Multi-channel Audio Production with a Focus on Recording, Mixing and Live Events

Christopher Steven McDonnell

A thesis submitted to the University of Huddersfield in partial fulfilment of the requirements for the degree of Master of Arts

The University of Huddersfield

November 2020

Copyright Statement:

Abstract

The advent of multi-channel audio in the mid-twentieth century has created an interesting way to play back and diffuse audio which can enhance listener experiences through working perceptual cues and immersion. In the past this has mostly been done with experimental music, but the projects explore how this format could be used with commercial music.

Four production projects are presented that explore different approaches to the use of space for commercial multi-channel music. Project 1 explores how testing a large set of microphone techniques can help make production decisions to match the mood of acoustic songs for 5 channel surround sound. Project 2 explores how the dynamic use of space can be used to create an imaginary space where gestures are created to enhance and match the style of the music for Higher Order Ambisonics. Project 3 explores how off-the-shelf PA and live mixing technologies can be used to mix a band in a live setting with the use of a custom versatile software controller. Project 4 explores a different approach to immersion where the band surround the listener to create a practice room layout for Higher Order Ambisonics.

The submitted works are accompanied by this commentary which begins with a review of the current research into ambisonics, producing music for multi-channel, and the technologies for recording, mixing and live sound multi-channel production. The text then discusses the projects individually and how the previously mentioned research and techniques were applied to align with the spatial aims of each project. I conclude this commentary by discussing the effectiveness of using multi-channel audio as a basis for each projects intentions and how the use of multi-channel could be used in future personal projects to expand on the works submitted.

Table of Contents:

List of Images:

List of Tables:

Appendix List:

List of submitted work

Project 1 - Charlie Hulejczuk - Surround Sound Recording

Cursed - 5.1 Surround - 5 minutes 16 seconds

Blessed by the Blossom - 5.1 Surround - 4 minutes 41 seconds

Waste Our Time - 5.1 Surround - 5 minutes 25 seconds

The Duchess of Westdene - 5.1 Surround - 5 minutes 18 seconds

Project 2 - create.evolve.destroy. Studio - 5th Order Ambisonics

Continuous mix - 5th Order Ambisonics - 22 minutes 11 seconds

Track list - *tent.legs.*, *small.pizzalike.ladder.monk.* & *fuck.the.grid.*

Project 4 - The University Funk Band - 5th Order Ambisonics

Hear Me Calling Out Your Name - 5th Order Ambisonics - 4 minutes 23 seconds

1 - Introduction

Since the development of mono audio in the late 19th century, music production has evolved to different formats, from stereo to surround sound and more recently 3D environments. Despite the widespread implementation for cinema, multi-channel technologies have not been as widely integrated into popular music production. With the current way of streaming popular music, musical releases in multi-channel are limited to mostly re-releases of classic albums in five channel surround that often do not use space in an experimental way. Throughout the thesis, the theory and applications of multi-channel production will be investigated and referenced. These include the technology and approaches of recording, mixing and live use. This was approached practically and investigated the use of space from a number of different angles to see how the music would come across as a result of the spatial elements.

While there is expansive research into the development of the technology to produce multi-channel audio and the perception of sound in multi-channel, research into popular music production for multi-channel, particularly ambisonics, is still an emerging area. James Bagshaw from The University of Hull is currently researching into popular music production for the output format of ambisonics with the aim to find a spatial language and workflow. Sébastien Lavoie from The University of Huddersfield conducted a PhD into the live performance aspects of spatialising electronic dance music in ambisonics. This thesis aims to contribute to this area showing how space can be approached from a range of angles to produce contrasting styles of music and musical intentions.

More recently, surround sound releases for popular music mixes have become more frequent. Producer Steven Wilson has been remixing the albums of bands such as King Crimson, Jethro Tull and Tears for Fears since 2009 on to 5.1 channel surround. Wilson's approach in these remixes is to remain in the logical sense of having a frontal image that adheres to the standards of music production and not experiment too much with the placement of sound (Bacon, 2019). Additionally in 2017, mixer Giles Martin reissued The Beatles' Sgt. *Pepper's Lonely Hearts Club Band* (2017) in 5.1. Certain points in the album such as *"A Day in the Life"* (2017, disc 5, track 17), took full advantage of the surround sound that mirrored the famous use of dynamically moving the lead vocal from side to side in the first verses. But again, the surround sound album mostly adheres to the standards of a frontal focussed mix. This approach from Wilson and Martin of cautiously mixing in surround is the most common approach to producing surround sound for commercial releases today. This thesis and the four accompanying personal projects explore the use of space in a range of ways and contrasting directions for different musical approaches, intentions and output formats.

This commentary is accompanied by four contrasting personal projects that highlight different approaches to space in multi-channel from a popular music production point of view for a variety of different output formats. These output formats are five channel surround sound, Higher Order Ambisonics (HOA) and live audio. The projects that investigate multi-channel are grounded in the Collins (2010) approach of the creative practice and deductive research approach. In this regard, the spatial elements of my work were recursively reflected based on the propositional ideas discussed between the artists and myself for each project. This recursive reflection was repeated until the original propositions were best highlighted for each project. Throughout the projects, space was used in a variety of ways such as creating realistic impressions of space and creating imaginary environments with dynamic movements.

Throughout, these are the key aspects of what I am exploring:

- To explore how the use of space in multi-channel can be an effective technique to enhance and highlight the compositional ideas in a set of contrasting projects put forward from the artists and not be distracting to the listener.
- To study different capture methods and find how they can be used to ensure creative mix possibilities in the various multi-channel playback formats.
- To explore how the use of space can be used to create either real life spaces or create something imaginary and non-realistic.

Project 1 was collection of recordings of an acoustic singer-songwriter that explored how testing a large set of possible recording techniques in order to make production decisions. The testing allowed me to make evaluations in order to help me match the style and mood of contrasting songs to the microphone techniques and positioning in the recording space. The reflection of style and mood matching aimed to increase the intimate connection with the artist. The output format of project 1 was 5 channel surround.

Project 2 was a mixing project of an experimental prog-rock band that explored the dynamic use of space to create non-realistic and imaginary spaces to enhance and fit the musical ideas put forward from the band. This exploration of using space dynamically in multi-channel popular music to create musical intrigue is less explored than other research areas of popular music. In the late 1960s and early 1970s, the likes of Hendrix's *"Crosstown Traffic"* (1968, track 3) used space in a dynamic way where elements would be moved from side to side of the stereo field.

This approach of dynamic movement has since fallen of favour (Dockwray & Moore, 2010). Now elements are mostly statically placed. This mixing project explored how to find a middle ground where the dynamic use of space was an element of interest in the mix, but not becoming distracting to the listener and not detracting from the music. The output format of project 2 was 5th order ambisonics.

Project 3 was a live event that explored how off-the-shelf PA and mixing equipment could be used to produce a live band on to a custom 3D speaker array with height. It is common practice in the acousmatic approach of spatialisation where custom hardware and controller interfaces are created for the sole purpose of diffusion. For example, the sound system at the University of Sheffield uses a custom “M2 Diffusion” interface for spatialisation (Moore, Moore & Mooney, 2004) that uses faders to control the level of one loudspeaker. This project was supplemented with the creation of a bespoke panning software that allowed the control of multi-channel panning in a commercial mixing console. As stated, the output format of this project was the live concert with the bespoke speaker array but there is a binaural recording with video.

Project 4 was a mixing project of a pop-funk band that explored an alternate approach on immersion where the band surround the listener in a stage perspective mix (Waldrep, 2007, pp. 3). This approach contrasts the typical pop mixing approach of using a frontal image. The stage perspective approach aimed to create a practice room layout that also added an element of separation and clarity to the arrangement. This project aimed to add clarity to the arrangement by separating the instruments but still be conventional enough for a pop-mix in terms of the spatial placements. The output format of project 4 was 5th order ambisonics.

This thesis aims to contribute to the research into multi-channel production of popular music by using different spatial approaches practically. The different projects highlight how the use of space can be used and approached in different ways to create different spatial impressions.

2 - Literature Review

This section assesses the current writings and methods in regards to the production of multi-channel music as well as the theory and applications of ambisonics.

In each subchapter, I discuss a different aspect of multi-channel production that is relevant to my masters research. Various recording, mixing and live audio research methods and techniques will be discussed, as well as how they influenced my decision making in the practical work.

2.1 - Ambisonics

Ambisonics is a spatial audio format proposed by Michael Gerzon in 1973 that allows for the use of full spherical information, therefore unlocking sonic possibilities not possible with the standard stereo, mono of surround plane (Gerzon, 1973, pp. 2). Nishiyama, Nagata, Ogata, Hasue & Kashiwagi (2018) call Ambisonics a “full-sphere spatial format” and explain how it is an appropriate medium for use with Virtual Reality (VR) and composition with the intent of full immersion. Ambisonics allows the reproduction of fully 3D (periphonic) audio, making it an important creative medium.

Based on an early 1930s stereo patent by Alan Blumlein (1958) that featured multiple directional microphones in combination with an omni microphone to locate sounds in a space, Michael Gerzon joined with Peter Fellgett’s Ambisonic team in late 1973 to develop the Soundfield Microphone. This microphone consisted of four, very closely spaced cardioid or sub-cardioid microphone capsules arranged in a tetrahedron. This microphone, like others later invented such as the RØDE NT-SF1 and the Sennheiser AMBEO, records 1st order ambisonics in A-Format. This A-Format however is not intended for listening and must be converted into a second set of audio signals known as B-Format. Depending on the hardware, this may be done either in the hardware of the microphone unit itself or in software. When decoded, the B-Format represents four signals (Zotter & Frank, 2019, pp. 9-12):

- W represents the pressure signal corresponding to the omnidirectional pickup with equal sound from all four directions at equal gain and phase.
- X is the front to back directional information
- Y is the left to right directional information
- Z is the up to down directional information

The X, Y and Z information comes from simulating bi-directional figure-of-eight pickup patterns. Signals in B-Format can then be used in a number of ways, such as in a virtual microphone simulation (with pickup patterns including omnidirectional, cardioid, figure-of-eight or anything in between), or simply be decoded to speakers positioned in both the horizontal and vertical planes.

The “order” in ambisonics plays a role in how detailed the spatial field is. A helpful visualisation of this concept is to think of the sound sphere being comprised of many pixels. The higher the order, the more pixels and therefore more clarity. With 1st order ambisonics, the spatial resolution is low. Because of this, recordings can create blurry images with a lack of definition in the direction of the sound source. Logically so, the higher the order, the higher the number of channels. 1st order ambisonics consists of 4 channels, when increasing the order, the number of channels also increases. 2nd order ambisonics consists of 9 channels, 3rd order consists of 16 channels. This can increase to 7th order ambisonics which consists of 64 channels. Due to software limitations 7th order ambisonics is currently the highest available to the public. When working in a High Order Ambisonics (HOA), this audio can be decoded down to a lower order for easier delivery of content, most commonly being 1st to 3rd order (SSA Plugins, 2020). For the ambisonic mixing projects (2 and 4), 5th order ambisonics was used due to the increased resolution in the localisation of the instruments in the soundfield. This allowed me to more accurately place sound and create more accurate gestures when the positions of the sounds were changed dynamically.

Encoding a signal into ambisonics can be done in one of two ways. As previously described, a sound source can be recorded into ambisonics with a microphone such as the Soundfield, or a sound could be encoded into the ambisonics format in point production using software. This is commonly achieved through the use of an ambisonic panner. This software panner allows the user to place the sound in a sphere by allowing control of the azimuth and elevation angle. The azimuth angle relates the horizontal angle (-180º to 180º), and the elevation angle relates to the vertical angle (-90º to 90º). An azimuth and elevation angle of 0º would position the sound directly in front of the listener. Most ambisonic encoding software will allow the user to set the order of ambisonics being used. This encoded signal will then be in a separate audio file with multiple channels to represent the position in the sound sphere. This file however is not intended to be listened to, much like A-Format, it must be first decoded. The main advantage of ambisonics is that the audio files are versatile and can be decoded to multiple different listening systems. The encoded audio file can be decoded to a binaural stereo signal for headphone listening or to a speaker system by telling the decoder software the azimuth and elevation angles of the speakers. This speaker system can range from anything to a small studio to large systems with hundreds of speakers.

When creating ambisonic loudspeaker environments, it is challenging to place speakers equally spaced in both the horizontal and vertical planes. It is recommended that all loudspeakers should be placed more than 2 meters from the listener and no closer than 1 meter to a wall (ITU, 2015). In most conventional spaces, this is not possible so bespoke spaces have to be made. Eric Benjamin (2008, pp. 2) explains how 3D arrays can be squashed horizontally and vertically to fit in domestic spaces while still keeping a fixed distance from the listener position. This can lead to issues where there is an inconsistent audio level in the sphere due to the different gaps between the speaker positions, ultimately leading to an inaccurate image. There are several possible successful shapes for speaker arrays. Polyhedra Ambisonic arrays are speaker arrays where the horizontal layout is mirrored to the vertical plane. Regular layouts for Polyhedra arrays are mirrored squares, hexagons or octagonal. Cylindrical arrays place the speakers in two circles where the horizontal speakers are mirrored in the vertical plane creating a cylinder of speakers. Image 1 shows this setup where the bottom circle of speakers arranged in a hexagonal shape are below the listener position and the top circle is above at equal spacing.

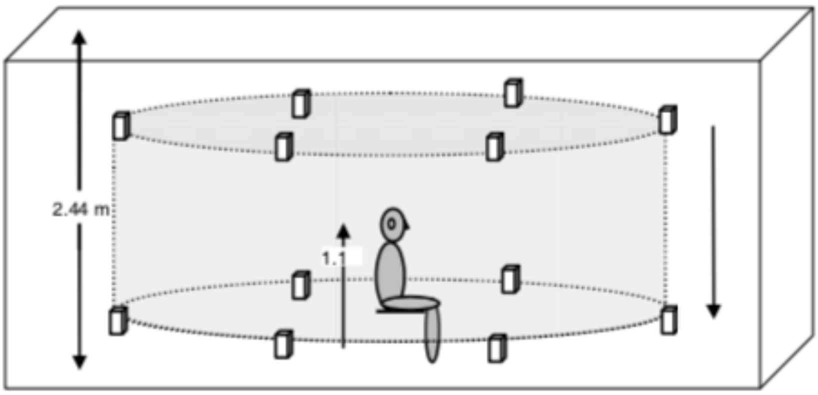


Image 1 - Cylindrical Loudspeaker Array (Benjamin, 2008, pp. 4)

SPIRAL studios at The University of Huddersfield has an example of a polyhedra speaker array that was used for the mixing of Projects 2 and 4. The studio incorporates three rings of eight Genelec 8240a speakers arranged in an octagon. There are no speakers below the listening height. The first ring is at the listener's ear height (1.4 meter from floor), the second ring is at 2.55m from the floor, the third ring is at 3.7m from the floor and there is a final speaker positioned directly above the listener position 4.4m from the floor. Additionally four subwoofers are placed in the left, right, front and rear positions that take the summed output of each side of speakers. Due to its cylindrical design, for ambisonic uses, a delay has to be added at each speaker to make the arrival times of each speaker to the listener position the same. This has to be done in software or in a dedicated decoder as there is no speaker processor available in the room. This added delay creates a hemispherical soundfield (Mac, 2009) with equal distances between the listener and speaker positions. Appendix 1 shows the azimuth and elevation angles of each speaker position and the delay amount required. Appendix 2 is a diagram showing the positions of the speakers provided by The University of Huddersfield.

Acoustics and their effect on localisation is an important aspect to consider in sound design, mixing, or any other activity that requires spatial accuracy. Thresh, Armstrong and Kearney (2017, pp. 1-9) experimented with the accuracy of perception angles on both a 3D speaker environment and a virtual speaker setup on headphones using 1st, 3rd and 5th order ambisonics. This was tested with pink noise bursts lasting 700ms that were placed at specific azimuth and elevation angles that the listener had to position using a three dimensional mouse. It is important to state that these tests used a generic Head-Related Transfer Function (HRTF) for the virtual speaker listening test on headphones. On headphones across all three orders, sounds coming from behind the listener position had a larger error in positioning due to the sounds sharing the same time and level arrival time. This area is also known as the “cone of confusion” (Yost, 2007). This error was particularly present while listening on headphones. With the “cone of confusion” outliers kept in the data, the range of errors decrease as the ambisonic order increases on speakers. However, on headphones, strangely, the opposite effect happens where increasing the ambisonic order increases the range of errors in position perception. With the “cone of confusion” outliers removed however, both the errors in headphones and speaker position perception decrease as the ambisonic order increases. This could mean that when creating ambisonic mixes and placing sounds in “cone of confusion” area could lead to confusion when comparing on headphones and speakers. Appendix 3 and 4 are graphs taken from these tests that show the angle error in box plot form. These tests show that using a generic HRTF for binaural listening can drastically change the perceived angle when listening at higher orders. They additionally show that there is little benefit to localisation between the 3rd and 5th order results when listening on a speaker system. However, when conducting my own tests where I directly compared the ambisonics orders in SPIRAL, I found that using the 5th order gave the best accuracy of localisation and allowed very direct placement of sound in the soundfield. This may be due to the nature of SPIRAL studio with its 25 speakers being processed into the hemispherical setup allowing a greater level of localisation. When increasing to a 6th or 7th order, I found that the difference was very little and given the additional channel count required for the higher order ambisonics, and the size of projects required during mixing, I decided that 5th order would give me the best results in terms of localisation and availability of computer power.

For headphone listening of 3D audio with ambisonics, Head Related Impulse Responses (HRIRs) and a time domain equivalent of Head Related Transfer Functions (HRTFs) are used to filter a sound in a similar ways that ears do when hearing sound from a specific direction (Thresh et. al, 2017, pp. 1-9). This combination of functions and responses can increase the accuracy of perception compared to results achieved using a generic set of functions. This is because of slight differences between every persons' head and ear shape and size. However, generic HRTFs can be useful for creating a general universally accurate binaural experience. These are created through recording impulse responses with microphones placed in a persons' ears, leading to an accurate representation of natural location cues and can later be used to simulate sound location when using headphones. These filters can also be used to simulate a virtual speaker array by recording impulses of speaker positions. Similarly this is done by recording impulses of speaker positions.

While accurate spatial localisation is the ultimate goal of ambisonics, in situations where either physical or computational practicality is essential, Mixed-Order Ambisonics is a useful compromise. Ambisonic production allows for Mixed-Order Ambisonics that can allow a higher resolution of horizontal spatial resolution with a lower order for elevated sound sources. This lower resolution in the vertical plane can be achieved because human hearing has less spatial resolution of sound from either above and below when compared to the horizontal plane. A Mixed-Order ambisonic system would be used when a speaker system does not have the same resolution in the horizontal and vertical planes. This would either be because of physical limitations in space or a personal choice for when vertical resolution is not as important as horizontal resolution. For reference, a normal ambisonics system with equal resolution in the horizontal and vertical planes in this format would be referred to as a Fully Periphonic (3D) system. The first mixed order ambisonic scheme introduced was the #H#P scheme that bridges the space between horizontal signals and fully periphonic signals. The H value represents being the horizontal order, and the P value represents the superimposed periphonic signal. This method has the advantage of a reduced channel count if sound sources are on the horizontal plane, however when sources increase in elevation, they can become smeared due to the lack of resolution in the vertical plane. To combat this, Chris Travis (2009, pp. 1-6) proposed The #H#V Mixed-Order notation scheme where H and V show the horizontal and vertical resolution respectively. For example, a 4H1V signal would be a mixed order signal with the horizontal resolution four times as accurate as the vertical resolution. This can be mixed in a number of ways but in the scenario where H = V, this would be a regular ambisonic system with the same resolution in the horizontal and vertical planes. When comparing the #H#P and #H#V schemes, Travis describes how when both are at a similar horizontal resolution the localisation in the horizontal plane is high, but with the #H#P scheme the resolution "degrades quite rapidly with elevation" when compared to the #H#V scheme. This would be undesirable when accurate movement in the vertical plane is required. This #H#V scheme would be used when the loudspeaker density is not sufficient to cover a full sphere. However, with the #H#V scheme, it is impossible to know the number of component channels unlike when using a normal ambisonic format or the #H#P format. In conclusion, Travis (2009) explains that the mixed-order schemes gives ambisonics freedom from the horizontal only system without the difficulties in incorporating a fully periphonic ambisonic speaker system. Due to the amount height speakers and the vertical resolution of SPIRAL studio a mixed-order system was not required for my ambisonic mixing projects (2 and 4).

Another aspect to bear in mind when mixing in ambisonics is the harmonic material. The accuracy of localising sounds from specific directions can change depending on the frequency. Lee (2018) states that “lower frequencies tend to be localised lower [in height] regardless of the source”. While experimenting with locating sound in the vertical plane, Roffler & Butler (1968, pp. 1255-9) found that as frequency increased, the judged image height increased. A sound played at a lower frequency will be perceived at the same height no matter where the actual loudspeaker is placed in the vertical plane. Image 2 below is taken from their findings and shows this. In a similar experiment, it was found that sounds under 500Hz were very challenging to accurately locate, especially when the position of the speaker was one meter or higher from the listener position (Ferguson & Cabrera, 2006, pp. 169-70).

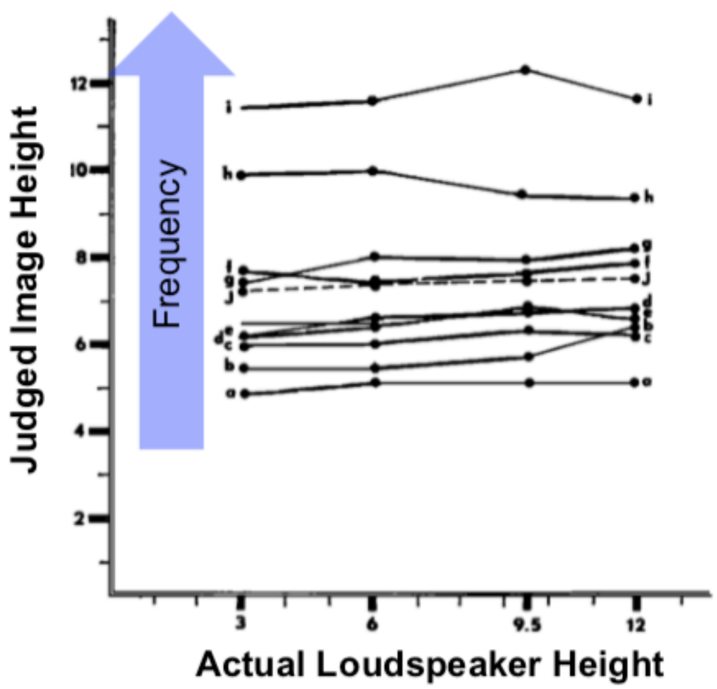


Image 2 - Loudspeaker Height vs. Judged Image Height (Roffler & Bulter 1968)

Similarly to these finding, frequency content can affect Apparent Source Width. Apparent Source Width is a subjective attribute of acoustical width of a sound source in a concert hall or opera house that is perceived by a listener (Sato & Ando, 2001). It is based upon the spectral components of the source as well as cues from Interaural Level and Time Differences. Apparent Source Width increases as frequency of the sound source decreases and vice versa. Because of this, the accuracy of the localisation of low frequency sound sources can be less than the accuracy of high frequency sounds. When combining this theory with the tests conducted by Roffler and Butler (1968), it shows that localising sound sources comprised of low frequencies can be harder when compared to localising sound sources with high frequency content. These findings highlight the issue of how a lack of localisation of low frequency content could lead to a blurred spatial image. For the most accurate perception of direction for low frequencies, it is advised to avoid placing these sounds in the vertical plane as well as avoiding the “cone of confusion” previously described. I conducted my own personal tests for frequency localisation and height perception where I determined what spatial impressions could be made while mixing in ambisonics. These tests are discussed in the ambisonic tests section of project 2 (Section 5.1).

One of the many advantages of 3D audio is its potential for interactivity. To simulate a virtual speaker array, head-tracking can be used to simulate head rotations within the array. This head tracking can be done in different ways. In Virtual Reality headsets, gyroscopes, accelerometers and a camera can track the position and angle and accurately place sound in the binaural field. When head tracking in a studio listening environment, a head tracking headpiece can be used to track the listener’s head position which would then control the rotation of the sound sources. The additional head tracking can also aid in perception angle as it can reduce the confusion between front and back placed sound (Thresh et al. 2017). None of my projects intended to have any elements of interactivity in them. It was intended that the listener would remain stationary in the listening position and slightly rotate to possibly focus on one specific part. However, in reviewing Project 4, it was found that possibly adding a binaural version with head tracking could have aided in the experimental nature of the mix and made the representation more realistic.

Ambisonic formats are becoming more popular to the general public with accessible tools to create ambisonic audio and easy ways of listening. The Facebook 360 Spatial Workstation (facebook360, 2020) is software that is used for producing spatial audio for 360 video and Virtual Reality (VR). This software exports the created audio to a binaural stereo signal or can be exported into a range of formats such as ambisonics. In the software you can add B-Format audio or single source audio that can be positioned at a point so that when the viewing angle of the 360º virtual camera is moved, the audio will move in relation to the viewing angle. In other words, the soundfield is dynamic rather than static, leading to many possible creative uses in audio-visual media. The video can then be exported and can be uploaded to Facebook or YouTube with all of the information required for playback on other devices. Depending on the playback device, head tracking can be used to increase the listener immersion or the gyroscopes in a mobile phone can be used to move the video angle. Other mainstream social platforms such as YouTube also allow users to upload 180° or 360° video footage. In addition, 3D audio technology is steadily being introduced into video games with VR games as well as options for binaural headphone audio.

Later sections will go into more detail on the different aspects of recording and mixing with ambisonics as well as how ambisonics was used during the personal projects I completed.

2.2 - Multi-channel Recording Techniques

This section discusses the recording techniques used to capture sound through stereo, surround and 3D ambisonics techniques. Each technique will be discussed regarding the advantages and disadvantages of how they are used to capture the desired subject and space for specific contexts.

2.2.1 - Stereo Microphone Recording Techniques

It is important to establish the following stereo microphone recording techniques as they are the basis of the multi-channel recording techniques later discussed. In this section, I will be referring to the term “stereo microphone techniques” as the use of two microphones to achieve a coherent stereo image. Stereo recording techniques are used in all modern recordings to provide a sense of the soundfield from left to right and a sense of depth and distances (Owsinski, 2014, pp. 87-98). Our ears use three methods to determine localisation of sound. These are amplitude, time and phase differences (Corbett, 2014, pp. 126). There are three main stereo microphone techniques that I will discuss that are used to capture stereo audio: Coincident/Near-Coincident, Spaced Pair and Mid-Side.

The XY array (or coincident pair) recording technique uses two cardioid or hyper-cardioid microphones that are time coincident meaning that the sound source arrives at both microphones at the same time. The difference in amplitude between the two microphones provide the perceptual cue to the direction of the sound source, also known as the Inter-aural Level Difference (ILD). Image 3 shows the most common use of the XY array where two cardioid microphones are angled at 90° with their pickup patterns crossing at the front and pointing towards the sound source. While 90° is the most common angle used for XY, the angle difference can range between

70° and 130° where the perceived stereo width would either narrow or widen respectively (Corbett, 2014, pp 126). In either a monitoring or playback situation, the two captured signals are panned left and right to create the stereo image. (Huber & Runstein, 2012, pp. 145). By using hypercardioid microphones, the overlap of the two captures will be decreased which will in turn increase the perceived image width. However, Corbett states that the sounds coming from the centre will be picked up with a "greater off-axis collocation and slightly less amplitude" (2014, pp. 128) due to the hypercardioid's more directional pickup. The XY array recording recording technique has the advantage of being mono compatible due to there being no time arrival differences between the two microphones meaning that they are phase coherent and will not cancel each other out when summed to mono. However, due to the lack of time difference, the perceived stereo width will be narrow when compared to other stereo microphone techniques such as AB (spaced pair) or mid-side.

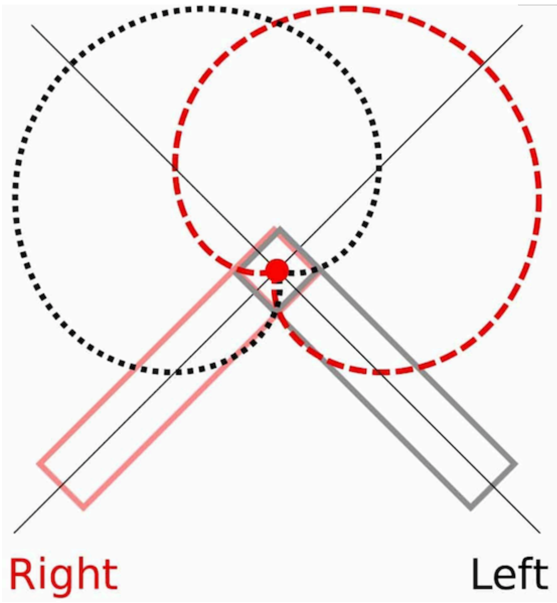


Image 3 - XY Recording. Retrieved on November 15th, 2020 from:
https://thehomerecordings.com/midside-or-xy/

For an increased width perception using cardioid microphones, near-coincident techniques such as ORTF could be used. This technique places the two microphones at an angle of 110° where the capsules are placed 17cm apart. NOS is similar where a 90° angle and spacing of 30cm is used. This gives a transparent stereo image due to the spacing between the two microphones being similar to the ears on your head while still being mono compatible. Near-coincident techniques have the advantage of good ambient reproduction with precise image positioning due to the introduced width (DPA, 2018).

Secondly, the AB recording technique (or Spaced Pair) uses two microphones (most commonly omnidirectional) where their capsules are spaced between 40 and 60cm apart, meaning that the arrival time of a sound source between the two microphones will be different. This difference in time provides the perceptual cue to the direction of the sound source which is known as the Inter-aural Time Difference (ITD). By increasing the distance between the two microphones, the perceived width will increase, however if the spacing is too wide, there will be a lack of definition from sound sources in the centre. Due to the time difference, this technique can become the least mono compatible due to phase differences and comb filtering. It is possible to minimise mono compatibility artefacts during the recording phase but impossible to remove all artefacts (Corbett, 2014, pp. 134). The spaced pair technique does however have the advantage of recording a wide and expansive stereo image that will be more enveloping than the previously mentioned techniques. In the 1950s, engineer Kenneth Ernest Wilkinson expanded the AB method to make the Decca Tree recording technique (Polymath, 2012). The Decca Tree technique uses three omnidirectional microphones to capture the left, right and centre of the soundfield. It is commonly used for orchestral recordings where the tree is placed above the conductor, offering a sound balance close to what the conductor hears (DPA, 2018). By adding a centre microphone, the issue of a lack of definition in the centre can be removed when the left and right microphones are placed too far apart is avoided. Because of this, the distance between the left and right microphone can be increased much wider (up to 2m) which creates a precise stereo image and enveloping capture.

The third main stereo technique is the Mid Side (MS) recording technique. Similar to the XY coincident pair technique, MS uses two microphones that are placed in the same position. However, the “mid” microphone points directly at the sound source and is most commonly a cardioid microphone. The “side” microphone is figure-of-eight and is aimed so that the null point is pointing towards the source. Image 4 shows the combination of the cardioid and figure-of-eight microphones in the Mid Side configuration. In post, the side microphone is duplicated and one side is panned hard left and the other is polarity inverted and panned hard right. The stereo width can be increased or decreased by adjusting the amplitude of the two side signals. When summed to mono, the two side signals are cancelled out leaving only the Mid signal (Keller, 2019). Mid Side recording has its advantages over XY recording because the stereo width can be adjusted in post production. Additionally, MS does not lead to any issues with mono compatibility unlike the spaced pair technique.

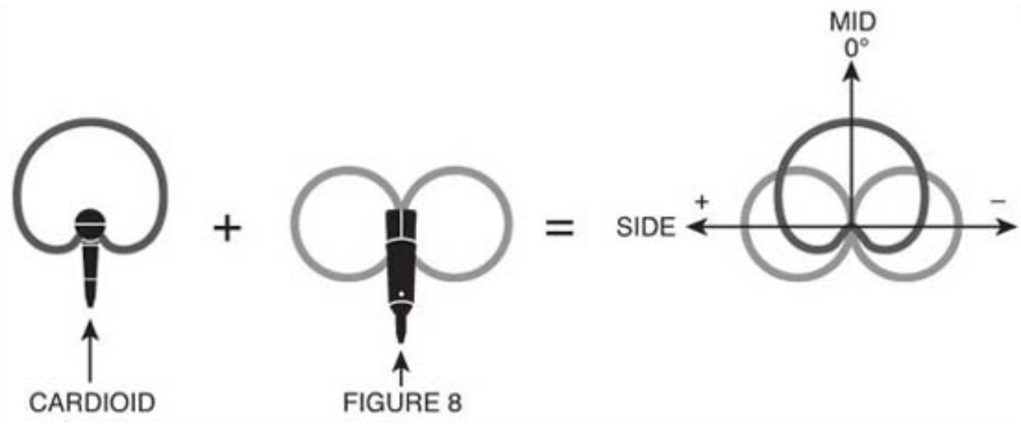


Image 4 - Mid Side Recording. Retrieved on 15th November, 2020 from: https://thehomerecordings.com/midside-or-xy/

These three main techniques can be used for capturing both single sources, ensembles as well as capturing the ambience of a room to create a stereo image.

2.2.2 - Surround Sound Recording Techniques

The stereo microphone techniques mentioned above were modified and developed so that they could successfully capture surround sound. "Surround Sound Recording Techniques" in the upcoming section refers to recording techniques used to capture 360º horizontally with no height information as the ones discussed are intended for playback on a surround sound speaker setup such as 5.1.

Mid Side was supplemented with a rear facing cardioid microphone to become Double Mid Side (DMS), originally invented by Curtis Wittig and Neil Muncy (Bartlett, 2009, pp. 477-8) (Rumsey, 2001). This rear facing microphone is used to capture indirect signals from the source and when combined with the original two signals from Mid Side it gives a "full surround image" (Wittek, H., Haut, C. & Keinath, D. 2006). The three captured signals can be decoded in numerous ways, but a common tool used is the Schoeps Double Mid Side decoder plugin (Schoeps, 2018). Different decoding options are available due to the coincident setup of this technique. This decoder plugin takes the three microphone signals and outputs them to 5 channel surround sound. The plugin allows the user to tweak the level of each signal in post as well as change the perceived angle and pickup pattern by changing the level and phase between each signal. The technique can lead to a realistic capture or reproduction of a space and "offers immense flexibility in post-production to create front and rear sound stages with adjustable width, prominence and focus" (Robjohns, 2017). Image 5 is a bird's-eye view of the Double Mid Side recording technique. The blue section refers to the frontal facing cardioid microphone. The orange refers to the side facing figure-of-eight microphone. The green refers to the rear facing cardioid microphone that makes this a surround sound capture technique.

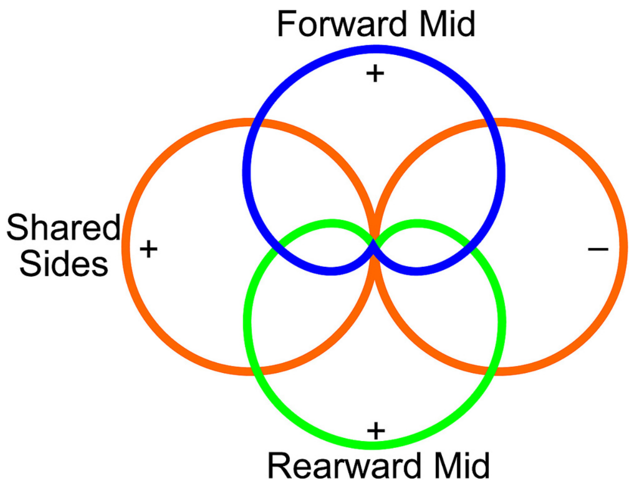


Image 5 - Double Mid Side Recording. Retrieved on 15th November, 2020 from:
https://www.soundonsound.com/techniques/double-mid-sides-array

Another widely used technique for surround sound recording is Decca Tree Surround (DTS) which is derived from the AB and Decca Tree recording techniques. Decca Tree Surround expands on the common recording technique of the Decca Tree. Decca Tree Surround takes the three frontal omnidirectional microphones of the standard Decca Tree setup and adds an additional two rear facing omnidirectional microphones. This gives full 360° horizontal coverage around the listening position. The frontal microphones are used to capture the main direct sound and the rear microphones are used to capture ambience of the space (DPA, 2018). In post production, each microphone signal is sent to its own discrete speaker in a 5 channel surround sound setup. The channel separation can be increased by changing the rear microphones to a cardioid pickup pattern. By doing this, however, the overall surround coverage can decrease and leave gaps in the soundfield between the rear and front facing microphones. By decreasing the distances between the rear and front microphones, the issue of gaps can be removed. The Decca Tree Surround technique has a lower rate of localisation when compared to coincident arrays (DPA, 2018) such as Double Mid Side, meaning that it may be more suited for capturing general room ambience where tight localisation is not as important. Image 6 is a promotional picture of the DPA Decca Tree Surround Mount.

Image 6 - DPA Decca Tree Surround Mount. Retrieved on 15th November, 2020 from:
https://www.dpamicrophones.com/accessories/surround-decca-tree-mount

Both of these techniques are examples of “Combined Front and Rear Arrays” (Los Sanderos Studios, 2018) where they are designed to capture both the frontal direct source as well as ambience in the rear.

To contrast this, The Hamasaki Square technique is an ambience array that is used to capture the ambience of the space with minimal direct sound of a concert hall (Hamasaki & Koichiro, 2003). This setup consists of four figure-of-eight microphones facing outwards with the microphone nulls pointing forwards towards the sound source. Image 7 shows an overhead diagram of this. This minimises the capture of direct sound while maximising the reverb capture of the space. As this setup is used to capture indirect sound, it would normally be combined with another surround sound capture array such as the Double Mid Side or the Decca Tree Surround.

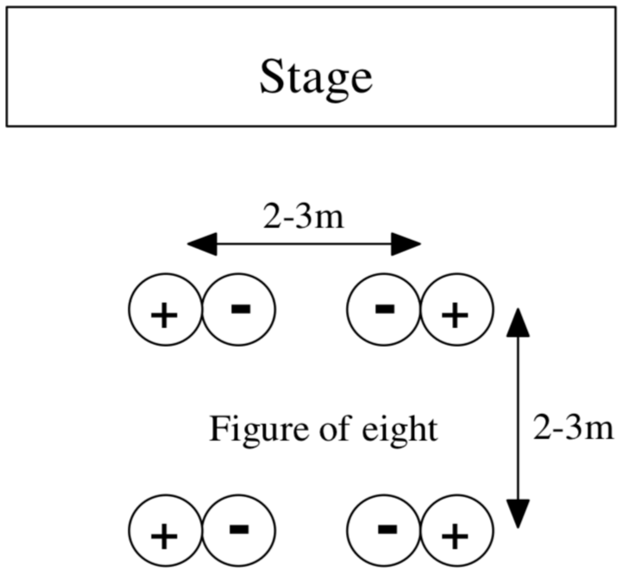


Image 7 - Hamasaki Square Setup. (Hamasaki, 2003, pp. 4)

The techniques listed above offer a variety of advantages and disadvantages in terms of capturing a realistic impression of the space, ambience and localisation. These difference became important to take into consideration when making selections for usage in the practical projects. There are many other surround sound recording techniques such as the Ideal Cardioid Arrangement that uses five cardioid microphones in a similar setup to the Decca Tree Surround, but the distances for the rears can be adjusted to change the ambient response. Other techniques are not discussed in this thesis as they were not used for any of the recording projects.

2.2.3 - 3D Recording Techniques with Height Channels and Ambisonics

Similar to surround sound recording techniques, the height channels are generally used for envelopment rather than localisation of sound due to it being harder and less accurate to localise sounds coming from the rear and above. Corresponding to stereo and surround sound techniques, there are examples of both non-coincident and coincident (ambisonic/binaural) techniques. These height channels do not have to be reproduced exclusively onto a 3D speaker system with speakers above the listener ear height and are commonly introduced on to a 5.1 surround or stereo system. When adding additional signal microphones however to capture height, care must be taken to make sure no excessive comb-filtering or strange artefacts are being introduced (Bowles, 2015). Due to the additional channels, these phase and comb-filtering issues can be easily introduced into the recording.

