



University of HUDDERSFIELD

University of Huddersfield Repository

Pavlitski-Bell, Mortimer

An Exploration of Creative Audio Spatialisation Tools for Ableton Live

Original Citation

Pavlitski-Bell, Mortimer (2021) An Exploration of Creative Audio Spatialisation Tools for Ableton Live. Masters thesis, University of Huddersfield.

This version is available at <http://eprints.hud.ac.uk/id/eprint/35494/>

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: E.mailbox@hud.ac.uk.

<http://eprints.hud.ac.uk/>

*AN EXPLORATION OF
CREATIVE AUDIO
SPATIALISATION TOOLS FOR
ABLETON LIVE*

Mortimer Pavlitski-Bell

Music MA | The University of Huddersfield

ABSTRACT

Electronic music composers working within Ableton Live lack integrated spatialisation tools that give global control over spatial behaviour. Popular spatialisation tools like GRM Tools Spaces (2011) and Ableton Live's Surround Panner are tied to specific speaker layouts presenting several drawbacks. Firstly, the tools cannot be chained together as is standard practice with stereo plugins, limiting their creative potential. Secondly, Ableton Live channels are restricted to stereo, making the setup of these tools a complicated and slow process, requiring many additional channels to route spatial audio signals. Other spatialisation tools such as the IEM Plug-in Suite (2020) and Envelop (2020) use ambisonics to enable the chaining of effects but are not sufficient for composers, primarily due to their utility-focused nature or unintuitive user interfaces.

This thesis proposes a solution utilising the Max for Live device format and 5th order ambisonic audio encoding to decouple the spatialisation from a specified speaker layout and enable chaining of spatial effects. The new tools integrate effects into the spatialisation process and enable a more rapid workflow for composers. Audio examples demonstrate the creative potential of the tools.

“Come with us now on a journey through time and space...”

– The Mighty Boosh

Acknowledgements

Special thanks to Alex Harker and Sam Findlow.

TABLE OF CONTENTS

Abstract.....	1
<i>Section 1: Introduction.....</i>	<i>7</i>
<i>Section 2: Aims.....</i>	<i>8</i>
2.1 Artistic Vision & Inspiration	8
2.2 Research Aims.....	9
<i>Section 3: Methodology.....</i>	<i>10</i>
3.1 Creative Research	10
3.2 Iterative design process.....	13
3.3 User Feedback	14
<i>Section 4: Literature Review</i>	<i>16</i>
4.1 A Lack of Creative Spatialisation Tools?	16
4.2 Why create new tools now?.....	19
4.3 What technologies are best suited for new tools?	20
Localisation.....	20
Panning.....	20
Object-Based Panning	21
Sound Field (Ambisonic format).....	22
Binaural.....	23
<i>Section 5: Project Development.....</i>	<i>25</i>
5.1 Technical Challenges.....	25
Why Max for Live?.....	25

Audio Bridge	25
Encoder and Decoder	26
Hosting VSTs	26
5.2 Creative Decision Making	27
Optimising Spatial Clarity	27
Relationship of Artistic Intention	28
5.3 Parameters, UX & UI Design	29
What parameters should be exposed to the user?	29
How should they be exposed to the user?	29
Are the controls Intuitive to navigate?	32
How does that fit in within the 169px height limit?	33
Section 6: Outcomes	35
MTools for Live (MT4L)	35
MT4L Pan	35
MT4L Master-Binaural	36
MT4L Master-Spiral	36
MT4L Granulator	37
MT4L TapDelay	38
MT4L Smudge	38
MT4L AmbiDelay	39
Section 7: Conclusion	41
7.1 Critical Discussion	41
Artistic Intentions	41
Development Process	42
Comparison with Pre-Existing Works	42

User Feedback	43
Audio Examples	44
7.2 Future Work	45
Further Composing	45
Creative Potential	46
Further Research	46
<i>Bibliography</i>	47
<i>Discography</i>	50
Research Inspiration Playlist	50
<i>Software References</i>	51
<i>Appendix A</i>	53
<i>Appendix B</i>	54

SECTION I: INTRODUCTION

This thesis documents the creation of a suite of ambisonic Max for Live devices designed for rapid creation of sonically complex spatial scenes. This document opens with, an introduction, the aims of the research and methodology followed by a literature review and breakdown of the project's development and closes with the project's outcomes, evaluation and conclusion.

The research was undertaken by Mortimer Pavlitski, a Music MA student at the University of Huddersfield under the supervision of Alex Harker. Mortimer's undergraduate degree involved, DSP programming in Max and C++, multichannel Max for Live devices, electronic music composition, a procedurally generated video game and placement at IRCAM and Ina-GRM. Interested in spatialisation, he wants to encourage spatial audio composition by creating easy to use spatial tools.

SECTION 2: AIMS

2.1 ARTISTIC VISION & INSPIRATION

Since the early explorations of spatial audio by Varese, Stockhausen, Schaeffer, Poullin, Bayle, Chowning and onwards (Kendall, Cabrera 2011) many tools have enabled the artistic shaping of sound in space. Modern tools like SPAT (2019) and GRM Tools Spaces (2011) enable great flexibility for accurate panning of multiple sound sources but are time-consuming to use when creating complex spatial scenes. Beyond simple jitter and smoothing, they lack features that can move sound sources intelligently and with respect to a soundscape as a whole. Composer Robert Normandeau talks about the difficulty of editing sound trajectories in the software Zirkonium “*At the moment, the only way to design trajectories in Zirkonium is to write every movement line by line, which is not adequate for complex movements.*” (2009, pp.284). This project investigates a more rapid, behaviour-based solution for creating such spatial variance.

I want music to take me on a journey through space as well as traditional musical development through time. The track ‘//’ by Second Woman (2017) opens with an iterative sound object that decays away with detailed timbral nuance but lacks spatial movement. I can hear how the sound wants to shoot off on different paths around the room. Similarly, the opening of ‘Screen’ by Visible Cloaks (2017) musically explodes out in all directions but only in one dimension spatially, left-right stereo. I want to feel connected to the life of a sound by the way it moves around. These are the kinds of tracks that inspire me. Spatialisation is another dimension to express artistic intent. What would it be like to stand in the eye of a tornado or rainstorm not made of rain?

This project aims to create a suite of Max for Live tools that give me artistic control over the spatialisation of a sound. I am creating these tools for Ableton Live as this is the DAW I use most in my creative practice. The created devices should work well within the Ableton Live environment to enable mapping and recording of parameters and the saving of devices as a presets. Seamless integration can expand the capabilities of the spatial tools further, as they work in combination with future and pre-existing tools like LFO and Envelope Follower. Envelop for Live (E4L) already provides devices that satisfy

some of the project's aims *E4L Delay Boids* and *E4L Brownian Delay*. However, they fall short for me as a creative due to the underdeveloped user interfaces. The spatial controls have highly technical labels with no hierarchy of importance and minimal grouping, making the interface's challenging to understand and difficult to navigate quickly.

In the studio, I want to have a creative relationship with spatialisation where sounds can move independently of my control. The author of one of my inspirational tracks Joshua Eustis of the band *Second Woman* states "*exploring time distortion, time dilation, space dilation*" (Wilson, 2018) as a mission statement for their earlier records. I would like to build on that by investigating not just how a sound's timbre develops over time but also the relation to its spatial trajectory. Sound sources in the natural world move all around us just as we move through them; I want to explore this creatively in my music. My work intends to simulate some of the emergent phenomena allowing me to mimic the rich spatial movement of the natural world. Leveraging the power of a modern computer, it is possible to intelligently generate limitless variation, reducing the time taken to create interesting results. Automating the panning positions of the sound allows me to focus on the macro variation, rather than manually controlling an excessive number of parameters. Removing the concern for details allows me to focus on the artistic qualities of the overall soundscape.

2.2 RESEARCH AIMS

I am to create and evaluate a suite of Max for Live devices that

- Create new sonically interesting spatial scenes.
- Give macro control over the spatialisation, minimising the time required to produce different versions of complex results.
- Simulate some of the natural world's intricate spatial movement, including bird flocking or animal herd behaviour.
- Have a clear and intuitive UI design that gives meaningful feedback about how different parameters affect the sound, allowing users to be in a creative headspace while navigating the underlying DSP engine.
- Create tools that integrate fully into the Ableton Live environment enabling an uninterrupted flow of work.

SECTION 3: METHODOLOGY

I have used a creative research methodology guided by an iterative design process supplemented with user feedback. I have decided to use an iterative design process to enable the flexibility to explore new ideas as Smith & Dean state “This feedback loop between speculation and experimentation is fundamental to research in many disciplines, and is also appropriate for research in the creative arts.” (2009, pp.154). The user feedback will reduce personal bias during development.

3.1 CREATIVE RESEARCH

Working in the *Spatialisation and Interactive Research Laboratory* (S.P.I.R.A.L) at the University of Huddersfield, I began to distil what I wanted to achieve sonically in space after questioning how a sound could move around the studio. I started by sketching out some of my ideas on paper (see figures 3.1 to 3.7) and consolidating the rest of my inspiration for the research into a playlist of music (see *Discography*), and a list of more philosophical themes to explore: *trajectories, lifespan, space, time, evolution, decay, systems and process, the natural world and scientific phenomena*. Using my sketches as a base, I refined the ideas further until I was left with a list of more concrete spatial effects to explore.

** representing my original selection of effects to create.*

*** Spatial Path Trajectory Delay**

Delay sound on a trajectory around the sphere to specified positions with unique effects at each point.

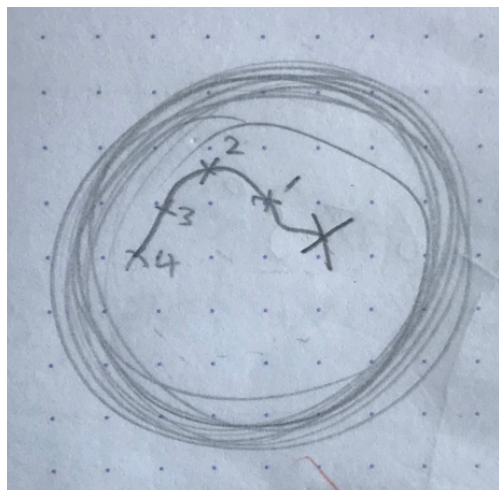


Figure 3.1 Initial sketch of Spatial Path Trajectory Delay

*** Spatial Boid Swarm Granulator. (Reynolds, 1987)**

Using Reynolds Boid algorithm as a base, simulate heard behaviour by panning grains of a granulator to their position in the sphere.

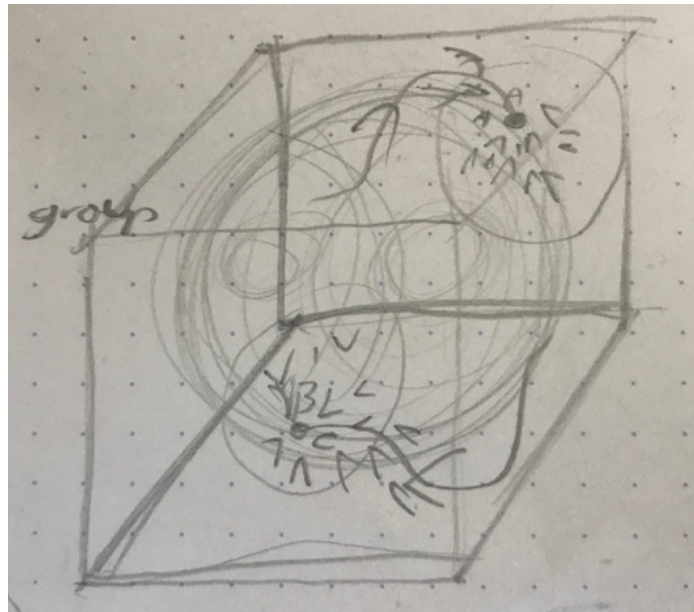


Figure 3.2 Initial sketch of Spatial Boid Swarm Granulator

Envelope Follower/MIDI Trigger Stochastic Panning.

Pan a sound to a random point on the sphere using an envelope follower or MIDI trigger.

*** Spectral-Spatial Smearing.**

As a sound moves different frequency lag behind it smearing the sound across the sphere.

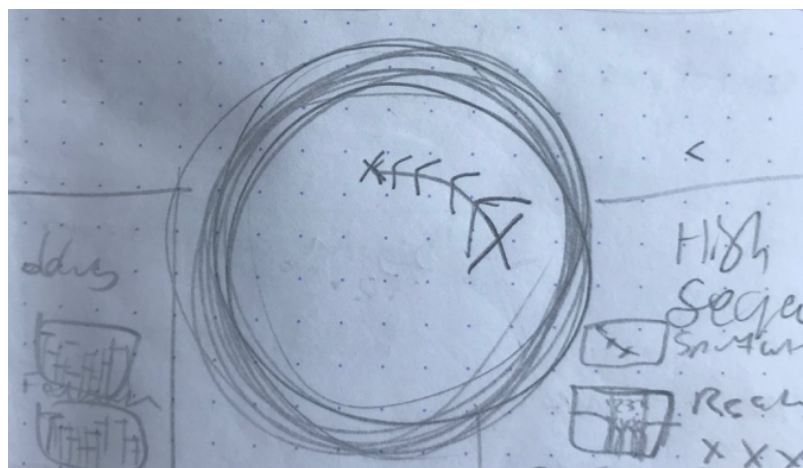


Figure 3.3 Initial sketch of Spectral-Spatial Smearing

*** Elastic Spatial Ripple Filter.**

Send ripples of resonant filters around the sphere filtering the sound.

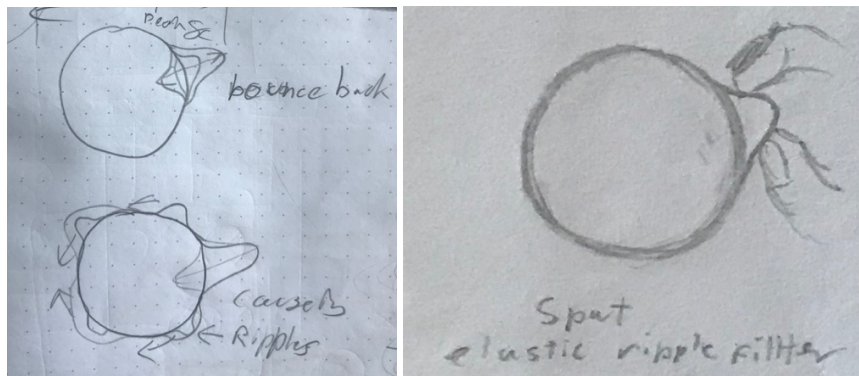


Figure 3.4, Figure 3.5 Initial sketch of Elastic Spatial Ripple Filter

Spatial Shepard Tone Texture Granulator. (Shepard, 1964)

A texture created by granulation and spectral processing that constantly moves through the sphere creating an endless sense of moving through space.

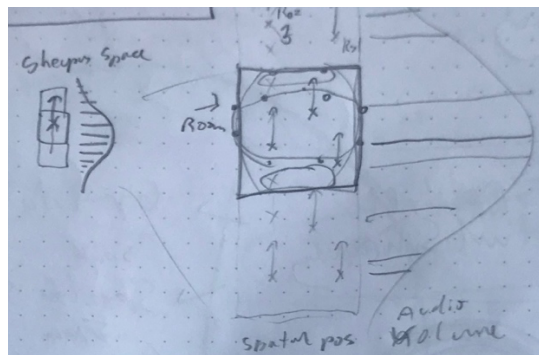


Figure 3.6 Initial sketch of Spatial Shepard Tone Texture Granulator

Spatial Path Reverb.

Create a reverb tail along a desired trajectory of the sphere.

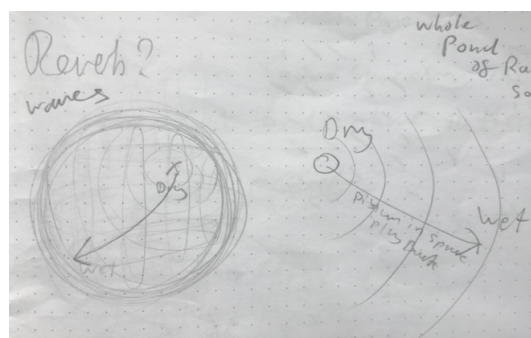


Figure 3.7 Initial sketch of Spatial Path Reverb.

