

# **FOUND IN SPACE**

JOURNEY TO THE THIRD DIMENSION  
OF MUSIC SPATIALITY  
IN THE DOLBY ATMOS FORMAT

PAUL ANTHONY NOVOTNY

A DISSERTATION SUBMITTED TO  
THE FACULTY OF GRADUATE STUDIES  
IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN Music,  
YORK UNIVERSITY  
TORONTO, ONTARIO

September 2025

© Paul Anthony Novotny, 2025

## ABSTRACT

In about 1550, Adrian Willaert was the first composer to stage his ensembles with voices placed physically apart. In 2019 Dolby Atmos for music was introduced and began the delivery of mainstream three-dimensional (3D) popular music to audiences. Today, this new generation of streamed media uses embedded metadata to render a compelling 3D music sound field from spatially aware playback devices that convey convincing virtual spaces and places.

Offering songwriters, composers and music mixers a 3D spatial-sound tool-palette, Atmos is capable of replicating Willaert's staging using an Atmos multi-speaker system or with binaural playback in standard stereo headphones.

This technology provides an abundance of creative potential for 3D music composers and recordists. However, practitioners and pedagogues are not fully aware of the substantial metaphysical componentry involved for construction of a non-conflicting and compelling 360-degree Atmos sound field. As such, innumerable questions arise. This author focuses on two: 1) What is the substance of compelling spatiality in immersive 3D Dolby Atmos music recordings? 2) How can we find it?

With interdisciplinary research and scientifically reasoned approaches, *Found in Space* locates and semantically describes the substantial metaphysical componentry involved in 3D music spatiality. The story unfolds in three separate episodic journeys.

Episode 1 describes a compositional ethos for 3D music spatiality. Episode 2 introduces 3D music spatial theory with an analysis synergy and ear training exercises. Episode 3 describes the philosophy and benchmark for a precise eco-friendly, large scope/small scale, sole operator Dolby Atmos 7.1.4 Hyper Near-Field Tiny Studio.



To conclude, all metaphysical components are gathered onto a one-page diagrammatic spider sheet to reveal the holistic DNA of 3D music spatiality. United by a central principle, the journeys seek the understanding of positional relationships, cognitive perspectives, proportional properties and character attributes.

*Found in Space* uncovers salient literacy and concepts specifically intended for the precise understanding, and description of recorded music's 3D spatiality.

Intended for composers, songwriters, mixing engineers, and pedagogues who teach the art and craft of 3D composition in the Dolby Atmos format, an endless frontier of creative possibilities can be *Found in Space*.

Dolby Atmos 3D Immersive Music

3D Spatial Music Theory

3D Spatial Audio Semantics

3D Immersive Sensorium and Metadata

3D Music Information Retrieval (MIR) and Metadata

Environmentally Sustainable Dolby Atmos Music Studio Design

Dolby Atmos Hyper Near-Field Tiny Studio

## **DEDICATION**

Dedicated to the next generation of 3D music and sound explorers,

and my great friend and composer Dr. Doug Wilde.

Doug has been a musical inspiration for my research.

He tasked me to co-produce, perform on, mix, and master a recording of hexatonic and tonal compositions that bear his unique compositional chemistry.

One of his compositions, presently entitled “Tune 1,” became a workhorse for the

development, testing, perception, and practice of

3D spatiotemporal orchestration and spatial theory.

“Doug, you are one of the finest composers any musician could know.”

Thank you!

## ACKNOWLEDGMENTS

With utmost gratitude, thank you to my lovely wife Ann Heinonen Novotny for her understanding and Herculean amount of patience as I pursued this research degree. At the onset of this journey my good friend Dr. Tom Reynolds said to me, “undertaking a doctorate strongly affects your closest family,” and he was correct.

Ann and I have travelled to Sweden and Boston to attend Art of Record Production conferences, and Oslo to attend the 2024 Innovation in Music conference, and recently to Victoria BC, Canada for the 2025 Society for Music Production Research (SMPR) conference.

She has been my resolute supporter during this seven-year journey, which has enhanced our life experience immeasurably. We have made good friends in the international community of popular music researchers, and many have taught me about life, my topic and research practices.

Harry Lewis has been my confidant, profoundly sage wordsmith and associate copyeditor who has spurred me onward when things became difficult. His patience for listening, reading, and understanding my research, and then teaching me how to write it down has given me an enduring gift that I am eternally grateful for.

Dr. Rob Bowman has been the central mentor for all of my doctoral work. He has challenged me to dig deeply while gifting me with his buoyant inspiration and encouragement. For Rob, I always tried to put love into my work because it is my honour to be the recipient of his wisdom and commitment to continued learning. When Rob and my other supervisory committee members—Dr. Denis Martin, and Professor Al Henderson—visited my Dolby Atmos Hyper Near-Field Tiny Studio, they experienced the joyous immersion of a compelling 3D music experience and all three departed in a lively emotional state.

Witnessing this made me feel like I was giving them something back. They have been extremely encouraging, quick to respond, helpful and generous with their time. I also have sincere appreciation for my examining committee, Dr. Mark-David Hosale, Dr. Eldad Tsabary, and especially Dr. Randolph Peters, who bestowed his extraordinary expertise.

Others who I must acknowledge for their inspiration and help are: Emma Novotny; Adrian Novotny; Stanley Heinonen; Helga Plumb; George Novotny; Barry McNicol; Matthew Lee; Henry Lee; Sara Heinonen; Jim Moske; Paul Barton (Lenbrook Industries, PSB speakers); James Tanner and Ian Cahoun (Bryston); Dave Dysart and Dave Morrison (ISO acoustics); Graham Collins and Nick Lobodycz (Long & McQuade Pro); Rune Jacobson (Merging Technologies); Krišjānis Geidāns (Sonarworks); Warren Beck and Toronto AES; Jim Norris; John Harris (Harris Institute); Bill King (7 Arts and Billboard Magazine); Dan Gumble (Headliner Radio); Professor Barry Elmes; Professor Micheal Coghlan; Professor Wells Gordon; Professor Steven Ward; Jeff Riches; Claude Desjardins; Claudia Theriault; Dr. Paul Hoffert; Brenda Hoffert; Joyce Zemens-CM; Jeff Tremblay; Royce Williamson; Fergus Hambleton; Terry Clarke-CM; Joe Sealy-CM; Phil Dwyer-CM; Stefan Bauer; Robi Botos; Ewen Farncombe; Lorne Nehring; Daniel Barnes; Norbert Botos; Chad Irshick; Richard Chycki; Frank Morrone; Damian Kearns; Andrew Hermant; Mariana C Hutten Czapski; Ron Searles; Hayward Parrot; John Wheels Hurlbut; Doug McClement; Ryan Mc Cambridge; Josh Bowman; Terry Medwedick; Edwin Segeren; Atom Egoyan; Amin Bhatia; John Herberman (Toronto Ravel); Tibebe Woldemichael; John Weir; Keith Loach; Michele Mayer; Michael Francis Nunan and Carson Mifsud (Eventide Microsystems); Justin Gray; Morten Lindberg; Marcela Rada; Jinting Zhao (She-Q Audio); Rob Wells; Greg Wells; Sam Hocking; Dr. Hyunkook Lee; Mark Wendl; Tim Canfer; Dr. Tom Reynolds; Dr. Jesse Feyen; Dr. Jan Gullö (J-O); Dr. David Thyén; Dr.

Gary Bromham; Dr. Nyssim Lefford; Dr. Zack Bressler; Dr. William Moylan, Dr. Andrew Bourbon; Dr. Dan James; Dr. Norman Ludwin; Dr. Rob Toulson; Dr. Mike Exarchos (AKA Stereo Mike); Dr. Jo Lord; Dr. Justin Paterson; Dr. Paul Thompson; Dr. Nikhil Deshpande and Emily Rittenhouse (Eventide Audio); Professor Russ Hepworth Sawyer; Professor Claus Sohn Anderson; Professor Kirk McNally; Dr. Simon Zagorski-Thomas; Dr. Shara Rambarran; Dr. Joe Bennett; Dr. Maria Karam; Emma Brooks (Dolby); Innovation in Music; Society of Music Production Research (SMPR); Art of Record Production (ARP); Routledge Taylor & Francis UK; John Krivit; Advisory Educator Groups—Eventide; The Les Paul Foundation; OekSound; Audio Movers; Bettermaker, and all the other guests who auditioned 3D music in the Hyper Near-Field Tiny Studio, as shown below.



**Figure 1:** The Hyper Near-Field Tiny Studio at mix position,

Photography by Paul Novotny, 2025.

## TABLE OF CONTENTS

|                                                                                                  |       |
|--------------------------------------------------------------------------------------------------|-------|
| ABSTRACT .....                                                                                   | II    |
| DEDICATION.....                                                                                  | IV    |
| ACKNOWLEDGMENTS .....                                                                            | V     |
| TABLE OF CONTENTS.....                                                                           | VIII  |
| LIST OF TABLES .....                                                                             | XIV   |
| LIST OF FIGURES .....                                                                            | XV    |
| ABBREVIATIONS .....                                                                              | XVI   |
| ETHICAL CONDUCT .....                                                                            | XVIII |
| THE AUTHOR.....                                                                                  | XIX   |
| PREFACE.....                                                                                     | XX    |
| EPISODE 1—PILOT .....                                                                            | 1     |
| CHAPTER 1: FRAMING THE JOURNEY .....                                                             | 1     |
| 1.1 INTRODUCTION.....                                                                            | 1     |
| 1.2 INSPIRATION AND RATIONALE .....                                                              | 2     |
| 1.3 STRUCTURE, INTENTION AND ORGANIZATION .....                                                  | 3     |
| 1.4 PURPOSE, APPROACH, AND METHOD .....                                                          | 4     |
| 1.5 OBJECTIVES AND STYLE.....                                                                    | 5     |
| 1.6 LITERATURE REVIEWS, EPISODES 1 AND 2.....                                                    | 7     |
| 1.6.1 <i>A Scientific Approach</i> .....                                                         | 7     |
| 1.6.2 <i>The Ethos</i> .....                                                                     | 8     |
| 1.6.2.1 <i>The Sweet Anticipation of “Rocket Man”</i> .....                                      | 10    |
| 1.6.2.2 <i>Spatial Context</i> .....                                                             | 11    |
| 1.6.3 <i>Literature Review for Episode 2</i> .....                                               | 13    |
| 1.6.3.1 <i>Psychoacoustic Literature and Binaural Rendering (APL)</i> .....                      | 14    |
| 1.6.3.2 <i>William Moylan, Francis Rumsey, Analysis and Evaluation of Reproduced Sound</i> ..... | 16    |
| 1.7 LITERATURE REVIEW, EPISODE 3 .....                                                           | 18    |
| 1.7.1 <i>Loudspeakers and Rooms</i> .....                                                        | 19    |
| 1.7.2 <i>Blue Sky Speakers, PSB Speakers, and Floyd Toole</i> .....                              | 20    |

|                                                                                                               |           |
|---------------------------------------------------------------------------------------------------------------|-----------|
| <b>CHAPTER 2: EXAMINE MILESTONES IN MULTICHANNEL AUDIO.....</b>                                               | <b>22</b> |
| 2.1 MILESTONES IN MULTICHANNEL LISTENER EXPERIENCE .....                                                      | 22        |
| 2.2 DOLBY ATMOS IN MAINSTREAM MUSIC, MEDIA, AUTOMOBILES, AND EDM .....                                        | 23        |
| <b>CHAPTER 3: CHOREOGRAPHING CONTEXT, EXPECTATION, EXPERIENCE, AND EMOTION .....</b>                          | <b>24</b> |
| 3.1 CONTEXT, EXPECTATION, EXPERIENCE, AND EMOTION IN 3D MUSIC .....                                           | 24        |
| 3.2 PRACTITIONERS OF EMOTION .....                                                                            | 27        |
| 3.3 GENEVA EMOTIONAL MUSIC SCALE (GEMS) .....                                                                 | 28        |
| 3.4 COMPOSING WITH SPATIAL CONTEXT .....                                                                      | 29        |
| 3.4.1 <i>Spaces and Places</i> .....                                                                          | 29        |
| 3.4.2 <i>The Political Economy of Egalitarian and Elite Spaces, and Places</i> .....                          | 30        |
| 3.4.3 <i>The Political Economy of Multichannel Virtual Spatiality</i> .....                                   | 32        |
| 3.4.4 <i>Spatial Recontextualization</i> .....                                                                | 34        |
| 3.4.5 <i>Spatiotemporal Orchestration (of Space and Time)</i> .....                                           | 35        |
| 3.4.6 <i>Notation of Spatiotemporal Orchestration</i> .....                                                   | 38        |
| 3.5 FREE- AND CLOSED FIELD SONIC EXPERIENCES.....                                                             | 43        |
| 3.5.1 <i>Acoustic Ecology</i> .....                                                                           | 43        |
| 3.5.2 <i>Acoustemology</i> .....                                                                              | 43        |
| 3.5.3 <i>Techoustemology</i> .....                                                                            | 44        |
| 3.5.4 <i>Perceiving Clairaudience</i> .....                                                                   | 45        |
| 3.5.5 <i>Perceiving Infrasound</i> .....                                                                      | 45        |
| 3.6 LISTENING REGRESSION .....                                                                                | 48        |
| 3.7 RUSH—ACOUSTIC, AND PSYCHOLOGICAL IMMERSION .....                                                          | 50        |
| <b>CHAPTER 4: EXAMINE EXPERIENCES IN MUSIC SPATIALITY .....</b>                                               | <b>53</b> |
| 4.1 FORMING SPATIAL COGNITIVE IMAGE SCHEMATA .....                                                            | 53        |
| 4.1.2 <i>Early Experiences in Music Spatiality</i> .....                                                      | 55        |
| 4.1.3 <i>Hyper Near-Field Sound Energy</i> .....                                                              | 56        |
| 4.2 SPATIAL GAMUT, ATTRIBUTES, SUPER COMPONENTS, AND THE 9.1.6 OBJECT-BED .....                               | 58        |
| <b>CHAPTER 5: EXAMINE, ORGANIZE AND PINPOINT THE COMPONENTS .....</b>                                         | <b>61</b> |
| 5.1 ORGANIZE THE COMPONENTS .....                                                                             | 61        |
| 5.2 <i>William Moylan’s Organization of Spatial Perspectives, Properties, Variables, and Attributes</i> ..... | 61        |
| 5.3 <i>NASA: Lost in Space</i> .....                                                                          | 63        |

