EDO#Links_and_articles
[13] https://en.xen.wiki/w/12edo
[14] More information for PRNGs https://docs.oracle.com/en/java/javase/17/core/pseudorandom-number-generators.html#GUID-08E418B9-036F-4D11-8E1C-5EB19B23D8A1

the same result when you call the generated pseudorandom list in another digital environment.

> What is needed is a sequence of numbers that has the properties of RNs but that is the same every time the program is run; this allows coding errors to be found and the same results to be produced when the same code is run on different computers. (Dunn & Shultis, 2023, p.56)

Although I have presented arguments above that it is possible to control the production of random numbers, I would say that my approach to its use in the production of musical content is more sceptical. What do I mean by scepticism here? Helping to read predetermined scale lists instead of directly contributing to the production of melodic/motivic values in musical content may mean less and more efficient intervention in the creative process.

4.2.2 Discovering micro and macro tonal scales

In the specified scale lists, the intent is to create a predetermined musical scale list. As we mentioned above, the equal division of octave approach is to divide the octaves into specific sections. 12edo is one of the most dominant examples of this. In this context, decreasing and increasing the division value will take us to the world of different scale lists. If we increase the number of divisions - 22edo[15] - this will get us to microtones, if we decrease it - 8edo[16] - this will get us to macrotones.

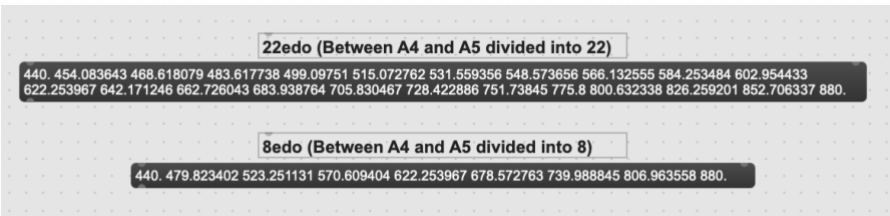

Figure 4.5: An example for 22edo and 8edo scale

[15] https://en.xen.wiki/w/22edo
[16] https://en.xen.wiki/w/8edo

4.2.3 Formula

EDO scales are basically based on how many steps the distance between two octaves is. It should be noted here that the number of steps is calculated on the basis of the frequency ratio. This is important for the following reason. MIDI protocols have not developed a step for each step that EDO scales calculate. Because the octave is divided into 12 equal parts, and if we divide it into 22 equal parts, it will not be applicable for many MIDI notes. The number of steps is determined as n and which step is to be calculated as k[17].

$$c = 2^{k/n}$$

According to the formula here, 3/12 (k/n) will give us step 3 of 12-edo. It should be noted here that the greatest strength this formula will provide to the modular interface will be the active control over the main character of scale. Imagine, in a generative structure, note values that are constantly changing, moving according to MIDI values, this will give you quite a wide range of possibilities. However, if these changing note values and the ratios on which they are based change, more unpredictable possibilities will emerge.

Figure 4.6 shows the application of the EDO formula in the max/MSP interface. At this point, each time the division number is determined, it assigns an index (n) value to the uzi object. Every time the uzi receives an index value, it instantly sends a list from 0 to n. Each element of this list represents the step (k) available to us. These available step values (k) are divided by the division number (n) each time. These two values are sent to the pow object. The current pow value is set to 2. In other words, every time a value is assigned, 2 is assigned as a power value with this value. This process is listed with the zl.group object. However, it should be noted that the calculations here are determined as frequency ratio (between 1 and 2). This is not an audible frequency range. At this point, each element in this list is multiplied by its fundamental frequency value, and the n-edo scale is fully calculated.

