MIXING SCENARIOS FOR B+ FORMAT AUDIO

Submitted in partial fulfillment of the requirements for
the Master of Music in Music Technology
in the department of Music and Performing Arts Professions
in the Steinhardt School of Culture, Education, and Human Development

New York University
Advisor: Dr. Kenneth Peacock
Technical Advisor: Paul Geluso

May 5, 2009
# Table of Contents

Acknowledgments iii

Abstract 1

I. Introduction and Rationale 2
   1. Introduction 2
   2. Rationale 2
   3. History 5

II. B+ Format Demystified 10
   1. B Format Ambisonics 10
   2. B+ Format Explained 16
   3. Advantages and Disadvantages of B+ Format Audio 18
   4. Uses for B+ Format Audio 19

III. Creation of Adaptive Matrix 21
   1. Explanation and Method for Adaptive Matrix Creation 21
   2. Stereo Adaptive Matrix to B+ Format 24
   3. 5.1 Surround Sound Adaptive Matrix to B+ Format 30
   4. Multitrack Session Adaptive Matrix to B+ Format 33
   5. Adaptive Matrix Summary 35
IV. Informal Evaluation (Listening Test) 36

1. Listening Test Procedure 36
2. System and Setup Validation 38
3. Listening Test 40
4. Listening Test Results 42

V. Conclusions 44

Selected References 46

Appendix A: Speaker Location Impulse Responses 47
Appendix B: Sine Sweep Analysis (20Hz-20kHz) 54
Appendix C: Listening Position Selected Tone Analysis 57
Appendix D: Survey Questions 62
Acknowledgments

Thank you to Dr. Peacock and Paul Geluso for your time and advisement at New York University. A special thank you to my mixing partner, Chris Camilleri shared many of his professional mixes. In addition, a special thank you to the musicians: Mr. Rob Mathes, Alvin Shelby, Tricia Walker, Davis Raines, Natalie and Brian Langlinais, Bury Me A Lion, and Britney Boras for their music and multi-track recordings. For assistance and support, thank you to the faculty and staff at New York University. To my wife, family, friends, and educators, thank you for your encouragement.
Abstract

Mix engineers prepare music and program material for a variety of different playback mediums. Stereo and 5.1 Surround Sound is an “industry standard” but additional formats are being experimented in search for the most “surround” and “immersive” reproduction environment. This thesis purposes a mix methodology through creating of an adaptive matrix in ones DAW (Digital Audio Workstation) that allows an engineer to use stereo, 5.1 surround sound, and multi track session material and mix them to B+ Format Audio. The mix is accomplished using the current technology with the robust DAW’s offered engineers. In addition, the mix is achieved by avoiding any difficult encoding/decoding schemes or “mysterious” black-box plug-ins, giving the mix engineer and producer inclusive control. The thesis provides the mix engineer with steps to create a mix format without additional equipment. The thesis suggests that mix engineers re-think an approach as they mix content for surround sound and immersive environments.
I. Introduction and Rationale

1. Introduction

Mix engineers have many different formats that they are required to mix. It is important to keep the requirements in mind, when mixing for a particular format or suggesting a particular format. There are two main formats that are mainly requested. One is stereo, and the other is 5.1 Surround Sound. Engineers accommodate formats by recording an ensemble in stereo, and/or 5.1, or in a similar surround format, using plug-ins to “upmix” and “downmix” the program material. Others may use delay, reverb, routing, and matrixing to provide the program material in the desired format.

Content can be prepared for a CD, DVD, theatrical release, web page, video game, sound installation, or any other possible myriad or experience. Various content release formats and distribution formats places added pressure on producers and mix engineers. The content must be played back “properly” on the required format for release or distribution. To convolute the issue, the debate continues on how to best “surround” or “immersive” an environment, and the debate is always in fluctuating. A mix prepared and mixed for 5.1 Surround Sound playback, needs to be consistent not only in 5.1 but 7.1. Considering sound art and other surround technology, the options become broader. Add shrinking budgets for engineers; it becomes increasingly difficult to accommodate all the different formats.

2. Rationale

The motivation for this research results from the Literature Review (Appendix E), Author Experience (Appendix D), interest in the surround sound mixing environments, and added inter-
est in potentially “new” and “exciting” surround formats or possibilities extended to the mixing environments. Manufacturers of professional audio equipment offer features to aid in the process of mixing for 5.1 Surround Sound. Surround Panner joysticks, upmix and downmix plug-ins are being utilized. The FCC mandated a switch to digital television broadcasting. This mandate increases the need for multi (surround) channel audio streams.

Surround theater systems are installed in homes and in many listening areas. The desire and need for content is growing. Even though according to some accounts, the consumer market has been slow to adopt surround sound formats, there is increasing interest in surround including MP-3 and computer games as well as in concert venues and home theaters. The interest in surround sound format continues to grow. One of the debates is “real” or “immersive” is the current surround sound technology. Competing technologies exist, and companies producing the content are in favor of the surround technologies. Due to constraints beyond an engineer’s control in areas of budgets, finance, distribution, tiered pricing, and tiered listening, the technology is not able to be explored or exploited.

When certain technologies are employed by the engineer, one of two scenarios occurs. First, the devices, encoding schemes, and magic boxes or other devices are extremely “clunky” and not intuitive. This causes the engineer to be segregated by the technology or device because it’s unrelated and/or unfamiliar. The engineer is unfamiliar with a rule, scheme, or process that the designers found intuitive. Secondly, the technology appears bright and shiny such as with a new “plug-in.” The engineers are tempted to want to investigate the latest technology. Though this is appealing and beneficial, once the engineers examine and use the new device or plug-in, it
begins to feel that the encoding/decoding is achieving sophisticated results, not being able to know exactly what is occurring in the plug-in, the engineer becomes apprehensive. Unfamiliar with what is causing problems, the engineers are unable to accurately represent their mix and opts out of the new devices or plug-in. At times, it is recommended to stay with what is known and work with the plug-in and devices that the engineers are familiar. Based on constraints, the engineers re-think the process. Out of frustration, the voice of the engineer expresses his or her disappointment, “I do not have the time or the resources to explore and re-think the process.” The operative word in this statement is time. The engineers not only compete with the expenses and resources for adding new technology and equipment but compete with the “learning curve production possibility constraints” in both time and resources. There is a need for increased productivity without adding costs in new equipment while effectively utilizing scarce resources to manage overhead, reduce costs with increased productivity, and generating satisfactory results and performance.

The motivations for this research is to explore a “surround/immersive” environment technology that is underutilized in B+ Format Audio, and to present the research in a useful and familiar environment to the mix engineer while adding no additional equipment. Utilizing the equipment that the engineer is accustomed can be exploited in new ways to achieve amazing results by means of the Adaptive Matrix accomplished in the mix environment. In turn, the engineer may suggest, “I have both time and resources to explore and re-think the process.”

To demonstrate this process, different audio formats employ the Adaptive Matrix System. First, stereo wav files are processed through this Adaptive Matrix to be played back in B+ For-
mat followed by 5.1 Surround Sound (six discrete channels) processed through its own Adaptive Matrix System. Secondly, a multi track recording session demonstrating the possibilities of mixing a session especially for B+ Audio Format is employed.

3. History

Ambisonics commonly referred to as B Format is not a new idea or is it new technology. Peter Fellgett, Duane Cooper, and Michael Gerzon, were “independently” working on this technology during the early 1970’s (Benjamin, Chen, Native B-Format Microphone 2). This technology resonates with hi-fi and audiophile enthusiasts, but never became “widely” used. Engineers liked the immersive qualities it provided while other engineers felt there wasn’t enough extreme localization. Other products were competing and were developed such as Dolby. B Format Ambisonics also referred to as B Format did not become an industry standard. It did, however, continue to be discussed. B Format is employed in commercial instances and audiophiles are again curious as to the ‘intrigue’ of Ambisonics.

B Format Ambisonics is comprised in two possible microphone formations. Use of a Soundfield Microphone comprised of a tetrahedral array built from cardioid microphones. The other option is to use two or three figure-of-eight microphone in conjunction with an omnidirectional microphone (Benjamin, Chen, Native B-Format Microphone 2). The tetrahedral Soundfield array’s capsules are positioned in four directions; LFU-Left Front Up, RFD-Right Front Down, LBD-Left Back Down, RBU-Right Back Up. Eric Benjamin and Thomas Chen call the four directions, “Native B Format Microphone Array” using an omnidirectional microphone on the W plane, figure-of-eight microphone on the X plane, and a figure-of-eight microphone on the
Y plane (Benjamin, Chen, Native B-Format Microphone 2-3). The emphasis of the Native B Format Microphone Array is that the microphones needs be as coincident as possible. An example is provided in Figure 1 and Figure 2 below. In this thesis, the term “Native B-Format Microphone Array” as Thomas Chen and Eric Benjamin purpose is considered standard terminology when referencing this specific microphone array.

