You may print or download ONE copy of this document for the purpose of your own research or study.

The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following:

- Copyright owners are entitled to take legal action against persons who infringe their copyright.
- A reproduction of material that is protected by copyright may be a copyright infringement.
- A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.
Spectromorphology and Spatiomorphology: Wave Terrain Synthesis as a framework for controlling Timbre Spatialisation in the Frequency Domain

This thesis is presented for the degree of

Doctor of Philosophy

Stuart George James

Edith Cowan University
Faculty of Education and Arts
Western Australian Academy of Performing Arts
2015
Chapter 4: Sound Shapes and Spatial Texture

4.1 Introduction

Chapter 4 is concerned with the performative aspects of the two models evaluated in Chapter 3, and the variety of sound shapes that emerge from the symbiosis of WTS and timbre spatialisation. As illustrated in Figure 81, the laptop performer is not intentionally and consciously responsible for the location of every point in space, and neither do they directly control the spectral processing, but rather the current state and evolving state of the terrain and trajectory, which subsequently determine the result of the spectral process. As a result of this, the performer has the opportunity to control the nature of the spatiomorphology and spectromorphology generated.

Figure 81. A block diagram of Model B outlining the processes that relate to WTS and those that relate to timbre spatialisation. The user interaction above clearly shows that the performer is only responsible for controlling the WTS section of the instrument.

In an extreme case, if we took a 4096-sample length FFT size, and were performing timbre spatialisation across 32 speakers, accounting for azimuth, elevation and distance for each frequency band, this would result in 196,608 simultaneous control parameters.
This chapter is structured in three parts: visualisation, generative sound shapes, and the morphology of sound shapes.

Many of the sound shapes created by Model A and B are difficult to describe without the use of visual illustrations, as they often consist of complex spatial movements. For this reason an emphasis has been placed on the importance of visualisation for the laptop performer. The first part of this chapter discusses the many kinds of visual tools and diagrams used that assist in illustrating the process. One of these visualisations is a frequency–space representation that shows how different frequencies move and interact across space. This visualisation is ideal for presenting the nature of sound shapes generated.

The exploration of sound shapes is a two-part process that involves the creation of static frequency–space distributions, and then involves the morphology of these distributions in order to animate them. For this reason, the remainder of the chapter is split into the generation and morphology of sound shapes.\(^\text{171}\)

The generation of sound shapes is primarily concerned with the distinct shapes that emerge through specific combinations of terrain and trajectory curves. This section systematically and explicitly shows terrain and trajectory pairs in such a way as to illustrate the pertinent features of Models A and B. The morphology of sound shapes is concerned with how the laptop performer is able to shift smoothly between several sound shapes, and apply transformational processes to the terrain and trajectory curves. This process is central to the performer’s ability to control the spectromorphology and spatiomorphology of sound shapes generated by timbre spatialisation.

\(^{171}\) This kind of classification is relatively common for spatial trajectories and movements. OMPrisma is one such application (Schumacher & Bresson, 2010). The HoloEdit application categorises functions into three types: generative (i.e. circular, random, etc.), spatial transformative (interpolation, symmetry, translation, etc.) and temporal transformation (acceleration, time stretch, etc.).
4.2 Visualisation

The interaction between instrument and performer is a particularly important dynamic. From the very beginning of this research project, it was evident that a visual interface for *timbre spatialisation* would assist in illustrating the spatialisation of spectra.

WTS also benefits from visualisation of the *terrain* and *trajectory* curves. These are two independent structures running concurrently, one being a haptic rate and the other an audio rate system. Due to the extent of non-linearities possible through this interaction, as will be seen in the *sound shapes* generated, the performer needs a visual aid to indicate the current state of the *terrain* and *trajectory* curve. This is particularly important when transformation of both of these structures occurs, as this is the visual guide that informs the performer of how this morphology unfolds, and serves as a confirmation of the associated auditory *sound shapes* generated. The visualisations are helpful for the reader too, however Chapter 4 is aimed at discussing the performative approaches to the instrument built, and so the discourse is centered around performative issues. One of these primary issues is the intuitive nature of the user interface, and the visualisations are important in establishing this.

For the reader, without the dimension of time, all of these diagrams do not encode all of the complexities that emerge throughout their evolution. Therefore observing how the visualisations change through time has been a problem. However, there are examples in Chapter 4 that have been made specifically to show evolution over time, and how the visualization changes accordingly. Descriptions are also provided, and most of Chapter 4 is written specifically in such a way as to explain the visualisations, and how these relate to specific sound shapes generated when using WTS to control *timbre spatialisation*. 
In conjunction with visual interfaces for displaying the terrain and trajectory structures, a number of other visualisations can be helpful. Loudness metering from the master output bus, multichannel spectrograms or sonograms, and frequency–space plots are some of the visualisations used.

4.2.1 Visualising the Terrain

In these software implementations, the terrain curve is a haptic rate multidimensional data array, which in these implementations is stored and transmitted in the form of video. Visualising the terrain surface is of importance for all implementations in order to make it explicit how WTS is deriving its resulting audio signal. Most of the timbre spatialiser instruments simply display a 2D representation of the terrain in the form of a density plot (pictured in Figure 82a); however virtual 3D renderings were also explored as a way of further enhancing the impression of the sound shapes being generated.

![Figure 82a](image1.png) A density plot of a terrain surface showing the height map with hue colour coding.

![Figure 82b](image2.png) A non-real-time 3D rendering of the terrain in Figure 82a plotted in Mathematica with the same hue colour coding.

Virtual 3D plots of the terrain surfaces have been explored, including the Mathematica plot in Figure 82b. Real-time virtual 3D implementations in MaxMSP used jit.gl.mesh (Figure 83b) and though this is a fairly concise way of generating
virtual 3D surfaces using the hardware on-board graphics processor, there are limited options for the nature of the colour coding used (Taylor, 2012). Later software revisions utilised the *jit.gl.render* object, providing considerably more customisability in terms of the nature of the colour rendering used. The colour scale was experimented with, as exemplified by the gold and red shown in Figure 83c, but unlike 2D representations that rely on specific differences of colour to highlight precise relationships, the choice of colour scale for 3D renders became less of an issue as the height and depth of the terrain were accentuated by shadows cast by virtual lighting sources when rendering virtual 3D objects using OpenGL. The *jit.gl.render* object expects up to a 12-plane matrix that describes topographical and visual characteristics of the structure to be rendered. The most relevant planes for the purpose here are planes 0–2, which correspond with the coordinates of $x$, $y$ and $z$ vertices, and planes 8–10, which correspond with red, green or blue components of the colour.

Although all of these methods proved to be successful, hue colour coding was preferred in later software revisions as it ties more concisely with the diagrams and colour scales used throughout the exegesis.
4.2.2 Visualising the Trajectory

The trajectory is a two- or three-multichannel audio signal representing a 2D or 3D spatial trajectory. Visualising the trajectory can be achieved by writing this two-channel audio signal to a video matrix using \texttt{jit.poke~} (Figure 84a), or a three-channel vector of audio samples can be rendered in 3D using \texttt{jit.gl.render} (Figure 84b). Rendering in OpenGL provides the facility for viewing the elevated dimension giving an impression of height in the trajectory curve.

Figure 84a. A 2D rendering of modulated random walks using \texttt{jit.poke~}, \texttt{jit.matrix}, and \texttt{jit.glop}.

Figure 84b. A 3D rendering of modulated random walks using \texttt{jit.gl.render} and \texttt{jit.gl.mesh}.
Trajectories plotted using both *jit.poke*~ and *jit.gl.render* can be made to prolong the temporal history of the terrain by fading out pixels gradually. This was explored using *jit.glop* to gradually fade instances that have appeared previously, giving a smoother display of how the trajectory changes and shifts over time. This is important as the visual ‘memory’ serves as a reference for morphology. This can similarly be achieved using *jit.gl.render* by adjusting its fourth argument for erase_colour.

A provision was made for *Model A* allowing the trajectory plot to be encoded in colour, which would ultimately highlight the spectral bins and how they aligned with the movement of the trajectory. This required the building of an audio rate hue–saturation–value to red–green–blue colour space converter,\(^\text{172}\) as the colours needed to be rendered at the same resolution as the trajectory. This became an important visual diagnosis tool, as it is possible to see how spectral bins align with the trajectory. It is particularly important in diagnosing whether or not a trajectory is synchronised with the FFT frame, because if it is synchronised the colours will not shift (Figure 85a). Trajectories that do shift across FFT frames often will give rise to an effect described as *spatial texture* (Figure 85b).

4.2.3 Multichannel Loudness Metering

In live multichannel contexts the diffusionist is often responsible for a master bus that consists of many channels of audio. Logistically it does not make sense to adjust every one of these levels independently, but rather use a single fader on the master output level. The `live.gain~` object (Figure 86) is a software option that can be configured for a customised number of channels. The object defaults to 0 dBFS (decibels relative to full scale), so provision was made in the software to set the object to $-\infty$ dBFS by default, in case there is a substantial amount of signal level already passing through the patch. Effective monitoring is always important in order to gauge peak level, and whether any digital clipping is evident on output.
4.2.4 Visualising the Frequency–Space Domain

Spatialising spectra across space may be considered analogous to the painting of sound across a virtual canvas. Visualising the \textit{sound shapes} that emerge can be helpful in providing another perspective on the process that is unfolding.

Kim-Boyle (2006) used OpenGL renderings in \textit{MaxMSP} to depict the kinds of movements generated by the Boids and some other algorithms, for \textit{spectral spatialisation}. Figure 87 shows how Kim-Boyle maps individual bins of the FFT to different colours—lower frequency bins to reds, higher bins to violet and those between, the orange through indigo colour spectrum.

Figure 87. Several time-displaced frames of frequency–space distribution as shown in the work of David Kim-Boyle (2006, 2008).
Kim-Boyle explains that though this is useful for tracking the movements of certain spectral bands, it requires a 12-plane *Jitter* matrix to be used, which can slow down performance—a significant consideration if the visual information is to be used for real-time performance.

Kim-Boyle outlines the problem of enormous amounts of data presented using OpenGL renderings for visual feedback. This applies to a certain kind of movement such as that generated by the Boids algorithm, but Kim-Boyle suggests that visualisation is not required for every aspect of this movement.

There are few commercial applications designed to represent this kind of frequency–space visualisation. One exception is the *Flux Pure Analyzer Essential*\(^{173}\) software with its multichannel option. This tool plots the spectrum of a sound in terms of its spatial distribution supporting up to eight discrete channels. This is possible from a DAW using the *Flux SampleGrabber*\(^{174}\) plugin. By default, 8.0 is mapped in this way:

Channel 1 – Left (L)  
Channel 2 – Centre (C)  
Channel 3 – Right (R)  
Channel 4 – Left surround (Ls)  
Channel 5 – Right surround (Rs)  
Channel 6 – Left back surround (Lbs)  
Channel 7 – Right back surround (Rbs)  
Channel 8 – Centre surround (Cs)

As the software developed in this project is mapped to a circular clockwise configuration of speakers, the mapping must conform to the following order:

Channel 1 – Left (L)  
Channel 2 – Centre (C)  
Channel 3 – Right (R)  
Channel 4 – Right Surround (Rs)  
Channel 5 – Right Back Surround (Rbs)  
Channel 6 – Centre Surround (Cs)  
Channel 7 – Left Back Surround (Lbs)  
Channel 8 – Left Surround (Ls)

---

174 Retrieved 10\(^{th}\) Jan 2015 from https://www.fluxhome.com/download
The *Flux Pure Analyzer Essential* plots psychoacoustic informed renderings by calculating the relative balance and position of different frequency components, allowing the *sound shapes* to be visualised, as shown in Figure 88.