|                                                                                                      |           |
|------------------------------------------------------------------------------------------------------|-----------|
| <b>EPISODE 2 .....</b>                                                                               | <b>64</b> |
| <b>CHAPTER 6: 3D SPATIAL MUSIC THEORY AND SEMANTIC ANALYSIS SYNERGY .....</b>                        | <b>64</b> |
| 6.1 OVERVIEW .....                                                                                   | 64        |
| 6.2 EPISODE 2 STRUCTURE.....                                                                         | 65        |
| 6.3 INTRODUCTION.....                                                                                | 66        |
| 6.3.1 <i>Seven Principles for Perception of 3D Spatiality in Music</i> .....                         | 66        |
| 6.4 ANALYSIS SYSTEM EVOLUTION .....                                                                  | 67        |
| 6.5 RELATIONSHIPS, PERSPECTIVES, PROPERTIES AND ATTRIBUTES .....                                     | 68        |
| 6.5.1 <i>Relationships Determine Perspectives and the Circular Recording Method</i> .....            | 68        |
| 6.5.2 <i>Properties Determine Attributes</i> .....                                                   | 69        |
| 6.5.3 <i>The Subject-Listener</i> .....                                                              | 69        |
| 6.6 HUMAN INDUCTION AND PERCEPTION .....                                                             | 70        |
| 6.7 INFLUENTIAL LITERATURE AND CONCEPT MAPS .....                                                    | 71        |
| 6.7.1 <i>Cognitive Image Schemata and Metaphor</i> .....                                             | 72        |
| 6.7.2 <i>Evaluating Proxemics</i> .....                                                              | 72        |
| 6.7.3 <i>Evaluating Spatial Quality with Semantic Language (Rumsey)</i> .....                        | 73        |
| 6.8 ASSOCIATING THE SPATIAL ATTRIBUTES OF CONCERT HALLS AND THEATRES.....                            | 76        |
| 6.9 AUDITORY SPATIAL IMPRESSION (ASI) AND FIVE DIMENSIONAL SUPER COMPONENTS .....                    | 77        |
| 6.10 ANATOMICAL ASSOCIATIONS.....                                                                    | 79        |
| 6.11 NATURAL SOUND-BOX EXPERIENCE INGRAINS A PALPABLE SENSE OF SPACIOUSNESS .....                    | 81        |
| 6.12 WILLIAM MOYLAN’S CONCEPT MAP AND FRAMEWORK .....                                                | 82        |
| 6.12.1 <i>Moylan’s Concept Map</i> .....                                                             | 83        |
| 6.13 THE 3D CONCEPT MAP: A THESAURUS OF CONCEPTS, VOCABULARY, AND TERMINOLOGIES .....                | 84        |
| 6.13.1 <i>Moylan’s Concept Map Houses the Five-Facet Typology</i> .....                              | 91        |
| 6.13.2 <i>The Five-Facet Typology with Semantic Descriptions of the Substantial Components</i> ..... | 92        |
| 6.14 IMPROVED PERCEPTION WITH EAR TRAINING EXERCISES.....                                            | 92        |
| 6.15 THE ANALYSIS PROCESS .....                                                                      | 93        |
| 6.16 ANALYSIS 1: “RANDOMIZED,” <i>SPEICHER 120</i> .....                                             | 94        |
| 6.17 ANALYSIS 2: “OUTSKIRTS,” SAM HUNT .....                                                         | 96        |
| 6.18 ANALYSIS 3: VOCAL STAGING, THREE COMPARISONS .....                                              | 98        |
| 6.18.1: <i>“Heart of Rockn’ Roll,” Mixed by Bob Clearmountain</i> .....                              | 98        |
| 6.18.2: <i>“This is Home,” Mixed by Hans Martin Buff</i> .....                                       | 100       |
| 6.18.3: <i>“Jaat Du Maqabla,” Sidu Moose Wala (Mix Engineer Unknown)</i> .....                       | 100       |
| 6.19 ANALYSIS 4: “ROCKET MAN”, MIXED BY GREG PENNY AND BRYAN PENNINGTON.....                         | 101       |
| 6.19.1 <i>ITPRA Response Analysis of “Rocket Man”</i> .....                                          | 102       |
| 6.20 ANALYSIS 5: “ON THE RUN,” PINK FLOYD, <i>THE DARK SIDE OF THE MOON</i> .....                    | 105       |



|                                                                      |            |
|----------------------------------------------------------------------|------------|
| 6.20.1 Channel Usage .....                                           | 106        |
| 6.20.2 Storytelling .....                                            | 106        |
| 6.20.3 Scene by Scene Activity Breakdown .....                       | 108        |
| 6.21 EPISODE REVIEW AND SUMMARY .....                                | 110        |
| 6.22 EPISODE CONCLUSION: .....                                       | 111        |
| 6.23 NASA: FOUND IN SPACE .....                                      | 112        |
| <b>EPISODE 3 .....</b>                                               | <b>113</b> |
| <b>CHAPTER 7: THE HYPER NEAR-FIELD TINY STUDIO .....</b>             | <b>113</b> |
| 7.1 OVERVIEW .....                                                   | 113        |
| 7.2 INTRODUCTION.....                                                | 114        |
| 7.3 TINY STUDIO FIRST STEPS.....                                     | 116        |
| 7.4 WHY IS A TINY STUDIO NECESSARY .....                             | 118        |
| 7.5 PSYCHOACOUSTICS, SPEAKER, AND HEADPHONE WORKFLOWS .....          | 118        |
| 7.6 ATMOS ECONOMICS .....                                            | 119        |
| 7.7 AUDIO HISTORY, SPEAKER LOCATION, AND FORMAT EVOLUTION.....       | 120        |
| 7.8 ATMOS BASIC THEORY AND OPERATION .....                           | 120        |
| 7.9 A, B, AND QC CHAINS .....                                        | 121        |
| 7.10 DESIGN WITH DARDT AND DO MORE WITH LESS.....                    | 122        |
| 7.11 DECORRELATED SOUND FROM A CORRELATED SPEAKER LAYOUT .....       | 122        |
| 7.12 THE COMPONENTS OF A CORRELATED PLAYBACK SYSTEM.....             | 126        |
| 7.13 SMALL CONTROL ROOM ACOUSTICS AND THE “MY ROOM PRINCIPLE.” ..... | 127        |
| 7.14 HYPER NEAR-FIELD ROOM GEOMETRY AND ROOM VOLUME .....            | 131        |
| 7.15 STANDING WAVES AND THE SUBWOOFER POSITION .....                 | 133        |
| 7.16 DOLBY ROOM CURVE .....                                          | 133        |
| 7.17 TIMBRALLY NEUTRAL CANADIAN LOUDSPEAKERS .....                   | 134        |
| 7.18 SPEAKER ISOLATION .....                                         | 137        |
| 7.19 MEASUREMENT TOOLS .....                                         | 137        |
| 7.20 SOUND PRESSURE LEVEL (SPL) SPECIFICATIONS .....                 | 137        |
| 7.21 AUDIO OVER INTERNET PROTOCOL .....                              | 138        |
| 7.22 ROOM TUNING AND CALIBRATION.....                                | 138        |
| 7.23 EQUIPMENT LIST (HN-FTS).....                                    | 141        |
| 7.24 ECONOMIC COMPARISON .....                                       | 143        |
| 7.25 EPISODE SUMMARY AND CONCLUSION .....                            | 145        |

|                                                                              |            |
|------------------------------------------------------------------------------|------------|
| <b>CHAPTER 8: <i>FOUND IN SPACE</i>, REVIEW AND SUMMARY .....</b>            | <b>148</b> |
| 8.1 ORGANIZATION OF THE APPEALS, SUMMARY OF THE METAPHYSICAL COMPONENTS..... | 148        |
| 8.2 THE OBJECTIVE .....                                                      | 148        |
| 8.3 THE NEED .....                                                           | 149        |
| 8.4 PRINCIPLES FOR SPATIAL ANALYSIS (REVIEW).....                            | 150        |
| 8.5 THE ACOUSTEMOLOGICAL COMPONENTS .....                                    | 150        |
| 8.6 THE TECHOUSTEMOLOGICAL COMPONENTS .....                                  | 151        |
| 8.7 THE SUPER COMPONENTS—SW, ASW, ES, AEW, LEV.....                          | 152        |
| 8.7.1 <i>A Highly Unusual Use of Vocal Source Width (SW)</i> .....           | 153        |
| 8.8 THE ORDER OF OPERATION FOR 3D MUSIC ANALYSIS AND PRAXIS .....            | 153        |
| 8.8.1 <i>Cross-Domain Rationales</i> .....                                   | 154        |
| 8.9 3D DIAGRAMMATIC SPATIAL SUBSTANCE SUMMARY MAP, (SPIDER SHEET).....       | 154        |
| 8.9.1 <i>The 3D Diagrammatic Spatial Substance Map</i> .....                 | 156        |
| 8.10 THE HYPER NEAR-FIELD TINY STUDIO.....                                   | 156        |
| 8.11 THE ECONOMICS OF THE HN-FTS.....                                        | 157        |
| 8.12 GENERAL DISCUSSION, IMPACT, AND CONTRIBUTION OF KNOWLEDGE .....         | 157        |
| <b>CHAPTER 9: CONCLUSION.....</b>                                            | <b>159</b> |
| 9.1 THE SKEPTICS .....                                                       | 159        |
| 9.2 ENDING THE REGRESSIVE LISTENING TREND.....                               | 160        |
| 9.3 THE DEVELOPMENT OF OPEN-SOURCE 3D AUDIO .....                            | 161        |
| 9.4 PREPARING THE NEXT GENERATION OF 3D MUSIC CREATOR .....                  | 161        |
| <b>BIBLIOGRAPHY · EPISODE 1 (PILOT) .....</b>                                | <b>163</b> |
| <b>BIBLIOGRAPHY · EPISODE 2 .....</b>                                        | <b>170</b> |
| <b>BIBLIOGRAPHY · EPISODE 3 .....</b>                                        | <b>174</b> |
| <b>WEBSITES VISITED · EPISODE 1.....</b>                                     | <b>178</b> |
| <b>WEBSITES VISITED · EPISODE 2.....</b>                                     | <b>179</b> |
| <b>WEBSITES VISITED · EPISODE 3.....</b>                                     | <b>179</b> |
| <b>WEBSITES VISITED · CONCLUSION.....</b>                                    | <b>180</b> |
| <b>DISCOGRAPHY.....</b>                                                      | <b>180</b> |

|                                              |            |
|----------------------------------------------|------------|
| <b>GLOSSARY .....</b>                        | <b>181</b> |
| <i>Acoustics, Conceptual, Physics:</i> ..... | 181        |
| <i>Anatomy:</i> .....                        | 183        |
| <i>Cognitive Science:</i> .....              | 184        |
| <i>Proxemic Zones:</i> .....                 | 184        |
| <i>Musicological:</i> .....                  | 184        |
| <i>Speaker Sphere: (9.1.6)</i> .....         | 185        |
| <i>Technological:</i> .....                  | 186        |
| <i>Theatrical and Stage:</i> .....           | 186        |
| <b>IMAGE GALLERY .....</b>                   | <b>187</b> |
| <b>MEDIA LINKS .....</b>                     | <b>188</b> |
| <b>FOUND IN SPACE .....</b>                  | <b>188</b> |
| <b>THE END.....</b>                          | <b>188</b> |

## LIST OF TABLES

|                                                                                                             |     |
|-------------------------------------------------------------------------------------------------------------|-----|
| <b>Table 1:</b> The ITPRA response system (Huron, 2006). .....                                              | 11  |
| <b>Table 2:</b> Geneva Emotional Music Scale (GEMS). .....                                                  | 28  |
| <b>Table 3:</b> Direct sound attributes, corresponding objective criteria, properties (Rumsey, 2002). ..... | 74  |
| <b>Table 4:</b> Indirect sound attributes, corresponding objective criteria, properties (ibid). .....       | 74  |
| <b>Table 5:</b> Floyd Toole’s Concept Map for evaluation of spatial quality (Rumsey, 2002). .....           | 75  |
| <b>Table 6:</b> IEC 60268, Concept Map for overall spatial quality. ....                                    | 75  |
| <b>Table 7:</b> Moylan's Concept Map. ....                                                                  | 83  |
| <b>Table 8:</b> The 3D Concept Map organized using William Moylan’s framework.....                          | 90  |
| <b>Table 9:</b> Moylan's Concept Map housing the Five-Facet Typology. ....                                  | 91  |
| <b>Table 10:</b> The Five-Facet Typology. ....                                                              | 92  |
| <b>Table 11:</b> ITPRA response analysis of “Rocket Man.” .....                                             | 104 |
| <b>Table 12:</b> Speaker distance to mix-position. ....                                                     | 124 |
| <b>Table 13:</b> DARDT-RT <sub>60</sub> times for this author’s Hyper Near-Field Tiny Studio (HN-FTS). .... | 131 |
| <b>Table 14:</b> Comparison of the Near-Field to the Hyper Near-Field. ....                                 | 132 |
| <b>Table 15:</b> Hyper Near-Field Tiny Studio speaker timing delays.....                                    | 139 |
| <b>Table 16:</b> Hyper Near-Field Tiny Studio equipment list (10_05_25). ....                               | 142 |

## LIST OF FIGURES

|                                                                                                                               |      |
|-------------------------------------------------------------------------------------------------------------------------------|------|
| <b>Figure 1:</b> The Hyper Near-Field Tiny Studio at mix position,.....                                                       | vii  |
| <b>Figure 2:</b> <i>Found in Space</i> , cover. ....                                                                          | xvii |
| <b>Figure 3:</b> Spatiotemporal orchestration notation using numerical coordinates. ....                                      | 40   |
| <b>Figure 4:</b> 360° coordinates of azimuth, elevation, and distance. ....                                                   | 41   |
| <b>Figure 5:</b> The ASCENDO 80” cinematic subwoofer is custom designed for infrasound. ....                                  | 48   |
| <b>Figure 6:</b> Impedance, wattage and frequency response (Toole, 2018). ....                                                | 57   |
| <b>Figure 7:</b> Lonely Galaxy Lost in Space. ....                                                                            | 63   |
| <b>Figure 8:</b> Metaphoric theatrical associations. ....                                                                     | 79   |
| <b>Figure 9:</b> Anatomical associations. ....                                                                                | 80   |
| <b>Figure 10:</b> This author’s reimagined virtual Dolby Atmos sound-box for music. ....                                      | 81   |
| <b>Figure 11:</b> Marc Romboy in his studio. ....                                                                             | 94   |
| <b>Figure 12:</b> “Outskirts,” Sam Hunt. ....                                                                                 | 96   |
| <b>Figure 13:</b> “Outskirts” channel breakdown. ....                                                                         | 97   |
| <b>Figure 14:</b> Vocal Staging, Huey Lewis, Peter Gabriel, Sidu Moose Wala. ....                                             | 98   |
| <b>Figure 15:</b> Vocal staging: 1) Bob Clearmountain: 2) Hans Martin Buff; 3) mixer unknown. ....                            | 98   |
| <b>Figure 16:</b> “Rocket Man,” <i>Honky Château</i> , Elton John. ....                                                       | 101  |
| <b>Figure 17:</b> <i>Honky Chateau</i> , iTunes review. ....                                                                  | 101  |
| <b>Figure 18:</b> “On The Run”, <i>The Dark Side of The Moon</i> , 50 <sup>th</sup> Anniversary, Remastered, Pink Floyd. .... | 105  |
| <b>Figure 19:</b> “On The Run,” channel usage. ....                                                                           | 105  |
| <b>Figure 20:</b> Scene by scene activity breakdown. ....                                                                     | 108  |
| <b>Figure 21:</b> Gemini-7 – Earth-Sky View – Limb – Outer Space. ....                                                        | 112  |
| <b>Figure 22:</b> Hyper Near-Field Tiny Studio (HNFTS), photography by Paul Novotny, 2022. ....                               | 125  |
| <b>Figure 23:</b> Hyper Near-Field Tiny Studio, sole mix-position, photography by Paul Novotny, 2022. ....                    | 126  |
| <b>Figure 24:</b> “MyRoom Principle”—studio 1 with hybrid surfaces. ....                                                      | 129  |
| <b>Figure 25:</b> Hyper Near-Field Tiny Studio, dimension of standing waves. ....                                             | 133  |
| <b>Figure 26:</b> HN-FTS Aggregate room response, 12_19_23, measured by Edwin Segeren. ....                                   | 141  |
| <b>Figure 27:</b> Near-Field economic estimate. ....                                                                          | 143  |
| <b>Figure 28:</b> Hyper Near-Field economic comparison. ....                                                                  | 144  |
| <b>Figure 29:</b> Near-Field economic breakdown by chain. ....                                                                | 144  |
| <b>Figure 30:</b> Hyper Near-Field economic breakdown by chain. ....                                                          | 145  |
| <b>Figure 31:</b> 3D diagrammatic spatial substance map, (spider sheet). ....                                                 | 156  |
| <b>Figure 32:</b> Building the HN-FTS, wiring EDAC connectors. ....                                                           | 187  |
| <b>Figure 33:</b> Building the HN-FTS, patch bay with connectors. ....                                                        | 187  |

## ABBREVIATIONS

(selected)

|          |                                                     |
|----------|-----------------------------------------------------|
| 3D       | 3-dimensional                                       |
| AIDT     | Average Interaural Time Delay                       |
| ASA      | Audio scene analysis                                |
| ASW      | Apparent Source Width                               |
| CBA      | Channel Based Audio                                 |
| DFC      | Diffuse Field Correlation                           |
| ER       | Early Reflection                                    |
| FCC      | Frequency-dependant cross-correlation               |
| GEM      | Geneva Emotional Music Scale                        |
| HOA      | Higher-order Ambisonics                             |
| HRTF     | Head-related transfer function                      |
| IACC     | Interaural cross-correlation                        |
| ICTD/ITD | Inter-channel time difference                       |
| IPD      | Interaural phase difference                         |
| ITD      | Interaural time difference                          |
| LEDT     | Lateral early decay time                            |
| LEV      | Listener envelopment                                |
| LEF      | Lateral energy fraction                             |
| LF       | Lateral fraction                                    |
| LG       | Lateral gain                                        |
| LR       | Late reflection                                     |
| KBPS     | Kilobytes per second                                |
| MF       | Mid-frequency                                       |
| MIR      | Music information retrieval                         |
| MRIR     | Multichannel room impulse response                  |
| MS       | Mid-side (microphone technique)                     |
| OLE      | Overall listening experience                        |
| PICC     | Perceptual interaural cross-correlation coefficient |
| PLP      | Preferred listening position                        |
| $R$      | Measurement of spatial impression                   |
| SASC     | Spatial audio scene characterization                |
| SBA      | Scene-based audio                                   |
| TOA      | Third-order Ambisonics                              |
| VOG      | Voice of God (speaker position)                     |
| XTC      | Crosstalk cancellation                              |
| $\tau$   | Time constant                                       |

# Found in Space

Journey to the Third Dimension of Music Spatiality in the Dolby Atmos Format

Composition Ethos

Spatial Music Theory and Semantic Analysis Synergy

Philosophy and Benchmark for a Hyper Near-Field Tiny Studio

The Metaphysical Componentry in a Compelling Spatial Music Sound Field

**Figure 2:** *Found in Space*, cover.

## ETHICAL CONDUCT

No Artificial Intelligence (AI) was used in the authorship or research of this dissertation. However, this author recognizes the possibilities of AI for the future of ethically efficient learning using neuroscience-based protocols, metadata channels and devices. Please see Dr. Poppy Crum (Dolby Laboratories former Chief Scientist) in conversation with Dr. Andrew Huberman, Neuroscientist at Stanford University School of Medicine (Huberman, 2025).