[17] https://en.xen.wiki/w/EDO#Formula

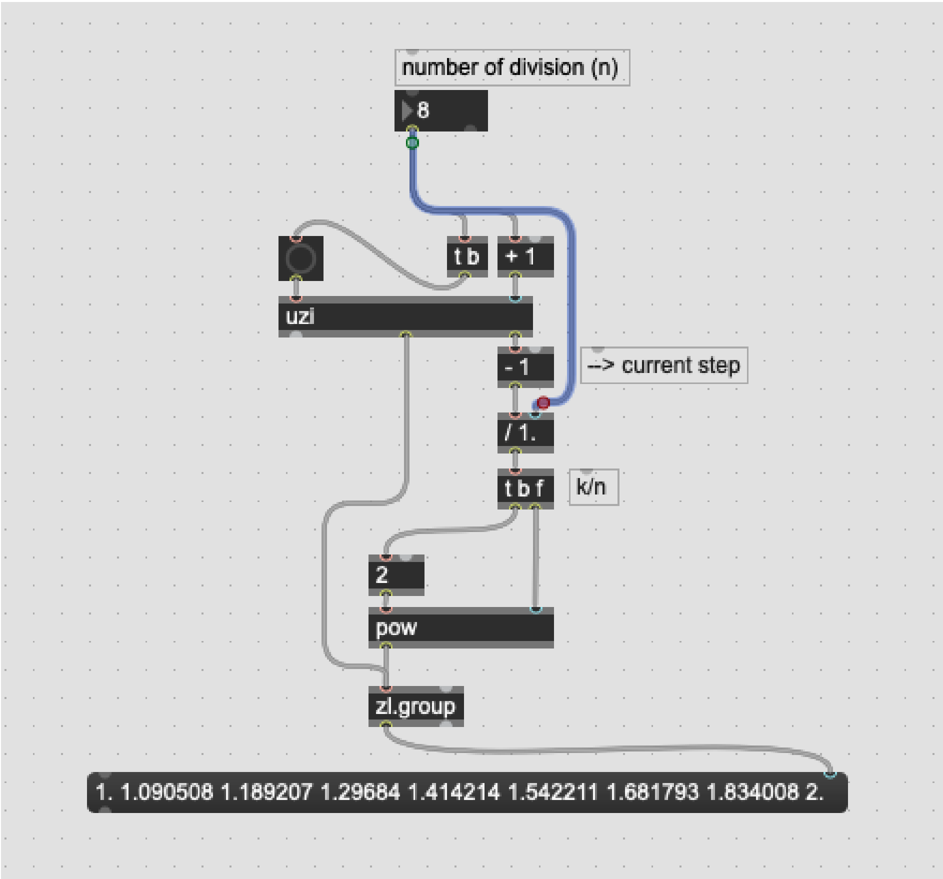

Figure 4.6: Implication of formula in max/MSP

4.3 Sound Sources

4.3.1 Multichannel (MC) based additive synthesis

Additive synthesis, which is one of the most basic sound synthesis approaches, is one of the most comprehensive sound sources of the modular interface at this point. In this approach, the basic element consists of the sum of sine waves. In this approach, I consider the frequency domain approach important, which shows which frequencies have which amplitudes, based on the information of the frequency collections whose snapshots have been taken. In this approach, while the x-axis shows the frequency values, the y-axis shows the amplitude information of those frequency values (Cipriani & Giri, 2010, p.192). Here, the fundamental frequency is determined by the EDO scale module. Since the visual representation gives strong clues to the sound space in this sense, I used certain visual sonograms in the modular interface. One of them is the frequency domain visualizer programmed on the multislider[18], shown in Figure 4.7.