![Figure 88. Two time-displaced frames of frequency–space distribution plotted using Flux Pure Analyzer Essential software using the multichannel option. These are two sound shapes from my work Veden Ja Tulen Elementit (2013).](image)

My implementation of a frequency–space visualisation in *MaxMSP* resulted in two different visualisations for *Model A* and *B* originally introduced at the end of Section 1.2.4. This process involved colour coding the spatial cues with the associated frequency. This process was implemented using *jit.poke*~ and, like in Section 4.1.2, converts RGB colour values at audio rate. This process was achieved using Java code (Appendix C). Some examples of this rendering are provided in Figure 89. This enabled the laptop performer to visualise the frequency–space distributions generated by *Models A* and *Model B*. It is hoped to extend this software in future to be able to display this colour-coded trajectory in three dimensions.
4.2.5 Multichannel Spectrograms and Sonograms

For the purposes of evaluating how sound shapes evolve over time, it is necessary to look for visualisations—both real time and non-real time—that illustrate a history of frequency–amplitude. Examples of such tools include Metric Halo Spectrafoo,¹⁷⁵ UI Software Metasynth, IRCAM Audiosculpt,¹⁷⁶ Izotope RX4,¹⁷⁷ INA-GRM Acousmographe¹⁷⁸ and Pierre Couprie’s EAnalysis.¹⁷⁹ Surprisingly there did not appear to be many tools on the market allowing for precise spectrogram analyses of multichannel audio files—that is, audio files with more than two channels.

Non-real-time multichannel spectrograms were rendered using IRCAM Audiosculpt software, as shown in Figure 90. It was in such multichannel spectral analyses that one could visualise spatial texture, due to complex interactions of terrain.

---

¹⁷⁶ Retrieved 10ᵗʰ Jan 2015 from http://forumnet.ircam.fr/product/audiosculpt/
and trajectory structures. One can see in these analyses the movement of different spectral bands across an eight-loudspeaker system, tracing the various non-linear spatial movements generated.

**Figure 90a.** An eight-channel spectral analysis showing the creation of spatial texture. This is generated on Model A using a chaotic trajectory that is geometrically modulated over a linear terrain function.

**Figure 90b.** An eight-channel spectral analysis showing the creation of spatial texture as well as more distinct and directional spatial movement across the system. This is generated on Model A.

As a real-time solution in MaxMSP, a multichannel sonogram was also added to the interface window of both Models A and B in order to monitor the spectrum of sound reproduced in each loudspeaker. The advantage of the sonogram is that it reports a history of changes in the spectrum over time, meaning it is possible to visualise the nature of the spectromorphologies as they evolve. Although these diagrams did not give a sense of the way in which frequencies interact across space, their temporal and unfolding qualities are clear. Figure 91 shows two very distinctive kinds of spectral shapes, some of the many that emerge through this process.
4.3 The Generating of Sound Shapes

This section is concerned with the language of sound shapes that may be derived by Models A and B as a result of the performer’s interaction with WTS. This involves a systematic evaluation of the kinds of sound shapes created by both models in separation along with documentation outlining the kinds of terrain and trajectory curves used for reproducing such cases. The intention here is to establish a meaningful inter-relationship between the topographies and geometries used and their resulting sound shapes, and to outline those cases that are more intuitive. The necessity for deriving a meaningful symbiosis between WTS and the subsequent distribution of spectra is an important one. This relates to how predictable a series of outcomes are.

The trajectory and terrain are not independent in determining an outcome. Rather, due to the lookup process they both interact and are interdependent in determining the outcome (James, 2005). Therefore, exploration of the many different kinds of terrain and trajectory curves as applied to timbre spatialisation was considered of high importance, as it is significant in determining both the spatial distribution of spectra and the way in which this evolves. Although it is not possible to show the many different sound shapes possible, this section outlines some of the more pertinent cases that show a diversity of frequency–space distributions.

Figure 91a. A distinct sound shape evolving over time where there appears to be two different processes running concurrently.

Figure 91b. A distinct sound shape evolving over time.
What follows is a categorical list of terrain curves, and both synchronous and asynchronous trajectory curves, and then an evaluation of the sound shapes created through the combination of these structures. Sound shapes are evaluated separately for Models A and B in terms of their frequency–amplitude plots and in terms of psychoacoustic attributes.

4.3.1 Generative Terrain Curves

Previous research resulted in a catalogue of mathematically derived wave terrain surfaces (James, 2005, Appendix A). Terrain curves used for WTS included piecewise functions (Mitsuhashi, 1982), polynomials (Borgonovo et al., 1984, 1986; Mikelson, 2000), fractal and iterative maps (Di Scipio, 2003; James, 2003; Mikelson, 2000), additive and Chebyshev functions (Nelson, 2000), elliptic functions (Catagna et al., 2002), dynamical systems of equations (Mikelson, 2000), Perlin noise functions (James, 2005),180 recurrence plots (James, 2005), surgical biomedical data (Jovanov et al., 1999; Wegner, 1998), topographical data taken from analyses of global seafloor and land topography (Thibault & Gresham-Lancaster, 1992a, 1992b), and video data (Dannenburg et al., 2003; Dannenburg & Neuendorffer, 2003). Other terrain curves ranged from finite solutions to constrained algebraic, trigonometric, logarithmic/exponential and complex functions, through to OpenGL NURBS surfaces. Many of these techniques fall into a category of video synthesis. Andrew Benson presents many more ideas for generative video in the various MaxMSP patches he has published online.181 Other methods such as texture synthesis use algorithms to construct large digital images from a small digital sample image by taking advantage of its structural content. This is used in many fields including digital image editing, 3D computer graphics and post-production in the film industry.

180 Another technique (invented by Ken Perlin) for generating all kinds of random periodic textures such as artificial terrain, clouds, water, wood, fire and marble.
181 Retrieved 10th Jan 2015 from http://cycling74.com/category/articles/jitter-recipes/
A number of techniques classed as ‘computational geometry’ involve the generation of cellular texture. Such techniques include Voronoi diagrams, Delaunay triangulation and convex hulls.\textsuperscript{182}

When applying these topographies for \textit{timbre spatialisation}, the performer is interested in the variety of curves possible as every unique topography results in its own intrinsic spatiomorphological and spectromorphological features. Figure 92 illustrates a small number of the infinite variety of \textit{terrain} curves possible. Most of these are mathematically derived except that Figure 92i is a digital photograph.

\textbf{4.3.2 Generative Trajectory Curves}

The exploration of \textit{sound shapes} is ultimately defined by the geometries exhibited by a \textit{trajectory}—a multichannel audio rate signal. The \textit{trajectory} determines not only the way in which the spatialisation reveals sections of the \textit{terrain}’s topography, but it can determine audio rate modulations of the frequency–space distribution, allowing for some complex non-linear results.

\textsuperscript{182} See Nylander (2004).
Figure 92a. Ramp functions.

Figure 92b. Curved functions.

Figure 92c. Polar/spiral functions.

Figure 92d. Modulo functions.

Figure 92e. Complex-domain curves.

\[ f(x,y) = \text{Im}(\sqrt{x + iy}) \]

Figure 92f. Voronoi diagrams.

Figure 92g. Fractal turbulence.

Figure 92h. Noisy functions.

Figure 92i. Real-world functions.

Trajectories as documented in my previous research in WTS have been categorised typically according to time domain characteristics such as their level of periodicity, quasi-periodicity or whether they exhibit more chaotic or random qualities.
as shown in Table 10. Many of the trajectories here are derived algorithmically, but some repeating trajectories could be generated from existing data sets. Calculation of these trajectories involves a range of computations including algebraic, trigonometric, iterative, procedural and vector-based processes. All of the curves render coordinates in three dimensions where possible, so that for scenarios like Model B, the trajectory can derive azimuth, distance and elevation cues.\footnote{\cite{James2005}}

Table 10

Various Kinds of Trajectory Curves (James, 2005)

<table>
<thead>
<tr>
<th>Evolution</th>
<th>Curve types</th>
<th>Graphical representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>Point</td>
<td><img src="image1" alt="Constant Point" /></td>
</tr>
<tr>
<td>Periodic</td>
<td>Lines, curves, Lissajous figures, Bresenham’s algorithm</td>
<td><img src="image2" alt="Periodic Lines" /></td>
</tr>
<tr>
<td>Quasi-Periodic</td>
<td>Invariant tori, spirals, spirographs, real-instrument musical signals, turtle graphics, phase–space plots, pseudo-phase–space plots</td>
<td><img src="image3" alt="Quasi-Periodic Invariant Tori" /></td>
</tr>
<tr>
<td>Chaotic</td>
<td>Strange attractors, webs and wreaths, environmental sound signals, mazes, traditional space-filling curves, L-systems</td>
<td><img src="image4" alt="Chaotic Strange Attractors" /></td>
</tr>
<tr>
<td>Stochastic</td>
<td>Random walks, noise, jitter</td>
<td><img src="image5" alt="Stochastic Random Walks" /></td>
</tr>
</tbody>
</table>

\footnote{\cite{James2005}} Noise generators, Lissajous figures and oscillators can easily be extended to three dimensions. Many strange attractors already involve three dimensions, such as the Rossler and Lorenz strange attractors.
Other kinds of approaches for generating trajectory curves, such as particle systems, involve the moving of particles through space using vector fields, vector math and quaternions. This is often classed as a field called ‘kinematics’ (Pecino, 2014). Such vector-based systems also give rise to behavioural systems such as flocking and swarming algorithms like the Boids algorithm. Some of these are discussed as transformational trajectories in Section 4.3.2.3.

As WTS is being used to control an FFT-based process, the temporal aspects of the trajectory and its synchronicity with the FFT process over time describes the nature of how a sound shape evolves. This made it apparent that some trajectories are derived that are synchronous to the FFT, and others are asynchronous. By colour coding the points of the trajectory to different frequency bins it is possible to observe visually this synchronicity in a way that monochromatic illustrations do not allow.

To serve as a more specific breakdown of trajectory types in relation to FFT processes, the following sections briefly describe both synchronous and asynchronous trajectories, as well as the space-filling qualities of trajectory curves before discussing a range of sound shapes.