As a “strongly visual-spatial” thinker (Silverman, 2005), I have chosen to use visual metaphors to inspire the sonic processes. Visual metaphors also make it simpler to convey the spatial process when verbally describing and when representing graphically in the user interface. There are, however, disadvantages to this approach. Primarily, overestimating the audio spatial resolution by thinking sound can match our visual perceptual capacities. Humans are better equipped to perceive motion within particle systems in the visual domain than the sonic domain. Chion (2016) explains the eye’s ability to perceive trajectories in greater detail than we can hear. The visual approach did prove problematic as it is easy to exceed the audio spatial resolution. To minimise this, the tools were adapted as the project progressed. (read more in section 5.2).

I knew there would not be enough time to develop all of my ideas and that the devices may undergo drastic changes as the research carried forward. I intended to create around four or five of the effects from the list, plus the accompanying utility effects needed to make them function; a basic panner for testing, an ambisonic decoder for use with S.P.I.R.A.L and a binaural decoder to enable working from home. I began by creating the utility effects to help work out any unforeseen technical issues there may be and to enable the development of the other effects. I then began translating my sketches into basic DSP block flow diagrams and GUI (graphical user interface) wireframes. I started to prototype the ***Spatial Boid Swarm Granulator*** first, as it was the most technical and similar to my previous work with game engines and particle systems. I also knew some of the other effects could be adapted from the same engine and would be a solid foundation to begin my research.

3.2 ITERATIVE DESIGN PROCESS

In order to conduct rigorous research, I used an iterative design process affording the flexibility to trial different ideas. I originally intended to inform my design process by simultaneously composing with the devices. This proved difficult in practice due to several factors, time constraints, maintaining multiple versions and composing with buggy tools, all made it impractical to compose at an early stage. Later in the project, I began creating musical studies with the devices. To maintain a creative outlook earlier in the development, I committed to musical testing sessions once a fortnight that steered the development focus of the tools. I would go through all the tools trying everything possible and methodically question what works well, what could be improved to gain better spatial clarity, and how can the devices better meet my research aims. This formed a cyclical design process.

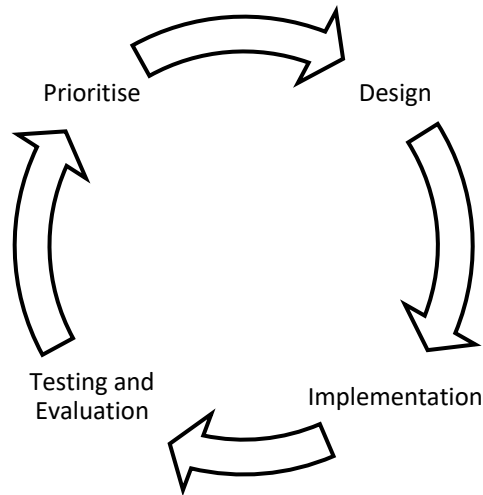


Figure 3.8 diagram of the iterative design process.

In order to maintain an iterative design process, the programming had to be neat and well abstracted, allowing testing of different variations and combinations of DSP and graphical components. Apel, Batory, Kästner and Saake (2013) describes this as the separation of concerns design principle (Goderis, 2008) (Jackson, Kang, 2010). This design methodology decreases the braided complexity of code, allowing for better reasoning about each code block's purpose and how they interact with each other. This makes the addition of features and bug fixing feasible without time-consuming refactoring. I have abstracted code based on reusability, limiting duplication, keeping code as modular as possible and have a strict separation of the GUI and DSP achieved by tagging user interface values with their destination parameter. Utilising Max 8's multichannel patching system proved essential for consolidating the signal routing between blocks. These concepts allow for better reasoning about individual GUI and DSP components helping to maintain an iterative design process that facilitates quick iteration and evaluation of different options.

3.3 USER FEEDBACK

I have also conducted qualitative user feedback to supplement the testing and evaluation stages of my iterative design process. I conducted an interactive presentation as part of the Huddersfield University CCL (Creative Coding Lab) where an early version of the tools was presented. The attendees suggested a few points of feedback mostly about user interface design. I also sent a copy of the devices out via email and posted a link on an online forum near the project's end, hoping to receive additional

Mortimer Pavlitski

feedback. However, there was minimal response possible due to the coinciding of the COVID 19 Virus. Informal conversations with users did provide some meaning full feedback about the implementation and alternate design ideas. (see section 5 for more details).

SECTION 4: LITERATURE REVIEW

The following section has been split into three parts in order to consider prior knowledge from multiple angles. Part one takes a compositional perspective to justify why there is a need for new tools, as many tools already exist for spatial audio but do not satisfy the creative aims of this project. The second part addresses a broader social relevance question answering what has prompted the research and why now is the right time to create new tools. The last part takes a technical approach to determine what technologies are best suited to create the tools.

4.1 A LACK OF CREATIVE SPATIALISATION TOOLS?

Many tools exist for adding spatialisation to musical compositions (360° Ambisonics Mixing Tools | Waves, 2019), (AmbiX, 2019), (SSA Plugins, 2019), (Audio 360, 2019), (O3A Manipulators, 2019) (Facebook 360, 2020), (SPAT, 2019), (GRM Tools Spaces, 2011), (Ambisonic Toolkit (ATK) for Reaper, 2015), (IEM Plug-in Suite, 2020). These spatialisation tools are mostly utility-focused: ambisonic decoders and encoders, panners, spatial imaging and realistic room reverb simulations; they are required for spatial audio and provide only the foundation for spatialised music. They do not give a composer creative spatial control over traditional musical effects such as delay, chorus and pitch shift. There are some exceptions, IEMs Delay has an ambisonic rotation inside the feedback path of a delay. GRM Tools Spaces also includes interesting spatial effects; however, they are limited to random panning within a speaker layout and cannot be chained. There are also a few tools that can apply the same effect to multiple channels, like the mc object set in Max, but they cannot affect each audio channel uniquely with respect to its localisation in space. These tools present a gap in technology for creating sonically interesting spatial scenes with integrated effects like pitch shift, chorus and spectral processing.

Spatial tools need to be integrated into a composer's DAW to eliminate the need for routing large numbers of channels between applications. Routing audio between applications for spatialisation is not suitable for a composer's workflow as it can introduce glitches, takes time to set up and is restricted by arbitrary channel limits that may not be enough for many spatial audio workflows. A composer's flow state would also be weakened by having to work between multiple programs simultaneously. An external application would also have limited integration with the DAW, missing out on parameter

mapping, recording and device chaining. Reaper has full support for ambisonic VSTs however other DAWs like Logic, Pro Tools, FL Studio, Cubase and Ableton Live struggle to offer much beyond low spatial resolution 2D panning or just stereo in the case of Ableton Live. While working with Multichannel plugins is possible in Ableton Live, it is not practical as it requires the use of many tracks and the manual routing of pairs of audio signals. Max for Live presents a way of automating the audio routing process in Ableton Live using the Live Object Model.

Several Max object suites exist for working with spatial audio including, SPAT5 (2019), Hoalibrary (2019), ICST (2019) and Cosm (2008). These objects allow for the creation of more artistic tools but require a composer to code their own effects, which is beyond the scope of most non-coding composers. Composers need fully developed DSP effects with a user interface; Max objects alone are not sufficient.

Manually programming the panning information of a hundred sound sources poses a large time commitment for a composer. Panning software like SPAT (2019) and GRM Tools Spaces (2011) offer the ability to precisely control the position of a sound in space and can create complex spatial scenes. However, recording or hand drawing automation for many individual sounds can be repetitive and time-consuming. This time commitment becomes a more significant problem every time the composer wants to change the sound's overall movement as each pan position needs reworking. This commitment could influence a composer's creative decision making if there are time constraints on their project. I believe there is a need for a more rapid solution for panning groups of sound sources to afford a composer more time to trial different variations of complex spatialisation.

Developments in technology enable more rapid interaction with spatialisation. Normandeau provides an example "In the analogue days, there was no way to record the movement of a fader or rotary button. The recording of these gestures made by the composer was then a major change in the way they figured their relationship with sound material." (2009, pp.278). The computer's ability to record panning information performed by a composer changed the relationship composers have with spatialisation by allowing more expression to be captured and saving composers time. I propose to take this a step further and allow the computer to pan sounds intelligently on their own. More

computationally powerful computers are capable of simulating some of the beautiful moment in the natural world. Reducing the level of control over individual pan positions enables a more rapid interaction with spatialisation and affords control over a group of pan positions, giving better reasoning about a soundscape as a whole.

Nature provides a model for automating the movement of panned sound sources. Flocks of starlings display a fascinating emergent behaviour known as flocking. Reynolds (2019) presents a way to simulate these patterns of movement digitally with his boids algorithm, which is based on only three simple rules (separation, alignment and cohesion). Shiffman (2012) implements the algorithm to allow other forces to be integrated with the boids system like flow fields simulations imitating weather systems. This develops on the random panning and jitter of GRM Tools Spaces creating a more natural humanised moment of sound. In the natural world animals of prey chase after groups of other animals where one animal's sound triggers the movement and sounds of another; creating long chains and recursive loops where every sound is spatially connected. Payne and Roger (1962) explain how birds of prey use sound to locate and hunt causing other birds to flee creating the chasing behaviour. It is these organic behaviours in nature that create a rich soundscape of complex sonic and spatial interaction that can be simulated to fulfil some of the research aims of the project.

Envelop (2019) provides a suite of Max for Live devices that present interesting musical potential, including one that implements Reynolds boids algorithm (2019). However, many of the effects are not suited to the majority of composers as the effects are either creatively oversimplified or have an overly complicated user interface. E4L Boids Delay allows each tap out of the delay to be panned to a different position using Reynolds boids algorithm. However, this device falls short for me as a creative in three major ways. Firstly, the boids section of the interface contains 15 dials with shortened technical names that are hard to navigate effectively when trying to compose. Secondly, the dials are all the same size and offer no prioritisation or categorisation to help users understand what effect a parameter will have on the spatialisation. Lastly, the top-down graphical display does not offer much of a visual representation of the space, due to the display taking up less than half of the 169-pixel limit for a Max for Live device.

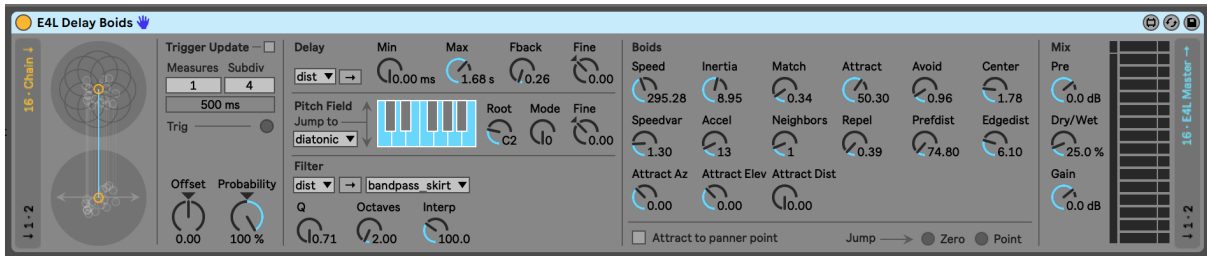


Figure 4.1 E4L Delay Boids Interface

I want to reimplement the idea of using Reynolds boids algorithm to pan sound sources with a more intuitive UI design. Canziba (2018) suggests several design fundamentals and principals to follow when designing a user interface. Bjørn (2017) suggest some novel control ideas that may be better suited for spatialisation parameters. Flowing these design fundamentals and principals along with the Max for Live Production Guidelines (Cycling '74, 2020) will enable my research aim of having a clear and intuitive UI design.

4.2 WHY CREATE NEW TOOLS NOW?

New tools would enable composers to keep up with the rising demand for spatial audio. The need for more spatial audio has been triggered by a growing trend towards more immersive experiences. Spatial audio is a key feature for enhancing a listener’s immersion within a given context. “Immersive sound can give the listener an experience of being there through sound. Compared to vision, sound provides a fully immersive experience and can be perceived from all directions simultaneously.” (Roginska, Geluso, 2017, pp.1). The COVID 19 virus has globally increased dependence on online resources with many industries reinventing themselves in a digital space; spatial audio provides a way to enrich our digital connections with each other and our entertainment.

Composers could use the tools to create content for many of the new spatial audio platforms. Many affordable commercial VR (virtual reality) headsets have come to market in the last five years increasing the popularity of binaural audio at home and in VR experience centres (Schutze & Irwin-Schütze, 2018) (Wikipedia, 2020). Several cinemas have opened in the UK supporting Dolby Atmos. Triple-A title games are also beginning to support these higher-resolution spatial formats like Shadow of the Tomb Raider supporting Dolby Atmos. Streaming services like Netflix, Vudu, iTunes and Amazon's Prime Video also support Dolby Atmos. Mainstream social media platforms YouTube and Facebook both support 360°

video streaming with 1st and 3rd order ambisonics, respectively. New headphones like the Apple Ear pod pros have built-in head tracking that enables better localisation for binaural audio, a component missing on traditional headphones. These readily available technologies give the general public many new affordable ways to experience spatial audio.

4.3 WHAT TECHNOLOGIES ARE BEST SUITED FOR NEW TOOLS?

To create the tools, a variety of spatial technologies could be used. This section will explain the reasoning behind the initial choices of technology.

LOCALISATION

To understand what technologies are best suited for creating new spatial tools, a basic knowledge of how we localise sound must be established. Our sense of aural localisation denotes the direction, distance and movement of a sound source (Blauert, 1997). Howard and Angus (2013) explain how the direction of a sound source is perceived by the volume of a sound at each ear known as ILD (Interaural Level Differences), the sounds arrival time at each ear known as ITD (Interaural Time Differences) and filtering caused by the physical anatomy of the listener's shoulders, head and ears known as a HRTF (Head Related Transfer Function) (Everest, Pohlmann, 2015) (Rumsey, 2001). Roginska and Geluso (2018) describe how perceived distance is conveyed by, the volume, high-frequency damping, R/D ratio (the ratio of reflected to direct sound) and the ITDG (Initial Time Delay Gap). Lastly, to deduce movement, we can consider the Doppler effect and Parallax effect. All of these acoustic phenomena can be simulated with panning algorithms recreating our sense of localisation digitally.

PANNING

Loudspeakers can recreate our sense of localisation through the panning of a phantom image. Two common ways of achieving this phantom image are Intensity panning and HAAS effect panning (Gardner, 1967) (Izhaki, 2018). Intensity panning simulates ILD by adjusting the left and right channel volumes with an equal power curve. Whereas HAAS effect panning also known as the precedence effect, uses delay-lines to simulate the real world ITDs that occur as a result of spatial localisation. HAAS effect panning is a less common technique due to the technical difficulty and undesirable introduction of destructive phase interference that causes comb filtering in a mono downmix. More advanced

panning algorithms take into account some of the other acoustic phenomena. CircularDoppler (2020) simulates the doppler effect giving a listener better judgment of movement. SPAT (2019) and GRM Tools Spaces (2011) also implement room modelling to addresses the direct / reflection ratio and ITDG as well as the directionality. Room modelling can be achieved using IR convolution reverb or a network of delay-lines and all-pass filters.

Having more than a stereo pair of speakers in the horizontal plane will widen the range of movement a sound can be panned around a listener. Positioning speakers behind a listener makes use of their HRTF and can be used to simulate a sound from behind or to the far side. Adding more speakers in the vertical plane unlocks the potential to pan a phantom image in 3D. With enough speakers, it is possible to localise sound anywhere around us. Humans have a limited sense of localisation in the vertical plane, as we cannot rely on ILD and ITD to detect the height of a sound, only the filtering caused by our physical anatomy (Blauert, 1997, pp.44) (Roginska, Geluso, 2018, pp.223). Vertical localisation accuracy can be compromised due to a psychoacoustic phenomenon that causes low frequencies to be perceived lower down in space and high frequencies to be perceived higher in space (Roginskaa, Geluso, 2018, Chapter 12) (Kendall, Cabrera 2011).

The accuracy of a panned image can be significantly increased by introducing more speakers. Izhaki (pp.191, 2018) remarks that the accuracy of a phantom image is greatly improved by a third centre speaker when localising a central point source. Roginska and Geluso (2018) explain the progression of different speaker setups in the horizontal plane that optimises localisation and immersion. By increasing the density of speakers, the perceived phantom image gains better localisation and an improved ability to create an immersive sound field; Roginska and Geluso also argue immersion to be a more important factor in perceived sound quality for untrained listeners (2018, pp.215).