Suggesting an additional figure-of-eight microphone with the Native B-Format microphone array that will be used for height on the Z axis (plane), gives a heightened elevated axis.
(up/down) sensation. However, the more additional microphones that are added for increased sensation and manipulation, the more difficult to keep the microphone array coincident.

Research in Ambisonics and surround sound technologies involves the pursuit of research to gain an understanding of how the ears hear and interpret sounds. Then the natural conclusion is for the engineer to imagine how to create sound in a loudspeaker environment to better simulate those sounds that are actually heard and occurring in the real world environment or to manipulate the sound to achieve certain results in a particular environment under various conditions. Michael Dickreiter gives a clear definition on how the ears hear sounds, and its correlating axis. In reference to the horizontal plane, it is the time difference of the sound reaching each ear that gives us the direction from which the sound occurred. In the vertical plane, it is the angle at which the sound occurred which is perceived as sound coloration due to its angle. Loudness, timbre, and diffused sound with direct sound ratio play an integral part in this whole process (Dickreiter 78). When sound moves up the vertical plane, two of these key cue properties remain unchanged to a certain degree. What does change is the timbre due to the diffracting effect caused by the head and ear lobes (Dickreiter 78-81). The intrigue is considering different axis in sound reproduction to “better” represent the listening environment.

Utilizing creative license, playback engineers present the listener with what they heard, or what the engineer wants the listener to hear. Based on Streicher & Everest’s “Two by Two Matrix,” often the scenarios are either “You are there” or “They are here.” Inside of these two qualifiers, there is a choice between “Re-Creative” or “Creative” (Streicher & Everest 4.4).
In evaluating current standard surround sound technologies, **Table 1: Advantages of B+ Format Audio Compared To Limitations Of 5.1 Surround Sound** is presented based on several references as well as the experience of the thesis author (Appendix G). A comparison is made between current 5.1 surround sound technologies with the advantages of B Format Ambisonics and where applicable with B+ Format. Table 1 demonstrates various drawbacks to current surround sound technology and how to overcome drawbacks using B Format Ambisonics and B+ Format. The mix engineer can develop program content for the playback system without the need for additional equipment. Table 1 is not intended as a standalone list of all limitations and advantages of 5.1 and B Format Ambisonics, which is beyond the scope of this thesis.

**Table 1: Advantages of B+ Format Audio Compared To Limitations Of 5.1 Surround Sound**

<table>
<thead>
<tr>
<th>5.1 SURROUND SOUND</th>
<th>B-FORMAT (AMBISONICS)</th>
<th>B+ FORMAT AUDIO</th>
</tr>
</thead>
<tbody>
<tr>
<td>surround sound channels are often not as full range as front channels-</td>
<td>sounds panning through can loose some quality</td>
<td></td>
</tr>
<tr>
<td>left &amp; right surround channels are made of long arrays of speakers-</td>
<td>sounds don’t localize to a point as is observed in front channels</td>
<td></td>
</tr>
<tr>
<td>“fine line between encompassing and distracting” the listener</td>
<td></td>
<td></td>
</tr>
<tr>
<td>needs 6 channels</td>
<td>only requires four channels</td>
<td></td>
</tr>
<tr>
<td>no height information</td>
<td>includes height information</td>
<td>includes height information with additional stereo channels for increased localization and direction</td>
</tr>
<tr>
<td>at least two mixes required</td>
<td>only one mix required</td>
<td>only one mix required</td>
</tr>
<tr>
<td>higher data density: lossy compression required</td>
<td>lower data density: lossless compression can be used</td>
<td>lower data density: lossless compression can be used</td>
</tr>
</tbody>
</table>
### 5.1 SURROUND SOUND

<table>
<thead>
<tr>
<th>Imaging only accurate across front stage</th>
<th>Imaging equally accurate in all directions</th>
<th>Added stereo overlay can help improve image or allow for unique manipulations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poor inter-speaker imaging (dependent upon hardware and routing)</td>
<td>Imaging between speakers</td>
<td>Imaging between speakers</td>
</tr>
<tr>
<td>Speakers must be placed in special positions</td>
<td>Speakers can be placed almost anywhere</td>
<td>Speakers can be placed almost anywhere—but with some attention to the stereo overlay and the B format speakers</td>
</tr>
<tr>
<td>“Sweet spot” only covers a relatively small area within the speakers</td>
<td>“Sweet spot” large enough to enjoy image outside of the array</td>
<td>“Sweet spot” large enough to enjoy image outside of the array</td>
</tr>
</tbody>
</table>

(Holman, 198)

(Elen, Ambisonics: The Surround Alternative, 3)
II. B+ Format Demystified

1. B Format Ambisonics

B Format is also termed B Format Ambisonics and is not to be confused with B+ Format audio. B Format Ambisonics was developed in the early 1970’s. In the early 1970’s, Michael Gerzon, and others independently of Gerzon, developed a method to capture all sounds inside a sound-field, specifically lateral information and height information. Gerzon’s argument was that “only in this way can proper spatial imaging be recreated in the listening room” (Streicher, R., & Everest, A The New Stereo Soundbook 13.12).

Two methods exist in recording material for B Format Ambisonics. The first uses a special Ambisonic microphone, such as the SoundField microphone. This microphone has specially located subcardioid capsules positioned on a tetrahedral array. The second method as defined by Streicher and Everest, explain the basic components as the measure of absolute pressure (omnidirectional) and three pressure-gradient (bidirectional) vectors (Streicher, R., & Everest, A The New Stereo Soundbook 13.12). This is found in a single microphone such as the SoundField microphone. Figure 3: Four Channels of Ambisonic Signal, exhibits signals W,X,Y,Z. Figure 4: 3D View of B-Format offers a three-dimensional view of Ambisonics.

The second method used to capture material in Ambisonics is by focusing or directing several (multi-microphone technique) different microphones to create the array. Research is inconclusive as to whether the use of the SoundField microphone or an array of microphones will produce better results.
Following the guidelines set forth by Gerzon, creating the multi-microphone array for Ambisonic considers bidirectional microphones for each of the axis X-Forward/Back, Y-Left/Right, Z-Up/Down and an omnidirectional microphone for W (Figure 3 & Figure 4). The difficulty of course in this technique is the requirement for the microphone capsules to be as coincident as possible.

When recording in Ambisonics with multi-microphones, the user handles all the encoding and decoding either on the mixer or in software. The different sums and differences in the signals gives the microphone array its directionality. This is true for microphones of uniform directivity with frequency at low and middle frequencies (Chen, Native B-Format Microphone 5). Table 2: Numerical Method demonstrates the proper sum and difference needed to decode the proper axis.

**Table 2: Numerical Method**

<table>
<thead>
<tr>
<th></th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>LFU + RFD + LBD + RBU</td>
</tr>
<tr>
<td></td>
<td>*sum of all capsules</td>
</tr>
<tr>
<td>X</td>
<td>LFU + RFD - LBD - RBU</td>
</tr>
</tbody>
</table>
The process that Gerzon developed was termed: “Ambisonics” based on a higher order of mathematics. Regardless, of whether using a SoundField microphone or an array to make up B-format, four consistent stages exist. 1) Transduction (all four sound components of SoundField microphone or B-Format array); 2) Encoding (encoding into current medium); 3) Decoding (decoding to occur upon playback from medium); 4) Reproduction (reproducing the sound field by means of properly placed speakers) (Streicher, R., & Everest, A The New Stereo Soundbook 13.13). For proper representation of the sound field, it is crucial that the loudspeakers for playback are positioned properly.

Playback for Ambisonic systems is robust. One common argument is that the more speakers you have for playback, the better the representation of the sound field. The sound field is developed by using a minimum of four speakers, the commonly used six or eight speaker hexagonal formation, or additional loudspeakers may be used in ring formations above and below the listener.
A variety of playback scenarios exist for B-Format Ambisonic. A few key items remain the same. The listening position is always located in the middle of the playback space. As few as four speakers can be used. Ambisonics records height information, in its simplest incarnation, an Ambisonic playback system places a single speaker overhead, and three others equally spaced around the listener to form a tetrahedron (Streicher, R., & Everest, A The New Stereo Soundbook 13.15).