Non-coincident techniques, or spaced arrays, have spacing between the microphones which can reduce the crosstalk between the signals. Similar to when recording in surround sound, having microphones spaced too far apart can lead to a “spotty” image with gaps. By placing an upwards facing cardioid or hyper-cardioid microphone coincidently with the microphones of a Decca Tree Surround array height perception can be achieved. However, as Bowles (2015) discusses, due to the lack of time arrival, differences between the two signals can be difficult to discern. PCMA-3D proposed by Hyunkook Lee (2014) uses 5 coincident pairs of forward facing cardioid microphones plus a super-cardioid microphone pointing upwards to capture the height. The super-cardioid microphone reduces the capture of direct sound in the height channels leading to a greater separation between the two. Additionally, Lee has experimented with increasing the distance between the two coincident microphones for the rear channels to make them near coincident and thus introducing some decorrelation between the two signals (SCHOEPS Mikrophone, 2018, 32:07-32:45). Similar to PCMA-3D, Bowles (2015) proposed his own array that uses 4 omni microphones in the front left, front right and rear positions, a directional microphone in the centre position and four hyper-cardioid microphones coincidently mounted with the omni microphones.

This technique requires M/S processing at each coincident pair in order to split the front and rear lobes of each height microphone. The Bowles Array and PCMA-3D capture information from the source, rear reflections, as well as ceiling reflections with sufficient isolation from floor reflections without excessive comb filtering due to the coincident pairs. 2L-Cube, proposed by Morten Lindberg (2l, n.d.), uses 9 omni-directional microphones arranged in a 1m cube. Similar to PCMA, the height microphones are placed at the same position as the direct microphones but this time one meter above, providing a higher level of difference between the two signals. However, due to the spacing of the pairs, this microphone placement could lead to comb filtering when listening in a 5.1 environment but this issue can be removed when listening in an environment with height speaker channels.

As previously explained, ambisonics is a spatial format that can used to create a full sphere of spatial information. Sometimes this sphere can be created through binaural recording of a single-point array in which there is no difference in the time arrival on all axises. One example of this is recording though a binaural head such as the Neumann KU100 Binaural Dummy Head. This device captures two channels, one of the left ear position and one of the right, leading to an immersive binaural listening experience. One disadvantage of this recording method is that the recording must be played back on headphones only, due to the need for independent playback of the left and right channels without any crosstalk.

Different order microphones exist such as the Sennheiser AMBEO VR and the Soundfield microphones that records in 1st order ambisonics. The Sennheiser AMBEO VR microphone, similar to the Soundfield Microphone, has capsules arranged in a tetrahedron. It records into A-format which then is decoded into a second set of signals in B-Format (W, X, Y and Z) (Zotter & Frank, 2019). As this microphone only captures in 1st order ambisonics, the overlap between the channels is excessive. Bowles (2019) explains that because of this overlap, localisation of direct signal is harder when compared to a surround sound array with spacing between the microphones (such as Decca Tree Surround). Bowles discusses that these single point arrays are an “excellent choice for where portability and complete mixdown compatibility are desired such as recording outdoor soundscapes or background ambience for film”. These 1st order Soundfield Microphones have the advantage of being phase coherent unlike the 3D spaced arrays such as 2L-Cube.

For greater accuracy in direction and perception, microphones have been developed to capture in higher order ambisonics. The em32 Eigenmike by mh acoustics is capable of capturing fourth order (25 channels) ambisonics (mh acoustics, 2017). By using signal processing in either the monitoring or post-recording stage, specific directions of the soundfield can be focussed in on, which can lead to many different options in post production. This can range from microphone array simulation, acoustic echo cancellation, or highly accurate source localisation. The higher order capture leads to a much more accurate representation of the space that can then be decoded to different systems such as binaural headphones or large speaker systems.

Some of the surround and ambisonic recording techniques mentioned above were assessed during tests where the different arrays were placed at different distances from a speaker to draw creative decisions to be used for the final recording sessions for each project. The following individual chapters for the projects go into more detail for the multi-channel recording techniques used.

2.3 - Multi-Channel Mixing

This section discusses the use of multi-channel audio in regards to studio mixing for 5.1 surround sound, ambisonics and the technologies surrounding these aspects. Each section will draw on example of multi-channel commercial mixes as a framework for discussion of the potential challenges with mixing in multi-channel, the use of different mixing techniques and how space is used to enhance the music.

2.3.1 - The Use of Space in Production

This section discusses the use of space while mixing in regards to the sound stage, authenticity, functional staging and sonic cartoons in music production.

Moore (2001) introduced the idea of the ‘sound-box’ where sound textures are organised in the three dimensions of depth, width and height from the listeners’ point of view. In this model, sound textures are mentally placed by the listener in their relative position to auditory cues. For instance, most rock music purposefully creates a sense of musical depth through the illusion of sounds originating from different distances and a sense of horizontal location to provide the stereo image (Moore, 2001, pp. 121). The ‘sound-box’ concept can be quite limited in terms of analysis of music production due to how it only deals with spatial formation of recorded sound, but it is a good starting tool to establish the perceived width, depth and height of sound (Kraugerud, 2017).

Dockwray and Moore (2010) discussed the idea of a “normative mix” by analysing music released between 1965 and 1972 with the sound-box method. From the mid 1960s, mono was being shifted to stereo meaning there was more scope for the placement of sound and gave more range of sound-box configurations. In this analysis they analysed music and discussed different sound-box configurations until the “normative positioning mix” was established in 1972. In the analysis they categorised the songs into four main mix configurations. The “cluster mix” relates to when the main element of the mix is separated from the other parts that are clustered in one position of the sound-box. The “triangular mix” differs from the “cluster mix” where the overall width has been increased. Two main “triangular” mix setups can vary to where there are centralised vocals and off centre snare and bass to off centred vocals and central snare and bass. The “dynamic mix” refers to when sounds are dynamically moved in the soundbox. Examples of this “dynamic mix” are early Pink Floyd tracks such as *“Set The Controls For The Heart Of The Sun”* (1968, track 3) and Hendrix’s *“Crosstown Traffic”* (1968, track 3). This approach of dynamically mixing elements quickly ran out of fashion. The “diagonal mix”, which became the standard as the “cluster”, “triangular” and “dynamic” mix approaches fell out of practice, relates to where the vocals, bass and snare are often grouped in the centre and other instruments such as guitars are panned left and right. The “diagonal mix” approach became standard after 1972 and is still heard in most popular music releases today. With the four personal projects submitted, I wanted to move away from the adoption of the “normative mix” and create interesting spatial profiles that could enhance the music of the recorded artists by the use of space. For instance, the mixing project for create.evolve.destroy (project 2) explored the use of a “dynamic mix” to create spatial 3D gestures of music. However, the band and I still wanted the music to adhere to general mix rules and not detract from the music. Therefore the spatial elements were based on the normative mix as a reference point and then expanded to highlighting musical elements with the use of spatial gestures. The different approach of spatial music in project 4 where the funk band surround the listener aimed to explore how separating the band could aid in clarity of the arrangement. This approach moves away from the “normative mix” of clustering sounds together and relates to the “triangular mix” approach but the end result of the funk mix further separates the instruments.

Moylan (2012, pp. 164) defines the sound stage as a “singular area occupied by all of the sound sources of the music” that is defined by two factors. The lateral width from the furthest left to furthest right sound and the dimension of depth defined by the closest sound source to the nearest. In stereo mixing, this sound stage can change throughout the music which can bring the listener to different relationships with the music but can also stay the same to establish a fixed context for the music. This soundstage can be expanded in surround sound by adding additional lateral direction in the front and back as well as additional width. In the context of my projects, the surround sound mixes of Charlie Hulejczuk (Project 1) incorporated a stable sound stage that lets you connect to Charlie throughout the pieces of music, whereas the 3D ambisonic mix of create.evolve.destroy. (Project 2) has a constantly changing soundstage where width and depth change to add impact and emphasis to the changing parts and playing styles throughout the mix. Due to the common lack of home multi-channel systems, producers tend not to push the boundaries set by stereo music mixing in terms of placement of sound and hence not explore what can be achieved in multi-channel audio (Robjohns, 2001), but given that I had available facilities, I wanted to push these boundaries.

Relevant to Project 1, Moore (2002) discusses the aspect of authenticity in terms of not what is being authenticated but rather who. In the regard of authenticity, Moore discusses authenticity in three perspectives. “First person authenticity” relates to how an audience may engage directly with the originator (or performer) where they convey the impression of integrity in what they are performing (pp. 211-4). “Third person authenticity” occurs when a performing artist represents the experiences of others. Moore explains how an artist may never be completely independent without influence from other musical artists therefore needing to draw experiences from others. (pp. 214-218). “Second person authenticity” occurs when the audience creates authenticity by basing their own experiences on what is being presented by the artist (pp. 218-20). With these perspectives on authenticity, the recording of Charlie Hulejczuk (project 1) explored how to convey his own personal experiences and his performances in an authentic way and create an intimate connection between Charlie and the listener. While Moore’s perspectives on authenticity somewhat overlap, the approach for project 1 best relates to the "first person authenticity” approach due to highly personal aspect of Charlie’s writing.

Another concept of authenticity that was considered for this project was how authentic the production decisions were. Originally, I intended the final versions of Charlie’s songs to be one take. However, due to slight performance issues and background noise, some takes had to be sliced together and audio restoration software had to be used.

“Functional staging” (Zagorski-Thomas, 2006) (Zagorski-Thomas, 2010) refers to a style of staging in music production where the treatment of sound is based upon audience expectation and the expected listening environment rather simply than the aesthetic choices of the producer. He explains that functional staging relates to the perceivable size of the space and the perception of distance rather than the lateral placement. Zagorski-Thomas (2010, pp. 255-7) gives the example of functional staging in rock music in which a contrast of aesthetics can occur. The playback of the studio albums would likely be on small speaker systems in a home environment but the audience expectation may be to be listening at a large scale venue such as a stadium. Because of this expectation, additional reverbs are often added to simulate a rock show and to recreate the sound and atmosphere of a live event. The techniques used to create a functional staging can range from realistic recreations of a concert hall to highly stylised forms of staging which could be characterised as acoustic cartoons (2010, pp. 256). These cartoons can come from emphasising certain aspects such as over-compressing or exaggerating specific frequencies and reverb effects. Similarly, Zagorski-Thomas' concept of "sonic cartoons" (2014, pp. 49-70) highlights how over-exaggerating the representation of sound can influence the listener's interpretation. Zagorski-Thomas gives the example of over-emphasising the low frequencies of rock music to create an unrealistic impression of the instruments. However, this over-emphasis has now become the norm for the production of rock music. For the recording project with Charlie Hulejczuk (Project 1), the concept of Zagorski-Thomas' functional staging is used in relation to the way that the performances and spaces were captured. The intention of this project was to accurately reproduce the recording space to give a realistic representation of the performance. In this regard, the recording locations were captured so that when played back on speakers, the listener would have a sense of a real space. There was an element of changing the levels subtly on some of the microphone arrays to emphasise certain words. This approach created the perception of realism but was not entirely a purist approach for recreating a real life concert space. Similarly, the intention of Project 4 was to place the listener amongst the funk band in a practice room scenario where the band are separated and surround the listener rather than create a wall of sound of each instrument. To contrast the concept of functional staging, the production style of create.evolve.destroy (Project 2) is more artificial in the way that space and the placement of sound is used dynamically to create artificial and non-realistic listening environments. Whereas the production style does relate to a large and epic space from the use of creating a 'wall of sound', no real life listening scenario would have individual elements of the band rotating around the soundfield.

With the above points of sound staging, functional staging and sonic cartoons in mind, each project responded to these based on the style of music in project. For instance, the recordings of Charlie in project 1 attempted to create a realistic impression of a performance and to create a natural sounding reproduction of Charlie's acoustic playing. In this regard, no over-emphasising of frequency ranges or dynamics occurred so no sonic cartoons were made. To contrast the authentic production style of project 1, the 5th order ambisonics mix of create.evolve.destroy. in project 2 over-emphasised bass frequencies, had constant modulation of effects and dynamic movements of space creating an unrealistic space.

2.3.2 - 5.1 Surround Sound Mixing

In this subchapter, Surround Sound will refer to 5.1 Surround Sound which is five full range speakers and one subwoofer. All self-assessed and mixed audio was listened to on a Genelec Surround System setup in the ITU standard (±30º, 0º, ±110º).

Waldrep (2007) discusses ways of delivering 5.1 surround mixes and the two perspectives of mixing. The first perspective is "5.1 stage perspective mixing" (pp. 3) which places the listener amongst the musicians and creates a fully immersive surround mix. Similar to Zagorski-Thomas' idea of "sonic cartoons" (2014) this technique doesn't necessarily recreate the live space, but rather enhances the impact of the music being played. The other perspective Waldrep discusses is "audience perspective mixing" (pp. 6) which places the listener in the live space through placement of the direct sounds of the performance in the front LCR speakers and the reverberation in the rear LR speakers. Project 4 experimented with the stage perspective approach where alternate mixes were created to place the listener amongst the players.

There is a lot of debate surrounding the proper use of the centre channel. Some argue that its primary use in commercial music is to provide "hard centre anchoring" (The Recording Academy's Producers & Engineers Wing, 2004, pp. 4-5) of the key components such as the lead vocals or solo instruments. By placing signals into the centre channel, the placed sound has more stability than if one was to place it in the front left and right speakers to create a phantom centre (Moylan, 2015, pp. 250-2). This idea was born from the role of the centre speaker in cinema-based surround systems in which dialog is kept mostly in the centre speaker to improve clarity. Others recommend to not rely on the centre channel as most home systems have a small centre speaker leading to a compromise in frequency and level (White, 2002). In the context of Project 1, Charlie's vocal is sent to the centre channel instead of the front left and right for stability in the image and clarity.

In most commercial multi-channel mixing, the rear speakers are predominantly used for ambience and effects returns (The Recording Academy's Producers & Engineers Wing, 2004, pp. 4-6), but some producers use the rear speakers to place musical content for the "stage perspective mix" which favours musical experience over realism (Waldrep, 2007, pp. 3). Another usage for them is to create a sound that appears to "float directly in front of the listener's face" (The Recording Academy's Producers & Engineers Wing, 2004, pp. 4-6) by duplicating the centre channel into the rear speakers. However, when sound is placed in multiple speakers, many more phantom images are created than on a stereo system. These additional phantom images can lead to instability of imaging and confusing sonic impressions (Moylan, 2015, pp. 250–2). Image 8 shows phantom images that occur in a five channel setup. As shown, when sound is placed in the middle of circle (highlighted by "X") a blurred image could be created. This would become a greater issue in an incorrectly setup studio where there are differing level and time amounts between each speaker.

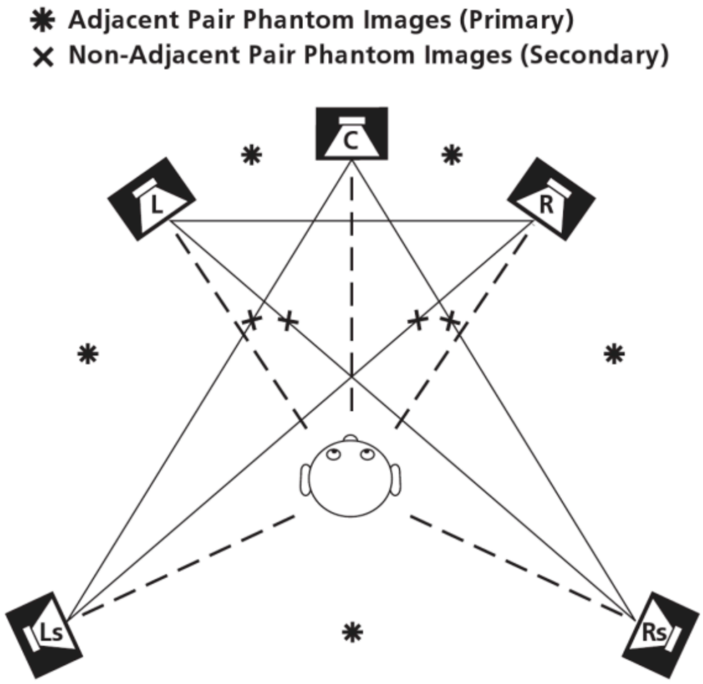


Image 8 - Surround Phantom Images. (Moylan, 2015, pp. 252)

The Low Frequency Effect (LFE) is an audio track specifically designed for low frequencies typically ranging from 20Hz to 120Hz that is sent to a subwoofer. In a speaker setup, this would be referred to as the .1 in 5.1, meaning that there are five full range speakers to one subwoofer. This is not to be confused with a typical subwoofer output on an amplifier or speaker management system where this may take a combination of all the other full range channels and output the low frequency content to the subwoofer. It is often recommended that the use of the LFE is to be approached with caution due to the irregular setup of most home systems where the room acoustics and frequency range could affect the low frequencies of a subwoofer (White, 2002). Mike Thornton (2017) believes that the LFE channel should be mainly used for effects and not for musical content so that nothing critical to the mix would be missed on incorrectly setup systems. However, in the same interview Alan Sallabank argues that the LFE channel should be used to its full effect and as an extension of sound in the low frequencies because it can add a sense of realism and immersion for the listener. These two contradicting methods make it difficult for producers to take full advantage of surround sound systems which has led to a lack of commercial releases in surround sound as a whole. Each of my projects were mixed with the impression that the listening environments would be correctly set up and reproduced on a full frequency range speaker setup, therefore the LFE and subwoofers were used to their full effect for both effects and musical content.

As each of my projects were different in terms of production goals, throughout the projects, the subwoofer was used in different ways. In Project 1, the subwoofer was not used at all due to there being no significant low frequency content. In Projects 2 and 3, the subwoofer channels were used to its maximum capacity where both critical elements of the mix are reproduced in the subwoofers as well as it being used as a low frequency extension of the bass and drum instruments. In Project 4, the subwoofer was used minimally for frequency extension of the bass guitar and kick drum.

In an article for Sound on Sound, Robjohns (2001) gives the advice to produce surround sound music “cautiously” due to the nonstandard nature of home surround sound setups in terms of correctly set levels and equal distances between each speaker. Many home setups are not correctly setup, so having intricate detail and levels in mixes is not recommended. However, in the same article, producer Rik Ede who often mixes for 5.1 surround sound says how this cautiousness is a waste of potential for surround sound. For his own personal mixes, Rik Ede describes how “there are no established rules in surround” and how he doesn’t just use the frontal speakers to create “a wall of sound” with reverbs and delays in the rears. As the intention for the create.evolve.destroy studio mix (Project 2) was to find a way in which the dynamic use of space could be used to enhance the musical ideas from the band, a somewhat imaginary space was created from not being cautious when placing the audio in the soundfield. This experimental style of using space dynamically contrasted the other mixing projects (1 and 4) where realism was the goal and was achieved by having static placement of sound and the production of surround sound was more cautious.

Reverbs can be used in surround sound to convey a realistic sense of space or to enhance the musical aesthetics. Stereo reverbs for music are often placed in the rear two channel of surround sound (Robjohns, 2001), with an added delay of around 100-200ms to add more separation between the front and rear speakers (Waves, 2017). Producer Douglas Murray (Farley, 2012) gives the advice to “have the early reflections and reverb bloom outwards from the direction of the source signal in nature” to give a realistic impression of reverb. There are surround reverb plugins such as the Waves R360 reverb (Waves, 2019) that work by taking a signal that has been positioned in a surround panner and using an algorithm to create the reverb. During Projects 2 and 4, I only used stereo reverbs that were placed in the soundfield by the IEM stereo encoder. This was because it gave me more freedom in terms of where the reverb could be placed.

In most modern Digital Audio Workstations (DAW) the software provides the user the ability to mix in surround sound with a surround panner. For example, in Logic Pro X (Apple, 2019) the built-in surround panner gives the ability to place both mono and stereo sources in a surround soundfield. For mono sources, the interface allows the user to place the sound at a specific angle, while the diversity amount defines the amount of crosstalk to surrounding speakers relative to the angle position. Image 9 is the user interface of the Logic Pro X surround panner for a mono source. For stereo sources, there is an additional width control to determine the spread between the left and right signals. There is also control for the amount of diversity for each of the left and right signals. Image 10 is the user interface of the Logic Pro X surround panner for a stereo source. This level of control is standard for DAWs and similar parameters are present in the built-in surround panners found in Ableton Live, Cubase and Pro Tools.

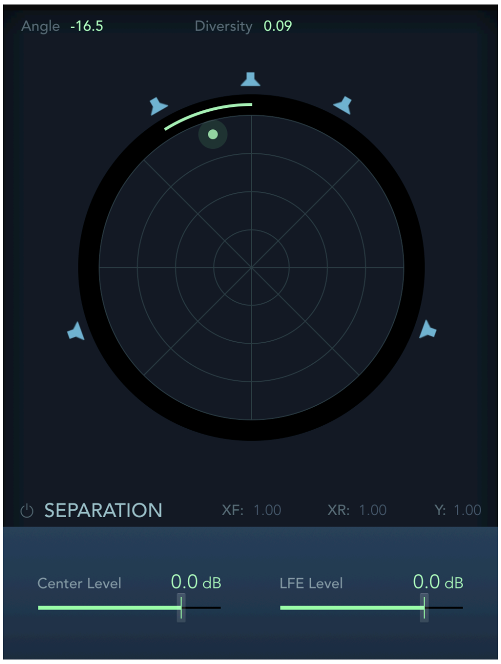


Image 9 - Logic Pro X Surround Panner (Mono Source)

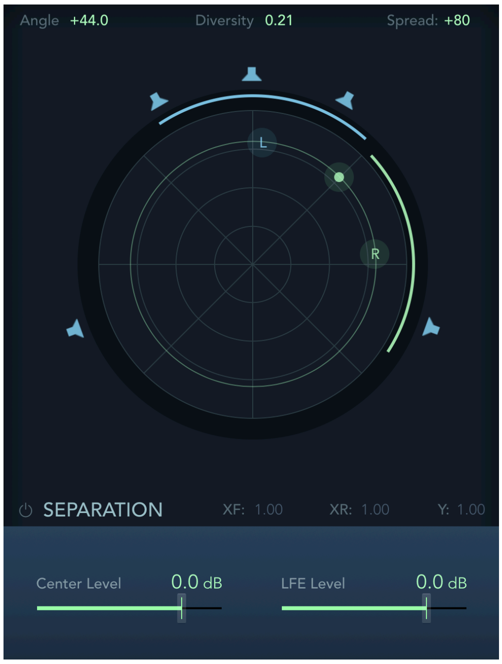


Image 10 - Logic Pro X Surround Panner (Stereo Source)

2.3.3 - Spatial Audio Mixing

Spatial Audio in this context will refer to when the output format contains height information. "Channel-based audio" (Ghanekar, 2017) is where each track is associated with a specific speaker and the content is generally made for a specific target loudspeaker layout such as 5.1 surround sound or a 22.2 speaker system with height channels. "Object-based audio" (Ghanekar, 2017) is where audio objects are independent of speaker setups. This means that the audio content can be mapped onto any speaker setup. "Object-based audio" contains the audio as well as panning information and is used in Dolby Atmos where several sound objects are delivered in a file format that also contains the placement information (Thornton, 2019). “Object-based audio” has the advantage of being more transferable to different speaker configurations but requires a decoder at the playback location.

As previously explained, Ambisonics is a spatial audio format allowing the reproduction of 3D sound in a sphere. Ambisonics allows the mixed audio to be easily decoded to different listening environments such as headphones (binaural) or spherical speaker systems where the distance between the listening position and each speaker is equal. The position of the sound is not related to the position of a speaker but rather represented as B-format signals with relative amplitude and phase information (Bleidt, R., Borsum, A., Fuchs, H. & Merrill Weiss, S. 2014, pp. 7). In regards to mixing, a common way for the user to place a sound in the sphere is to use a plugin that allows for an azimuth and elevation angle which would then be decoded to the specific listening setup. The IEM plugin suite (IEM, 2019) is a stereo encoder that can be used to place a sound and can specify the resolution order. Image 11 is the user interface of the IEM stereo encoder plugin that shows the azimuth and elevation amounts as well as the B-Format WXYZ amounts below.

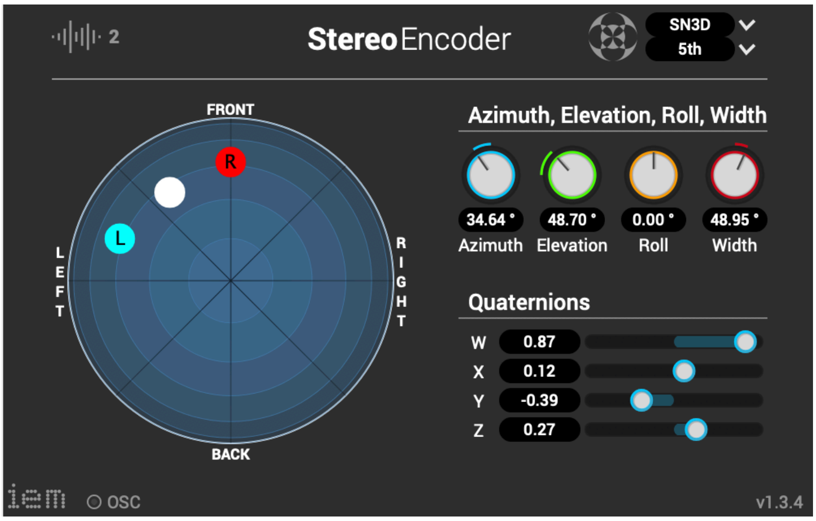


Image 11 - IEM Stereo Encoder

Dolby Atmos is an example of object-based audio where the audio files contain metadata that describes the positioning of those audio files. Dolby Atmos allows for 128 audio tracks (plus the previously mentioned metadata) where each can either be assigned to a transitional format speaker (like channel-based audio) or to an audio object that will then be positioned. By default, Atmos is set up in a 10 channel format containing seven full range speakers at ear height, one subwoofer and two height channels (Dolby, 2016). This setup would be written as 7.1.2. Dolby recommend using at least a 5.1.2 system, but currently setups can be as large as 24.1.10. To make this technology more available to the general public, Dolby has created speakers with up firing drivers that are aimed upwards to reflect sound off of the ceiling to give the impression of height. This method removes the requirement for mounting speakers directly in the ceiling. In the production phase, the audio objects are given an apparent source location represented by a 3D rectangle with the co-ordinates of the speaker positions. During playback, the Dolby Atmos system renders everything in real-time and places the audio objects relative to the given speaker positions (Dolby, 2014). This real-time rendering allows the flexibility of playback to different speaker systems.

Spat Revolution from Flux:: in collaboration with ircam is an object-based spatialisation tool that can be used for many output formats from stereo, to 5.1, to a scalable Atmos system or to Higher Order Ambisonics. The software allows for the input to come from any type of multi-channel order such as Channel-Based Audio, A/B format and Higher Order Ambisonics. Because of these options, a mix using Spat Revolution can be used for many different formats meaning that it is very versatile. Sabuni Cannone in an interview (Niklasson, 2020) explains how the object-based system allows for a greater “freedom in movement of sounds” and that a “mix in 5.1…will work in 7.1, in 20.2” meaning that it’s a “huge time saver” and easy to “explore new creative ideas”. Spat Revolution has the advantage over other spatialisation softwares such as the IEM plugin-suite as it can be used in both studio and live situations. In the studio setting, the software works as a DAW send and receive and in live scenarios audio can be sent over different protocols such as MADI and Dante to a computer for low latency spatialisation to then be sent to the speakers directly.

Wave Field Synthesis (WFS) is a spatial audio rendering technique used to simulate virtual acoustic environments through the use of a large number of loudspeakers. The main aim with WFS is to accurately place sound in the virtual space which can then move through space in many possible defined spatial pathways. By placing speakers incredibly close together, sound is placed physically rather than the more standard approach of creating phantom images between speakers. For a true, non-aliasing for the full frequency spectrum of human hearing, the speaker drivers would have to be around 2cm apart. However, Spors & Rabenstein (2006) state that a distance between 10 and 30cm has proven sufficient for reproduction due to the lack of sensitivity in spatial aliasing from the human ear. The high amount of speakers allows for a more stable image and drastically reduces a sweet spot in the listening room. Due to the way that WFS simulates acoustical environments, the acoustics of the listening area must be suppressed so that the audio coming from the speakers is not influenced by the listening environment.

Depending on the placement, these speaker arrays can sometimes involve decorrelation. Decorrelation is defined as a "process whereby an audio source signal is transformed into multiple output signals with waveforms that appear different from each other, but which sound the same as the source" (Kendall, 1995, pp. 71-87). Decorrelation has a dramatic impact on the perception of sound imagery. Decorrelation minimises the change in sound imagery when the material is moved from one reproduction setting to another (Kendall, 1994, pp. 319-26). Decorrelation was not used throughout any of the projects but would be something I would consider for future projects.

2.3.4 - The Use of Space in Commercial Multi-Channel Examples

This section will assess commercial multi-channel mixes in the format of surround sound CDs and DVDs, re-mixes/re-masters of albums that were up-mixed from stereo to surround sound as well as ambisonic pieces where available. While some albums are either specially released or re-released in surround sound, there is a distinct lack of readily available surround sound music mainly due to the lack of correctly set up personal listening systems.

Steven Wilson commonly re-releases mixes of albums in 5.1 that are generally praised by fans and critics (Bacon, 2019). His approach is to take the original recording stems and firstly recreate the stereo mixes accurately and then begins to mix in surround. Since starting in 2009, Wilson has remixed many classic albums from artists such as King Crimson, Jethro Tull and Tears for Fears. In 2014, Tears for Fears re-released their album Songs from the Big Chair (Tears for Fears, 1984) in 5.1 surround format (Tears for Fears, 2014). In an interview, mixing engineer Steven Wilson discusses how the original 1984 tape analog multitrack was used to accurately reproduce a new stereo version to keep the sonic quality as close to the original 1984 release, before being expanded to 5.1 (Sinclair, 2014). The re-release of the album brought additional separation which was absent in the original stereo mix, as mixing engineer Steven Wilson was able to place certain elements in the rear speakers. This personal comparison was achieved by comparing both the original stereo and surround sound versions from the 2014 issue. For instance, in *"Everybody Wants to Rule the World"* (2014, track 3), the backing vocals are positioned in just the rear left and right speakers, meaning they are much more audible and clearer than in the original 1985 mix. This re-release is in both the "stage" and "audience perspective mix" (Waldrep, 2007) categories previously discussed, as it mainly utilises the front three speakers for the direct sound as described in the Audience Perspective Mixing, but does occasionally use the rear speakers to fully immerse the listener. The LFE channel is used sparingly as just a slight extension of the low frequencies that are present in all other full range speaker channels. Spatially, the surround sound mix is still generally frontal with the surround speakers being used for effects such as reverbs and delays as well as non-critical musical parts such as secondary guitar parts and backing vocals which can be particularly heard on *“Shout”* (2014, track 1). This mix of separation and mix of stage and audience perspective mixing uses the 5.1 medium to its full potential and made the album re-release feel fresh.

Often for DVD releases of concerts that are mixed for surround sound, the rear speakers are only used for envelopment rather than for localisation of direct sound. Throughout Hans Zimmer’s Live in Prague (2017) and Kraftwerk 3-D (2017), the rear speakers are dedicated to crowd sound and reverb whereas the front left, centre and right speakers are used for the main instruments. The centre channel was used in conjunction with the front left and rear speakers where often a dry solo instrument would be directly sent here. This was particularly apparent during Live in Prague (2017) where on *“Pirates of the Caribbean”* (2017, track 6) a dry solo violin is placed on its own in the centre channel and the ensemble and reverbs are placed in the other frontal speakers. From listening and referencing these concerts I aimed to not use the very back of the image for any lead instruments as I felt that placing instruments away from the frontal image may cause a distraction when realism was the goal. For instance, for The University Funk Band mix (Project 4), I chose not to place any sound directly behind the listening position and used the front for the lead vocals and rhythm sections.

In regards to ambisonic musical mixes, there are not many which are publicly available. However, lots of binaural music is released publicly because of the easy delivery format of just two channels. Yosi Horikawa experiments with the use of capturing a reproducing sound in surround and has released EPs and albums intended as binaural listening. His EP, *Wandering* (2012), uses binaural audio to achieve an immersive mix using real life sounds. This EP uses 360° of the horizontal plane with no height. This is particularly present in *“Letter”* (2012, track 3) which is based around the sound of a pencil on paper that rapidly moves around the entire 3D soundfield. As higher frequencies localise much better than lower frequencies (The Recording Academy's Producers & Engineers Wing, 2004), the pencil movements can be accurately positioned in the binaural soundfield. Similar to Tears for Fears’ *Songs from the Big Chair* (2014) surround mix, the most important instruments such as bass and percussion are kept central but *Wandering* (2012) is far more experimental with the use of space and creates a much more immersive experience and compliments the composition and intentions.

The production of popular music in ambisonics is still quite limited with little to no commercial releases and there is little research into this area. Malecki, Piotrowska, Sochaczewska &

Piotrowski (2020) conducted a case study into producing electronic music for ambisonics. In this case study, they used 5th order ambisonics to place stereo stems in the soundfield. For the mix, Malecki et al. intended to create a similar listening experience as the stereo mix. Because of this, they adopted a frontal mix approach with no “gravity” towards one specific point in the sphere and avoided source movements. During the listening tests, they concluded that the ambisonics version had more preference in regards to spaciousness when compared to the stereo mix. They additionally highlighted that the use of ambisonics in electronic music could be used to bring it to “another dimension”. For the ambisonic mixing of create.evolve.destroy. (Project 2), I adopted a similar approach of the spatial elements not becoming distracting for the listener and detracting from the composition. With this in mind, there is no direct “gravity” to one specific point in the soundfield and there is a balance between the left/right and front/back present throughout the mix. However, I opted to use space dynamically to add musical gestures and interest to their music. I was careful not to overcomplicate any of these movements and have them become distracting but instead highlight parts of the composition. For example, these spatial gestures could highlight arrangement section changes in the music and accentuate certain musical moments such as drum fills or solos. The use of ambisonics for the style of create.evolve.destroy. worked well due to the complex and experimental nature of the band.

Other research into popular music production in ambisonics is limited. Sebastien Lavoie who recently completed a PhD at The University of Huddersfield (Lavoie, 2019) in his personal practice of live performance in spatialising electronic dance music for various formats including binaural ambisonics. His research is from the point-of-view of a performer using tools to create spacial movements live for electronic music. James Bagshaw is currently conducting a PhD at The University of Hull where he is researching into popular music production for ambisonics. For his research he is discussing workflows on the creation of popular music in ambisonics in regards to the recording and mixing options and boundaries (Bagshaw, 2019).

2.4 - Multi-channel Audio in Live Situations

This section explores the ways that multi-channel audio is used in live situations, this includes the discussion of how multi-channel speaker arrays were originally created for live playback and how multi-channel audio is used in live scenarios today.

2.4.1 - History of Multi-channel Speaker Systems in the context of Electroacoustic Music

Whereas commercial PA systems have been using either a mono or stereo setup for live events since the 1920s (insure4music, 2017), the earliest uses of multi-channel sound in a live situation dates back to 1951 (Fielder, 2016, pp. 3). Composers such as Pierre Schaeffer and Pierre Henry pioneered technological advancements through musique concrete that allowed their compositions to be played back on multiple speakers that expanded on commercial PA systems. Schaeffer and Henry played back recordings on magnetic tape through four speakers placed throughout a room. Their speaker configuration consisted of two speakers in front of the audience (left and right), a speaker at the rear centre of the audience and the final speaker above the audience projecting downwards creating a tetrahedral shape. The magnetic tape was separated into five tracks. The first four tracks were each sent to its discrete speaker (Track N to Speaker N) and the fifth channel could be switched between the four speakers dynamically (Fielder, 2016, pp. 4). The setup was expanded on in 1956, where Karlheinz Stockhausen used the same array but with an additional speaker placed in the rear to create a rear stereo image, creating an immersive quadraphonic surround image (Smalley, 2000).

Another significant development in the world of experimental music occurred in Brussels at the 1958 World's Fair where 425 loudspeakers were placed around the building. Composer Edgard Varèse composed the electronic piece "*Poème Électronique*" to be diffused throughout all 425 loudspeakers (Cogan, 1991, pp. 26-35). Other notable advancements in multi-channel audio for live reproduction include François Bayle's "orchestra of loudspeakers" (Desantos, 1997, pp. 11-19) in 1973. This large array of loudspeakers notably varied in size and therefore power and frequency response. Because of this, compositions could be diffused across the various speakers which would change the colour of the sound due to filtering and differing frequency responses. Experimental setups like these paved the way for future advancements in multi-channel speaker arrays.

More recently, the Birmingham ElectroAcoustic Sound Theatre (BEAST) was developed by Jonty Harrison in 1982. The concept of BEAST derives from Bayle's idea of the use of differently sized speakers with contrasting frequency responses and characters. Through angular variation (i.e. speakers facing different directions such as facing towards or away from the audience) and vertical suspension, the 100 speakers grant the composer a wide variety of choices and opportunities to playback their music. The ZKM Sound Dome in Karlsruhe, Germany consists of 47 Meyersound speakers arranged in a dome that allows "sound movements to be realistically represented from any place in the hall" (ZKM, 2020). Due to the dome shape of the speaker array, compositions in ambisonics can be created at the ZKM. Composer Fernando Lopez-Lezcano used this space to compose a full 3rd Order Ambisonics piece in 2014 (Lopez-Lezcano, 2014). The Sonic Arts Research Centre (SARC) (qub, 2020) is another diffusion system which has loudspeakers placed above as well as below the listening position to create a full 3D speaker system.

The Huddersfield Immersive Sound System (HISS) was formed in 2008 at The University of Huddersfield. The HISS is predominantly used for stereo diffusion concerts and started as the “shoe-box” (Fielder, 2016, pp. 11) setup with a central punch speaker in addition to also feature several “colour” speakers. These indirect speakers would point sound at the walls of the venue. These “colour” speakers have a specific frequency response to give direct characteristics like those used in Bayle’s “orchestra of loudspeakers”. These include car stereo speakers and Bellecour omnidirectional speakers. Later in 2017 the system was expanded by adding D&B e8 speakers were used as mounted speakers on a truss system to add the impression of height. The system today features over 65 discrete speakers for stereo diffusion concerts and live performances.

Sazdov, Paine & Stevens (2007) state that engulfment is perceived as feeling covered in sound rather than just being enveloped or surrounded by sound. In this context, Sazdov et al. defines engulfment as a spatial attribute that is unique to 3D sound where the listener is “covered in sound”. While envelopment could be explained as sound surrounding the listener in 2D, engulfment would be used to explain how sound is heard in 3D. In the listening tests he conducted, it was found that perceptually, engulfment was higher when spatialisation scenes were presented on elevated speaker systems with multiple planes rather than horizontal only. From this, Sazdov also concluded that the additional height planes “presents the composer with extended compositional possibilities” (Lynch & Sazdov, 2011).

The speaker array I used for the create.evolve.destroy. concert was based on the HISS but was modified to fit my needs and requirements. Instead of using “colour speakers” all speakers were directional point source. This was due to there not being a requirement during the concert and also due to the Yamaha CL5 console being limited by processing channels for frequency coloration.

2.4.2 - Types of Multi-Channel Speaker Systems

Multi-channel speaker arrays that are created using eight loudspeakers can be categorised into two main types of systems. The first of these is referred to as the “double diamond” setup (Fielder, 2016, pp. 11). Double diamond refers to the two diamond shapes created by the two sets of four speakers shown in Image 12 below. This setup consists of eight loudspeakers that are placed equidistantly with a front and rear centre speaker. This setup is similar to the surround sound setups of 5.1 and 7.1 due to it having a frontal centre speaker. This setup is ideal for creating circular movements with mono audio (Mooney, 2005, pp. 226).

The other setup is referred to as either the “shoe-box” setup or commonly referred to by Jonty Harrison as “big stereo” (Fielder, 2016, pp. 11). This setup also consists of eight loudspeakers but placed in four stereo pairs (front, wide, side and rear L-R). The “big-stereo” setup is ideal for the diffusion of stereo pieces that require clear stereo imaging. Image 13 below shows the “big stereo” setup and the set of four stereo pairs that mirror each other.