OBJECT-BASED PANNING

Object-based panning is a widely used technique for keeping track of panning information. The format is based on the recording of individual sound objects and their panning information. Each audio track has an accompanying track of pan position information that specific playback systems use to pan the tracks in real-time. Object-based panning has gained popularity in cinema and home theatre

environments with the use of Dolby ATMOS (2020) and DTS X (2020) systems. Although object-based panning can pan a phantom image across a 3D space, there are many inherent disadvantages. Object-based panning does not scale up to allow control over hundreds of sound sources as different formats limit the number of objects that can be panned to around 128 channels in the case of Dolby ATMOS. It also lacks the portability of other formats usually requiring a remapping of the speaker positions for specific playback systems, which restricts the distribution of the music created. Lastly, spatial effects are not possible as the panning is always last in the DSP chain, limiting the creative control a composer can have over a moving sound. These concerns make object-based panning unsuitable for my research.

SOUND FIELD (AMBISONIC FORMAT)

Ambisonic audio takes a different approach to keeping track of panning information where the panning is encoded into the audio. The encoded audio then has to be decoded for a given playback system. The encoded audio has greatly improved portability, as the use of an encoder and decoder decouples the spatial audio from the speaker positions, eliminating the need for remapping at different locations. Ambisonic audio also has the advantage of scaling well to larger room sizes (Malham, 2011).

Ambisonic audio comes in different orders. First-order, known as B-format, can be conceptualised as an extension of the mid/side format. B-Format makes use of four channels in total, W (omnidirectional or mid component), X (left/right or side component) plus an additional two channels Y (up/down component) and Z (forward/back component). B-format is well documented and standardised; however, disadvantaged by blurry localisation due to the low channel count (Roginska, Geluso, 2018, pp.286).

Increasing the ambisonic order increases the localisation accuracy at the cost of more channels. Second, Third and HOA (Higher Order Ambisonics) encode additional spatial information using more channels (see fig 4.2). Adding channels increases the complexity with diminishing returns due to the exponential channel growth, where the number of channels for the encoded audio is equal to $(\text{Order} + 1)^2$. However, HOA presents an optimised format with the minimum number of channels required for the highest spatial resolution at each order using spherical harmonics (Roginska, Geluso, 2018, pp.286).

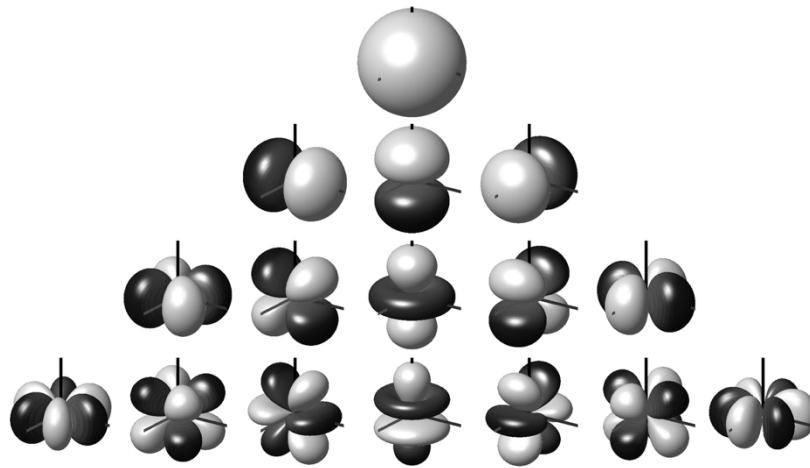


Figure 4.2 Spherical harmonics up to 3rd order. (Zotter, Frank, 2019, pp.68)

There are two main standards for ambisonic format audio with different channel ordering and normalisation. The first, proposed by Furse and Malham (Malham, 2003) known as FuMa, supports up to third order and is the most widespread format. The second, AmbiX, is a newer format supporting any order using SN3D normalisation and ambisonic channel numbering (ACN) (Nachbar, Zotter, Deleflie, Sontacchi, 2016) (Hodges, 2018). The AmbiX format is also supported by both YouTube 360 videos (Google, 2020) and Facebook 360 videos (Facebook 360, 2020).

The AmbiX format is suitable for my research because it has the advantage of higher spatial resolution. 5th HOA affords greater portability while retaining a high spatial resolution and enabling chaining of effects, making it a suitable alternative to object-based panning for my research. I will use 5th HOA to create my tools to utilise the S.P.I.R.A.L studio's capability of reproducing the upper half of a 5th order signal. S.P.I.R.A.L has a non-equidistant speaker layout and will require volume and time delay adjustments to accurately reproduce an ambisonic signal. A higher spatial resolution could be achieved using 6th order or more HOA, but the higher channel count and speaker requirements make it unpractical for this creative based research.

BINAURAL

The binaural listening format provides a way to listen to spatial audio outside of the S.P.I.R.A.L studio. The format enables a highly portable way to listen to a full spatial audio signal using only a pair of

headphones. Headphones also eliminate crosstalk interference inherent with speaker systems as the sound travels directly to the listener's ears reducing localisation blur (Blauert, 1997) (Izhaki, 2018, pp.191). The binaural format uses convolution to simulate HRTF acoustic localisation phenomena with an impulse response sample known as a HRIR (Head-Related Impulse Response) (Roginska, Geluso, 2018, pp.88). Encoded ambisonic audio can also be decoded to a binaural format to allow further portability of an ambisonic sound field.

While binaural audio offers detailed spatialisation with many benefits through a regular pair of headphones, there are some drawbacks. To be able to move your head within the sound field additional head tracking and real-time decoding is needed, this can be CPU intensive for a high-quality convolution decoder and require additional hardware. The HRIR used for encoding a binaural signal is specific to every person, so a generic IR must be used as a compromise lowering the spatial resolution. (IRCAM, 2002). Another factor that affects the spatial resolution of binaural audio is the frequency response of the playback headphones, as they are generally less transparent than a high-quality pair of studio speakers. For these reasons, I will not be relying on the binaural format alone and instead use it as a complementary listening format for working outside the studio.

SECTION 5: PROJECT DEVELOPMENT

This section describes the development of the tools and how important decisions were made. This section has been split into three parts, the technical implementation, creative decision-making, and user interface design.

5.1 TECHNICAL CHALLENGES

WHY MAX FOR LIVE?

In order to meet my integration aims, I had to decide what format was best suited to the creation of my tools. A standalone application would not integrate well with the Ableton Live environment, as it would require a third-party audio bridge and an OSC or MIDI bridge to enable recording and mapping of parameters. After ruling out a standalone application, two options remained Max for Live and VST3 or a similar plugin technology. Plugins have the advantage of being portable across DAWs. However, Ableton Live cannot chain multichannel plugins together without repetitive routing of stereo signals to return tracks. To connect a multichannel plugin, pairs of outputs and inputs must be routed via individual tracks. For a 36 channel ambisonic signal 18 tracks are needed per connection. The time required to route all of the tracks does not satisfy the rapid creation aspect of my aims, leaving Max for Live as the only option. Using Max for Live's LOM (Live Object Model), I created a multichannel routing bridge that automates the routing connection process, enabling device chaining. Max for Live also has the advantage of recently introduced multichannel support. Multichannel patch cables and a set of multichannel objects streamline the creation of multichannel effects in Max for Live. These tools afford better code encapsulation and modularity in line with my methodology. For these reasons, I have chosen to create a suite of Max for Live devices instead of a standalone application, VSTs or similar plugin technologies.

AUDIO BRIDGE

The audio bridge was essential to enable the routing of ambisonic signals between any two devices in Ableton Live to fulfil the project's aims. Creating the ambisonic audio bridge was convoluted due to the way the LOM exposes the device routing options. There are several obstacles to making a connection as devices cannot query their own identification number needed to connect to their input. Confusion is also caused by the naming scheme of the LOM where an object id is not related to an identification

number for a device IO's available routing options. However, it is possible to get the identification number by requesting information from other devices and comparing all the names and paths on a track with the list of available routings. It became evident it was going to take me some time to understand how to code a fully automated bridge with minimal bugs. To continue with the rest of the research, I coded a simple bridge that would only work under certain conditions. I then came back to this problem near the end of the project after I had done more testing to remove some of the bugs and make the connections more stable when a user is adding, deleting and moving the devices around.

ENCODER AND DECODER

The next decision to be made was what decoder and encoder to use for working with ambisonic signals. There were two viable Max object packages, HOA Library, ICST or the option of hosting a plugin using the VST~ object. Committing to an encoder and decoder at the start of the project would limit options and flexibility, so the encoder and decoder stages were modularised within the code. By abstracting the encoder and decoder into a modular block, it was possible to trial different encoders and decoders throughout the project. I settled on the HOA Library (2019) for the binaural decoder as it includes an adjustable HRIR length to give users the option to trade some of the localisation accurately for better CPU usage. The binaural decoder was vital when working from home at the end of the project because of COVID 19. For the S.P.I.R.A.L decoder, I decided to use the HOA Library (2019) decoder as it can compensate for the non-equidistance speaker arrangement in the studio and had better localisation than the IEM decoder.

HOSTING VSTs

As movement is one of the key themes I wanted to explore, I had to work out how to rotate the ambisonic sound field as many of my initial ideas required rotation. Zotter (2019) provides a rotation matrix for rotating in the Z-axis up to 3rd order. This method can be extended for rotation in the Y and X axes using successive static 90-degree rotations. However, Zotter's method can cause specific rotations to be unobtainable in a situation known as gimbal lock. Perumal (2014) suggests using quaternion rotation that makes it possible to rotate the ambisonic sphere freely. This is achieved using 4-dimensional complex numbers (quaternions) that can rotate around the 3D situations where gimbal lock would occur (Mahé, Ragot, Marchand, 2019). I decided to use the IEM Scene Rotator that enables up to 7th order rotations. The complex nature of high order ambisonic rotation is beyond my research

scope, which is why I justify having a 3rd party plugin as a dependence; it abstracts out a highly technical problem from my creative based research (Choi, 1999). IEM's plugins also have the advantage of being lightweight and open source, making them more suitable than other plugins for the project. I also used a few externals for various number conversions and embedded a small snippet of code for doing fast onset detection from SubtleSonic EnvelopeShaper 1.0 (McCulloch, 2012) (see Appendix A).

5.2 CREATIVE DECISION MAKING

OPTIMISING SPATIAL CLARITY

Once the initial prototypes had been created, I realised I had overestimated the number of discrete sound sources I could localise. My initial intention for the granulator was to have 300 boids representing the pan position of 300 grains. However, encoding 300 signals from the granulator was too CPU intensive to be practical, and there were diminishing returns on the number of discrete sound sources I could perceive. I decided to reduce the number of boids to 30 and allow multiple grains to be played at the same location. This provides a more continuous sound at each location, making it easier to localise while still having enough positions to create a diffuse sound field.

Similarly, I had intended the delay device to have 64 tap outs arranged in different paths. This was too many taps for three reasons. Firstly, any per tap effects would be duplicated 64 times and be too CPU intensive to use. Secondly, due to the close angles of the panned taps, the perceived localisation would be compromised, creating one larger and more diffuse image (Kendall, Cabrera 2011). Lastly, adding controls for 64 taps presented a significant challenge for designing a meaningful user interface. For these reasons, I decided to limit the number of taps to 16. A stereo input with eight taps for each side integrates better with Ableton Live's stereo tracks and simplifies the user interface, affording more precise control over the tap outs.

To improve the localisation of the devices, I tested different effects. The addition of a waveshaper after the granulator filter produces additional high-frequency content enabling more accurate location (Blauert, 1997) (Kendall, Cabrera 2011). Adding a pitch-shift effect to each tap out of the delay helps to decorrelate the delayed sound and localise the taps at discrete positions in space. Artistic judgment

was also utilised when choosing effects like the reverb on each tap out of the delay and spectral gate added to the *Spectral-Spatial Smearing*.

RELATIONSHIP OF ARTISTIC INTENTION

After prototyping the original ideas for the *Spatial Path Trajectory Delay*, I decided to remove the ambisonic feedback loop and develop it into a separate device. The two stages of the delay were doing fundamentally different things, panning tap outs of a stereo delay and creating an ambisonic feedback tail. By splitting the device up, the underlying DSP engine could be presented behind an easier to navigate user interface. This decision was supported by user feedback when presented at the CCL interactive presentation.

Once the ambisonic delay had been developed into a device, it could then be chained with other effects using its ambisonic input, creating many more sonic possibilities. Some of the ideas for the *Envelope Follower/MIDI Trigger Stochastic Panning* were incorporated into the ambisonic delay device resulting in the variation section. However, splitting the delay increased the development time and in order to focus the research and trial modified versions of the other prototype devices, I decided not to create the *Elastic Spatial Ripple Filter*.

I decided to limit the modulation options (LFO's, envelopes, step-sequencers) added to the devices. I made this decision as I did not want to complicate the devices that are already quite technical to understand. I also wanted to encourage the use of the wide selection of previously created modulation device include with Ableton live and created by the Max for Live community. Many devices can be downloaded and mapped using Ableton's mapping system. This integrates my devices within the Max for Live environment further, and it encourages the linking of other devices in new ways allowing for more sonic possibilities.

My original methodology mentions using the *Spatial Boid Swarm Granulator* as a starting point for other devices. However, instead of repurposing the granular engine, I decided to incorporate some of the other ideas into the same device. Using Shiffman's (2012) model for applying forces, I explored a range of spatial behaviours in combination with the Reynolds boids algorithm. I added an attractor to the boids system that could be chased or avoided giving an impactful macro control over the movement of the boids; which was crucial for satisfying one of my aims. The more philosophical inspiration I had

about life, decay and evolution was incorporated like a particle system. Boids have a life span and decay amount and are reborn at a set location in space. These two behaviours allow for the setup of many complex interactions and trajectories, encapsulating some of my ideas for the *Spatial Shepard Tone Texture Granulator*. I also experimented with a flow-field based wind system for the boids to follow. However, due to the complexity of generating 3D and 4D flow-fields, it became apparent that the user interface would become very complicated. I did not want to sacrifice the simplicity of the user interface in the same way E4L Boids Delay did and decided to remove the flow-field to keep in line with my user interface design aims.

5.3 PARAMETERS, UX & UI DESIGN

The user interface plays an essential role in meeting three of my research aims: macro-control minimising the time required to produce different versions, clear and intuitive UI design to give feedback on how parameters affect the sound and integrate fully into the Ableton Live environment. After reading the Max for Live Production Guidelines (Cycling '74, 2020) I came up with a series of questions that directed the design process of the user interfaces.

WHAT PARAMETERS SHOULD BE EXPOSED TO THE USER?

I wanted the composer to have as many controls as possible to shape the sound and not limit the creative potential of the suite. However, there would not be space to expose every control and would cause too much confusion, so a few exceptions were made in unique circumstances where exposing additional controls would have been detrimental to the user experience. For example, the granulator decay dial increases the chance for a boid to enter a dead state. Then a random amount of time after it will respawn back. The specific chance weightings are all fixed as they have a minimal effect on the overall sound. The result encapsulates life, decay, and death all behind one easy to understand dial; this reduces the tab's complexity, which already has many controls.

HOW SHOULD THEY BE EXPOSED TO THE USER?

I didn't want the interface to fall into the trap of having all sliders or all dials, as these controls may not be best suited for every parameter, especially where multiple parameters are related. Bjørn (2017) showcases a wealth of user interface options which have been carefully considered for every

parameter. The **Spatial Path Trajectory Delay** originally featured a per tap out digital pin matrix in the style of the VCS3 for rewiring the order of the effects (figure 5.1). However, as I settled on what effects would be in the final version, a sensible fixed order developed, and it seemed unnecessary to give the user control over it. The final effects section replaces the pin matrix with a small bypass toggle consistent across all devices (figure 5.2).

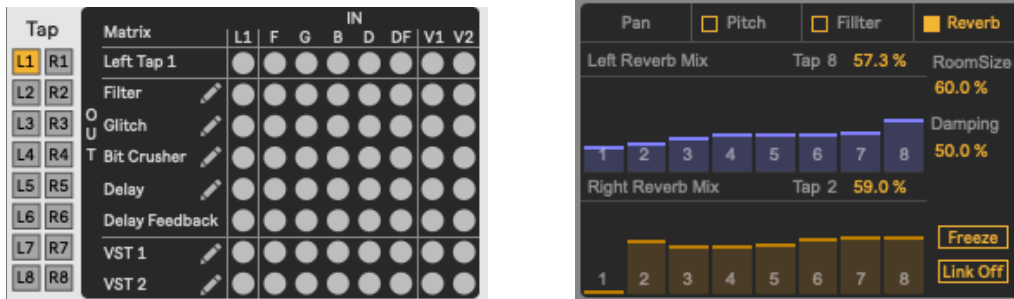


Figure 5.1, Figure 5.2, The stereo tap delay effects section.