Trajectory curves were coded in Java due to its support for turning on and off various routines using the switch call.\(^\text{184}\) Although gen~ has a switch function, it is limited to two possible states, whereas Java leaves the option for many different states. The advantage of the switch call is that it will turn on the routine that is called ‘only’, turning off all others, saving computational resources for other tasks. The implementation in Java includes one audio inlet for synchronising the trajectory generators with the FFT frame, and has three audio outputs that conform to a stream of Cartesian spatial coordinates \([x, y, z]\). Other inlets allow the user to modulate the state of

\(^{184}\) See Appendix D for the Java code.
the curve generated, either at control rates or audio rates. Each trajectory is coded with its own block procedure:

```java
    case 16:
        //Random Walk
        {
            for(i = 0; i < o1.length;i++)
            {
                len = (float)(sr / in2[i]);
                if (j == 0)
                {
                    a = d;
                    b = e;
                    c = f;
                    d = (float)Math.random();
                    e = (float)Math.random();
                    f = (float)Math.random();
                    x = (d - a)*(j/len)+a;
                    y = (e - b)*(j/len)+b;
                    z = (f - c)*(j/len)+c;
                    j = j + 1;
                }
                else if (j < (int)len)
                {
                    x = (d - a)*(j/len)+a;
                    y = (e - b)*(j/len)+b;
                    z = (f - c)*(j/len)+c;
                    j = j + 1;
                }
                else
                {
                    j = 0;
                    x = (d - a)*(j/len)+a;
                    y = (e - b)*(j/len)+b;
                    z = (f - c)*(j/len)+c;
                }
                o1[i] = x*(float)2-(float)1;
                o2[i] = y*(float)2-(float)1;
                o3[i] = z*(float)2-(float)1;
            }
            break;
        }
```

### 4.3.2.1 FFT Synchronous Trajectories

FFT synchronous trajectory curves are those that remain in formation from one FFT frame to the next, and create steady-state or static sound shapes. These trajectories are derived using the FFT frame index as a phasor (refer to Section 3.1.3). Consequently, the trajectory signals have a derived frequency as calculated using
Equation 14, where $f$ is the frequency, $SR$ the sampling rate and $W$ the FFT frame size in samples (and see Table 11):

$$f = \frac{SR}{W}$$

(14)

Table 11

*Frequencies for Synchronous Trajectories for Different FFT Frame Sizes*

<table>
<thead>
<tr>
<th>FFT size</th>
<th>256-sample FFT</th>
<th>512-sample FFT</th>
<th>1024-sample FFT</th>
<th>2048-sample FFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synchronisation frequency at 44.1 KHz sampling rate</td>
<td>172.265625 Hz</td>
<td>86.1328125 Hz</td>
<td>43.06640625 Hz</td>
<td>21.533203125 Hz</td>
</tr>
</tbody>
</table>

Periodic curves follow repeated paths. The geometry expressed by these shapes can be any periodic curve, whether continuous or discontinuous in the time domain. This research project has explored geometric archetypes such as lines, circles, squares and other polygons, spirals and a variety of curves such as the rose curve and hypotrochoids. Some two-dimensional functions have been derived from audio oscillators such as sawtooth, square and pulse waves that are set to a fixed fundamental frequency, and where one signal of the two is displaced in phase. In scientific terms this kind of plot is described as a pseudo-phase–space plot.

The colour coding of this trajectory always remains the same from one FFT frame to the next. If the colours line up to a similar point in space across successive FFT frames, it means the *sound shape* will exhibit a non-evolving but highly correlated sound shape. The *sound shapes* derived from this specific class of trajectory do not exhibit any perceived spatial motion; rather they describe a static and motionless distribution of frequency across space.
4.3.2.2 FFT Asynchronous Trajectories

FFT asynchronous trajectories are time domain geometries that are not synchronised to the frequency domain; this is ultimately a question of the continuity of individual FFT bins from frame to frame of the FFT. There are many different instances that give rise to the exploration of this discontinuity. Such trajectories can be classed in two categories: low- and high-frequency asynchronous trajectories. Table 12 outlines several different trajectories and their classification according to low- and high-frequency asynchronicity.

Table 12

<table>
<thead>
<tr>
<th>Low-frequency asynchronous trajectories</th>
<th>High-frequency asynchronous trajectories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random walks</td>
<td>Noise generators (white/pink noise, etc.)</td>
</tr>
<tr>
<td>Low-frequency oscillators</td>
<td>High-frequency oscillators</td>
</tr>
<tr>
<td>Continuous differential equations</td>
<td>Iterative function systems</td>
</tr>
<tr>
<td>Asynchronous vector-based process</td>
<td></td>
</tr>
</tbody>
</table>

What is common to all of these curves is that FFT bins coincide differently along the path of the trajectory over successive frames. We can observe this in Figure 93 for both a continuous differential equation and white noise: in both, there is no repeating cycle in the way FFT bins align.

As a consequence of this asynchronicity, these trajectories result in spatial texture. Spatial texture as described by Smalley (1997, p. 124), is concerned

---

185 One of the primary factors enabling me to explore WTS was the fact that it was driven at audio rate. Almost all other spectral spatialisation techniques have focused on event-based systems in combination with smoothing functions like vectral~, so a major difference here was the need to explore systems that are synchronous and do not require these means. If anything, vectral~ and its rampsmooth function would smooth over the precise audio rate fluctuations generated through low- and high-frequency modulation and transformation. As a consequence of this it was necessary to consider, if there was going to be
primarily with how the spatial perspective is revealed through time. Smalley divides
spatial texture into two categories: non-contiguous space (which is revealed when
spectromorphologies are presented in different spatial locations such that two
successive events are not considered near neighbours), and contiguous space (which is
subdivided into spread settings and trajectories).\textsuperscript{186}

Both low- and high-frequency asynchronous trajectories exhibit spatial texture,
but this manifests in different ways. At slow rates, spatial texture unfolds at a rate that
is perceptible. In this way the intrinsic spectromorphologies and their associated
spatiomorphologies unfold very clearly in sequence. Once the frequency of these
asynchronous trajectories shifts predominantly into the audible frequency spectrum, the
spatial texture tends to sound more like a pulsation for fast periodic signals, or an

\textsuperscript{186} Smalley does state that the differences between contiguous and non-contiguous spaces are not definite and that a space can be non-contiguously erratic at a low level but contiguous at a higher level.
interweaving and fluttering for non-periodic signals reminiscent of the signal
decorrelation techniques explored by Kendall (1995) and the audio rate pulsations
evident in works such as Karlheinz Stockhausen’s *Sirius* (1975–77). Such fast rate
asynchronous trajectories also tend to be effective in creating more immersive sound
shapes.

A special case arises for periodic trajectories if one tunes the frequency of this
trajectory orbit to the frequency of the FFT divided by an integer multiple, for example,
1/5, 1/8 or 1/40. This means that one cycle of this periodic orbit will stretch across
several FFT frames resulting in a repeated cyclical spatial texture. In this way it is
possible to create longer repeating textures that last 5, 8, 40 or any number of successive
FFT frames. However, most other asynchronous trajectories result in non-repeating
structures, and therefore non-repeating and ever-evolving spatial texture.

4.3.2.3 Space-filling Properties

Another feature of the trajectory that drastically influences the creation of sound
shapes is its space-filling property. This property relates to the extent of the terrain that
is revealed by the trajectory, as the values that are read from the terrain depend on the
trajectory’s path of orbit. Some trajectories may only cover a small fraction of a two-
dimensional terrain curve, and consequently do not reveal the extent of topographical
features the terrain holds. This issue was outlined in the context of dynamical WTS:

*WT synthesis* technique has been largely restricted to the world of simple
oscillator types due to the way in which trajectory orbits have been derived; a
situation that—on the surface—seems to render the possibility of using a terrain
data array impractical. For example, a small trichromatic 24-bit image file of
dimensions 320 × 240, requiring 230,400 bytes of memory, seems far too large a
wavetable if the result is much like a simple oscillator. This is largely
problematic when the trajectory orbit is periodic, as only a small percentage of
the terrain data is accessed for the resulting sound. (James, 2003)

The most effective audio signals used as space-filling curves tended to be white
noise. Most other curves, with the exception of those that are generated as a set of
independent curves, tended to adhere to a local region. Other curves may at times exhibit space-filling properties but unfold very slowly, as in the case of random walks and Brownian motion. However, other kinds of space-filling curves such as the Peano, Moore, Hilbert and Lebesgue curves will fill up spaces quickly and efficiently, and are designed specifically to evenly fill the space the curve occupies.

Other curves with potential space-filling characteristics, although this is dependent on variables, include spirals, the rose curve, hypotrochoids, iterative functions, and stochastic and random processes.

**4.3.3 Sound Shapes**

The following two sections are aimed at illustrating the performative approaches to creating *sound shapes* using both Models A and B. These are discussed separately as the sound shapes generated by both models are different, and require their own specific evaluation in terms of the resulting frequency–amplitude curves and psychoacoustic attributes. Each section includes a categorical table of *sound shapes*.

**4.3.3.1 Model A**

When evaluating the *sound shapes* created through the performative control of WTS, it is important to discuss outcomes that are most intuitive, in order to establish predictable results. For *Model A*, the most linear and predictable scenario is when using a terrain function that describes the angles of azimuth around a centre point (Equation 15 or 16). In this situation the laptop performer can specify any trajectory, and its movement will correspond with azimuth cues that follow the trajectory path synonymously around a centre point. In this way the height of the terrain \( z \) is proportional to the angle of displacement around the central point or origin:

\[
f(r,\theta) = \theta \quad \text{(15)}
\]

\[
f(x,y) = \tan^{-1}\left(\frac{2y-1}{2x-1}\right) \quad \text{(16)}
\]
We normalise this function to ensure that the height of the terrain \((z)\) is within the range \(0–1\):

\[
f(x,y) = \frac{\tan^{-1}\left(\frac{2y-1}{2x-1}\right)}{2\pi} + 0.5
\]  

(17)

We can generate a 3D plot of this terrain function, and we can see that the hue across the surface is synonymous with the colour wheel referred to in Chapters 1 and 3.

*Figure 94a*. A 3D plot of a polar curve with colour coding representing the height of the terrain, which is subsequently mapped to azimuth.

*Figure 94b*. A density plot of the polar curve where colour represents the height of the terrain; curve-specific colour degrees are associated with specific points of azimuth.

If we take a trajectory that consists of a constant or stationary point—what we might describe as point source—in this instance this point will reproduce the full spectrum of a sound in one location in the sound scene, as shown in Figure 95.
Extending this point into a line, what would be a trajectory that traverses the line at the frequency of the FFT frame, we end up with a spread of frequency dependent on the highest and lowest angles of azimuth read from the terrain. The spectral bin alignment would depend on which bin lands on what specific angle of azimuth. The trajectory traverses the line only in one direction. An example is shown in Figure 96.
When we extend this to a bidirectional line—that is, a line that the trajectory traverses in one direction and then the other direction—we end up folding the frequency bands in the resulting sound shape. The range of azimuth is still determined, as before, by the maximum and minimum azimuth cues that are read from the terrain. We would have a similar ‘foldover’ resulting from an elliptical orbit if it stayed on one side of the terrain and did not cross the centre point, as an ellipse has a forwards and backwards motion. The resulting spectral distribution reflects symmetry across the frequency range shown in Figure 97.