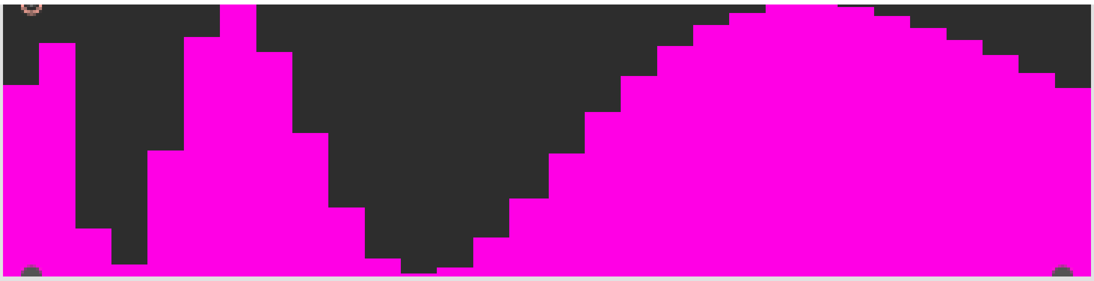

Figure 4.7: Demonstration of multislider sonogram

4.3.1.1 Advantages of the MC library of Max/MSP

The additive expression mechanism here is developed on multichannel (MC) objects integrated into the max/MSP interface. The most important reason for this is that the addition processes are very optimized and fast. Optimization here means that adding 30 sine waves one by one will push the audio processing into a less optimized area. More importantly, a new mechanism will have to be integrated if a separate operation is desired for each added partial. If we consider a pure sine wave, it consists of a single frequency wave. However, as frequencies are added, the waveform becomes more complex. Here, when the fundamental frequency (lowest component frequency) is multiplied by integer multiples, we get harmonics (Cipriani & Giri, 2010, p.195). Here the power of MC objects will come into play. MC objects provide us with options for harmonic series that we can create with the fundamental frequency. In this sense, a very strong argument for MC objects will come into play. The @chans[19] argument determines how many waves are added to the system at once. The mc.sig~ object, which is one of the basic parts of my site, takes arguments that can control all these added waves at the same time. These arguments are illustrated in Figure 4.8.

[18] https://docs.cycling74.com/max7/refpages/multislider
[19] https://docs.cycling74.com/max8/refpages/mc.sig~#Arguments

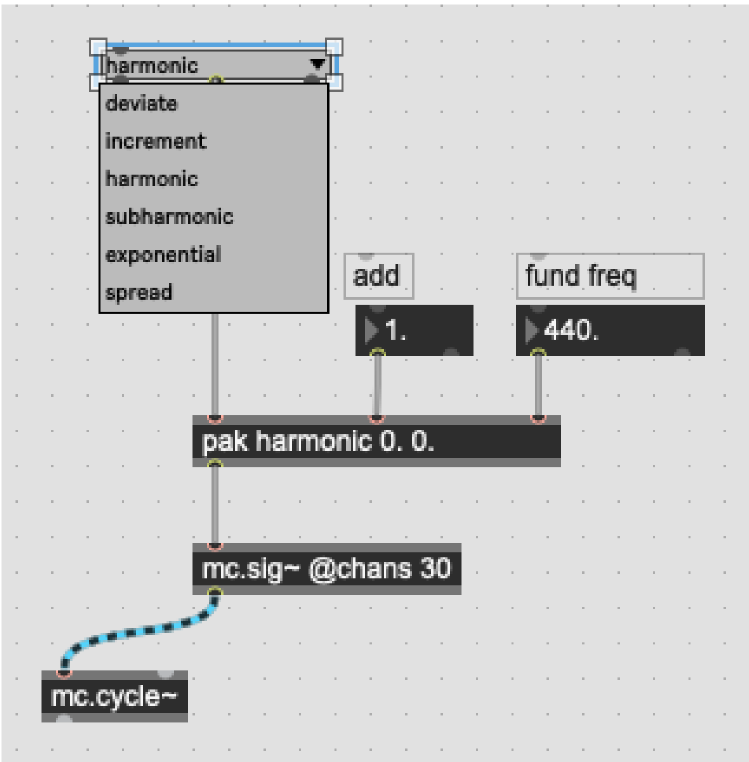

Figure 4.8: mc.sig~ arguments' option to apply entire channels

4.3.1.2 Distinct panning

We mentioned above how to process all generated waves (here 30 waves). However, they were concerned with the quantitative elements and amplitudes of the waves. So where will all of these waves be heard in space? This is where the mc.mixdown~ object comes into play. With the first argument it takes, this object determines how many separate channels the generated waves will be sent and where the sent channels will be located in space. It makes the positioning over values between 0-1. This could be positioned somewhere between 0-1 in individual space if 30 channel waves were generated. Here again, the power of creating the desired number of waves and assigning a value to all of them is used in the mc.sig~ object. The panning relative here is described in Figure 4.9.