## THE AUTHOR



Paul Novotny in his Dolby Atmos Hyper Near-Field Tiny Studio

Photography: Bill King, used with permission.

## PREFACE

*Found in Space* reveals the substantial metaphysical components involved for the creation of a compelling sound field in 3D recorded popular music using the Dolby Atmos format. Combining fixed multichannel audio, audio-objects and metadata, the Atmos format enables music playback to scale up and down various multi-speaker systems in a variety of listening scenarios. It provides mechanisms for the creator to define how their mix should respond to these scenarios. The Atmos playback renderer has awareness of the speakers in a calibrated listening space. Creator defined metadata specifies audio-object location within a virtual 360-degree sound field.

Dolby Atmos separates itself from fixed multi-speaker Surround Sound formats with its ability to locationally animate dynamic audio-objects within a 3D sound field. *Found in Space* focuses on mainstream 3D popular music in the Dolby Atmos format, heard in the Apple Spatial Music ecosystem of media and devices.

Since June of 2025, Dolby Atmos and the Apple Spatial Audio Format (ASAF) started using the Apple Positional Audio Codec (APAC) so Apple devices could encode head-tracked, multi-source audio with an iPhone 16. The four-microphone device uses the APAC to encode a 3D audio sound field using 1<sup>st</sup> order ambisonics and then embeds that spatial audio into media using the Apple Projected Media Profile (APMP). An iPhone 16 movie file has eleven audio channels. [Please see](#) Dr. Deep Sen, the Lead Immersive Audio Architect at Apple describe these groundbreaking advancements for the Apple Spatial Format [\(Sen, 2025\)](#). Research and development of immersive media, devices and metadata is progressing rapidly as the industry understands and describes what contextual information can be embedded into consumer media.

Michael Francis Nunan—Senior Architect of R&D at Evertz Microsystems—is advancing discoveries about the digital twinning of human context as metadata, while media and devices are rapidly evolving toward a new *immersive sensorium*.

In her discussion with Dr. Andrew Huberman (a neuroscientist at Stanford University), Dr. Poppy Crum—former Dolby Laboratories Chief Research Scientist—discusses how Artificial Intelligence (AI) and digital twinning can model physicality, systems and processes to enhance ethically efficient learning using neuroscience-based protocols, metadata channels, devices and applications (apps).

The fast-evolving research of a *contextually aware* and *interconnected digital ecosystem* presents the “adjacent possible evolution” (Kauffman, 2020, p.9); (Kauffman, 2013, p.610) of life-like immersive media, professional and consumer devices, in combination with experiential learning. Dr. William Moylan is central to my research because his life’s work has exhaustively explored the description of metaphysical spatial componentry that he calls “an illusory element” of recorded music (Moylan, 2020). He determined that perspectives, properties and attributes are essential principles for the understanding of spatial concepts in the stereo format. I extended his framework to 3D music which required the addition of “relationships” as a necessary principle.

*Found in Space* extends stereo music spatial analysis to 3D and semantically describes its salient metaphysical componentry. Dr. Moylan is currently working with Hans Martin Buff (Peter Gabriel-i/o, 2024 Grammy Award winning mixer/spatiotemporal orchestrator) to describe Buff’s Atmos praxis. In a recent conversation with Michael Francis Nunan, he stated that “what’s next for the new *immersive sensorium* is to undertake the equivalent research to that of *Found in Space*, but in the domain of visualisation and cinematography.”

# EPISODE 1—PILOT

## Chapter 1: Framing the Journey

---

### 1.1 Introduction

My life's journey began in the year 1958, born in the city of Peterborough, Ontario, Canada, of a musical family. My father of Czech heritage was a professional pianist and my mother, also from Czechoslovakia, a singer. They emigrated to Canada in 1951 from Prague, a city where Mozart and Beethoven lived on Nerudova Street—named after Czech writer Jan Neruda. I'm a professional musician, composer, and audio producer, operating in Toronto since 1980. In essence, my vocation involves making decisions about how to compose, organise, and express the space and time (spatiotemporality) of audio-objects heard in recorded music and sound.

With years of practice, I honed an ethos that defined my music's character, context, creative essence, and construction of expectations for an optimal listening experience. Beginning in 2004, I've worked with music and broadcast audio in the 5.1 Surround Sound format and learned that, in contrast to stereo, the added dimension of rear depth is a potent element for recorded music. It enhanced my music listening experience with a sense of envelopment that touched off deep-seated emotions. In 2020 I began a transition to the three-dimensional (3D) Dolby Atmos format, also known as immersive Spatial Audio.

With hindsight, my journey to the third dimension of music spatiality actually began at age eleven while playing the piano, and now, combining professional experience with scholarly research of concepts, context, theories, vocabulary, terminology, philosophies and benchmarks, I

am concentrating on a comprehensive understanding of the third dimension of music spatiality heard in the Dolby Atmos format.

## 1.2 Inspiration and Rationale

This dissertation's title—*Found in Space*—is inspired by the whimsical 1965 science-fiction television series *Lost in Space*, which chronicled the exploits and adventures of a pioneering expedition that, due to overpopulation on Earth, sought to colonize a family on an Earth-like planet orbiting Alpha Centauri, but they became lost in outer space and crash landed onto an alien world. The lineage of *Lost in Space* reaches back to *Swiss Family Robinson* and Daniel Defoe's *Robinson Crusoe*. Reminiscent of Defoe, *Found in Space*, within its episodic structure, may befittingly utilise his writing style in an attempt to imbue an uneasy rhythmic cadence, as it clarifies 3D music complexities via these uncharted journeys to spatiality.

Endeavouring to orient stereo music creators toward their creation of a compelling 3D music listening experience, *Found in Space* is an auditory quest that documents how practical heuristic knowledge, scholarly research, and determined resourcefulness have combined to find and reveal the substantial metaphysical componentry of an unconflicted 3D music sound field.

The TV show *Lost in Space* recounts expeditions that unfold in *outer space*, while I recount conceptual expeditions that emerge from the *inner mind-space* of 3D music.

Stereo music creators that transition to the Dolby Atmos format can quickly become lost in myriad possibilities for art and craft, technological understandings, and semantic descriptions of spatial componentry. Even eminent 3D music creators struggle to explain it by using abstruse terminology, all the while, their creation of compelling 3D music happens intuitively through a combination of inherent musical sensibility and hard-earned technical prowess. By necessity, an

elemental understanding of 3D music spatiality requires the pinpointing of musicological and technological components to form a precise literacy for what I deem to be a new theoretical idiom called 3D Spatial Music Theory.

*Found in Space* proposes a way to decode and legitimize popular music 3D Spatial Theory. To do this, interdisciplinary concepts and philosophies must be synthesized to establish a holistic rationale. I borrow central ideas from David Huron and David Byrne about music composition, listening experiences, emotions, environments and their spatial context. These ideologies can help a composer to shape their 3D music creation with persuasive expression, reliable transference, a compelling Overall Listening Experience (OLE), and the awakening of far-reaching emotions. Also, pedagogues can interpret and extend these ideas to engender semantic descriptions and novel epistemologies for their teaching of musicological 3D spatial theories.

### 1.3 Structure, Intention and Organization

*Found in Space* intends to find the metaphysical substance of 3D music spatiality and form a sturdy foundation for successive journeys beyond this effort. The reader may not find a traditional sense of flow because the episodes unfold as three separate expeditions, each with a unique focus on answering the central questions. The episodic nature begets measured but deliberate use of repetition as a way to carry forward previously introduced concepts that must cross over episodes. To conclude, a criterion [summary](#) of all the salient metaphysical componentry established in each episode is gathered onto a one-page diagrammatic spider sheet. The episodes can be read in, or out of order and this document's navigation of understandings is reinforced with hyper-linking to related matters.

Defining music's spatiality requires suppositional concepts and context that intersect acoustics,<sup>1</sup> induction, cognition, musicology,<sup>2</sup> and technology,<sup>3</sup> which I sub-categorize into the silos of Acoustemology and Techoustemology<sup>4</sup> as a way to frame the criterion for summary.

A [glossary of selected abbreviations](#), is located in the front matter and separate [bibliographies](#) are organized by episodes [1](#), [2](#), [3](#), and found in the end-matter. References that are shared between episodes are duplicated in their respective episodic bibliographies.

## 1.4 Purpose, Approach, and Method

This dissertation's purpose is pedagogical and is aimed toward professional creators, students, researchers and developers. Obtained from interdisciplinary investigations, scientifically reasoned approaches, metaphoric learning<sup>5</sup> and adjacent extension, the findings are expressed with subjective intertextuality, specifically selected to establish the idiom's dialect.

*Found in Space* blends professionally acquired heuristic knowledge with scholarly research to form theories that stimulate ready understandings, known as *a priori* (prior) “cognitive image schema,”— “the basis of rational thought by way of metaphorical extension,” (Moore, Schmidt, Dockwray, p.85). Essential semantic language is etymologically assessed and

---

<sup>1</sup> Acoustics is an offshoot of physics that studies the waveforms of vibration, sound, ultrasound and infrasound.

<sup>2</sup> The Austrian musicologist Guido Adler first defined the rigorous aim of musicology as “The Science of Music,” or “Musikwissenschaft” (Beard., Gloag, K., 2016, p.xi.).

<sup>3</sup> With over a century of recorded music production and delivery to consumers, 3D Spatial Theory is founded with the belief that musicology and music technology are inextricably intertwined. Paul Théberge argues “any discussion of the role of technology in popular music should begin with a simple premise: without electronic technology popular music in the 21st century is unthinkable” (Théberge, 2001, p.3).

<sup>4</sup> Acoustemology and techoustemology are explained in the glossary. Acoustemology is a term that stems from acoustics and sonic ecology, and techoustemology is a complimentary term. ([Please also see chapter 3.52 and 3.53](#)).

<sup>5</sup> In his 2022 book entitled *A History of Inspiration through Metaphors of Learning*, Robert Nelson argues, “We mishandle educational practice through a focus on technical process and machinery rather than the imaginary animating vision that propagates inspired study through metaphor” ... “metaphors are intrinsic to all our educational ambitions” (Nelson, 2002, p.1).

carefully selected from closely associated disciplines as a way to describe “what matters most” (Nelson, 2022, p.2) about the topics.

## 1.5 Objectives and Style

**Episode 1**, (The Pilot Episode) contains five chapters of examination that propose tangible foundational forces for the governance of 3D popular music spatial theory. It is written in the first person and establishes my ethos of rationales, metaphoric concepts, contexts, vocabulary, terminologies and interdisciplinary theories about 3D spatial music composition,<sup>6</sup> spatiotemporal orchestration,<sup>7</sup> spatial context,<sup>8</sup> spatial creativity,<sup>9</sup> philosophy, psychoacoustics, acoustics, and biological sound induction. Additionally, Episode 1 [chapter 3.7](#) shares a captivating anecdote—provided by Rich Chycki—explaining the first time that Rush guitarist Alex Lifeson experienced Rich’s Atmos mix of “Tom Sawyer.”

**Episode 2** is written in the third person style and introduces 3D Spatial Music Theory with a procedure for 3D music spatial analysis. It is a three-step musicological system which I call a Synergy. The word synergy<sup>10</sup> implies that three components—a Concept Map, a Five-Facet Typology, and 3D ear training exercises<sup>11</sup> can work together synergistically toward improving the understanding and perception of 3D spatial music.

---

<sup>6</sup> I propose this to be a distinct approach to compositional art, craft, and pedagogy.

<sup>7</sup> I propose this to be a distinct practice of orchestration.

<sup>8</sup> I propose this to be a distinct musicological approach to 3D Spatial Theory. It involves Acoustic Ecology (Schafer, 1994), Acousmatics (Schaeffer, 2017), (Cox, Warner, 2004), ethnography, social, and cultural context (Byrne, 2012).

<sup>9</sup> I propose that spatial creativity can abstract and distort spatiality to create virtual sound fields of spaces and places. It can be informed by concepts of expectation from David Huron (2006), and environmental context from David Byrne (2012).

<sup>10</sup> I am aware of the overuse and stigmatization of the word synergy, but I could not identify a better word to accurately encapsulate this 3D analysis system.

<sup>11</sup> Ear training for 3D music improves the brain’s ability to differentiate spatial auditorial contrasts.



**Episode 3** is also written in the third person and introduces an innovative philosophy and benchmark for an environmentally sustainable Dolby Atmos Hyper Near-Field Tiny Studio (HN-FTS). The HN-FTS is the auditorial research laboratory that I built and used for *Found in Space*.

The HN-FTS philosophy describes how to achieve an economical large scope of premium operation within a hyper-small scale. This studio/lab enables accurate subjective induction of 3D spatial “perceptual elements,” which I organize as 1) Positional Relationships, 2) Cognitive Perspectives, 3) Proportional Properties, and 4) Character Attributes. The HN-FTS benchmark is a form of technological luthiership that is capable of articulating precise 3D spatiotemporal orchestration of audio-objects.

Imported from the global Tiny Home movement (Carrizosa, 2021), the studio’s ecological emphasis employs socially aware technological construction that emphasises buy-local, low-energy consumption, repair and reuse, premium long-life quality, and an economical large scope of operation, within a small scale. This HN-FTS philosophy and benchmark is intended for academic institutions, professional studios, sole operators, and researchers.

All of the episodes embody a holistic ontology for deep understanding of 3D spatial music metaphysical componentry, expressed herein as cross-domain appeals of the ethos, logos and pathos.<sup>12</sup> *Found in Space* expresses its logos scientifically to reveal the innerworkings of the respective topics. The pathos of this dissertation can be experienced by listening (on the Apple Music Streaming service) and perceiving song audio scenes which are analysed between [Chapters 6.16](#) and [6.20](#). The songs are found in a specific [Apple Music playlist](#) that I curated for

---

<sup>12</sup> Aristotle taught that a speaker’s ability to persuade an audience is based on how well the speaker appeals to that audience in three different areas: Logos, Ethos, and Pathos.” These areas combine to form “the componentry of appeals” in the “Rhetorical Triangle” (Lutzke, Henggeler, 2009, p.1).

ready access to what I consider to be spatially compelling 3D music. I update this playlist regularly and invite the reader to listen in as I constantly seek compelling new 3D music releases.