Left, the early pin matrix design. Right, the final design with bypass toggles.

I designed custom user interfaces with the JSUI object to create a more intuitive user experience. After studying Logic X’s Delay designer plugin, I was considering switching the **Spatial Path Trajectory Delays** delay time controller from a multislider (figure 5.3) to a drag window (figure 5.4). Both ideas were presented at the CCL interactive presentation, where attendees unanimously agreed the drag window was much easier to understand. Listening to the feedback, I switched the interface to the drag window. The drag window has the disadvantage of taking up more horizontal space, to compensate, related parameters have been placed under and above. The per tap out delay time values were also condensed into one display that changes to the current tap out when clicked or changed to save space and minimise clutter.

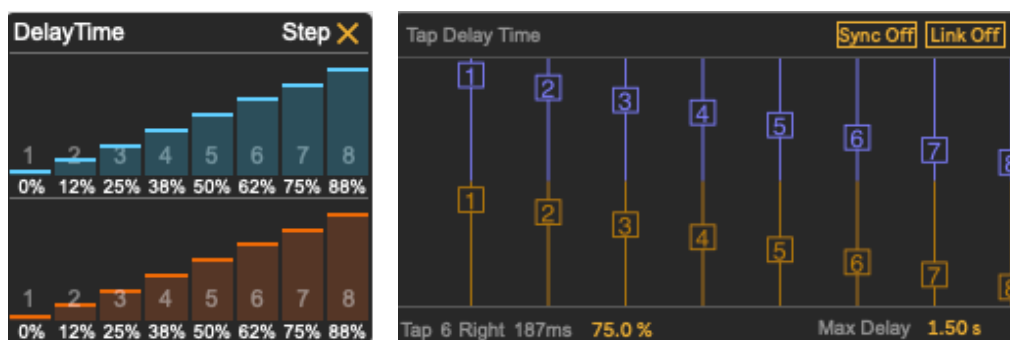


Figure 5.3, Figure 5.4, The stereo tap delay’s delay time.

Left, the early multislider. Right, the final design with bypass toggles.

Several other custom user interfaces were also created, including one to control the *Spatial Path Trajectory Delays* tap out pan positions (figure 5.5). The pic slider object included with Max has only one control point, and the nodes object does complex interpolation causing CPU lag when all the pan position parameters are being automated, making both unsuitable for controlling multiple pan positions.

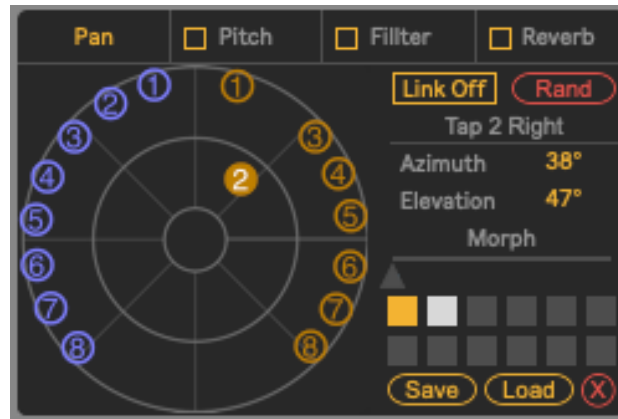


Figure 5.5, The stereo tap delay's pan positions.

In some cases, alternate user interface options for the same controls are provided for increased flexibility and better integration with Ableton Live. For example, a Free/Sync toggle switches a dial between time in milliseconds and divisions of a bar. This gives the user more flexibility when using the tools. I also offered alternatives for panning between polar and cartesian coordinates to enable front/back and left/right mapping options and azimuth and elevation. The panner can be used as an XY pad which is necessary for mapping many external controllers enabling rapid control.

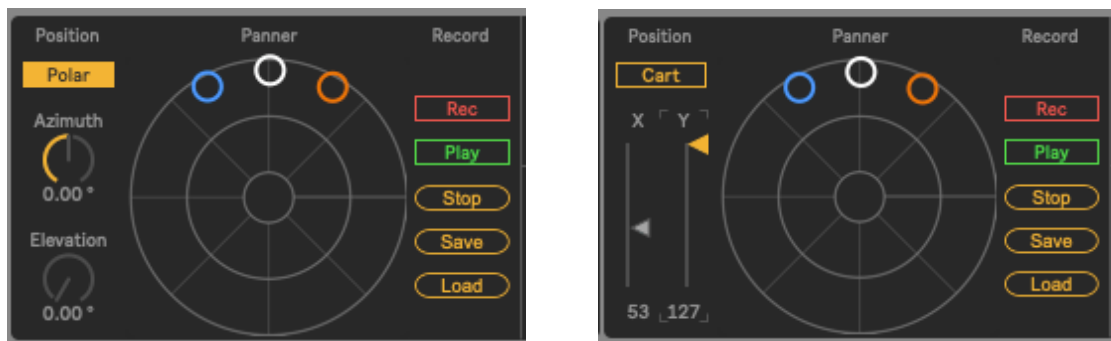


Figure 5.6, Figure 5.7, The stereo panner. Left, in polar mode. Right in cartesian mode.

ARE THE CONTROLS INTUITIVE TO NAVIGATE?

The interface design relies on aesthetics to make the devices intuitive. Canziba (2018) states seven design principals for good interface design, alignment, hierarchy, contrast, repetition, proximity, balance and space. In combination with the basic elements of visual design, lines, shapes, colours, font/typography, textures and form. I have focused on the size, colour, consistency, proximity, lines, layout and grouping to subtly guide the eye of the user. Small details like the consistency of toggle locations across different tabs reduce the amount of effort required by the user to navigate the controls. This can be seen in the stereo link toggle at the bottom left of the *Spatial Path Trajectory Delays effect* section (figure 5.8).



Figure 5.8, Each tab of the stereo tap delay's effect section.

I have followed the Max for Live guidelines to integrate the tools better with Ableton Live. The Max for Live guidelines recommends using the included Ableton Sans to maintain a consistent look and feel and integration with Ableton Live. I have also opted for a 9.5 font size as a minimum as this is used in various Max for Live packages as the default size. Although the Max for Live guidelines recommends using Lives native colour themes, I decided to create my own theme to impart additional meaning through the colour. Following the familiar Max for Live LCD style with a few additional colours helps set the suite apart from other devices. Separation is important as the devices have their own system for routing signals and do not function as a standard stereo device. Use of complementary colours blue and orange, help denote there grouping as left and right channel controls throughout the devices, respectively. The other colours have been used for their commonly known function, like red meaning record in most music software and hardware.



Figure 5.9 suite colour pallet.

Grouping plays an important role in telling the user how different parameters will affect the sound. Visually separating the devices' controls guides the user to their specific function and creates consistency across the suite. Dark grey panels indicate the relationship controls have to the underlying DSP blocks. These blocks also denote the underlying DSP signal flow of the device, starting from left and moving to the right. Most devices are split into four sections, input, the main DSP effect, and the additional effects and output (which can be seen in figure 5.10).

I have used a combination of Canziba's (2018) design principles to make controlling the left and right channels explicitly clear in the stereo tap delay (which can be seen in figure 5.10). Three things tell the user which side a control belongs to. First, it is written in the name of the parameter or by the units in the case of the delay time. Secondly, the colour, all the controls for the left side are in blue and all for the right are in orange. Lastly, the interface is designed with a horizontal split where the top half of the interface always controls the left and bottom half always controls the right. The vertical separation made the interface much harder to design and may appear like a small detail, but it makes it much clearer when navigating between the different sections of the devices. Controls that break the vertical separation and colour scheme then intuitively control both sides.



Figure 5.10, The stereo input tap delay interface.

HOW DOES THAT FIT IN WITHIN THE 169PX HEIGHT LIMIT?

Max for Live devices have a vertical restriction of 169 pixels. I started with a fixed grid system, but this restricted space and did not abide by the proximity and size design principals.



Figure 5.11, Figure 5.12. Left, an early sketch of the fixed grid interface.
 Right, a digital render of the fixed grid interface.

To fit all the controls in the limited space without the devices being unnecessarily wide, it became evident I would need additional window space. I decided on a tabbed view as pop-outs are not recommended by the Max for Live user guidelines. Using a tabbed view allowed more space for controls without sacrificing the intuition of the devices and helped to further group controls. However, strictly no sub-tabs are used only alternative views to eliminate menu diving, which adds to a device's complexity and takes a user time to navigate. I experimented with wireframe paper sketching to enable rapid prototyping of the layout as suggested by Canziba (2018, pp.56). Figure 5.13 shows an early sketch of the stereo tap delay (more interface sketches are provided in Appendix B).

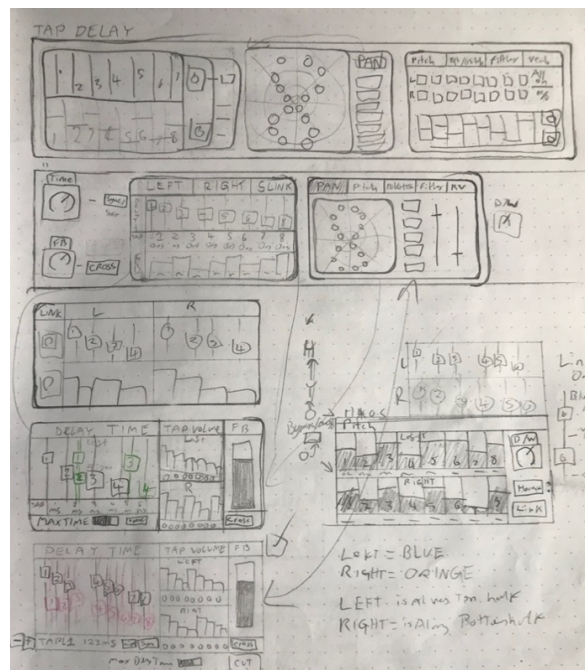


Figure 5.11, an early sketch of the stereo input tap delay interface.

SECTION 6: OUTCOMES

MTOOLS FOR LIVE (MT4L)

The outcome of the research is a suite of Max for Live devices capable of creating and processing 5th order ambisonic signals. The suite has been named *MTools for Live* and abbreviated to **MT4L**. The devices route 5th order ambisonic signals automatically to the next **MT4L** device on the track. If the device is last in the chain on a track, it will route the ambisonic signal to the currently active **MT4L Master**. The **MT4L Master** decodes the ambisonic signal for playback over headphones, in the case of **MT4L Master-Binaural**, and playback in S.P.I.R.A.L in the case of **MT4L Master-Spiral**. There must be only one instance of an **MT4L Master** loaded for the suite to function properly, as it also serves as a hub for controlling the routing of ambisonic signals. All **MT4L** devices can be added, deleted and moved at any time and will reconnect automatically.

*Note: The included ambisonic audio files can be played back using the **MT4L Player** device as Ableton Live has no practical way of playing back ambisonic audio files. This device is not part of the research and only serves as a convenient listening format.*

MT4L PAN

MT4L Pan brings stereo audio into the ambisonic environment and enables any track to be spatialised. The device takes a stereo input and encodes it to the ambisonic sphere. It also features pan position, rotation, stereo width adjustment and a record feature for the panner.

Ambisonic_Audio_Test_File.aif demonstrates the panning effect.

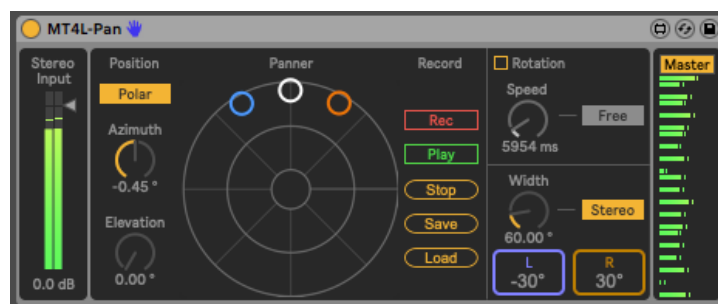


Figure 6.1, Screenshot of the MT4L Pan device.

The record system is ideal for capturing gestures in the moment, supplying composers with a hands-on way to interact with the spatialisation process. Pressing the record button will prime the device to start recording, once the mouse has been clicked down the recording will begin and is automatically stopped once the mouse is released. The gestures can also be saved and reloaded for use another time or with the **MT4L Smudge**. A more optimal solution may have also included a way to quantise the recording to sync playback with Ableton Live.

MT4L MASTER-BINAURAL

The **MT4L Master-Binaural** device will receive ambisonic audio from the last **MT4L** device on every track and decode the ambisonic audio for playback over headphones. The device also features quality adjustment that lowers the decoder quality for lower CPU usage, rotation and flipping of the ambisonic sphere. This device gives composers a way to work with their ambisonic compositions when away from the studio.

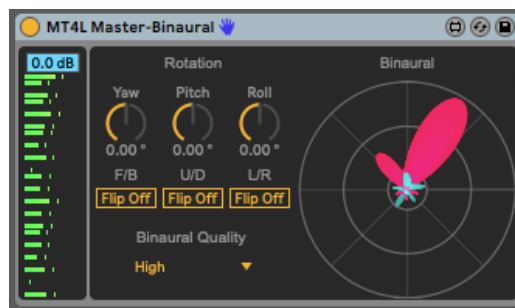


Figure 6.2, Screenshot of the *MT4L Master-Binaural* device.

MT4L MASTER-SPIRAL

The **MT4L Master-Spiral** device will receive ambisonic audio from the last **MT4L** device on every track and decode the ambisonic audio for playback in the S.P.I.R.A.L studio. The device also features optimisation modes, rotation and flipping of the ambisonic sphere. This device gives seamless integration into the S.P.I.R.A.L studio with no additional bridges or conversions necessary. Composers can work within the Ableton Live environment uninterrupted.

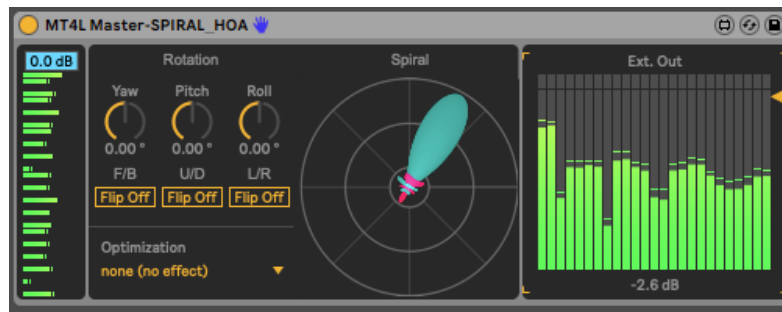


Figure 6.3, Screenshot of the MT4L Master-Spiral device.

MT4L GRANULATOR

The **MT4L Granulator** device granulates from a buffer that can be recorded into from a stereo input or loaded into from a file. Individual grains are panned to positions in the ambisonic field that correspond to the position of a boid in the simulation. The device also features a per grain effects section.

MT4L Granulator_Bots.aif demonstrates the Granulation effect.



Figure 6.4, Screenshot of the MT4L Granulator device.

The feature-rich granulator and combined effects section provides many sonic possibilities from classic granular effects to unique sounds. The effect section only supports additive randomisation. Including a small selection of random distributions would have unlocked more creative possibilities. The boid section exposes controls that change the overall movement of the panned grains. The boid behaviour, decay and attractor, controls provide many options to set the system as desired. When the decay parameter is in effect, the boids will randomly start respawning. The rebirth and attractor position then control the start and endpoint, allowing trajectories to be specified. The easy to navigate user interface saves the composer time, giving them more freedom to trial variations of complex sound fields. The device can create anything from swarms of sound flying around the ambisonic sphere to diffuse static drones.

MT4L TAPDELAY

The **MT4L TapDelay** device takes a stereo input and pans the tap outs of a delay line to different positions on the ambisonic sphere. The device also features a per tap effect section.

MT4L TapDelay_Immi.aif demonstrates the TapDelay effect.