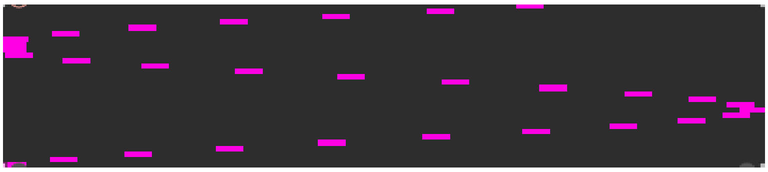

Figure 4.9: Illustration of 30 channels pan position

4.3.2 Multichannel based granular synth module with gen~ object

The second sound source in the modular system is the granular synthesis module. Granular synthesis can basically be defined as the mechanism of dividing a complex waveform into thousands of small pieces with a duration of 1-100 ms and processing each of them into these sub-forms. In this sense, grain is the name given to these small pieces of 1-100 ms each.

> By combining thousands of grains over time, we can create animated sonic atmospheres. The grain is an apt representation of musical sound because it captures two perceptual dimensions: time-domain information (starting time, duration, envelope shape) and frequency-domain information (the pitch of the waveform within the grain and the spectrum of the grain). (Roads, 2002, p.87)

It should be noted that splitting and processing the grains in an optimal way will make the sound output more audible and efficient. Here I am using the max/MSP MC objects library to manipulate the waveform. Information about the efficiency and optimization of MC objects is given above. There are two main elements in this mechanism. The first of these is a module that divides the waveform into grains and assigns their quantitative values, the second is the module that reads these modules at the desired speed and manner and generates triggers for audio output. In these two modules, the mc.gen~[20] object is used. The reason for this comes from the power of the modular approach we mentioned above to exceed the limits. The modules produced for each action not only increase the possibilities but also facilitate the optimized operation of the system. In Figure 4.10, granules engine density value and reading speed are adjusted. This is provided over the wave that the mc.sig~ object produces 30 channels.

[20] https://docs.cycling74.com/max8/refpages/mc.gen~?q=mc.gen

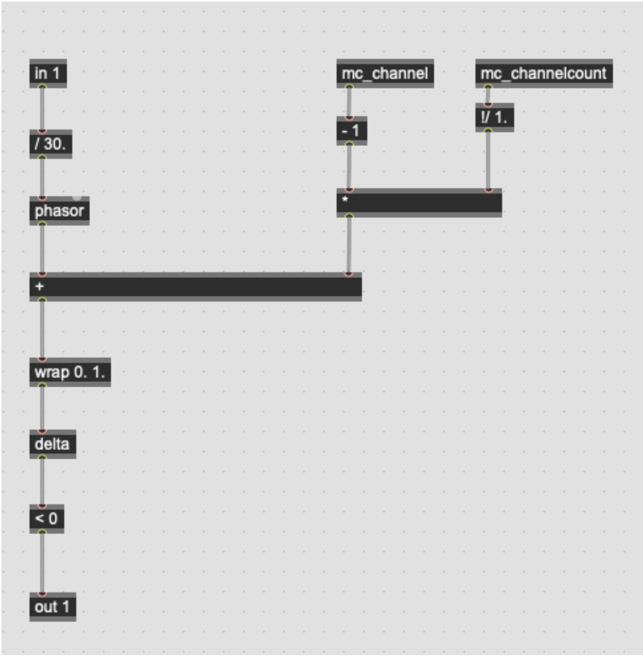

Figure 4.10: Readback and trigger module with mc.gen~

It should be said that all grains produced here are in raw form. Raw means that the grains obtained as a result of dividing the given waveform into sub-parts, have a high risk of producing clips according to their speed and pitch values when replayed. For example, if one grain ends at n and another begins at n +/- a, the transition will not be smooth.

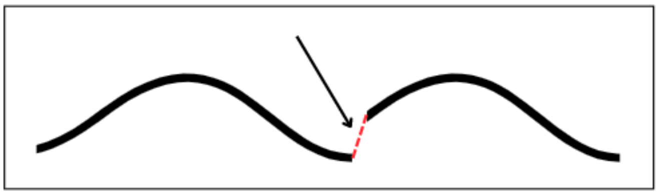

Figure 4.11: unmatched two grain

However, the waveforms of grains may differ from each other. It may be a fixed waveform that does not change over the grain duration, or it may be a time-varying waveform (Roads, 2002, p.90). So, the precaution to prevent clips will require an adjustment for all of them. Suppose you have 2000 grains. Is it possible to make separate settings for all of these? At this point, my answer is not yes or no. Here it is important to mention *two buffer~* approaches that I use in the system. In this approach, the given waveform will be divided into grains and loaded into a created buffer~. The grains loaded here will be identified through the mc.channel[21] object shown in Figure 4.10. All these grains will then be reloaded into a newly created buffer, each passing through a gaussian envelope generator.