There are other popular methods for full sound field reproduction in Ambisonics. These playback environments are usually categorized as follows; Cuboid, Octahedral, and Birectangular. Figure 5 Cuboid 6 Octahedral, 7 Birectangular, illustrate these methods, respectively.
Figure 6 Cuboid

Figure 7 Birectangular

Figure 5 Cuboid, 6 Octahedral, 7 Birectangular (Furness, R., K. Ambisonics-an overview 185)
Playback systems configured as Figure 5 Cuboid, 6 Octahedral, or 7 Birectangular, or another incarnation such as rings of speakers made up of six, or eight and possibly more speakers are positioned equally in a ring above the listener and on the floor below the listener. Research is active in this field defining the most immersive and representative method for playback in Ambisonics. Becoming familiar with B-Format Ambisonics directs the user toward UHJ (Universal HJ) encoding scheme. This hierarchical encoder is robust and the standard when using Ambisonics in the studio. UHJ encoding carries the same information as the B-Format signal, four channels and on the decoding side is adaptive.

UHJ encodes the incoming sound field into four signals, the same that are carried by the B-Format signal. This has significant advantages for broadcast, cable, and television. Through a process of multiplexing the audio signals on the decoding of UHJ, the signal can be decoded into a variety of formats with inclusive compatibility. An example of these possibilities are mono, stereo, 2-Channel UHJ, 2/12-Channel UHJ, 3-Channel UHJ, 4-Channel Speaker-Emphasis UHJ, and 4-Channel Full-Sphere UHJ (Furness, R., K. Ambisonics—an overview 184-186). The intention is achieve extensive compatibility for various playback systems. The BBC is one organization that records and produces numerous types of concerts, radio dramas and other events exploring this technology.

With current Ambisonic technology, the ability exists to replay Ambisonic material in current 5.1 surround sound environment and encode the Ambisonic program material on a DVD. The main limitation to UHJ technology is that you need a separate decoder. G-Format (G-Format Hierarchy) is a DVD-Audio specification flag that would allow future DVD players to
play this G-Format on the players (Elen, R. "G+2"A compatible, single-mix DVD format for ambisonic distribution 1).

An extension of the “G+2” technology allows for the decoding of the Ambisonic signal for playback on a 5.1 surround sound system while also creating a 2-channel UHJ signal from the Ambisonic signal (Elen, R. "G+2"A compatible, single-mix DVD format for ambisonic distribution 1). There is an argument that program material recorded in B-Format Ambisonics and played back through the G+2 system for playback in 5.1 surround sound has more “pleasing” results and immersive qualities. In addition, it is argued that it is preferred over the same material recorded in 5.1 surround sound and played back on the same system. This debate and argument along with the presentation of the research is beyond the extent of information presented in this thesis. However, the debates and arguments along with the research and technology are important in serving the B-Format Ambisonics environment.

2. B+ Format Explained

B+ Format audio has not achieved the standardization of B Format Ambisonics. Rather it is a cousin to B Format technology with advancements. B+ Format is thought of as an extension to B Format Ambisonics. Based on Ambisonics as its principle, B+ Format offers advantages over Ambisonics and current surround sound technologies.

B+ Format is based on the Ambisonic specification, meaning that it works with the standard four channel material that is referenced as W, X, Y, and Z. B+ Format audio is B+ because of the addition of a stereo (Left/Right) overlay in addition to the already defined four channels (W, X, Y, Z) (Chen, Working in B+ Format (2007). B+ Format renders two additional channels...
(the stereo overlay) to the Ambisonics four channels, totaling six audio signal channels. The additional stereo overlay separates the direct and diffused sound. Typically, the stereo Left/Right would demonstrate the direct sound, whereas the other remaining four channels would demonstrate the diffused/ambient sound of the sound-field upon playback.

Thomas Chen, who purposed working in B+ Format, designed a philosophy exhibiting the working of B+. This design is illustrated below (Chen, Working in B+ Format 1):

→ Ambient capture an entire acoustic event in a convincing fashion that typical stereo does not.
→ Accurate sounds captured from their direct locations.
→ Equality system will suit various types of music, not a select few.

(Chen, Working in B+ format 1)

Proponents for B+ Format audio suggest that B Format Ambisonics does not offer proper image localization. With the added benefit of the stereo over playback system in B+ Format, this localization issue is recovered due to the no height information that is produced in the stereo Left and Right channels and used for directional cues. The rest of the system captures the height information, giving the listener the localization and focus through the added stereo overlay and the immersive quality of the three dimensional space.

It is important to remember, B+ Format direction sounds drive the directional, timbral, and instrument cues. The direct sound arrives at the listener earlier than the ambient channel information (Chen, Working in B+ Format 2). This is implemented by inserting time delays for the
ambient channels inside of a multi-track session. A time delay system on regular playback envi-
ronment avoids ambient information reaching the listener before the stereo direct sound.

3. Advantages and Disadvantages of B+ Format Audio

B+ Format audio offers unique advantages. The biggest advantage with B+ Format audio
is that it can be decoded into many channels. The content is recorded onto six channels. The
number of decoding channels and the size of decoding representation is dynamic and available to
the end user.

Consider current optical formats, DVD, DVD-A, SACD, all of which allow for six chan-
nels of audio compression and/or storage. The speaker arrangement doesn’t dictate how the
audio will be played back. The decoder is the determining factor on the listening environment,
and this decoder can be customized to the environment and taste (Chen, Working in B+ format 2-
3). B+ is not as restrictive as current surround sound technology in defining a relatively small
“sweet-spot” for optimum listening. B+ Format, as well as, Ambisonics is renowned for the en-
larged optimum listening position.

B+ Format is considered to be “compatible” with current delivery methods and technol-
go. The data can be stored on current optical media. Users can augment their home environ-
ment to accommodate listening with addition speakers. B+ Format does not require new devices,
black boxes, and encoders. It is available to the masses currently. With “slight” adoption need
from technology companies and manufacturers, implementation can be accomplished with the
greatest of ease.
4. Uses for B+ Format Audio

Uses for B+ Format audio branch into media consumption, distribution and display. The first groups to gravitate toward this format are sound designers and sound installation artists. B+ gives them greater control and helps remove the guess-work from their productions. Commercial DVD releases benefit from including a stereo, 5.1 surround-sound, and B+ Format audio. This is implemented by recording the soundtrack in B+ Format and extrapolating the 5.1 and stereo mix from those B+ Format channels. If desired, B+ Format can supply an entirely different mix. Theater and cinema are not the only entities that can take advantage of B+ Format. Music artists who desire working and mixing in 5.1 surround sound can take advantage of the B+ Format. A whole new sonic palate is available to them. This palate is available to the mix engineer and is the focus of the remaining portion of this thesis. Illustrating how the B+ Format audio can be implemented into a production with the current technology and mixing utilities is the focal point. Engineers have at their disposal the necessary equipment. Aside from additional speakers, no extra equipment is needed to mix, audition, and prepare media for this medium.

The remaining emphasis in this thesis regards the technical specification for playback demonstration based on Michael Gerzon’s article “Practical Periphony: The reproduction of full-sphere sound.” Defining Periphony as “recording and reproduction of sound via loudspeakers from a full sphere of directions, including both all horizontal directions and all elevated and depressed directions also…” (Gerzon, M. A. Practical periphony: The reproduction of full-sphere sound 1). The result is a full-sphere twelve speakers at the face-centers, or twenty speakers at the...
vertices. Gerzon refers to this as, “impractical periphony.” Figure 8 Impractical Periphony presents an example:

This is not an easy implementation in a home listening environment or in professional recording studios. The impracticality of the reproduction environment creates a speaker system that is difficult to utilize. Used in this research is the design of an “Ambisonic Cube” playback system. This is an arrangement of eight speakers in a cube. The speakers positioned with four speakers above the listener in a cube formation, located at forty-five degree angles facing down toward the inner-center of the cube, and another four speakers on the bottom of the cube facing in the opposite direction, at forty-five degree angles pointing to the center and top of the cube. Additionally, the stereo overlay is at the correct monitor positioning for the mix engineer. Not moving the stereo monitor speakers but placing the Ambisonic Cube around that position. An example of this environment is shown in Figure 8a: “Practical Periphony”.