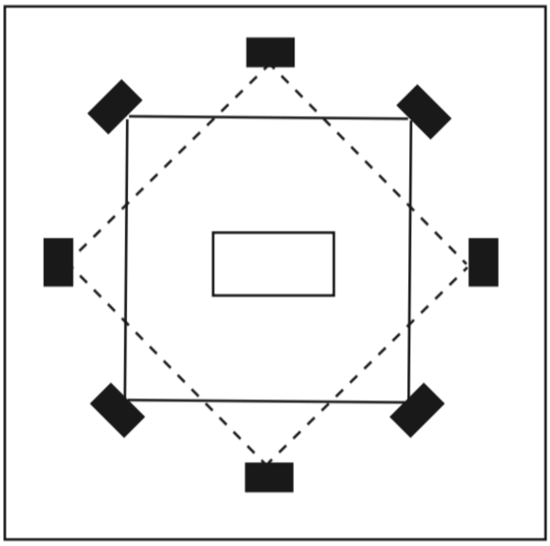


Image 12 - Double Diamond Setup
(Fielder, 2016, pp. 22)

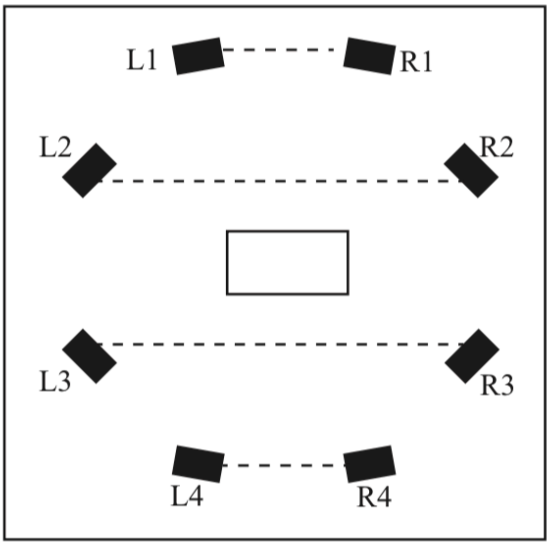


Image 13 - Shoe-Box (Big Stereo) Setup
(Fielder, 2016, pp. 22)

For the concert, I chose to use a modified version of the “shoe-box" setup with an additional centre speaker to fill in the gap for the frontal image. This type of system was chosen so that certain channels such as the electric guitar amp could be positioned with a wide stereo image while having the bass and drums be anchored in the middle with the centre speaker. This would have been more challenging to achieve with the “double diamond” setup.

2.4.3 - Commercial uses of Multi-channel Speaker Systems

Outside of the world of diffusion for electroacoustic music, multi-channel arrays are also used for live commercial music events. The first of these was Pink Floyd's "Games for May" concert in 1967 at the Queen Elizabeth Hall in London, England. The concert utilised a quadrophonic loudspeaker setup with a controller dubbed the "Azimuth Co-ordinator" made by Bernard Speight at EMI's Abbey Road Studios (Cunningham, 1997). The "Co-Ordinator" was controlled by keyboardist Rick Wright wherein sound could be moved from speaker to speaker. It worked by using four potentiometers (one for each speaker) that were controlled using a single joystick (Calore, 2009). Throughout the set, sound effects such as footsteps, backwards cymbal crashes, and Roger Waters' manic laughter were moved throughout the hall. This setup was expanded on in 1969 with a new "Azimuth Co-ordinator" that featured two joysticks to position sound on to four and six channel speaker setups. This new controller allowed Rick Wright to move two sounds at the same time. Drummer Nick Mason described the show as "…one of the most significant shows we ever performed" (Calore, 2009). This azimuth controller was the inspiration for using a PlayStation 4 controller to pan sounds in the 3D concert of create.evolve.destroy. (Project 3). Using the triggers I had the ability to change what channel on the desk to position and the joysticks allowed for 360° placement.

Since then, other artists have implemented multi-channel setups in their live situations. Notably, this includes Frank Zappa's "Yellow Shark" concerts in 1992 which used a six-channel speaker setup. Lehnert (1994) described how the overall purpose of using the multi-channel setup was to make the audience feel "(electronically) surrounded by the musicians". The music from the Yellow Shark album was specifically designed for a six-channel speaker system for this audience immersion but also to help create greater separation of the complex instrumentation (Rundel, 1993). Audience members reported experiencing a clearer sound compared to that of standard stereo speaker arrays (Michie, 2003). Separation of instruments was used in the live concert to allow the audience to hear the complex rhythms and melodies.

More recently, Hans Zimmer Live (2016-17) used a quadrophonic array for their arena tour. Many pieces in the second set such as The Dark Knight Suite and the Interstellar suite were augmented through spatialisation of sound effects and electronic tracks. Live engineer Colin Pink operated the system using the two inbuilt joysticks of the DiGiCo SD7 mixing console. This controlled the level of the tracks in the two rear left and right speaker hangs (Duff, 2017).

In addition, Kraftwerk have been touring with a full 3-D surround sound system that utilises height channels as well surround sound since 2012. Their tours deploy a 3D speaker system using d&b's Soundscape (d&b, 2019). The d&b system allows the four members of the band to access the spatial renderings of the PA system from stage using a MIDI link. When reviewing the event, Curran (2017) described the concert as a “breathtaking exhibition” of musical performance as installation art.

I took the idea of using controllers to move sound around for the live concert of create.evolve.destroy. By taking influence from the Azimuth Controller that Pink Floyd used for their quadraphonic show as well as the surround pan controllers on the DiGiCo SD7 which enabled quick panning movements, I designed a software controller in Max that could control the Yamaha CL5 mixing console to change the position of each input channel independently. This was supplemented with a PlayStation controller to be able to dynamically position two channels at the same time.

2.4.4 - Current Technologies Surrounding Multi-Channel Events

As well as d&b Soundscape, a recent technology from L-Acoustics titled “L-ISA” enables artists and mixing engineers to deliver multi-dimensional sound experiences to audiences of all sizes. The usual deployment of this setup is to have five frontal speaker hangs at the front of the performance space which allows a wider and more accurate representation of amplified sound relative to the stage when compared to a regular stereo PA system (L-Acoustics, 2019). Having more speaker hangs in a concert venue creates a great sound consistency throughout the venue. This means there would be a more consistent loudness (±6dB) and frequency response across the venue space compared to that of a standard stereo PA system. The system can be expanded from a normal frontal image by adding surround systems, thus creating 360º horizontal coverage as well as height speakers overhead for full 3D. The system is controlled using a software panner that allows the user to change the parameters of the panning, width, distance and height of discrete channels independent from one another. As the system is scalable from small studio setups to arenas, pre-production is often done in a studio setting and then only minimal tweaks are required when scaled to a full arena 3D system (L-Acoustics, 2019). In 2018, the full L-ISA surround system was used at the Royal Albert Hall by Alt-J. The system featured an array of three central speaker hangs with two additional hangs used as the main frontal speakers. To deal with the surround sound elements, ten additional speaker systems were placed around the hall (Campos, 2018). In a review for the event, Richards (2018) described the event as “even more all-encompassing” and the system allowed the band to “enter newer, weirder and more experimental places". The L-ISA system continues to be used by many other artists such as Bon Iver, Childish Gambino and Ennio Morricone. While many tours are deploying these surround sound speaker arrays, various venues are permanently installing them. The d&b Soundscape system is now installed at Leeds Playhouse, UK (d&b, 2020) and the L-Acoustics L-ISA system has been recently installed at EartH (earthackney, 2020) in London, UK.

As previously explained, Flux Spat Revolution developed with ircam can be used as a tool for spatialising audio for different listening scenarios such as on headphones and speaker systems. Spat Revolution can also be used in live situations as well. Similar to the design of L-ISA and Soundscape, Spat Revolution uses the same design of assigning speaker positions and then placing the sound in relation to the speaker positions in an object-based setup. Spat Revolution also allows for the configuration of multiple "rooms" (flux.audio, 2018) meaning that it can support multiple speaker setups such as 5.1, binaural and many other speaker setups at once. For the 2018 production of Cirque du Soleil, they used Spat to recall a standard stereo speaker system, a 5 channel directional surround system for localisation as well as a circular surround system for immersion (FLUX:: Immersive, 2020, 41:40 - 46:30). The software can also take in multiple different input streams that can be translated to the different "room" configurations which includes Higher-Order Ambisonics up to the 7th order. Spat Revolution also allows for various different panning techniques such as vector-based amplitude panning (VBAP), distance-based amplitude panning (DBAP) and ambisonic equivalent panning (AEP). These different panning options are useful in live situations that do not have the traditional sweet spots found in studios (flux.audio, 2020).

The idea behind these technologies is to enhance the musical content and be more experimental with space. My aim for the concert was to enhance the music of create.evolve.destroy. by placing sound to fully surround the audience. By having the 3D speaker array, I was able to have an equal loudness across the event space. In conclusion, this deployment of the 3D setup helped both musically and technically.

The following sections go into the individual projects and how they used recording, mixing and live techniques to take advance of the creative possibilities of working with space in multi-channel audio.

3 - Methodology

Candy and Edmonds (2018, pp. 63) explain that practice-based research in creative arts discourse emphasises the “creative process and the works that are generated” where the “artifact plays a vital part in the new understandings about practice that arise”.

Based on this, my research conducted into the production of multi-channel audio over the course of the four projects was practice-based. The four separate project artefacts highlight how the use of space is used in different aspects of producing multi-channel audio from a recording, mixing and live sound point of view. The research also took a creative practice and deductive approach based on Collins’ approach stages (Collins, 2010). From this approach, a recursive reflection on my work took place with repeating stages of the production process to make sure the spatial intentions of each project were met. This Collins methodology approach allowed me to determine the success of the creative outcomes of each project by critical reflection on the work from myself and the recorded artists.

The recursive reflection on the testing and spatial elements of each project maps on to Collin’s five approach stages for a deductive methodology (2010, pp. 42):

1. Writing a testable proposition with a testable concept or variables
2. Indicate how this proposition can be tested
3. Test the proposition
4. Study the outcome of the research which will either confirm the original proposition or how it needs to be changed
5. If necessary, modify and repeat the process of testing and studying until the proposition is met.

Each sub-chapter of this section will discuss how each project explores space and how the project maps on to the Collins stages. As each project has creative outcomes to produce multi-channel audio with different intentions, the methodology generally followed these stages that align with the Collins stages.

- Developed intentions and a proposition for each project by communicating with the artists and discussed how multi-channel space could be used (Stages 1 and 2).
- Conducted practical tests on the discussed proposition (Stage 3). These tests would be designed to compare different spatial elements from recording and mixing examples.
- Reviewed the results of the tests and then communicate with the artists again on possible changes required before moving on to the final products (Stage 4).

- Personally reviewed the spatial elements throughout the projects and at key moments in the projects by showing the work to the artists for their input (Stage 5).

Stages 3 to 5 were part of the recursive reflection and were repeated until the best possible attempt to demonstrate the proposition was achieved.

3.1 - Project 1 - Charlie Hulejczuk - Surround Sound Recording Methodology

The proposition for this project was that testing a large set of possible surround sound recording techniques could help make decisions to match the style and mood of contrasting songs. The testing portion of this project is borrowed from other contexts of microphone shoot-outs such as Hyunkook Lee's tests on the psychoacoustics of 3D recording (Lee, 2018). While his tests are more grounded in the scientific perception of sound, the tests I conducted are similar in the way that many variables are tested to find which techniques could best fit a specific musical purpose. The thorough testing pre-production of this project allowed me determine the spatial elements of each track before the real recordings took place and find how the mood of each track could be conveyed.

At the start of the project, Charlie and I discussed how recording in multi-channel could create this realistic impression of an acoustic performance and match the style and mood of each track (Stage 1). We listened to his demo bedroom recordings and chose four tracks that would contrast each other in style and mood that would best highlight the different spatial elements. I then designed a set of tests that would allow me to gather a large set of recordings of different microphone techniques at different distances and heights. These tests allowed me to compare the different possibilities and create "mock mixes" of each track so Charlie and I could get a spatial impression of how each track would sound before the final recording session were conducted (Stage 2). The tests that allowed me to compare different surround sound microphone techniques were then completed. In the tests, four main techniques were tested and positioned in between 3 and 10 different locations in the recording space (Stage 3). At first these were personally assessed in a surround sound listening environment (Stage 4) and repeated if required due to technical issues or if any new ideas developed (Stage 3 - repeated). These assessments determined how the different techniques could be used to convey different moods and styles in Charlie's work. After all the microphone tests were completed, Charlie and I assessed the recordings and created "mock mixes" to determine spatial impressions and how the mood of each track could best be conveyed. The final recording sessions were then organised and completed where the location and recording techniques were chosen based on the recording tests (Stage 4). The recordings were then personally assessed based on technical aspects such as spatial imaging as well as creatively assessed to see if the original moods were being reflected in the recordings (Stage 5). Any recordings were then re-done if any issues arose (Stage 4 - repeated). Once Charlie and I were happy with the recordings, mixes were completed of each track and assessed on their spatial elements until the intentions and proposition was best demonstrated (Stages 4 and 5). This cyclical, recursive reflection allowed me to produce the four tracks in different styles to best convey Charlie's compositional ideas.

While working with Charlie, I defined myself as a Collaborative Producer (Burgess, 2013, pp. 14) where I worked with Charlie and contributed to the sonic aspects of the final product. In this role I had no impact on the songwriting and explicitly left him to perform his music in his own way. We worked together at the start of the project to establish the style and mood of each track to be recorded. For the final mixes, I did combine different takes to favour better performances or to remove any background sounds that were not possible to remove with the use of audio restoration software. My ideas of the use of space in terms of the recording location, microphone techniques and balance between dry and reverberant sound shaped the final sound and mood of the four tracks. In this role, my aim was to take Charlie's ideas in songwriting and performance and be able to best convey this in space.

3.2 - Project 2 - create.evolve.destroy. Studio - 5th Order Ambisonics Methodology

The proposition for this project was that the dynamic use of space in 3D with height can be used in a more experimental manner to create non-realistic and imaginary spaces to enhance and fit the musical ideas put forward. Due to the experimental prog-rock nature of the band, this allowed for a more experimental style of the placement of sound dynamically. The common recommendation for the production of surround sound audio is to tread with caution in terms of the placement of sound (White, 2001) due to the lack of correctly set up home listening environments. Additionally, from Dockwray and Moore's (2010) discussions on configuring the sound-box, after the 1970s, the production in terms of space for commercial music adopted the "diagonal mix" approach and left behind the more dynamic use of space found in Hendrix's *"Crosstown Traffic"* (1968, track 3). This adoption of the "normative mix" (Dockwray & Moore, 2010, pp. 188) in music production has led to a lack of experimentation in using space dynamically where sounds are generally placed statically throughout the mix. The band and I agreed that the mix should use space dynamically but should still adhere to general mixing standards of a mostly frontal mix and the dynamic elements should not become overly distracting to the listener.

At the start of this mixing project, the band and I discussed the way that space could be used dynamically in 3D to create musical gestures to enhance their musical ideas (Stages 1 and 2). As the band have two contrasting styles of playing, this suggested two different spatial approaches. When the band play together, this suggested a more typical “wall of sound” approach that would be contrasted when the band play separately which would be reflected by separating the instruments in the horizontal and vertical planes. From this, I tested certain elements of ambisonics such as the order for resolution, localisation of different frequency ranges as well as height perception. I also created some demos of movements based on some basic stems sent from the band (Stage 3). These tests and demos allowed me to show the band what could be achieved and then began collecting the electronic audio and recording the guitar, bass and viola parts. In this stage, some of the original electronic audio was replaced with new recordings such as the cymbals for a more natural and contrasting sound (Stage 4). Once all of the audio was gathered, the mixing stage began where I was given complete control of the spatial elements as long as the dynamic gestures were not becoming distracting for a listener. Throughout the mixing process, these ideas were shown to the band for their feedback and the mixes were repeated until the original intentions were completed (Stages 4 and 5). Also at this stage, the mix was tested on different playback systems to see if the ambisonic mix was portable. The process of assessing the mixes with the band and on different playback systems meant that the spatial elements of the mix were constantly being assessed by myself and the band to make sure they were not becoming distracting and that the original intentions of the project were being met.

For this mixing project, I defined myself as a facilitative producer (Burgess, 2013, pp. 14). In this role of a facilitative producer I maximised the band’s writing ideas and was given room for my own personal creativity in regards to spatial and sonic mixing. The band gave me creative freedom in terms of the sonic and spatial features of the mix but within the constraint that they still wanted the mixes to be coherent and similar to their style of music. With this in mind, referencing artists such as Godspeed You! Black Emperor and Polyphia was important to make sure that sonically the instruments sounded as they should.

3.3 - Project 3 - create.evolve.destroy. 3D Concert Methodology

The proposition for this project was that off-the-shelf PA and mixing equipment that is not specifically designed for it can be used to mix in 3D for a live concert with the addition of bespoke controller software.

As this project started after the studio mix of create.evolve.destroy. (Project 2), the band and I agreed that we should attempt to recreate the ambisonic studio mix in a live setting. These similarities relate to the two different spatial impressions of creating a "wall of sound" and the separation approach depending on how the band were playing. The main task of this project was finding a way for the technology that was available to me to be used in a way to mix in multi-channel (Stage 1). There are multi-channel live technologies such as L-Acoustics L-ISA and d&b's Soundscape, these are very inaccessible for general use without extensive training. The mixing console I chose to use was the Yamaha CL5 which is limited to stereo or 5.1 surround sound and has no option for extending this to fit my requirements of a bespoke speaker array setup with height speakers. Therefore, I created a piece of software using Max to control the mixing console that allowed me to pan sounds to direct outputs that is similar to the channel-based mixing approach (Stage 2). I designed the controller in a way that any custom speaker layout could be setup, much like the IEM AllRA ambisonic decoder. This meant that I could test the controller before the event took place and the speaker array was built. I therefore tested the controller in SPIRAL Studio which let me check how accurately sounds could be moved around the speaker array (Stage 3). This testing stage was vital as it brought up issues that would have been unforeseeable before the real speaker array was built and the concert had to be done. After the speaker array was built, I then retested the controller with the new speaker positions and show ed the band the live ideas that could be achieved and get their input and suggestions (Stage 4). Here any final adjustments took place to the controller before moving on to the soundcheck phase. Unlike a typical concert, the soundcheck was an extended period of time that allowed multiple passes of each track to make adjustments to both the PA sound and their monitors. Due to the complexity of the concert and my role as a mixer for not only level but also the spatial positioning, it was important that everything was set as perfectly as possible before the gig took place so this process of repeating the tracks in the soundcheck was vital (Stages 4 and 5). The extensive sound check and the testing of the bespoke software controller allowed me to be able to review the spatial elements and make adjustments that would just not be possible in the live scenario.

3.4 - Project 4 - The University Funk Band - 5th Order Ambisonics Methodology

The proposition for this project was that placing the listener amongst the band in a stage perspective mix could help with the separation of the instrumentation in a dense arrangement of a pop funk band. Similar to Project 2, the mix was driven by creating the stage perspective mix and intended to move away from the normative mix approach but some of the tested mix positions relate to the approaches discussed by Dockwray & Moore (2010) such as the "triangular" and "diagonal" mixes.

When discussing with the frontman of the band Charlie, we thought that creating a mix where the band surrounds the listener would be a good opportunity to create a practice room layout and allow for a greater separation between the instrument parts (Stages 1 and 2). Similar ambisonic choices were made in Project 2 but as realism was the goal, Charlie and I decided that the recorded sounds would remain static and without too much elevation above ear height. The band was mostly recorded separately due to timing constraints but as all of the parts were isolated, it gave me greater flexibility in the placement of sound. This experimentation would not have been possible if the band were recorded all at once already in their relative positions. After recording, I did basic mixes and created several different spatial mixes with different positions that contrasted each other. These contrasting positions highlighted the different possibilities of the band surrounding the listener and what could possibly work in this scenario (Stage 3). These tests were then showed to Charlie and we agreed on one of the position setups to continue with into the final mix (Stage 4). The mix positions were chosen based upon being able to highlight the arrangement of the track but not being too far removed from a typical pop funk mix. After the positions were mostly chosen, the final mix was worked on and then assessed on the typical mix criteria of balance and tone as well as the spatial elements with Charlie until the original proposition was best demonstrated (Stages 4 and 5). While the perspective mix was achieved and did bring a better separation between the instruments, the end result was not as strong as the other projects due to the experimental way that space was used in this project.

Similar to the mixing project for create.evolve.destroy. (Project 2), I defined myself here as a facilitative producer (Burgess, 2014, pp. 14). Here I had no impact on the composition or arrangement but was given creative control on the overall tonal and spatial aspects of the final product. Throughout the mixing process, artists such as Marvin Gaye, Vulfpeck and Pomplamoose were referenced due to their similar arrangement styles in both the rhythm and lead sections.

4 - Project 1 - Charlie Hulejczuk - Surround Sound Recording

When listening to the final audio files or any other referenced audio, it is highly recommended that this should be done on a correctly setup 5 channel listening environment in line with the ITU specification. The listening projects are routed this way so no further setup is required.

Charlie Hulejczuk is a singer-songwriter who plays acoustic guitar and sings. His style is similar to Blair Dunlop's *"356"* (2016, track 3) and Nick Drake's *"Place to Be"* (1972, track 2). This project is a four track EP where the output format is 5 channel surround in the ITU format. For each of the four tracks, the location and microphone techniques were chosen to suit the genre and style. No additional effects other than EQ and compression were using while mixing.

My role during this project can be defined as a Collaborative Producer (Burgess, 2013, pp. 14). In this role I had no impact on the songwriting or compositional style but instead controlled the recording sessions in terms of capture techniques to change the sonic aspects of the final product. This allowed Charlie to perform his songs in his own way. The only impact I had on the actual performances was choosing which takes to use for the final mixes based on the performances and if any background noise was interfering with the recording capture. In the role as a Collaborative Producer, I aimed to shape the sound aesthetics of each track by deciding on microphone types and placements to achieve results that suited the style. The tracks that were used in this project were chosen by Charlie and myself. The chosen tracks contrast each other in their style and mood where the level of intimacy changes based on the compositions.

The outcome of this project was to create spatial mixes that matched the style and mood of each of the four contrasting tracks. Sonically, I intended the tracks to be intimate and close sounding to further match the mood of Charlie's writing and performances. These decisions were made from testing a large range of possible surround sound recording techniques at different positions in the room. Throughout the project, the spatial elements of the recordings were critically evaluated and decisions were made based upon Charlie and my intentions of each track. This recursive reflection method was borrowed from Collins' five stages of research for creative practice (Collins, 2010).

4.1 - Recording Tests

All marker references throughout the following sections relate to a Reaper project that has the different recording audio for comparison with the intention of being listened on a 5 channel surround sound speaker system. This project can be found at *USB Files/02-Charlie Hulejczuk/01-Pre-production*. Appendix 5 is the table that shows the full list of markers that are referenced within this chapter with the technique and notes describing the setup and distance measurements. For easier comparison, all audio that is referenced in this chapter are taken from the recordings of Track 3 *"Waste Our Time".* The full recording session featuring all four tracks recorded with each technique can be found at *USB Files/06-Appendices/01-Virtual Source Recording Project.*

The recording tests were designed in a way for easy direct comparison of the microphone techniques at predetermined distances and heights. It was decided that recording short snippets of each of the four tracks in a dry booth to be played back in a reverberant space would be the best way to assess the numerous techniques (Markers 1-4). This was recorded with close Neumann microphones to capture the detail of the guitar and voice and then blended together to make a mono audio file. Appendices 6 and 7 are photos from this recording session. As the same audio was played through the Genelec 8040a speaker in the same position of the room, there would be no alterations in dynamics, performance or tempo meaning that all audio files would be perfectly aligned in a Reaper project that would allow easy and direct comparison.

These recordings were then played back in St. Paul's Hall (RT60 = 2.1sec) at The University of Huddersfield through a Genelec 8040a loudspeaker as a Virtual Source Recording (VSR) (Lee, 2016). This allowed me to hear how each track would sound in a reverberant, real-life space. Below, Image 14 is the floor plan of St. Paul's Hall. The star at the left shows the position of the speaker in the room. All microphones (unless stated otherwise) were recorded using a D.A.V. Electronics BG8 pre-amp into an RME Fireface 800 at 44.1kHz/24Bit into Reaper.

Four techniques were assessed in this recording session, Decca Tree Surround, Double Mid-Side, the Sennheiser AMBEO Microphone and Hamasaki Square. These techniques were chosen due to their contrasting setups, capture of direct and indirect sound and the coincident vs. non-coincident setups. Similar techniques such as the OCT Surround and Fukada tree were not tested due to their similarities to the other techniques that would create similar results. The function of these tests was to gather audio recordings of contrasting microphone techniques to determine how the mood of Charlie's music could be reflected in the spatial recordings. The chosen techniques were tested between 3 and 10 times at different distances and heights to determine spatial impressions of technique and how the distance from the source could change the recording outcomes. The horizontal distance measurement refers to the distance between the Genelec 8040a tweeter and closest microphone capsule. The height measurement is the vertical distance between the floor and lowest microphone capsule. Additional techniques were tested such as omni outriggers and hyper-cardioid microphones facing away from the speaker, however these techniques are not discussed in detail but can be heard in the full recording project.

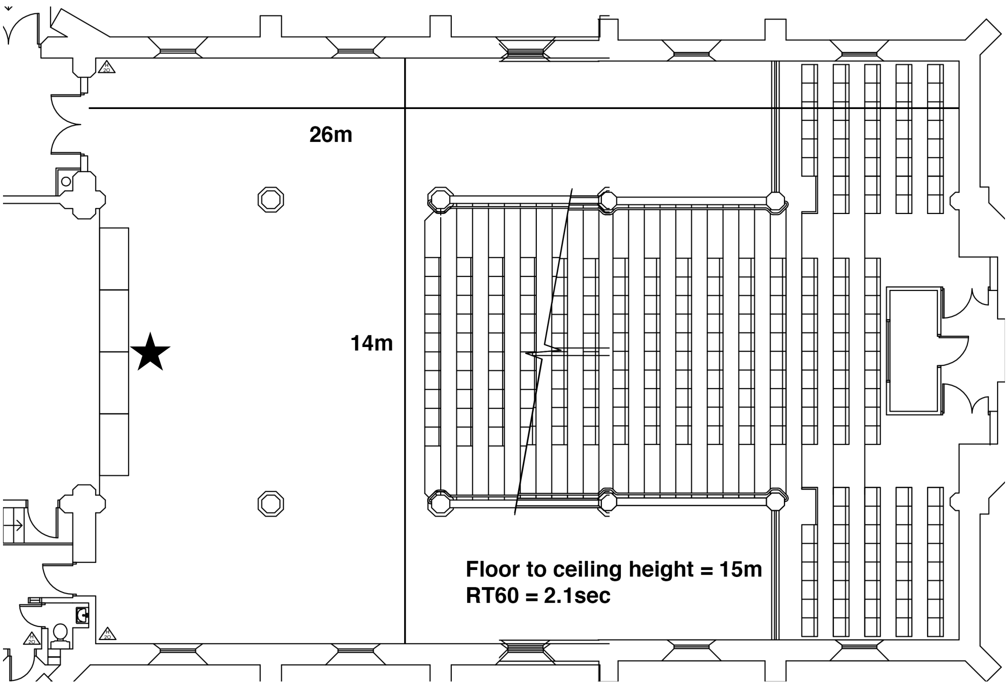


Image 14 - St. Paul's Hall Floor Plan

4.1.1 - Decca Tree Surround

Decca Tree Surround (DTS) was tested using five DPA microphones in either a full omni, full cardioid or mixed omni/cardioid setup. The omni microphones were DPA 4006s and the cardioid microphones were DPA 4011s. In the mixed setup, the three frontal microphones were omni and the rear facing microphones were cardioid. Each of these three arrays were tested at three horizontal distances (0.5m, 2m and 5m) and two heights (1.5m and 2.5m). Appendix 8 shows a diagram of the bird's-eye view of the DTS setup. Appendix 9 shows an image of Decca Tree Surround setup in St. Paul's Hall during the recording session.

For listening, each microphone speaker was sent to the equivalent discrete playback speaker.

When using DTS with five omni microphones (Marker 5), I found the capture to produce an enveloping reproduction of the recording space as described by DPA (2018). However, due to the omni microphones, there is a lack of isolation between the channels but this does mean there are no gaps in the soundfield between any of the microphones. When using DTS with the mixed omni/cardioid setup (Marker 6), I found the capture of the rear ambience to be more direct and prominent. However, using cardioid microphones for the rear capture led to a strange perceivable gap in the soundfield between the front and rear. This made it hard to blend these signals together and made the recordings less enveloping and a less realistic impression of the recording space. This effect may be removed however by moving the rear facing cardioid microphones closer to the frontal microphones to narrow the distance and remove the gaps. DTS with five cardioid microphones (Marker 7) only increased the gaps in the soundfield leading to an even less realistic impression of the recording space. Instead of there being an enveloping capture of the space, I perceived this setup to be multiple mono room microphones rather than a recording technique to create an enveloping reproduction of the space. Because of these reasons when using DTS my preference was towards using the five omni microphone setup for this project.

Additionally with the use of the cardioid microphones, the noise floor increased slightly increased and when the array was further away (5m), the signal to noise ratio was much less when compared to the closer distances. Because of this, de-noising would be required in post which could lead to strange artefacts being introduced to the audio files. To be able to reduce this, I tested the mixed omni/cardioid array with the Neumann KK184 digital microphone which did decrease the noise level in the recordings.

All other examples now show the differences between the different distance and height location using the full omni DTS setup using Track 3 *"Waste Our Time".* This track was chosen due to the louder strumming style of guitar playing that more clearly captures the ambience of the space.

At the closest tested distance of 0.5m and height of 1.5m (Marker 8), a strange occurrence happens where the centre microphone picks up the direct sound with little room reflections and the front left and right microphones are picking up a lot more reflections. Because of this, a strange image is created that has a blend of direct and indirect sound from the speaker. When the array was moved back to 2m (Marker 9) this issue decreases where the capture between the microphones is much more balanced. Moving the array further away to 5m (Marker 10) mainly captures the ambience of the room with little to no directionality or localisation. When tested at the same distances but at a 2.5m height (Markers 11 - 13), the direct capture decreases giving a more reverberant capture as expected. Based on these tests, my personal preferences led to placing the array at a distance of around 2m from the source and at a height of around 1.5m to give a good blend of direct and indirect sound so that there is still some form of localisation while giving a good spatial representation of the space.

Blending the original dry audio and the recording with five omni microphones at a medium distance gives a good balance of dry and ambience so when this microphone array was used in the final recording sessions it will be supplemented with close microphones on the guitar and voice. One main drawback of this array is the size of the frame. This means that it would not be suitable in a small space such as a small club room or for recording a live performance where sight lines from a conductor or audience are important.

4.1.2 Double Mid Side

Double Mid Side (DMS) was tested using two Schoeps CCM4 cardioid microphones and one Schoeps CCM8 figure-of-eight microphone attached to a Rycote shock mount for consistent placement and so it was easy to move and position. Image 15 is an annotated image that shows this setup.

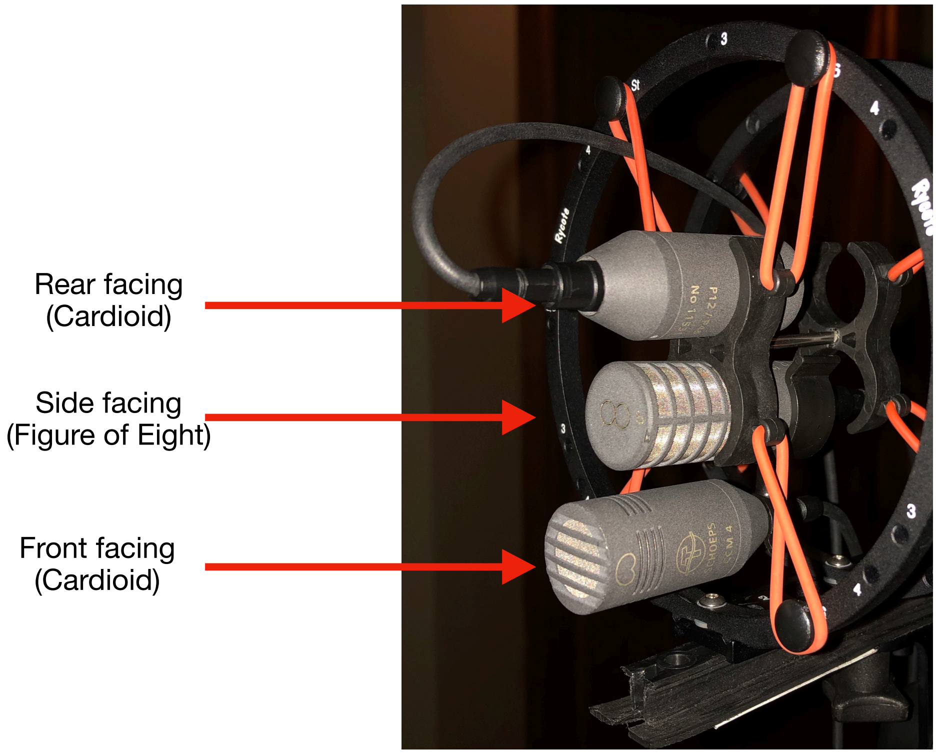


Image 15 - Annotated Double Mid Side Setup

The array was tested at 5 horizontal distances (0.5m, 1.5m, 3m, 5m and 7.5m) and at two different heights (1.5m and 2.5m) at each of those distances. An additional test was conducted in which the array was placed behind the speaker 1.5m away from the back of the speaker and at 1.5m high from floor level where the front facing microphone was pointing into the hall to see how capturing multi-channel from an indirect location would sound. Appendix 10 shows a bird's eye view of the Double Mid Side diagram.

All of the audio examples were exported to a 5 channel audio file using the default "5-ch" preset of the Schoeps Double MS decoder plugin.

At a horizontal distance of 0.5m and a height of 1.5m, this array creates a similar effect to the Decca Tree Surround array in terms of the ratio between direct and reverb capture. When moving the array to a horizontal distance of 1.5m, the image shifts to more side and rear reflections and becomes slightly darker in tone. This could be due to the positioning in the hall in relation to the side pillars. When the array is placed at 3m or further (Markers 16 - 18), the capture is very reverberant and does not display any significant perceivable differences when comparing the different horizontal distances.

While increasing the height of the array at each horizontal distance does increase the reverb (Markers 19 - 23), I intended to use this array to capture both direct and reverberant sound in the chosen rooms. Keeping this array at a height of approximately 1.5 was more in line with my aims for this project.

While increasing the height of the array at each of horizontal distances to 2.5m, the reverb capture naturally increases when compared at the lower heights of 1.5m (Markers 19 - 23). I decided that when the array was being tested at 2.5m, the overall difference between direct and reverberant sound was too different to capture Charlie's guitar and vocals in a way that they sounded like in the recording spaces. Therefore, during the recording sessions when I used the double mid side array, I placed the array at a similar height to Charlie's head height.

When the array was placed behind the speaker (Marker 24), the output is a very unusual capture in which both the front and rear facing microphone are capturing indirect sound. Without any direct signal, the capture is very reverberant and an omni microphone placed behind the speaker most likely would have yielded a similar result. For this project, I do not think this position would be suitable to yield any usable recording to localise the direct signal. This test however was useful to see how a surround microphone array would capture indirect signals even though it gave no workable outcome.

My personal preference was to use this array at a similar horizontal distance to the Decca Tree Surround array at an approximate horizontal distance of 1 to 2m away from the source. One advantage of this array is the compact size and how it only requires one microphone stand to support the mount. Because of this it can be used in smaller venues and when travelling to different places with limited equipment while still capturing an authentic and realistic representation of the space.

4.1.3 Sennheiser AMBEO

The Sennheiser AMBEO microphone was tested in this scenario to see how the ambisonic capture could be decoded to 5 channels.

For these audio examples, the recordings have been exported using the Sennheiser AMBEO plugin (Sennheiser, 2019) *to convert to B-format and then to 5 channel using the ATK FOA Decode 5_0* (Ambisonic Toolkit, 2016) *set to the focus mode.*

Overall, my perception of the Sennheiser AMBEO microphone recordings is that they are much brighter than the recordings any of the other arrays. This can be particularly heard at the horizontal distance of 0.5m (Marker 28). When increasing the horizontal distance, this high frequency boost decreases, producing a recording with less coloration of the guitar and vocals in the higher frequency range. However the Sennheiser AMBEO microphone did not capture low mid frequencies as accurately as that the other arrays such as the Double Mid Side.

The intention of this project was to capture Charlie in a way that sounds like the recording space without having to use a lot of post-production techniques. However, from the recording tests, I came to the conclusion that a lot of post-production would be required to balance out the frequency spectrum of this microphone. This microphone was tested to see how ambisonic microphones could be used and decoded to 5 channel surround. Overall, I did not find that the microphone accurately represented the recording space, due to the lack of mid frequencies and the high frequency boost.

I still wanted to try this microphone during one of the final recording sessions to see how the Sennheiser AMBEO could work when placed close to Charlie, specifically in between the acoustic guitar sound hole and Charlie's mouth. The AMBEO microphone was used in the first recording session for Track 1, *"Cursed"*. In post-production, EQ was used to flatten the microphone's frequency response but this track was re-recorded without the use of the Sennheiser AMBEO due to issues with the stereo capture of the acoustic guitar.

4.1.4 Hamasaki Square

The Hamasaki Square array was tested using four Schoeps CCM8 microphones setup in the square array. The array was tested at two horizontal distances of 5m and 7m in front of the speaker as well as surrounding the microphone as an experimental setup.

This technique is intended for capturing only reverb while still maintaining a good stereo image. The advantage of having two pairs at different horizontal distances is that it can create the effect of the reverb immersing the listener due to the different arrival times. For positioning this array, I tried sending both the front and rear microphone signals to just the rear speakers but decided that positioning the front left and right microphones in the front speakers and rear left and right microphones in the rear speakers gave the most aesthetically pleasing response. With this setup the recorded sound seems to bloom from the front and travel around the listening position. This is due to the two stereo pairs being placed at different distances, so when played back the closer microphones are heard before the microphones at the back of the array. When the array is placed at a horizontal distance of 3m or 5m from the speaker, there is only a slight difference in that the source sounds slightly more distant. (Markers 25 and 26).

I also tested the array in a position where the microphones surround the speaker. This position has an advantage of capturing some direct sound due to the microphones that are in front of the source being close enough for pickup (Marker 27). The microphones behind the source captured a reverberant sound, but mainly the early reflections. This may be due to the positioning of the microphones in the room in relation to reflective surfaces in St. Paul's Hall such as the pillars and walls around the stage. If this setup was to be during any of the recordings, I would move Charlie further into the hall performance space to avoid this issue.

I came to the conclusion that I would use the Hamasaki Square technique at a horizontal distance of 5m to capture the natural reverb of the room without any direct sound. I would combine it with at least one other array such as the Decca Tree Surround array at a mid-distance so that there could be a blend of direct and reverberant sound in the recording. Careful considerations would have to be taken into account while mixing so as to not drown out the detail of the acoustic guitar in the performance. Additionally, as the microphones will be placed far away and capturing a fairly quiet source, additional noise may unfortunately occur because of necessary gain staging.

4.1.5 - Recording Tests Review

As previously mentioned, these recording tests allowed me to not only gain a greater understanding of surround sound microphone techniques but be able to make plans on what to use for the final recording sessions with Charlie. The extensive collection of audio files with easy comparison due to the perfect alignment in the Reaper project meant that I could create mixes of the different recording techniques and plan what to use.

While discussing with Charlie on the recordings, we listened to his bedroom demos and discussed how the space could be used to reflect the mood of each track. When listening to the recording tests, we could create mock mixes of the tracks by mixing the different techniques. From these mock mixes we already had a spatial impression of how each track was going to sound at the end of project before the recordings had even taken place. These mock mixes were referenced throughout the recording and mixing process to make sure the intentions were being met. These recording tests were vital in the pre-production phase as without them, the real recording sessions with Charlie would have likely been repeated many times and would have most likely led to a worse spatial representation of the mood of Charlie’s work.