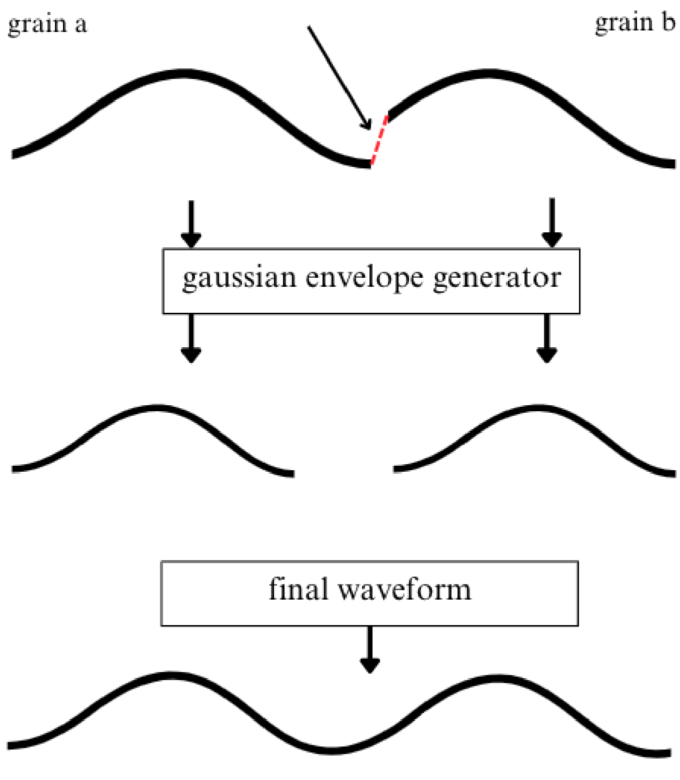

Figure 4.12: Matched grains via gaussian envelope

[21] If used within a patcher inside mc.gen~, the mc_channel operator will return the current channel index. Otherwise, it always returns 1.

4.4 Controlling Mechanism

Since all digital music interfaces are tools produced in a virtual environment, musical parameters play an important role. All parameters in the created musical interface can be controlled on the computer and with external controllers. Even if the systems created are highly generative structures, certain values must be entered as data. Parameter controls have a very strong effect on the character of the sound atmosphere created by the musical interface. For this reason, the control mechanism will need to be established and adjusted for each parameter to suit where it will intervene. All parameters of the system can be controlled separately, as well as all parameters can be controlled via a single channel/tool. At this point, I can say that I have adopted a hybrid approach in the interface created here. Besides a structure where most of the parameters can act jointly from a pre-scaled system, some individual parameters are left manually.

4.4.1 Interpolation Among Pre-sets via FluCoMa Tools

How to control the parameters of all modules in the modular interface creates a personal preference model. Imagine a synthesizer model, although it is a single product, it has millions of users. However, the musical result varies from person to person. At this point, we said that how to map parameters is of critical importance. In the modular interface created here, the parameters are designed as "modes". In other words, modes are the result of all parameters moving simultaneously from one direction to the other. So, how will it be possible to move from one mode to the other when all the parameters are changing at the same time? This is where the concept of interpolation comes into play. In this sense, interpolation[22] is a method of generating predictive new data sets for transitioning between given abstract data regions.

Here, two different tools will come to the fore; controlled parameters and control interface. In this sense, I used the MPLRegressor toolkit of the Fluid Corpus

[22] https://www.indeed.com/career-advice/career-development/how-to-interpolate

Manipulation Project (FluCoMa)[23], which develops important and powerful tools on machine learning, to communicate between these two main elements.