Figure 8: “Impractical Periphony”

Figure 8a: “Practical Periphony”
III. Creation of Adaptive Matrix

1. Explanation and Method for Adaptive Matrix Creation

The B+ Adaptive Matrix Systems is based on the B+ Format specification, a stereo overlay with a B Format Ambisonics playback system. The B Format Ambisonics playback environment being considered and under test is part of the B+ Format Adaptive Matrix and is based on an Ambisonic Cube. The four speakers positioned in a cube formation above the listener, pointing down at forty five degrees angles toward the center of the cube, and four speakers below the listener pointed up at forty five degree angles toward the inner of the cube.

The use of the Ambisonic Cube with the additional stereo overlay creating B+ Format is a result of Michael Gerzon’s article, “Practical Periphony: The Reproduction Of Full Sphere Sound.” In addition to the article, the thesis example illustrates and provides the matrix system to be employed in its “simplest,” “rudimentary,” and “practical” forms. Similarly, this footprint is practical for the home listening environment and for sound installation in a given space. Following the practical implementation is an appropriate solution for the tests and for the creation of the matrix system.

The B Format Ambisonics environment can utilize additional speakers. The theoretical advantage to B Format Ambisonics is that the more speakers that are used for playback, the more immersive and “real” the sound. Additional speakers improve the localization and listening experience. The flexibility to increase or decrease the number of speakers provides a listening environment with scalable abilities that function minimally or as a complex design.
The principles explained can be applied to most any Digital Audio Workstation (DAW). Digidesign Pro Tools LE 8, HD 7.3, and Apple Logic 8 are used and will illustrate images. The methodology can be applied to the digital and analog mixing environments.

It is recommended that a user record test material using an ambisonic microphone, or a multi-microphone array to record in B Format. This gives the engineers the experience of building the decoder inside of their DAW, and familiarizes the engineers with the encoding-decoding process before they begin creating the Adaptive Matrix for their mix program material. This is not required and is not the necessary foundation for this thesis. However, the microphone array can “easily” be constructed by referring to Figure 3. Using an omnidirectional microphone for the W position, a bi-directional microphone facing left-right for the Y position, a bi-directional microphone facing forward-backward for the X position, and another bi-directional microphone facing up-down for the Z position, the microphone array is constructed. Special care is in use to formulate this multi-microphone array as coincident as possible. The multi-microphone array bestows possibilities for unique microphone patterns and for exploration of further manipulations.

The Adaptive Matrix principles are analyzed consistently throughout the different Matrix systems. The theory of “sum” and “difference” principles are applied in the analog domain and the digital domain. It is possible in the analog domain, and through the use of high track counts and computer processing, convenient and productively efficient in the digital domain.

Table 3 displays the “Sum” and “Difference” technique used for playback in the Ambisonic Cube. This table does not consider the B+ stereo-overlay. To the right of each defined axis
either a “+” or “−” exists. The “+” or “−” denote each axis either combining or subtracting from each other. Since each axis will enter into a summing bus before it is outputted to its desired speaker, all signals are “summed” in a bus or auxiliary channel. To accommodating for the difference (-) signal, the signal will have a polarity inverse applied in each instance of a difference signal.

Table 3: Signal(s) Sum & Difference Formulas

<table>
<thead>
<tr>
<th>INCOMING SIGNALS</th>
<th>DEFINED SPEAKER POSITION</th>
<th>ABBREVIATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>W+ X+ Y+ Z+</td>
<td>Left Front Upper</td>
<td>LFU</td>
</tr>
<tr>
<td>W+ X+ Y- Z+</td>
<td>Right Front Upper</td>
<td>RFU</td>
</tr>
<tr>
<td>W+ X+ Y+ Z-</td>
<td>Left Front Lower</td>
<td>LFL</td>
</tr>
<tr>
<td>W+ X+ Y- Z-</td>
<td>Right Front Lower</td>
<td>RFL</td>
</tr>
<tr>
<td>W+ X- Y+ Z+</td>
<td>Left Rear Upper</td>
<td>LRU</td>
</tr>
<tr>
<td>W+ X- Y- Z+</td>
<td>Right Rear Upper</td>
<td>RRU</td>
</tr>
<tr>
<td>W+ X- Y+ Z-</td>
<td>Left Rear Lower</td>
<td>LRL</td>
</tr>
<tr>
<td>W+ X- Y- Z-</td>
<td>Right Rear Lower</td>
<td>RRL</td>
</tr>
</tbody>
</table>

(York University, 1-2)
A total of ten independent output options are needed. Eight of the outputs are used for the Ambisonic Cube and two additional outputs are used by the stereo overlay. Use a single interface that offers ten channels of independent output, multiple interfaces either “aggregated” together that are available in Mac OSX, or use a single interface sending all ten channels out in different formats. An interface sending eight channels of output via eight analog outputs and two channels via S/PDIF is an example. Another digital/analog (D/A) converter may be required for use with the desired speakers. Interface and manufacture documentation is verified to determine if the desired devices may be used in the manner that is being formulated.

2. Stereo Adaptive Matrix to B+ Format

The first step in creating an Adaptive Matrix is to define the ten outputs and the designated speaker inside of the DAW. The output to speaker configuration will be a 1:1 connection.

**Figure 9: Master**

**Fader** illustrates the ten master faders that will be used to drive all ten speakers in the B+ Format
speaker array. Each master fader is sending a signal to the defined output channel. Two channels are used for the stereo overlay (Left/Right) and the eight remaining master channels are used for the Ambisonic Cube. For simplicity, stereo master faders could be used for each speaker pair; Stereo, Front Upper, Front Lower, Rear Upper, Rear Lower, resulting in five master faders. Both approaches are effective. The demonstration illustrates the amount of control a mix engineer has over the matrix.

A series of auxiliary channels, feed the master faders. These auxiliary channels function as the summing channels for the B+ Format Matrix.

A total of ten auxiliary sends are needed, though not used for every auxiliary channel. The output for each auxiliary channel is not needed in this setup. The auxiliary sends control the amount of signal level sent to the master

Figure 10a: Decoder
fader channels. These are manipulated independently to taste. **Figure 10a: Decoder** illustrates the routing used for the W, X, Y, and Z planes.

The decoder is the “life blood” or “heart” of the adaptive matrix. First, notice that no outputs are defined because all the auxiliary (aux) channels are functioning as a summing bus and utilizing the auxiliary sends to route to the master faders. In recalling the signal summing and difference that is to occur based on **Table 3**, all channels but W, with W being omni-directional does not facilitate needing a polarity inverted instance of itself, have a polarity inverted instance of themselves. To accommodate the polarity inverse, an EQ plug-in is inserted on that auxiliary (aux) channel. This is illustrated in **Figure 11: Polarity Inverse**. Also notice the “Short Delay” inserted on each channel. This is simply to function as a device to time-align the Ambisonic Cube to the stereo overlay.

Thomas Chen suggests for B+ Format audio that the stereo overlay will function more as the dry/direct signal, and the B Format playback system for the ambience, whether real or artificial (Chen, Working in B+ Format 2). The functionality of direct sound needs to reach the lis-
tener before the ambient signal. A series of delays are used to time-align the Ambisonic Cube to the stereo overlay. Bumping the delay times of the Ambisonic Cube a few milliseconds allows for the stereo overlay (direct sound) to reach the listener before the Ambisonic Cube. This renders a more stable image in the front consisting of the dry/direct content. For an example, time-align the stereo overlay to the Ambisonic Cube by inserting a delay on each of the B Format aux channels of 5ms. This allows for the stereo overlay consisting of the dry/direct sound to reach the listening position first. Then bump the 5ms to around 8 to 9ms depending on the listening environment, size of Ambisonic Cube, and the program requirements.

A stereo track is created and a stereo audio file is placed on the correlating stereo track. To gain the use of the B Format cube and the stereo overlay, the use of Mid-Side (M/S) “matrixing” is employed. This is incorporated inside of the digital domain-negating the need for any additional equipment or outboard routing.

Mid-Side, or commonly called “M/S” is an intensity stereo microphone technique that involves two coincident microphones. These microphones are positioned on top of each other. The microphone supplying the M (Mid) signal is positioned facing the sound source. The polar pattern of the Mid (M) microphone can be any available pattern (cardioid, omni, Figure 8) whereas the Side (S) microphone must be (Figure 8) and rotated ninety degrees from the Mid microphone (Dickreiter 90).

Table 4: M/S and L/R Relationships illustrates the description of M/S and L/R and their mathematical relationships (Dickreiter 84).
M/S technique is utilized in recording and in mix and mastering technique. M/S allows the Mid (M) signal to be separated from the Side (S) signal.