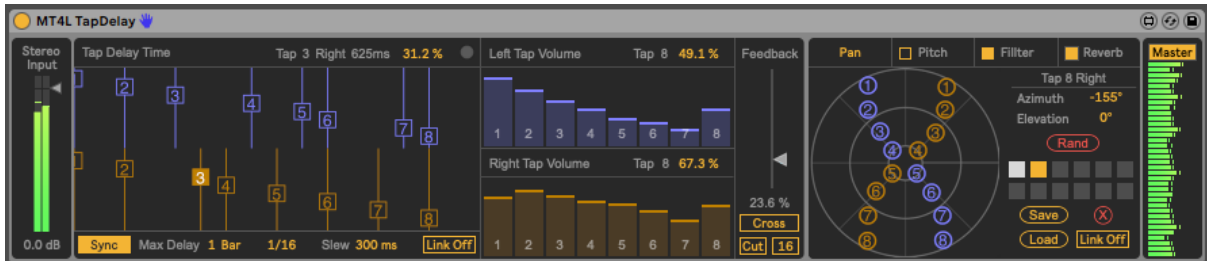


Figure 6.5, Screenshot of the MT4L TapDelay device.

As well as being a feature rich delay, the device can produce highly customised realistic and unnatural reverbs. Each tap out can be treated as an early reflection with customisable delay times, high-frequency damping via the filter effect and pitch shift variation. The feedback control can create late reflections, and the reverb effect can create a diffuse tail allowing for realistic and unnatural spatial reverb.

When the feedback is nearing 100%, the device becomes a looper, utilising ascetically pleasing tanh (hyperbolic tan) limiting saturation in the feedback path to keep the volume level. When enabled, the cut toggle will prevent the device's input from being recorded into the loop. To clear the loop, briefly reset the feedback to 0%. The underlying loop is unaffected when editing the delay times and effects of each tap out, so it is always possible to return to the original loop. This device configuration gives composers the freedom to develop musical ideas over time, exploring different controls to create naturally evolving and complex textures.

MT4L SMUDGE

The **MT4L Smudge** device takes a stereo input and pans the frequency bands of a filter bank to different positions on the ambisonic sphere using the Spatial Lag controls. The device also features a per band spectral gate and per band spectral delay.

MT4L Smudge_4thTemple.aif demonstrates the Smudge effect.

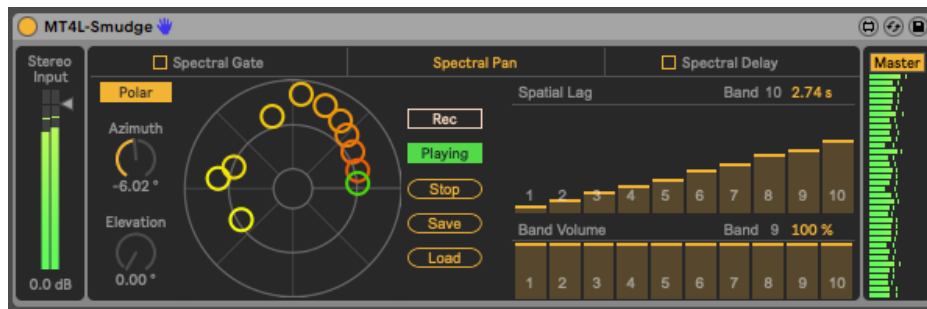


Figure 6.6, Screenshot of the MT4L Smudge device.

Gestural recordings saved from **MT4L Pan** and **MT4L Smudge** can be loaded into either device, giving composers a way to develop recordings made with **MT4L Pan** and downgrade **MT4L Smudge** gestural recordings to reduce CPU usage.

Spectral pan gives composers the ability to smear apart frequency bands of a sound across space. The effect has many creative uses, for example, anticipating the arrival of frequency content, creating decaying tails or even tearing a sound entirely apart. The spectral delay can tear apart frequency bands in the temporal domain, exposing a second dimension of sonic manipulation. The feedback and pitch shift sections of the spectral delay can build-up harmonic content, even to the point of chaos. Using the spectral gate gives the reverse effect, affording composers curation of the frequency content, revealing new musical ideas and transforming the original sound. When approaching the device to create and destroy frequencies, endless spatial sound design possibilities present themselves. However, there is no way to change the cross over frequencies for each band, which may narrow the device's creative potential.

MT4L AMBIDELAY

The **MT4L AmbiDelay** device is an ambisonic delay, taking an ambisonic input and producing an ambisonic output. The device also features a filter and rotator inside the feedback path of the ambisonic delay. Parameters with a random control (*denoted by a red percentage directly below the control*) can be randomly adjusted with the variation section. The variation section can be triggered by pressing the big red button manually or automatically with, onset detection, a MIDI note-on message, or synced to a beat of Ableton Live.

MT4L AmbiDelay_Twili.aif and MT4L AmbiDelay_Toms.aif demonstrate the AmbiDelay effect.

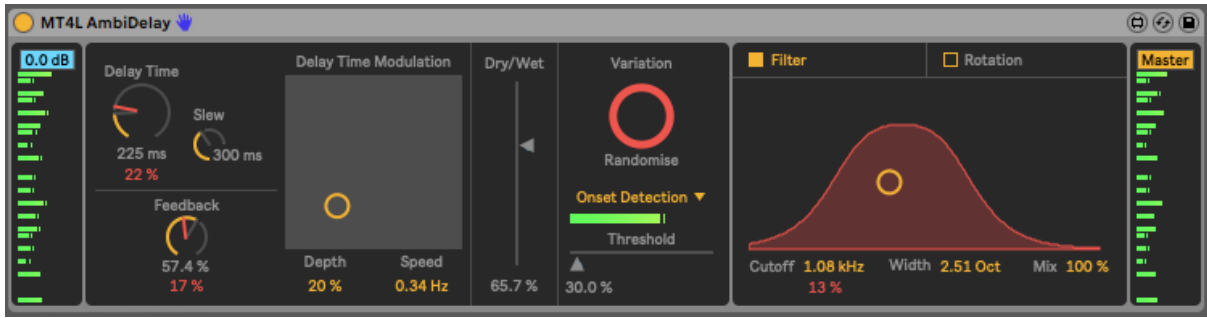


Figure 6.7, Screenshot of the MT4L AmbiDelay device.

The device takes an ambisonic input supporting all of the other **MT4L** devices. The **MT4L AmbiDelay** can add environmental context to the other effects mimicking the movement and variation of nature. The delay features dub style saturation and slew providing vintage sound characteristics. The integrated randomising system can automate the creation of organic moving trajectories, transforming a single impulse into an iterative evolving moving sound object, opening a spatial world similar to Second Woman's '//' (2017). Dials may not be best suited for the rotation amount interface. Decoupling the amount of rotation and direction may have provided more intuitive control of the spatialisation.

SECTION 7: CONCLUSION

7.1 CRITICAL DISCUSSION

ARTISTIC INTENTIONS

The created suite of devices ***MTools for Live (or MT4L)*** satisfies the main aims of this research. The suite presents new, sonically interesting spatial scenes using integrated ambisonic encoders and effects sections that were not available without custom programming. They collectively offer a new way of working with spatialisation that integrates well within the Ableton Live environment, enabling an uninterrupted flow of work. Specific control parameters, like the ***MT4L AmbiDelay*** trigger button and ***MT4L Granulator*** attractor position, provide macro control over the overall spatialisation enabling rapid variation of complex results. The ***MT4L Granulator*** also simulates some of the intricate spatial movement of the natural world and encapsulates some of the philosophical concepts about life, death and decay that I wanted to explore. Using Canziba's (2018) design principles, the devices are mostly intuitive to interact with and give meaningful feedback about how different parameters affect the sound. Although the complex nature of the underlying DSP is not always apparent, and I would have liked more user feedback to refine the interfaces.

The suite does fall short in some areas. The high CPU usage of the effects, caused by large numbers of ambisonic encoders, does impose a barrier to entry limiting which composers would be able to use this, although this can be improved by increasing Ableton Lives latency buffer size. While I consider the ***MT4L Granulator*** effect successful, I would have liked to further develop the spatial interactions with flow fields, machine learning, and genetic algorithms to explore other forces that could manipulate the spatialisation. I also believe the ***Elastic Spatial Ripple Filter's*** absence reduces the usefulness of the ability to chain effects, as the ***MT4L AmbiDelay*** is the only ambisonic input effect. The ***MT4L Smudge*** filter bank reconstruction is far from transparent and could be improved with convolution filters. I also intended to implement a custom distribution random number generator to enable more sophisticated randomness among the devices. The inclusion of delay line effects impacts the delay time, and no compensation for this has been applied, which slightly alters some of the delay times in the suite. Most of these problems could have been solved with more time or a more optimised development process.

DEVELOPMENT PROCESS

The development of the tools depended on a sensitive balance of time management skills, realistic ambition for the outcomes and good mental health. I underestimated the amount of time it would take to realise my goals and overestimated the condition of my mental health. All three were adjusted in order to finish the project to a standard I was happy with. The original timeline plan was revised multiple times as the technical implementation of the devices kept pushing back the creative and artistic exploration. In particular, the ambisonic audio bridge between devices took months rather than days to understand and implement.

The development process imbalance was compounded by an oversight in my methodology as my iterative design failed to consider how the project will conclude. The design cycle has no clear ending for the development of the tools and caused the addition of more features late in the project, without consideration for the time cost to implement and maintain them. This led to a reduced amount of time spent on musical output with the tools. While I think it would have been possible to create tools and music, I stand by my choice to focus the research on the tools, as they are the component that fundamentally satisfies the aims of this research.

COMPARISON WITH PRE-EXISTING WORKS

MT4L offers more creative ways of interacting with spatialisation than pre-existing ambisonic tools. The IEM plugin suite and ATK for Reaper feature utility-focused tools. **MT4L** develops on their technical capabilities with more artistic ideas about spatialisation, panning grains of a granulator with a boids simulation or panning different bands of a filter bank. IEM DualDelay features a rotator in the ambisonic feedback path of a delay, **MT4L** develops on this concept using the random variation system. **MT4L** also has an automated audio routing system, which enables more rapid routing of ambisonic signals than both IEM and ATK as they rely on Ableton Lives limited multichannel support. The E4L suite does offer some creative ideas about spatialisation and an automated routing system.

MT4L offers several advantages over the similarly scoped E4L device suite. Firstly, the **MT4L** suite works in a higher ambisonic order offering better spatial resolution in 5th rather than 3rd order, despite the higher resolution increasing CPU load. Secondly, the improvements of the interface design which are most evident when comparing E4L Delay Boids and the **MT4L Granulator**. The **MT4L Granulator** denotes

the importance of an interface element by its size. The most important parts of the interface are given the full height of the device. Figure 7.2 shows the full height display of **MT4L Granulator** instead of the half-height display of E4L Delay Boids (*which can be seen in figure 7.1*). The boids section of the **MT4L granulator** does various things to help the user navigate the controls that the E4L Delay Boid does not; related controls are grouped, controls are colour coded to the display, controls are labelled, controls have a clear description in the Info View section of Ableton Live, the interface uses lines to separate controls and the controls use a variation of interface elements not just dials. These differences show a development from the E4L suite to the **MT4L** suite.

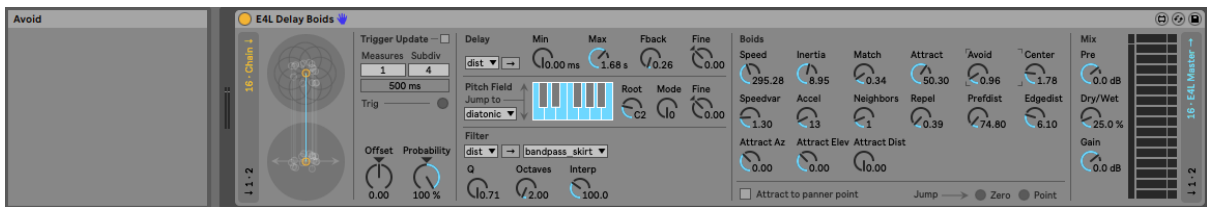


Figure 7.1 E4L Delay Boids interface

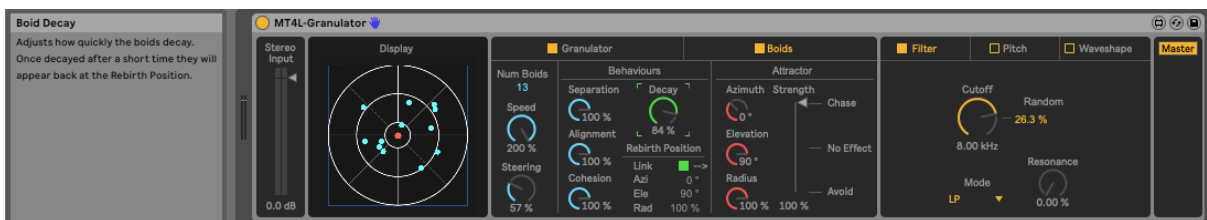


Figure 7.2 MT4L Granulator interface.

USER FEEDBACK

An oversight in the methodology caused limited user feedback response. Assuming people would be as enthusiastic as me to try the tools exposed a bias I had towards the research, leading to an underwhelming uptake in users of the suite. Perhaps a more proactive approach to finding users earlier in the project would have resulted in a larger user base, although the Covid 19 virus may have impacted the user uptake. In future work, I would aim to create a community of users earlier in the project, even while the tools were not functional. I would also create an online social presence; a forum could help create a dialogue between users where more feedback could be obtained.

Users of the software had an overwhelmingly positive response to the tools supporting their creation. Most of the feedback gained was through informal presentations of the tools. Users mostly commented on the interface and not the underlying technology, highlighting the importance of user interface design. If users had spent more in-depth time, these results may have been different. Many users continued to ask for the *Elastic Spatial Ripple Filter*. In reflection, my judgment could have been supplemented by users when considering not to create this device.

AUDIO EXAMPLES

AmbiDelay-Toms.aif starts to explore the **MT4L AmbiDelay's** ability to affect a sound over time and through space. The audio features a tom sound directly in front of the listener being delayed, filtered and panned away from the centre. Several of the tom hit delays are not panned enough to perceive movement due to the distribution of randomness used to generate the rotation amount. Switching from a uniform to a bimodal or inverted normal distribution randomness would solve this issue. *AmbiDelay-Twili.aif* develops on *AmbiDelay-Toms.aif* using more complex harmonic content, starting to present an exciting theme that could be explored further into a musical piece. Neither examples make use of the device's ability to take an ambisonic input and lack contextual grounding. An underlying sound texture could develop the piece and help position the sound objects within a sonic environment. Although leaving the arrangement sparse guides listeners to hear the spatialisation clearly before other elements are introduced. They both also develop on the work of *Second Woman* (2017) by recreating similar organically developing sound objects with the previously desired spatialisation element. (read section 2 for more details).

TapDelay-Immi.aif showcases the spatialisation technique of the **MT4L Tap Delay** and its ability to affect individual taps with unique effects. The effect can be clearly heard towards the end of the sound file as progressively more differences become apparent between the taps. With unique pitch shifting and filter cut-off's for individual taps, the sound file does demonstrate the primary ability of **MT4L-TapDelay**. The example could be developed further with more nuance in timing to make an excellent introduction to a musical piece.

Smudge-Temple.aif does demonstrate the **MT4L-Smudge** effect but only in short movements. Using longer spatial time delay and bigger gestures could help to highlight the effect. The original audio is filtered, making the spatialised audio harder to localise diminishing the overall effect. Additional higher frequency sound content could be introduced to increase the perceived spatial fidelity. There is also an appealing development when the spectral gate gets applied to the sound.

Granular-Bots.aif does not fully demonstrate all **MT4L granulator** has to offer and is somewhat lacking in conviction. The limited sound content is unappealing, and while the spatialisation is intriguing added care to the filtering and manipulation of sound parameters could provide more organic development. Similar ideas could be developed further by starting with more diverse sound material in the granulation buffer.

Collectively the sound examples give a limited glimpse at the creative potential of the MT4L suite. While the examples do not represent my initial intention for an EP of music or even musical studies, they portray the primary abilities of each of the devices in the suite.

7.2 FUTURE WORK

FURTHER COMPOSING

I would like to go back to my initial inspirations and clone some of the original tracks as studies with spatialisation. Specifically, I would start by recreating Second Woman's '/' using the randomising features of the **MT4L AmbiDelay**. I would also like to take inspiration from Curtis Road's (2002) ideas on microsound and start putting more care and effort into developing a sound over time. I would also like to explore the use of field recordings and found sound with the suite to create a rich aural backdrop contextualising the **MT4L AmbiDelay** sound objects, akin to the music of Visible Cloaks. I would create a diffuse drone texture using the **MT4L Granulator** boid behaviour parameters, place found sounds in the ambisonic sphere using the **MT4L TapDelay** as a looper and use the **MT4L Smudge's** spectral panner to blend sounds around the ambisonic sphere.