## 1.6 Literature Reviews, Episodes 1 and 2

---

### 1.6.1 A Scientific Approach

The literature of interdisciplinary discourse for *Found in Space* cuts across many domains, but musicology, human biology, psychoacoustics, engineering, and technology provide the primary scopes of reliable reference for scientific theories about 3D spatiality in recorded music. *Found in Space* applies a flexible but consistent admixture of citation styles to reflect its interdisciplinary scientific research.

A scientific approach toward understanding the elements of music was first identified by Guido Adler in 1885. He described musicology as the “science of music” (Beard, Gloag, 2016 p.xi). A scientific approach for learning is explained as “the systematic study of the structure and behaviour of the physical and the natural world through observation, experimentation, and the testing of theories against the evidence obtained” (*New Oxford American Dictionary*, 2018).

Science involves verifiable subjective observations via experiences, whilst subjective observations can also explain the objective grammar of relationships, perspectives, properties, attributes, concepts, practices and benchmarks.

Edwin Pfanzagl-Cardone is an Austrian sound engineer who researches and writes about 3D audio. His 2023 book entitled *The Art and Science of 3D Audio Recording* delves into theory and practical implementations of 3D microphone technique. He asserts that the term 3D is incorrect because it excludes the fourth dimension of time. This observation has led me to coin

the term “Spatiotemporal Orchestration” as a new space and time-based 3D compositional practice. Pfanzagl-Cardone prefers the term 3D over the term ‘immersive audio.’ I agree because the short-cut term “immersive” obscures the difference between 3D acoustic immersion and psychological immersion, both of which are significant in 3D spatially recorded music.

Dr. Bosun Xie possesses a bachelor’s degree in physics as well as Master and Doctor of Science degrees in acoustics. He has a great interest in classical opera, technology and science. His highly scientific book entitled *Spatial Sound, Principles and Applications* is rich with mathematical formulae. In addition, he provides research about subjective and textual analysis. He observes that “spatial sound analysis is a subjective and objective exercise.” He also argues that “spatial sound is an enhanced and immersive set of audio techniques which provides sound in three-dimensional virtual space” (Xie, 2023, p.1). Because music is heard by people, by necessity, popular musicology involves subjective intertextuality and semantic description of contextual meaning, whilst accepting that objective comparisons can also be subjectively perceived as requisite systematic science.

### 1.6.2 The Ethos

I argue that an ethos for 3D popular music spatial theory begins with the primary considerations of music composition and listener experience. For this approach, David Huron’s 2006 book entitled *Sweet Anticipation, Music and the Psychology of Expectation* will provide significant value. Huron describes a “psychological theory of expectation” which he calls “the ITPRA theory” (Huron, 2006, p.5). He presents a logical discussion about composing music’s temporally driven expectations and resulting emotional responses, which can lead the listener to

a reconfigured mind-state of “flow” (Csikszentmihalyi, 2014, p.279). Many times, I’ve witnessed a flow-state *transfix* guests in my Atmos 7.1.4 3D Hyper Near-Field Tiny Studio.

The positive psychological mind state of “flow” (ibid) has been extensively explored by Hungarian American psychologist Mihaly Csikszentmihalyi. He “discovered that people find genuine satisfaction during a state of consciousness called flow” (Csikszentmihalyi, 2004).

Csikszentmihalyi has studied the science of happiness with the mind state of flow, as it controls our consciousness. “In this state they are completely absorbed in an activity, especially an activity which involves creative abilities” (ibid).

Three primary signs of flow exist; 1) distractions are excluded from consciousness, 2) self-consciousness disappears, and 3) the sense of elapsed time disappears. The Atmos format can stimulate these outcomes because it creates strong acoustically immersive experiences that result in rich musical enjoyment, enabling the psychologically immersive mind-state of flow to arise.

David Huron joins cognitive science with the palpable and visceral heartbeat of musical experience. His ITPRA theory affords music scholars a scientific way to conceptualize compositional tension and release in 3D spatiotemporal music. An ITPRA analysis of Elton John’s “Rocket Man” is found in [Chapter 6.19](#).

### 1.6.2.1 The Sweet Anticipation of “Rocket Man”

In a Dolby Professional video interview,<sup>13</sup> which I regard as essential literature for this topic, eminent music mixer Greg Penny describes how the Atmos format can enliven the experience and emotion of a song with anticipation. He describes mixing Elton John’s 1972 song, “Rocket Man,” the first song ever to be mixed in Atmos in 2017. He did it with Dolby’s senior applications engineer Bryan Pennington. Penny describes how he created what he calls “the stun gun effect [...] I’ve got you and then the release after it’s over” (Episode 1, Dolby, 2020, 2:22). Penny says that he tries to “identify emotion in the track, so the machine (Atmos technology) goes away” (ibid, 4:35). He speaks about how his Atmos mix unfolds over ‘space and time’ to have emotional impact.

*“There’s a lot of pop tracks that sort of start on stun and stay there.”*

(ibid, 3:48)

*“Rocket Man” has “a concentration of the lyric and the vocal in the front with piano, and things build, and then come back down again, it’s a really effective way to work in Atmos.”*

(ibid, 4:20)

I interpret Penny to be speaking about his deployment of tension and release over time in his 360° sound field. Penny is actually speaking about the anticipatory stages of “*what, why, where, when, and how,*” within his art and craft.

David Huron describes psychological music expectation and emotion using his scientific ITPRA theory which organizes the human response system by epoch and biological function.

---

<sup>13</sup> Video research literature: available on YouTube, please see Episode 1, websites visited, (Dolby, 2020).

The temporal aspect is described as the event epoch. The first response stage is the imagination (I) of *what*, *why*, and *where* events could unfold. It is a pre-outcome future-oriented behavioural motivation that enables deferred gratification and sets up expectations. The second pre-outcome stage of tension (T) brings optimum arousal and attention in preparation for anticipated events. The prediction (P) response stage is post-outcome, and it provides negative or positive reinforcement to encourage formation of accurate expectations. The post-outcome reaction (R) response is a fast response that assumes a worst-case assessment of the outcome.

I argue that in a joyful 3D music experience the reaction stage can also produce a best-case assessment. The final post-outcome stage is the appraisal response (A). This is where assessment of the final outcome occurs. Please see **Table 1**: The ITPRA response system, and the [ITPRA analysis of “Rocket Man”, Chapter 6.19](#).

| Response System          | Epoch        | Biological Function                                                                                      |
|--------------------------|--------------|----------------------------------------------------------------------------------------------------------|
| (I) Imagination response | Pre-outcome  | Future-oriented behavio[u]ral motivation; enables deferred gratification.                                |
| (T) Tension response     | Pre-outcome  | Optimum arousal and attention in preparation for anticipated events.                                     |
| (P) Prediction response  | Post-outcome | Negative/positive reinforcement to encourage the formation of accurate expectation.                      |
| (R) Reaction response    | Post-outcome | Neurologically fast responses that assume worst-case assessment of the outcome.                          |
| (A) Appraisal response   | Post-outcome | Neurologically complex assessment of the final outcome that results in negative/positive reinforcements. |

**Table 1:** The ITPRA response system (Huron, 2006).

### 1.6.2.2 Spatial Context

In his 2012 book entitled *How Music Works*, David Byrne describes theories about how spatial context determines creative design. He points to differences between free field (outdoor)

and closed field (indoor) spatial contexts. He also describes how they directly affect music creation, composition and instrument design.<sup>14</sup> In the domain of acoustic analysis, a free field outdoor (borderless) spatiality can be closely recreated indoors using an [anechoic chamber](#) because it has no ability to reflect indirect sound (echo or reverb) from its closed field surfaces.

Anechoic chambers are used to test loudspeaker and microphone designs. This acoustic outcome brings forward an essential awareness for the crafting of a critical balance<sup>15</sup> between [direct](#) and [indirect sound](#). In critical balance these components form the Auditory Spatial Image (ASI), the essential proportional property for creation of a compelling 3D music experience.

Critically balancing direct and indirect sound affect perceptions of intelligibility, intensity, and impact, but also codify spatial contexts and listener proximity to the ensemble and persona of the music's protagonist, all of which affect song meaning. Byrne's arguments are reinforced by Dr. Bosun Xie who discusses sound field spatial context from the perspective of audio analysis.

*“Free Field is an important concept in acoustic analysis. It refers to a special sound field in a uniform and isotropic medium in which the influences of boundaries are completely negligible (i.e., the absence of reflections from the boundary). Cases of the free field under actual conditions are rare. Sound Fields can be approximated as a free field only in an anechoic chamber or a local region, which is sufficiently high above ground. However, the concept of the free field specifies an ideal standard condition for acoustic analysis.”*

(Xie, 2023, p.804)

---

<sup>14</sup> This is a good example of combined acoustemological and techoustemological discourse. It indicates how interdisciplinary observations can reinforce new cross-domain (adjacent) rationales and possibilities.

<sup>15</sup> The critical balance is also known as the reverberation radius of a closed field venue. When recording 3D spatial indirect sound reverberation, the directionality of a sound source and microphone pick-up pattern are the essential components. The omni pattern is desirable because it captures a full spectrum from the environment.

Extending from Xie's argument, my Hyper Near-Field Tiny Studio ([Episode 3](#)) relies on a very low (closed field) RT<sub>60</sub> (control room reverb time decay to -60 dB) and implements hyper near (1.2m) distancing of speaker-to-listener sound as a way to achieve its flat aggregate response curve at the sole operator's hyper tuned listening position,<sup>16</sup> making it reliable for 3D Spatial Music analysis.

### 1.6.3 Literature Review for Episode 2

**Episode 2** proposes a semantic analysis synergy to subjectively describe the substance of immersive acoustic spatiality. It involves a *3D Concept Map*, a *Five-Facet Typology* and *Ear Training Exercises*. The Synergy is inspired by William Moylan's pioneering research (Moylan, 2020) and introduces a practical text-based system that provides reliable concepts and semantic language for the analysis of music's 3D spatiality. It is intended for creators, musicologists, 3D Music Information Retrieval (MIR) scientists and the adjacent possible future development of contextual metadata channels for media and devices. The Synergy embraces subjective analysis that identifies 1) Relationships, 2) Perspectives, 3) Properties and 4) Attributes as central codifiers for specific understandings of 3D music meaning and spatiotemporal praxis.

Positional relationships describe listener and audio-object associations as lateral or contralateral (clockwise or counterclockwise). They are formed with instrumental staging and spatiotemporal orchestration. Cognitive perspectives can be [egocentric or allocentric](#). Properties are the proportional objective grammar of spatiotemporal orchestration, art and craft. Attributes describe and characterize musical traits in relation to song meaning, experience and emotion.

---

<sup>16</sup> Low RT<sub>60</sub> is achieved via a hyper-near critical listening distance between the listener and speakers. This effectively removes room sound from the speaker sound heard at the sweet spot listening position.



*“[M]usical ‘meaning’ cannot be limited to translatable signification. In music we look not only for understanding but also enjoyment.”*

(Middleton, 1990, p.247)

Richard Middleton’s philosophy aligns with my ethos—3D music analysis can textualize, characterize and gauge the subjective listener’s optimal playback experience as a translatable signification that is rich with meaning and precipitous to music’s scale of enjoyment and emotional responses. Extending from Middleton, Marcel Zenter’s (well-researched) concept map for the Geneva Emotional Music Scale (GEMS) will soon be discussed in [Chapter 3.3](#).

#### 1.6.3.1 Psychoacoustic Literature and Binaural Rendering (APL)

New research and literature about 3D Spatial Audio Scene Characterization (SASC) is oncoming from the Applied Psychoacoustics Laboratory (APL) under the directorship of tonmeister Dr. Hyunkook Lee at Huddersfield University UK. The APL is bridging a gap between academia and the audio industry. APL’s binaural renderer called Virtuoso is now in wide use by eminent professionals worldwide and Duopan was recently released. It is a psychoacoustic stereo panner for headphone and speaker playback.

In a 2024 empirical research paper, authors Pawel Antoniuk, Stawomir Zieliński, and professor Hyunkook Lee presented ‘a new method for the automatic estimation of [...] “ensemble width”’ in Head Related Transfer Function (HRTF)-convolved binaural music recordings uploaded to Freesound, YouTube, and Vimeo. HRTF is a term that explains human induction of the 3D sound field. It involves interaural timing differences (ITD), interaural level differences, (ILD), and interaural cross-correlation (IACC), which is frequency spectrum based.

The rapidly expanding area of [music information retrieval \(MIR\)](#) is challenged to perceive, analyse and communicate the 3D spatial relationships, perspectives, properties and attributes of music recordings. This task is referred to as Spatial Audio Scene Characterization (SASC). Antoniuk, Zieliński, and Lee report a growing need for the development of 3D analysis systems for SASC in binaural signals. It is from their work that I will propose [five Super Components](#) of 3D spatiality. I believe that these metaphysical Super Components can be highly assistive for efficient creation of [auditory spatial impressions](#) involved in a non-conflicting 3D music sound field. Antoniuk, Zieliński and Lee argue that traditional ensembles are staged in front of the subject-listener “whereas for pop, dance, or electronic music recordings, sound sources tend to be distributed (by mixing engineers) in a much wider way, even to the extent that the listeners are [circularly] surrounded by musicians (360° stage scenario).”<sup>17</sup> These authors espouse concepts from Francis Rumsey’s scene-based paradigm (Rumsey, 2002, p.1) and further suggest that music’s spatial content can be organized in three levels of abstraction: “low, mid, and high” (Antoniuk, Zieliński, Lee, 2024, p.2). They employ terminologies to describe the direct sound physical width of instruments and ensembles. Source width (SW) refers to an instrument such as the drum kit, and ensemble width (EW) refers to the physical width of the entire band. These audio-objects involve direct sound and should not be conflated with terms for

---

<sup>17</sup> (Lee, Millns, 2017). To date, Justin Gray’s 2025 release entitled *Immersive* is the best example I’ve heard of the circular object to subject relationship which involves an egocentric perspective. The listener is placed in the middle or inside of the instrument or ensemble section (strings, brass etc.,) while a circular microphone array functions as the listener.

apparent width. Apparent Source Width (ASW)<sup>18,19,20</sup> and Apparent Ensemble Width (AEW) involve indirect reflected sound.

I speculate that these metaphysical Super Components could bring new agency to Dolby's renderer and guide creators toward a clean 3D sound field with substantial immersive weight and listener envelopment (LEV). They could also improve the conception of ensemble staging, levels of abstraction, critical wet/dry balances, and allocentric proxemics of audio-objects, heard as instrument-to-instrument relationships. SW, EW, ASW, and AEW signify a listener's relationship, proximity, and perspective<sup>21</sup> to audio-objects. Abstraction of these conceptual metaphysical components can involve the onstageing or offstageing of audio-objects placed anywhere within the virtual 360° sound-box.

#### 1.6.3.2 William Moylan, Francis Rumsey, *Analysis and Evaluation of Reproduced Sound*

Throughout his career Dr. William Moylan has tried to understand how recording shapes the song. His book entitled *Recording Analysis: How the Record Shapes the Song*, "is about how the sounds of the recording process add dimensionality and richness to the musical and socio-psychological experience of records" (Moylan, 2020, p.xxv).

In Chapter Eight Moylan focuses on "The Illusion of Space as an Element of Recording." I adopted Dr. Moylan's philosophical approach and organizational structure to design my 3D analysis synergy. He provides a framework of stereo concepts, vocabulary and terminologies which I have adapted and extended to form a broad concept map for 3D music. Moylan describes

---

<sup>18</sup> (Käsbach, Hahmann, May, Dau, 2016)

<sup>19</sup> (Sato, Ando, 2002)

<sup>20</sup> (Barron, Marshall, 1981)

<sup>21</sup> Subjective listener relationships and perspectives will be fully explained in Episode 1.

spatiality as “an illusory element in every recording and as such recordings re-invent the relationships and dimensions of our natural world” (Moylan, 2022). He determines the following three categories of spatial properties: 1) Angular direction and width; 2) Span of distance location and depth; and 3) Dimensions of the environment.