The Mid (M) signal is sent to the front stereo field-acting as a direct/dry signal, and the Side (S) signal is sent to the Ambisonic Cube.

First the stereo signal is sent through a M/S decoder. **Figure 12: Stereo Mid Signal** illustrates the summing of the Left and Right channels into a single mono signal channel. Busses are used to separate the Left from the Right channel, with each of their outputs sending a signal to a new bus, which acts as the summing bus for the Mid (M) signal. In this example, route the output of the Stereo Mid (St/Mid) to the Left and Right stereo overlay speakers. It was a preference to

<table>
<thead>
<tr>
<th>SIGNAL</th>
<th>DESCRIPTION</th>
<th>MATHEMATICAL RELATIONSHIPS</th>
</tr>
</thead>
</table>
| L, R   | L(left) and R(right) are used independently | L = (M + S) * \( \sqrt{2} \)  
|        |             | R = (M - S) * \( \sqrt{2} \) |
| M, S   | M(mid, mono, sum,) and S(side, difference) are the microphone signals from the MS microphone technique. M & S signals can be converted to L R signals. | M = (L + R) * \( \sqrt{2} \)  
|        |             | S = (L - R) * \( \sqrt{2} \) |
place the mid signal directly in front of the listening space to act as the direct sound. The auxiliary (aux) sends allow the mid signal to be sent to one, two, or to several of the axes in the Ambisonic Cube. The X and Y plane are receiving the desired amount of Mid signal.

Working with the Side signal, key concepts to remember are the mathematical relationships. Mid = L + R and Side = L - R. An EQ plug-in is inserted on the R2 (right 2) track. Another summing bus is used to sum the Side signals and route it to two different busses for further manipulation. Figure 13: Stereo Side Signal details the approach. The Side signal is split into two different busses for sending to the Ambisonic Cube via auxiliary sends. Send options are available for all axes in the Ambisonic Cube. The sends are sent in the desired mix amount. In this example, the output of the St S+ and St S- are utilizing the output as well feeding to the W bus, and the master fader for
analog output 7-8. This example is used to facilitate the experimentation adding support to the image where it was lacking.

By following the fore mentioned process, a stereo mix is transformed into a B+ Format compatible program material. Any sound that is better suited as more “diffused” can be sent to the W bus which feeds all channels in the Ambisonic Cube. This is noteworthy, giving the sensation of more diffused sound as opposed to direct sound. B+ Format offers new excitement and new experiences to the most familiar mix. A word of caution, the W bus is feeding equal amount of signal to all speakers in the Ambisonic Cube. This could result in uneven level balances between the Ambisonic Cube busses. It is suggested based on the W bus being the element that adds the diffused sound, send program content to the W bus at levels -3dB (3dB down) compared the send levels of other busses to the Ambisonic Cube busses.

3. 5.1 Surround Sound Adaptive Matrix to B+ Format

The main advantage to B+ Format Audio is the ability to mix content for this format without the need for additional equipment. As long as an engineer has the ability to output ten channels simultaneously and ten speakers, that all that is needed for mixing to B+ Format. The other important advantage to B+ Format over other current surround formats such as 5.1 Surround Sound concerns production and mix capabilities. As more mixing and production is completed on tighter budgets in various studio configurations, the tools the engineer may be using could be limited in functionality due to “Lite” (stripped down) versions of software or other hardware limitations. The engineer needs to purchase expensive “expansion” capabilities to the
desired software and/or upgrade hardware, or needs may dictate additional out-board equipment
with a steep learning curve in time and resources.

Content mixed and produced for 5.1 Surround Sound can “run through” the Adaptive Ma-
trix System allowing it to be reproduced in the B+ Format playback environment.

The same decoding scheme is used as was il-
Illustrated in Figure 10a: Decoder which is feeding the master fader (output)
control for all ten speakers as explained in Figure 9:
Master Fader. Figure 14: Discrete 5.1 Surround
Content illustrates how a simplified or scaled down
version of software is cap-
able of providing discrete 5.1 Surround Sound channel program material into B+ Format Audio playback.
The six discrete channels, L (Left), C (Center), R (Right), LFE (LFE), LS (Left-Surround), RS (Right-Surround), are inserted into their own discrete channel.

By using the auxiliary sends, the 5.1 Surround Sound mix can be sent to the Adaptive Matrix system/Decoder. **Figure 15: 5.1 Matrix Routing** provides an example and the basic methodology to mixing 5.1 content to B+ Format.

The approach is similar to the stereo adaptive matrix with additional extensions. In considering the 5.1 Surround Sound mix, assume that what is mixed Left, Center, and Right are specific in their nature to the placement inside of the surround sound - sound-field. The matrix is changed on these three channels and utilizes the auxiliary sends and the channel output. The Left, Center, and Right channels were positioned in the stereo-overlay in their respective space (L, C, R) by using the channel outputs. This allows the front of the image to be dynamically adjusted through the use of the channel faders into the stereo-overlay of the B+ Matrix. Using the
auxiliary sends on these three channels controls the diffused sound allowed to enter into the Adaptive Matrix and ultimately the B+ Format playback system by adjusting the send level of the W channel.

Another process investigated the LFE channel. Since B+ Format, as the standard specifies, has no subwoofer channel, one approach is to supply a more “diffused” effect for the LFE by sending to the W bus positioning on a specified axis, or on all axes. This is program dependent and mixed to preference. Care must be exercised. LFE material is broad in dynamic range. Avoid overloading the channels the LFE is sending.

The discrete 5.1 Surround Sound channels could benefit by using M/S encoding and decoding, similar to stereo. The six discrete channels are paired up into five stereo pairs: (1) L and Center channels, (2) Center and Right channels, (3) Right and Right Surround channels, (4) Left Surround and Right Surround channels, and (5) Left Surround and Left channels. This allows for each “stereo-pair” to be processed by an M/S encoder and decoder allowing independent Mid and Side channels for each stereo-pair. Routing the M/S processed signals to the Adaptive Matrix and B+ Format playback allows for increased precision and placement of program material inside of the sound-field.

4. Multitrack Session Adaptive Matrix to B+ Format

It is “tricky” to define specific rules or procedures that must be followed. However, an outline can be generated providing concepts that consistently produce positive results. Considering the same Adaptive Matrix as outlined in Figure 10a: Decoder and Figure 9: Master Fader, interesting results occur. The use of multiple outputs through conventional IO, as well as, auxil-
iary sends into the matrix, with extensions of current modern mixing practices, such as parallel compression, create positive results.

Assuming hardware and software limitations are considered equal, take full advantage of the robust routing a software DAW provides. With discrete 5.1 Surround Sound channels, utilize the auxiliary sends as well as utilize the output routing for each channel. This allows an engineer to tailor the mix specifically to B+ Format. Sounds are located in specific speaker regions in addition to the Ambisonic Cube. This gives a desired sound source the desired directionality and localization intention, and also introduces a more diffused effect using the Adaptive Matrix to employ B+ Format playback. The result is a “more” immersive environment and representative mix.

An advantage of introduced a more involved sound-field is lessening the number of sounds competing for sonic space in one or two speakers. More advanced mixing techniques that are common place today such as Parallel Compression, is where a track is “mult’ed” and one track is highly compressed. The other track is slightly compressed or not compressed. These tracks are mixed together and positioned in the stereo field. The dynamic range is retained by using the Adaptive Matrix and B+ Format sound-field while introducing the heavily compressed version of a track in a certain region inside of the sound-field and the non-compressed version in another region in the sound-field. A more “diffused” effect is achieved by the Matrix for sounds that do not have the desired localizability.
5. Adaptive Matrix Summary

Adaptive Mix Matrix can be used in various situations including stereo, 5.1 Surround Sound, and in Multi-track sessions. Each mix scenario provides its own specific extension to the matrix system while allowing the content material to be either “conformed” or “up-mixed,” as in the case of stereo and 5.1 Surround Sound, or mixed, in the case of a Multi-track session, into B+ Format playback.