One outlier that was surprising to me was the combination of the Decca Tree Surround at a far distance and the Hamasaki Square. Originally I believed that this approach would not have enough direct sound and too much reverb to be suitable for my intentions in the recording sessions. The centre microphone of the Decca Tree Surround at around 3m from the speaker was still picking up enough detail in the guitar finger picking sections that could be used to highlight the intimacy that was intended with Charlie’s work in this project. This approach was used in Track 4 with these distances in mind. However, in the recording session the Decca Tree Surround array was moved slightly closer to increase the detail just slightly. This technique would not have been considered without the testing approach I completed during the pre-production phase.

The main negative of testing in this way was the lack of close microphones to capture the detail in Charlie’s voice and guitar. As these close microphones were not tested, issues arose in the real recording sessions such as imbalanced stereo images and phase issues due to the different arrival times. These issues did lead to one recording session having to be re-done but the way the recording tests were conducted allowed for a much more in depth pre-production phase that I believe led to a stronger final product.

4.2 - Individual Track Recording and Mixing Details

Based upon the recording tests and communicating with Charlie to discuss the style and meaning behind each track, the microphone techniques and recording locations were chosen. These decisions were not set and could be changed during the recording sessions but generally these pre-made decisions were kept with minor adjustments to positioning and microphone choice. In each of the final mixes, a stable soundstage (Moylan, 2012) was created by placing Charlie in front of the listener position with the surround sound microphone techniques enveloping the listener. This meant that the recordings were a realistic representation of the space and it felt like listening to a performance with high amounts of detail being reproduced from the close microphones. In regards to Zagorski-Thomas' "Functional Staging" (2006) concept, the final mixes are intended to create a realistic representation of the performance which is reproduced in the five channel listener environment. This functional staging approach was suitable for this project due to the intended listening perspective of an acoustic singer-songwriter.

The choices for surround sound capture in the final recordings were all based on the tests conducted and based on personal choices from myself and Charlie for accurate capture of the space and style of the songs. Additionally to the room capture, close microphones were used to capture the detail of the guitar and voice of the performances. A mono cardioid condenser microphone was used on the voice and varying stereo microphone techniques were used on the guitar such as XY, ORTF and MS.

Throughout the mixing process, the spatial impressions were personally evaluated and compared to what Charlie and I had discussed at the beginning of the process and to the mock mixes created from the recording tests. Throughout the mixes, only EQ and compression was used on the microphone signals, no additional effects were used. While the original intention was to create this realistic impression of the recording space, it was discovered that having slight and subtle level change in some of the microphones would emphasis certain lines and phrases. This synthetic change in the mix was most frequently used to increase the omni room signals throughout Track 1 - *"Cursed"*.

As reference tracks, the likes of Jon Boden's *"Under Their Breath"* (2010, track 11) and Villager's *"The Waves - Live at RAK"* (2016, track 9). These tracks were the basis for the guitar and vocal tones as they are authentic and represent a true to life performance. As this project intended to be an authentic reproduction, referencing similarly mixed music that isn't a "sonic cartoon" (Zagorski-Thomas, 2014, pp. 49-50) with over-exaggerated characters was important.

4.2.1 - Cursed

The first track on the E.P., *"Cursed"*, is a slow finger-picked, intimate song. I therefore intended the overall sound of the vocals and acoustic guitar to be very close rather than distant. I wanted the use of the rear speakers to communicate a contrasting, rich and long reverberant sound which would compliment the frontal detail. To find a suitable space to record in, I played the dry recording in contrasting places including St. Paul's Hall, Phipps Concert Hall, and a medium-sized practice room. These were all at The University of Huddersfield. Phipps Concert Hall is a large wooden room at The University of Huddersfield which features a short but dense reverb. Phipps Concert Hall has an advantage of having acoustic curtains along the walls, so the spatial characteristics can be changed to fit the need of the recording session.

St. Paul's Hall was chosen due to its long reverb time which would be used to emphasize Charlie's vocals when he sings louder. As he increased in volume, naturally the level of the reverb in the room increased and I wanted to record in a space that could show this. I intended for this track to have really intimate moments that would draw the listener in and to contrast this with immersive sections that would surround the listener with a realistic portrayal of the sound in the room, including the sound reflections of the recorded space. By using a mixture of close and far microphones, this balance of intimacy and immersion was achieved.

The first recording of this track had an issue where due to an error in the placement of a mid-side microphone setup to capture the acoustic guitar closely, the stereo image was unbalanced and was pulling the guitar to the left. When trying to boost the right channel of the side microphone, this caused the vocal to move to the right of the frontal stereo image. These issues can be heard here: *USB Files/02 - Charlie Hulejczuk/03-Additional Files/01-Cursed Recording*. Because of this issue with stereo balancing, a new recording was carried out in the same space. Table 1 shows the microphones used in this repeated session. Instead of a mid-side setup for capturing the guitar, it was changed for an ORTF configuration using Schoeps CCM4 (cardioid) microphones. This setup allowed me to more accurately monitor the stereo image during the recording session while still ensuring a wide and stable stereo image. Based on the recording tests previously conducted, I opted for a double mid-side configuration for the medium room. Originally this was placed close to Charlie (approx. 1m away) but was then moved further back to change the balance between dry and reverberant capture. In the recording session I found that I wasn't picking up enough reverb of the space by just the DMS setup but I didn't want to move it because of the previously explained balance. To capture more reverb of the space, I put up an additional two omni outriggers at a distance of approximately 6m away and 8m wide. This choice was based on previous recording session in St. Paul's Hall. While this second recording didn't have as strong of a performance compared to the first, the stereo imaging issues with the close microphones in the first recording were unfixable so the second recording was chosen for the final version. Using an ORTF microphone configuration to capture the acoustic guitar gave a well-balanced stereo image without any balance issues. Appendix 11 shows a diagram of the recording session to represent the positions of each microphone. Appendices 12-13 are photos from the *"Cursed"* recording session.

Input	**Position**	**Microphone**	**Notes**
1	Vocal Close	Neumann U87 (cardioid)	
2	Guitar Close - Left	Schoeps CCM4 (cardioid)	ORTF Configuration. Pointing at 12th fret, approximately 1m away
3	Guitar Close - Right	Schoeps CCM4 (cardioid)	
4	DMS - Front	Neumann KMD w/ KK184 (cardioid)	Coincident array. Positioned approximate 2.5m away
5	DMS - Side	Neumann KMD w/ KK120 (figure-of-eight)	
6	DMS - Rear	Neumann KMD w/ KK184 (cardioid)	
7	Omni Outrigger - Left	Neumann KMD w/ KK131 (omni)	Positioned in front of far pillars 6m away
8	Omni Outrigger - Right	Neumann KMD w/ KK131 (omni)	

Table 1 - "Cursed" Re-recorded Channel List

Of the four tracks, Charlie and I intended for this one to be the most intimate. With this in mind, while mixing, I mostly used the close microphones to represent the detail in the finger-picked guitar playing and intimate lyrics. The omni outriggers were sent to the rear left and right speakers that subtly changed in level throughout the mix to emphasis certain phrases. Using the Schoeps Double MS decoder plugin, I added more focus to the front by narrowing the pickup pattern and increasing the level of the front mid microphone. The versatility of using the DMS technique and the plugin allowed me to reflect the mood of the piece by focusing the detail of Charlie's performance. The rear speakers were kept at a relatively consistent level throughout the track with subtle cuts or boosts to emphasize certain points in the song. For example, the line "playing out inside my head" at 2:40 when Charlie sings the loudest, the rear speakers are brought up slightly to make the line sound larger and more reverberant than the others. Contrary to this, throughout the final verse at 3:45, the omni outrigger channels are reduced slightly to emphasize the intimacy in the lyrics and performance. While the intention of this of this project is realism and not altering the recording where they sound unnatural, the artificial change in the level of the capture enhanced the performance. I found that by keeping the close guitar microphone signals lower in the mix led to a more mellow guitar tone and let the vocal be the forefront of the mix which Charlie and I thought would work best based on the style and lyrical content.

4.2.2 - Blessed by the Blossom

“Blessed by the Blossom” features two main contrasting sections. The first section has a soft finger-picked guitar with similarly soft and intimate vocals, whereas the second section has fast strumming guitar with a loud and full vocal performance. With this in mind, I decided to record in a space that would be suitable for this. I did not want to have the performance drowned in reverb, but also did not want a completely dead and dry space. As there was no space like this on the university campus, I used a small club room at Holmfirth Civic Hall in Huddersfield. This space was narrow but long, and had a low ceiling (4.7m x 8m x 3m). This space gave a very short, dense reverb that I thought would suit this track.

As this space was not close to the university, I had to take the amount of equipment I could transport into consideration. With this in mind, I limited myself to eight microphones. This meant that using a lot of room microphones was not an option. Double Mid Side was chosen for general room capture as it is a small and coincident array that can be setup quickly and with only one microphone stand. XY was chosen for the acoustic guitar as previous sessions had shown it work well and could be setup on one microphone stand. I positioned the far room omni outriggers at 6m away from Charlie in order to capture the short and dense reverb of the room. Table 2 shows the microphone list used during the recording session. Appendices 14-17 show the floor plan and photos of the recording session.

Input	Position	Microphone	Notes
1	Vocal Close Mic	Neumann U89 (cardioid)	
2	Guitar Close Mic - Left	DPA 4011 (cardioid)	XY Configuration. Pointing at 12th fret, approximately 1m away
3	Guitar Close Mic - Right	DPA 4011 (cardioid)	
4	DMS - Front	Schoeps CCM4 (cardioid)	Coincident array. Positioned approximate 3m away
5	DMS - Side	Schoeps CCM8 (figure of eight)	
6	DMS - Rear	Schoeps CCM4 (cardioid)	
7	Far Room - Left	DPA 4006 (omni)	Positioned at the back of the room, approximately 6m away
8	Far Room - Right	DPA 4006 (omni)	

Table 2 - “Blessed by the Blossom" Channel List

As I had not recorded in this room before, this recording session did take longer to setup due to the unfamiliar acoustics and reflection of the room. Multiple positions of the Double Mid Side array were tested to balance the capture of direct and indirect sound. Additionally, the DMS was tested in multiple places to try to capture more room ambience during the loud strumming section. However, this led to less capture of the detail of the guitar and the reliance on the close microphones had to be increased to try and match what was being captured during the first section. This tonal imbalance was unsuitable for a natural recording so I decided that keeping the microphones in the same spot throughout the track was the most sensible option. Moving the microphones would have also meant that two takes would have to be edited together which would have made the performance less authentic and natural.

The Double Mid Side capture is the main focus throughout the mix as it captured the desired blend of the detail of the performance and the room ambience. Using the Schoeps Decoder, I could increase the focus from the front by increasing the level of the front mid microphone and narrowing the polar pattern of the centre channel to become more direct.

The mix of this track was the most challenging of the four tracks due to varying dynamics throughout the track as well as the room and traffic noises coming through on the sensitive microphones. Because of this, iZotope RX (iZotope, 2013) was used to remove any unwanted sound from neighbouring rooms and traffic. One main challenge was to achieve consistency in the acoustic guitar level during the first half of the song but was achieved using multi-band compression on with a larger ratio in the mid range to have consistent levelling throughout. For the final strumming section, the levels of the vocal and acoustic guitar close microphones were increased slightly due to how the louder playing was increasing the amount of reverb in the room. This level increase of the close microphones ensured the detail could still be heard.

4.2.3 - Waste Our Time

The intention of this track in the E.P. was to have the most immersive capture of the space due to it being the loudest of the four tracks and containing the most powerful vocals. With this in mind, I decided that St. Paul's Hall would be the most suitable of the available spaces to me due to the large space and long reverb time.

For the capture of the acoustic guitar, a mid-side setup was used positioned approximately 50cm away from Charlie to capture a natural stereo image that could be sent to the front left and right speakers. For vocal prominence and detail, a Neumann U87 was used.

To capture the reflections in the room more prominently, I chose to use cardioid microphones as the rear setup of the Decca Tree Surround array instead of all omni microphones. This captured the rear reflections more directly while still blending in with the other three microphones in the array. During the tests for different pickup patterns in the recording tests for Decca Tree Surround, I found that the DPA 4011 cardioid microphones naturally have a lower sensitivity to noise and this higher noise floor (20dB A-weighted). To avoid this issue, I chose to use the Neumann Digital Microphones with the KK184 capsule (13dB A-weighted). This meant that a clearer and less noisy was captured without the need for de-noising during the post production phase.

An additional stereo pair was set up 5m away from Charlie using two AKG C414 microphones set to hyper-cardioid. These were pointing away from Charlie at a 45° towards the ceiling to capture reflections of the back of the hall. During the recording I found that these microphones were capturing the reverb well but during the mix, these were sparingly used. Appendices 18-20 show the floor plan and photos of the recording session. Table 3 is the recording inputs from the recording session.

Input	Position	Microphone	Notes
1	Vocal Close Mic	Neumann U87 (cardioid)	
2	Guitar Close - Mid	Schoeps CCM4 (cardioid)	Mid Side Configuration. Pointing at 12th fret, approximately 50cm away
3	Guitar Close - Side	Schoeps CCM8 (figure of eight)	
4	DTS - Front Left	Neumann KMD w/ KK131 (omni)	Not using DPA Decca Tree Mount but positioned microphones as if it was being used Positioned approximately 2m away
5	DTS - Front Right	Neumann KMD w/ KK131 (omni)	
6	DTS - Centre	Neumann KMD w/ KK131 (omni)	
7	DTS - Rear Left	Neumann KMD w/ KK184 (cardioid)	
8	DTS - Rear Right	Neumann KMD w/ KK184 (cardioid)	
9	Room - Left	AKG C414 (hypercardioid)	Positioned 5m away at 2.5m height. Pointing away from Charlie at 45°
10	Room - Right	AKG C414 (hypercardioid)	

Table 3 - "Waste Our Time" Channel List

Throughout the mix, the levels of each array stay relatively consistent with slight cuts and boosts on the close microphones when required to bring more detail to the performance. The vocal microphone was sent to the central speaker to give a hard anchor to the lead vocal without having to use any phantom imaging (The Recording Academy's Producers & Engineers Wing, 2004).

This track utilises the rear speakers more than any other track, making the natural reverb of the space is more prominent. This worked well for this track due to it being the loudest and most powerful track in the E.P., meaning that the listener can be fully immersed and placed into the recording space.

One issue regarding compression of the acoustic guitar arose while mixing. The volume at which Charlie plays the guitar contrasts drastically in the verses and choruses. To attempt to level this out, heavy compression was firstly used to create a consistent level. However, this was creating an effect of changing the balance between direct and reverberant signals between the close microphones and the Decca Tree Surround array. Because of this, the level of the guitar throughout is less consistent than what I was aiming for. In future projects, an addition of a very close microphones such as the DPA 4099 may add additional control in terms of balance of the direct and reverberant signal.

4.2.4 - The Duchess of Westdene

As the musical style of this track is very intimate and personal, my intention was to record this in a drier space to give the listener the impression of closeness and intimacy to Charlie. Phipps Concert Hall at The University of Huddersfield was chosen for this recording session. I used the acoustic curtains to reduce the amount of reflections in the room, so that a dry recording could take place.

I decided that the main array for this track should be the Decca Tree Surround as an all omni array with added close microphones in case of needing additional guitar and vocal detail. The sound of the room would be captured using the Hamasaki Square array. The Decca Tree Surround array was chosen as I felt that the recording tests showed that this technique had the most suitable representation of the recording space when compared to the other tested techniques. Table 4 is the microphone list used during the recording session. Appendices 21-23 show the floor plan and photos of the recording session.

Input	Position	Microphone	Notes
1	Vocal Close Mic	Neumann U87 (cardioid)	
2	Guitar Close - Left	Neumann KK184 (cardioid)	ORTF Configuration. Pointing at 12th fret, approximately 1m away
3	Guitar Close - Right	Neumann KK184 (cardioid)	
4	Decca Tree Surround - Front Left	DPA 4006 (Omni)	Used the DPA Decca Tree Surround mount where mounts are at max length (60cm) Positioned approximately 2.2m away
5	Decca Tree Surround - Front Right	DPA 4006 (Omni)	
6	Decca Tree Surround - Centre	DPA 4006 (Omni)	
7	Decca Tree Surround - Rear Left	DPA 4006 (Omni)	
8	Decca Tree Surround - Rear Right	DPA 4006 (Omni)	
9	Hamasaki Square - Front Left	Schoeps CCM8 (figure of eight)	Positioned in 8m away, 2.5m in height
10	Hamasaki Square - Front Right	Schoeps CCM8 (figure of eight)	
11	Hamasaki Square - Rear Left	Schoeps CCM8 (figure of eight)	
12	Hamasaki Square - Rear Right	Schoeps CCM8 (figure of eight)	

Table 4 - "The Duchess of Westdene" Channel List

While setting up, I found that I could place the Decca Tree Surround array further away than where I had been placing it in St. Paul's Hall and still achieve a good balance between the direct sound of Charlie and reverberant sound of the room. I created an environment that reduced the denseness of the reverb by using acoustic curtains around the walls up to where the rear microphones were placed. The Hamasaki Square was placed quite far away (8m) to capture the room sound of Phipps Hall to be mixed with the other microphones. As this song is very quiet and the microphones were placed far away, a lot of gain was required which meant that de-noising was needed in post-production to get the clean and natural sounding recording I was aiming for. iZotope's RX (iZotope, 2013) was used for the de-noising where a noise print was taken from each microphone signal and subtracted leaving a much lower noise floor.

One disadvantage of recording in Phipps Hall is that the room will occasionally creak because of wind. This is unavoidable and happens at random. Because of this, certain takes were ruined and it was very difficult to get one continuous take for the whole song. Several takes were spliced together because the alternative option of using iZotope RX (iZotope, 2013) to remove the creaking of the room was leaving strange artefacts that were causing issues with imaging. Because of these creaks, the Hamasaki Square array was placed at a very low level in the mix which was a shame as based on the recording tests Charlie and I thought the use of Hamasaki Square as a pure reverb capture of the space would have worked for this track. The rear facing microphones of the Decca Tree Surround array led to a good capture of indirect sound though.

As the microphones were placed further apart from each other in this session and were picking up a drier sound than in the other tracks, time aligning was required to remove any phase issues between the close and room microphones. This alignment was achieved by using alignment from a clap recording and the AutoAlign plugin from SoundRadix (2019). The end result is a detailed intimate recording that uses the surround speakers sparingly to immerse the listener so they can imagine they are watching the performance.

4.3 - Discussion

The original proposition of this project was that testing a large set of possible surround sound recording techniques could help make decisions to match the style and mood of contrasting songs. The overall, intention of this project was to represent the performance space and give an authentic reproduction of the performance. Zagorski-Thomas' "functional staging" (2006) concept was suitable for this project due to how a listener would intend on hearing a performance of an acoustic singer-songwriter in an acoustic space. Additionally, Moore's "1st perspective of authenticity" (2002) was considered in the representation of intimacy in the recording techniques. The aim with this representation of authenticity was to create a personal engagement between the listener and Charlie directly. This was achieved by creating the realistic representation of Charlie in the recording space and maintained intimacy by using close microphone techniques that would draw the listener into Charlie's performance and compositions.

The pre-production testing of surround sound recording techniques allowed me to make predetermined decisions that were the basis of the spatial elements in the final mixes. Each of the four tracks has its own sonic identity that is reflected in the spatial elements and the way that the mood is represented. Combinations of microphone techniques were chosen based on the pre-production phase to be represent the mood of each track. For instance, in tracks 1 and 4 where intimacy was the goal, closer techniques that could focus on the detail in Charlie's finger picking guitar playing were chosen. To contrast this, tracks 2 and 3 utilised microphone techniques that would highlight the natural reverb of the space that would represent the more powerful compositional style of those tracks.

The contrasting use of space is best highlighted in Track 1 *"Cursed"* and Track 3 *"Waste Our Time"* that create differing levels of intimacy. In *"Cursed"*, the close mic'ing techniques add to Charlie's intimate playing and lyrics, and the slight changes in room reverb levels on specific phrases adds emphasis. Surround sound is the most apparent in Track 3 *"Waste Our Time"* where the rears are the loudest and the room reverb envelops the listener.

The methodology I used allowed me to make effective production decisions due to the amount contrasting microphone techniques and the variables of distance and height tested. Because of the extensive pre-production testing process, spatial impressions of each track could be made with the mock mixes and be assessed before the final recording sessions took place. This methodology approach allowed myself, as a producer, to be able to make specific decisions in pre-production and only have to make small changes in the final recording sessions to get the intended sound. This extensive pre-production not only saved time in the recording sessions but also ensured that recording sessions with Charlie weren't being repeated unless absolutely necessary. This meant that Charlie only had to perform his songs a handful of times and didn't have to repeat takes due to technical flaws due to lack of planning on my part. Overall this led to stronger performances and better representation of Charlie's music in the spatial elements.

Additionally, because of the number of positions and variables that were tested in the pre-production phase, certain preconceptions about what techniques would be suitable for this project were eliminated. For example, going into this project, I thought combining too many spatial arrays that are intended for reverb capture would lead to a lack of intimacy with Charlie due to there being no detail in the guitar and vocal parts. However, from the pre-production testing, I found that Decca Tree Surround which was placed far away from the source and combining it with Hamasaki Square could still have the detail of Charlie's performance while picking up a reasonable amount of reverb in line with the intentions of each track.

For similar future projects, it would be interesting to try how to create a realistic impression of a performance featuring multiple acoustic musicians and see if similar impressions of intimacy could be achieved. In this, both of Waldrep's stage and audience perspectives (2007) could be experimented with to create different spatial impressions. By recording multiple acoustic musicians, considerations on the placement of them would have to be taken into consideration so that, dynamically, everything is balanced in the room. At the start of the project, I decided not to use height as I felt it was not adding too much to the level of immersion and the additional setup time was not worth the slight changes. It would be intriguing to test more 3D microphone techniques that include height capture to compare how much height can be used to increase the level of immersion and help match the mood of the compositions. Another spatial technique that was not considered for this project that would be interesting to use would be Wave Field Synthesis (WFS). This was originally not considered due to the lack of access to a WFS listening setup. By using WFS, it may be possible to better represent the acoustic space due to the closely spaced speakers and the lack of phantom imaging.

Overall this project successfully used the extensive pre-production process to make spatial decisions to represent Charlie's compositions and created the intended intimate mixes to make the listener have a personal connection with Charlie.

5 - Project 2 - create.evolve.destroy. Studio - 5th Order Ambisonics

When listening to either the full mix project or any of the audio examples referenced throughout, it is important to know that by default the SPIRAL 5th Order decoder is loaded and turned on and sent to the master output. Additionally a blank IEM Allra plugin is also inserted on the master channel so that the listener can import their own decoder settings. If not using either of these setups, the listener must import their own decoder. Throughout this project it is highly recommended listening to all audio on a full range speaker system with subwoofers to increase the low frequency range due to the heavy bass frequency content throughout the examples. The mix was created on a hemispherical speaker setup and therefore for listening, speakers below the listener ear position are not required.

create.evolve.destroy (c.e.d.) are a four piece experimental electronic progressive rock band. The band consist of four members: Yan who triggers drum and synth parts from an Ableton laptop rig, Jack who plays viola which is processed through a laptop, Chris who plays electric bass guitar, and Pete who plays electric guitar. The band mainly takes influence from Polyphia and their track *"O.D."* (2018, track 2) that is progressive metal-based in its guitar and bass arrangements. They also take influence from 65daysofstatic's *"Unmake the Wild Light"* (2013, track 7), a track that has electronic synth elements a lot of time signature changes, complex drum programming and is in glitch style.

This project follows the Collins (2010) approach of creative research where the musical elements were reflected throughout the project. In the pre-production phase of mapping out the projects and gathering the electronic sounds from Yan, certain elements were re-recorded to have a better overall sound. For example, the electronic cymbal crashes were replaced with real cymbals that sounded less mechanical and brighter. In the mixing phase, the spatial elements were recursively reflected by myself and the band to make sure the original propositions and intentions were being met and the spatial gestures were highlighting certain aspects of the band's composition such as build sections.

The intention of this project was to use space dynamically in 3D with height in a more experimental manner to create imaginary spaces to enhance and fit the musical ideas put forward from the band. By having the height dimension, this allowed for a more experimental way of mixing with placing instruments above the listener. Throughout the project, the spatial elements were assessed by myself and the band to make sure their original musical compositional ideas were being reflected in the space and the dynamic spacial elements were not becoming distracting to the listener. We agreed that the spatial elements should not be forefront of the mix but rater enhance some of the musical moments such as section changes throughout the continuous mix. At the beginning of this project, tests in ambisonics were conducted to determine the workflow. These tests included ambisonic orders, frequency localisation and height localisation. These tests gave me a framework of what would work in terms of placement and moving sounds dynamically throughout the mix. Additionally the mix was created in SPIRAL but was additionally tested on other 3D speaker arrays to check for portability and consistency.

This project is a three track continuous mix E.P.. The three tracks are separated by two ambience sections that help the flow in between the tracks. The recorded audio was typically recorded in a typical stereo studio setup with a mixture of mono and stereo microphone techniques apart from a viola solo which was recorded with multi-channel microphone techniques. The E.P. was mixed in 5th order ambisonics using Cuckos' Reaper (Version 6.0; Cuckos, 2019) and predominantly using the IEM Ambisonics Plug In Suite (IEM, 2019) for panning and positioning the audio. The Slate Digital Bundle (Slate Digital, 2019) EQs and Compressors were the main audio processing plugins along with other plugins from companies such as Flux Audio (Flux Audio, 2019) and DMG Audio (DMG Audio, 2019).

Normally the intention of ambisonics is for the final mix to translate to any multi-channel speaker system which is still the intention of this project. However, for the full effect of the mix, a hemispherical ambisonic speaker array with full range speakers and subwoofer is highly recommended. This project was mixed in SPIRAL Studio at The University of Huddersfield which features 25 Genelec 8240a loudspeakers (frequency range of 41Hz - 23kHz) and 4 Genelec 7270A subwoofers (frequency range of 19Hz-100Hz). There are no sub channels in the audio files so the speaker setup would require a speaker processor or active crossover to deal with low frequency management. As SPIRAL studio is setup in a three tiered polyhedra setup an ambisonic decoder created by Oliver Larkin was used that delayed and EQ'd the speakers to create a perfectly spaced hemispherical speaker array. When listening in SPIRAL, this plugin is recommended for accurate listening and is preloaded into the listening project. The plugin itself can be found on the USB file in this file path to listen to any other audio at different ambisonic orders *00 - Help Files/02 - Plug-ins/Oli Larkin's HACK*.

For this mixing project, I defined myself as a facilitative producer (Burgess, 2013, pp. 14). In this role of a facilitative producer I maximised the band's writing ideas and was given room for my own personal creativity in regards to spatial and sonic mixing. During the recording process we worked together to get the tonal and dynamic sound they intended, but was given full artistic freedom in terms of multi-channel production for the placement of sound. Moving away from the convention of “cautious” surround mixing (White, 2002), I didn’t want to establish any prior rules concerning the positioning of the instruments in the soundfield. This way spatial gestures such as sweeping position changes of the electric guitar could be featured in the mix.

5.1 - Ambisonic Tests

Before starting any mixing, I tested how different ambisonic factors would influence the localisation of sound in both the horizontal and vertical planes and how a range of frequencies would change the perception. As SPIRAL is setup in a polyhedra way where it is a stacked ring setup, it was important to make sure that the mixes would be phase coherent. When using the IEM Allra Decoder I found that because of the different arrival times it was creating a smeared image that was particularly apparent on pad sounds. I therefore used Oli Larkin’s SPIRAL decoder that adds a delay to each output to make the arrival times equal. This delay makes SPIRAL into a hemispherical ambisonic array.

When testing different ambisonic orders in regards to the accuracy of localisation and reducing any strange artefacts such as smearing, I found that 5th Order Ambisonics was ideal. These tests were similar to the tests by Thresh et al. (2017) but instead of using pink noise bursts, I used audio files that would be used in the mix projects. I compared how specific audio tracks such as hi-hats and guitars could be moved around the array. As SSA Plugins (2020) recommended using at least 3rd Order Ambisonics I started here and ignored 1st and 2nd. I found with 3rd order, the localisation was high but creating some smearing artefacts particularly when increasing the elevation above 45°. When increasing to 4th order this resolution again increased but with 5th order I found that I could very accurately place sound anywhere in the array with a high amount of accuracy. As I was intending to move sounds around and create musical gestures in time with the music I knew that this high accuracy was vital. When testing with 6th and 7th order I found the difference to be very minimal and not worth the additional intense computer power that would be required to process the additional channels. 5th Order Ambisonics gave me a sweet spot between localisation accuracy and level of computer power required for the full mix. In regards to a Mixed-Order Ambisonics setup, as SPIRAL studio has a high amount of speakers at a reasonable distance, and with the immediate access to Oli Larkin’s ambisonic decoder for SPIRAL, I decided that Mixed-Order was not required. Therefore a Fully Periphonic ambisonic setup was used where the final output file has the same resolution in both the horizontal and vertical planes.

As I previously mentioned, sounds with high frequency content such as hi-hats could be well perceived from any point in the array such as behind or above and no issues were created when fast movements were automated in. However, when testing sounds with bass frequencies, I found that the localisation would not be as accurate. When testing a solo'd bass guitar paying low notes, I found that the elevation could be increased slightly to approximately 15º while still being localised accurately but when increasing above this elevation the image became unclear and smeared. This result is in line with the tests from Roffler & Butler (1968) where they tested different frequencies at different loudspeaker heights to see how accurately listeners would localise. When this bass guitar was played with the other audio, this localisation decreased and became harder. Because of these issues, I decided to only place these low frequency sounds at the from of the array (±30º azimuth) and at ear level (0º elevation) and not create spatial movements and gestures with them.

Going into this project, I knew that a full speaker system with subwoofers would be required for this project. As ambisonics is intended as multi-useable format on different playback systems such as headphones, I tested throughout the mixing process that the mixes were translating moderately well to headphones. I also found that head tracking was not required for this project as I mixed it in a way that clearly represents all the instruments in a specific place. While the mixes were translating to headphones, there was a lack of power from the use of subwoofers in the speaker setup so it is recommended that this should be listened on speakers. The mix was tested on several different speaker setups such as the HISS system using a combination of Meyer and d&b speakers as well as a small setup using L-Acoustics speakers.

5.2 - Recording

The recording of the three tracks were all done separately using studio techniques of overdubbing and double tracking when required. Only the solo viola for Track 1 *"tent.legs"* was recorded with the ambisonic format in mind.

As the band have only ever performed as a live act, I worked closely with Yan to map out the projects and import the sounds that he would normally trigger from his laptop. We spent time programming the various time signature and tempo changes so that his pre-programmed audio would sync up with the click track for recording purposes. When c.e.d. perform live, Yan's triggered drum and synth audio is sent from an audio interface on a stereo left and right track. For a studio mix, this would be less than ideal as it wouldn't give a lot of scope for changes in terms of tone, level and placement. I therefore worked with Yan on separating these channels into stems so that I could have better control. Appendix 24 shows the audio from Yan's laptop I received for each track in the E.P.. The synth channels were mainly synth bass parts, but occasionally they were used for spoken word parts and effects.

The programmed drums had the advantage of being able to be converted to MIDI from the audio files without any issues. This let me resample the audio for different drum sounds that better fit the mix. However, as the drums samples are MIDI triggered, the human elements of timing and dynamic changes were lacking because the programming was overly quantised. After listening to the drum tracks, I decided that re-recording the cymbals and some of the hi-hats should be done to get an overall better tonal sound and human feel. The re-recorded live cymbals were recorded in a small drum recording booth with a pair of Coles 4038s and a mono Earthworks QTC40. The mono QTC40 was later scrapped because the Coles gave me the sound I was aiming towards so only the Coles were used in the final mix. The recorded cymbals helped in the mix as they gave me more tonal control and felt more human-like than the previous audio that was fully synced to the grid. The differences between the replaced cymbals can heard on the USB here: *03 - create.evolve.destroy. Studio Ambisonics and Mixing/02 - Recording Sessions/02 - Cymbal Replacement.*

The bass was recorded in a standard studio setup in which a Direct Input (DI) signal from a Rupert Neve RNDI was taken from Chris' pedal board and then re-amped later. All tonal and level decisions were mainly the result of processing from the Darkglass Microtube B7K on Chris' pedal board. The DI signal was then re-amped using a Matamp Bass Head Amplifier into a 4x12 Cabinet. A total of six microphones were placed on the cabinet and in the room for tonal options arising from mixing the different signals. These six signals gave me a lot of opportunities to sculpt the bass sound depending on the section. Appendices 25 & 26 show an annotated picture of the Matamp and the microphones used as well as a table with each microphone and the tonal characteristics achieved from them.

The Electric guitars were double tracked so that no artificial doubling was required during the mixing phase to have a stereo image. Pete's 8 string guitar was processed through several pedals such as the Wampler Dual Fusion and Zvex Fuzz Factory before being re-amped through a Peavey amplifier. The amplifier was mic'd up using a Royer R121 directly on the speaker cone for a warm, rounded tone and a Shure SM57 was used to capture the high frequency detail of distorted tone.

The viola parts were recorded over the course of two sessions. The first was for all the main parts throughout the three tracks of the E.P. The viola was recorded in Phipps Hall at The University of Huddersfield. As with double or quad tracking guitars, I wanted to use a similar technique with the viola to make it sound as if more than one viola was playing at the same time without having to add chorus or doubling effects. I decided that triple recording the viola would work well so that I could have left, centre and right recordings to position in the soundfield. Appendix 27 shows the microphones used in the first viola recording session and Appendix 28 shows a bird's-eye view diagram. I instructed Jack to stand in three different positions when recording the viola parts. By combining the three microphone signals of the three positions, a stereo image was achieved. This could then be placed in the soundfield using the IEM stereo encoder (IEM, 2019) that has a width control. For certain sections, such as the plucked intro to Track 3 *"fuck.the.grid"*, only the centre position was used. If the three positions were combined, the slight performance issues of the percussive plucking were causing an unstable image. This issue was not as obvious and perceivable in the bowed sections.

The second recording session was for the freely timed viola solo section of Track 1 - "*tent.legs."*. This was recorded in St. Paul's Hall at The University of Huddersfield. The idea of this session was to record the solo with ambisonics in mind. Therefore, microphones were placed at different heights and distances which could then later be relatively positioned in the soundfield. This was done so the natural reverb of viola in the room would fully surround the listener whereas in previous sections the whole band had been surrounding the listener. A Double Mid Side setup utilising the Neumann Digital Solution microphones was placed at head height approximately 2m away from Jack to capture a balance of direct and indirect sound. This microphone capture was then decoded to be placed in the front, sides and rear at ear level in the mix. A Hamasaki Square microphone array using Schoeps CCM8s for a general reverb capture was placed approximately 5m away from Jack and at 2.5m in height. This was then placed in the mid elevation (+45º) in the four corners. Finally, a pair of Neumann Digital Omni Outriggers were placed 8m away at 4m in height for a pure reverb capture of the hall. This was then placed in the mix at the rear left and right at a high elevation (+70º). By recording height channels, the localisation of direct sound and of early reflections was preserved (Miller, 2006). As heard in the audio example, when the three arrays are combined and played back, it gives a highly immersive and detailed capture of the solo. Appendices 29-34 are the recording documents from the viola session including the microphone list, a bird's-eye view diagram and photos from the recording session. The individual recording files from this session can be heard in the project found on the *USB: 03 - create.evolve.destroy. Studio Ambisonics and Mixing/02 - Recording Sessions/01 - Solo Viola.* It should be noted that there is no de-noising or processing of any kind on any of these channels.

5.3 - Creative Mix Decisions

All time stamps relate to the continuous mix of all three tracks

The style of mixing for this project is experimental in the way that it uses 3D space dynamically which moves away from the tradition of statically placing sounds. The spatial musical gestures that reflected the composition of the band enhanced musical moments throughout the continuous mix. Throughout the mixing process, the spatial elements were assessed by myself and the band to make sure that the original compositional ideas were being correctly reflected in the spatial ideas of the mix. Additionally, the band gave suggestions on possible gestures to create but also highlighted when certain spatial elements, such as overly complex movements, were distracting from the original music. In regards to Zagorski-Thomas' functional staging (2006), this project adopts a mixture of rock music staging and creating an imaginary sound space that does not reflect a real life performance scenario. In the rock music staging, the production of the instruments and reverb effects create a similar sense to a live concert. Additionally, there is a low frequency emphasis on the bass guitar, kick drum and bass synth instruments that create a "sonic cartoon" (Zagorski-Thomas, 2014). This was intentional to bring extra power to the mix and make use of the subwoofers in the mixing environment.

The band take influence from many different styles of music such as progressive rock, electronic rock, math rock and drum & bass. The Prodigy's *"Firestarter"* (1997, track 8) and Billie Eilish's *"you should see me in a crown"* (2019, track 4) were used as the main references for the hard hitting, transient heavy electronic drums and synth production. Karnivool's *"We Are"* (2013, track 4) was referenced for the deep distorted bass guitar production. Godspeed You! Black Emperor's *"Moya"* (1998, track 1) and *"Blaise Bailey Finnegan III*" (1998, track 2) were reference tracks for blending in string parts into richly detailed band arrangements with drums and many guitar parts.

This mix was created in SPIRAL studio at The University of Huddersfield but was tested for portability on headphones, the speaker setup created in Project 3 and a small 3D L-Acoustics system. The use of subwoofers does greatly impact the listening of this project so that is recommended. By testing the mix on different systems, the sonic characteristics of the music could be assessed but also the spatial elements could be tested to see if they were translating.

Space was used throughout the mixing process to enhance the music of create.evolve.destroy. and highlight musical moments. For example, when sections change from a breakdown to a verse section, this would be reflected in the spatial elements. For instance in Track 2, *"small.pizzalike.ladder.monk.",* at 13:14 where all instruments are playing in a low register, this is reflected by placing all the parts at 0° elevation. During the transition section (13:32 - 13:36) the instruments are scattered around the soundfield to represent the unpredictable playing style and then when the section changes back in to the main melody of the track (13:36), the instruments are once again separated. Another strategy for using space throughout the mix was to highlight solo sections by creating spatial gestures that reflected the playing style. For example, the guitar sweep solo (18:33 - 18:50) is highlighted by creating an arc over the top of the soundfield that moves in time with the guitar sweep going up and down.

As the band have two contrasting styles of playing, this suggested two main different spatial approaches. When the band play together in unison during the breakdown sections, this suggested a typical "wall of sound" approach. To contrast this, the band often play complex and contrasting parts typically with counter melodies and polyrhythms. This playing style suggested separating the band around the soundfield to help with clarity and allow the listener to discern the different instruments and what they are playing. This separation was achieved both horizontally and vertically. Each of these approaches had their issues when creating them for an ambisonic mix. With the intention of using space to enhance the musical ideas of the band, I didn't want the spatial elements to become distracting for a listener but instead add to their performances and compositions.

The first mix approach utilises frontal speakers to showcase the band in a typical wall of sound when they are all playing similar things in a tight groove. This can be heard throughout, but especially in the final heavy section of Track 1 *"tent.legs"* (9:06 - 9:20) where the band is locked in a groove and playing as one ensemble. This wall of sound approach created its own problem in the mix. As ambisonics positions sound to a relative spatial position rather than a physical speaker, phantom images and comb filtering effects were being created. I found that by placing sound and then slightly changing listening position would create a somewhat smeared image. Using 5th order ambisonics definitely helped in this regard as the placement was very stable and did not change too much depending on slight variations in the listening position. The placement of the sound in these wall of sound sections was generally placed in the relative position of the speakers in SPIRAL. However, when testing the mix for portability on different systems such as headphones and the speaker array built for Project 3, I found that the same smeared image was being created again. While this effect was subtle some of the power was lost. To reduce this effect, slight changes in the mix were made to even out these differences. The compromise was to reduce the power in the SPIRAL mix slightly to make the mix more portable. Due to the lack of subwoofers and externalisation when listening on headphones, the same effect of power is not as strong because due to the style of music from the band, an extended low frequency range is required. This recursive reflection of listening to the mix on different systems ensured that the mix would translate to different speaker setups.