By extending the trajectory out to a circle, the bidirectional fold flips over to the other side of the sound scene. If the origin of the circle is exactly aligned with the centre of the terrain, the frequency distribution will be an even spread across the speaker configuration. If the circle were slightly off centre, the frequency distribution would distort—that is, will become slightly non-linear. This is shown in Figure 98.
Mathematically it makes sense that the circle generates a linear frequency distribution over the four speakers, because if we substitute the parametric functions into the Cartesian terrain function, we end up with a linear function:

\[
f(x) = \frac{\sin(2\pi x) + 1}{2}
\]

\[
f(y) = \frac{\cos(2\pi y) + 1}{2}
\]

\[
\tan^{-1}\left(\frac{2\left(\frac{\cos(2\pi t) + 1}{2}\right) - 1}{2\left(\frac{\sin(2\pi t) + 1}{2}\right) - 1}\right) + 1
\]

This can be reduced or simplified to the linear function:

\[
f(t) = t + 0.5
\]
When we consider the archetypes that result in a linear spectral distribution, the terrain function is always a structure that is the mathematical antithesis of the trajectory’s topography. A circular trajectory traversing over a polar ramp function effectively cancel each other out, resulting in a linear function. The ramp function with a linear trajectory is of course another linear distribution, as is any function that is the ‘inverse’ of the other, such as a circular sinusoidal trajectory over an arcsin terrain function. The non-linear and it’s inverse properties in these cases cancel each other out,
resulting in linearity. However, again the terrain used in these examples above is additionally intuitive given its association for the performer with the angle of azimuth. Notice how in the following example we can no longer assume that the position of the FFT bin in the trajectory will correlate with the same position in the frequency–space plot.

**Figure 99a.** A linear trajectory superimposed over a density plot of a ramp terrain curve.

**Figure 99b.** The linear trajectory colour coded to show the alignment of spectral bins.

**Figure 99c.** The sound shape created by the terrain and trajectory in Figure 99a.

**Figure 99d.** The sound shape shown as frequency–amplitude curves for each speaker. This is the same as Figure 98d.

If we slightly contort this terrain so that we bend it out of its linear shape, we skew the allocation of frequency bins to points of azimuth.
Returning to a circular trajectory in Figure 101, as soon as we use a different terrain—that is, any terrain surface other than Figure 94a—we introduce further complexity and non-linearity into the system. A ramp function will introduce one fold across the entire speaker system as we observed earlier. If the terrain also exhibits undulation, the resulting frequency–space distribution will exhibit further folds. This is illustrated in Figure 101.
Figure 101a. A circular trajectory superimposed over a density plot of a more complex terrain curve.

Figure 101b. The circular trajectory colour coded to show the alignment of spectral bins.

Figure 101c. The sound shape created by the circular trajectory in Figure 101a reading a series of azimuth values from the terrain. This generates a series of folds in the spectrum.

Figure 101d. The sound shape in Figure 101c shown as frequency–amplitude curves for each speaker. This shows a series of symmetrical folds throughout the frequency spectrum.

It is clear that with more undulations the folds increase, but it is important to note that the irregular nature of undulations in the terrain and/or trajectory will result in a much more complex and non-linear frequency–space distribution. A more harmonically complex terrain structure with a regular periodic trajectory is illustrated in Figure 102.
Figure 102a. A circular trajectory superimposed over a density plot of a terrain with increasingly more undulations.

Figure 102b. The circular trajectory colour coded to show the alignment of spectral bins.

Figure 102c. The sound shape created by the circular trajectory in Figure 102a reading a series of azimuth values from the terrain pictured. This generates increasingly more folds in the spectrum than compared to Figure 101c.

Figure 102d. The sound shape in Figure 102c shown as frequency–amplitude curves for each speaker. This shows an increasingly larger number of symmetrical folds throughout the frequency spectrum compared to Figure 101d.

It is important to note here that although the continuous lines shown in the frequency–space plot are synonymous with the nature of a coherent sound shape, in reality we are looking at a discrete set of points that describe this shape as illustrated in Figure 103. The frequency–space plots from this point will express these shapes in terms of points. Although the sound shape shown in Figure 102c may be perceived as a fused and coherent image, it is in fact a rendering of a series of discrete points shown in Figure 103.
These *sound shapes* exhibit more and more folds, as the frequency–space plot heads towards the limit of potential folds within the discrete point size of the FFT frame. After this point, the results exhibit a form of aliasing of the frequency plot at which point we start to arrive at discontinuous distributions of spectra as can be seen when using photographs, which tend to have noisy components as illustrated in Figure 104.

We also find discontinuous frequency–space plots when using terrain surfaces that are not continuous, but step-wise as illustrated in Figure 105. This may show whether or not adjacent bins are ‘connected’ or ‘disconnected’ across the listener field, a notion that stems back to Smalley’s (1997) concept of *non-contiguous space*. This synonymously depends on the continuous or discontinuous nature of the terrain and trajectory functions used.
All the examples discussed so far have used periodic trajectory curves, and consequently these have all demonstrated static steady-state sound shapes. If we begin to introduce some of the other categories of trajectory such as those that are asynchronous, this introduces a time-evolving aspect to the sound shapes. Asynchronicity provides a means of a shift from FFT frame to the next, of one frequency–space distribution to another. Many of these kinds of trajectories give rise to spatial texture.
A low-frequency random walk would give rise to shifting sound spectra as illustrated in Figure 106. The relatively slow movement of the trajectory ensures that momentary states and shifts are clearly delineated. This kind of result is similar for all LFO (low-frequency oscillation)-based signals, which would either present a repeating quality if they are periodic, or a continually evolving quality if the signals are generated using a random process.

High-frequency asynchronous trajectories exhibit different kinds of characteristics. For example white noise has a number of specific properties that makes
it successful not only because of its non-repeating and random qualities, but also because of its space-filling qualities as can be seen in Figure 107a. Points are never the same from FFT frame to frame. Instead, each frame consists of a newly randomised frequency distribution, and over several frames, this is effective in creating immersive sound shapes. In Figure 107b we see the spectral distribution for a single FFT frame, followed by the points plotted for 10 consecutive FFT frames in Figure 107c, which are generated within the space of a fraction of a second.

Figure 107 illustrates that for a random distribution we create a circumspectral effect over a series of FFT frames. This could be described as contiguous spatial texture. Rather than a perception of the steady-state or instantaneous frequency–space distribution, the brain tends to assimilate these instantaneous and successive FFT frames into a single continuous sound shape. For the random trajectory distribution shown in Figure 107a, the associated spectral distribution in Figure 107b and 107c shows a relatively even spread of the frequency spectrum across all points of azimuth around the speaker configuration. The resulting perceived effect is that the sound scene has no specific sense of directivity, but rather the separate random streams of sound spectra merge to form a single spatial sound image across the entire speaker array. The speed and multiplicity of movements alert the listener to movement within this image, but discerning the precise movement of individual spectra for the listener is extremely difficult.
Despite using an effective space-filling curve such as white noise, we can still see that terrains that exhibit discrete and discontinuities surfaces produce isolated zones of sound, or non-contiguous spatial texture, across the listener area (shown in Figure 108), whereas continuous terrain curves tend to exhibit sound shapes with wider angles of azimuth, providing a more circumspectral, and contiguous, experience.
The resulting shapes of Model A can be categorically separated into single point source, multiple point source and spectral diffusion with various numbers of spectral folds, spatial width, and low- and high-frequency spatial texture. These categories are documented in Table 13, which shows how these are generated.
### Table 13

**Model A Sound Shapes**

<table>
<thead>
<tr>
<th>Sound shape</th>
<th>Generation method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single point source (traditional panning)</td>
<td>Any constant trajectory (fixed point) over any continuous or discontinuous terrain surface</td>
</tr>
<tr>
<td></td>
<td>Any continuous or discontinuous trajectory curve over a flat (topographically constant) terrain surface</td>
</tr>
<tr>
<td>Multiple point source (diffusion of bands of frequency as Normandeau explored)</td>
<td>Any stepped and discontinuous trajectory curve (within the time of an FFT frame) over any continuous terrain curve. The number of distinct bands is dependent on the number of discontinuities within the time of the FFT frame.</td>
</tr>
<tr>
<td></td>
<td>Any continuous trajectory curve over a discontinuous and stepped terrain curve. The number of distinct bands is dependent in this case on the number of discontinuities in the terrain.</td>
</tr>
<tr>
<td></td>
<td>Any discontinuous trajectory over a discontinuous terrain curve. The number of distinct bands is dependent on the sum of the discontinuities in both the terrain and trajectory.</td>
</tr>
<tr>
<td>Spectral diffusion (no folds)</td>
<td>Unidirectional trajectory curve (within the time of an FFT frame) over a unidirectional terrain curve (i.e. linear trajectory over linear ramp terrain curve)</td>
</tr>
<tr>
<td>Spectral diffusion (1 fold)</td>
<td>Bidirectional trajectory over a unidirectional terrain curve</td>
</tr>
</tbody>
</table>
Unidirectional trajectory over a bidirectional terrain curve

Spectral diffusion (multiple folds)

When a unidirectional trajectories path of orbit causes more than one oscillation OR
When a trajectory of more than one oscillation orbits a unidirectional terrain surface OR
When both trajectory and terrain curves have more than one oscillation within the time of an FFT frame

Spatial width

The width of the sound shape is determined by the highest and lowest points of traversal by the trajectory over the terrain surface. This determines whether the sound shape is reproduced across the entire speaker configuration or a smaller subset of this. The potential of spatial width of a given terrain surface is dependent on the global maxima and minima of the surface. In Model A by ‘scaling’ the terrain surface one can adjust the spatial width. Similarly tilting a terrain surface about the centre may also result in a spectral divergence.

Low-frequency spatial texture

Produced with asynchronous trajectories that exhibit oscillation or rates of change generally below the audible frequency range. These changes occur slowly enough for the listener to discern the unfolding of sound shapes. The result is a kind of animation of a sequence of spectral sound shapes. The temporal unfolding of the trajectory is sonified in an abstracted way.
High-frequency spatial texture

Produced by asynchronous trajectories that are well within the audible frequency range. Such trajectories explored include noise oscillators, particularly white noise, iterative function systems and other chaotic systems, oscillators, and other sound sources. As the distribution of spectral bands is much faster, psychoacoustically these changes tend to fuse into a single image. The implications of these strategies are effective in creating decorrelated distributions of spectra similar, in a way, to the signal decorrelation techniques described by Kendall (1995). These are effective in creating immersive sound shapes. White noise statistically favours all possible outcomes within a domain range and therefore is effective in generating even and immersive distributions of frequency; however these spectral distributions can be biased to favour certain frequencies depending on the trajectory used.

All the sound shapes described so far result in an equal-weighted distribution of spectra across the speaker configuration. This is because the model only maps the height of the terrain $z$ with a point of azimuth. As this process only determines changes in azimuth, ideally the reproduction of different frequencies in terms of their loudness would be the same across the system.\(^{187}\) Although the process itself is not applying a spectromorphology it is subject to the non-linearities of the listening experience.

Speaker angle, placement, listener position, room acoustics, and the convolutions of the pinnae, head and shoulders will all contribute to a spectromorphology of the perceived sound shape based on the directional properties of individual spectra. One can also appreciate that the listener’s own experience, in terms of the directivity of individual spectra, will be enhanced by the spatial resolution of the speaker system used. For more

\(^{187}\) This is stated in reference to a global perception of the system; however this cannot possibly take into consideration other variable non-linearities such as speaker frequency response, room acoustics, the Haas effect, air absorption and the further non-linearities of the pinna, head and shoulders.
complex distributions, the higher the spatial resolution of the speaker system, the more convincing and detailed the *sound shape*.