In B+ Format the stereo content presented through the matrix system is “forgiving.” The content was played back in B+ Format without any noticeable anomalies. In 5.1 Surround Sound, there is a difference. Depending on the program material, surround sound recordings are generally done either in a “documentative/re-creative” fashion, or in an attempt to sculpt the sound using the surround format. The material in B+ Format is represented without strange or adverse stress. The use of automation and sound sculpting in 5.1 caused odd image shifts and unpredictable behavior in B+ Format playback. These are resolved through the use of additional filtering and processing. The process of running this mix through the Adaptive Matrix will create surprises that are manageable.
IV. Informal Evaluation (Listening Test)

1. Listening Test Procedure

The Adaptive Matrix was tested by selecting a pool of professionals for qualitative analysis of the Matrix system. The evaluation process included a number of participants who were either considered professional mix/recording engineers with surround sound mixing experience and expertise or professional musicians. Musicians were classified as listeners and were not classified as professionals or mix engineers. Because of their extensive musical background, they were utilized in the testing process.

The listening tests provided feedback to the success or failure of music being presented in its native format as compared to the B+ Format environment. In each case, the content evaluated was not mixed or “prepared” for playback via the B+ Format Adaptive Matrix, but was demonstrated by program material that was intended for stereo and 5.1 Surround Sound. Program pieces were placed through the matrix system and the evaluating participants experienced the content perceived from its native format either in stereo or 5.1 Surround Sound and made a comparison with the demonstrated B+ Format. The participants were each evaluating the material in real-time. The participants A-B the native program material to the B+ Format playback system and moved between each of the playback formats in real time.

The evaluations were conducted from April 19, 2009 to April 26, 2009, in a studio environment. The playback environment consisted of a stereo “nearfield” monitoring setup consisting of JBL LSR 4328. A 5.1 setup that used the stereo “nearfield” monitors JBL LSR 4328 and the remaining 5.1 surround channels, Center, Left-Surround, Right Surround, were Behringer
Truth 2031A speakers. The Sub-woofer channel being produced by a Behringer B2092A. The B+ Format speaker playback system utilized the JBL LSR 4328 speakers as the stereo Left/Right overlay. The remaining eight speakers formed the Ambisonic Cube and were M-Audio Studio-ophile AV20. These were lightweight and catered to being “flown” in the air above the listener pointing on a center position in the middle of the cube pointing at a forty five degree angle. This avoided harm to the listener and avoided damage to the speakers in case of a fall.

The stereo playback speaker location and position was held constant and consistent in the playback systems. The 5.1 Surround-Sound was set up with the stereo speakers as the Left/Right channels and adhered to the ITU 5.1 specifications. The B+ Format playback system utilized the same stereo Left/Right channels as the stereo and 5.1 systems, however the Ambisonic Cube maximized the largest “ideal” listening position which caused the Ambisonic Cube to be set up on the perimeter of the 5.1 Surround-Sound.

To accommodate a maximum of ten simultaneous channels of output, two audio interfaces were used. A Digidesign Digi 003 was used as the main hardware interface which allowed for eight channels of analog output. The additional two channels were gained by taking the S/PDIF outputs on the Digi 003 and running them into the second interface, an Apogee Ensemble. The Apogee was routed where the input of the S/PDIF were outputted to the analog 1-2 outputs on the Ensemble unit. Pro Tools was outputting to the eight analog outputs, gaining two additional outputs by outputting from the S/PDIF into another interface and routing to that second interface outputs. The stereo program material was routed straight from the Digi 003’s S/PDIF into the Ensemble’s S/PDIF input which outputted to the analog output 1-2. The B+ Format Ma-
The Ambisonic Cube was created by positioning the top four speakers on microphone stands pointing to the “center” of the cube angled down at a forty five degree angle. The bottom four speakers of the cube were placed at a forty five degree angle to the “center” of the cube. The stereo overlay was then fed by using the S/PDIF (master fader(s) outputting to S/PDIF from the 003) to feed the Ensemble for the stereo Left and Right channels. The Digi 003 and the Apogee Ensemble were both clocked via word-clock to an Apogee Big Ben which served as the master clock for the two hardware devices.

Pro Tools was the application used to playback the stereo and B+ Format Adaptive Matrix content. Logic Pro 8 was the application used to playback the 5.1 Surround Sound content. The Apogee Ensemble was used as the sole interface for the 5.1 Surround Sound content. The routing for the Ensemble was Analog 1-2 were the stereo Left/Right channels. The 3-4 were the Center/LFE channels and 5-6 were the Left-Surround/Right-Surround channels. Six discrete surround tracks were placed in Logic Studio 8 and each had its discrete output channel. When the six discrete channels (5.1 program content) was evaluated using the B+ Format Adaptive Matrix, those same six discrete program channels were placed on six tracks in Pro Tools which then routed to the Adaptive Matrix for routing, processing and playback.

The use of both software applications and hardware devices allowed for testing, creating the matrix system, and for evaluating the system by means of A-B in real time.

2. System and Setup Validation

The studio environment in which the B+ Adaptive Matrix System evaluation process occurred was a studio space approximately ten feet wide and twelve feet long. The Ambisonic Cube was created by positioning the top four speakers on microphone stands pointing to the “center” of the cube angled down at a forty five degree angle. The bottom four speakers of the cube were placed at a forty five degree angle to the “center” of the cube. The stereo overlay was then fed by using the S/PDIF (master fader(s) outputting to S/PDIF from the 003) to feed the Ensemble for the stereo Left and Right channels. The Digi 003 and the Apogee Ensemble were both clocked via word-clock to an Apogee Big Ben which served as the master clock for the two hardware devices.

Pro Tools was the application used to playback the stereo and B+ Format Adaptive Matrix content. Logic Pro 8 was the application used to playback the 5.1 Surround Sound content. The Apogee Ensemble was used as the sole interface for the 5.1 Surround Sound content. The routing for the Ensemble was Analog 1-2 were the stereo Left/Right channels. The 3-4 were the Center/LFE channels and 5-6 were the Left-Surround/Right-Surround channels. Six discrete surround tracks were placed in Logic Studio 8 and each had its discrete output channel. When the six discrete channels (5.1 program content) was evaluated using the B+ Format Adaptive Matrix, those same six discrete program channels were placed on six tracks in Pro Tools which then routed to the Adaptive Matrix for routing, processing and playback.

The use of both software applications and hardware devices allowed for testing, creating the matrix system, and for evaluating the system by means of A-B in real time.

2. System and Setup Validation

The studio environment in which the B+ Adaptive Matrix System evaluation process occurred was a studio space approximately ten feet wide and twelve feet long. The Ambisonic Cube was created by positioning the top four speakers on microphone stands pointing to the “center” of the cube angled down at a forty five degree angle. The bottom four speakers of the cube were placed at a forty five degree angle to the “center” of the cube. The stereo overlay was then fed by using the S/PDIF (master fader(s) outputting to S/PDIF from the 003) to feed the Ensemble for the stereo Left and Right channels. The Digi 003 and the Apogee Ensemble were both clocked via word-clock to an Apogee Big Ben which served as the master clock for the two hardware devices.

Pro Tools was the application used to playback the stereo and B+ Format Adaptive Matrix content. Logic Pro 8 was the application used to playback the 5.1 Surround Sound content. The Apogee Ensemble was used as the sole interface for the 5.1 Surround Sound content. The routing for the Ensemble was Analog 1-2 were the stereo Left/Right channels. The 3-4 were the Center/LFE channels and 5-6 were the Left-Surround/Right-Surround channels. Six discrete surround tracks were placed in Logic Studio 8 and each had its discrete output channel. When the six discrete channels (5.1 program content) was evaluated using the B+ Format Adaptive Matrix, those same six discrete program channels were placed on six tracks in Pro Tools which then routed to the Adaptive Matrix for routing, processing and playback.

The use of both software applications and hardware devices allowed for testing, creating the matrix system, and for evaluating the system by means of A-B in real time.
cube were fastened to the same microphone stands, but at the base, approximately one foot from
the ground pointing upwards toward the center of the cube and were angled upwards at a forty
five degree angle. The top four speakers of the cube were eight feet high and ten feet apart from
each other. The surround system was positioned in the room following the ITU specifications.

All speakers were calibrated by sending a 1kHz tone to each speaker and setting it to 83
dB SPL. The calibration in all measurement tests were used by placing the SPL meter in the cen-
ter listening position.

Sine sweeps were performed on each speaker by using a software application Fuzz
Measure v.3.1.1. An AKG C414 XL-II microphone using omnidirectional polar pattern was po-
sitioned in the desired listening spot. The sine-wave sweeps that were played through each
speaker were considered “Full-Range,” meaning 20Hz-20kHz. The duration of the sine sweep
for each speaker lasted one second. The response of the speaker under test was then recorded by
the AKG C414 microphone as displayed in Appendix B: Sine Sweet Analysis. The AKG C414
microphone was connected to an Apogee Duet interface which was recording at 44.1kHz 24 bit.