His concept map concisely organizes: 1) Levels of perspective; 2) Spatial properties; and 3) Variables/Attributes, to describe: A) Individual Sound Source; B) Composite Texture; and C) Overall Sound.

To facilitate an analysis synergy for 3D spatial theory, I’ve expanded Moylan’s framework to generate a seven-page thesaurus-like reference guide. It’s a worksheet that exhaustively maps 3D scenarios, concepts, vocabulary, and terms. It reveals seminal semantic language which required simplification. That resulted in a distilled and practical Five-Facet Typology that aims to guide 3D music analysis efficiently and effectively.

Extended research about prior concept maps revealed that in 2002 Francis Rumsey published his system of *Spatial Quality Evaluation for Reproduced Sound*. “Rumsey’s aim is the clarification of semantic language” (Moore, Schmidt, Dockwray, 2011) and his research helped guide the concepts, vocabulary and terminologies for my 3D Concept Map and Five-Facet Typology. Other significant research about music spatiality involves Brøvig-Hanssen and Danielsen (2013), Blauert (1983, 1997), Dockwray and Moore (2010), Meaux, E. and Marchand, S. (2021), Moore (1992), Holman, T. (2008), Patterson and Lee, (2021), Pawlak, Lee, Mäkivirta, and Lund (2023), Rumsey (2002, 2001) and Denis Smalley (Smalley, D., 2007, 1997, 1986), Pfanzagl-Cardone (2023), and Xie, (2023).

## 1.7 Literature Review, Episode 3

**Episode 3** proposes a Hyper Near-Field Tiny Studio (HN-FTS) philosophy and benchmark to enable 3D music analysis and mixing. It relies on negligible indirect sound influence from closed field boundaries (walls), enabling a ‘low level and short reverb time decay at -60 dB (RT<sub>60</sub>)’ and results in a very [flat aggregate room response curve](#). The HN-FTS employs an intimate critical distance to the speaker in a small-scale studio environment to express 3D sound with a high degree of intelligibility, and impact, perhaps comparable to listening on a big set of open headphones with natural crosstalk.<sup>22</sup> Dr. Bosun Xie’s 2023 book entitled *Spatial Sound* is the most formidable scientific literature on the topic that I have found and he notes that anechoic conditions are advantageous for recorded audio and loudspeaker analysis.<sup>23</sup> His encyclopedic book provides over 1000 bibliographic references, presenting a lifetime of research on spatial sound. Xie observes,

*“During the past 40 years, especially since 1990, spatial sound has been developed greatly”  
“Current spatial sound techniques differ considerably from those in 1970 in many  
aspects of basic physical, auditory principles, and technical means.”*

(Xie, 2023, p.xvi)

Xie argues that spatial sound research is interdisciplinary and deals with “acoustics (physics), psychology, and physiology of hearing, electronics and signal processing, computers, and even the art of music” (Xie, 2023, p.xv). “Physical and auditory analysis of a sound field is

---

<sup>22</sup> The comparison of open headphones with natural crosstalk was expressed by acoustic engineers Edwin Seegeren and Terry Medwedick in public discussion at a Toronto AES, HN-FTS presentation at Toronto Metropolitan University in October 2023.

<sup>23</sup> Anechoic free field listening conditions are used for speaker design analysis, meanwhile, pop music mixers favour a dry control room with a low RT<sub>60</sub>. Richard Chychi mentioned this to me in conversations about Atmos mixing. Highly diffused spaces such as George Massenburg’s Blackbird studio are also effective music mix rooms.

the foundation of spatial sound” (ibid). He observes that, “special topics on spatial sound have been covered in some books, such as *Spatial Hearing* by Prof. Jens Blauert (1997), *Analytical Methods of Sound Field Synthesis* by Dr. Jens Ahrens (2012), *Ambisonics* by Dr. Franz Zotter and Mathias Frank (2019), *Sound Visualization and Manipulation* by Profs. Yang-Hann Kim and Jung-Woo Choi (2013), and *3D Sound for Virtual Reality and Multimedia* by Dr. Duran R. Begault (1994),” (Xie, 2023, p.xv).

Xie acknowledges that literature covering “relatively complete topics on spatial sound are rare” (Xie, 2023, p.xvi). He highlights Prof. Francis Rumsey’s 2001 book entitled *Spatial Audio* as a unique offering that well supports college and university courses in music technology, sound recording, multimedia, and their related fields. Specifically, regarding semantic concept maps for music spatiality, the work of Moylan and Rumsey is central to *Found in Space*.

### 1.7.1 Loudspeakers and Rooms

A wide-ranging body of literature and research was essential for the HN-FTS project. It is a sole operator control room that places speakers hyper near to the subject-listener which creates a hyper tuned sweet spot for critical music analysis, mixing and mastering. Research for this room began in March 2020, a time when there were approximately 30 Dolby Atmos studios in the world. By spring of 2022 there were still less than 200 Atmos music studios globally and in 2024 Dolby was tracking over 1200 Atmos music studios worldwide.<sup>24</sup> Prior to my research, a precise philosophy and benchmark for a small-scale Atmos studio did not exist. My exploration began with Dolby’s resources, where I found little information or encouragement for a hyper

---

<sup>24</sup> This statistic comes from Emma Brooks, Technical Solutions and Studio Enablement, Dolby Laboratories (Brooks, 2024).

near-field Atmos control room size. Like the crew members in *Lost in Space*, I persisted with blind faith and intuition hoping that a hyper near studio design for 3D spatial music could succeed. I was encouraged by Pascal Sijen’s guiding philosophy for Blue Sky studio monitors.

### 1.7.2 Blue Sky Speakers, PSB Speakers, and Floyd Toole

In 2008 Pascal Sijen published the design philosophy for Blue Sky studio monitors, speakers which were used at Skywalker Sound (Sijen, 2008). Serendipitously, I have worked with Blue Sky speakers for fifteen years and found them to be very reliable for transference of my 5.1, Surround Sound broadcast and music mixes. I was a believer in the Blue Sky philosophy and solution for full-range small studio/monitor interaction, which Sijen articulates below.

*“True full-range monitoring is the phrase which describes our goal. Not just full-range monitoring for those willing to spend huge amounts of money on large in-wall monitoring systems, but true full-range for all applications, from the desktop on up. The technologies we employed to achieve this goal are based on well understood principles of physics, are relatively simple to implement and deliver superior real-world results.”*

(Sijen, 2008, p.4)

Sijen’s philosophy led me to believe that an economical small Dolby Atmos control room capable of delivering premium results was possible. In many ways his philosophy is supported by Dolby’s Senior Field Applications Engineer Bryan Pennington who argues that “Atmos works with all speakers and price does not denote quality” (Dolby-D, 2020). This understanding gave me the confidence to select economical and timbrally neutral, Canadian made, PSB Alpha 3 loudspeakers as a modern upgrade to the Blue Sky speaker line, which has now become an inactive brand.

I reached out to PSB owner and designer Paul Barton and told him about my hyper near-field project; he was intrigued and became an advisor. During our conversations he afforded me advice about loudspeakers, rooms, binaural headphone playback, Atmos soundbars he was working on, and how the biological human hearing system works. He added significant real-world context to the literature that I was reading.

Paul explained that Dr. Floyd E. Toole was his mentor—I was already reading Dr. Toole’s book, *Sound Reproduction: The Acoustics and Psychoacoustics of Loudspeakers and Rooms* (Toole, 2018). Paul told me of their professional and personal history. Toole’s book is highly regarded by studio designers around the world, and it provided durable research about loudspeakers, listening position, and room behaviour. Other texts from Philip Newell, Alton Everest, Kenneth Pohlmann, the AES-E-Library, *Sound on Sound* magazine, and Dolby Professional Support contributed essential science, while hands-on knowledge was extracted from on-line video interviews with experienced Atmos practitioners. I synthesized all the research to form my philosophy, build the room and create a novel benchmark. It is an Atmos control room milestone with results that have surpassed expectations.

Dolby Labs (Emma Brooks) has reviewed aspects of my HN-FTS philosophy and benchmark which have been technically verified by the eminent Toronto, Los Angeles, New York City based electro-acoustician Edwin Segeren.



## Chapter 2: Examine Milestones in Multichannel Audio

---

### 2.1 Milestones in Multichannel Listener Experience

Dolby Atmos and popular music’s 3D spatial theory stand on historic theatrical milestones in multichannel audio that seek compelling listening experiences and enlivened emotions. The attraction of multichannel audio is that with each speaker the resolution of audio dimension increases. Multichannel speaker installations express a “range of available space” called “the spatial gamut” (Holman, 2008, p.1). First invented by Clément Ader in 1881, two-channel transduction using earphone speakers attracted audiences to the Paris Opera (*Scientific American*: 1881; Hertz: 1981; Rumsey, 2001). In France the format was named *Théatrephone*, and in the UK, *Electrophone*. *Théatrephone* used telephone lines that ran through sewer gutters in the streets of Paris to carry the audio pickup of two phones placed in front of an orchestra to a physically disassociated location for playback through two other phones with their playback speakers held against the ears. The listener experienced an interaural “in the head” stereo audio image of the music, staged in its naturally decorrelated opera hall environment.

In 1933 Alan Blumlein invented stereo binaural-sound and matrixing (Blumlein, 1958; Gerzon, 1992). The origin of multichannel (multitrack) surround sound goes back to 1941 when the Walt Disney film *Fantasia* was released in *Fantasound*. It involved nine record tracks, four optical playback tracks, and up to 80 speakers. The L, C, R program tracks relied on a control track to vary channel volume. Multitrack/panning & overdubbing were also invented in this era (Klapholz, 1991; Torick, 1998). In 1952 *Cinerama* involved seven-channel surround sound and in 1971 Quadraphonic sound was introduced. Pink Floyd’s legendary Quad recording entitled *The Dark Side of The Moon* was released in 1973. Soon after, in 1979 *Superman* and *Apocalypse*

Now introduced Dolby 5.1 Surround Sound to mainstream theatre goers (Holman, 2008; Dientsfry, 2016). Even though Dolby Atmos (2012-13) is nascent technology, already the Audio Engineering Society (AES) and Society of Motion Picture and Television Engineers (SMPTE) have standardized the 7.1.4 multichannel format<sup>25</sup> for archival of 3D audio masters.

## 2.2 Dolby Atmos in Mainstream Music, Media, Automobiles, and EDM

Extending from film, Dolby unveiled Atmos music sound on November 24<sup>th</sup>, 2019. It provides consumer 3D music playback via binaural headphones, home theatre, automobile, and multichannel studio control room speaker systems.

The format is quickly stimulating interest among veteran (stereo) professionals and pedagogues, and it is widely used in theatre and cinema, videogames, VR, TV, EDM dance clubs,<sup>26</sup> and music streaming playback. Atmos music creation requires significant economic investment and training along with a sustainable business plan.

For record labels and devout music enthusiasts the Atmos format seems to represent the future of recorded music consumption, production, and distribution. Atmos is backwards compatible with previous Dolby formats like Dolby Digital and Pro-Logic, and Atmos is endorsed by music labels Universal Music (UMG), Warner, Sony, and retailers Apple, Tidal, Amazon, and Netflix.

---

<sup>25</sup> 7.1.4 requires 12 speakers in total, on three planes. The inferior (lower) plane locates the Low Frequency Effects (LFE) speaker. The standard mid-plane is approximately 120 centimeters above the floor and involves 7 speakers: Left, Center, Right, Left Side, Right side, Left side rear surround, Right side rear surround. The superior (height) plane involves 4 speakers: Top front right, Top front left, Top rear right, and Top rear left.

<sup>26</sup> Dolby Atmos is used at the Ministry of Sound in London, UK. <https://www.ministryofsound.com/>

## Chapter 3: Choreographing Context, Expectation, Experience, and Emotion

---

### 3.1 Context, Expectation, Experience, and Emotion in 3D Music

David Byrne's theories about cultural and spatial context (Byrne, 2012, pp.13-30) add rich understandings that are "central to making sense, [and] to knowing" (Rice, 2018, p.1) about music's 3D environmental nature. In this regard, Byrne's research is in the domain of acoustemology. He provides foundational observations about creativity and context which help to rationalize expectations for crafting answers about the elements involved in the *what*, *when*, and *where* of audio-object orchestration, and the artful consideration of *why*.

Huron describes *what*, *when*, and *where* as "mental codes" (Huron, 2006, p.360). He uses the term loosely and assumes that "auditory expectations are manifested as distinctive brain states" (ibid). Brain-states involve *a priori* "cognitive image schemata," (Johnson, 1987, p.109); (Fauconnier, 1997, p.37); (Moore, Schmidt, Dockwray, 2011, p.84) that empower the psychological "flow state" (Csikszentmihalyi, 2004).

Huron argues, "for example, expected sounds might be represented by distributed action potentials that are only implicitly coded" (Huron, 2006, p.360). I relate his example directly to 3D spatial sound and interpret the meaning of "distributed action potentials that are implicitly encoded" to be understood as metaphoric audio-objects, coded to be genuinely believable by the listener so as to meet their expectations of pre-formed (*a priori*) 3D cognitive image schemata and expectation.

I contemplate David Byrne's theories about cultural and spatial context to conceive of the positional relationships, cognitive perspectives, proportional properties and character attributes

of audio-objects which can then be deployed while considering Huron's ITPRA theory. Pink Floyd, *The Dark Side of The Moon* "On The Run" exemplifies how context and ITPRA work together to express song meaning and story tell. I argue that these concepts can help the creator to understand and activate listener expectations, environmental contexts, subjective meanings and resultant emotions.

Quoted by Huron, musicologist Leonard Meyer's seminal 1956 book entitled *Emotion and Meaning in Music* argues that "the principle emotional content of music arises through the composer's choreographing of expectation" (Huron, 2006, p.11). Meyer did not use those exact words but his argument reinforces David Huron's [ITPRA theory](#), arguing that emotions created by expectations elicit responses involving "five functionally distinct neurophysiological systems: [I]magination, [T]ension, [P]rediction, [R]eaction, and [A]ppraisal" (Huron, 2006, p.25). Each of these responses "involve both physiological and psychological changes" (ibid, p.15). He applies the theory to music and aims to "better understand how listeners form music-related expectations, and how these might account for various emotional responses" (ibid, p.5). "Accurate expectations are adaptive mental functions that allow organisms to prepare for appropriate action and perceptions" (ibid, p.12).

Recorded music expectations are formed by the performance, songwriter, composer, arranger, orchestrator, and recording/mixing engineer as events in time, virtual<sup>27</sup> spaces, and places. The 3D Dolby Atmos format affords creators with the ability to transform the sound of

---

<sup>27</sup> "[O]ur creative thoughts and imagination (i.e., the virtual) can be either transformed or *nearly* transformed into reality and actuality through digital means" (Rambaran, 2021, p.1).

real social and cultural “spaces and places” (Whitely, Bennett, Hawkins, 2004)<sup>28, 29</sup> into virtual sonic environments that precipitate palpable listening experiences and evoke instinctive visceral emotions. By applying ITPRA responses, Huron offers “psychologically based insights into such venerable topics as meter, syncopation, cadence, tonality, atonality, and form” (Huron, 2006, p.5). He argues that the shaping of compelling emotional content, perceived as experiences in recorded music involves “...codes of practice, heuristic rules of thumb, and speculative theories [that] have been passed from teacher to student across the generations” (ibid, p.10). Across these generations the element of temporality<sup>30</sup> has remained constant, but now in mainstream recorded popular music released in the Dolby Atmos format, the element of two-channel stereo spatiality has evolved into a locationally dynamic 4D temporal sound field with highly detailed spatial resolution.