As mentioned in Chapter 3, there is a way to introduce *spectromorphology* by narrowing or widening the window function used for each speaker in the algorithmic spatialisation process. A window size that is double, or three or four times the length will only result in an increased sense of spatial width, but window sizes that are fractional lengths between, or that are smaller than the window matched to the given number of loudspeakers will result in non-linear-weighted distributions. One example of this is shown in Figure 109 and more examples that more extensively demonstrate this are included in Appendix D.

4.3.3.2 Model B

While *Model A* exhibits some inherent non-linearities in the way *sound shapes* are generated through the combinations of distinct terrain and trajectory contours, *Model B* presents a very intuitive approach to this interaction. *Sound shapes* generated using *Model B* are made apparent through the revealing if a frequency–space
distribution by the path of the trajectory contour. Figure 110a and 110b show two plots of trajectories that reveal colour coding that is determined by the terrain curve (see Figure 64a). Some other examples of this were shown in Figure 64b and 64c.

One can see that the detail of the frequency–space distribution is dependent on the space-filling properties of the trajectory curve, as it is this curve that determines what aspects of this spectral distribution are revealed, and made audible. *Model B* makes a distinct connection between distance and azimuth (and also elevation cues in 3D implementations) with respect to frequency bands. This is dependent on the temporal evolution of the trajectory that describes distance, azimuth and, when necessary, elevation cues. The colour retrieved from the lookup table corresponds with the associated frequency. *Model B* ultimately assigns a frequency band with a spatial coordinate. The plots in Figure 111 show how with asynchronous trajectories the spectral distribution can be revealed over time.
In the case of a random distribution like white noise, this trajectory is asynchronous to the FFT frame, and will reveal different aspects of this frequency–space distribution from FFT frame to frame. There are some instances where spectral interpolation across several FFT frames may be effective, using the \textit{vectral} object; however this also sacrifices the detail of \textit{spatial texture} that results between consecutive frames.

Unlike the linear frequency–amplitude curves seen with \textit{Model A} (in Figure 109a), the frequency–amplitude curves associated with \textit{Model B} are rarely linear as shown in Figure 112. These curves often feature strong energy on certain bands of frequency, which is dependent on the rate of change of the trajectory curve. In other words, stationary points in the \textit{terrain} or \textit{trajectory} are the reason for this accumulation of energy in certain regions of the frequency spectrum.
**Model B** invariably reproduces considerable timbral variation depending on a number of factors including the terrain contour and the non-linearities of the panning algorithms used. The addition of distance encoding, as well as directivity and spatial blur in AEP and DBAP panning, introduces a number of non-linearities in the way the relative loudness of different spectral bands are reproduced in such a model. All of the following documentation describes the creation of sound shapes using AEP, as the concepts covered apply easily to DBAP too.
Distance in AEP is determined by three parameters: centre size, centre attenuation (dB) and distance attenuation (dB). Some distance curves with various parameter settings are shown in Figure 113.

Calibration of the distance as applied to timbre spatialisation can be achieved using the combination of a white noise trajectory over a simple linear terrain function. Figure 114a shows the standard frequency–space visualisation used for Model B, and Figure 114b shows the ideal position of a listener (centre), where the distance of low frequencies highlighted in red are more distant than the mid-range frequencies in green that should sound perceptively louder. The high frequencies in blue again are some distance from the listener.

By reading the resulting frequency–amplitude curves from this process, it is possible to determine how much frequencies that are further away from the centre point are attenuated as a result of their relative distance from the listener, as shown in Figure 115. These frequency–amplitude curves can be used to calibrate the distance rolloff...
curve and centre size of AEP. The combined use of the centroid smoothing and a linear-phase low-pass filter can also help to smooth out the peaks in the SPF in order to gauge better the rolloff in each instance. These smoothed frequency–amplitude plots are shown in Figures 115d, e and f. With a centre size of one and a rolloff of 3 dB, the impression of distance is subtle but evident. The use of the low-pass filter can also remove the comb filtering effects of the SPFs that result from computing the histogram.

![Figure 115a](image1.png)  
**Figure 115a.** A frequency–amplitude plot over 10 FFT frames using Model B with Ambisonic equivalent panning (AEP) centre size 0.05, centre attenuation 0 dB and distance attenuation 0.2 dB (spectral centroid width set to 0.9).

![Figure 115b](image2.png)  
**Figure 115b.** A frequency–amplitude plot over 10 FFT frames using Model B with AEP centre size 0.05, centre attenuation 0 dB and distance attenuation 3 dB (spectral centroid width set to 0.9).

![Figure 115c](image3.png)  
**Figure 115c.** A frequency–amplitude plot over 10 FFT frames using Model B with AEP centre size 0.05, centre attenuation 0 dB and distance attenuation 6 dB (spectral centroid width set to 0.9).

![Figure 115d](image4.png)  
**Figure 115d.** The frequency–amplitude plot in Figure 115a with a linear-phase spectral low-pass filter applied.

![Figure 115e](image5.png)  
**Figure 115e.** The frequency–amplitude plot in Figure 115b with a linear-phase spectral low-pass filter applied.

![Figure 115f](image6.png)  
**Figure 115f.** The frequency–amplitude plot in Figure 115c with a linear-phase spectral low-pass filter applied.

The *sound shapes* that emerge through this process have some similarities in classification, yet the kinds of frequency–space distributions are completely different from those in *Model A*. For example, for a trajectory that is a constant point, rather than
all frequencies being fixed to a point of azimuth, *Model B* refers to a specific frequency and a spatial location as determined by the position of the point in relation to an ideal listening position—that is, centre of the spatial field. All of the sound shapes described here assume a basic four-channel speaker arrangement, but they can be easily adapted for more complex arrangements. For the purposes of this section where I wish to highlight the emergence of sound shapes through the interaction of *terrain* and *trajectory* curves, I will adopt the use of a compact diagram showing the *terrain*, *trajectory*, listener position and speaker positions as shown in Figure 116a. *Model B* is the only model where we can simplify this diagrammatically. This is also for the purpose of space.

For a single point we reproduce only a single band of frequency as shown in Figure 116. The perceived direction of this point of frequency is determined by the position of the trajectory, such that in the diagram this appears to come from just to the left and behind the listener. Note the accumulation of energy that occurs here, such that the amplitude ratio of this frequency is over 30 times, correlating with an increase in level of approximately 30 dB.

*Figure 116a.* A trajectory at a constant position of (0.35, 0.35).

*Figure 116b.* The resulting frequency–amplitude for four speakers that accumulate over one frequency band.
If the trajectory is spread across the listener field in a circular fashion such that the radius is equidistant about the centre listener position, then frequencies read at their respective lookup positions appear to come from all around the listener as shown in Figure 117a. Note the range of frequencies reproduced is also affected by the highest and lowest undulations of the terrain curve. So, for example, if the terrain is completely flat, again only a single frequency band is reproduced. If the trajectory passes revolves around the central listener position, this single frequency is reproduced in all directions about the listener at equal weight as shown in Figure 117b.

![Figure 117a](image1.png) A circular trajectory passing over a terrain where frequencies (shown in colour) are distributed spatially.

![Figure 117b](image2.png) A circular trajectory passing over a terrain that consists of only one frequency (in yellow).

The distance correlation is constant in Figure 117a, and as the terrain also includes a different value of frequency about this circular path, we end up with a constant distribution of frequencies across the entire frequency spectrum after applying this terrain and trajectory to Model B. We notice here that there are four bands of frequency discerned by the speaker with which they coincide. In this case the Ambisonic equivalent panner determines the relative weights of sound components to
their respective speaker. We can see in Figure 118 that the colour coding of the four speaker channels correlates too.

A common problem with using histograms is that the resultant graph is not continuous, but disjointed, whereas the original terrain and trajectory curves used for deriving SPFs are smooth and continuous. This has the effect of the SPF sounding comb filtered. The use of centroid smoothing and a linear-phase low-pass filter assists in smoothing out these irregularities in order to arrive at a result that timbrally sounds more natural in relation to the original source sound used for timbre spatialisation.

Although the weights of the original curve sit at unity gain, or 0 dBFS, after the smoothing process (spectral centroid smoothing and linear-phase filtration) the bands shift in level to a generalised weighting of four or an increase of +12 dB. In the scheme of things this is substantial, so the smoothing algorithms adopt an auto-normalise option that recalibrates automatically for this large level of difference. Incidentally this is the one and only instance where both models create an equivalent kind of sound shape.

![Figure 118a. The spectral processing functions (SPFs) for four speakers after computing the histogram for the terrain and trajectory in Figure 117a.](image1)

![Figure 118b. The SPFs in Figure 118a after spectral centroid smoothing and linear-phase filtration have been applied. This distribution shows similarities to Figure 98d.](image2)
Figure 118c. The frequency–amplitude plots in Figure 118b summed together showing an amplitude ratio of four or an increase of 12 dB.

Figure 119a. A vertically symmetrical terrain curve, with a vertically and horizontally symmetrical trajectory, and a vertically and horizontally symmetrical speaker configuration.

Figure 119b. Sonograms of the terrain and trajectory in Figure 119a through Model B showing a steady-state and unchanging sound shape over four speaker channels. The frequency spectrum is duplicated in adjacent speakers.

Figure 119c. The frequency–amplitude curves for all four speakers after spectral centroid smoothing and linear-phase filtration have been applied.

Figure 119d. The sum of all frequency–amplitude curves in Figure 119c showing the intensity of various bands of frequency.
A very unique outcome arises for Model B when the speaker arrangement, the terrain and the trajectory curve are symmetrical about the vertical or horizontal axes, resulting in the same SPF being produced in multiple speakers. As soon as there is asymmetry in one of these structures, all speakers will reproduce different SPF functions. We can observe this outcome in Figure 119 with a square trajectory placed over a linear ramp terrain curve. This produces two duplicates of SPFs across the quadraphonic speaker configuration. The spectrum is also symmetrical above and below a centre frequency.

Figure 120a. A vertically symmetrical terrain curve, and vertically and horizontally symmetrical trajectory.

Figure 120b. Sonograms of the terrain and trajectory in Figure 120a through Model B showing spectral symmetry.

Figure 120c. The frequency–amplitude curves for all four speakers after spectral centroid smoothing and linear-phase filtration have been applied.

Figure 120d. The sum of all frequency–amplitude curves in Figure 120c showing the intensity of various bands of frequency.
The same symmetry applies if the square were in a different orientation. In Figure 120 the vertices of the square cause some stationary points resulting in an increase in energy on certain frequencies in the SPFs generated.

However, this symmetrical relationship is broken as soon as one of these elements is asymmetrical. In these situations the spectrum changes in all speakers, nullifying any duplication. In Figure 121 the SPFs for all speakers are different.

**Figure 121a.** A vertically symmetrical terrain curve, with a vertically and horizontally asymmetrical trajectory, and a vertically and horizontally symmetrical speaker configuration.

**Figure 121b.** Sonograms of the terrain and trajectory in Figure 121a through Model B showing a different spectrum in all four speakers.

**Figure 121c.** The frequency–amplitude curves for all four speakers after spectral centroid smoothing and linear-phase filtration have been applied.

**Figure 121d.** The sum of all frequency–amplitude curves in Figure 121c showing the intensity of various bands of frequency.

This scenario does not apply to terrain surfaces and/or trajectories that are not symmetrical over the horizontal or vertical axes. Sound shapes generated by non-
symmetrical relationships result in all speakers having vastly different timbres as shown in Figure 122.