To contrast the wall of sound approach, I chose to separate the band when they are performing counter melodies so that the listener can easily distinguish between the different parts in the busy arrangements. This can particularly be heard during the long build section of Track 3 *“fuck.the.grid.”* (18:50 - 20:40). During this section, the band start together but as their parts change, each instrument moves away from the centre until gradually coming back together for the ending section. This approach of separation added an element of clarity between the complex parts that would not have been possible mixing on stereo as all the parts would have been clustered in one spot. When mixing in this way however, I had to take careful consideration in to the positioning of the separated instruments. While showing early drafts to the band, they commented on that while the additional separation was helping clarify the parts, the separation was becoming distracting for the listener as there was no direct focus to one area of the soundfield. Because of these comments, the separation was narrowed in some of the sections. For example, in the section after the breakdown in “small.pizzalike.ladder.monk” (13:36 - 14:38), originally the bass, guitar and viola were equally separated in three corners but was changed to create multiple wide stereo images that still gave focus to the front of the mix.

The dynamic spatialisation of sources was mostly achieved with mono sources. For instance, the hi-hat roll in *“tent.legs”* (4:04 - 4:16) was achieved by altering the position of just the hi-hat channel. As the electronic drums were mixed in a way where they were individually placed in the soundfield and then bussed to an ambisonics mix bus, this meant I had overall level control while still having the flexibility of dynamic movement of the individual parts. This way of individually placing certain elements that were intended for dynamic movement was used for all of the drums and percussive parts, as well as the solo sections for parts such as the guitar solo in *“small.pizzalike.ladder.monk”* (14:40 - 14:56). When stereo or multi-channel signals were spatialised, I found that occasionally, due to the width, some comb filtering artifacts were being made. I found that this would especially be apparent when increasing the height above +45º elevation as well as when dynamically moving these multi-channel files. To reduce this effect when spatialising stereo signals, the width was either decreased or the position was kept closer to 0º elevation. In the breakdown section of Track 2 *“small.pizzalike.ladder.monk”,* every instrument hits the stabs in unison. For this section, I positioned every instrument at the front to create a wall of sound and to blend all of the instruments as one (11:36 - 11:53). As the instruments change what they are playing, they move away from each other and begin rotating in the soundfield around the 2nd and 3rd height rings in SPIRAL Studios (11:53 - 12:45). The bass changes first, then the guitar, and finally the viola where they “chase” each other around the top of the soundfield. When they are all playing the new part, the recorded parts spin at an equal angle difference (120º) which I intended to reflect the fast and looping playing style. The section ends with a breakbeat drum section with seemingly improvised instrumentation until resolving back into the verse section (12:45 - 12:52). During the breakbeat section, the instruments are spun around the soundfield to emphasize the unpredictability of the playing until rapidly changing back to previous spatial positioning.

Throughout the continuous mix, spatial movement gestures were used to help transition into the new sections. During these transitions, I wanted to added a spatial gesture to these rolls to highlight the changes in position of the other instrumentation. After the first heavy section in Track 1 *“tent.legs.”*, there is a hi-hat pattern that becomes busier during the build up (4:04 - 4:16). I chose to pan this across from the front left to the front right of the soundfield to accompany the musical build as well as to increase the high frequencies. This emphasised the speed increase of the hi-hat rhythm. During the final roll, each triggered hi-hat sample rapidly moves from left to right while getting wider. This was done to transition into the next section where the drums come in and the electric guitar is wider. This was influenced by Nine Inch Nail’s *“Copy of A”* (2013, track 2) where in the final section the hi-hat part alternates between left and right. At a point of showing the mix to the band, they commented how while this part is interesting to listen to, it was becoming a bit distracting hearing the hi-hat bounce from the far left and right of the soundfield. As both myself and the band thought this spatial gesture should still be highlighted, I reduced the width of the automation so the hi-hats still moved in time but didn’t get as wide as they did in the previous version. Additionally, in Track 3 *“fuck.the.grid.”* after the intro build there is a filtered synth that transitions the song into the first heavy section with full drums (17:40 - 17:50). During the build, the guitars and bass stay fairly central and slightly widen as the section builds. When the band stop playing and the filtered synth is introduced, it starts from behind and above the listener position and then sweeps around the wide left and right until ending in the centre directly in front of the listener position. This was done to add a huge impact to the first initial drum hit of the next section. Due to the way of recording the band with close mic’ing techniques to create dry recordings, the dynamic movement was easily achieved and was clear when listening on the different speaker arrays.

Having highly detailed dry recordings of the viola allowed me to use spatial effects in a way that isn’t possible when the band play live. During live concerts, Jack’s viola is processed through an

Ableton Live rig that has various delays and reverbs which are not normally separated. Because of the separated outputs in this project, in Track 3 *"fuck.the.grid",* there are many contrasting delays that are placed in the highest ring and voice of god speaker. Fluctuating the level of each delay which all have different timings and tonal qualities creates a texture of polyrhythmic delays that are perceptually challenging to follow (16:10 - 16:35). This was suitable in this section as there is no direct sense of time and the alternating delays in the different spatial positions enhance the existing music.

Reverbs were generally placed slightly wider and higher than the dry instruments. For example, the short plate reverb (2.1 sec) on the electric guitar was placed so that the guitar seemed to fill a slightly larger part of the soundfield. Multi-channel reverbs were tested with in the pre-production phase, however I felt that these were no where near as useful for the style of music in this project. These multi-channel reverbs had to still being upscaled to 5th order ambisonics. I mostly used stereo reverbs to create the sounds heard in the reference tracks including classic big plate reverbs and emulations of the Lexicon 224 Digital Reverb. These stereo reverb emulations were much more effective in the mix than the multi-channel reverbs. These stereo reverbs were then placed in the ambisonic soundfield using the IEM stereo encoder. When placing these sounds I found that having them placed slightly wider than the dry signal gave the best realistic sound. When moving the sound dynamically, the reverb channels were also automated to reflect the change. At times, the reverb effect would be placed in a separate location in the soundfield. In the band section of Track 1 *"tent.legs."* (2:44 - 4:04), the dry electric guitar is placed at ±100° azimuth angle and at 10° in elevation. The reverb was placed at ±120° azimuth angle and at 25° elevation. The slight difference in azimuth and elevation allowed for a slight separation that would not have been possible for a stereo mix. To counter the use of having the dry and reverberant sound in a similar location of the soundfield, occasionally the gated reverb of the snare was placed on the opposite side of of the dry signal in the soundfield. For the final section of Track 1 *"tent.legs."* (8:24 - 9:20), the dry snare is positioned directly in front of the listener position at 0° azimuth angle and 0° elevation, whereas the gated reverb is placed at ±160° azimuth angle and 80° elevation. The effect of having the two signals in contrasting locations meant that the snare sounded as big as possible which added to the intensity of the music. As the frontal image was occupied by all of the other instruments in that section, placing the gated reverb at the back gave it move space. This effect isn't that obvious but if the listener slightly moves away from the centre to the rear, it becomes more so.

5.4 - Discussion

The original proposition of this project was that the dynamic use of space in 3D with height can be used in a more experimental manner to fit the musical ideas put forward from the band. Throughout the continuous mix, space was used to highlight the band's ideas and create movement in the piece both horizontally and vertically that would not have been possible in using stereo or five channel surround. The additional vertical plane allowed for a greater separation between the instruments that aided in clarity in the dense mixes. By using the methodological approach of recursively reflecting on the spatial elements of the continuous mix, and by showing these ideas to the band for their feedback, I ensured that the experimental spatial gestures were not detracting from the band's work. This project shows how the dynamic use of space can be used to highlight musical moments and how the additional width and height dimension can create an imaginary, non-realistic space which reflected the style of create.evolve.destroy.'s music.

Section changes throughout the mix were highlighted in the spatial elements. For instance, breakdown sections would have contrasting spatial mixes that highlighted the heavy playing style of the band that would then be contrasted with the more melodic sections where the instruments would be separated around the soundfield. This approach of highlighting section changes is best shown in Track 2 between 13:00 and 13:45 where the spatial elements are constantly changing in response to changes in the musical material.

At times, separating the instruments did make the mix somewhat fall apart due to the lack of a frontal image. In earlier versions of the mix, the bass guitar and kick drum were placed in different locations in the soundfield where the kick was placed centrally (0º azimuth) and the bass was spread at the front (±30º azimuth). While the image was still frontal, this was leading to a lack of glue in the low end. To remove this issue, the bass guitar and kick drum were mostly placed in the same location at the front of the soundfield.

The main challenge of this project was creating a very powerful mix on a large ambisonic speaker system. The powerful mix was achieved by highlighting the band's playing styles in the spatial elements by having sections of separation be contrasted by the wall of sound approach. This is best highlighted after the introduction in Track 3 *"fuck.the.grid."* where in the introduction (16:34 - 17:42), elements are separated and moving around the soundfield. Then when the next section comes in at 17:43, a wall of sound is created in the centre where all the instruments a densely packed created the intended powerful mix.

In future projects it would be interesting to mix in a space that has additional speakers pointing from below the listening position to increase the sound sphere to include negative elevation. In this scenario, it would be interesting to separate the kick and bass instruments more as they predominantly acquired the same frontal space. Having subtle changes in negative elevation between these two instruments could increase in the separation and have a cleaner low end with less cluttering.

While this mix proved to be portable and translated well to other speaker systems with low frequency extensive, the mix did not translate as well to headphones. As one the intentions with ambisonics is to make a fully portable mix for contrasting setups, the end product is let down in this regard. Having the mix translate better to headphones would also have made it more accessible for general listening away from complex speaker setups. Early in the mixing phase, I did make the decision to fully utilise SPIRAL studio and the subwoofers available to me as that is what the band and I intended to do based on the referenced music material.

The end result of the 3D ambisonic mix, in my opinion, is a highly immersive and interesting listening experience that highlights the music of the band that would not have been possible on other systems such as stereo and 5 channel surround. The spatial gestures were carefully crafted to highlight musical moments and then reflected on by myself and the band to make sure they were not becoming distracting. The end result compliments the band’s music and makes use of the full ambisonic soundfield and the dynamic of use space.

6 - Project 3 - create.evolve.destroy. 3D Concert

The third project was a live concert of the experimental electronic progressive rock band create.evolve.destroy which featured a hemispherical speaker array. The event took place in Phipps Hall at The University of Huddersfield. For this concert, I designed the speaker array, created a 3D panner software controller to easily move the sounds around the hall, and worked with the band to optimise their setup to give me many options for creative live spatial mixing. This project is a technical demonstration of using off-the-shelf equipment for the use of 3D live mixing.

The intention of this event was to find a way how to use off-the-shelf PA and mixing equipment that was available to me to be able to create a live mix in 3D. As the technology available to me had no direct way of mixing on to a bespoke speaker array with height speakers, a workflow was created that routed the desk in a way to allow for direct outputs to the speakers and for the desk to be controlled by a custom external panner software. This software allowed me to dynamically place sound on to the 3D speaker array and be moved using a PlayStation controller meaning I could move two sources at the same time.

Similar to the production style of Project 2, I aimed to have the two spatial styles of separation and a “wall of sound” be represented throughout the concert. Throughout the pre-production phase, preset ideas were made based on the ambisonic mixes and by listening to demo tracks of the music not recorded for Project 2. These ideas were quickly implemented into the live set by programming the positions using the controller software and then being able to recall them on the desk. The extensive sound check time allowed me to repeat sections to make spatial changes and have most of the show programmed allowing me to use the software and the PlayStation controller for dynamic changes throughout the concert.

All time stamps relate to the binaural and video recording of the concert which can be found at: USB Files/06 - Appendices/02-Full create.evolve.destroy. Concert Binaural Recording with Video

6.1 - Array Design

Appendices 35-43 show diagrams of the speaker placements as well as annotated images of the spaces to get a clearer idea of the setup and speaker placements. The intention of the speaker array was to create a fully immersive sound system that could immerse the audience at all angles. The array was built with all point source speakers where the lowest speakers were mounted on speaker stands and the height speakers were either hung or mounted on to a truss system. As this event had a full live band with projected visuals, the placement of speakers was taken into consideration so that the speaker array was not detracting from the band's performance aspects.

The main design of the array was based on the Huddersfield Immersive Sound System (HISS) setup used for Electric Spring 2019 at The University of Huddersfield. The Electric Spring setup used Harrison's "shoe-box" configuration (Fielder, 2016, pp. 22) with symmetrical speakers for stereo pairs. Some modifications were required to fit the requirements of the band. As this was a concert with performers, a 4x2m stage was required meaning that the frontal left and right speakers would have to wider than what was used in the Electric Spring setup. Based on doing standard stereo concerts in Phipps Hall with a similar spacing between the left and right speakers, I knew that a central fill speaker would be required to help fill this gap at the front and help with coverage at the front of the stage for the audience. Another consideration was the live visuals that the band project on to a screen behind them. Because of this, no speakers could be hung in front of the screen but could still be mounted to the truss leaving a gap in the front height plane. Careful levelling between the mounted speakers and the main left and right speakers had to be done so when sound sources were moved they would seamlessly move across the soundfield. To test this, I used my 3D panner in SPIRAL Studios and turned off speakers to check for level changes. For the event, a Yamaha CL5 mixing console was used for all sound treatment and positioning. A Yamaha RIO3224-D I/o Rack interface was used for stage and speaker I/O and the speakers were made up of a mixture of Meyer UPJs, Meyer M1D subwoofers and d&b E8s.

The first speaker ring (blue circles in appendix 35), which was at the ear height of the audience, was based on Jonty Harrisons' "shoe-box" setup in which eight speakers (speakers 1-8) are placed in four stereo pairs. These pairs include a main L-R (±35º), front-side L-R (±70º), rear-side L-R (±130º) and rear L-R (±150º). I chose this setup as opposed to the "double-diamond" setup so that I could have stereo pairs that could easily accommodate the stereo signals of the keyboards and programmed drums. This made panning easier and which was important when performing quick spatial gestures as the band performed.

I added a central speaker (speaker 9) to these main eight speakers, referred to as a "punch". This was based on what has previously been used with the HISS. This was added because the main L-R speakers were wide (approximately 4m due to the staging), leaving a gap in amplification at the front. Including a punch speaker solved issues that arose from this gap and the speaker could be used at key points during the concert to add additional impact in the heavy sections. For example, during the end section of Track 1 *"tent.legs"* (08:16 - 09:10), the kick and bass were added to the central punch speaker to create an intense low mid boost. I did consider the addition of a rear punch speaker but as I intended the rear of the array to be used for effects and not direct sound localisation, as well as the limitations of the number of outputs available from the hardware, I decided only to have a rear stereo pair.

All of the speakers in the “shoe-box” configuration and the central punch were Meyer UPJs (frequency response 55Hz-20kHz). These were placed on speaker stands, and the tweeters of the speakers were two meters high from the floor. For the main L-R, a Meyer USW-1P subwoofer (frequency response 32-200Hz) was added, and for all the other speakers a Meyer M1D subwoofer (32-180Hz) was added for low end extension. Using the Yamaha CL5, a low pass filter per speaker was inserted on the subwoofers at 80Hz (-12dB/oct) and a high pass filter was added on the Meyer UPJ speakers at 80Hz (-12dB/oct). This crossover filtering was added so a low-mid build up would not occur where the frequency ranges of the speakers overlap (55-200Hz).

The next ring (green circles in appendix 35) was 1m higher than the Meyer ring and hung from suspension wires connected to the trussing of the system. These six hanging speakers (speakers 10-15) were used as the first height layer. I could not hang speakers at the front due to the band having visual projection, so these six speakers were suspended above the front-side L-R, rear-side L-R and rear L-R Meyer speakers.

The next ring (red circles in appendix 35) was 2.5m higher than the Meyer Ring and was mounted to the truss. These six mounted speakers (speakers 16-21) were used as the second height layer. These speakers were arranged in three stereo pairs. A front L-R (±25º), a side L-R (±90º), and a rear L-R (±155º).

The final ring (red circles in appendix 35) was at the same height as the previous ring but was mounted as overhead speakers, or Voice Of God (VOG) speakers. This ring consisted of four speakers (speakers 22-25) placed in a quadrophonic setup as a front L-R (±45º) and a rear L-R (±135º). The intention of this setup was to add drastic height perception to certain elements of the performance such as synth tracks and speech.

All of the speakers attached to the truss were passive d&b E8 speakers (frequency response 62Hz-18kHz) which were powered using a d&b D10 amplifier. Lee (2016) states that low frequency information (sub 125Hz) cannot be accurately perceived in the height direction. Because of this, all of the speakers mounted to the truss that were acting as height channels were high-passed to

125Hz as I did not want any low frequency information sent to these speakers which could lead to confusion in height localisation.

There was an equal distance between each consecutive speaker in each ring so that when a channel on the desk was panned using the controller software there was a consistent level in the central sweet spot. For example, for a sound that was being panned around the third level (speaker 16-21), the level would stay the same due to the equal distance between consecutive speakers.

The overall intention of the speaker array was to create a dome with equal spacing, but due to the square trussing, the distances between the central mixing point and each speaker varied. A distance measurement was taken and a delay was added so that sound arrived at the central mixing point in phase no matter which speaker it was coming out of. Table 5 below shows the distances of each ring and the delay time that was added. Due to the speakers in ring 3 being the furthest from the centre point (5.6m), this was used as the reference point.

Ring	Distance from centre	Delay time (ms)
1 (Main Eight)	4.5m	3.21
Punch	3m	7.58
2 (Hanging)	4.3m	3.79
3 (Mounted)	5.6m	0.00
4 (VOG)	3m	7.58

Table 5 - Delay Times

All sounds were processed using the Yamaha CL5 mixing console. Normally at concerts in which diffusion is performed, the desk can be routed in a way that allows fader control of stereo signals and sound levels for each discrete speaker. The multi-instrument band required 20 stage inputs, meaning that this method of single fader control for each input going to each speaker would be physically impossible as the Yamaha CL5 only has 34 faders. I would have required over 500 faders to do one fader per channel per speaker. I therefore utilised the 24 mix buses and 8 matrices on the Yamaha CL5 console. The 20 inputs could be sent to each mix or matrix independently, giving me full level control from each input channel to each speaker. This method was the only viable and reliable option to run this event given the equipment I had available to me. Another method was tested using Reaper and ambisonic routing, but this method heavily relied on software and introduced latency, so I decided to use the mixing console to mix the event. As the Yamaha CL5 console is designed to work for live sound events it has built in EQs, compressors and effects. This allowed me to quickly and easily process the inputs of the band and the outputs to the speakers without any additional software or hardware. Appendix 34 shows the desk routing settings used on the Yamaha CL5 mixing console.

6.2 - Controller

I made a 3D panner software in Max 8 (cycling74, 2019) in order to quickly and easily control the positioning of each channel. I wanted the controller software to be able to do three main things:

1. Be useable on any custom speaker setup
2. Be reliable and have consistent level panning around the speaker array
3. Have the ability to control multiple channels at the same time.

The idea came from Pink Floyd's Azimuth Co-ordinator that used two gyroscopes to control the level of sound playback for a quadrophonic speaker array (Calore, 2009). My controller expands on this concept by allowing the user to change the level of input channels to a large scale speaker array. The use of a PlayStation controller allows the user to change the position of two input channels at the same time using the two joysticks on the controller.

I based the user interface functionality on the IEM suite plugins which uses a simple XY grid to produce azimuth and elevation co-ordinates. The patch works by importing a file containing the speaker X and Y co-ordinates and creating a graphical representation for the position of the input channels that can be moved. This ability to import any setup of speakers meant that the patch was modular and could be used for any setup of speakers regardless of the azimuth and elevation amount. The distance between the channel XY values and the mix XY values is then calculated and scaled to accurately position the channel in the hall. These values are sent as a MIDI message that is sent to the Yamaha CL5 console to change the level of the input channel going to the mixes. Image 16 below is the user interface for the controller software.

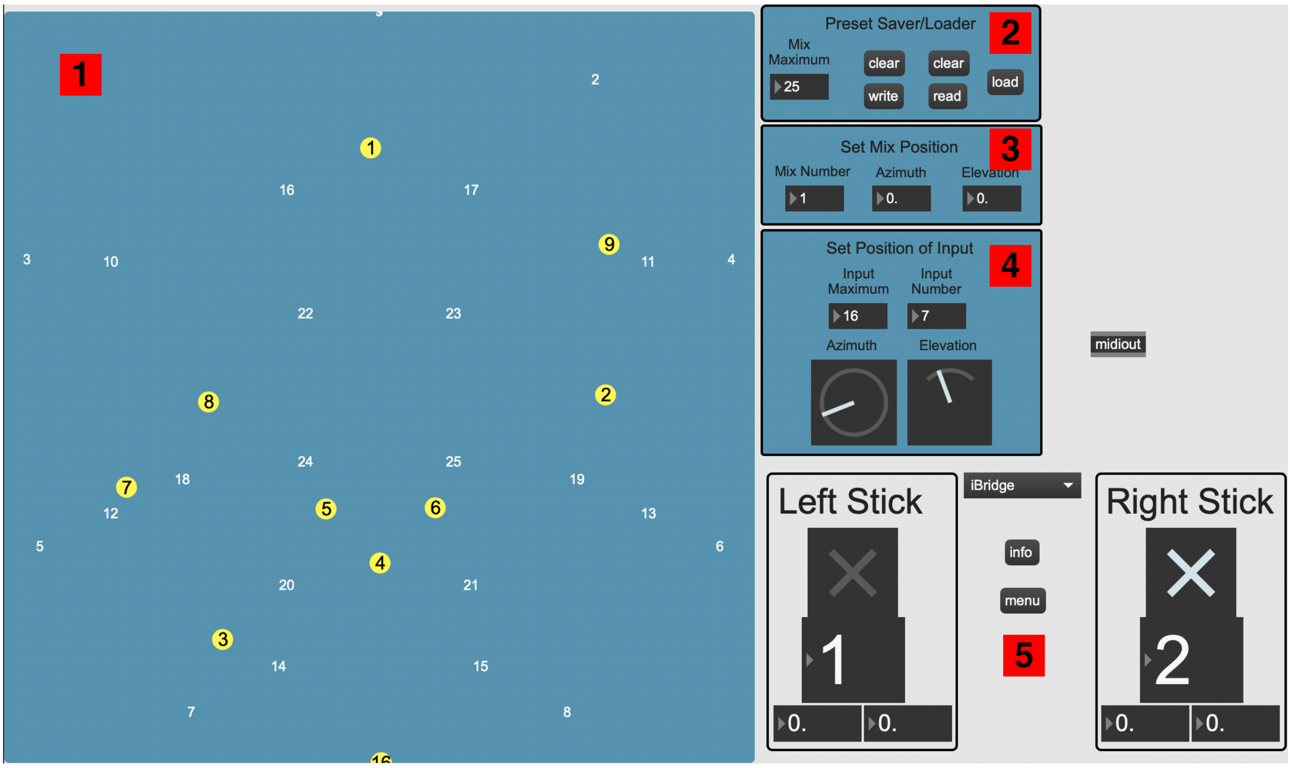


Image 16 - Controller Software

Section 1 is the main controlling section. The white numbers represent the speaker positions and the yellow circles represent the channel position. The user can move around any of the yellow circles to change the position of the input channels using a trackpad or a mouse.

Section 2 is used to import and export a text file that contains the X and Y positions of the speakers. The X and Y values are scaled between 0 and 1 throughout the patch so that a co-ordinate of 0.5, 0.5 is always the centre. For example, the line of text “1, 0.213 0.09;” would set the X and Y value of Mix 1 to position 0.213 by 0.09 respectively. In this section, the user can also set the maximum number of speakers. For instance, this setup has 25 speakers. The design of this patch can technically be scaled to any amount of speakers but would introduce more latency to the MIDI messages due to the way that the values are calculated and exported.

Section 3 is used to change the speaker positions (white numbers) without having to edit the text file. The user can select the speaker number and change the azimuth and elevation. This way the user can either prepare a text file to import into the patch or use this section to input all the values and then export it using the write function in section 2.

Section 4 allows the user to set the maximum number of input channels (yellow circles). In this case it is 16. The user can use the two dials to change the azimuth and elevation of each channel. This part of the patch also converts the Azimuth and Elevation values to XY values for the nodes object in Max 8. This is used for when the user wants to accurately change the value of the azimuth or elevation with affecting the other. For example, the user could rotate the channel using the azimuth dial without changing the perceived height.

Section 5 is the visual representation of the PlayStation controller input. The large toggle switches are highlighted when the thumb sticks are controlling the position. The number boxes show which input channel will be moved by the left thumb sticks. The two smaller number boxes simply show visual feedback of an XY value of the stick position to check it is working correctly.

The PlayStation controller can also change the values of the channel that the thumb sticks are controlling. The trigger keys L1 and L2 increase and decrease the value of the left stick respectively, and R1 and R2 do the same for the right stick. To see videos of me using the PlayStation controller to position sounds on the speaker array used for the concert, see below the videos show a screen capture of the software as well as live camera footage of the desk and the controller. The videos with binaural audio captured with the Neumann KU 100 showing the controller working and can be found *USB Files/04 - create.evolve.destroy. 3D Concert/02 - controller software/02 - Videos.* The first video shows me using the user interface of the patch to move a hi-hat around the array. The second video shows me using the PlayStation remote to move two sources, a distorted guitar and a hi-hat channel. In the second video, the left stick is controlling the relative position of the hi-hat and the right stick is controlling the position of the distorted guitar.

The patch calculates the XY distance difference between each speaker and each input. This distance is then scaled to a relative value for the mix level. This value is then sent to the Yamaha CL5 console to change the mix send amount. For instance, if the yellow circle for Channel 1 was on top of Speaker 1, it would send a value to send Channel 1 to the specific Mix/Matrix at 0dB. When the input channel position is halfway between the speaker positions, an approximate panning law of -4.5dB is applied. Because of the imperfection in the patch, there was a slight level variation of approximately ±2dB but there was still a strong localisation of the location of the positioned channels.

This value for the channel's mix level is then transferred into a System Exclusive (SysEx) message that the Yamaha CL5 console uses to change the mix send values. An example of a message would be like this:

240, 67, 16, 62, 25, 1, 0, 73, 0, (17)[1], 0, (3)[2], 0, 0, 0, (4, 124)[3], 247

1 - The mix number

2 - The channel number

3 - A value that represents the mix level in dB as a hexadecimal value. This mix level value ranges from 0 which would represent -∞dB on the channel fader to 1023 which would represent +10dB on the channel fader

While this patch works for basic panning, I would like to fix a few bugs and implement a few new features to improve functionality. One main flaw with the patch is that when the PlayStation controller is being used for position changes, the MIDI messages can get backlogged, creating a long latency. While Max 8 could create and output the MIDI messages quick enough, the hardware of the MIDI interface and the Yamaha CL5 console cannot process the messages at the pace Max is outputting them. To fix this, I could add a change threshold option to the XY tracking of the PlayStation controller thumb sticks so that small changes do not create new MIDI messages.

Another new feature that would be useful is a size control. This would allow the user to either have a narrow position, meaning that it would only output to one speaker, or have a large position where the channel would be outputted to many speakers. This could be implemented by changing the minimum distance between the input and speaker position for when it starts to increase the mix value. Visually this could be represented by changing the size of the yellow circle.

Another feature would be to have stereo mirroring. This would be useful so that on stereo channels the two inputs could be accurately stereo spread by both mirroring the X or Y value respectively. This feature could be used in a diffusion based situation where the console operator could change the position of the stereo sound files.

Many other features of the PlayStation controller could be made use of, such as the gyroscope and the touch bar to add additional levels of control. In addition, the controller buttons could be mapped to trigger desk events such as tap tempo, scene changes and effect changes.

6.3 - Technical Decisions

While planning for the event, changes were made in regards to the way that the band would be sent to the mixing console. Usually the band would perform with the following channel list. Table 6 is the channel list that create.evolve.destroy use for their standard stereo consoles.

Channel no.	Instrument
1	Laptop L (drums, synth, effects)
2	Laptop R (drums, synth, effects)
3	Keys L
4	Keys R
5	Bass DI
6	Bass Amplifier
7	Electric Guitar Amplifier
8	Acoustic Guitar DI
9	Viola/Ambiences L
10	Viola/Ambiences R

Table 6 - create.evolve.destroy Standard Channel List

Having mixed the band for the studio ambisonics project and in live events with a stereo PA, I knew that I wanted to create separation of the sounds coming from the two laptops. Yan's laptop is used for the drum and percussion tracks, synth tracks and effects such as spoken word. For stereo concerts in the past, he would output a mix of everything in stereo which had been pre-made in their practice sessions. This method has sometimes caused issues due to the lack of control over the levels of individual drum types. I worked with Yan on stem separation so that I could have the best opportunity for tonal control as well as position independence for each track. Appendix 44 shows a table of the inputs from Yan's Laptop. For instance, by separating the drums in Track 4, *"when.the.fall.bites."* allowed me have the kick drum in the central punch speaker (9) and the spoken word poem in the Voice of God speakers (22 - 25) which would not have been possible with the standard stereo output that is normally used (24:30 - 25:30). These channels were then sent from Ableton Live (Ableton Live, 2019) to an RME Fireface 800 and RME ADI-8 into the analog inputs of the Yamaha RIO 3224. Below is the channel outputs from Yan's laptop.

In the past, Jack's laptop had been used for viola processing as well as playing the ambiences in between each track. Due to having to EQ the viola on the desk to remove problem feedback frequencies, using his laptop in this way would have lead to issues with levelling and tonal shifts between the viola parts and the ambiences. To remove this issue entirely, I decided to have a separate laptop for playing back the ambiences so that the only thing coming from Jack's laptop was his processed viola. For further control of the dry and processed viola, it would be better to have separate outputs. For instance, at the start of the final track, *"fuck.the.grid."*, Jack plays with a lot of delay on the picked viola. It would have been interesting to be able to have independent position control of the dry and delayed viola as I was able to do in the studio ambisonics mix.

The same setup as their usual channel list was used for the bass, guitar, and keys. Table 7 is the channel list used for the multi-channel concert.

Channel no.	**Instrument**
1	Kick
2	Snare
3	Hi-hat
4	Tom 1
5	Tom 2
6	Cymbal 1
7	Cymbal 2
8	Synths 1
9	Synths 2
10	Ambiences
11	Bass DI
12	Bass Amplifier
13	Electric Guitar Amplifier
14	Acoustic Guitar DI
15	Viola L
16	Viola R
17	Ambience Front Left
18	Ambience Front Right
19	Ambience Rear Left
20	Ambience Rear Right

Table 7 - create.evolve.destroy Full Channel List

Further changes were made to their usual setup for on-stage sound and monitoring. For band-based concerts that have taken place in the hall in the past, the main issue with monitoring has been getting other parts such as vocals and guitars to be heard over the acoustic sound of the drum kit. As the band does not have any live drums, this main issue was already removed. While the band do not have a standard band setup, I still wanted to minimise the sound in the room coming directly from the stage such as amps and monitoring. To remove the sound of amplifiers from the room, I placed the amps in a separate store cupboard. This allowed me to have more control of the overall sound space with the 3D speaker array. While this cupboard was not soundproof, I still could have the amplifiers at a much quieter level than if they were on stage and the amplifiers could not be heard in the hall over the speaker array. Therefore, the only sound coming from stage was the viola and the monitors. During sound check, I encouraged the band to have quiet monitoring but it also turned out that the sound from the rear speakers in the 3D array naturally acted as monitoring for the band. Feedback was minimised due to the lack of acoustic sound on stage, but a graphic EQ was used to remove problem frequencies in the viola such as 500Hz and 1kHz. The use of quiet monitoring and very minimal acoustic sound on stage allowed me to have more control and opportunities be more creative in terms of mix choices and positioning.

As the Yamaha CL5 mixing console is digital, I was able to store presets that could be triggered to change not only the level of the input channel, but also the spatial position of each channel in the speaker array. Listening to demo recordings of each of the tracks that were going to be being performed at the event allowed me to make position decisions ahead of time. During the soundcheck, I used the controller software to quickly position each channel in the array and then made any necessary adjustments using the rotary controls on the console. These positions were then stored as scenes on the desk so that I could trigger changes for different sections in each track. Over 40 scenes were created in the soundcheck that allowed for complex spatial changes to be triggered throughout the show. This allowed me to have rapidly changing positions of the input channels that would not have been possible if using an analog mixing console. The scenes were changed by pressing a user defined key on the console that when pressed switches the desk settings to the next scene instantly. While this made the soundcheck process long and tedious at points, it was necessary to get the best possible result for the concert. However, at one point in the concert, an unforeseen glitch occurred and instead of going to the next scene when the user defined key was pressed, the desk skipped forward several scenes to a completely different section.

6.4 - Creative Mix Decisions

The mix intention of this event was to take the spatial elements created in the studio mix and recreate this live on a 3D speaker array. A similar approach of separation and a wall of sound frontal mix was taken into account for how the band were playing. This was made possible due to the extended width and height dimensions which would not have been possible with a standard stereo setup. During the sound check however, I found that having a wide separation of the instruments was creating a strange disconnect between the performers and the music. For instance, having the electric guitar coming from above when seeing the performer on stage was causing more of a distraction rather than an enhancement. Because of this, the separation was tamed in comparison to the ambisonic mix but was still present at key points. Throughout most of the show, a frontal image was created by placing the incoming sound in the front of the speaker array. However, the extended width and height allowed me to add separation between the instruments that aided in the clarity of the busy compositions. This allowed the rear speakers be used for effects such as reverb.

Having the drum tracks separated gave me a high level of control of positioning in the speaker array. For example, in the first verse of Track 2, *"small.pizzalike.ladder.monk"* (10:52 - 11:20), I used the controller app to move the rhythmic snare patterns that are heard in the intro. I rotated these around the voice of god speakers. This was done to add fluid movement to the rhythmic section while leaving space for the other instruments around the main ring of speakers.

In the past, the ambient sections in between each track would usually be played from Jack's Laptop. For this event I created quadrophonic mixes to be played back and live diffused using stems that the band provided. They were mixed in quadrophonic instead of full 3D due to the limitation of channel inputs of the desk and Rio rack. The positioning of the elements was based on the studio mixes in Project 2. During the ambient sections, I could then change the level of each speaker much like in diffusion, adding a live element to these sections. For instance, in the final ambience section before the final track, I was changing the level of each of the rings to create a sense of change in height (36:35 - 37:43). This approach was a mixture of the two methods of live diffusion and making pre-defined spatial elements discussed in Jonathan Fielder's paper (Fielder, 2016). The shoe-box design of the speaker array allowed for simpler diffusion of the quadrophonic mixes due to how I pre-mixed the files with stereo pairs in mind. Image 17 shows how I positioned the quadrophonic ambience mixes.

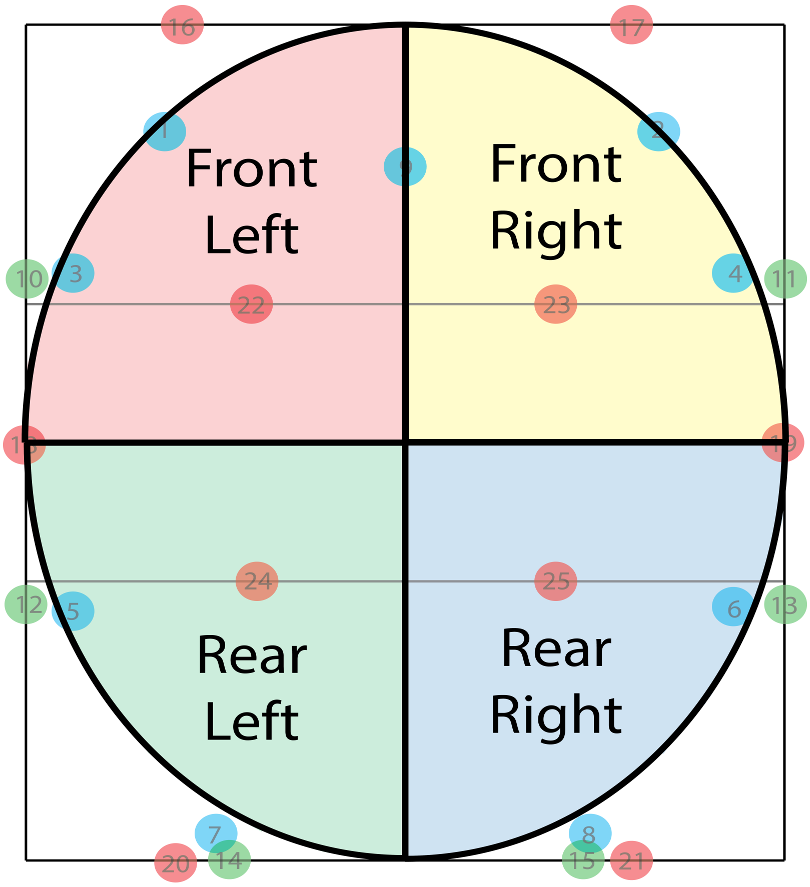


Image 17 - Ambience Track Positions

Throughout the concert I was conscious of creating moments that would sound narrow that could then later be expanded to the entire speaker array for effect. For example, the start of Track 5, *"blood.thru.the.green."* begins with a filtered drum breakbeat sound and then introduces a bass riff that is heavily effected with distortion and a DigiTech Whammy Pedal. To bring a dramatic impact to this section, I decided to only output the drum breakbeat sound out of the central punch speaker (9) but to have the bass in every speaker in the speaker array (30:12-30:40)**.** This created a large contrast in the immersion of the event at this point. When the rest of the band start playing, the bass is then brought back to the front creating space for the other instruments (34:40 onwards).

Usually on the Yamaha CL5 console, internal effects can be accessed using the mix and matrix buses. As all the mix and most of the matrix buses were being used as output channels, the use of effects such as delays and reverbs was limited. Because of this only one simple mono delay and a stereo reverb could be applied throughout the concert. The reverb was used on most of the instruments to aid the natural dryness of the acoustic DI and amplifiers. The reverb was placed in the higher rings of the array so that it would give space for the dry instruments and cluttering in the first ring of speakers could be avoided. The delay was used sparingly on certain inputs such as the viola and guitar parts. The delayed signal could then be moved around using the controller software. By using one hand to increase the level going to the delay and the other to control the position of the delayed signal using the PlayStation controller, I could add a very dynamic and strange sound in key sections. One of these was during the first half of Track 4, *"when.the.fall.bites.",* where Jack is playing a plucked viola (24:35 - 25:45). In this section I changed the level of the viola going into the delay and the position of the signal, creating contrasting polyrhythms in different spaces. Using a user defined key set up on the mixing console to be a tap tempo input, I could also change the delay time quickly and be in time with the music rather be an estimated millisecond time.

Similar to the studio mix, I decided to create two main images throughout the concert where in one scenario the band would be presented in a "wall of sound" where all the instruments are clearly represented as a frontal image and the other scenario where the instruments are separated add intelligibility between the complex rhythms and melodies put forward from the band. For example, at the end of Track 1, "*tent.legs."*, (08:55 - 09:10) it made sense for the band to presented as a large wall-of-sound frontal image as the band are playing in a tight heavy rhythm together. I achieved this by essentially creating a mono image of the instruments by sending equal amounts to the frontal speakers (Speakers 1, 2, 9, 16 and 17) and utilising the surround and height speakers for reverb and percussion tracks. To contrast this, at times I also separated the parts by placing each part in its own area of the array. For example, in the breakdown section of Track 2, *"smallpizza.like.ladder.monk"* (11:55 - 13:10), they start by playing on the same beats but then gradually change parts with the guitar changing first, then the bass and then the viola. To highlight this separation I began moving the parts in a circular motion in the upper area of the speaker array until 13:08 when they begin playing in unison once again. This circular motion was created by using the PlayStation controller to rotate the sound sources to control the desk. In my opinion, the use of having both of these approaches gave different spatial impressions of the contrasting sections throughout the concert. If just one of these approaches was used throughout the show, the spatial ideas would have become uninteresting and as the aim of the concert was to expand on the band's ideas use space, it would have failed in this aspect. One issue with trying to create a separated mix was leaving gaps in the soundfield. At times, my approach of separation was probably too extreme. By placing instruments in the wide left and right left a gap in the front. To overcome this issue, having the ability to place the sounds in multiple locations would have been useful. For example, having a static location at the front and having a moving location for spatial interest.