Atmos uses a multichannel 7.1.2 channel bed with an additional 128 unfixed audio-objects for accurate playback location when rendered by decoders in individual playback systems. Music’s 3D spatial environment portrays audio-object location in relation to the listener’s full body anatomy [medially, anteriorly, vertically, laterally, posteriorly, superiorly, and inferiorly](#), via the three dimensions of width, depth, and height ([X,Y, Z](#)), and the 4<sup>th</sup> dimension of time. Induction of the phenomena involves complex human auditory physiology that localizes

---

<sup>28</sup> “As an increasing number of popular music researchers now acknowledge, the search for social and cultural meanings in popular music texts inevitably involves an examination of the urban and rural spaces in which music is experienced on a day-to-day basis” (Whitely, Bennett, Hawkins, 2004, p.2)

<sup>29</sup> Dolby Atmos can express 3D spaces in countless places and spaces via binaural headphones and speaker playback systems, which affect the listener’s interpretation of meaning in relation to their locational context.

<sup>30</sup>When analysing recorded music using “audio scene analysis” (Bregman, 1993) and studying scores, temporality suspends to create a static scene-based paradigm, necessary for textualization of the music and its sonic constructs. Additionally, Stan Hawkins refers to ‘temporal-specific listening’—as recorded music “is foreclosed by temporality; it’s sense of being in the here and now can indeed propel into the then and there. Yet, it can also take us back in time” (Hawkins, 2016, p.4).

distinct sound sources<sup>31</sup>. They form the listener's perceptions of musical experiences that alter their mind state and trigger evolutionary behaviours that bring forth expectations and emotional responses.

*“From an evolutionary perspective, the capacity to form accurate expectations about future events confers significant biological advantages. Those who can predict the future are better prepared to take advantage of opportunities and sidestep dangers. Over the past 500 million years or so, natural selection has favo[u]red the development of perceptual cognitive systems that help organisms to anticipate future events. Like other animals, humans come equipped with a variety of mental capacities that help us form expectations about what is likely to happen.”*

(Huron, 2006. p.12)

### 3.2 Practitioners of Emotion

David Huron asks, “How do emotions arise?” What purposes do they serve? What accounts for the distinctive feelings we experience?” He sees storytellers as “practitioners of emotion” (Huron, 2006, p.10).

*“Playwrights, novelists, poets, film directors, musicians, choreographers, comedians, and theatrical magicians all have a professional interest in what distinguishes delight from boredom. Therapists, game designers, carnival operators and traffic engineers have good reasons to try to understand what causes people to be surprised or fearful. Even advertisers and politicians have practical motivations for understanding how the flux of events shape human emotional experiences.”*

(ibid)

---

<sup>31</sup> See Wenzel, Begault, Godfroy-Cooper, (2018, pp.5-39) Chapter 1, “Perception of Spatial Sound” for a detailed explanation of the human hearing system and perception.

Huron views emotions as “motivational amplifiers” (ibid) because they materialize from experiences. Whether utilizing an intuitive awareness of psychological expectation, or an understanding of David Huron’s ITPRA theory during creation, eminent 3D music creators routinely ask *what, when, where, why, and how* do audio-objects enliven emotional outcomes.

### 3.3 Geneva Emotional Music Scale (GEMS)

Marcel Zenter, Didier Grandjean, and Klaus R. Scherer conducted four interrelated studies of music-induced emotions that underpin the Geneva Emotional Music Scale (GEMS.) Their work resulted in a 9-factorial model. In his 2021 TEDx talk Zentner asserts that classical music, which often lasts longer than a 3:00 minute pop song, can take you on an extended emotional journey. Notwithstanding, pop songs can also provide compelling journeys. GEMS broadly categorizes nine spheres of music induced emotion: Please see **Table 2:** Geneva Emotional Music Scale (GEMS).

|   |                                                |
|---|------------------------------------------------|
| 1 | Wonder/amazed/moved                            |
| 2 | Transcendence/overwhelmed/awed/ [3D immersion] |
| 3 | Nostalgia/longing/dreamy/melancholic           |
| 4 | Tenderness/romantic/sensual                    |
| 5 | Tranquility/peacefulness/serene/meditative     |
| 6 | Joy/bouncy                                     |
| 7 | Power/strength/triumphant                      |
| 8 | Tension/agitated/edgy/suspense                 |
| 9 | Sadness/blue/sorrowful                         |

**Table 2:** Geneva Emotional Music Scale (GEMS).

Zenter argues that the GEMS system enables organization of music into scientifically researched categorizations of emotion, evoked by subjective listener music experience. The

GEMS Scale is an emotional concept map that could be integrated into 3D popular music spatial theory as it relates to context and ITPRA. That idea is beyond the scope of this dissertation; however, it could be the intention for a future episode of *Found in Space*.

### 3.4 Composing with Spatial Context

---

#### 3.4.1 Spaces and Places

During the stages of music composition and pre-production, theoretical conceptions of spatial context often formulate in the creator's mind. They intend to set music's holistic sound which can imagine the acoustic space or recorded place of the music. *The Oxford English Dictionary* describes the word 'context' to mean "The whole structure of a connected passage regarded in its bearing upon any of the parts which constitute it" (Oxford, 2024). That description can apply to music's spatial context, understood as phenomenological sonic "circumstances that form the setting for an event, statement, or idea, and in terms of which it can be fully understood and assessed" (*New Oxford American Dictionary*, 2018).

Music's spatial context is a powerful determinant that subconsciously engages the listener. It is built upon pre-formed cognitive schemata about the sound of spaces and places. David Byrne argues, "context largely determines what is written, painted, sculpted, sung, or performed" (Byrne, 2012, p.13).

Byrne exemplifies the significance of spatial context arguing that orchestral music sounds best indoors, and percussive groove music carries well outdoors. He argues that African hand drums were "carefully fashioned, selected, tailored, and played to best suit the physical, acoustic, and social situation" (ibid). "The music perfectly fits the place where it is heard, sonically, and structurally" (ibid). "Africans, whose spiritual music is often rhythmically complex, may not



associate the music that originates in these spaces [stone-walled gothic cathedrals ... monasteries and cloisters]<sup>32</sup> with spirituality; they may simply hear it as being blurry and indistinct” (ibid).

The interactions between music and acoustics are the “acoustic ecology” of music. The domains of “soundscape” and “acoustic ecology” theorise human encounter with sound. R. Murray Schafer studies this in his 1977 book entitled *The Tuning of the World* (Schafer, 1977). Soundscape and acoustic ecology ([see chapter 3.5.1](#)) often inform popular music and its spatial context. A good example is Pink Floyd, *The Dark Side of The Moon*, “On The Run.” This musical soundscape utilizes 3D spatial context to tell a story about the fear of flying and death while running through an airport terminal. The 3D spatial context in “On The Run” is compelling, and it amplifies the story because it functions as a mimetic sonic metaphor involving allocentric object-to-object, and egocentric subject-to-object perspectives. The 3D sound field expresses auditory codes of space and place, heard as human experiences in society, culture, and political economy. See [Analysis 5](#): “On The Run” by Pink Floyd, *The Dark Side of The Moon*.

### 3.4.2 The Political Economy of Egalitarian and Elite Spaces, and Places

Quoted by Byrne, in Alan Lomax’s book entitled *Folk Song Style and Culture*, Lomax argues that African hand drumming music “emanates from and mirrors egalitarian societies” (Byrne, 2012, p.16). In my opinion Byrne’s take on Lomax requires some scrutiny. For example, on page 161 Lomax speaks of “African villages linked by decent ties and common bonds” and the “African emphasis on overlap, with its egalitarian structure.” Lomax admits that he avoids investigation of “the relationship between egalitarian Pygmy bands of the Congo and the

---

<sup>32</sup> “The reverberation time in those spaces is very long—more than four seconds in most cases—so a note sung a few seconds ago hangs in the air and becomes part of the present sonic landscape” (Byrne, 2012, p.16)

setting[s] in which Bach and his fellow composers worked” (Lomax, 1968, p.161). To me, the word “settings” means context that involves cultural and auditorial components. These contexts carry the sonic codes of societal political economy and class structure within the sound of their music performance spaces and places.

Egalitarian societies are based on human equal rights and Simon Behrman’s on-line article entitled *The Political Economy of Classical Music*, argues “the story of classical music is inseparable from the rise of capitalism.” “Capitalism first created the space in which such music could flourish, and then took it away, leaving behind a frozen, formalized tradition.” (Behrman, 2022). Wealthy elite people and corporate entities built their privileged-class spaces for music performance, commercialised the concert experience, and created a class structure based on payment for accessibility to music. The social elite class often associate with grand spatial contexts heard in sanctuaries, monasteries, stone-walled foyers, glass-walled atriums, and wood and cloth clad concert halls.

The age of commodified indoor music experiences began in concert halls and sanctuaries that featured enveloping signatures of acoustic spatial context which has now resulted in celebrated European concert halls, such as Municipal House, Smetana Hall, in Prague, located on Náměstí Republiky, next to the powder gate. The style of architecture is ornate art nouveau and the exterior of the building features allegorical art and stucco.

Behrman recounts, “The rise of the middle classes, and the expansion of leisure time available to them and to the upper middle classes of their societies, conditioned the emergence of classical music” (Behrman, 2022), which was composed for these spaces. “By the late eighteenth and nineteenth centuries, many great opera houses and concert halls were being built, and [symphonic] orchestras were created with the backing of commercial enterprises” ... “For

example, the first major concert hall in London was Hanover Square Rooms, which Italian entrepreneur Sir John Galliani built and ran with composers Johann Christian Bach (son of Johann Sebastian) and Carl Friedrich Abel” (ibid).

In 1958, Milton Babbitt published an article for High Fidelity magazine entitled “Who cares if you listen?” (Babbitt, 1958). His article argues, “in order for a composer to maintain their artistic integrity in the face of populist culture, he insisted, they must withdraw from the commercial world ...” (ibid). Arguably this is a purist stance and would disassociate an entire canon of symphonic music from its intended spatial context.

In her 2013 doctoral dissertation, Sarah May Robinson investigates “chamber music performances outside of traditional concert venues in the U.S. in the 21st century” (Robinson, 2013). She “traces the use of non-traditional venues throughout history from Bach’s coffee house concerts to the gallery and loft concerts, which first emerged in the 1950s and 60s” (ibid). Some of the topics she addresses are venue noise levels, amplification and acoustics, audience engagement, and seating arrangements. She argues, “alternative venue atmosphere is making its way into the concert hall” ... “Ensembles that develop in alternative venues tend to keep [their] casual presentation style when they move into more formal spaces” (Robinson, 2013, p.109). Creators of 3D spatial music can optimize the listeners experience of environment, space and place by thinking carefully about the cultural codes contained in their choices for 3D spatial contexts and reverbs.

### 3.4.3 The Political Economy of Multichannel Virtual Spatiality

Techoustemological elite concert hall spatial context has crossed over to the domain of popular music production. Hardware and software can be designed, constructed, and configured

to recreate indirect sound (reverb) that models the acoustic properties and attributes of socially elite concert halls or egalitarian public places. Producers routinely describe the lush multichannel decorrelation of digital reverberators as sounding “expensive.”

The Bricasti Model 7 Stereo Digital Reverb Processor sounds expensive and is expensive. It comes with presets for halls, chambers, and ambient spaces that are often named after the physical spaces and places they seek to emulate—Chicago Hall, Vienna Hall, and Berliner Hall are a few examples. The reverberation presets are synthesized in real-time using impulse responses<sup>33</sup> taken from a variety of elite and egalitarian spaces and places. They provide music creators with virtual emulations of environmental indirect sound and when multiple instances are spatiotemporally cascaded during the Atmos music mix, highly convincing 3D spatial contexts can be created. These virtual spaces recreate the acoustic ecology of where music happens in local scenes, communities, societies, and cultures. This is why spatial context is essential in popular recorded music—it subconsciously directs the listener to locate themselves within their subjective music listening experience.

The spatial context of formal (Elite), alternate (Egalitarian), and abstracted spaces and places can be expressed vividly in the Dolby Atmos format, but Atmos mix engineers and recordists must understand that within their sound field, transportation of music performance to an unfavourably small or large spatial context can create compositional conflict and acoustical problems. Reverberant smearing or a loud noise floor<sup>34</sup> that masks quiet dynamics can challenge a composer’s expectations of their spatial context.

---

<sup>33</sup> Impulse responses are actual reverberations, recorded in spaces and places that are convolved with direct/dry sounds and artificially synthesized reverberations to form spatial contexts, heard as reflected (indirect) sound.

<sup>34</sup> Noise floor can increase to a problematic level with multiple overdubs of musicians in the same space.

Byrne argues, “It’s usually assumed that much of Western medieval music was harmonically ‘simple’ (having few key changes) because composers hadn’t yet evolved the use of complex harmonies” ... “in this context [stone-walled gothic cathedrals ... monasteries and cloisters] there would be no need or desire to include complex harmonies, as they would have sounded horrible in such spaces” (Byrne, 2012, p.17). My interpretation of Byrne’s argument is that 3D composers can create harmonic density specifically acclimatized to their imagined space and place. This compositional tactic is the lodestar of an unconflicted sound field involving compelling spatiality with immersive weight. Critically balanced spatial context determines the holistic compositional character of the Overall Listening Experience (OLE.)

#### 3.4.4 Spatial Recontextualization

Pierre Schaeffer (1910-1995) is the founder of *Musique Concrete*. He was a theorist of music listening, trained as a radio engineer and announcer who coined the terms “reduced listening” or “acousmatic music” (Cox, Warner, 2004, p.76). Acousmatic music is “a noise that one hears without seeing what causes it” (ibid). This is a significant form of music de-contextualization. “[M]uch of what we thought was heard was in reality only seen, and explicated, through the context” (ibid). Acousmatic music is the heartbeat of hip-hop music sampling and loop creation, a practice that de-contextualizes recorded music, separating it from its original era, genre, social scene, and culture, for re-contextualization into a new form. Spatial context contained within acousmatic samples and loops carry music meanings that take on new codes of meaning within a re-contextualized work. In 3D music, direct sound audio-objects are effectively heard as acousmatic sound sources (audio-objects) that can inherit spatial contextualization, or re-contextualization to be heard in a new form with new meanings.

Changing the spatial context of music is significant for the listener's sense of location in a space or place.

*“The defining characteristics of an environmental context allow us to discriminate between the same source heard in different size enclosures or out-of doors.”*

(Wenzal, Begault, Godfroy-Cooper 2018, p.30)

### 3.4.5 Spatiotemporal Orchestration (of Space and Time)

Born across centuries, Western theories for musicology,<sup>35</sup> music context,<sup>36</sup> listening,<sup>37</sup> and analysis<sup>38</sup> of composition, arranging, and orchestration techniques are well developed, but now, 3D spatiotemporal orchestration expands the traditional notions of orchestration to create and recreate environments as spaces and places. It involves audio-objects, perhaps heard as musicians that can be located on or off the traditional concert hall stage and located anywhere in the 360-degree theatre environment (sound-box) to surround the listener. Please see [Chapter 6.8](#).

#### **Figure 8.**

In his book entitled *Surround Sound Up and Running*, Thomlinson Holman shares history for this art and craft. He reminds us that “the use of spatial separation as a feature of

---

<sup>35</sup> “In an article entitled ‘Umfang, Method und Ziel der Musikwissenschaft’ (‘The Scope, Method and aim of Musicology’), published in 1885 in the first issue of *Vierteljahrsschrift für Musikwissenschaft* (Adler 1885; see also Bujic’ 1988, pp.348-55), Adler outlined a separation between the historical and systematic dimensions of music, with rigour of the exercise reflected in the term *Musikwissenschaft* (science of music). (Beard, Gloag, 2016, p.xi)

<sup>36</sup> “For Adorno the relationship between the musical work and its context cannot be reduced to a direct parallel between the two domains; rather, the musical content, its material, bears the immanent trace of the antagonisms and problematics of the world.” (Beard, Gloag, 2016, p.25)

<sup>37</sup> “[A]ural ethics” emerge in the space and encounter between music and listener. This strategy requires the exercise of an ‘attentive listening’ that is alert to music’s inner details, its contexts, and multiple meanings” (Cobussen and Nielsen, 2012; Beard, Gloag, 2016, p.93). Also see *Auditory Technologies and the Regression of Listening*, Culture Industry, iPod, iTunes, Soundtrack, Silent Film, (Steele, 2008).