**Figure 122a.** An asymmetrical and nonlinear terrain curve, with a vertically and horizontally asymmetrical trajectory, and a vertically and horizontally symmetrical speaker configuration.

**Figure 122b.** The frequency–amplitude curves of the terrain and trajectory in Figure 122a through Model B showing a different spectrum in all four speakers. These spectral processing functions have had spectral centroid smoothing and linear-phase filtration applied.

**Figure 122c.** The sum of all frequency–amplitude curves in Figure 122b showing the relative intensity of various bands of frequency.

Most of the trajectories discussed up to this point generally have poor space-filling properties. Noisier signals increase the potential for describing a sound shape in more detail. This is also compromised by the number of values that can be accounted for within the space of an FFT frame. Figure 123 shows a high-frequency asymmetrical trajectory used in Model B over the same terrain curve as seen in Figure 122a, resulting
in a much more detailed series of SPFs that take into account more distance information across the resulting sound shape.

**Figure 123a.** A noisy high-frequency asynchronous trajectory passed over a non-linear terrain curve.

**Figure 123b.** The frequency–amplitude curves of the terrain and trajectory in Figure 123a through Model B showing a different spectrum in all four speakers. These spectral processing functions have had spectral centroid smoothing and linear-phase filtration applied.

**Figure 123c.** The sum of all frequency–amplitude curves in Figure 123b showing the relative intensity of various bands of frequency.

The spatial resolution of these sound shapes can increase drastically with larger numbers of loudspeakers. In Figures 125, we see the same contour distributed between 1, 2, 8 and 32 speakers. The higher the number of loudspeakers, the more separation is applied to those spectral bands assigned to particular speakers. This enables the frequency response curves to represent the states ‘in between’. As the number of
speakers increases we observe increasing detail in each subsequent area of the spatial field determined by the set of SPF functions.

<table>
<thead>
<tr>
<th>Figure 125a</th>
<th>Figure 125b</th>
<th>Figure 125c</th>
<th>Figure 125d</th>
</tr>
</thead>
<tbody>
<tr>
<td>A frequency–amplitude curve applied to one loudspeaker.</td>
<td>A frequency–amplitude curve applied to two loudspeakers.</td>
<td>A frequency–amplitude curve applied to eight loudspeakers.</td>
<td>A frequency–amplitude curve applied to 32 loudspeakers.</td>
</tr>
</tbody>
</table>

In general the sound shapes created by Model B (see Table 14) fall into categories of frequency-specific shapes or more broadband spectral shapes. Notions of spatial perception such as spatial direction, distance, width, depth, height and immersion are relevant in their classification. For asynchronous trajectories spatial texture is still a relevant perceptible parameter as well.
Table 14

*Model B Sound Shapes*

<table>
<thead>
<tr>
<th>Sound shape</th>
<th>Generation method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single frequency</td>
<td>Any constant trajectory (fixed point) over any continuous or discontinuous terrain surface. This generates perceptively a single directional cue. Any continuous or discontinuous trajectory curve over a flat (topographically constant) terrain surface. This creates a spread of a single frequency according to the spatial movement of the trajectory.</td>
</tr>
<tr>
<td>Multiple single frequencies</td>
<td>Any stepped and discontinuous trajectory curve (within the time of an FFT frame) over any continuous terrain curve. The number of distinct bands is dependent on the number of discontinuities within the time of the FFT frame. This results in specific directional points of frequency. Any continuous trajectory curve over a discontinuous and stepped terrain curve. The number of distinct bands is dependent in this case on the number of discontinuities in the terrain. This results in a spread of different frequencies dependent on the spatial movement of the trajectory. Any discontinuous trajectory over a discontinuous terrain curve. The number of distinct bands is dependent on the sum of the discontinuities in both the terrain and trajectory. This may produce both directional points of frequency as well as spread frequency in the sound shape.</td>
</tr>
<tr>
<td>Spectral quality</td>
<td>The terrain surface determines the spectral qualities of <em>timbre spatialisation</em>. The topographical nature of this curve describes where certain frequency components of a sound will be located within a sound scene. As a terrain surface consists of many hundreds of thousands of points, a trajectory is used to highlight specific frequencies along its path. These frequencies are also associated with their spatial coordinate in the sound scene. The undulating nature of a terrain surface or the discontinuous nature of this spatial sound shape is sonified.</td>
</tr>
<tr>
<td>Spectral width</td>
<td>The global maxima and minima of the terrain surface determine the maximum and minimum frequencies that are reproduced in a specific <em>sound shape</em>.</td>
</tr>
<tr>
<td>Spatial direction</td>
<td>The trajectory determines the spatial direction of specific frequency components.</td>
</tr>
<tr>
<td>Spatial width</td>
<td>The spatial width of the sound scene and sound shape is determined by the geometric width of the trajectory signal.</td>
</tr>
</tbody>
</table>
Spatial depth

The spatial depth of the sound scene is determined by how many concurrent layers of spatial cues are highlighted. Trajectories that consist of a single continuous path often have a small amount of spatial depth, whereas space-filling curves or discontinuous trajectories may introduce several spatial layers at varying distance that will enhance the impression of depth in the sound scene.

Spatial height

For systems where elevated speakers are used, the same principle of spatial width can be applied to the elevated dimension. For trajectories that incorporate movement vertically will also impart spatial height in the rendering of sound shapes.

Immersion

Immersion is most prevalent of trajectories that spatially surround the listener, and evoke all of the various cues above, including spatial width, depth and height. Immersion can also be enhanced through the use of movement and spatial texture.

Spatial texture

In terms of spatial texture, Model B really only works with high-frequency asynchronous trajectory signals such as noise generators (particularly white noise), iterative function systems and other chaotic systems, oscillators, and other sound sources. The asynchronous interaction of these signals with the FFT process result in micro-movements in the resulting spectral and spatial image, giving it a greater sense of spatiality.

Pitch shift (Doppler)

Slowly moving trajectories result in a pitch shift for continuous terrain curves. Random walks and low-frequency oscillators are quite effective in these kinds of applications.

4.4 The Morphology of Sound Shapes

So far we have been mostly concerned with the variety of terrain and trajectory curves and the implications these have in generating fixed frequency–space diffusion in the case of synchronous trajectories, and the emergence of spatial texture as a result of asynchronous trajectories. This section in contrast deals with the morphology of these sound shapes through interpolation and transformation processes. In a performative sense, when we consider Models A and B, these processes of interpolation and transformation are applied to the terrain and trajectory structures, and finally result in a smooth morphology of sound shapes through time.
The issue of temporality is an important one. Laptop performers, whether they are reading a score or responding in a more intuitive way to sounds around them, often want to be able to intentionally control the outcome with enough knowledge about the ultimate direction of their musical gesture.

For sound synthesis applications, WTS depends on transformational processes in order to produce time-varying changes in the resulting sound. For control applications, this is necessary for creating evolving and animated sound shapes (James, 2014). Wishart (1996) explains that the use of spatial movement of sound objects provides the definition and transformation of musical landscape. The interpolation of terrain curves allows for a smooth transition from one terrain to the next, and the transformation of these curves gives the performer various options for contortion and evolution of these topographies. This directly influences the spectromorphologies and spatiomorphologies created as a result of these transformations.

Controlling the interpolation of the trajectory gives the performer the ability to control the level of FFT synchronicity or asynchronicity by moving between these states. Similarly, this interpolation can be applied to high-frequency or low-frequency spatial texture. Further transformations can be applied to trajectories in the time domain or frequency domain to directly control both the spectromorphology and spatiomorphology of sound shapes.

4.4.1 Terrain Morphology

The morphology of terrain curves involves two processes: interpolation and transformation. How these result in spectromorphology and spatiomorphology depends on the mappings used in Models A and B. Although these sound shapes are animated, rather than steady state, their derivation still follows the same processes as documented in Sections 4.2.3.1 and 4.2.3.2.
4.4.1.1 Interpolation

Processes of interpolation involve taking a minimum of two possible terrain curves and cross-fading between these. This concept originally arose when using different terrain maps for describing the algorithmic process of different sound synthesis techniques (James, 2005). Cross-fading between several terrain curves allowed for a morphology of sound synthesis. As the models explored in this research project are mapped to the spatialisation of spectral bands, this interpolation can be used for smoothly shifting between different sound shapes. The jit.xfade object in MaxMSP is used to perform a linear cross-fade, and is shown in Figure 126.

![Figure 126](image)

*Figure 126. The use of jit.xfade to morph between two terrain structures used during a 2012 performance at the Make It Up Club in Melbourne.*

A problem that originally emerged during my research in WTS was related to the audio rate lookup process when using motion video as a dynamic terrain curve (James, 2005). Reading from these dynamic terrain curves resulted in an audible artefact due to the instantaneous frame changes in the video signal. A suggested

---

188 As the terrain curves are buffered into a Jitter matrix, this update process between consecutive frames of matrices occurs much more slowly than audio-based processes. One of the issues with extracting audio
workaround was proposed involving interpolation (James, 2005). This process has been re-worked, and some further suggestions are made here for improving its reliability.

A system like this needs to interpolate ‘across’ frames, in a way like the OLA method used in the STFT used to smooth out the segmented results of frequency domain analysis and resynthesis methods. However, unlike the FFT implementations in MaxMSP that are clocked, multi-signals like video are not always guaranteed to arrive on time, or arrive at all! Synchronicity was an ongoing issue here as there is no direct way of ensuring the exact synchronicity and timing of video frames. They tend to fluctuate, and the \textit{jit.fpsgui} object that is responsible for reporting the frames per second shows the amount of variability that can occur. On a Macbook Pro 2.7 GHz Intel Core I7 with 16 GB of 1200 MHz DDR3 memory, the frame count was 49–50 fps (frames per second) from a \textit{jit.grab} object buffering frames at 300 × 300 pixels. With an additionally \textit{jit.pwindow} displaying the video signal, frame rates became substantially more unstable and dropped to 32–42 fps. It is clear that with more complex video processing tasks, and larger arrays and buffering routines, it could be more unstable in an environment like MaxMSP, a relatively high-level programming language, which may compromise processes for the sake of real-time priorities in some cases.

From a robust engine point of view it was decided the timing would be dictated by the audio engine: a phase-driven system. All other processes are clocked around this robust synchronised engine, meaning that read and interpolation events and video frame buffering are all synchronised. Each frame of video is buffered before running into the interpolater, so that only one frame is allowed into the 3D video buffer at a time. The disadvantage here is that video frames may be dropped as a result, although there is a trade-off in that video frames are not stored at multiple times in the same place, which increases the overall efficiency of the system. The system proposed here involves a from a multi-signal, such as using \textit{jit.peek\~} to extract audio from a video signal, is the high potential to create discontinuity artefacts.
circular video buffer consisting of two 3D matrices that each store eight consecutive frames of video as shown in Figure 127a. \(^{189}\) The trajectory, pictured in Figure 127b shows a circular trajectory translated smoothly across eight frames. The cross-fade ramp time in samples is determined by dividing the sampling rate by the number of frames; that is, 10. So each frame is cross-faded in the space of 4410 samples at a sampling rate of 44.1 KHz. Due to real-time performance, latency is a critical factor here. For this reason, and also reasons of efficient use of memory, the size of these 3D matrices was kept minimal. \(^{190}\)

![Figure 127a. A circular video buffer consisting of two 3D matrices 300 × 300 × 8.](image1)

![Figure 127b. The circular trajectory orbit that also slowly passes through the eight frames in Figure 127a.](image2)

The advantage of this process is that sound shapes will evolve smoothly, and still exhibit the spatial texture and temporal detail dictated by the evolution of the trajectory. If the interpolation were introduced at any later stage in either model, this temporal detail would be smoothed and potential lost. The MaxMSP implementation is shown in Figure 128.