6.5 - Discussion

This event showed that with the addition of controller software and forethought into details of routing, off-the-shelf PA and mixing equipment can be used for 3D live events without the need of expensive third-party dedicated hardware. This was achieved by routing the mixing console in a way so that the mix buses (which are normally used for monitoring or effects) to be used as direct outputs to the speakers. This allowed me to do amplitude based panning by changing the channel level to each mix. With the use of the controller software, I could quickly positioning sounds on to the speakers without having to use the rotaries on the desk and allowed for dynamic movement during the concert.

The original intention was to use the off-the-shelf equipment available to me in a way to create a live 3D mix. Multi-channel events are becoming more popular recently with the installation of 3D speaker arrays in the likes of the L-ISA system in EartHackney and d&b's Soundscape in The Royal Albert Hall. These venues however do use expensive equipment that is fairly inaccessible for general use. My event showed that 3D mixing of a live band can be done with digital mixing equipment that most universities and small venues would have access to.

The software controller I created was very good at being able to quickly place sound on to the speaker array. This was used in the soundcheck to quickly place all the inputs from stage. The addition of the PlayStation controller meant that I could easily move sounds without having to use the trackpad of my laptop and be looking at the screen. The downside to the controller however was in the latency. When moving two channels simultaneously, occasionally the MIDI control messages would become backlogged and have increased latency in moving the sounds. If I were to do an event similar to this in the future, I would investigate other protocols for controlling digital mixing consoles such as OSC that would be quicker and more reliability. The Yamaha CL5 does not currently support OSC so a different mixing console such as the DiGiCo SD series would have to be used.

While the extensive soundcheck time allowed me to experiment with the placement of sound, find the best approach and be able to make a complex show file with over 40 scenes allowing for many spatial changes throughout the show, the soundcheck was very long. Luckily the band were more than happy to repeat sections, test out ideas in regards to the positioning of sounds and make subtle tweaks to their monitoring. If a band that weren't as enthusiastic, this approach would not work. To overcome this, in the future I could use a technique called Virtual Soundcheck (VSC). With this, I could have recorded their set while working out their monitoring and then played back the recording over Dante without needing the band and finely tune the mix and positioning. This technique is commonly used in large live tours so that the engineers aren't relying on the artists if any tweaks are required and is especially used in tours that utilise multi-channel array such as Bon Iver using the L-ISA system (Bon Iver, 2019). VSC would have allowed me to make more small changes without having the band in the room. This would have worked especially well as all the timing is based on click tracks from laptops so the band's playing is similarly timed between each performance with little to no alterations.

While this event worked well for this style of concert with very little on stage volume, the function of the speakers was for spatialisation rather than reinforcement. If this style of multi-channel event was done again in the same space with a live drum kit, I think that the live acoustic sound of the drum kit would cause issues if placing the drums in an experimental way. Additionally in this regard, create.evolve.destroy. are an experimental band by nature and this allowed for a more experimental placement of sound sources in the speaker array. I don't think such experimental placement would be suitable for a standard rock concert that are mostly in mono.

7 - Project 4 - The University Funk Band - 5th Order Ambisonics

Similar to the ambisonic mixing project of create.evolve.destroy. it should be noted that the listening projects are by default setup to be listened in SPIRAL with the Oli Larkin decoder plugin preloaded. Additionally a blank IEM Allra plugin is also inserted on the master channel so that the listener can import their own decoder settings. If not using either of these setups, the listener must import their own. While subwoofers aren't as mix critical in the listening of this project, it is recommended that listening should be done on a full range speaker system.

The University Funk band are a group of students from The University of Huddersfield. For this project, the two singers Lucy and Charlie wrote an original song titled *"Hear Me Calling Out Your Name"* and it was recorded and mixed for the output format of 5th order ambisonics. The band took influences from Marvin Gaye's *"Ain't No Mountain High Enough"* (1967, track 1) for the unison male and female vocals and Vulfpeck's *"Running Away"* (2017, track 5) for the rhythm section grooves while writing this track.

Similar to Project 2, this mixing project follows the Collins (2010) approach. In this regard, the alternate spatial mixes were recursively assessed by myself and the band frontman, Charlie. In this assessment we decided on what placement setups worked best and would best create the practice room layout we were originally intending. Additionally in this recursive reflection, the mix was tested on alternate speaker setups to test for portability and consistency.

The proposition of this project was that placing the listener amongst the band in a stage perspective mix could help with the separation of the instrumentation in a dense arrangement of a pop funk band. As the intention with this project was to create a practice room layout where the members of the band surround the listener, functional staging (Zagorski-Thomas, 2006) was not considered for this project. This approach goes against how a listener would intend to hear this type of music due to how the listener is placed inside and amongst the band, and how the individual elements of the band are separated to surround the listener.

Between myself and the band frontman, we decided that creating a realistic impression of the band in a practice room would be suitable. In this regard, we decided that having the instruments be static and at ear height would be required for the realism aspect. To find the best approach, multiple spatial mixes were created in the pre-production phase to find what setup may work best. Originally we thought that recording the band together would yield the best result in terms of creating the practice room layout but came to the conclusion that recording the parts isolated from one another would give the best scope for experimentation in the spatial mixing elements.

A similar approach was taken to the mixing of this project in SPIRAL where Oliver Larkin's decoder was used to change the polyhedral shape of SPIRAL to have identical speaker arrival times in an equal spaced ambisonics array. In this project, I also defined myself as a facilitative producer (Burgess, 2013, pp. 14). Similar to Project 2, I had no impact on the songwriting but was given freedom in the mixing phase for spatial placement as long the stage perspective mix was not jarring to the listener.

7.1 - Ambisonic Tests

Most of the decisions in regards to the setup of ambisonics are similar to the decisions made for the ambisonic mix of create.evolve.destroy. in Project 2. 5th Order Ambisonics was chosen for the high accuracy of localisation and for mixing, SPIRAL studio was used with Oli Larkin's SPIRAL decoder to make the Polyhedra speaker design have identical arrival times at the listener position.

The main difference between these two projects is the nature in which space is used. In Project 2, space was used dynamically with gestures created where sound would travel across the soundfield. For this project, sounds are placed statically to realistically represent the location of the players in a practice room. Because of this decision, almost no direct sound is placed in the vertical plane except to add slight separation between percussive layers. I found that placing the reverb channels slightly above (elevation +15º) the listener position would lead to more envelopment and lead to a more cohesive representation of the practice room when compared to having all elements at listener height (elevation 0º). Because of this test, the vertical plane was predominantly used for indirect sound such as reverb, delay and chorus effects.

Additionally in contrast with Project 2, there is no sub frequency content from the bass (<40Hz) and therefore full extension of the bass frequencies was not required. Because of this, I found that I could place the bass guitar more to the sides (azimuth ±80º) without creating any strange smearing artefacts where localisation was becoming hard. This discovery led to more options and experimentation in the placement of the instruments around the listener position.

7.2 - Recording Sessions

Due to constraints in getting all the members available at the same time, the recordings took place in several sessions. Firstly the rhythm section of drums, bass and keys were recorded. This recording took place in Phipps Hall at The University of Huddersfield and all parts were recorded at the same time in one continuous take. The drums were recorded with standard studio recording techniques with the addition of a Sennheiser AMBEO microphone placed approximately one metre away from the kick drum at the same height as the toms. This was done so that I could capture the natural reverb of the room in ambisonics to then be encoded in the mix. The bass was only recorded through a DI to avoid bass bleed into the drum microphones. It was then later re-amped through a TC-Electronics BH550 amp head and a K-210 amp cabinet to achieve a well rounded, clean tone. The Roland stage piano was recorded by capturing both the audio and MIDI outputs on a generic electric piano sound so that I could choose to either keep the original Roland Keyboard sound or switch it to a keyboard sound that could better fit the final mix. I later used Keyscape (Spectrasonics, 2019) for a better and more authentic sounding Rhodes sound.

Secondly, the rhythm and lead guitar parts were recorded. This was done in a small recording room where a Fender Stratocaster was recorded through a Fender Blues Junior amplifier. The amplifier was recorded with a Sennheiser e609 dynamic microphone and a Royer R121 ribbon microphone. These two contrasting microphones provided tonal differences that could be used to contrast the verses and choruses and overall create sonic interest in the track. The Sennheiser e609 gave a bright tone that allowed the Nile Rodgers style strumming to cut through the mix whereas the Royer R121 resulted in a warm timbre that gave a well rounded tone to the slow strumming parts in the verses.

Next a percussion recording session took place where congas, tambourines and egg shakers were added to the arrangement. The congas were captured with two Sennheiser MD421s positioned close to the skin to capture transient detail and low end information. The tambourines and egg shakers were recorded separately, each with a Neumann U87.

The vocals from Charlie and Lucy were recorded at the same time but in different vocal booths for complete separation. Recording at the same time meant that they could work off of each other to match harmonies and have better performances. If both vocals were recorded separately, the performances may not have been as strong. Overdubs were recorded as required to fix slight performance issues, but the majority of the performances were done together.

Finally the saxophone and trombone were recorded, completing the arrangement. The two instruments were recorded in a medium sized dry recording booth at the same time. Each brass instrument had a Royer R121 ribbon close microphone, and a pair of Neumann U87s were used as a room microphones setup in XY configuration approximately 2m away from the players.

The separation that was achieved by recording all parts in isolation gave me as much scope as possible for positioning during the mixing stage. While some funk artists do record at the same time in the same room, having to up mix to ambisonics could have introduced issues with spatial imaging if it had done it this way. If certain elements such as brass and drums had been recorded at the same time, the bleed that would have been picked up in different microphones might have created strange imaging and phase issues. By recording everything in mostly dry environments this was avoided.

7.3 - Creative Mix Decisions

The sonic intention of the mix was to create a practice room layout so overproducing the instruments and featuring long reverbs in the mix would not have fit the goals of this project. References to the modern pop-funk production style of a small studio from the likes of Pomplamoose's *"Jamirobeegees Mashup: Stayin' Alive / Virtual Insanity"* (2018, track 2) and Vulfpeck's *"Running Away (feat. Joey Dosik, David T. Walker & James Gadson)"* (2017, track 5) were referenced for their relatively dry and punchy production style.

I found that by placing the individual elements of the drum kit separately was leading to the issue of a smeared image, comb filtering and a lack of glue in the drums. Because of this, I mixed the all of the drum signals, apart from the overheads, to one stereo bus. The drum mix was then sent to channels 1-2 of a bus and the overheads were sent to channels 3-4 of that same bus channel. This multi-channelled bus was then encoded to ambisonics using the IEM Multi-Encoder where I positioned the four channels to the position in the soundfield. Having the overheads separated from the rest of the drum kit allowed me to slightly increase the width and elevation to give an impression of height to the drum kit. Bussing this way allowed me to make quick changes in level and positioning of the kit that would not have been possible with encoding each close mic'd signal to ambisonics separately. The Sennheiser AMBEO microphone was encoded in a different way where the 1st ambisonic order recorded signal was upscaled to 5th Order ambisonics with the IEM ToolBox. This signal was then rotated to match the position of the drum kit. An additional time delay was added to phase align all of the drum signals to minimise any comb filtering artifacts. If using the AMBEO microphone again, it would be interesting to attempt different positions such as an overhead or behind the listener. These different approaches may have been able to give me different perspectives of the drum kit. In the final mix, the drum sound mainly relied on the close mics for detail as the AMBEO microphone was placed too far away to capture my intended overall drum sound. Due to the lack of monitoring available to me in the recording phase, this issue was not realised until the mixing phase.

All other instruments were recorded in either mono or stereo with close mic techniques so I did not have the same issue of a smeared image so placing them individually worked fine. Instruments that were recorded in stereo such as the congas and brass, these were routed to one bus and using the IEM Stereo Encoder placed in the soundfield with a width control. Unlike project 2, all elements in the mix were static so the comb filtering artifacts when sound was moved dynamically was not an issue in this mix project.

In regards to reverb, only stereo reverbs were used. All reverbs, apart from effect reverbs such as the gated snare reverb, was the Slate Digital VerbSuite Classics recreation of the Lexicon 480L Reverb to simulate a short room reverb of a practice room. This reverb was duplicated for each instrument so that the return channel could be encoded to the same azimuth angle of the dry channel. This method gave the advantage of having more control in the reverb compared to just using one reverb channel. On the stereo return channels, the azimuth angles were slightly wider and the elevation angles were higher at around +15-25°. This was done to give the dry channels a more natural sound in the surrounding setup. A dedicated ambisonic or multi-channel reverb such as the Waves R360 Surround Reverb plugin was not used. While the routing and positioning was simpler and less intensive on the computer, I was not pleased with the sonic results when compared to the Lexicon emulator plugin from Slate Digital. The main negative of the surround reverb was that it was challenging to create the impression of being in a small practice room due to the way the Waves plugin emulates the size and density of a space. The stereo Lexicon reverb was conjuring up the sense of a small practice room better when compared to the surround reverb plugin so the stereo reverb was chosen for this mix. For future mixes, it would be interesting to take surround sound or ambisonic impulse responses of the real life spaces I am trying represent in the mix.

After the recording phase, a quick mix for balance was completed and then four main placements experiments were made. These mixes had different setups for separating the band to surround the listener. The different versions had contrasting approaches to be able to make the best judgement before going through to the end mixing phase. The four setups were then assessed between myself and Charlie to see what gave the best impression of a practice room and not be strange to the listener. To listen to the different setups, a listening project can be found here with 30 second snippets of each track: *USB Files/05 - The University Funk Band/01 - Audio Examples.* (Markers are setup for quick and easy switching between the setup configurations). After Charlie and I determined what setup was most suitable we then made minor tweaks to the positioning for balance and then I completed the mix.

These different setups experimented with the placement of sound and looked at the way that sounds were sometimes placed in the 1960s and 70s. Following Dockwray and Moore's discussions about the sound-box and the normative mix (2010), their analysis of the different mix styles were implemented. The idea of the "cluster mix" was not looked at due to the intention of separation for this project.

The first spatial setup (Marker 1) featured the drums and bass in the front and centre with the rest of the band surrounding the listener. The intention of this setup was to have a more familiar frontal image for the listener with grounding the drums and bass at the front. However, this setup with the lead instruments on the left and the rhythm instruments on the right was creating an imbalance between the left and right. This setup wasn't completely disregarded as Charlie and I agreed that the familiar frontal mix was a good idea. This setup closely relates to the "triangular mix" (Dockwray & Moore, 2010) where the main part of the mix, the vocals are shifted to one side allowing space for the drums and guitar. Image 16 below shows the positioning of the instruments in setup one.

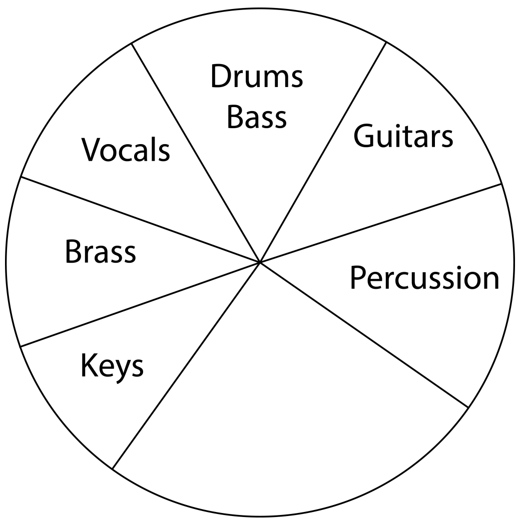


Image 18 - Mix position setup one

For the second mix setup (Marker 2), the idea was to have no frontal centre focus by leaving a gap at the front (±25° azimuth). This setup was quickly disregarded due to the constant changing direction of the main parts which was leading to constant rotation in the listening position.

Additionally as there was nothing in the front and had the lead brass lines and keys in the rear, this was creating a confusing perspective where it was difficult to determine where these instruments were coming from. To avoid the “cone of confusion” in future setups, either no instruments were placed in the rear (±150-180° azimuth) or had grounding of instruments at the front. From this setup, Charlie and I agreed that the separation of the drums and bass could potentially work for the mix. This mix setup does not closely identify with any of the sound-box approaches discussed by Dockwray & Moore (2010) due to the somewhat undefined positioning and lack of central focus. Image 19 below shows the positioning of the instruments in setup two.

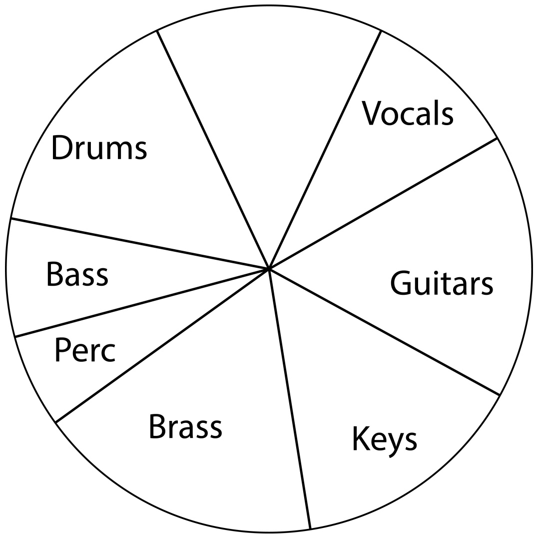


Image 19 - Mix position setup two

The third setup (Marker 3) returned to the frontal bass guitar and drum image but experimented with separating the vocals to the wide left and right (±60º) and the grouping of similar instruments. While this separation did help with having clarity between the two vocal parts when they sing together, in their solo sections it was again creating an imbalance between the left and right sections of the soundfield. Additionally with no lead parts in the centre position the spatial impression was again feeling unnatural. Again, the brass was placed at the back of the soundfield but the same “cone of confusion” effect was happening again. For future mixes, Charlie and I decided that the rear of the soundfield should not be used for this mix in any way. One takeaway from this test that we had not realised in the previous setups was that by having the drums and percussion be almost opposite each other helped bring separation between the two parts and brought clarity to the percussion. This setup returns to the “triangular mix” (Dockwray & Moore, 2010) approach where the drums and bass fill the central image and are the focus of the mix and the vocals and other instruments surround the soundfield. Image 20 below shows the positioning of the instruments in setup three.

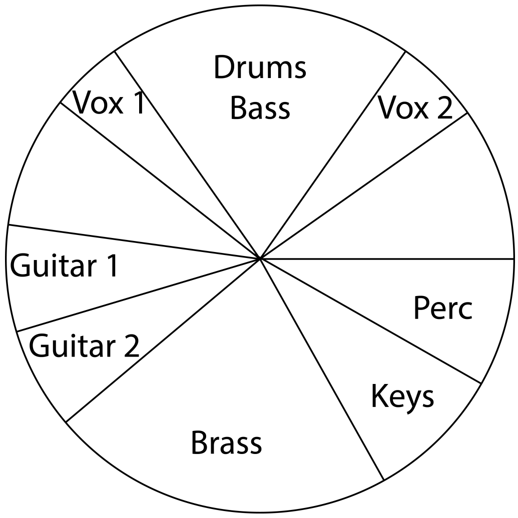


Image 20 - Mix position setup three

The fourth and final setup (Marker 4) took influence from the previous three setups and the findings from them. From the previous experimental setups, I decided that a frontal image was most suitable so the drums, bass and vocals were placed at the front. Taking influence from setup two where the bass and drums were separated, I positioned them separately at approximately ±30° in the left and right. The placement of the drums and bass allowed for a clearer low end and added clarity to the bass guitar part. The other tests determined that having a lead instrument in the centre would be useful to ground the listener in a position. At first, both vocal parts were positioned in the exact same position of 0° azimuth. However, in the sections where they sing in unison was detracting from the intended practice room layout due to the unnatural way they were overlapping. Because of this, the female vocal line was moved slightly to the right hand side to have subtle separation between the two parts. This made the vocals still be grouped together so in their individual sections they appeared to come from the same place but had subtle separation that aided in clarity when singing together. From setup three, I decided to place the percussion parts opposite of the drum kit on the right hand side at approximately 140° azimuth. This mirroring created a wide percussive image but helped with the clarity between the drums and percussion. Snarky Puppy's *"Tio Macaco"* (2014, track 7) was referenced to balance the busy percussion parts and create the wide stereo image.

During the final mixing phase I found that due to the percussion being placed towards the rear, the "cone of confusion" issue was again happening so the percussion was moved back towards the front at approximately 90° azimuth. Also from setup three, I decided that grouping the keys and guitar parts together as an intended "rhythm section" area. They were placed in the left hand side and mirrored the bass guitar part. While separation was the goal, grouping the keys and guitars together created a nice blend between the two. It's still possible to hear the differences between the two.

Finally, the brass group was placed on the left hand side of the soundfield. When the brass comes in, the listener might naturally rotate to listen to that part. The brass mainly plays after the choruses where there are no vocals. This means that as the listener naturally rotates, there will be no detraction from other lead melody parts. In the final section the brass is slightly lowered in volume to not over-power the vocal parts. This ensured that the vocals were still the most prominent part in the mix whenever they appeared. This mix setup relates to the "triangle mix" (Dockwray & Moore, 2010) again but this time the vocals are the central focus of the mix and the other elements surround the vocals without any clustering. Image 21 below shows the positioning for the final mix.

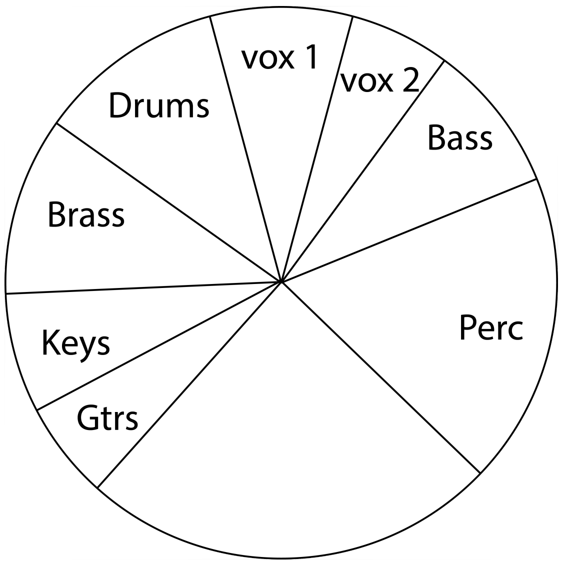


Image 21 - Mix position setup four

Overall this setup best reproduced the practice room layout by having the instruments surround the listener and create a somewhat familiar sense of a pop-funk mix by still having a fairly frontal image. After this setup was decided by myself and Charlie, the track went in to the final mixing stage to be finished. A typical mixing approach was conducted to balance out the instrumentation and keep the vocals to be the most present in the mix.

7.4 - Discussion

The original proposition of this project was that by pacing the listener amongst the band in a stage perspective mix could help with separating the band and add clarity to the mix. By creating the contrasting test mixes, alternate approaches were attempted to best find a way to represent the pop-funk band in this experimental spatial mixing approach. The assessment of these alternate mix positions led to me being able to make the decisions on which approach worked best for the intention of creating the stage perspective mix and best separate the instruments in line with the aims of this project.

By recording the instruments isolated from each other, this gave the best scope for experimentation on the placement of sound. If I had taken the approach of recording the band together in a room with predefined positions around a 3D microphone, there would have been less scope for experimentation in the positioning of the instruments in the mixing phase. This would have led to having to repeat the complex session of recording the entire band together. An approach that may have worked for this project would be to create these mock mixes with demo recordings of the band to find the most suitable layout and then record the entire band together in that layout. This approach may have had the same degree of separation but helped blend the instruments together due to them all being in the room together. This way, the spatial elements in terms of placement and reverb would already be present in the recording. In this mix, the 3D microphone could have been utilised more and give the directional cues in the mix, and then the close microphones could be used to reinforce the sound. By using a higher resolution 3D microphone such as the em32 Eigenmike by mh acoustics that records in up to 4th order ambisonics, more directional resolution could further separate the instruments. This would lead to having less of a smeared image when compared to the Sennheiser AMBEO microphone that only records in 1st order ambisonics. This approach of recording may have led to the final mix to be less jarring when compared to standard popular music productions due to the instruments all being recorded in the same room. This would be different to my production approach in this project where everything was recorded with close microphones and then used artificial reverb to give the impression that the instruments were recorded in a small practice room.

While separating the band this way did help with the intended separation and showed an alternate approach to spatial mixing, the final sonic result is not as strong as the other projects due to the experimental mix style of the spatial elements. Due to the diversion away from the “normative mix” (Dockwray & Moore, 2010), I was not as satisfied with the spatial elements of this project when compared to the other three projects. This was because I felt that it jarred with the expectations of hearing a pop-funk band due to the over separation of the parts. As the separation of the instruments did bring clarity to some of the parts such as the kick drum and bass guitar, I felt that due to the lack of “glue” in the mix (and particularly the low end) this further moves away from the expectations of hearing a pop-funk mix.

For future projects with this goal, maybe adding some form of interactivity to the product would help. In this regard, a 360º video of the band with a binaural render and head tracking could be used. This way additional separation and no centre point could be used which could further explore the original proposition and intentions. Due to the constraints of having the whole band be available at once, the 360º video was not possible for this project.

Overall this project shows that while separation can be achieved with this alternate spatial approach, additional elements are required to help in the way a listener would positively react to the music. However, the mixing project has shown that immersive multi-channel audio isn’t just limited to electroacoustic music but can be used for a pop-funk band mix.

8 - Conclusion

Each project showed how space and the different technological approaches could be used in contrasting ways for different spatial intentions such as realism and immersion. Project 1 highlighted how contrasting recording techniques can be used to reflect the mood of a track and create an intimate, personal connection with an acoustic performer. Project 2 shows that the space can be used to create dynamic movement of sound to effectively enhance the musical intentions of a band. Project 3 shows that off-the-shelf equipment can be used to create a mix of a band in a live event with the use of custom controller software. While Project 4 had the weakest result of the four, the project shows that space can be approached in alternate ways to separate the instruments of a pop-funk band to aid in clarity and separation but an additional element of interactivity may be required to achieve the best results.

The research aims established in the introduction were explored throughout the four personal projects to show how the use of space can be used for contrasting approaches and intentions. The first aspect I aimed to explore throughout the projects was to use space to enhance the compositional ideas from the artists. Throughout the four projects, the different compositional aspects were reflected in the spatial elements of the final mixes. For instance, Project 1 used space to reflect the mood of Charlie's work and increase the intimate nature of his work for the listener. Project 2 used space to create musical gestures that would highlight musical elements in create.evolve.destroy.'s music such as builds in the composition and section changes. Project 3 took the spatial ideas created in Project 3 and translated these to a live settings where they were controlled using the PlayStation controller. Project 4 took the compositional ideas of the funk band and added separation to the composition to bring clarity to the arrangement.

The second aim was to explore different capture methods to ensure creative possibilities during the mixes. This is best highlighted in Project 1 where contrasting capture methods were tested in the pre-production phase to best make decisions before the final recording sessions. This method allowed me to create mock mixes of each track to show to Charlie that highlighted how the mood of his writing could be best reflected in the surround sound recordings. Additionally, Project 2 featured a recording session that took the findings of Project 1 to make decisions for the viola solo in Track 1 *"tent.legs.".* For this solo, the recording techniques were chosen to be able to map on to an ambisonic setup with height.

The third and final aim was to explore how space can be used to create real life spaces or create something non-realistic. In this regard, realism was the goal for Projects 1 and 4 where a realistic impression of the performance space was reflected in the recording and mixing approach. In Project 1, this was achieved by carefully choosing the microphone technique, and in Project 4 this was achieved by creating a small practice room layout in the choice of the placement of the instruments and reverb. This realism goal was contrasted in Projects 2 and 3 where space was used to create an imaginary, non-realistic space where the instruments are dynamically moved around the soundfield in both a studio and live mix. This non-realistic representation of space worked for these projects due to the experimental nature of the band.

Additionally throughout the four projects, workflows were established to be able to best create the spatial elements given the technology available to me. This is best highlighted in Project 3 where the off-the-shelf live mixing equipment had to be used to 3D mix in a live setting. In this project, a workflow was created to ensure that the digital mixing console could be used for multi-channel mixing with the use of custom controller software. While the controller software had its minor flaws with latency, the ability to map any number of speaker positions (much like an ambisonics decoder) and be able to quickly position sounds on to the bespoke speaker array, ensured that the off-the-shelf equipment could be used in the live setting.

The Collins (2010) approach of recursively reflecting on the spatial elements allowed me to ensure that the propositions of each project were best highlighted with the use of space in multi-channel music production. By reflecting on the spatial elements made from the different recording, mixing and live approaches, I could ensure that the individual project aims were being explored and that the musical ideas of each band were being enhanced. This recursive reflection worked most successfully with the intentions of Project 1 where I could make spatial assessments of each track by creating mock mixes from the test recordings. This assessment allowed me to make spatial decisions on how to best represent the mood of each track. In Project 2, the recursive reflection of the spatial elements allowed me to judge how the experimental style of mixing was enhancing musical moments and the musical ideas from the band. By communicating with the band and showing them the spatial ideas, I could get their feedback on how well their musical ideas were being represented in the use of space. In Project 3, the extensive soundcheck time allowed me to reflect on the spatial elements of the live mix by repeating certain sections to create the complex show file on the digital mixing desk. This allowed me to recreate the spatial elements of the studio mix of create.evolve.destroy. in a live setting. The contrasting test positions in Project 4 allowed me to make judgements on what layout could be used for the final mix. By reflecting on these test positions I could make decisions on what layout would best represent the practice room layout while being as convincing as possible. Overall, the recursive reflection methodology approach throughout the projects allowed me to make spatial judgements throughout the production process and find ways to best represent the artists' musical ideas.

In future multi-channel production projects, I would like to explore different approaches of space that were not used in the projects. As previously discussed in Project 1, Wave Field Synthesis would be an interesting area to explore for creating a realistic representation of an acoustic space. Additionally different 3D production tools such as Dolby Atmos and Flux:: Spat Revolution would be interesting to explore for multi-channel mixing. Here I could explore how different mixing tools could achieve different musical results due to the expanded output formats and the built-in production options such the multi-channel room reverbs and interactivity options. In this regard, interactivity between the listener and the music would be an interesting area to explore to see how this would further impact the immersive nature of multi-channel music. As discussed, I believe an element of interactivity would have improved the result of Project 4. I would also like to be able to use different live mixing technologies such as L-Acoustics' L-ISA to be able to create complex movements and see how their workflow is different and how different spatial gestures could be created.

In conclusion, the four projects have highlighted some different approaches to space in the production of multi-channel audio and shown that space can be used in contrasting ways to best highlight the musical intentions of different artists.

References:

Bibliography:

2l. (n.d.). *3D Recording with the "2L-cube"*. Retrieved on November 12th, 2020 from http://www.2l.no/artikler/2L-VDT.pdf

Bacon, T. (2019). Steven Wilson: Remixing Classic Albums. *Sound On Sound.* Retrieved on November 12th, 2020 from https://www.soundonsound.com/people/steven-wilson-remixing-classic-albums

Benjamin, E. (2008). Ambisonic Loudspeaker Arrays. In *Journal of the Audio Engineering Society 125, San Francisco, CA, USA.* Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=14756

Bleidt, R., Borsum, A., Fuchs, H. & Merrill Weiss, S. (2014). Object-Based Audio: Opportunities for Improved Listening Experience and Increased Listener Involvement. In *SMPTE 2014 Annual Technical Conference & Exhibition.* Retrieved on November 12th, 2020 from https://www.iis.fraunhofer.de/content/dam/iis/en/doc/ame/Conference-Paper/BleidtR_SMPTE2014_Object-Based_Audio.pdf

Blumlein, A. D. (1958). British Patent Specification 394,325 (Improvements in and relating to Sound-transmission, Sound-recording and Sound-reproducing Systems). *Journal of the Audio Engineering Society, 6*(2), 91-98.

Bowles (2015). A microphone array for recording music in surround-sound with height channels. In *Journal of the Audio Engineering Society 139, New York, NY, USA.* Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=17986

Burgess, R. J. (2013). *The Art of Music Production: The Theory and Practice.* (4th ed.). Oxford. Oxford University Press.

Calore, M. (2009). *May 12th, 1967: Pink Floyd Astounds with 'Sound in the Round'*. Retrieved on November 12th, 2020 from https://www.wired.com/2009/05/dayintech-0512/

Campos, G. (2018). *Royal Albert Hall first in UK to deploy L-ISA sound system*. Retrieved on November 12th, 2020 from https://www.avinteractive.com/news/audio/royal-albert-hall-sees-uks-first-deployment-l-isa-technology-04-12-2018/

Candy, L., & Edmonds, E. (2018). Practice-Based Research in the Creative Arts: Foundations and Futures from the Front Line. *Leonardo, 51*(1), 63-69. doi: 10.1162/LEON_a_01471

Collins, H. (2010). *Creative research: the theory and practice of research for the creative industries* (1st ed.). London: Bloomsbury.

Corbett, I. (2014). *Mic It! Microphones, Microphone Techniques, and Their Impact on the Final Mix* (1st Ed.). London: Taylor & Francis Group.
Cogan, R. (1991). Varese: An Oppositional Sonic Poetics. *Sonus, 11*(2), 26-35.

Cunningham, M. (1997). Welcome to the Machine. The story of Pink Floyd's live sound: part 1. *Sound on Sound*, March 1997.

Curran, S. (2017). *Kraftwerk 3-D at the Royal Albert Hall, London: Reconfiguring what a pop gig can be - and how it can look.* Retrieved on November 12th, 2020 from https://www.independent.co.uk/arts-entertainment/music/reviews/kraftwerk-3d-tour-royal-albert-hall-review-performance-tickets-a7805136.html

d&b. (2019). *Kraftwerk 3D tours Europe in Soundscape*. Retrieved on November 12th, 2020 from https://www.dbsoundscape.com/global/en/stories/kraftwerk-3d-tour/

d&b (2020). *Leeds Playhouse upgrades with the d&b Soundscape.* Retrieved on November 12th, 2020 from https://www.dbsoundscape.com/global/en/stories/leeds-playhouse-upgrades-with-the-db-soundscape/

Desantos, S, Roads, C & Bayle, F. (1997). Acousmatic Morphology: An Interview with François Bayle. *Computer Music Journal, 21*(3), 11-19. doi: 10.2307/3681010

Dockwray, R. & Moore, A.F. (2010). Configuring the sound-box 1965-1972. *Popular Music, May 2010,* 181-197. Retrieved on November 12th, 2010 from https://www.researchgate.net/publication/231866249_Configuring_the_sound-box_1965-1972

Dolby (2014). *Authoring for Dolby Atmos Cinema Sound Manual*. Retrieved on November 12th, 2020 from https://www.dolby.com/uploadedFiles/Assets/US/Doc/Professional/Authoring_for_Dolby_Atmos_Cinema_Sound_Manual(1).pdf

Dolby (2016). *Dolby Atmos for the Home Theatre*. Retrieved on November 12th, 2020 from https://web.archive.org/web/20200521122348/https://www.dolby.com/us/en/technologies/dolby-atmos/dolby-atmos-for-the-home-theater.pdf

DPA. (2018). *Immersive Sound / Object-Based Audio - And Microphones*. Retrieved on November 12th, 2020 from https://www.dpamicrophones.com/mic-university/immersive-sound-object-based-audio-and-microphones

DPA. (2019). *Stereo Recording Techniques and Setups.* Retrieved on November 12th, 2020 from https://www.dpamicrophones.com/mic-university/stereo-recording-techniques-and-setups

Duff, S. (2017). *Exclusive: Behind the scenes at legendary film composer Hans Zimmer's world tour.* Retrieved on November 12th, 2020 from https://www.psneurope.com/live/hans-zimmer-2017-world-tour-digico earthackney (2020). *EartH Theatre*. Retrieved on November 12th, 2020 from https://earthackney.co.uk/venue/earth-theatre/

Farley, S. (2012). *Panning Reverb Returns. Guest contribution by Douglas Murray*. Retrieved on November 12th, 2020 from http://designingsound.org/2012/12/24/panning-reverb-returns/

Ferguson, S. & Cabrera (2006). Vertical Localization of Sound from Multiway Loudspeakers. *Journal of the Audio Engineering Society, 53*(3), 163-173.

Fielder, J. (2016). *A History of the Development of Multichannel Speaker Arrays for the Presentations and Diffusion Acousmatic Music*. Retrieved on November 12th, 2020 from jonfielder.weebly.com/uploads/1/2/3/0/12308331/fielder_pennycookcomp_multichannelarrays.pdf

flux.audio. (2018). *Spat Revolution with Jean-Michel Caron in Cirque's New Production - BAZZAR.* Retrieved on November 12th, 2020 from https://www.flux.audio/2018/12/21/spat-revolution-with-jean-michel-caron-in-cirques-new-production-bazzar/

flux.audio. (2020) *Spat Revolution*. Retrieved on November 12th, 2020 from https://www.flux.audio/project/spat-revolution/

Gerzon, M. (1973). Periphony: With-Height Sound reproduction. *J. Audio Eng. Soc.*, *21*(1), 2-10.

Ghanekar, S. (2017). *Object Based Audio: A Signal Processing Overview*. Retrieved on November 12th, 2020 from https://www.ee.iitb.ac.in/~eestudentrg/ppt/S_Ghanekar.

Huber, D. M. & Runstein, R. E. (2012). *Modern Recording Techniques* (7th Ed.). UK: Focal Press.

Hamasaki, K. & Koichiro, H. (2003). Reproducing Spatial Impression With Multichannel Audio. In *24th International Conference (Multichannel Audio, The New Reality)*, 19. Retrieved from: http://www.aes.org/e-lib/browse.cfm?elib=13880

insure4music. (2017). *A History of the PA System*. Retrieved on November 12th, 2020 from https://www.insure4music.co.uk/blog/2017/09/27/a-brief-history-of-the-pa-system/

ITU. (2015). *Report ITU-R BS.2159-6 - Multichannel sound technology in home and broadcasting applications.* Retrieved on November 12th, 2020 from https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-BS.2159-6-2013-PDF-E.pdf

Keller, D. (2019). *Mid/Side Mic Recording Basics*. Retrieved on November 12th, 2020 from https://www.uaudio.com/blog/mid-side-mic-recording/

Kendall, G. (1994). The Effects of Multi-Channel Signal Decorrelation in Audio Reproduction. In *International Computer Music Conference, Aarhus, Denmark.* 319-326.

Kendall, G. (1995). The Decorrelation of Audio Signals and Its Impact on Spatial Imagery. In *Computer Music Journal*, 19(4), 71-87.

Kraugerud, E. (2017). Meanings Of Spatial Formation In Recorded Sound. In *Journal of the Art of Record Production.* Retrieved from: https://www.arpjournal.com/asarpwp/meanings-of-spatial-formation-in-recorded-sound/

L-Acoustics. (2019). *L-ISA - Immersive Hyperreal Sound*. Retrieved on November 12th, 2020 from www.l-isa-immersive.com

Lavoie, S. (2019). My Practice of Live Performance of Spatial Electronic Dance Music. (Doctoral theses). Retrieved on 13th November, 2020 from http://eprints.hud.ac.uk/id/eprint/34954/1/Lavoie%20THESIS.pdf

Lehnert, R. (1994). Frank Zappa on CD (and LP), Part I-III The Yellow Shark. *Stereophile, 17*(2). Retrieved on November 12th, 2020 from https://www.stereophile.com/content/frank-zappa-cd-and-lp-part-i-iii-yellow-shark

Lee, H. & Gribben, C. (2015). Effect of Vertical Layer Spacing for a 3D Microphone Array. *Journal of Audio Engineering Society, 62*(12), 870-884.