CREATIVE POTENTIAL

There is a great potential for creativity with the suite now the tools have been created. Composers can explore spatialisation in new and improved ways. I want to continue exploring spatialisation with the tools using gestural controllers in live performance. The existing framework of the suite also lowers the barrier for additional effects to be created. The encapsulated audio bridge could lead to the creation of more ambisonic effects like the *Elastic Spatial Ripple Filter*. (*Specific device potential is discussed in Section 6*).

FURTHER RESEARCH

Future research could investigate the interconnection of the *MTools for Live* suite with existing technologies, like the E4L devices and IEM plugins. Although a limited ambisonic audio bridge between *MT4L* and E4L is possible, collaboration with the developers of E4L could provide a more stable solution. The creation of an *MT4L VST Host* would enable the chaining of ambisonic plugins in Ableton Live without the need for manual routing. This would further realise the goals of this research, making more creative spatialisation available for composers.

BIBLIOGRAPHY

- Apel, S., Batory, D., Kästner, C., & Saake, G. (2013). *Feature-oriented software product lines: Concepts and implementation* (1st 2013. ed.). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Bjørn, K. (2017). *Push Turn Move: Interface Design in Electronic Music*. Denmark: Bbooks Media.
- Blauert, J. (1997). *Spatial hearing: The psychophysics of human sound localization* (Rev. ed.). Cambridge, Massachusetts: MIT Press.
- Canziba, E. (2018). *Hands-on UX design for developers* (1st ed.) Packt Publishing.
- Chion, M., & Steintrager, J. A. (2016). *Sound: An acoulogical treatise*. Durham: Duke University Press.
- Choi, C., Ivanic, J., Gordon, M., & Ruedenberg, K. (1999). Rapid and stable determination of rotation matrices between spherical harmonics by direct recursion. *The Journal Of Chemical Physics*, 111(19), 8825-8831.
- DTS, Inc. (2020). *Welcome To DTS:X*. dts.com. <https://dts.com/media-center/welcome-to-dtsx-open-immersive-and-flexible-object-based-audio-coming-to-cinema-and>
- Everest, F. A., & Pohlmann, K. C. (2015). *Master handbook of acoustics* (Sixth ed.). New York, USA: McGraw-Hill Education.
- Gardner, M. B. (1967). The haas and/or precedence effects in sound localization. *The Journal of the Acoustical Society of America*, 42(5), 1179-1179.
- Goderis, S. (2008). *On the separation of user interface concerns: A programmer's perspective on the modularisation of user interface code*. Brussels: VUBPRESS.
- Hodges, P. (2018). Ambisonic Info | Ambisonic Software. <https://ambisonic.info/practical/software.html>
- Howard, D. M., & Angus, J. (2013). *Acoustics and psychoacoustics* (4th ed.). London; Amsterdam. <http://www.qualitative-research.net/index.php/fqs/article/view/1589/30955>
- Izhaki, R. (2018). *Mixing audio: Concepts, practices, and tools* (Third ed.). New York, NY: Routledge, Taylor & Francis Group.
- Jackson, D., & Kang, E. (2010). Separation of concerns for dependable software design. Paper presented at the 173-176.
- Kendall, G., & Cabrera, A. (2011, July 31st-August 5th). WHY THINGS DON'T WORK: WHAT YOU NEED TO KNOW ABOUT SPATIAL AUDIO . *International Computer Music Association*, Huddersfield UK, 7

Mortimer Pavlitski

Mahé, P., Ragot, S., & Marchand, S. (2019). First-order ambisonic coding with quaternion-based interpolation of PCA rotation matrices. *EAA Spatial Audio Signal Processing Symposium*, Paris, France, 7-12.

Malham, D. (2003). *Space in Music - Music in Space*.

Nachbar C, Zotter F, Deleflie E, Sontacchi A. (2016). *AMBIX - A SUGGESTED AMBISONICS FORMAT*.

https://iem.kug.ac.at/fileadmin/media/iem/projects/2011/ambisonics11_nachbar_zotter_sontacchi_deleflie.pdf

Normandeau, R. (2009). Timbre spatialisation: The medium is the space. *Organised Sound : An International Journal of Music Technology*, 14(3), 277-285.

Payne, Roger S. (1962). How the Barn Owl Locates Prey by Hearing. *The Living Bird, First Annual of the Cornell Laboratory of Ornithology*, 151-159

Perumal, L. (2014). Euler angles: Conversion of arbitrary rotation sequences to specific rotation sequence. *Computer Animation and Virtual Worlds*, 25(5-6), 521-529.

Reynolds, C. (2019). Boids (Flocks, Herds, and Schools: a Distributed Behavioral Model). <https://www.red3d.com/cwr/boids/>

Roads, C. (2002). *microsound*. Cambridge, Mass: MIT Press.

Roginska, A., Geluso, P. (2017). *Immersive sound: The art and science of binaural and multi-channel audio*. London: Routledge, Taylor & Francis Group.

Rumsey, F. (2001). *Spatial audio*. Oxford, Boston: Focal Press.

Schutze, S., & Irwin-Schütze, A. (2018). *New realities in audio: A practical guide for VR, AR, MR, and 360 video*. (1st ed.). Boca Raton: CRC Press, Taylor & Francis Group.

Shepard, R. N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America*, 36(12), 2346-2353.

Shiffman D. (2012). *The Nature of Code*. natureofcode. <https://natureofcode.com/>

Silverman, L. (2005). *Upside-Down Brilliance: The Visual-Spatial Learner*, The Institute for the Study of Advanced Development. Retrieved from <http://www.pegy.org.uk/Upside-Down%20Brilliance%20-A4%20pdf.pdf>

Smith, H., & Dean, R. T. (2009). *Practice-led research, research-led practice in the creative arts*. Edinburgh: Edinburgh University Press.

Wilson, S. (2018). *Second Woman: The studio secrets and Back To The Future philosophy of their intense new album*. <http://www.factmag.com/2017/05/13/second-woman-interview/>

Mortimer Pavlitski

Zotter, F., & Frank, M. (2019). *Ambisonics: A practical 3D audio theory for recording, studio production, sound reinforcement, and virtual reality* (1st 2019. ed.). Cham: Springer International Publishing.

Cycling '74. (2020). *Max for Live Production Guidelines*. maxforlive.com. <https://maxforlive.com/resources/M4L-Production-Guidelines.pdf>

IRCAM. (2002). *LISTEN HRTF DATABASE*. ircam.com. <http://recherche.ircam.fr/equipes/salles/listen/>

Oculus | VR Headsets & Equipment. (2020). <https://www.oculus.com/>

The University of Huddersfield. (2020). Spatialisation and Interactive Research Lab (SPIRAL). <https://research.hud.ac.uk/institutes-centres/cerenem/environment/facilities/studios/>

Wikipedia. (2020). *Comparison of virtual reality headsets*. wikipedia.org https://en.wikipedia.org/wiki/Comparison_of_virtual_reality_headsets

Zurich University of the Arts. (2019). *ICST Ambisonics*. zhdk.ch. <https://www.zhdk.ch/forschung/icst/software-downloads-5379/downloads-ambisonics-externals-for-maxmsp-5381>

Mortimer Pavlitski

DISCOGRAPHY

Visible Cloaks. (2017). Reassemblage. [WAV] US: Rvng Intl.

Second Woman. (2017). S/W. [AIFF] Austria: Spectrum Spools.

RESEARCH INSPIRATION PLAYLIST

Ashikawa, S. (1982). *Still Way* [Vinyl]. Japan: Sound Process.

<https://www.youtube.com/watch?v=MojW0fVaaFE>

Boards Of Canada. (2002). Geogaddi. [CD] UK: Warp Records.

Four Tet. (2012). Pink [MP3]. UK: Text Records.

Four Tet. (2017). New Energy [MP3]. UK: Text Records.

Jan Jelinek. (2001). Loop-finding-jazz-records. [CD] Germany: ~Scape.

Loscil. (2016). Monument Builders [MP3]. US: Kranky.

Second Woman. (2016). Second Woman. [WAV] Austria: Spectrum Spools.

Steve Hauschildt. (2011). Tragedy & Geometry. [CD] UK: Kranky.

Yoshimura, H. (1982). *Music For Nine Post Cards* [Vinyl]. Japan: Sound Process.

<https://www.youtube.com/watch?v=zyY7NU4cbtY>

Yoshimura, H. (1986). *Soundscape 1: Surround* [CD]. Japan: Misawa Home.

<https://www.youtube.com/watch?v=MojW0fVaaFE>

SOFTWARE REFERENCES

360° Ambisonics Mixing Tools | Waves. (2019). <https://www.waves.com/hardware/360-ambisonics-tools>

Ableton. (2018). Live 10 [MacOS]. Berlin. <https://www.ableton.com>

Ambisonic Toolkit (ATK) for Reaper. (2015). <http://www.ambisonictoolkit.net>

AmbiX. (2019). <http://www.matthiaskronlachner.com/?p=2015>

Audio 360. (2019). Facebook. <https://facebookincubator.github.io/facebook-360-spatial-workstation/Documentation/SpatialWorkstation/SpatialWorkstation.html>

CICM. (2019). *Hoalibrary High Order Ambisonics Library*. <http://hoalibrary.mshparisnord.fr/en>

Cockos Incorporated. (2020). *REAPER | Audio Production Without Limits*. Reaper. <https://www.reaper.fm/>

Cycling '74. (2018). Max 8 [MacOS]. San Francisco. <https://cycling74.com>

Cosm. (2008). W. Smith, G. Wakefield. <http://w2.mat.ucsb.edu/cosm/index.html>

Device Library. (2020). Max for Live. <http://maxforlive.com/library/index.php>

Dolby Laboratories, Inc. (2020). *Dolby Atmos*. dolby.com. <https://www.dolby.com/technologies/dolby-atmos/>

Envelop. (2020). Envelop <http://www.envelop.us/software/>

Facebook. (2020). *Facebook 360*. <https://facebookincubator.github.io/facebook-360-spatial-workstation/Documentation/SpatialWorkstation/SpatialWorkstation.html>

Google. (2020). *YouTube*. google.com. <https://support.google.com/youtube/answer/6395969?hl=en-GB>

IEM Plug-in Suite (2020) *Institute of Electronic Music and Acoustics*. <https://plugins.iem.at>

Ina-GRM. (2011). GRM Tools 3 Spaces [MacOS]. Paris. <https://inagrm.com/en/store/product/15/spaces>

IRCAM. (2019). SPAT [MacOS]. Paris. <https://www.flux.audio/project/spat-revolution/>

IRCAM. (2019). SPAT5 [MacOS]. Paris. <https://forum.ircam.fr/projects/detail/forummaxapps2-spat5/>

Mortimer Pavlitski

McCulloch, P. (2012). SubtleSonic EnvelopeShaper 1.0. maxforlive.com.

<https://maxforlive.com/library/device/1126/subtle-sonic-envelope-shaper>

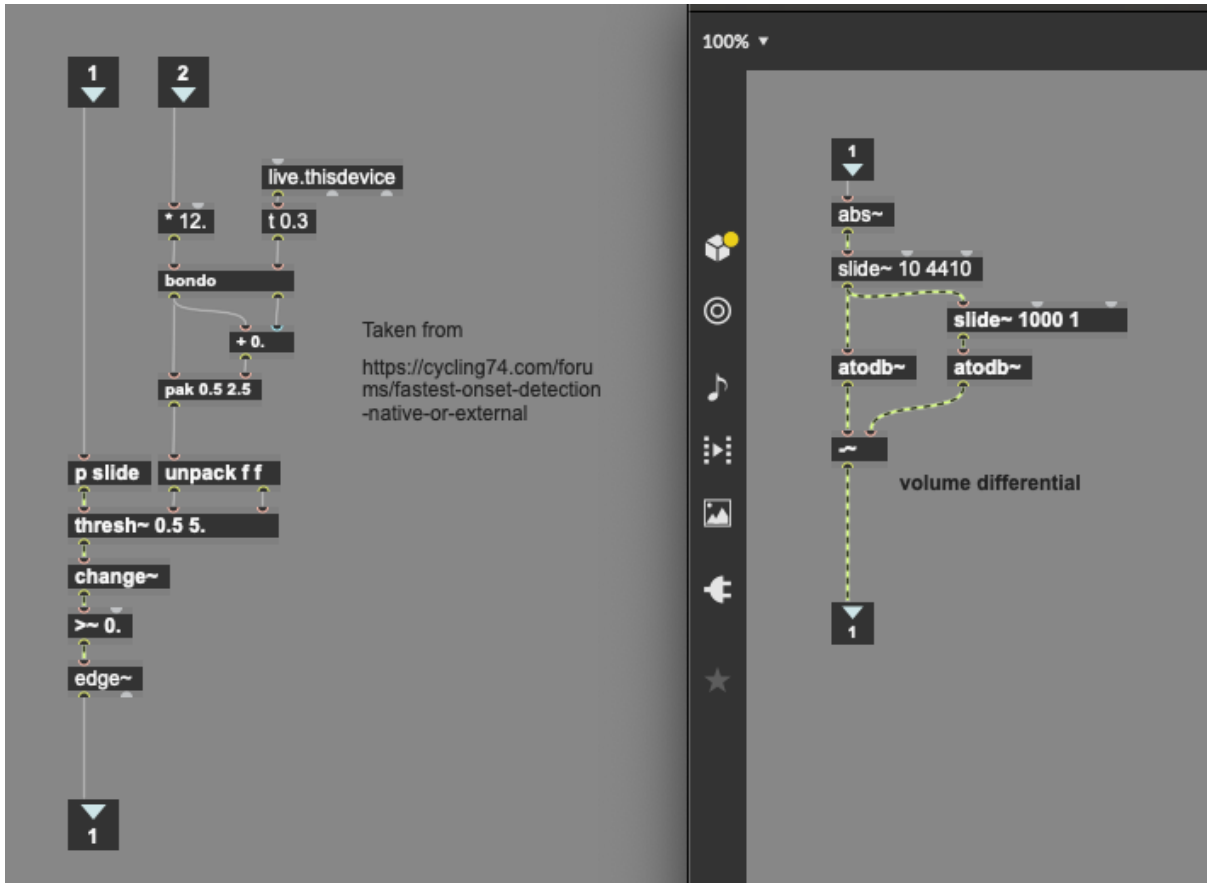
Monolake. (2020). *Circular Doppler*. maxforlive.com. <https://maxforlive.com/library/device/239/circular-doppler>

O3A Manipulators. (2019). <https://www.blueripplesound.com/products/o3a-manipulators>

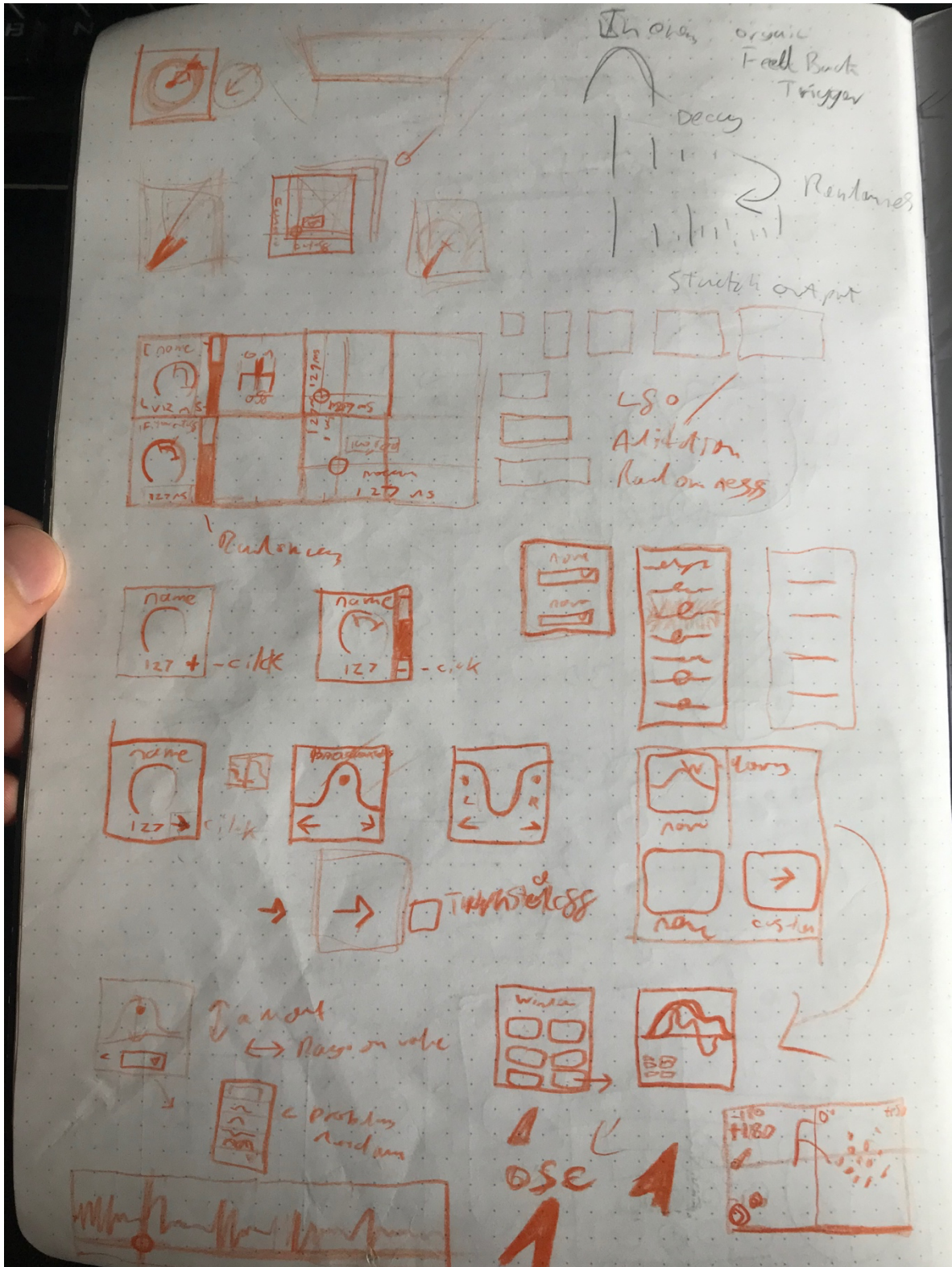
SSA Plugins. (2019). <https://www.ssa-plugins.com/>