---

\(^{189}\) The circular video buffer accounted for smoothly transitioning from the end of the matrix and back to the beginning again.  
\(^{190}\) A quantised video signal at 30 frames per second would require in this case 300 × 300 × 30 samples per second; that is, 2.7 million samples per second. Perhaps in the future, protocols and drivers will support a flexible system for this, and of course this kind of information is limited by bandwidth, but most technology now is easily capable of this.
4.4.1.2 Transformation

The notion of being able to manipulate sound shapes in terms of their spectromorphology and spatiomorphology requires the ability to contort and transform the nature of terrain contours. Note that in the case of both models the terrain contour directly influences the frequency–space distribution.

Previous research has also focused on transformational processes applied to both the terrain and trajectory structures (James, 2005; Mills et al., 1999; Roads, 1996). The transformations explored as part of this research involved the use of motion video as a transformational terrain surface, algorithmic terrain surfaces that allow for modulation\footnote{\textit{jit.bfg} in \textit{MaxMSP} includes a range of generative functions with the ability to control scale, rotation, translation and a parameter for modulating the nature of a contour.} and the spatial remapping of an image used as a terrain contour.\footnote{\textit{jit.repos} uses a method of table lookup that allows to spatially distort an image.}

Live motion video for control can be seen as an economical dynamical system. The easiest and most efficient method of input is the capture camera. In this way the artist has direct control over the nature of what is captured, and depending on the physicality of the objects captured, a degree of contortion may be applied. In this way morphologies are not dictated by computational and abstract means, but through the
physical interaction with objects in space. Video was often used as a means of testing the responsiveness of the models and certain routines, such as the interpolation of video frames in Section 4.3.1.1. However the noise introduced from a capture device made this less than ideal for the purposes of the systematic evaluation of sound shapes. Rather the project focused on smoother mathematically derived surfaces in order to arrive at more determined and predictable results.

Most commonly used processes were generative mathematical curves such as the Voronoi surfaces generated by jit.bfg and jit.expr. The jit.bfg object allows spatial transformations—that is, translation and rotation of the curve, manipulation of scale, and also a non-linear modulation of the contour allowing for dynamical surfaces that are reminiscent of the undulations in water or clouds of smoke.

Spatial distortion and contortion of a terrain structure through the process of 2D wavetable lookup has been another effective method of transformation; this is referred to as spatial remapping. By applying this it becomes possible to morph and contort the behaviour of another audio process; for example, these geometric transformations may be used to twist, push and pull the timbral and granular density and distribution of sound within a multichannel environment. This process of spatial contortion can be seen in Figure 129.

![Figure 129a. A spatially normalised image. Source: James and Hope, 2012.](image1)

![Figure 129b. A spatially contorted version of Figure 129a. Source: James and Hope, 2012.](image2)
4.4.2 Trajectory Morphology

As was found with the morphology of terrain curves, the ability to transform and explore the ‘inbetweenedness’ of different sound shapes relies just as importantly on the trajectory curve. Further, as the trajectory determines the emergence of temporal features, such as spatial texture, controlling the nature of the trajectory is seen as an important mode of expression for this process too. What follows is a discussion of strategies for the interpolation and transformation of the trajectory.

4.4.2.1 Interpolation

In order to explore morphology by interpolation, one needs a minimum of two curves and a method for interpolation. The interest here is in smoothly moving from one geometry to another. This can be achieved simply through interpolating the set of Cartesian points of one geometry to the Cartesian points of the next, or what is considered a continuous bijection from one curve to a different curve in mathematics. In its simplest form this process could be implemented using a linear cross-fade. Figure 130 shows the smooth morphology of the geometry of one shape into another, along with the colour coding of FFT bins in such a way that the topographical relationships can be observed. Although this may seem unfruitful out of context, the result after the table lookup process in WTS can be vastly different.
Mills and De Souza describe an orbit where one uses a compound path formed by a larger circular or elliptical path, usually at a slower sub-audio rate, as well as a local smaller and faster circular orbit. In this way the faster orbit is what we hear in the audible frequency range, and the slow orbit modulates the position of this faster orbit affecting timbral evolution of the resulting waveform with respect to time $t$ (Mills & De Souza, 1999). This process is achieved simply by adding two different trajectories together such that:

$$
\begin{align*}
    x(t) &= a_1 x_1(t) + b_1 x_2(t) \\
    y(t) &= a_2 y_1(t) + b_2 y_2(t)
\end{align*}
$$

(21)

Where $x_1(t)$ and $y_1(t)$ and are the periodic trajectories within the audible frequency range, and $x_2(t)$ and $y_2(t)$ are systems that are characterised by either quasi-periodic, chaotic or stochastic behaviours (James, 2005). This kind of compound trajectory has also been written about by Wishart (1996) who describes these visually, as shown in Figure 131.
This research project has covered two strategies for interpolating between trajectory curves. The first strategy used a two-dimensional linear panning function\textsuperscript{193} to smoothly transition between four different curves, as shown in Figure 132.

This strategy follows an additive synthesis or vector synthesis model, allowing the performer to control the relative balances of four different multi-parametric trajectory sources, and ultimately allows for the geometric morphology of these. This

\textsuperscript{193} The author experimented with interpolation functions based on the equal power panning curve and the linear panning curve by applying these functions over two dimensions (see Appendix C for formulae).
proved to be one of the most performative aspects of the instrument—the ability to shift from one state to another and explore these ‘inbetween’ states. By controlling the relative balances of determinate v. indeterminate, synchronous v. asynchronous and low-frequency v. high-frequency spatial texture, all of these intrinsic relationships create powerful and expressive dichotomies. In this way it is possible to navigate trajectories as though they all exist along one continuum, interpolating through a lattice of mathematical formulae.

This strategy was appropriate for interpolating between four trajectories, but the research project was also interested in navigating non-linear 2D arrangements of four or more trajectories. An alternative for interpolation in a 2D interface is the nodes object in MaxMSP. This object manages intensity levels for a number of different nodes that vary according to where a point lands in relation to a bounding circle around each spatial node as seen in Figure 133. Nevertheless, this object produces widely varying levels as it does not maintain equal intensity interpolation across the 2D plane.

This research adapted the DBAP panning technique (Lossius et al., 2009) to take multiple trajectory sources and pan across these input sources to generate only one output source. Each input source can be arbitrarily positioned within this virtual
navigable space as shown in Figure 133b. The advantage of DBAP over nodes is the ability for DBAP to adapt appropriate loudness curves where different sound sources might normally intersect as managed by Equation 22. This also ensures that loudness rolloff curves are extended for where sources do not intersect:

\[ I = \sum_{i=1}^{N} v_i^2 = 1 \]  

where \( v \) refers to the amplitude of each source and \( i \) refers to the source number, \( N \) being the maximum. This is a slight point of departure from the usual formula for DBAP panning where \( i \) is used to denote the speaker number. Navigation across many trajectories is possible as shown in Figure 134.

The nodes object can be integrated with the DBAP model for a more compact way of controlling and visualising these relationships. This does not inherently use
DBAP’s own interpolated weights for each node, but rather is used as a 2D graphical interface for control and visualisation as shown in Figure 134.

4.4.2.2 Time Domain Transformation

Although interpolation techniques are effective in allowing a performative exploration of the inbetweenedness of several trajectories, a single trajectory can itself be transformed using time domain or frequency domain methods.

There are a number of time domain transformations that have been explored as part of this research project. These include affine transformations such as scaling, translating or the rotation of trajectories. Trajectories can also be clipping, folded or wrapped at the bounds of the terrain lookup table (James, 2005). Besides these established transformation processes, this research has involved experimentation with a range of other time domain transformations. These include smoothing functions like rampssmooth~, foldover and wrapping with pong~, bit reduction with degrade~ and round~, and phase distortion with kink~. Feedback and crosstalk can also be effective ways of introducing distortion in the evolution of the trajectory shown in Figures 135c, d, e and f. Such methods can introduce non-linearities in the way in which the trajectory evolves over time. Low-pass v. high-pass filters can also serve as a sculpting tool for trajectories, respectively smoothing off sharp edges or sharpening them, as shown in Figure 135f.
Figure 135a. A square trajectory with no time domain transformation applied. FFT spectral bins are colour coded.

Figure 135b. The trajectory in Figure 135a transformed by affine transformation showing changes in its scale and rotation.

Figure 135c. The trajectory in Figure 135a transformed by its scale, rotation, time domain feedback and crosstalk.

Figure 135d. A 2D white noise trajectory transformed by its scale, rotation, time domain feedback and crosstalk.

Figure 135e. A random walk trajectory transformed by its scale, rotation, time domain feedback and crosstalk.

Figure 135f. An interpolated additive trajectory of 2D sawtooth, Lissajous, Hénon chaotic attractor, and random walks transformed by time domain feedback, crosstalk, low-pass filtration and bit reduction.

Figure 135g. A steady-state sound shape over a 16-channel sonogram using the trajectory in Figure 135a.

Figure 135h. A dynamic sound shape generated over a 16-channel sonogram using the trajectory in Figure 135b.
We can use all of these techniques as a means of shaping the geometry of the trajectory. Interesting results come from the deliberate aliasing of the trajectory waveform. *Waveshaping* distortion can also be used to change the harmonic content of a trajectory using Chebyshev functions. Transformations can be applied additively in series or parallel as shown in the transformation used in Figure 135. Here is illustrated the multichannel sonograms for each trajectory explicitly showing how the sound shapes evolves over time.

**4.4.2.3 Frequency Domain Transformation**

As WTS is being used to control a frequency domain process, and it is also synchronised to this process, frequency domain methods can be used for the
transformation of trajectory signals too. Inspired by systems like kinematics, particle systems and swarm systems, or those kinds of distributions determined by vector mathematics, this kind of transformation is concerned with the continuous nature of a set of discrete coordinates and their alignment to successive frequency bins of the FFT frame.\textsuperscript{194} Vector-based transformations may also be applied to trajectories such that each point of an existing geometry is subjected to an independent shift in geometric location. These kinds of transformations impart a different kind of interaction of frequency across space and tend to manifest differently with respect to time, sounding more continuous, smoother and evolving, rather than the time domain methods of transformation.

This kind of transformation as applied to \textit{Model A} can create highly immersive distributions as shown in Figure 136.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig136.png}
\caption{\textbf{Figure 136a.} A single FFT frame of a vector-based trajectory superimposed over a terrain. \textbf{Figure 136b.} 1 FFT frame of the sound shape created by Figure 136a. \textbf{Figure 136c.} 43 FFT frames of the sound shape created by Figure 136a.}
\end{figure}

\textsuperscript{194} Although the other methods use a time domain signal for determining evolution of the system, this method is in fact a frequency domain or vector-based method that determines the evolution of each coordinate in the vector independently.
Applications with Model B also show how this complex intersection of circular movements and random walks can manifest in various sound shapes creating clouds of sound reminiscent of wind or waves on the ocean. This is shown in Figure 137.