Lee, H. (2018). Psychoacoustics of 3D Sound Recording: Research and Practice. In *144th Audio Engineering Society International Pro Audio Convention, Milan, Italy*. Retrieved from https://pure.hud.ac.uk/files/13658202/HLee_Psychoacoustics_of_3D_Sound_Recording.pdf

Lopez-Lezcano, F. (2014). An Architecture for Reverberation in High Order Ambisonics. In *Journal of the Audio Engineering Society 137, Los Angeles, CA, USA.* Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=17432

Los Senderos Studios. (2018). *Surround Sound Microphone Techniques.* Retrieved on November 12th, 2020 from https://lossenderosstudio.com/article.php?subject=17

Lynch, H. & Sazdov, R. (2011). *An Investigation Into Compositional Techniques Utilizes For The Three-Dimensional Spatialization Of Electroacoustic Music*. Retrieved from http://www.ems-network.org/IMG/pdf_EMS11_lynch_sazdov.pdf

mh acoustics (2017). *Eigenbeam DataSheet.* Retrieved on November 12th, 2020 from https://mhacoustics.com/sites/default/files/Eigenbeam%20Datasheet_R01A.pdf

Mac, P. (2009). A Multi-Channel Experiment at Huddersfield University. *Audio Media Magazine*, June, 2009. Retrieved on November 12th, 2020 from https://www.genelec.com/multi-channel-experiment-huddersfield-university

Malecki, P., Piotrowska, M., Sochaczewska, K., & Piotrowski, S. (2020). Electronic Music Production in Ambisonics - Case Study. *J. Audio Eng. Soc., 68(*1-2), 87-94. doi: https://doi.org/10.17743/jaes.2019.0048

Michie, C. (2003). *Frank Zappa: You Call That Music?*. Retrieved on November 12th, 2020 from https://beatpatrol.wordpress.com/2009/03/23/chris-michie-frank-zappa-you-call-that-music-2003/

Mooney, J. (2005). *Sound Diffusion Systems for the Live Performance of Electronic Music. An Inclusive Approach led by Technological and Aesthetically Consideration of the Electroacoustic Idiom and an Evaluation of Existing Systems.* (Doctoral Thesis, The University of Sheffield). Retrieved from: etheses.whiterose.ac.uk/14886/1/425579_vol1.pdf

Moore, A., Moore, D. & Mooney, J. (2004). *M2 Diffusion - The live diffusion of sound in space.* In: Proceedings of the International Computer Music Conference 2004. International Computer Music Conference 2004, 01 - 06 Nov 2004, Miami, Florida, USA. International Computer Music Association

Moore, A. F. (2001). *Rock: The Primary Text - Developing a Musicology of Rock*. (2nd ed.). UK: Ashgate.

Moore, A. F. (2002). Authenticity as Authentication. In *Popular Music 21*(2). 209-223. Retrieved on November 13th, 2020 from https://www.jstor.org/stable/853683?origin=JSTOR-pdf&seq=1

Moylan, W. (2012). Considering Space in Music. In S. Zagorski-Thomas, K. Isakoff & S. Stévance (Eds), *The Art of Record Production: An Auditory Reader for a New Academic Field* (pp. 163-188). London: Routeledge.

Moylan, W. (2015). *Understanding and Crafting the Mix: The Art of Recording* (3rd ed.). US: Focal.

Niklasson, F. (2020). *Sabino Cannone Dives Deep Into Immersive with Spat Revolution*. Retrieved on November 12th, 2020 from https://www.flux.audio/2020/09/11/sabino-cannone-dives-deep-into-immersive-with-spat-revolution/

Nishiyama, T., Nagata, T., Ogata, S., Hasue, T., Kashiwagi, M. (2018). Recording and Mixing Techniques for Ambisonic Sound Production. In *2018 AES International Conference on Spatial Reproduction - Aesthetics and Science.* Retrieved from: http://www.aes.org/e-lib/browse.cfm?elib=19644

Owsinski, B. (2014). *The Recording Engineer's Handbook* (3rd Ed.). US: Course Technology Cengage Learning.

Polymath (2012). *The Decca Sound: Secrets Of The Engineers*. Retrieved on November 12th, 2020 from http://polymathperspective.com/?p=3219 qub. (2020). *Sonic Laboratory.* Retrieved on November 12th, 2020 from https://www.qub.ac.uk/sarc/facilities/soniclab/

Richards, W. (2018). *Alt-J, Royal Albert Hall, London*. Retrieved on November 12th, 2020 from diymag.com/2018/11/05/alt-j-royal-albert-hall-london-live-review

Robjohns, H. (2001, January). Mixing in Surround - Surround Sound Explained. *Sound on Sound*. Retrieved on November 12th, 2020 from https://www.soundonsound.com/techniques/surround-sound-explained-part-1

Robjohns, H. (2001, November). Mixing in Surround - Surround Sound Explained. *Sound on Sound*. Retrieved on November 12th, 2020 from https://www.soundonsound.com/techniques/surround-sound-explained-part-4

Robjohns, H. (2017). *The Double Mid-Sides Array*. Retrieved on November 12th, 2020 from https://www.soundonsound.com/techniques/double-mid-sides-array

Roffler, S. K. & Butler, R. A. (1968). Factors That Influence the Localization of Sound in the Vertical Plane. *Journal of the Acoustical Society of America, 43*(6), 1255-9.

Rumsey, F. (2001). *Spatial Audio*. (1st ed.). Oxford. Focal Press.

Sato, S. & Ando, Y. (2001). Apparent Source Width (ASW) of complex noises in relation to the intramural cross-correlation function. In *17th International Congress on Acoustics, Rome, Italy.* Retrieved from https://www.researchgate.net/publication/260356527_Apparent_source_width_ASW_of_complex_noises_in_relation_to_the_interaural_cross-correlation_function

Sazdov, R., Paine, G. & Stevens, K. (2007) Perceptual Investigation Into Envelopment, Spatial Clarity, and Engulfment In Reproduced Multi-Channel Audio. In *31st International Conference: New Directions in High Resolution Audio.* Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=13515

Sinclair, P. (2014). *EXCLUSIVE: Remixing Tears For Fears and the stereo channel saga*. Retrieved on November 12th, 2020 from https://www.superdeluxeedition.com/feature/exclusive-remixing-tears-for-fears-and-the-stereo-channel-saga/

Smalley, J. (2000). *Gesang der Jünglinge: History and Analysis*. Retrieved from: http://sites.music.columbia.edu/masterpieces/notes/stockhausen/GesangHistoryandAnalysis.pdf

Spors, S. & Rabenstein, R. (2006). Spatial Aliasing Artifacts Produced by Linear and Circular Loudspeaker Arrays used for Wave Field Synthesis. In *120th Convention.* France. Retrieved from: https://www-aes-org.libaccess.hud.ac.uk/tmpFiles/elib/20200906/13515.pdf

SSA Plugins. (2020). *Why Use Seventh Order Ambisonics?* Retrieved on November 12th, 2020 from https://www.ssa-plugins.com/use-seventh-order-ambisonics/

Thornton, M. (2017). *Heads or Tails - Is The Life Channel For Effects Only OR is It OK To Put Other Content In The LFE Channel?*. Retrieved on November 12th, 2020 from https://www.pro-tools-expert.com/home-page/2017/12/11/heads-or-tails-is-the-lfe-channel-for-effects-only-or-is-it-ok-to-put-other-content-in-the-lfe-channel

Thornton, M. (2019). *Object Based Audio Can Do So Much More Than Just Dolby Atmos? We Explore*. Retrieved on November 12th, 2020 from https://www.pro-tools-expert.com/production-expert-1/2019/3/15/is-there-more-to-object-based-audio-than-just-dolby-atmos-we-explore

Thresh, L., Armstrong, C., Kearney, G. (2017). A Direct Comparison of Localisation Performance when using first, third and fifth order ambisonics for real loudspeaker rendering. In *Journal of the Audio Engineering Society 143, New York, NY, USA.* Retrieved from: http://www.aes.org/e-lib/browse.cfm?elib=19261

The Recording Academy's Producers & Engineers Wing. (2004). *Recommendations For Surround Sound Production*. Retrieved on November 12th, 2020 from https://www.grammy.com/sites/com/files/surroundrecommendations.pdf

Rundel, P. (1993). *The Yellow Shark*. [CD - Liner notes]. US: Barking Pumpkin Records.

Travis, C. (2009). A New Mixed-Order Scheme For Ambisonic Signals. In *Ambisonic Symposium 2009, Graz, Austria.* Retrieved on November 12th, 2020 from https://iaem.at/ambisonics/symposium2009/proceedings/ambisym09-travis-newmixedorder.pdf/view

Waldrep, M. (2007). Creating and Delivering High-Resolution Multiple 5.1 Surround Music Mixes. In *31st International Conference: New Directions in High Resolution Audio*. Retrieved from: http://www.aes.org/e-lib/browse.cfm?elib=13951

Waves. (2017). *Mixing in Surround: DOs and DON'Ts*. Retrieved on November 12th, 2020 from https://www.waves.com/mixing-in-surround-do-and-dont

White, P. (2002, February). Mixing in Surround - Surround Sound Explained. *Sound on Sound*. Retrieved on November 12th, 2020 from https://www.soundonsound.com/techniques/surround-sound-explained-part-7

Wittek, H., Haut, C. & Keinath, D. (2006). *Double M/S - a Surround recording technique put to test.* Retrieved on November 12th, 2020 from *https://schoeps.de/fileadmin/user_upload/user_upload/Downloads/Vortraege_Aufsaetze/Papers/SCHOEPS_Double_MS_paper_E_2010.pdf*

Yost, W. A. (2007). *Fundamentals of hearing: an Introduction* (5th Edition). London Academic Press.

Zagorski-Thomas, S. (2006). Functional Staging Through the Use of Production Techniques in the late 20th Century African and Cuban Popular Music Recordings. In *2nd Annual Art of Record Production Conference,* Edinburgh. Retrieved on November 12th, 2020 from http://www.artofrecordproduction.com/aorpjoom/arp-conferences/arp-archive-conference-papers/18-arp-2006/101-zagorski-thomas-2006

Zagorski-Thomas, S. (2010). The Stadium in your bedroom: functional staging, authenticity and the audience-led aesthetic in record production. In *Popular Music 29*(2). 251-266 doi: 10.1017/S0261143010000061

Zagorski-Thomas, S. (2014). *The Musicology of Record Production.* UK: Cambridge University Press.
ZKM. (2020). *Infrastructure*. Retrieved on November 12th, 2020 from https://zkm.de/en/about-the-zkm/organization/hertz-lab/infrastructure

Zotter, F. & Frank, M. (2019). *Ambisonics - A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality* (1st ed.). US: Springer.

Discography:

65daysofstatic. (2013). Unmake the Wild Light. On *Wild Light.* [CD]. Germany: Superball Music.

Boden, J. (2010). Under Their Breath. On *Songs from the Floodplain* [CD]. UK: Navigator Records.

Cooper, A. (1972). Schools Out. On *Schools Out* [CD]. US: Warner Bros. Records.

Drake, N. (1973). Place to Be. On *Pink Moon* [CD]. UK: Island Records.

Dunlop, B. (2016). 356. On *Gilded* [CD]. UK: Gilded Wing Records.

Eilish, B. (2019). you should see me in a crown. On *WHEN WE ALL FALL ASLEEP, WHERE DO WE GO?.* [CD]. US: Darkroom/Interscope Records.

Gaye, M. & Terrell, T. (1967). Ain't No Mountain High Enough. On *United.* [CD]. US: Tamla.

Godspeed You! Black Emperor. (1998) Blaise Bailey Finnegan III. On *Slow Riot for New Zero Kanada.* [E.P.] US: Kranky.

Godspeed You! Black Emperor. (1998) Moya. On *Slow Riot for New Zero Kanada.* [E.P.] US: Kranky.

Hendrix, J. (1968). *Electric Ladyland.* [CD]. US: Reprise Records.

Hendrix, J. (1968). Crosstown Traffic. On *Electric Ladyland.* [CD]. US: Reprise Records.

Horikawa, Y. (2012). *Wandering*. [EP]. UK: First World Records.

Horikawa, Y. (2012). Letter. On *Wandering*. [EP]. UK: First World Records.

Kraftwerk. (2017). *3-D The Catalogue.* [Blu-Ray]. UK: Parlophone.

Karnivool. (2013). We Are. On *Asymmetry.* [CD]. Australia: Sony Music.

Nine Inch Nails. (2013). Copy Of A. On *Hesitation Marks*. [CD]. US: Columbia.

Pink Floyd. (1968). Set The Controls For The Heart Of The Sun. On *A Saucerful Of Secrets*. [CD]. US: Columbia

Pomplamoose. (2018). Jamirobeegees Mashup: Staylin' Alive / Virtual Insanity. On *Best of 2018.* [CD]. Self Release.

Polyphia. (2018). O.D.. On *New Levels New Devils.* [CD]. US: Equal Vision Records.

The Prodigy. (1997). Firestarter. On *The Fat of the Land.* [CD]. UK: XL Recordings.

Snarky Puppy. (2014). Tio Macaco. On *We Like It Here.* [CD]. US: GroundUp Music.

Tears for Fears. (1984). *Songs from the Big Chair.* [CD]. US: Universal.

Tears for Fears. (2014). *Songs from the Big Chair.* [Blu-Ray - CD]. US: Universal.

Tears for Fears. (2014). Everybody Wants to Rule the World. On *Songs from the Big Chair.* [Blu-Ray - CD]. US: Universal.

Tears for Fears. (2014). Shout. On *Songs from the Big Chair.* [Blu-Ray - CD]. US: Universal.

The Beatles. (1967). A Day in the Life. On *Sgt. Peppers Lonely Hearts Club Band.* [CD]. UK: Parlophone.

The Beatles. (2017). *Sgt. Peppers Lonely Hearts Club Band.* [Blu-Ray - CD]. US: Universal Music Enterprises.

The Beatles. (2017). A Day in the Life. On *Sgt. Peppers Lonely Hearts Club Band.* [Blu-Ray - CD]. US: Universal Music Enterprises.

Villagers. (2016). The Waves. On *Where Have You Been All My Life?* [CD]. UK: Domino Recording.

Vulfpeck. (2017). Running Away (feat. Joey Dosik, David T. Walker & James Gadson). On *Mr. Finish Line.* [CD]. US: Vulf Records.

Zimmer, H. (2017). *Live in Prague*. [Blu-Ray]. US: Eagle Records.

Zimmer, H. (2017). Pirates of the Caribbean. On *Live in Prague*. [Blu-Ray]. US: Eagle Records.

Videos:

Bagshaw, J. (Speaker). (2019). *James Bagshaw: Towards a Spatial Language for Popular Music in Surround Sound.* [Youtube Video]. Retrieved on 13th November, 2020 from https://www.youtube.com/watch?v=1ryc3nUyjs4

Bon Iver. (2019, July 31). Bon Iver: Autumn. [Youtube Video]. Retrieved on November 12th, 2020 from https://youtu.be/ne7VX3sb4rQ

FLUX:: Immersive. 2020, April 15). *Introduction to Spat Revolution* [Youtube Video]. Retrieved on November 12th, 2020 from youtube.com/watch?v=x3U3scY2_E8.

SCHOEPS Mikrofone. (2018, April 28). *Hyunkook Lee: Psychoacoustics of 3D sound recording techniques: Research and Practice* [Youtube Video]. Retrieved on November 12th, 2020 from https://www.youtube.com/watch?v=5SsQdfKAhVA

Software:

Ableton. (2019). Ableton Live 10 [Digital Audio Workstation]. (10.1). Retrieved from: https://www.ableton.com/en/shop/live/

Ambisonics Toolkit. (2016). The Ambisonic Tookit [Musical Plugins]. (1.0). Retrieved from: http://www.ambisonictoolkit.net

Apple. (2019). Logic Pro X [Digital Audio Workstation]. (10.4.8). Retrieved from: https://apps.apple.com/gb/app/logic-pro-x/id634148309?mt=12

Cuckos. (2019). Reaper [Digital Audio Workstation]. (6.0). Retrieved from: http://reaper.fm

cycling74. (2019). Max 8 [Visual programming software]. (8.0.2). Retrieved from: https://cycling74.com

DMG Audio. (2019). Equilibrium. [Musical Plugin]. (1.59). Retrieved from: https://www.dmgaudio.com/equilibrium

Facebook360. (2020). Spatial Workstation. [Musical Plugin]. (3.3.3). Retrieved from https://facebook360.fb.com/spatial-workstation/ On 1st September 2020

Flux Audio. (2019). Alchemist [Musical Plugin]. (3.0). Retrieved from: https://www.flux.audio/project/alchemist-v3/

IEM (2019). IEM Plugin Suite [Musical Plugins]. (1.11.0). Retrieved from: https://plugins.iem.at

iZotope (2013). RX 3 Advanced [Audio Repair Software]. (v3.01.734). Retrieved from: https://www.izotope.com/en/products/rx.html

Schoeps. (2018). Schoeps Double MS. [Musical Plugin]. (1.0). Retrieved from: https://www.plugin-alliance.com/en/products/schoeps_double_ms.html

Sennheiser. (2019). AMBEO A-B Format Converter. [Musical Plugins]. (1.2.2). Retrieved from: https://en-uk.sennheiser.com/microphone-3d-audio-ambeo-vr-mic

Slate Digital. (2019). Virtual Mix Rack. [Musical Plugins]. (2.4.4.3). Retrieved from: https://slatedigital.com

Spectrasonics. (2019). Keyscape. [Sample Library]. (1.1). Retrieved from: https://www.spectrasonics.net/products/keyscape/index.php

Waves. (2019). R360 Surround Reverb. [Musical Plugin]. (11). Retrieved from: https://www.waves.com/plugins/r360-surround-reverb

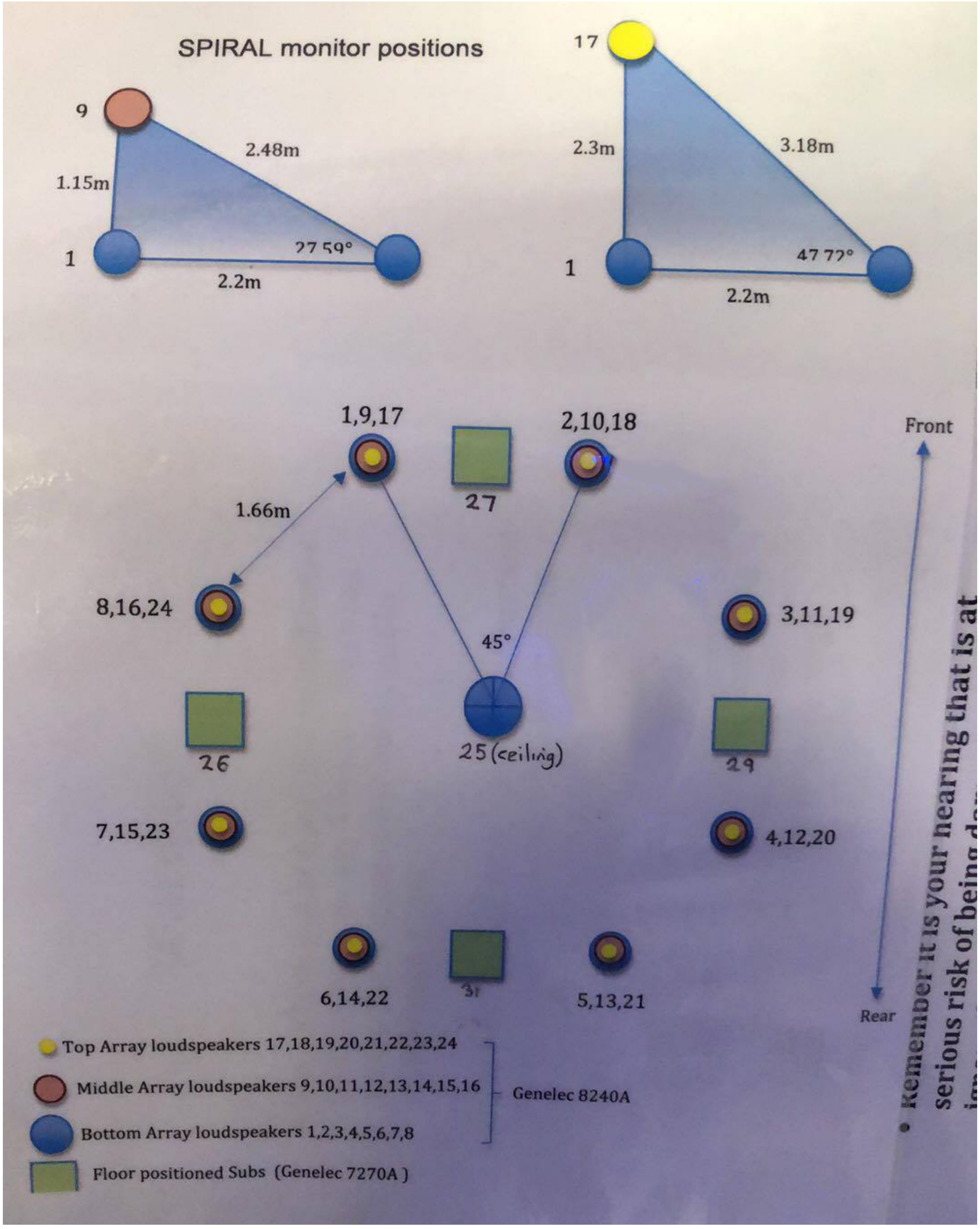
Appendices:

Literature Review Appendices

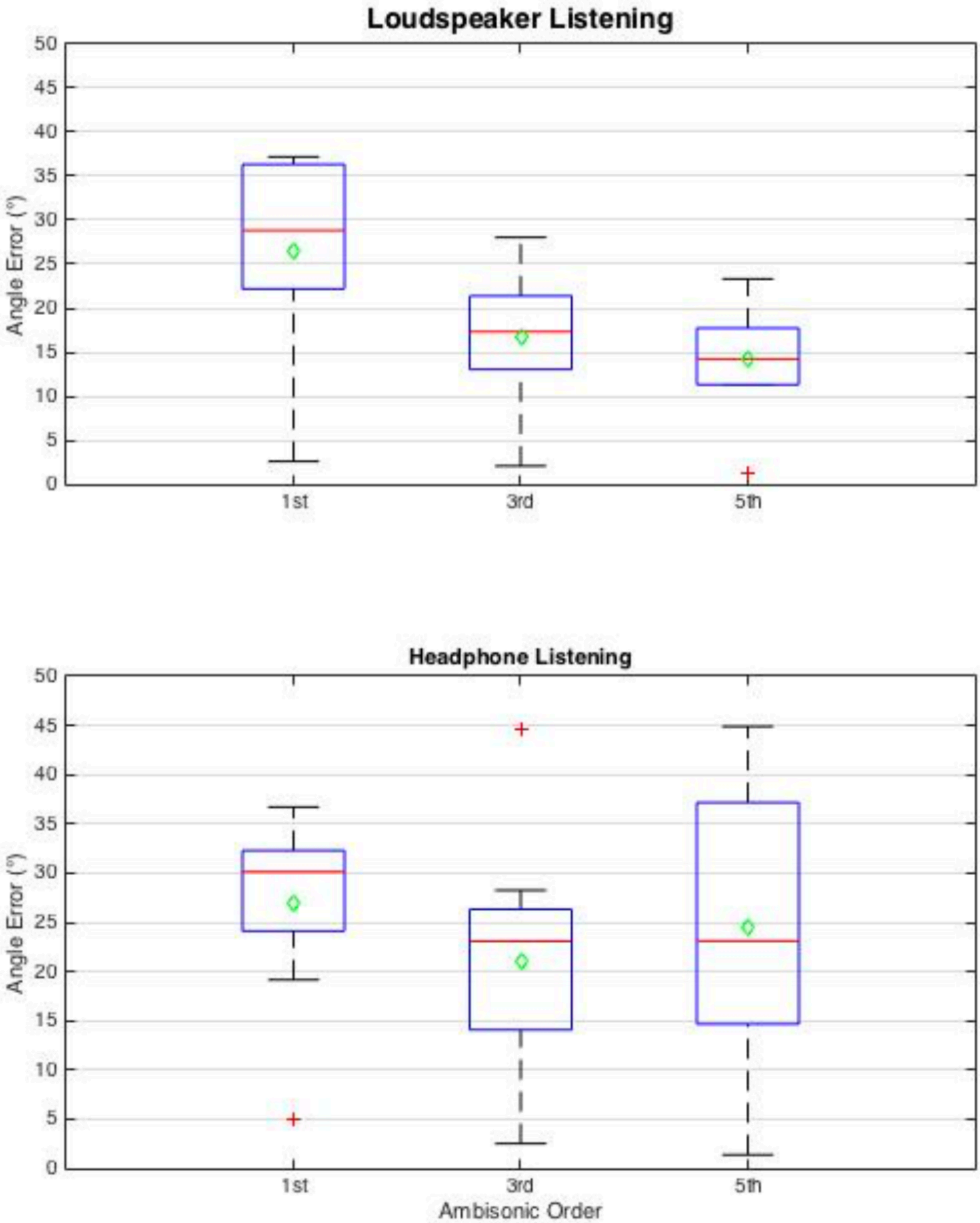
Appendix 1 - Spiral azimuth, elevation and delay times required.

Speaker No.	Position	Azimuth º	Elevation º	Delay (ms)
1	Lower Front-Front L	22.5	0	5.83
2	Lower Front-Front R	-22.5	0	5.83
3	Lower Front R	-67.5	0	5.83
4	Lower Rear R	-112.5	0	5.83
5	Lower Rear-Rear R	-157.5	0	5.83
6	Lower Rear-Rear L	157.5	0	5.83
7	Lower Rear L	112.5	0	5.83
8	Lower Front L	67.5	0	5.83
9	Mid Front-Front L	22.5	27.59	5.01
10	Mid Front-Front R	-22.5	27.59	5.01
11	Mid Front R	-67.5	27.59	5.01
12	Mid Rear R	-112.5	27.59	5.01
13	Mid Rear-Rear R	-157.5	27.59	5.01
14	Mid Rear-Rear L	157.5	27.59	5.01
15	Mid Rear L	112.5	27.59	5.01
16	Mid Front L	67.5	27.59	5.01
17	Upper Front-Front L	22.5	47.72	2.97
18	Upper Front-Front R	-22.5	47.72	2.97
19	Upper Front R	-67.5	47.72	2.97
20	Upper Rear R	-112.5	47.72	2.97
21	Upper Rear-Rear R	-157.5	47.72	2.97
22	Upper Rear-Rear L	157.5	47.72	2.97
23	Upper Rear L	112.5	47.72	2.97
24	Upper Front L	67.5	47.72	2.97
25	Top (Voice of God)	0	90	0
26	Sub Left	90	0	0
27	Sub Front	0	0	0
28	Sub Right	-90	0	0
29	Sub Rear	180	0	0

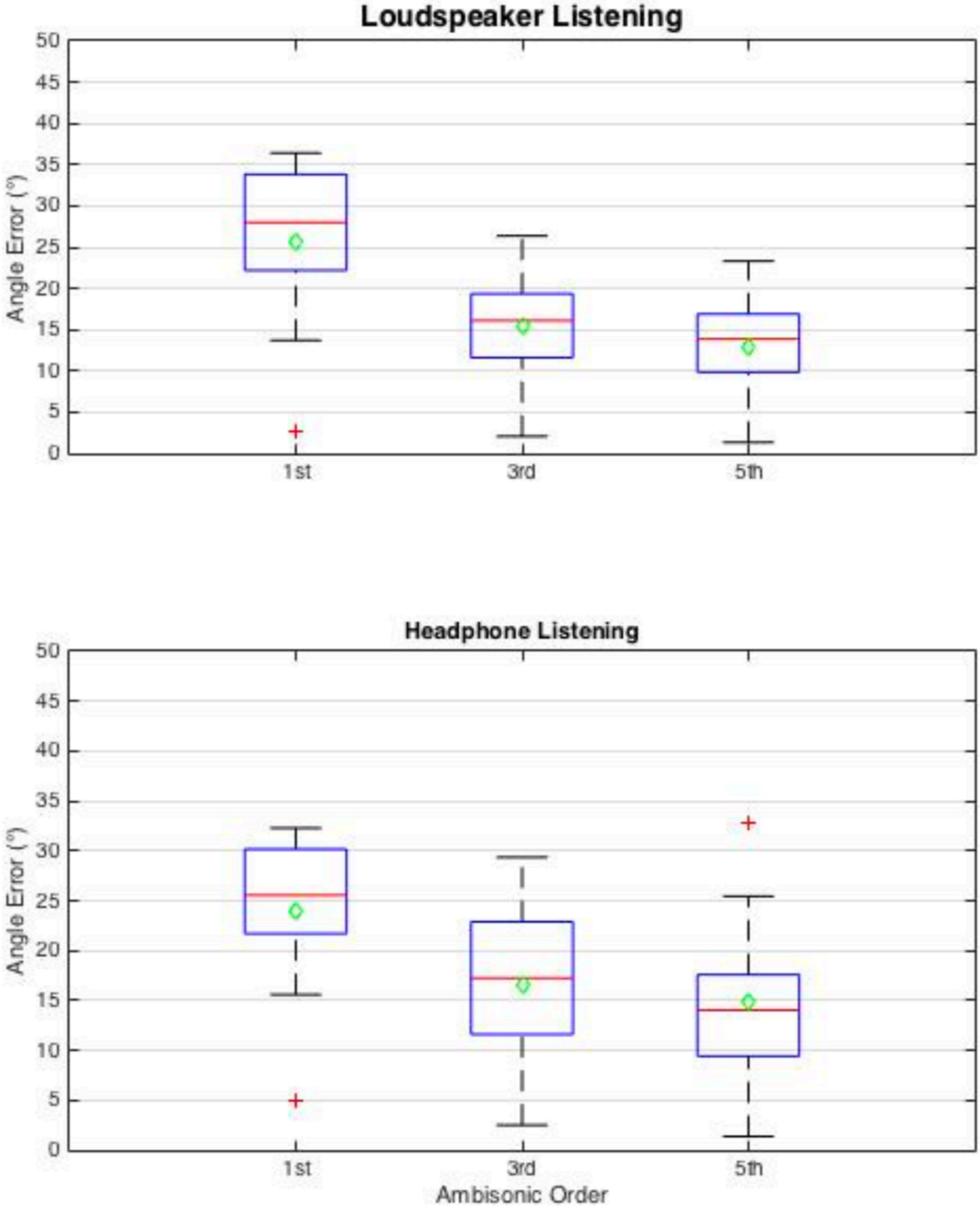
Appendix 2 - SPIRAL monitor positions diagram. From The University of Huddersfield.



Appendix 3 - Box-plot of Average Angle Error Across All Given Angles For Each Ambisonic Order (Thresh, Armstrong and Kearney, 2017)



Appendix 4 - Box-plot of Average Angle Error Across All Given Angles For Each Ambisonic Order With Removed Outliers (Thresh, Armstrong and Kearney, 2017)



Project 1 Appendices

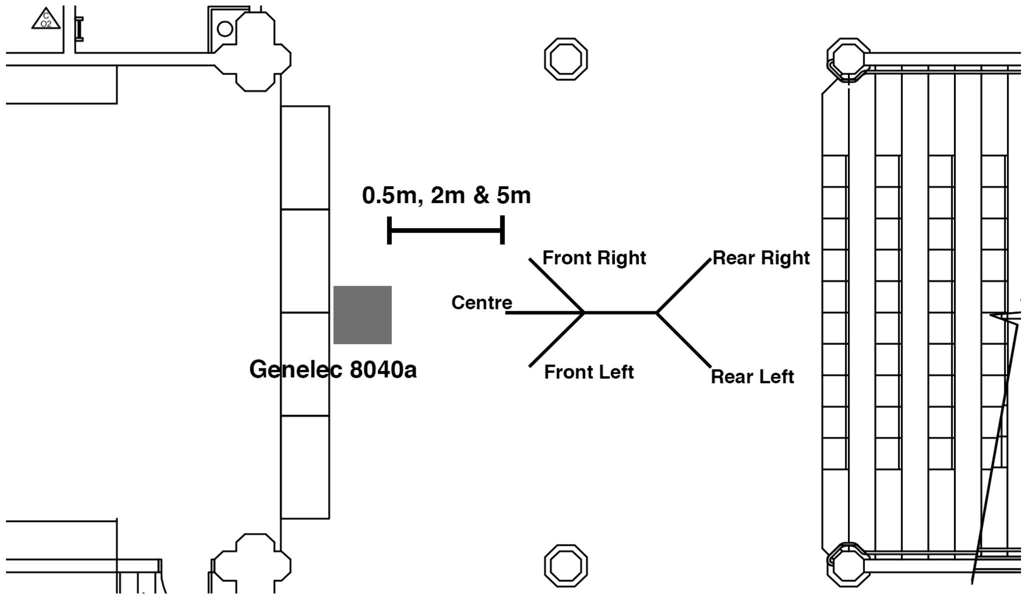
Appendix 5 - Project 1 - Charlie Hulejczuk Audio Example Markers in Reaper Project Descriptions

Marker	Technique	Notes
1	Dry Recordings	Track 1 - *“Cursed”*
2		Track 2 - *“Blessed by the Blossom”*
3		Track 3 - *“Waste Our Time”*
4		Track 4 - *“The Duchess of Westdene”*
5	Decca Tree Surround Setup	5o at 2m distance, 1.5m height
6		3o2c at 2m distance, 1.5 height
7		5c at 2m distance, 1.5m height
8		5o at 0.5m distance, 1.5m height
9		5o at 2m distance, 1.5m height
10		5o at 5m distance, 1.5m height
11		5o at 0.5m distance, 2.5m height
12		5o at 2m distance, 2.5m height
13		5o at 5m distance, 2.5m height
14	Double Mid Side Setup	0.5m distance, 1.5m height
15		1.5m distance, 1.5m height
16		3m distance, 1.5m height
17		5m distance, 1.5m height
18		7m distance, 1.5m height
19		0.5m distance, 2.5m height
20		1.5m distance, 2.5m height
21		3m distance, 2.5m height
22		5m distance, 2.5m height
23		7m distance, 2.5m height
24		1.5m behind the speaker
25	Hamasaki Square Setup	5m distance
26		7m distance
27		Surrounding the speaker
28	Sennheiser Ambeo Setup	0.5m distance
29		1.5m distance
30		3m distance
31		5m distance

Appendix 6 - Dry Recording Photo (Front View)

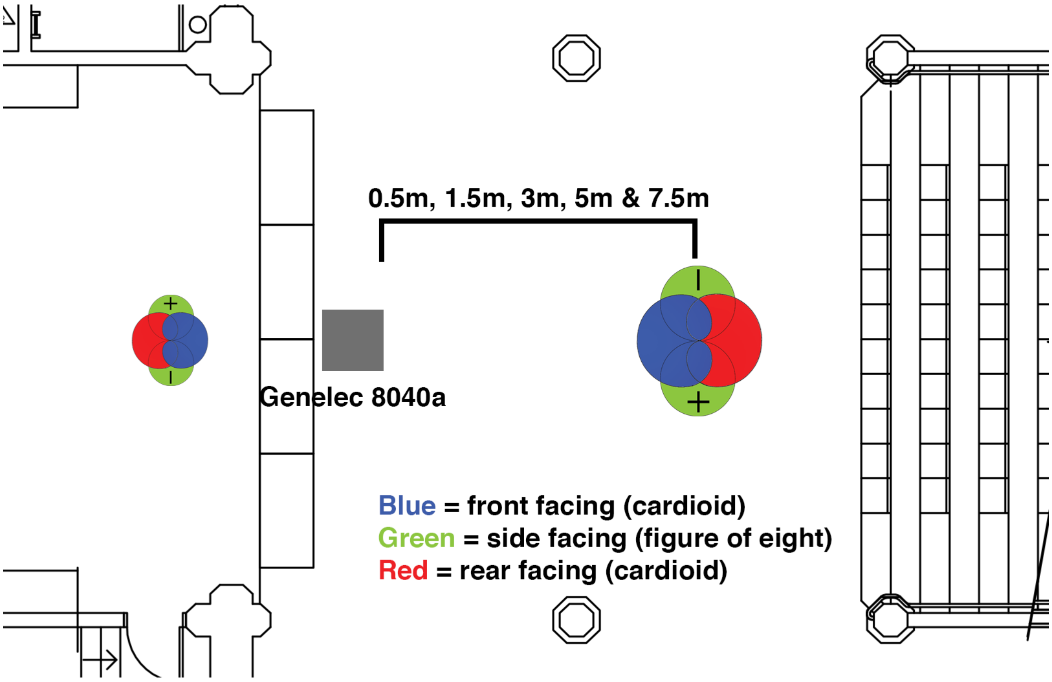
Appendix 7 - Dry Recording Photo (Side View)

Appendix 8 - Decca Tree Surround birds-eye view diagram

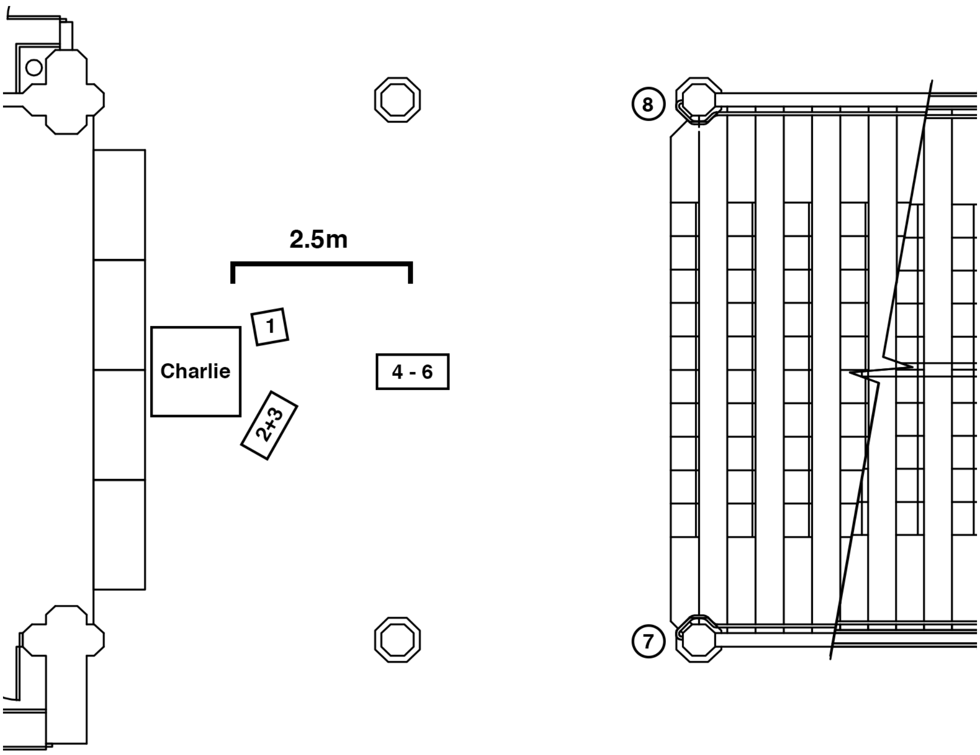


Appendix 9 - Decca Tree Surround Setup in St. Paul’s Hall

Appendix 10 - Double Mid Side birds-eye-view diagram



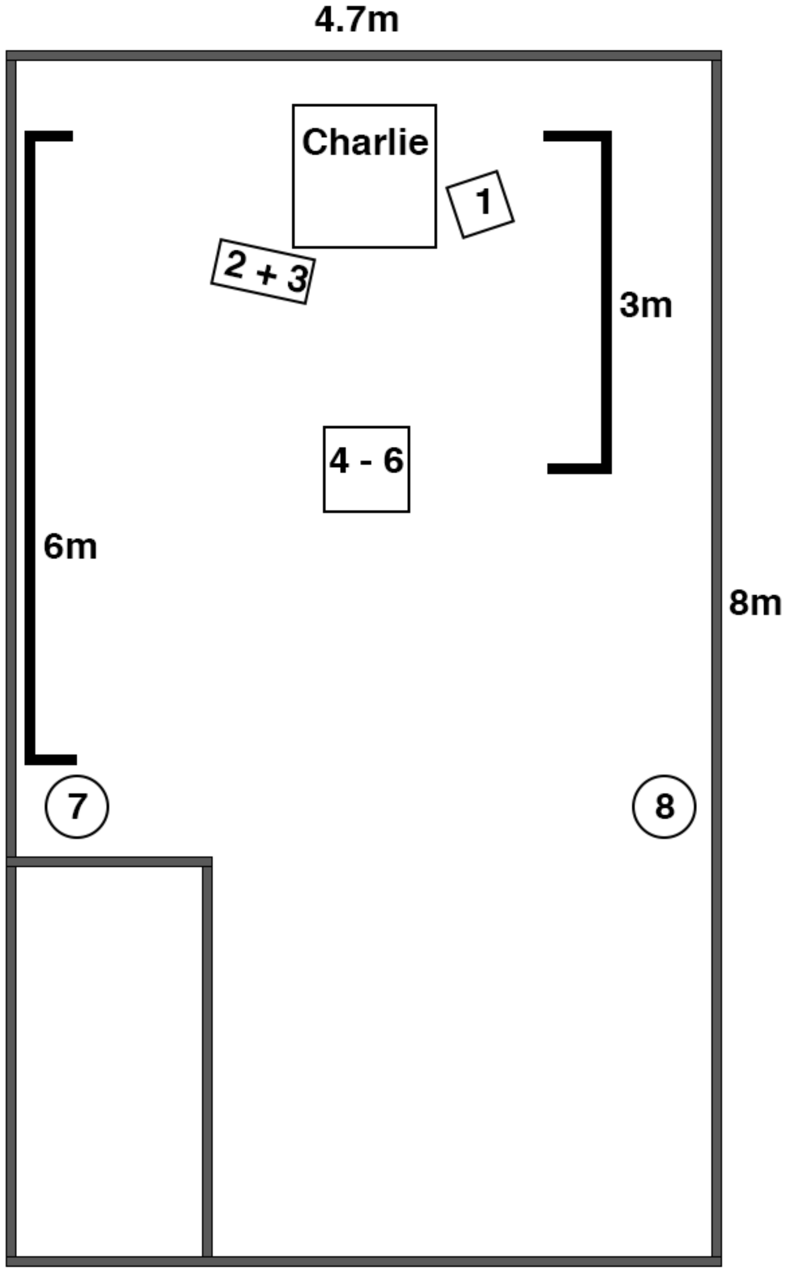
Appendix 11 - “Cursed” Re-recorded Floor Plan