Impulse responses were taken at all speaker locations (except the LFE). A balloon was
popped at each speaker location, and its impulse response was recorded through the same AKG
C414 microphone connected to an Apogee Duet interface. Each impulse response was recorded
at 44.1kHz 24 bit into Logic Pro 8. Using the Logic channel EQ analyzer, prints of each impulse
response were created to test for unexpected system problems and to understand the sonic print
of the listening environment. These impulse responses can be found in Appendix A: Speaker
Location Impulse Responses.
A series of five tones were played through each speaker. Starting with 250, 500, 1kHz, 2kHz, 4kHz were all played using a signal generator inside of the DAW and recorded in the same listening position as was calibrated by the SPL meter and the sine-wave sweeps. These samples were recorded using an Apogee Duet, recording into Logic Pro 8, recording at 44.1kHz and 24 bit. Each sample was then analyzed using RND Digital’s “Inspector Free” native plug-in to analyze the spectral content as perceived by the microphone in the listening environment. These results can be found in Appendix C: Tone Analysis.

These tests were conducted to objectively understand the environment in which the listening tests were to occur, for purposes of calibration, and to “remove” bias that entered environmentally or by equipment limitation, latency, or the like.

3. Listening Test

Pools of test subjects were selected based on either their professional mix experience or their professional/excellence in their respected musical field or instrument. Each test subject participated anonymously to help ensure the most accurate response to the listening test experience.

The listening tests were in two sections lasting approximately 15 minutes (total) for each participant. The first section allows the subject to solo/mute two tracks in Pro Tools to evaluate the stereo program material. One track is routed directly to the stereo Left/Right speakers, and the other track is routed to the adaptive matrix. The user was able to A-B the stereo content selections in real time by means of an Ethernet controller or the content is A-B’d for them while they listened, based on their preference. Each participant sat in the desired listening location and listened to the stereo program material presented as stereo material and routed using the Adaptive
Matrix System. Upon completion of the four stereo sound samples the second section of the test was started. The user then listened to a song that had been mixed for utilizing 5.1 Surround-Sound played through Logic Pro 8. The test subject then listened to the same six discrete tracks played through Pro Tools routing to the Adaptive Matrix System.

After each participant finished the listening tests, a survey was administered through Survey Monkey. This survey allowed the participants to rank their experience and perceptions based on a series of questions. The results of these questions are found in Appendix D: Survey Questions.
4. Listening Test Results

Several events occurred during the listening tests based on the four participants. All participants were able to differentiate between the test materials’ native format and the use of the B+ Format Adaptive Matrix playback system. All participants stated they would consider listening to music in B+ Format Audio if available. The basis for considering the music content in B+ Format was based on either classifying the experience as the sonic environment being captivating or the sound-field being very immersive.

When each participant was asked if they felt that B+ Format was an improvement over the native format, one said definitively “yes” while the rest were “undecided.” Each was asked what anomalies were perceived out of a collection of potential attributes. Those attributes were: image shifts, phase shifts, comb-filtering, and distortion. They detected image shifts, which could be expected, and this is not necessarily a negative attribute. However, phase shifts and comb-filtering were also selected. These could potentially be environmental or system based.

Concerning the 5.1 Surround Sound sample, the opinions were not clear and could be the result of the content material selected. The 5.1 test material was a mix done to utilize 5.1. Thus, the 5.1 delays, reverbs, and other processing were being used. It was a “representative” recording to determine perception of “creatively mixed” surround program material running through the Adaptive Matrix. Based on the content material, the results were mixed. The results were not conclusive in defining those participants evaluating the system whether they perceived this type of mix as inviting as the stereo program material. One theory as to why this may be the case is that the sample’s automated events and effects that jumped out in the matrix and disap-
peared at the same right, were suffering from comb-filtering. phase cancellations. or phase super-
impositions at the listening location. This may be more specific to environment than matrix re-
lated. Specific information regarding the listeners’ test results are presented in Appendix D:

Survey Questions.

In the case of the stereo sound samples of content, all participants stated that listening to the content through the B+ Format Adaptive Matrix described their experience as “listening to a live concert” and found it pleasing and intriguing.

Items that potentially attribute to the margin of error in the listening tests were the play-
back system speakers. Speakers of different make and from different manufactures, and those used in the Ambisonic Cube were not full-range as can be seen in the frequency sweep graph analysis in the Appendix B. Regardless of these issues all participants were engaged by the B+ Format matrix systems and expressed interest in exploring it further.
V. Conclusions

The B+ Format Adaptive Matrix proved an exciting extension and a new environment to work and listen. Using B+ Format playback systems can be utilized by sound artists and sound designers. This thesis outlines an approach for a an engineer to mix and create the program material for those installations and exhibits.

B+ Format renders a compelling sonic sound field. Engineers can mix material already produced or multi-track session directly into B+ Format without the need for additional equipment. The processes explores and suggests a reevaluation of how the engineer may mix surround sound material to create a sound-field as opposed to positioning sounds in speakers which delivers a multi-mono experience that is rather “lack-luster.” Again, one of the main advantages of the thesis is to provide a process and outline for an engineer to mix for the format in their own environment without additional equipment. This thesis allows engineers to re-think an approach as they mix content for surround or immersive environments regardless of their format and playback medium.

The “sweet-spot” of the listening position is enlarged in B+ Format. Other popular surround sound formats wrestle with such a small “sweet-spot” in relation to what is offered in Ambisonics and B+ Format Audio. This allows for the small movements of the listener to not cause adverse effects on the sound being perceived. The increased optimal listening location size allows more listeners to take part in the ideal listening experience.
If engineers consider the approach presented in this thesis with the mix experience in this playback environment, B+ Format may receive more attention. Continued tests with additional subjective and objective analysis may greatly influence the viability of the format based on the research and experience set forth in this thesis. It is the beginning of the investigation and starting the process of considering an approach to the format that has substantial and sonic worth which should be considered as a playback medium.
Selected References


Appendix A: Speaker Location Impulse Responses

Left Speaker

Right Speaker
Center Speaker

Left-Surround Speaker
Right-Surround Speaker

Left Font Upper Speaker
Right Font Upper Speaker

Left Font Lower Speaker
Right Font Lower Speaker

Left Rear Upper Speaker
Mixing Scenarios For B+ Format Audio

52
Right Rear Lower Speaker

<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>Gain</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>30.0</td>
<td>0.71</td>
<td>1.10</td>
</tr>
<tr>
<td>80.0</td>
<td>0.71</td>
<td>0.98</td>
</tr>
<tr>
<td>200</td>
<td>0.71</td>
<td>0.71</td>
</tr>
<tr>
<td>500</td>
<td>0.71</td>
<td>0.71</td>
</tr>
<tr>
<td>1200</td>
<td>0.71</td>
<td>0.71</td>
</tr>
<tr>
<td>3500</td>
<td>0.71</td>
<td>0.71</td>
</tr>
<tr>
<td>10000</td>
<td>0.71</td>
<td>0.71</td>
</tr>
<tr>
<td>17000</td>
<td>0.71</td>
<td>0.71</td>
</tr>
</tbody>
</table>

Note: The chart displays the frequency response of the speaker, with gain values in dB and slope values in dB/Oct.
Appendix B: Sine Sweep Analysis (20Hz-20kHz)

Left Speaker

Right Speaker

Center Speaker

Left Surround Speaker

Right Surround Speaker
Left Front Upper Speaker

Right Front Upper Speaker

Left Front Lower Speaker

Right Front Lower Speaker

Left Rear Upper Speaker

Right Rear Upper Speaker

Mixing Scenarios For B+ Format Audio

Miles Fulwider
Appendix C: Listening Position Selected Tone Analysis

Mixing Scenarios For B+ Format Audio  
Miles Fulwider  
57
### Mixing Scenarios For B+ Format Audio

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Right Front Upper</th>
<th>Left Front Lower</th>
<th>Right Front Lower</th>
</tr>
</thead>
<tbody>
<tr>
<td>250Hz</td>
<td>![Graph 250Hz]</td>
<td>![Graph 250Hz]</td>
<td>![Graph 250Hz]</td>
</tr>
<tr>
<td>500Hz</td>
<td>![Graph 500Hz]</td>
<td>![Graph 500Hz]</td>
<td>![Graph 500Hz]</td>
</tr>
<tr>
<td>1000Hz</td>
<td>![Graph 1000Hz]</td>
<td>![Graph 1000Hz]</td>
<td>![Graph 1000Hz]</td>
</tr>
<tr>
<td>2000Hz</td>
<td>![Graph 2000Hz]</td>
<td>![Graph 2000Hz]</td>
<td>![Graph 2000Hz]</td>
</tr>
<tr>
<td>4000Hz</td>
<td>![Graph 4000Hz]</td>
<td>![Graph 4000Hz]</td>
<td>![Graph 4000Hz]</td>
</tr>
</tbody>
</table>

Miles Fulwider
Mixing Scenarios For B+ Format Audio

Miles Fulwider

61
### Appendix D: Survey Questions

**B+ Format Adaptive Matrix Evaluation**

1. In listening to the program material were you able to differentiate between the Native format (i.e. stereo, 5.1 Surround/Sound) and the program's B+ Format matrixed version?