Figure 137a. A static line trajectory that has been transformed by a vector-based process that independently shifts all points with circular trajectories (each spectral bin is giving a separate trajectory).

Figure 137b. A circular trajectory with all points transformed independently in circular trajectories.

Figure 137c. The Archimedes spiral with all points transformed independently in random walk trajectories.

Figure 137d. A dynamic sound shape generated over a 16-channel sonogram using the trajectory in Figure 137a.

Figure 137e. A dynamic sound shape generated over a 16-channel sonogram using the trajectory in Figure 137b.

Figure 137f. A dynamic sound shape generated over a 16-channel sonogram using the trajectory in Figure 137c.
This family of transformations can give rise to a sense of spatial width, but the evolving nature of these systems will also in certain cases contribute to an exploration of spatial texture of contiguous space. The results can be reminiscent of a pulsing and shimmering sound scene when applied to Model A, and in Model B can provide a detailed revealing of the spectral map defined by the terrain curve subject to its space-filling properties. However, it is important to note here that each trajectory depending on its speed will reveal the spectral map at different rates, and in many different kinds of ways as a result of the inherent layers of geometries exhibited.

Swarm-based applications may tend to exhibit longer-term space-filling properties, but the points coalesce into groups. Multiple swarms working independently may result in spatial texture that would be considered non-contiguous, and yet each system of swarms may interact or be completely independent. The computational requirements for these behavioural systems tend to be substantially more intensive.

Below is a simple procedural section of Java code responsible for generating random circular orbits for all spectral bins. This is shown along with independently determined random walks. Colour coding is used to show how each spectral bin aligns with the trajectory, and we can see that each random walk has a designated colour:

```java
public void reset()
{
    for(var1=0;var1<1024;var1++)
    {
        var3 = (float)Math.random();
        myInitial[var1] = var3;
    }
}

//Circles
for(i = 0; i < o1.length;i++)
{
    var2 = (int)in1[i];
    var4 = (float)((myInitial[var2]) * (Math.cos(myNew[var2] + (p1 * myInitial[var2]))));
    var5 = (float)((myInitial[var2]) * (Math.sin(myNew[var2] + (p1 * myInitial[var2]))));
    o1[i] = var4;
    o2[i] = var5;
    o3[i] = (float)0;
    myNew[i] = Math.abs(myNew[var2]+(p1*myInitial[var2]));
}
break;
```
Other frequency domain transformations include other spectral processes such as spectral bin rotation. In conjunction with some of the other trajectory transformations this will result in some extremely radical effects. For example, applying this to Model B will instantly reassign certain frequencies with different spatial cues, and is useful in modifying the frequency–space distribution.

4.4.3 The Spatiomorphology of Sound Shapes

In both Models A and B, spatiomorphology and spectromorphology are inextricably linked. However in Model B it is possible to apply a spatiomorphology in separation by applying transformation to the coordinates of the trajectory path after the process of lookup. This process must occur after the terrain lookup process otherwise this will result in a change of frequency distribution as well.\textsuperscript{195} Such transformations are shown in Figure 138.

Rotation, contraction, expansion and translation of sound shapes are some of the transformations defined by Wishart (1996),\textsuperscript{196} and all of these transformations can influence the spatial impression of sound shapes across the listener space. In this way Model B is flexible in accommodating some of the many conventional needs of spatial distribution above and beyond the kinds of detailed spatial treatments for which timbre spatialisation is already known, including the introduction of spatial texture.

\textsuperscript{195} The reason for this is that the trajectory before transformation determines what points of spectra are included in the sound shape. If the trajectory is transformed before the lookup process occurs it will simply change the points of lookup from the terrain, resulting in a change in timbre, whereas if the trajectory transformation is done after the lookup process, then we create a spatiomorphology without affecting the timbral qualities of the sound shape, allowing for any kind of geometric transformation including contraction or expansion of the sound scene, rotation, translation, inversion and many other complex transformations that can be applied to a series of points.

\textsuperscript{196} Wishart describes multi-point spatial trajectories as frame-based transformations.
Figure 138a. A sound shape created with Model B using a 2D white noise trajectory.

Figure 138b. The geometric transformation of the sound shape in Figure 138a.

Figure 138c. A sound shape created with Model B using the Archimedes spiral as a trajectory.

Figure 138d. The geometric transformation of the sound shape in Figure 137c.

Wishart (1996) argues that certain sounds in the environment, like that of a fly, need spatial motion in order to be recognisable. He suggests that more generally, we may look on spatial movements as musical gesture, and consider the typology and implications of different types of spatial gesture and how the spatial motion of one sound object might relate to those of others. Wishart systematically documented many of the known movements that may apply along the horizontal plane (or pantophonic speaker systems). Some of these movements relate to transformations that are possible with the models presented in this research. Table 15 shows some of the scene
transformations documented by Wishart and how these may be applied in the models presented in this research.

Table 15

*Transformations Informed by Wishart*

| Scene contraction and expansion | Defining a scene contraction and expansion has been described in Wishart’s text describing a transition from *distal* to *proximate*, and vice versa respectively.

The scale of the sound shape correlates with spatial width. A scene contraction and expansion may be implemented in several different ways depending on the mapping strategy employed, as described in Chapter 3.

In *Model A* as discussed in Section 3.7, a scene contraction can be generated in two different ways: by reducing the scale of the trajectory signal or by flattening the terrain curve. This consequently has the effect of folding down all azimuth cues into a single point in space, so the sound will appear to contract. As there is no distance encoding in this mapping strategy, the virtual point of contraction will always sit perceptively as a point along this equidistant sound scene.

Early approaches to *Model B* required a very different approach in order to create a similar kind of effect. The multiplexing stage restricts the model in some ways, but by folding the multiplexed windows down so they finally cover the same area the resultant sound becomes identical in all speakers. From here the levels of all speakers need to be managed by another spatial process that will manage distance and azimuth cues. This is not completely ideal either, as one would prefer the model to adapt.
for these unique kinds of spatial transitions.

Model B on the other hand will generate a scene contraction and expansion by ‘scaling’ the distance cues directly. This means that by altering the nature of the spatialisation directly, it bypasses the usual scenario where WTS determines both spectromorphology and spatiomorphology in an intrinsically combined way. Instead this means that if the user transforms the spatial parameters directly they can influence a spatiomorphology without an associated change in the spectromorphology.

Rotation correlates with a perceived rotation of the auditory scene. This auditory scene change may be possible using mapping strategies one, two and three. A rotation of the auditory scene is possible by rotating the trajectory signal in both models. In Model B this spatio-transformation must be performed post-terrain lookup so that it does not affect the spectromorphology exhibited.

A clockwise rotation of this trajectory by $90^\circ$ or $\pi/2$ radians correlates with the same shift in the frequency–space distribution.

In both models the spectrum in this way is treated as a continuous domain, tiled from low to high frequency, and back to low, etc. As frequency bands approach the highest FFT bin, they naturally migrate again to the lowest band. Although the exact relationship between the terrain function and what is heard correlates superficially, they are in fact abstracted from one another in the sense that the ‘direction’ of the trajectory as it passes across the terrain function determines the order of subsequent frequency bins, and hence any phase shift applied is also subject to this clockwise or
Scene translation

This auditory scene change is only possible with Model B, and would involve a spatial transformation of trajectory translation in the $x$ or $y$ value. This is a spatio-transformation without any change in the spectromorphology; therefore the trajectory transformation is done post-terrain lookup.

Scene swing

This auditory scene change may be possible using Model A by inverting the colour spectrum about one fixed value. Alternatively this is possible with Model B, and would involve a spatial transformation of rotation at the corner of the trajectory plot of $\pi/2$ or $-\pi/2$ radians, depending on whether the transformation is clockwise or counter clockwise respectively. This is a spatio-transformation without any change in the spectromorphology; therefore the trajectory transformation is done post-terrain lookup.

Scene twist

This auditory scene change is only possible with Model B, and would involve a spatial transformation of trajectory scale from positive $[+]x$ and $[+]y$ to negative $[-]x$ and $[-]y$. This is a spatio-transformation without any change in the spectromorphology; therefore the trajectory transformation is done post-terrain lookup.

Scene spiral

This auditory scene is possible using Model B, and would involve a spatial transformation of trajectory rotation about the centre by $\pi$ or $-\pi$ radians as well as a change in magnitude from 1 to 0. This dual transformation will generate this spiral shift in ward, and the opposite a spiral shift outwards. This is a spatio-transformation without any change in the spectromorphology; therefore the trajectory transformation is done post-terrain lookup.

anticlockwise movement.
Scene distortion (contortion)

Any complex successive chain of transformation processes in series will arrive at extremely complex geometric shifts. The kinds of transformations seen to trajectories in the previous section indicate how varied some of these shapes can be. Some of these shifts are shown in Figure 137.

4.5 Chapter Summary

Chapter 4 has been concerned primarily with the language of sound shapes possible with Models A and B through the laptop performer’s interaction with WTS. Steady-state distributions are discussed before discussion notions of spatial texture that arise when the trajectory is asynchronous to the timing of the FFT frame. Both low- and high-frequency spatial texture are discussed for both models. Morphology of the WTS is also discussed—that is, the morphology of the terrain and trajectory that results in clearly articulated changes in spectromorphology and spatiomorphology. By controlling both of these aspects the laptop performer has the ability to morph sound shapes by contorting, distorting and transforming the terrain and trajectory. This mechanism consequently becomes quite manageable to control very complex distributions of
frequency bands across space. Visualisation of this process is also discussed as a means of illustrating the terrain, trajectory, loudness, frequency–space and frequency–amplitude–time. This information provides the laptop performer with a confirmation of these complex processes in visual form.
Chapter 5: Compositions

… algorithmic composing was to result not so much in a music of notes (the ‘lattice’ structure of quantised pitch, duration and intensity values) as in sound textures and complex sonic gestures defined compositionally by their timbre and internal development. Such was my method of conceiving ‘timbre’, here understood as the emergent sonic morphology, as ‘musical form’ itself, and ultimately the very object of composing … The idea was that both the micro- and the macro-level of music would emerge from a hidden, low-level (chaotic) dynamics. (Di Scipio, 2001, p. 249)

5.1 Introduction

This chapter discusses four of my musical works that have each served as a performative means of evaluating and improving the software throughout its development. The creative perspective gave a unique insight not only in terms of the necessity for the exploration of sound shapes, but also how these sound shapes would serve musical function within the context of electroacoustic performance practice. Although the majority of the creative folio explores versions of the software that do not reflect all of the later developments in Model B, these compositions proved invaluable in the critical evaluations that led to the development of both models.

From the beginning of this research project, it was my intention to submit a folio of new compositions and to undertake a series of performances that would test the models in practice at various stages of software development, culminating in a performance recital. Performative evaluation and experimentation of the models presented in Chapter 3 led to a discovery and exploration of many kinds of sound shapes derived from timbre spatialisation, as documented in Chapter 4. WTS provided a framework that allowed both the intuitive and flexible performative exploration necessary for creating sound shapes in real time, both in the studio and in live performance. In this way, timbre spatialisation allowed for exploration of spectromorphology and spatiomorphology on pre-recorded sound materials as well as