<sup>38</sup> “To go beyond mechanical conversations of musical notation into written words, analysis must uncover something beyond or behind the mere sonic surface. (Parker and Abbate, 1989, pp.1-2);” (Beard, Gloag, 2016, p.14) “[A]nalysis is an interaction between the music itself, music theory, aesthetics and history” (Beard, Gloag, 2016, p.14).

composition, probably started with the call and response form of antiphonal music in the medieval period, an outgrowth of Gregorian chant” (Holman, 2008, p.2). In about 1550 Flemish composer Adrian Willaert staged a chorus on the left and right sides for antiphonal singing. In about 1585, Giovanni Gabrieli, principal organist at St. Marks Basilica precisely directed musicians to locate in various proximities to the listening audience.

In 1830 *Symphonie Fantastique* by Berlioz instructed the oboist to be placed offstage as a way to express a distant sound.<sup>39</sup> In Symphony No.2 *Resurrection*, Gustav Mahler instructs 4 trumpets, 4 to 6 horns, bass drum with cymbals attached, triangle, and timpani to locate off the stage, usually in the balcony.

Gardener Read’s book entitled *Thesaurus of Orchestral Devices* is dedicated to Arthur Cohn and Nicolas Slonimsky, who he describes as “peripatetic encyclopedists and indefatigable begetters of pandects, lexicons, and thesauri” (Read, 1953, p.vi). American composer Aaron Copeland regards Read’s book as “a rare thing in musical scholarship—a “first time anywhere” (ibid, p.vii). Read explains the art and craft of orchestration as “composing for instruments.” (ibid, p.ix). Read says orchestration thinks first about,

*“colo[u]r, balance, and sonority—inventing a melodic line that finds expression only, for example, in the English horn, and in no other instrument; writing a sonority that can be achieved by no other combination than muted brass; or weaving an insistent rhythmic ostinato that reaches full expressiveness only to be muffled by the tenor drum.”*

(ibid)

---

<sup>39</sup> The components of “distant sound” involve favouring the energy of indirect sound in relation to its direct sound. Distant sounds have diminished high frequencies and their anterior (frontal) apparent source width narrows, while lateral (side) listener envelopment increases.

Traditional orchestration inherently involves an allocentric object-to-object (instrument-to-instrument) perspective. It is an exercise occurring between instruments and instrument groups, but spatiotemporal orchestration considers the egocentric listener-to-object perspective.

Read’s book is a 630-page tome of orchestration examples and was first published in the era of Walt Disney’s acoustically immersive *Fantasound* (1941) and the later format of *Cinerama* (1952). For *Fantasia*, Disney wanted conductor Leopold Stokowski to orchestrate the “Flight of the Bumblebee” melody to move around the heads of audience members seated in the theatre. In retrospect, this era of musicological and technological interdiscipline presented the first cinematic example of spatiotemporal offstage orchestration, but without the dimension of height. Now in the Dolby Atmos format, spatiotemporal offstage orchestration involving height is possible. Gardener Read’s book was first published in 1953, and Chapter 64 provides two examples of off-stage (offstage is my preference) orchestration and Chapter 11 provides off-stage terminology in English, Italian, French, and German, as it was used to express spatiotemporal location of instruments on a music score. Please see **Table 3: Multilingual Music Scoring Terms for off-stage Location of Instrumentalists**.

| English                  | Italian                     | French            | German                       |
|--------------------------|-----------------------------|-------------------|------------------------------|
| Behind the scene [stage] | A colotta                   | Dans la coulisse  | Auf dem Theatre              |
| In the distance          | Dietro la scena             | Derrière la scène | Auf der Bühne                |
| Offstage                 | In distanza                 | Lointain          | Aus der Ferne                |
|                          | In lontananza ed invisibili |                   | Hinter dem Orchester         |
|                          | Interna                     |                   | Hinter der Bühne<br>[scene]  |
|                          | Lontano                     |                   | In der Entfernung<br>[Ferne] |



|  |                        |  |                              |
|--|------------------------|--|------------------------------|
|  | Nella scena<br>lontano |  | In Entfernung<br>aufgestellt |
|  |                        |  | In weiter Entfernung         |
|  |                        |  | Isoliert postiert            |

**Table 3:** Multilingual music scoring terms for off-stage instrumentalists (Read, 1953, p.75).

See **Table 4** for Read’s selected examples of offstage score instruction by instrument.

| <b>Instrument</b>   | <b>Composer</b> | <b>Work</b>                                              | <b>Page</b> | <b>Measure</b> |
|---------------------|-----------------|----------------------------------------------------------|-------------|----------------|
| <u>Flute</u>        | Skilton         | <i>Suite Primeval, Part II</i>                           | 3           | 1              |
| <u>Oboe</u>         | Berlioz         | <i>Symphonie Fantastique</i>                             | 100         | 3              |
|                     | Prokofieff      | <i>Alexander Nevsky</i>                                  | 6           | [1] 1f         |
| <u>English Horn</u> | Prokofieff      | <i>Alexander Nevsky</i>                                  | 53          | 4b [34]        |
|                     | Wagner, R.,     | <i>Introduction to Act III,<br/>“Tristan und Isolde”</i> | 6           | [2] 10f        |

**Table 4:** Offstage orchestration examples organized by instrument.

Orchestrated offstaging is not new, but now with Dolby Atmos it’s possible to create hyper-real audio-object rotation and azimuthal tilt between the ear level and the top height speakers to enliven music’s spatiotemporal anticipation in exciting new ways. These forms of spatiotemporal orchestration can be organized in three levels of abstraction: “low, mid, and high” (Antoniuk, Zieliński, Lee, 2024, p.2).

### 3.4.6 Notation of Spatiotemporal Orchestration

Standing on the shoulders of about 140 years of multichannel audio history that has expanded the gamut of speaker system spheres, the advent of 3D spatiotemporal orchestration

requires orchestrators<sup>40</sup> to explore the cognitive perspectives of allocentric (instrument-to-instrument), and egocentric (listener to instrument/ensemble). These perspectives can be represented in fixed channels, or as dynamic audio-objects within the four dimensions of width, height, depth, and temporality. The time-honoured topics of music composition, harmony, timbre, percussion, syncopation, meter, dynamics, ensemble staging, arranging, orchestration, and now, spatiotemporal orchestration,<sup>41</sup> must all be in consonance to achieve compelling impact, intensity, and intelligibility within music's 3D sound world. This acoustically ecological sound field contains the sonic imprint of era, architecture, nature, culture, and all other contextual aspects of social gathering in spaces and places. Inappropriate spatial context experienced as spatial under- or over-envelopment creates an unflattering listening experience. It impedes music's communication, emotion and enjoyment.

Byrne's profundities about context and creation can be distilled into a maxim that engenders a wide range of contextual music production considerations. Rationalizing how *context determines creativity* effectively guides 3D music "practitioners of emotion" (Huron, 2006, p.1) to orchestrate music's five spatiotemporal queries: 1) *what*; 2) *when*; 3) *where*; 4) *why*, and also the technical crafting of *how*.

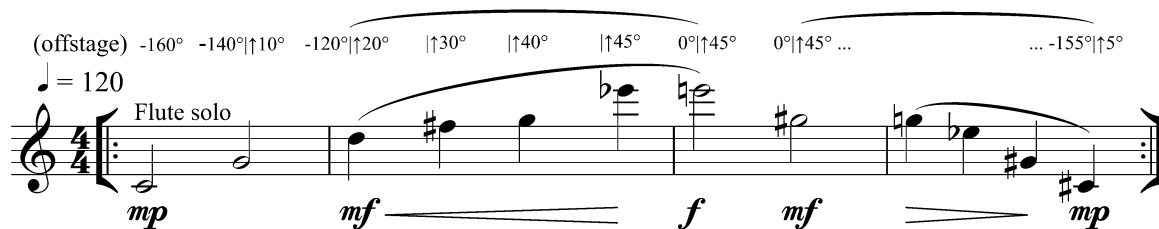
A practical way to notate spatiotemporal orchestration on a traditional music score can utilize the cartesian coordinate system of azimuth, elevation and distance. This idea was

---

<sup>40</sup> Gardener Read refers to orchestrators as "instrumentators" (Read, 1953, p.ix).

<sup>41</sup> Coined by this author, the musicological term 'spatiotemporal orchestration' reflects the traditional understanding of a 3D coordinate system but adds time as the fourth dimension. Edwin Pfanzagl-Cardone describes the need for a temporal dimension descriptor in 3D audio and suggests 4D, but he laments, "this change is not likely to happen." (Pfanzagl-Cardone, 2023, p.4) *Found in Space* maintains the term 3D, while acknowledging that 4D is a more accurate descriptor of cartesian X, Y, Z coordinates in combination with the temporal dimension.

described to me by the eminent (Toronto born) Atmos practitioner Justin Gray. The fourth dimension of time is communicated accurately with tempo and time signature. Mathematical coordinate notation uses arrows with numerical coordinates to indicate precise locations of unfixed audio-objects<sup>42</sup> in the 360-degree speaker sphere. See **Figure 3: Spatial Orchestration notation using coordinates**. See **Figure 4: the 360° coordinates: azimuth, elevation, and distance**.

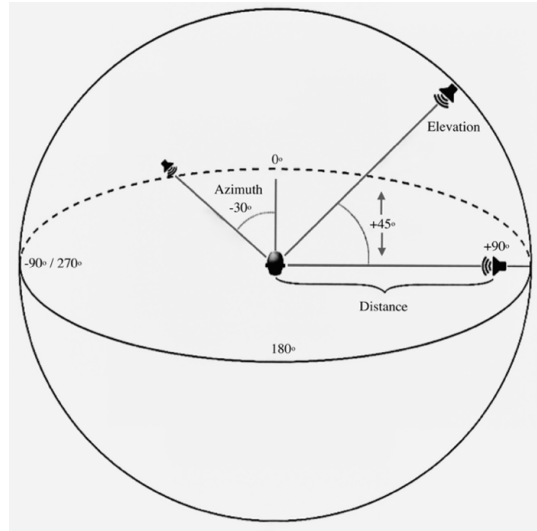


**Figure 3: Spatiotemporal orchestration notation using numerical coordinates.**

*“[A] listener facing directly forward is defined as 0° azimuth and 0° elevation.”*  
*“Azimuth is defined by the angle ( $\theta$ ) between the source location and the median plane at 0° azimuth (projected on to the horizontal plane) and elevation is the angle ( $\delta$ ) between the source location and the horizontal plane at 0° elevation (projected onto the median plane).”* *“Azimuths to the right of the listener are positive, to the left are negative, and the rear is defined as 180°. Elevations are positive for upper directions and negative for lower directions relative to the listener. Distance is defined as the radius ( $r$ ) projected along the vector formed by the azimuth and elevation of the source.”*

(Wenzel, Begault, Godfroy-Cooper, 2018, p.10)

<sup>42</sup> Unfixed audio-objects are commonly referred to as “dynamic objects.” In this case the word “dynamic” should not be conflated with the traditional understanding of gestural volume dynamics in music.



**Figure 4:** 360° coordinates of azimuth, elevation, and distance.

Wenzel, Durand, and Godfroy-Cooper note that a clockwise 360° spherical coordinate system can also be used, indicating 0° in front, 90° to the right side, 180° to the rear, and 270° to the left side. This notational system of spatial orchestration affords modern composers an efficient way to convey performer staging location via the traditional music score.

When recording 3D spatial music natively in Dolby Atmos, with microphones placed where the Atmos format designates speakers to be located, the notated coordinate system can indicate microphone placement in relation to the performer and speaker playback system.

The perceived proximity or auditorial depth of field (DOF) in relation to the performer can be textually expressed on the score using anthropologist Edward T. Hall's theory of proxemics. Hall undertook a lifetime of research about cultural perceptions of space<sup>43</sup> and theorized intimate, personal, social, and public as the four primary zones that humans experience in spaces (Hall, 1963). 3D music recordists and mixers can benefit from Hall's proxemic theory

---

<sup>43</sup> Space refers to indoor spaces or outdoor environments, not outer space.

because they indicate a critical balance between direct sound and indirect (reflected) sound.

Hall's theory presents an easy way to conceive and textualize proxemic location involving Source Width, Apparent Source Width, Ensemble Width, Apparent Ensemble Width, Listener Envelopment (the five Super Components) and the perception of depth of field.

Intimate (close) audio-objects would have a wide Source Width (SW) measurement, and public (far) audio-objects would have a narrow SW. Public audio-object Apparent Source Width (ASW)—generated by indirect reflected sound—would be wide. In a public proximity, due to logarithmic diminishment of direct sound energy, the reflected indirect sound energy would be predominant within the critical balance of direct and indirect sound energy.

Spatial orchestration can be communicated in other ways too. The coordinate system is mathematically accurate and supports dynamic audio-objects with the use of slurs in the formats of Atmos, Auro 3D, Sony-360, and Ambisonics.<sup>44</sup> Proxemics can be communicated textually using Edward Hall's theory, and fixed object location can also be communicated using Dolby's speaker nomenclature: L, C, R, Lw, Rw, Ls, Rs, Lss, Rss, Ltf, Rtf, Ltm, Rtm, Ltr, and Rtr. Any combination of these systems can be employed on a traditional music score when notating the art and craft of spatiotemporal orchestration for a free- or closed field music experience.

---

<sup>44</sup> Many other 3D audio formats do exist. They are not my focus here, but this 3D notation system will work in those formats as well. Ambisonics “was probably the first attempt at capturing 3D audio [...] in the 1970s” (Pfanzagl-Cardone, 2023)

## 3.5 Free- and Closed field Sonic Experiences

---

### 3.5.1 Acoustic Ecology

Free field spatial context occurs out-of doors and is contrasted by closed field indoor spatial context. Closed field spatial context is the norm in music studios and popular recorded music. R. Murray Schafer (1933-2021) is a Canadian composer, born in Sarnia, Ontario and lived much of his life in Bancroft, Ontario. He founded the “Acoustic Ecology” movement and coined the term “Soundscape” when he formed the World Soundscape Project. Similarly, to David Huron’s phrase “choreographer of emotion,” Schafer argues, “The musician is an architect of sounds” (Cox, Warner, 2004, p.30). *Acoustemology* is a closely related term that stems from the combination of acoustics and sonic ecology.

### 3.5.2 Acoustemology

Steven Feld is an anthropologist and ethnomusicologist who extended Schaefer’s work in acoustic ecology to conceive of the term acoustemology. The word is a portmanteau that combines “‘acoustic’ and ‘epistemology’ to foreground sonic experience as a way of knowing” (Rice, 2018, p.1).

*“[A]coustemology emphasizes the importance of sound in human experience, particularly in relation to place and notions of emplacement”*

(Rice, 2018, p.3)

*“In 1976, Feld went to live with the Kaluli, a small group living in the forests of the Great Papuan Plateau in the Southern Highlands province of Papua, New Guinea” ...*

*“Feld observed that the Kaluli had an acute acoustic sensitivity to the rainforest environment because the density of vegetation in the forest meant that much of the Kaluli surroundings were visually hidden, sounds were relied upon as primary indicators of presence. At the same time, the rainforest was sonically rich, [with] the*

*sounds of insects, birds and other wildlife, as well as those of running and falling water combining to create a particular density of sound. Shifting sonic presences and dynamics could communicate a great deal to knowledgeable Kaluli listeners. As Feld writes, “in Bosavi, ‘sounds are heard as time of day, season of year, vegetation cycles, migratory patterns, forest heights and depths.’”*

(Ibid, p.1)

### 3.5.3 Techoustemology

In 2004 Thomas Porcello introduced “a complimentary term to acoustemology. “Techoustemology” (Feld, 2018, p.5) is rooted in the social construction of technology and technoculture and “demands an accountability not only to the specifics of how technology is deployed but to a serious consideration of the resultant *sounds* of sociotechnical action, and how individuals and groups conceptualize, rationalize, and discursively render their own techoustemological, ideologies and practises” (Porcello, 2004, p.271). Porcello argues that “technological mediation almost invariably shapes our expectations of how sounds should sound and is rarely absent from what he calls ‘the sonic signal chain’ in all parts of the modern world” (Porcello, 2004, p.270; Feld, 2018, p.5). “Acoustemologies can seldom be extricated from the processes of technological mediation that make them possible” ... “acoustemology emerges at the interface of environments, materiality[’s], technologies, and nonhuman forms of life” (Feld, 2018, p.4). For the Kalui, “place is heard and felt synesthetically, kinaesthetically, and affectively” (ibid, p.3).