> The MLPRegressor is a neural network that can be used to perform regression. In machine learning, regression can be thought of as a mapping from one space to another where each space can be any number of dimensions. "MLP" stands for multi-layer perceptron which is a type of neural network. (Tremblay et al., 2022)

The controlled parameters are represented by 24 values of the modular interface. The controlling interface represents a pitchslider with x, y coordinates. In this context, when the desired "mode" is provided, this pin point will be marked in the data system. In other words, the x, y coordinates of the pitchslider and the values of 24 parameters will be assigned to this pin point.

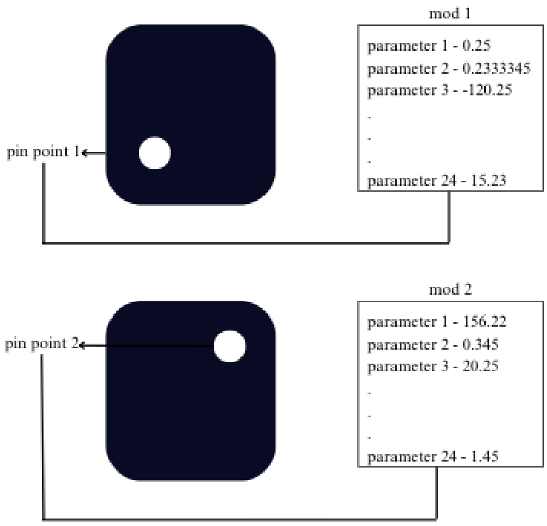

Figure 4.13: Communication process of two data sets

[23] https://www.flucoma.org/about/

Here, the task of the MPLRegressior~ object is to predict what the 2 input values (x, y) and 24 parameters (output) might be. In summary, an id is assigned to each pin point. Inside that id are both the x, y coordinates and the value of the 24 assigned parameters. At this point, if you want to switch between pin point 1 (mode 1) and pin point 2 (mode 2), MPLRegressor~ makes the most optimized interpolation possible between these two data sets.

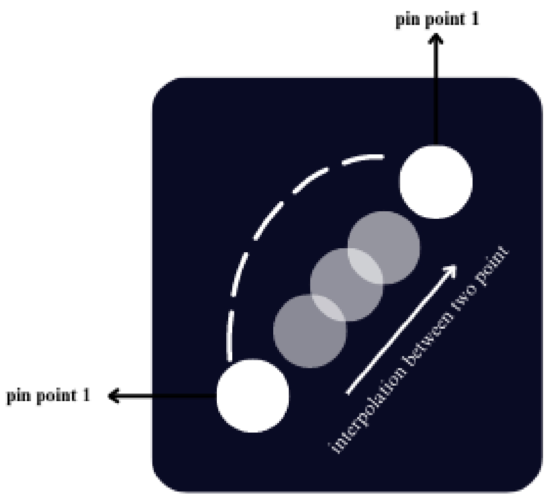

Figure 4.14: Interpolation between two modes

5 SOUNDS

The "modes" in this project represent abstract names given to the atmospheres evoked by the current set of parameters. Therefore, when the modular interface is used by others, it implies that the "modes" can be personalized. It should be noted that the two-dimensional plane provided by the pitchslider object gives hints about future possibilities. In spatial applications, this interface can enable individuals to design sounds specific to certain areas by integrating sensors into their surroundings. In this sense, it is important to emphasize that this project is still in the process of development for more interactive experiences. One of the future goals of this project is to involve the listener in the production process, thereby making the connection between the music medium, the producer, and the listener more visible. This interaction entails creating an online space where individuals can participate using their mobile phones or integrating their movements captured by cameras in specific areas into the system. However, as observed in previous projects, this interaction process tends to create a disjointed connection between the producer and the listener. In other words, the listener cannot accurately perceive how their movements affect the sonic atmosphere they engage with. It is not intended for the listener to possess technical knowledge, but rather to establish an organic connection between their movements and the evolving atmosphere. Therefore, this aspect is still in the developmental stage. The sounds included in this project can be seen as a preface to the interactive sonic atmosphere mentioned.

The sounds[24] of the modular interface have been added to the link using many pre-sets in collage.