APPENDIX A

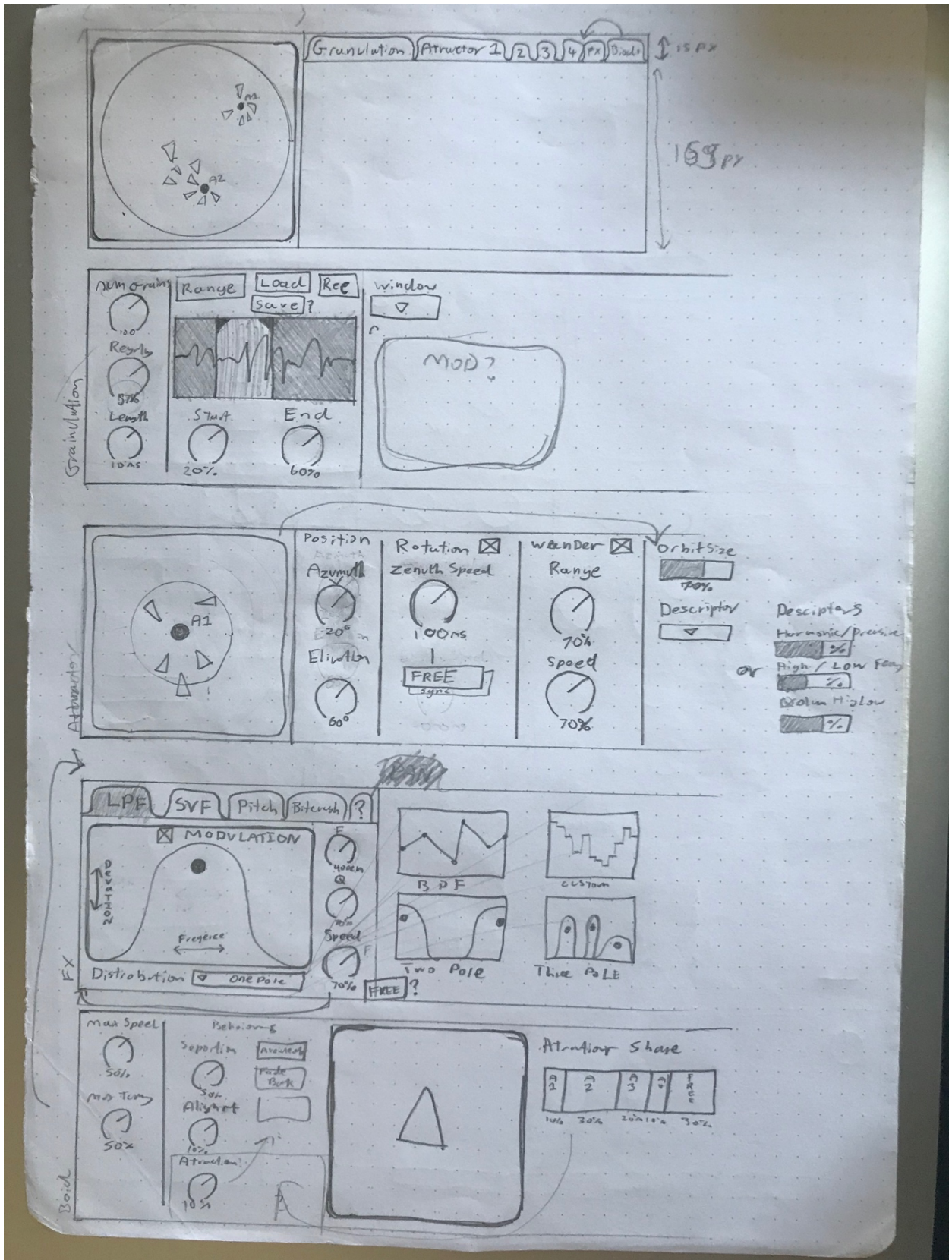
The code copied from SubtleSonic EnvelopeShaper 1.0. (McCulloch, 2012).



APPENDIX B



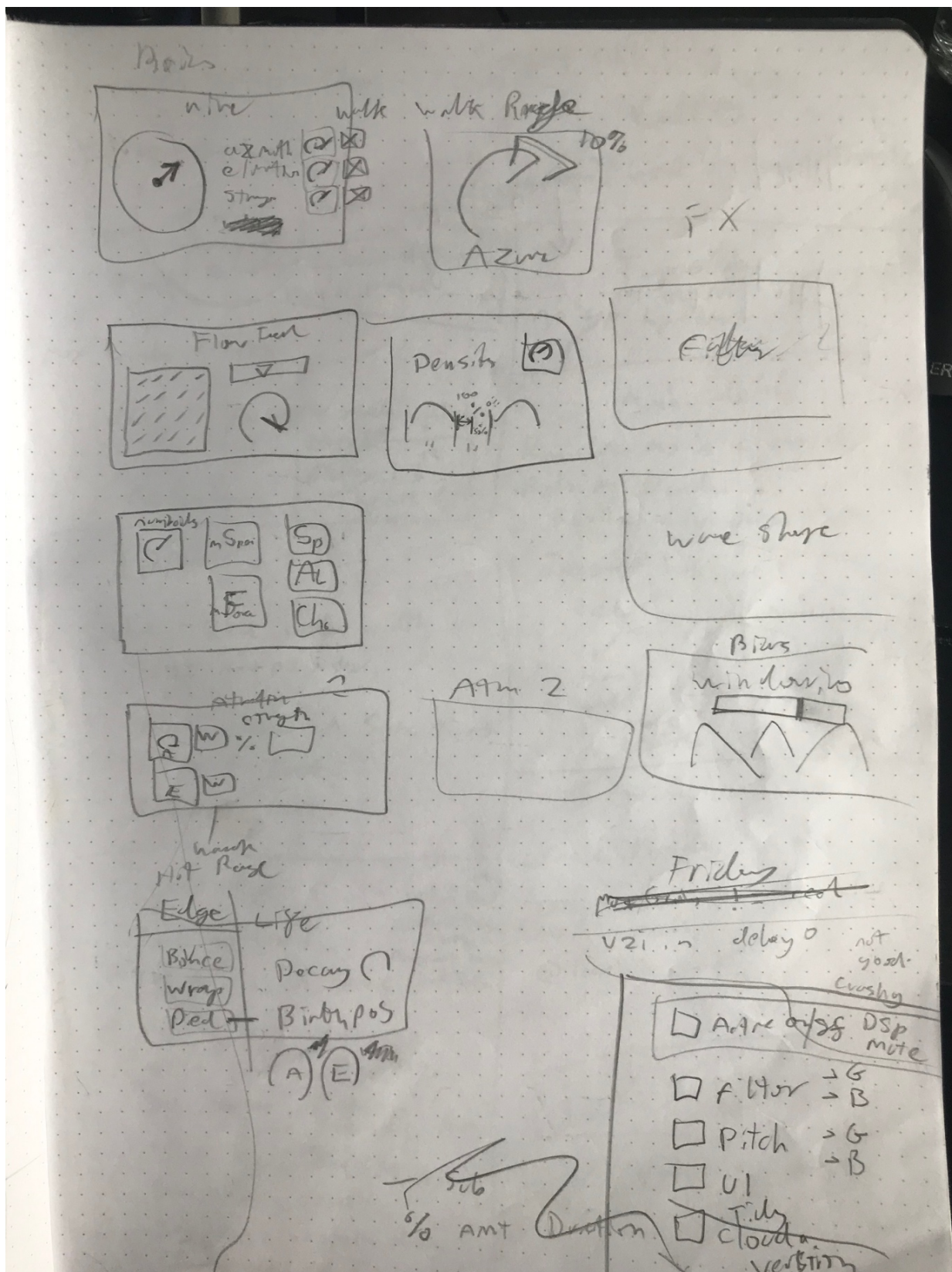
Early sketches of modular user interface components.



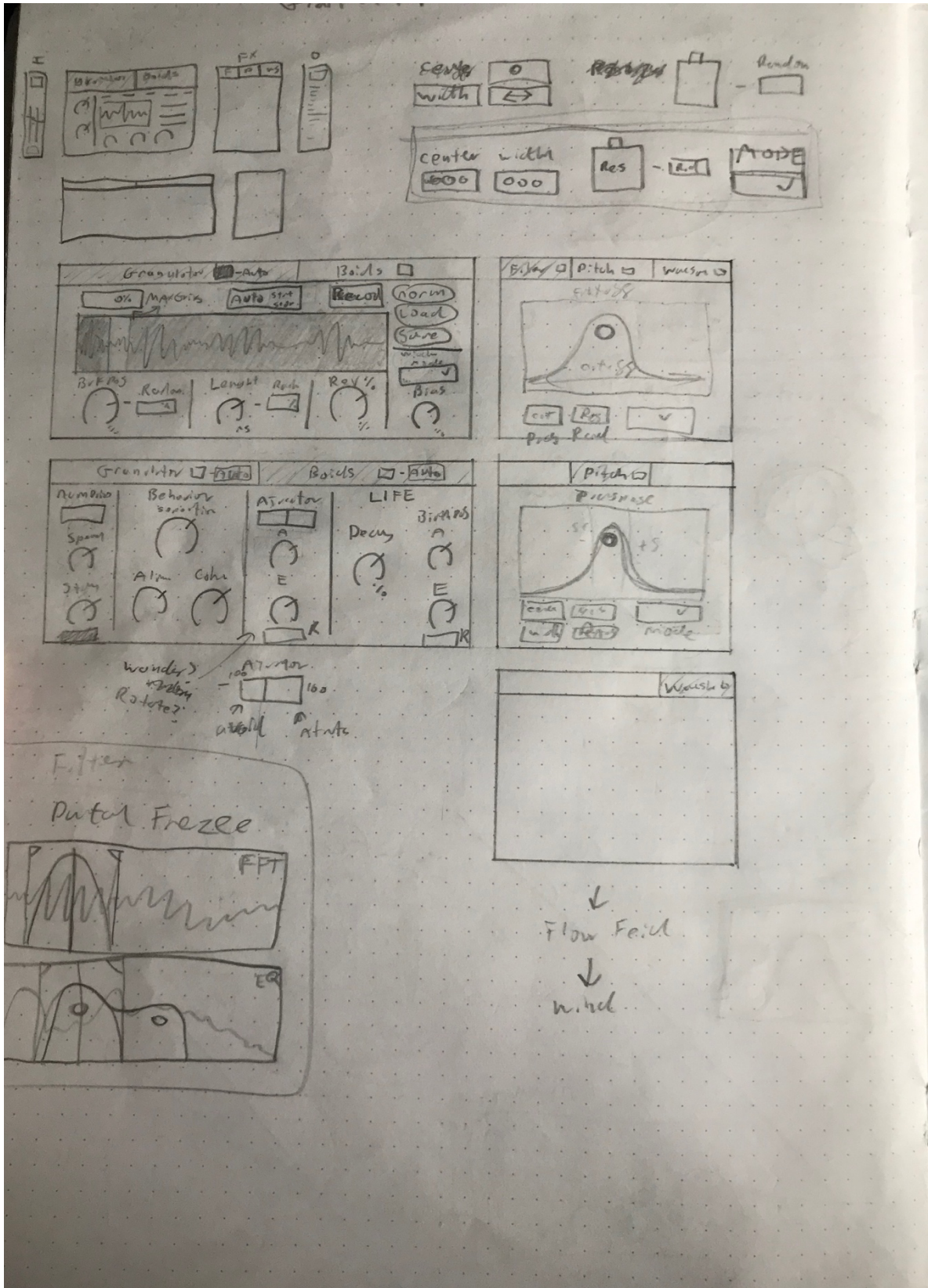
Early sketches of the MT4L-Granulator's graphical user interface.



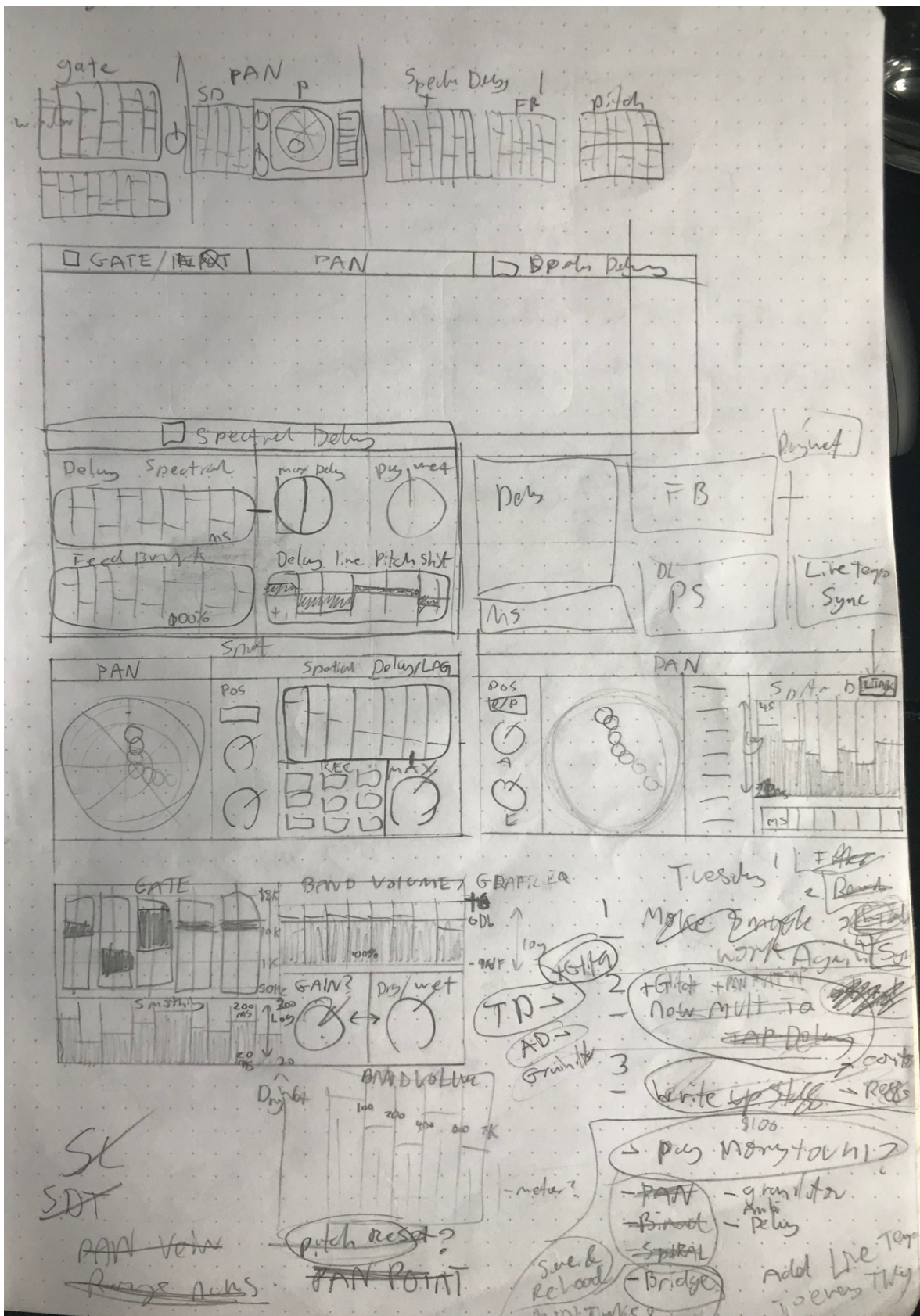
Early sketches of the MT4L-Granulator's graphical user interface.