Appendix 12 - “Cursed” Recording Session Image (Rear View)

Appendix 13 - “Cursed” Recording Session Image (Rear View)

Appendix 14 - “Blessed by the Blossom” Floor Plan

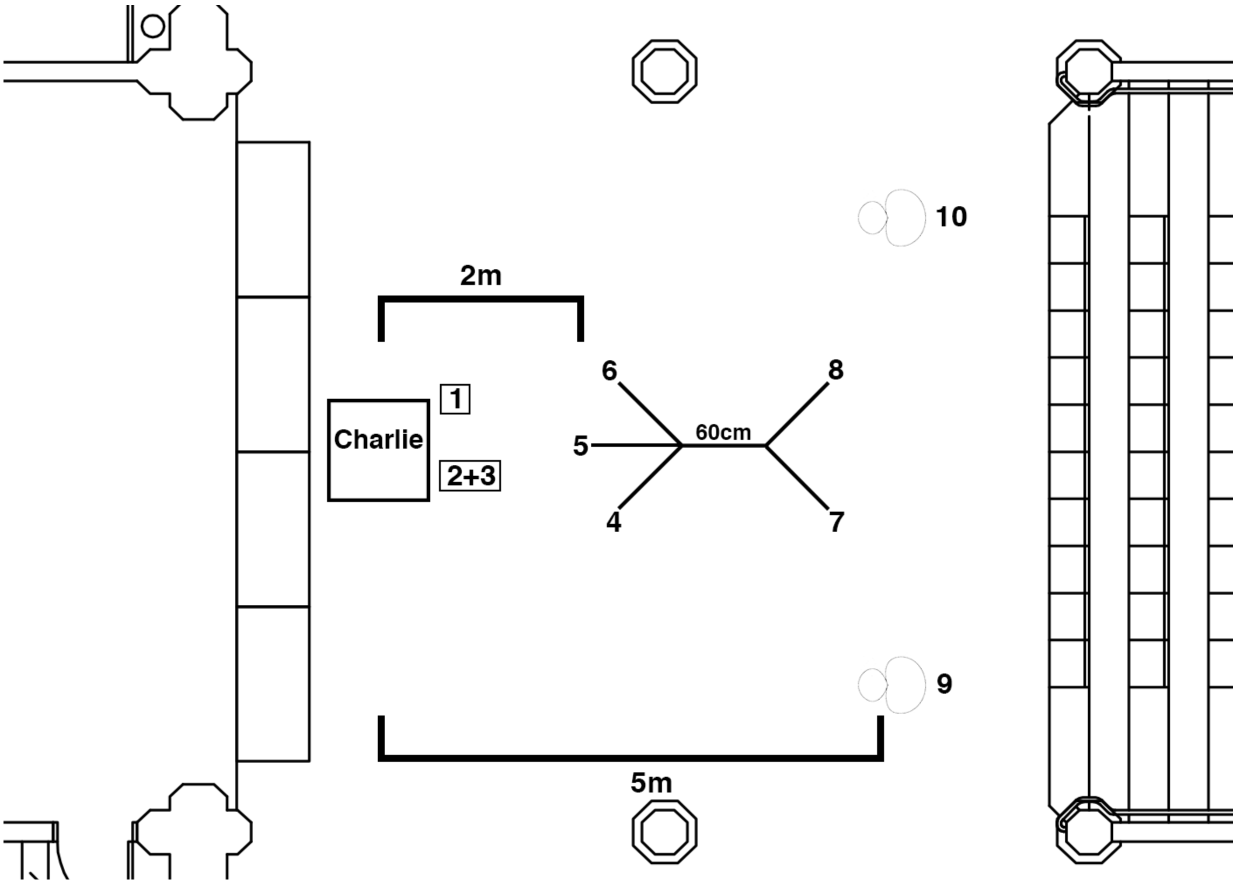


Appendix 15 - “Blessed by the Blossom” Recording Session Image (Side View)

Appendix 16 - “Blessed by the Blossom” Recording Session Image (Front View)

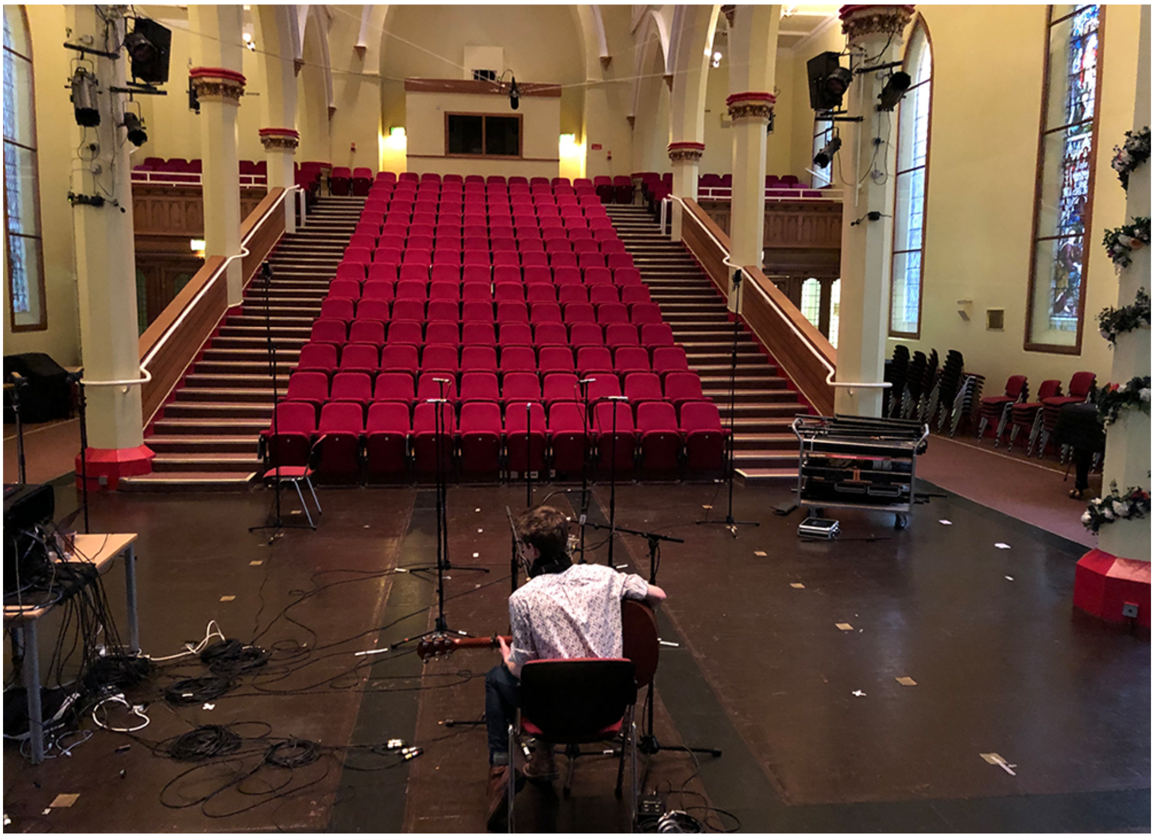
Appendix 17 - “Blessed by the Blossom” Recording Session Image (Rear View)

Appendix 18 - “Waste Our Time” Floor Plan

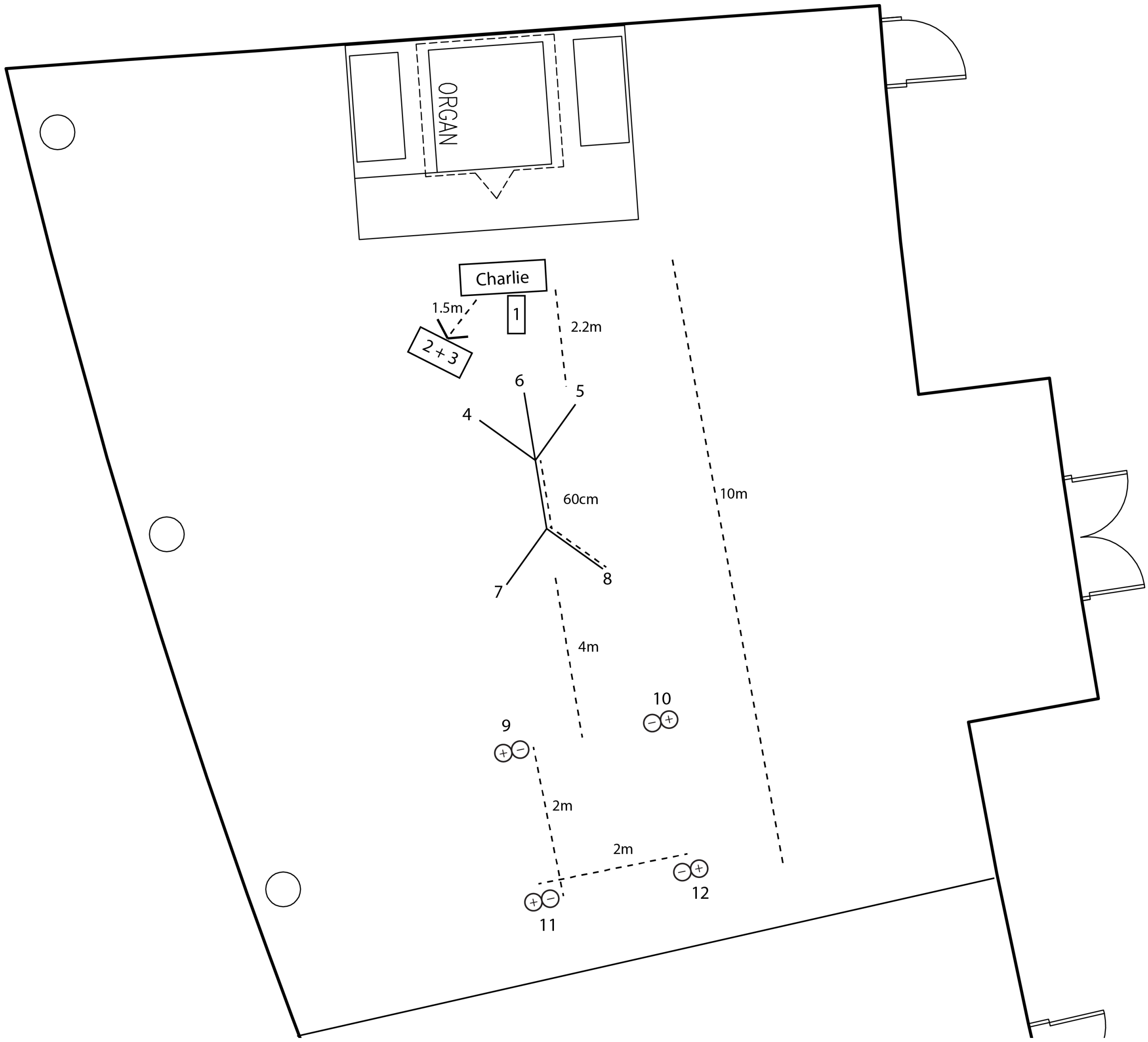


Appendix 19 - “Waste Our Time” Recording Session Image (Front View)

Appendix 20 - “Waste Our Time” Recording Session Images (Rear View)



Appendix 21 - “The Duchess of Westdene” Floor Plan



Appendix 22 - “The Duchess of Westdene” Recording Image (front view)

Appendix 23 - “The Duchess of Westdene” Recording Image (side view)

Project 2 Appendices

Appendix 24 - Percussion and synth channel list

Channel:	Instrument:		Channel:	Instrument:
1	Kick		7	Cymbals 1
2	Snare		8	Cymbals 2
3	Snare 2		9	Percussion 1
4	Tom 1		10	Percussion 2
5	Tom 2		11	Synth Channels 1
6	Ride		12	Synth Channels 2

Appendix 25 - Bass Amp Microphones

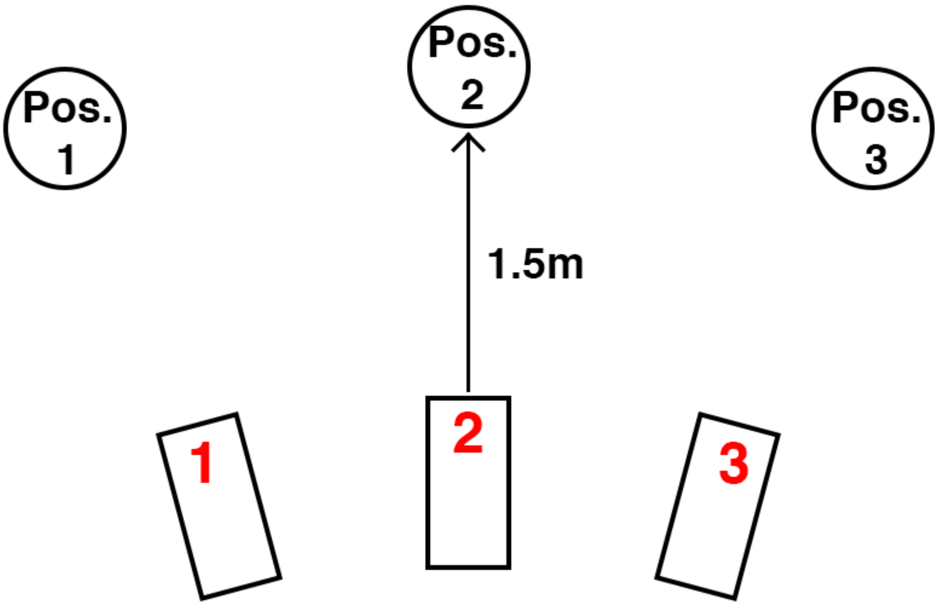
Channel	Microphone	Position	Notes
1	Audio Technica AE2500 (dynamic capsule)	Directly on the centre of the speaker cone	Well rounded and low
2	Audio Technica AE2500 (condenser capsule)	Directly on the centre of the speaker cone	Well rounded and less bright
3	AKG C12 (cardioid)	2cm off centre, pointing directly on	Gritty and bright
4	Sennheiser MD421	Directly on the centre of the speaker cone	Lows
5	AKG D112	On centre of the speaker cone pointing towards edge	Lows with a bit more top end
6	Neumann M149 (cardioid)	60cm away from cabinet	Distinctive bad tone
7	Coles 4038	60cm away from cabinet very close to the floor	Mid heavy and roomy

Appendix 26 - Annotated picture of the microphones used in the re-amping session

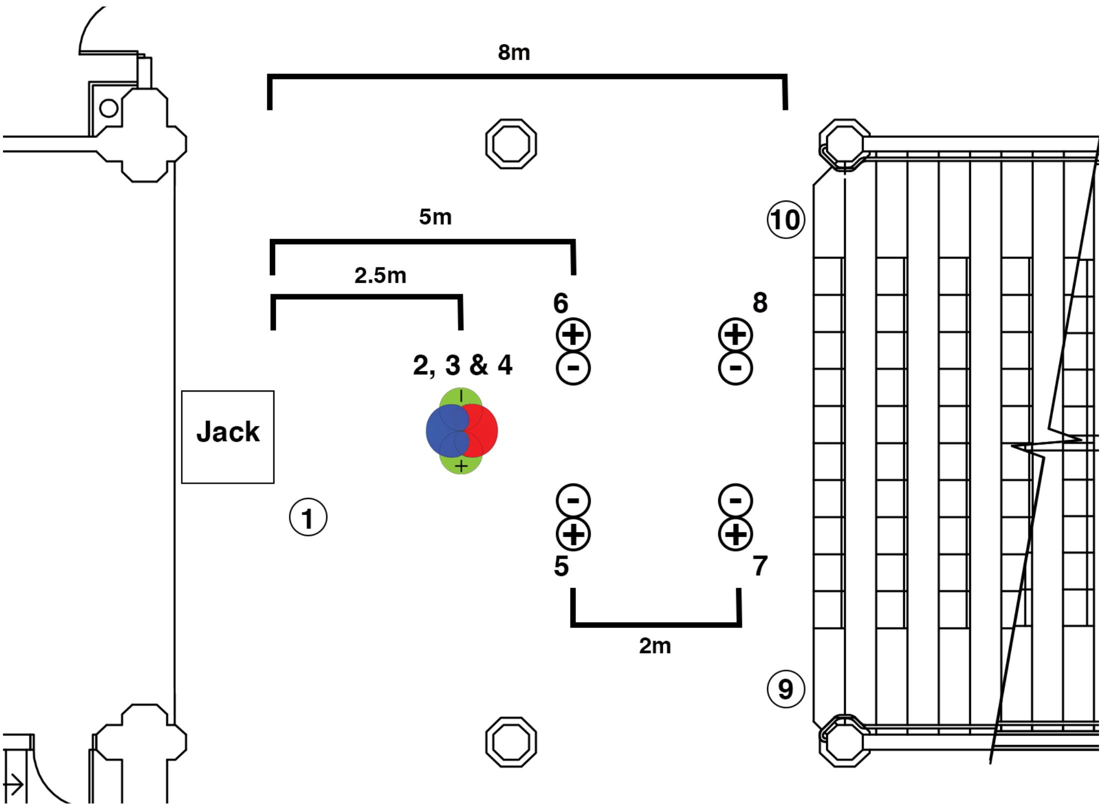
Appendix 27 - Viola Recording channel list

Channel	Microphone	Position
1	DPA 4060	Attached to the viola, under the bridge
2	AKG C414 (cardioid)	Room Left
3	AKG C414 (cardioid)	Room Centre
4	AKG C414 (cardioid)	Room Right

Appendix 28 - Viola Recording Positioning



Appendix 29 - Solo Viola bird's-eye view



Appendix 30 - Solo Viola recording plan

Channel	Position	Microphone	Notes
1	Close Mic	Neumann KMD w/ KK131 omni capsule	1m away
2	Double Mid Side - Front	Neumann KMD w/ KK184 cardioid capsule	2.5m away at viola height
3	Double Mid Side - Side	Neumann KMD w/ KK120 figure-of-eight capsule	
4	Double Mid Side - Rear	Neumann KMD w/ KK184 cardioid capsule	
5	Hamasaki Square - Front Left	Schoeps CCM8	Positioned 5m away at 2.5m height. Spacing 2m between each microphone
6	Hamasaki Square - Front Right		
7	Hamasaki Square - Rear Left		
8	Hamasaki Square - Rear Right		
9	Outrigger - Left	Neumann KMD w/ KK131 omni capsule	Positioned 8m away at 4m height.
10	Outrigger - Right	Neumann KMD w/ KK131 omni capsule	

Appendix 31 - “tent.legs.” viola solo front view

Appendix 32 - “tent.legs.” viola solo side view

Appendix 33 - “tent.legs.” viola solo double mid side setup

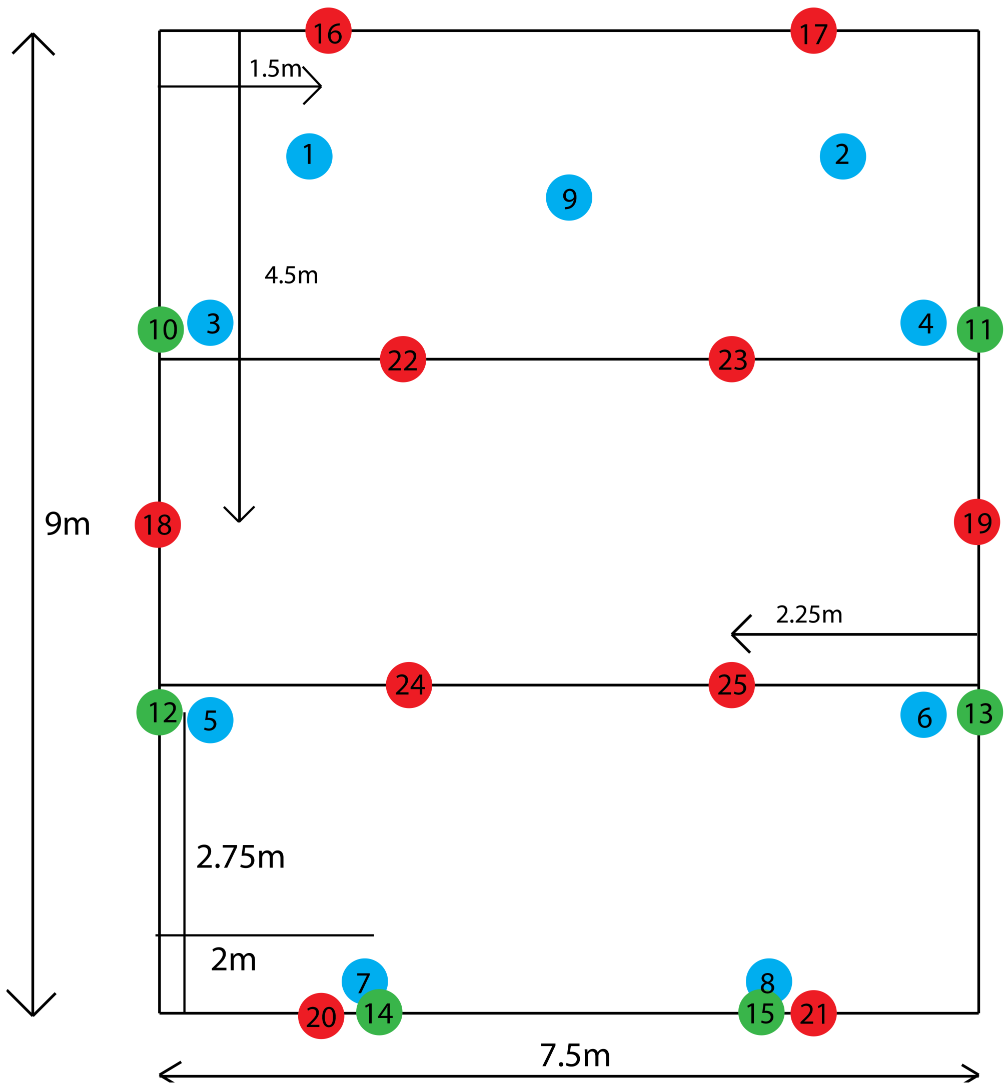
Project 3 Appendices

Appendix 34 - Yamaha CL5 desk routing

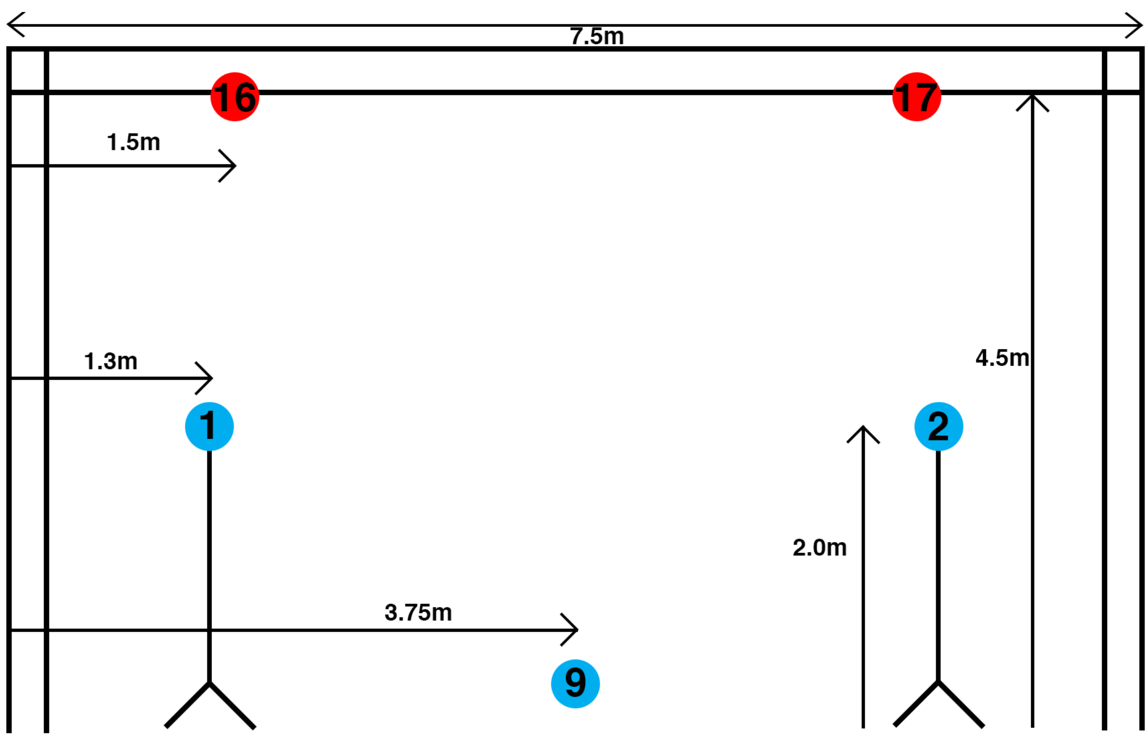
Mix	Speaker No.	Name
1		Stage Monitor L
2		Stage Monitor CL
3		Stage Monitor CR
4		Stage Monitor R
5	1	Ring 1 - Front L
6	2	Ring 1 - Front R
7	3	Ring 1 - Front-Side L
8	4	Ring 1 - Front-Side R
9	5	Ring 1 - Rear-Side L
10	6	Ring 1 - Rear-Side R
11	7	Ring 1 - Rear 1
12	8	Ring 1 - Rear 2
13	9	Ring 1 - Punch
14	10	Ring 2 - Front-Side L
15	11	Ring 2 - Front-Side R
16	12	Ring 2 - Rear-Side L
17	13	Ring 2 - Rear-Side R
18	14	Ring 2 - Rear L
19	15	Ring 2 - Rear R
20	16	Ring 3 - Front L
21	17	Ring 3 - Front R
22	18	Ring 3 - Side L
23	19	Ring 3 - Side R
24	20	Ring 3 - Rear L

Matrix	Speaker No.	Name
1		Subs
2	21	Ring 3 - Rear R
3	22	Ring 4 - Front L
4	23	Ring 4 - Front R
5	24	Ring 4 - Rear L
6	25	Ring 4 - Ring R

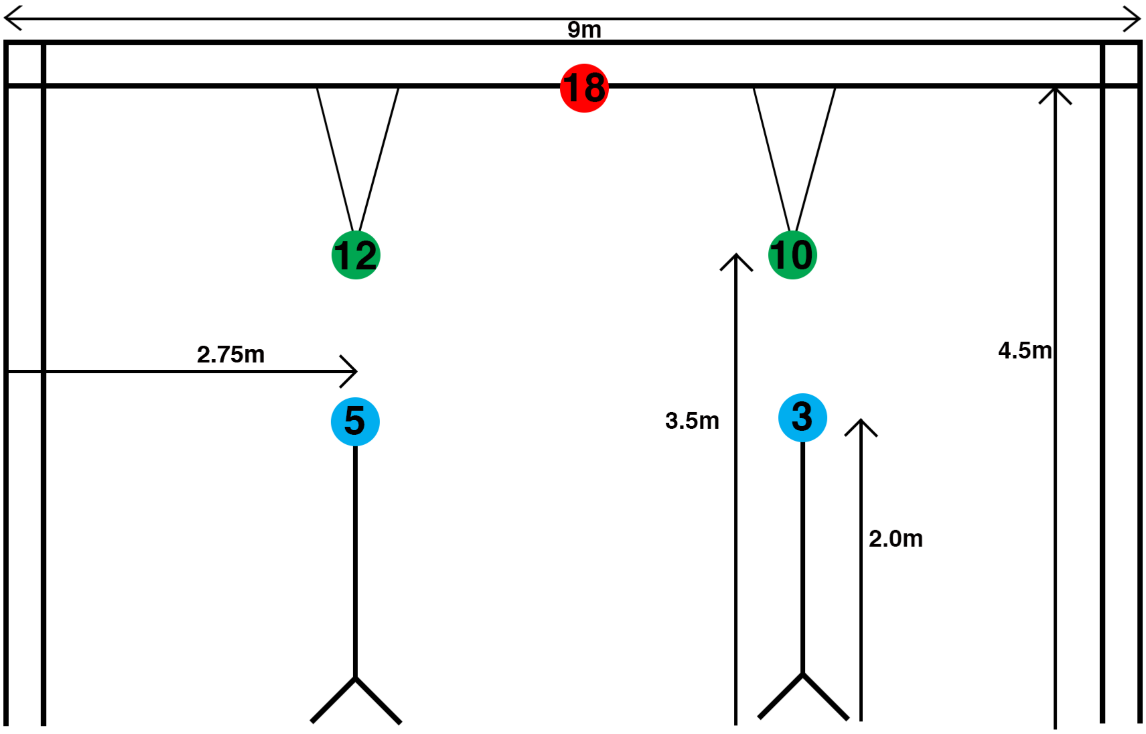
Appendix 35 - Top Down Speaker Positions



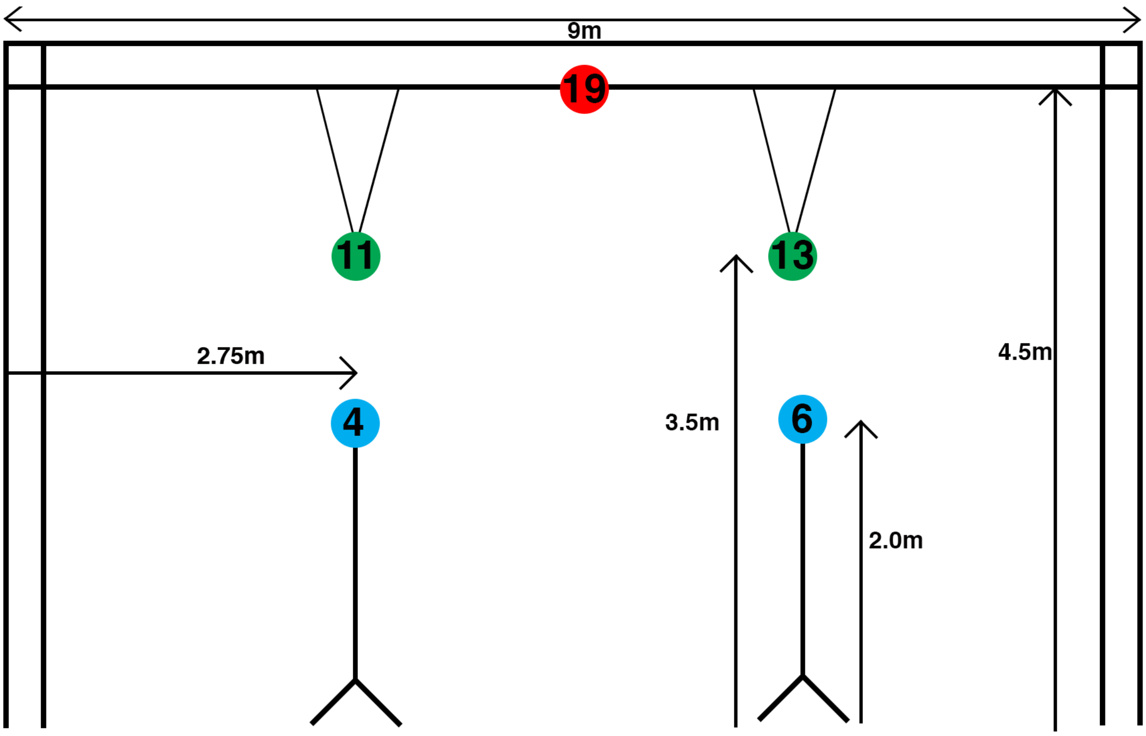
Appendix 36 - Front View



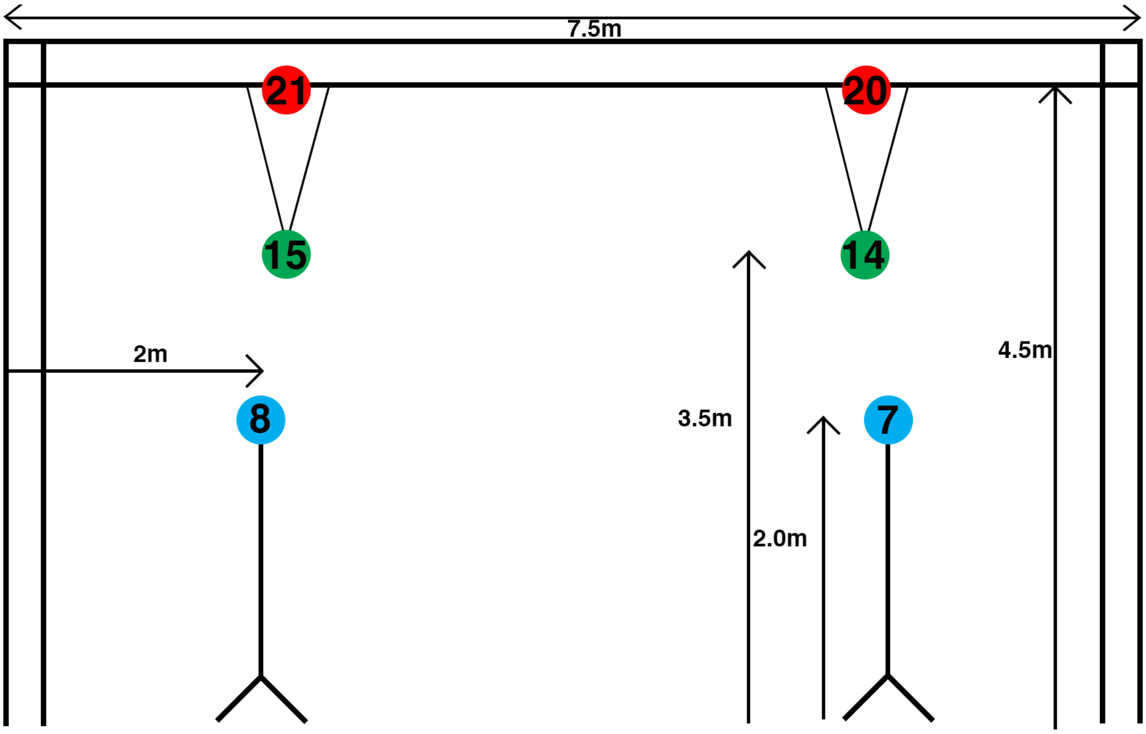
Appendix 37 - Left Side View



Appendix 38 - Right Side View



Appendix 39 - Rear View

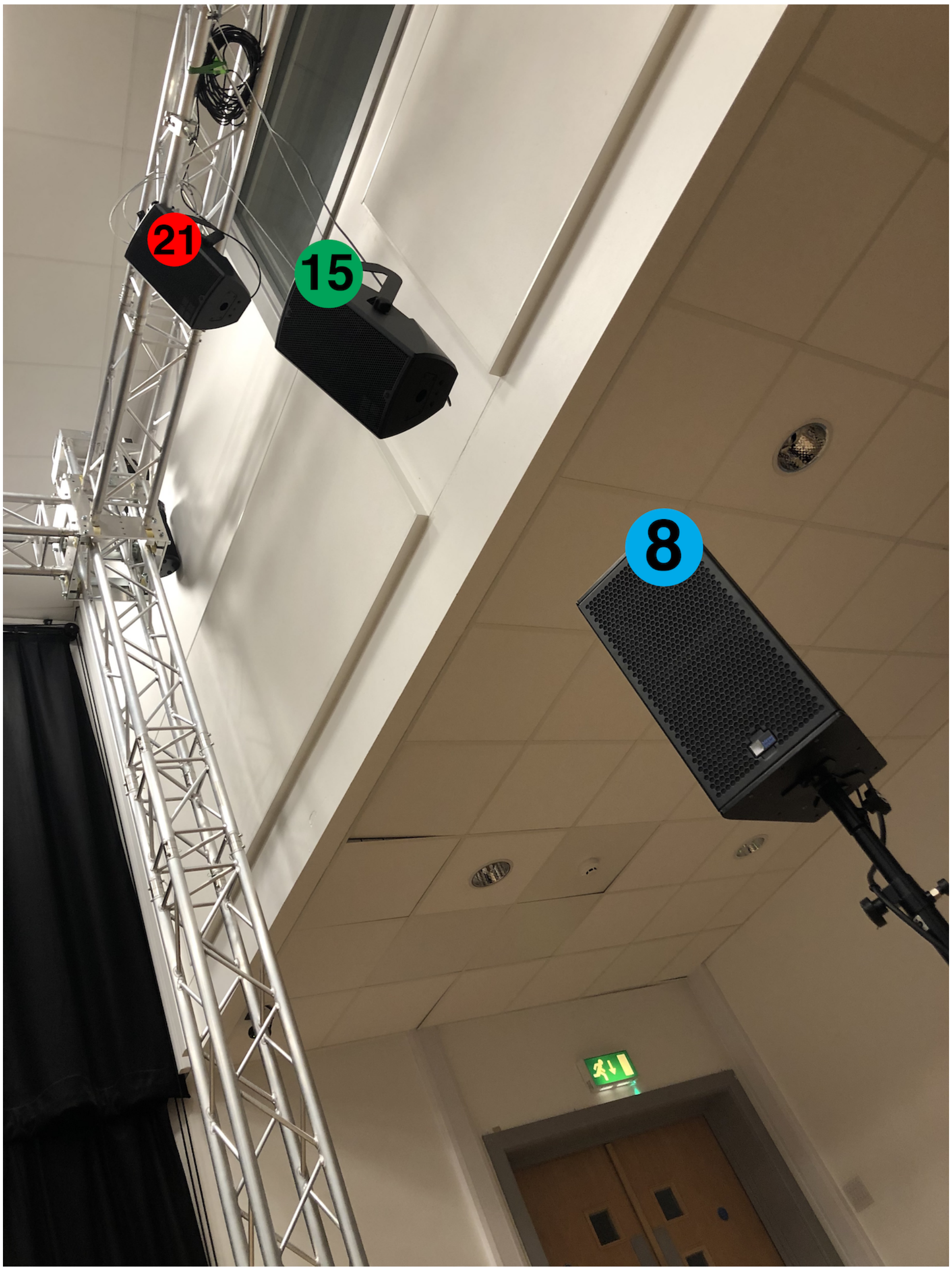


Appendix 40 - Front View Annotated

Appendix 41 - Side View Annotated

Appendix 42 - Rear View Annotated

Appendix 43 - Up View Annotated



Appendix 44 - Yan's Laptop Stems

Channel no.	Instrument
1	Kick
2	Snare
3	Hi hat
4	Tom 1
5	Tom 2
6	Cymbal 1
7	Cymbal 2
8	Synths 1
9	Synths 2
10	Ambiences