[24] The link for the final modular interface. https://youtu.be/gzHjXZHHAXA

6 CONCLUSION

In this study, we explored the domain of modular sound synthesis and control, aiming to provide flexibility, creativity, and effective parameter control. By focusing on two primary sound sources, namely multichannel-based additive synthesis and granular synthesis, we developed a modular interface that empowers musicians and artists to create dynamic and immersive sound atmospheres. The multichannel-based additive synthesis module offered a comprehensive sound source by combining sine waves, while the granular synthesis module divided complex waveforms into smaller grains for further manipulation. The integration of optimized and fast multichannel (MC) objects facilitated efficient summation processes and harmonic series generation.

To ensure smooth transitions between grains and prevent clipping, we employed a buffer~ approach. The waveforms were divided into grains and loaded into a buffer~, which were then reloaded with a Gaussian envelope generator. This approach provided optimized playback of grains, regardless of their speed and pitch values. Furthermore, we adopted a hybrid control mechanism that allowed both collective and individual parameter control. The system parameters were designed as "modes," representing the simultaneous movement of all parameters in a specific direction. To facilitate mode transitions while simultaneously changing all parameters, we utilized the MPLRegressor~ tool from the Fluid Corpus Manipulation Project. By predicting parameter values based on intermediate points, an optimized interpolation between data sets was achieved, enabling smooth and seamless mode transitions.

Through this research, we have demonstrated the significance of parameter mapping and interface control in the modular environment. By defining specific modes within the system, we facilitated optimized interpolation and smooth transitions between different sonic characteristics. The modular interface presented in this study serves as a powerful platform for musicians and artists, offering them the opportunity to create dynamic and impactful soundscapes.

In conclusion, this project contributes to the field of sound synthesis and control by combining additive and granular synthesis techniques within a modular interface. The integration of optimized multichannel objects, buffer~ approaches, and interpolation mechanisms enhances the creative potential and efficiency of the system, empowering users to explore the realm of generative music and unlock new sonic possibilities.

REFERENCES

Thompson, C. (2016, January 6). *How the Phonograph Changed Music Forever*. *Smithsonian Magazine.* https://www.smithsonianmag.com/arts culture/phonograph-changed-music-forever-180957677/

***A Brief History Of Sound Recording*.** (2022, July 7). Pirate Studios. https://pirate.com/en/blog/history-of-music-recording/

Aichele, M. (2020). *Understanding Time Signatures and Meters: A Musical Guide*. Liberty Park Music. https://www.libertyparkmusic.com/musical-time-signatures/

Fell, M. (2013). *Works in sound and pattern synthesis: Folio of Works* (thesis). University of Surrey, Surrey.

Hallam, S., Cross, I., & Thaut, M. H. (2014). *The Oxford Handbook of Music Psychology*, Second Edition. In Oxford University Press eBooks. https://doi.org/10.1093/oxfordhb/9780198722946.001.0001

Saunders, J. C. (2008). *Modular Music. Perspectives of New Music*, 46(1), 152–193. https://doi.org/10.1353/pnm.2008.0015

Oxenham, A. J. (2013). *The Psychology of Music* (Third Edition): The Perception of Musical Tones (D. Deutsch, Ed.; 3rd ed.), 1-33. Academic Press.

Dunn, W. F., & Shultis, J. K. (2023). *Pseudorandom Number Generators: Exploring Monte Carlo Methods* (Second Edition). In Elsevier eBooks (2nd ed., pp. 55–110). https://doi.org/10.1016/b978-0-12-819739-4.00011-1

Cipriani, A., & Giri, M. (2010). *Electronic music and sound design : theory and practice with Max/MSP (Vol. 1).* http://ci.nii.ac.jp/ncid/BB05768525

Roads, C. (2002). *Microsound*. In The MIT Press eBooks. https://doi.org/10.7551/mitpress/4601.001.0001

Tremblay, P.A., Green, O., Roma, G., Bradbury, J., Moore, T., Hart, J., & Harker, A. (2022) *The Fluid Corpus Manipulation Toolbox* (v.1). Zenodo. doi: /10.5281/zenodo.6834643

bpatcher Reference - Max 8 Documentation. **(n.d.).** https://docs.cycling74.com/max8/refpages/bpatcher?q=bpatcher