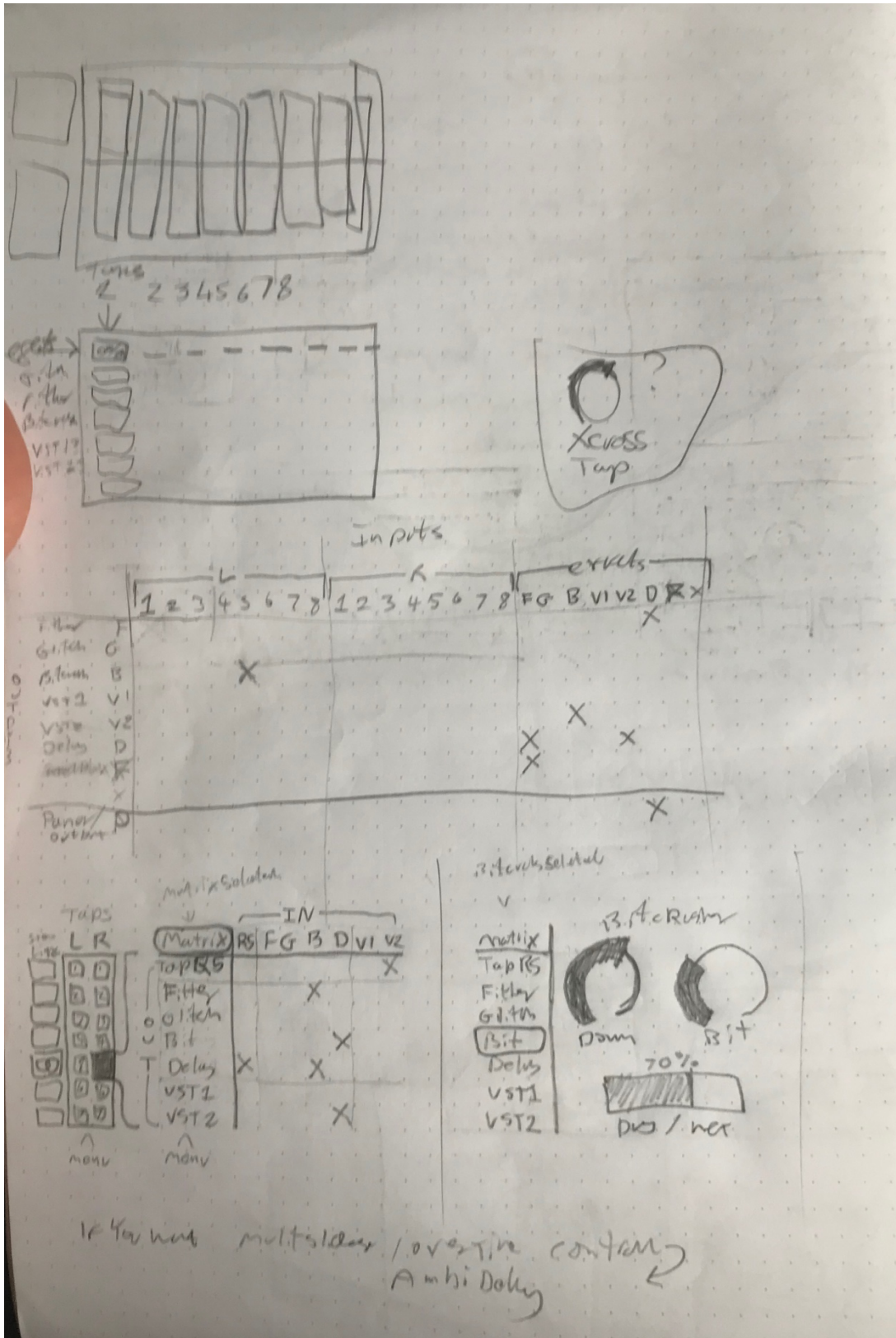
Early sketches of the MT4L-Granulator's graphical user interface.



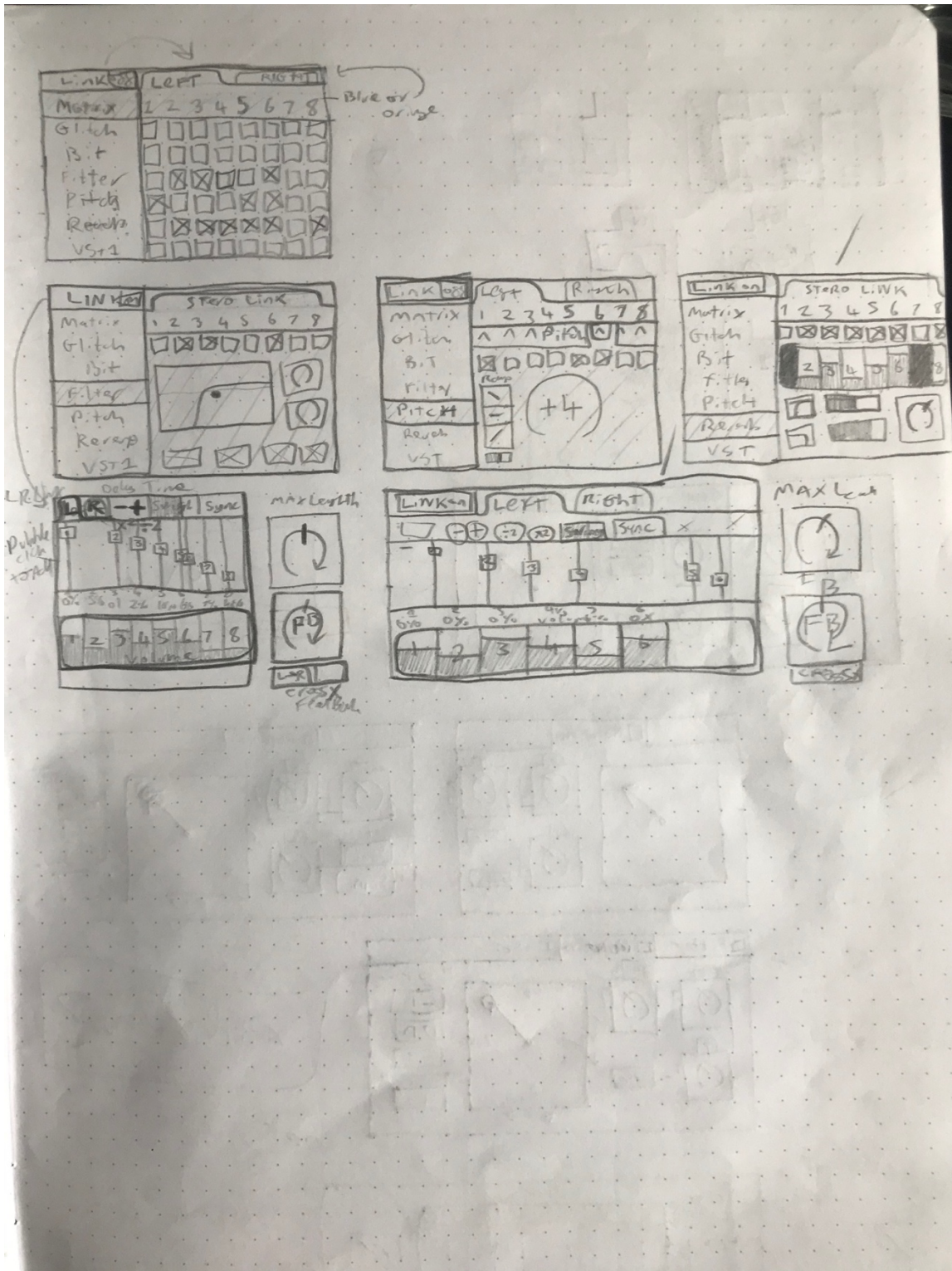
Later sketches of the MT4L-Granulator's graphical user interface.



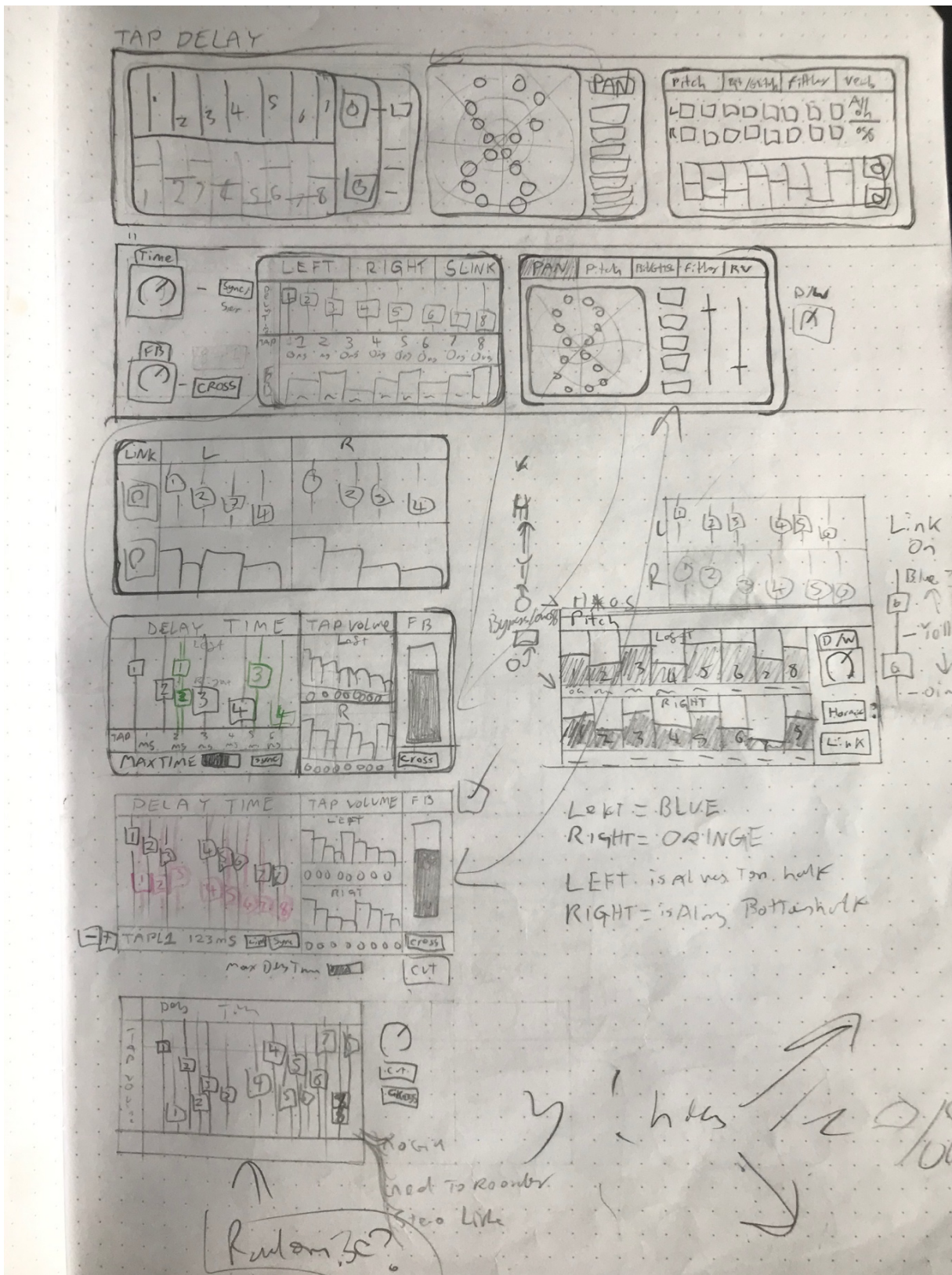
Sketches of the MT4L-Smudge's graphical user interface.



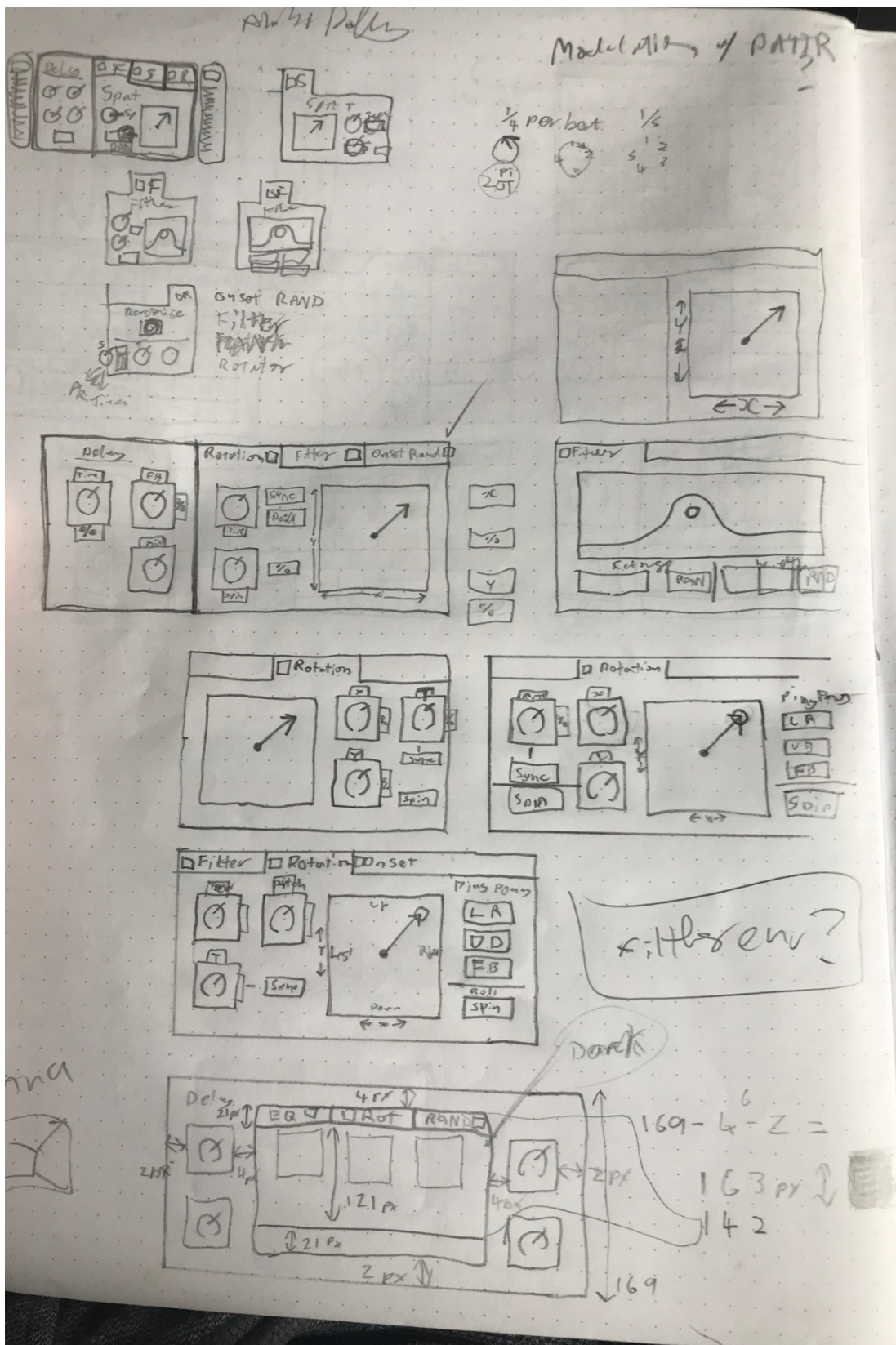
Early sketches of the MT4L-TapDelay's graphical user interface.



Midway sketches of the MT4L-TapDelay's graphical user interface.



Later sketches of the MT4L-TapDelay's graphical user interface.



Sketches of the MT4L-AmbiDelay's graphical user interface.