<table>
<thead>
<tr>
<th>Response</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>100.0%</td>
<td>4</td>
</tr>
<tr>
<td>No</td>
<td>0.0%</td>
<td>0</td>
</tr>
</tbody>
</table>

answered question 4

skipped question 0

2. Based on your listening experiences today would you consider listening to music in B+ Format?

<table>
<thead>
<tr>
<th>Response</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>100.0%</td>
<td>4</td>
</tr>
<tr>
<td>No</td>
<td>0.0%</td>
<td>0</td>
</tr>
</tbody>
</table>

answered question 4

skipped question 0

3. Based on the question previous please select the basis for your choice. If YES:

<table>
<thead>
<tr>
<th>Response</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enjoyed the music selection</td>
<td>0.0%</td>
<td>0</td>
</tr>
<tr>
<td>Sonic environment was captivating</td>
<td>50.0%</td>
<td>2</td>
</tr>
<tr>
<td>Sound-field was very immersive</td>
<td>50.0%</td>
<td>2</td>
</tr>
<tr>
<td>Just curious about B+ Format-Tell me more!</td>
<td>0.0%</td>
<td>0</td>
</tr>
</tbody>
</table>

answered question 4

skipped question 0

4. Based on the question previous please select the basis for your choice. If NO

<table>
<thead>
<tr>
<th>Response</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>I did not enjoy listening in B+ Format</td>
<td>0.0%</td>
<td>0</td>
</tr>
<tr>
<td>I see no reason logistically to listen to music in B+ Format</td>
<td>0.0%</td>
<td>0</td>
</tr>
<tr>
<td>B+ Format is distracting to me</td>
<td>0.0%</td>
<td>0</td>
</tr>
<tr>
<td>I could hear adverse effects/anomalies as result of the B+ Format Matrix system.</td>
<td>0.0%</td>
<td>0</td>
</tr>
</tbody>
</table>

answered question 0

skipped question 4

5. Did you perceive the B+ matrix converted Stereo example to be an improvement from the original source material?

<table>
<thead>
<tr>
<th>Response</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>25.0%</td>
<td>1</td>
</tr>
<tr>
<td>No</td>
<td>0.0%</td>
<td>0</td>
</tr>
<tr>
<td>Undecided</td>
<td>75.0%</td>
<td>3</td>
</tr>
</tbody>
</table>

answered question 4

skipped question 0
### 6. What anomalies occurred during the B+ matrix stereo playback example? (check all that apply)

<table>
<thead>
<tr>
<th>Anomaly</th>
<th>Response Percent</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase Shifts</td>
<td>25.0%</td>
<td>1</td>
</tr>
<tr>
<td>Comb-Filtering</td>
<td>25.0%</td>
<td>1</td>
</tr>
<tr>
<td>Image Shifts</td>
<td>50.0%</td>
<td>2</td>
</tr>
<tr>
<td>Distortion</td>
<td>0.0%</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>25.0%</td>
<td>1</td>
</tr>
<tr>
<td>No anomalies were perceived.</td>
<td>50.0%</td>
<td>2</td>
</tr>
</tbody>
</table>

**Response:**
- 4 answered question
- 0 skipped question

### 7. Were these anomalies presented by the B+ Format Matrix advantageous to the source material?

<table>
<thead>
<tr>
<th>Response</th>
<th>Response Percent</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0.0%</td>
<td>0</td>
</tr>
<tr>
<td>No</td>
<td>0.0%</td>
<td>0</td>
</tr>
<tr>
<td>Undecided</td>
<td>75.0%</td>
<td>3</td>
</tr>
</tbody>
</table>

**Response:**
- 4 answered question
- 0 skipped question

### 8. Was the B+ Format matrixed example a positive listening experience over the sources native format?

<table>
<thead>
<tr>
<th>Response</th>
<th>Response Percent</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>50.0%</td>
<td>2</td>
</tr>
<tr>
<td>No</td>
<td>0.0%</td>
<td>0</td>
</tr>
<tr>
<td>Undecided</td>
<td>50.0%</td>
<td>2</td>
</tr>
</tbody>
</table>

**Response:**
- 4 answered question
- 0 skipped question

### 9. Was the stereo B+ Format matrixed example(s) a sonic distraction and/or disorientating? (y/n) If yes: distraction? disorientation, or both?

#### Yes

<table>
<thead>
<tr>
<th>Distraction</th>
<th>Disorientation</th>
<th>Both</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0.0% (2)</td>
<td>100.0% (1)</td>
<td>0.0% (0)</td>
</tr>
</tbody>
</table>

#### No

<table>
<thead>
<tr>
<th>Not Distracting or Disorienting</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>100.0% (2)</td>
</tr>
</tbody>
</table>

**Response:**
- 4 answered question
- 0 skipped question
10. Did you perceive the B+ matrix converted 5.1 Surround-Sound example to be an improvement from the original source material?

<table>
<thead>
<tr>
<th>Response</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>50.0%</td>
<td>2</td>
</tr>
<tr>
<td>No</td>
<td>25.0%</td>
<td>1</td>
</tr>
<tr>
<td>Undecided</td>
<td>25.0%</td>
<td>1</td>
</tr>
</tbody>
</table>

answered question 4
skipped question 0

11. What anomalies occurred during the B+ matrix 5.1 Surround-Sound playback example? (check all that apply)

<table>
<thead>
<tr>
<th>Response</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase shifts</td>
<td>25.0%</td>
<td>1</td>
</tr>
<tr>
<td>Comb-filtering</td>
<td>50.0%</td>
<td>2</td>
</tr>
<tr>
<td>Image shifts</td>
<td>75.0%</td>
<td>3</td>
</tr>
<tr>
<td>Distortion</td>
<td>0.0%</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>50.0%</td>
<td>2</td>
</tr>
<tr>
<td>No anomalies were perceived</td>
<td>25.0%</td>
<td>1</td>
</tr>
</tbody>
</table>

answered question 4
skipped question 0

12. Were these anomalies presented by the B+ Format Matrix advantageous to the source material?

<table>
<thead>
<tr>
<th>Response</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>25.0%</td>
<td>1</td>
</tr>
<tr>
<td>No</td>
<td>75.0%</td>
<td>3</td>
</tr>
<tr>
<td>Undecided</td>
<td>0.0%</td>
<td>0</td>
</tr>
</tbody>
</table>

answered question 4
skipped question 0

13. Was the B+ Format 5.1 Surround-Sound matrixed example a positive listening experience over the source's native format?

<table>
<thead>
<tr>
<th>Response</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>50.0%</td>
<td>2</td>
</tr>
<tr>
<td>No</td>
<td>25.0%</td>
<td>1</td>
</tr>
<tr>
<td>Undecided</td>
<td>25.0%</td>
<td>1</td>
</tr>
</tbody>
</table>

answered question 4
skipped question 0

14. Was the 5.1 Surround-Sound B+ Format matrixed example(s) a sonic distraction and/or disorienting? (y/n) If yes: distraction? disorientation, or both?

<table>
<thead>
<tr>
<th>Response</th>
<th>Distractions</th>
<th>Disorientation</th>
<th>Both</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0.0% (0)</td>
<td>66.7% (2)</td>
<td>33.3% (1)</td>
<td>3</td>
</tr>
<tr>
<td>No</td>
<td>100.0% (1)</td>
<td>0.0% (0)</td>
<td>0.0% (0)</td>
<td>1</td>
</tr>
</tbody>
</table>

Not distracting nor disorienting

<table>
<thead>
<tr>
<th>Response</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>No</td>
<td>0</td>
</tr>
</tbody>
</table>

answered question 4
skipped question 0