I argue that Acoustemology and Techoustemology provide the most appropriate silos to house the semantic metaphysical componentry that I find to be substantial in 3D music spatiality.

### 3.5.4 Perceiving Clairaudience

Murray Schafer argues for the faculty of “clairaudience,” perceiving, as if hearing, what is inaudible (Schafer, 2004, p.29). Large Atmos installations do involve “clairaudience.” “In the West the ear has given way to the eye as the most important gatherer of environmental information” [but] “When man [woman, and mammals] was [were] fearful of the dangers in an unexplored environment, the whole body was an ear” (ibid, p.31).

As a primal defence mechanism in mammals<sup>45</sup>, somatic induction (involving the full-body somatosensory system) of sonic environmental context senses low frequency sound below about 20 Hz. Called infrasound, it is felt and not heard.<sup>46</sup> Infrasound can be expressed in large full frequency range Atmos theatres and dance clubs with high-power multiple sub-woofers and speaker arrays. This is a form of clairaudience, which is described as, "the supposed faculty of perceiving, as if by hearing, what is inaudible" (*New Oxford American Dictionary, 2018*). An optimal Atmos listening experience, at 85 [dB](#) SPL playback can involve clairaudience, induced via body vibration from infrasound.

### 3.5.5 Perceiving Infrasound

Canadian born Dr. Floyd Toole was the corporate Vice President at “Harman International” industries from 1991 to 2007. He directed the Harman Research and Development Group. In 2009 Harman audio engineers studied whole body vibration for automotive sound

---

<sup>45</sup> “The dynamic range of normal hearing in humans extends over 120 dB from detection of very soft sounds to the tolerance limit for loud sounds; the span of audible frequencies in humans is 20Hz to 20kHz. In contrast, other mammals, such as bats and dolphins, can hear sounds with ultrasonic frequency (greater than 20kHz), whereas whales and elephants can hear sounds with infrasonic or subsonic frequency (lower than 20Hz) (Shapira, Pujol, Plaksin, Kimmel, 2016)

<sup>46</sup> Alligators can communicate with infrasonic bellows that result in “water dances.” (Fairchild, 2020-2024)



systems and concluded that the “presence of whole [human] body-body vibration” could “reduce the preferred level of bass equalisation by as much as 3 dB” (Simon, Olive, Welti, 2009, p.1). ‘Preferred bass level’ equates to LFE/subwoofer calibration in a Dolby Atmos playback system and this study argues that it can be skewed with ‘whole-body vibration.’ This is a significant finding since Atmos is a calibration intensive format that is now being installed in automotive car cabins. Dolby Professional/Dolby Atmos for Cars reports that the 2025 Volvo “EX90 SUV will be outfitted with a 25-speaker Bowers & Wilkins Dolby Atmos playback system including four additional headrest speakers and four 3D speakers that emit sound from above.” (Dolby-A, 2024).

In his dissertation entitled *Auditory-Tactile Music Perception*, Sebastian Merchel argues that even with the Low Frequency effects (LFE) channel, expressed via a subwoofer, “the generation of tactile components is still very restricted” (Merchel, 2014). He observes that custom vibration tracks using a vibration actuator coupled to an *electro dynamic seat-mounted shaker* have been demonstrated in action-oriented DVD movies.<sup>47</sup> Depending on sound pressure level, the human ear can induce frequencies between about 20 Hz and 20 kHz, and spectral (frequency) range can be very different between music genres. For example, infrasound is highly involved in electronic dance music (EDM) and the last movement of Tchaikovsky’s 1812 overture, when the cannons fire.

*“Sound and vibrations are often perceived via the auditory and tactile senses simultaneously, e.g., in a car or during a rock concert. Even in a concert hall or a church, sound can excite suprathreshold vibrations in the ground or seats. If concert recordings are played back through headphones, this vibratory information is missing*

---

<sup>47</sup> See (Walker K., Martens W., L., 2006.)

*to date. The same holds true in the majority of cases for reproduction with multimedia or high-fidelity systems.”*

(Merchel, 2014, p.v)

Pop music mixers and mastering engineers often *high-pass filter* low frequency energy below about 30 Hz to prepare mainstream music for the average ability of consumer playback systems. Antithetically, when heard optimally in a calibrated full range speaker sphere, such as a cinematic theatre system, the Dolby Atmos format can express infrasound that is similar to what can be experienced in natural free field outdoor environments.

In 1999 three music enthusiasts, Norbert Heinz, Stefan Köpf, and Jurgen Scheuring founded their company called ASCENDO immersive audio, in Stuttgart, Germany. Ascendo has built the world's largest custom cinematic subwoofer. It is 80" with a 21" voice coil and has 70 mm of linear excursion. With a frequency response of 1Hz to 100 Hz its efficiency is 105 dB at 1 watt per meter. Its peak output is 148 dB and can be driven by a 20,000-watt power amplifier. To compare that with a common knowledge of power, 20,000 watts is equal to about 27 horsepower. The minimum enclosure size is 10m<sup>3</sup>. Ascendo makes a range of infrasonic subwoofers from 24" to 80". See **Figure 5**, The ASCENDO 80" subwoofer.



**Figure 5:** The ASCENDO 80” cinematic subwoofer is custom designed for infrasound.

### 3.6 Listening Regression

The Dolby Atmos format invites questions that are bound in the industrial construction of devices, formats, and listening practices. Innovations like the low resolution 128 kbps .mp3 file format have reduced mainstream audio quality and listening practice making music the background score of life via passive listening habits while on the go. This reduced listening function is akin to music in film media where it functions as accompaniment, compliment, and supplement.

For broadcast and streaming purposes Atmos is currently transported in an .mp4 container as a Dolby Digital Plus Joint Object Coding (DD+JOC) file, that allows for fifteen full range channels or audio-objects, plus the LFE channel. The data rates are, 384, 448, 576, 640, 768, and

1,024 kbps. On Blu-ray disc, Dolby Digital Plus Atmos supports data rates of 1,152, 1,280, 1408, 1512, 1536, and 1664 kbps. Dolby Atmos bitstreams are backward compatible with Dolby Digital Plus (Enhanced AC-3, DDP, DD+, E-AC-3, EC-3) decoders which have been in home theatre receivers for decades. Home theatre playback provides full cinematic dynamic range with a dialogue normal level of -31 LKFS<sup>48</sup>. TV speaker playback broadcast loudness is -24 LKFS, and portable devices afford -11 LKFS, which is typical of popular music. Dolby Atmos began as a film sound format but now it offers music the ability to employ 17 dB of dynamic range between the integrated loudness level of -18 LUFS and -1 True Peak (TP).

By fully incorporating artful 3D spatiotemporal orchestration and abundant dynamic range, the Atmos format is poised to reverse technologically induced (iPod/.mp3) listener regression.

In 1991 Theodore Adorno argued that “a culture industry is central to a consequential regression of listening in consumers for whom it provides commodities” (Steele, 2008, p.213). “Adorno traces the regression of [music] listening back to silent film musical accompaniment” (ibid; Eisler, 2003) which reduced music to a function of storytelling accompaniment.

It is heartening that the Dolby Atmos format for 3D music prioritizes dynamic range and a full spectrum. Atmos encourages active listening in a 3D virtual sound field that impels acoustic and psychological listener immersion.

---

<sup>48</sup> LKFS stand for Loudness, K-weighted, relative to full scale.

### 3.7 RUSH—Acoustic, and Psychological Immersion

Psychological immersion involves convincing localisation of audio-objects and palpable spatial contexts in the virtual 3D sound-box. It arouses cultural expectations that are capable of stirring emotions to cathartic levels by purging a listener of their previous state of mind.

During a conversation with eminent Atmos mixing engineer Richard Chycki<sup>49</sup>, he shared a convincing anecdote. Rich has mixed the legendary rock bands Aerosmith, Alice Cooper, Dream Theatre, Mick Jagger, and Rush. He described his Atmos approval playback of the song entitled “Tom Sawyer” to Rush guitarist Alex Lifeson. Lifeson is Rush’s key-man for mix approval. At the time, Amazon Music Unlimited was adopting Dolby Atmos and they asked Rush for an Atmos version of “Tom Sawyer.” This mix was the deciding factor that would determine if the band would adopt the format. It was Lifeson’s first Atmos playback experience and as Rich explains, it transitioned Alex’s mind state, perhaps to an immersive psychological “state of flow” (Csikszentmihalyi, 2004).

Alex arrived for playback. Rich cleared out of the listening space and left Alex to listen. Rich could see him grooving to the song but after the huge drum solo Alex became very still, so much so that Rich wondered if Alex liked what he was hearing. Alex remained unstirred until the 7/4 section and song’s ending fade out. Once the fade finished, he remained stationary. At this point Rich was convinced things were bad and Alex disliked the 3D music experience. Rich’s mind raced to thoughts about Rush’s adoption of the Atmos format now being doomed. Lifeson

---

<sup>49</sup> On September 30, 2024, at 11:00 am, in a zoom meeting, Richard and I talked about many aspects of making music in Atmos. He said one of the hardest parts of the process is achieving client and record label mix approval. From there he described playing his Atmos mix to RUSH guitarist Alex Lifeson for the first time. Rich visited me to hear my Dolby Atmos Hyper Near-Field Tiny Studio on Thursday October 10<sup>th</sup> and together we listened to some of his RUSH mixes. He deemed the studio to sound exceptionally good and was very pleased with how his mixes sounded. Rich also shared his RUSH story in a Dolby Professional video interview, (Dolby-B, Creator Talks, 2024).

remained still for what felt like an eternity. Eventually he turned around to face Rich, but still said nothing for an uncomfortably long time—finally, he spoke.

*“I’ve been playing this song for so long and this is the first time I’ve heard through speakers what I’ve heard in my head.”*

(Rushes Alex Lifeson, Dolby-B, Creator Talks, 2024)

Alex was speechlessly absorbed in the 3D music experience. Based on his lineage with the song, and then seeing him respond powerfully to the mix, to Rich, the importance and power of the Atmos format was confirmed. “Tom Sawyer” was released in Atmos on the *Moving Pictures* 40<sup>th</sup> Anniversary Edition, April 15<sup>th</sup>, 2022. Rich recounted that other artists routinely respond in similar ways when they first hear their music in immersive 3D audio, and I have also witnessed this.

The word “immersive” is common when describing 3D audio because manufacturers have adopted it, but Rich’s captivating story suggests that we should scrutinize the term further, as Edwin Pfanzagl-Cardone does. The noun immersion and its adjective immersive are often attributed to the Dolby Atmos format’s ability to provide auditorial immersion for the listener. However, psychological immersion describes a reconfigured state of consciousness. Lifeson’s experience indicates that psychological immersion caused his speechlessness.

Psychologically immersive experiences can be found in many of life’s activities. Mihaly Csikszentmihalyi’s interpretation of psychological immersion describes total absorption in an activity however one may even ponder similarities to other experiences of psychological immersion.

*“[B]oth cognitively and neurologically, the state of flow shares overlap with other altered states of consciousness, including meditative and psychedelic states [...] and so called peak or optimal experiences.” “[F]low is accompanied by a loss of self-consciousness.”*

(Kotler, Mannino, Kelso, Huskey, 2022, p.1)

Understanding the intentions and effects of acoustically and psychologically immersive music experiences are important considerations. Enlightened 3D music creators like Alex Lifeson and Richard Chycki see Atmos as a format that can reverse decades of listener regression which has been driven by an industry concerned with mass commodification of media, not the prioritisation of music’s sonic qualities and engaging listening experiences.

## Chapter 4: Examine Experiences in Music Spatiality

---

### 4.1 Forming Spatial Cognitive Image Schemata

Audio scene analysis (ASA) is a term introduced by Albert Bregman in 1993. It involves the study of sound behaviour in free- and closed field environments.

Rooted in Soundscape Theory (Fiebig et al, 2023, 1992), ASA trains the brain to perceive the re-creation of acoustic ecologies or the design of abstracted sonic environments in 3D music. I often find opportunities to observe nature's inherent construction of a sound field as I take sound walks focused on exploration of 3D acoustic ecology in environments.

Described anatomically, I seek to perceive sound posteriorly, laterally, anteriorly, and inferiorly (below), but with particular attention to the superior (height).

My favourite place to distinguish natural sound from above is at Tommy Thompson Park at Leslie Street Spit, in Toronto. It's a man-made headland that extends into Lake Ontario and is populated by over 300 species of birds. I listen and differentiate single birds, flocks of birds, and birds over the water and land. Turbo-prop planes are particularly fascinating audio-objects as they fly with regularity above the spit and approach from the East or depart from the West, located Billy Bishop Airport. I like to adapt my head-related transfer function by cupping my hands behind, above, or in front of my pinnae (outer ears) to perceive the interaural differences of sound energy, spectrum, and timing.

Sometimes I will plug one ear to perceive a monaural image of allocentric perspectives between audio-objects. Other times I alter my egocentric perspective by rotating or tilting my head upward or downward to hear the differences in bio-mechanical induction.



The metaphysical concepts contained in my analysis synergy for 3D music spatiality can also apply to the subjective identification of relationships, perspectives, properties and attributes heard in nature's spatiality. Audio scene analysis in all free- and closed field environments is an essential practice for 3D music creators. With understanding and perception of the following primary metaphysical concepts, 1) direct sound, 2) indirect sound, 3) relationships, 4) perspectives, 5) properties, 6) attributes, 7) audio source width, 8) apparent audio source width, 9) ensemble width, 10) apparent ensemble width, 11) listener envelopment, 12) proxemics, 13) impact, 14) intensity, 15) intelligibility, 16) pitch and height, 17) the spatial aggregate, and 18) the spatial impression of the holistic aggregate (Overall Listening Experience-OLE), we can exercise ASA and train ourselves daily. Humans have unique inductive hearing physiology and ASA utilizes our individually ingrained spatial cognitive image schemata which has accumulated in the brain since the first 28-30 weeks of gestation in the fetus.

*“Unlike the visual system where actual visual experiences begin after birth at term, the auditory system requires auditory experience with voice and language, music, and meaningful environmental sounds during the last 10 to 12 weeks of fetal life (28-30 weeks of gestational age) whether in utero or in a newborn intensive care unit ...”*

(Graven, Browne, 2008, p.1)

Closing our eyes helps the brain to focus on aural induction of 3D sound and when the eyes are open in outdoors the ocular and aural systems work together to locate sounds in 360-degrees. When first hearing music in Dolby Atmos, ear- and brain-training with the eyes closed is valuable. Hyunkook Lee has stated that effective auditory brain-training for 3D spatial music can take up to five or ten minutes, while the brain updates years of prior experience, stored as cognitive image schemata

#### 4.1.2 Early Experiences in Music Spatiality

My prior experiences and first curiosity about sound, music spatiality, acoustical, and psychological immersion started unwittingly at age eleven while playing the piano. Experiencing the spatial properties and attributes of sound created with the sustain pedal were mesmerizing. I could not explain what I was hearing, and I was not aware that foundational sonic schematic impressions were forming in my brain.

“Cognitive image schema” (Johnson, 1980; Lakoff, 1987; Leviton, 2006) identifies formative *a posteriori* world experience that is saved in the human brain.

*“The world is the totality of objects that can be known through experience (Erfahrung), known in terms of orderly theoretical thought on the basis of direct present (aktueller) experience”*

(Husserl, 1931, 2013, p.56)

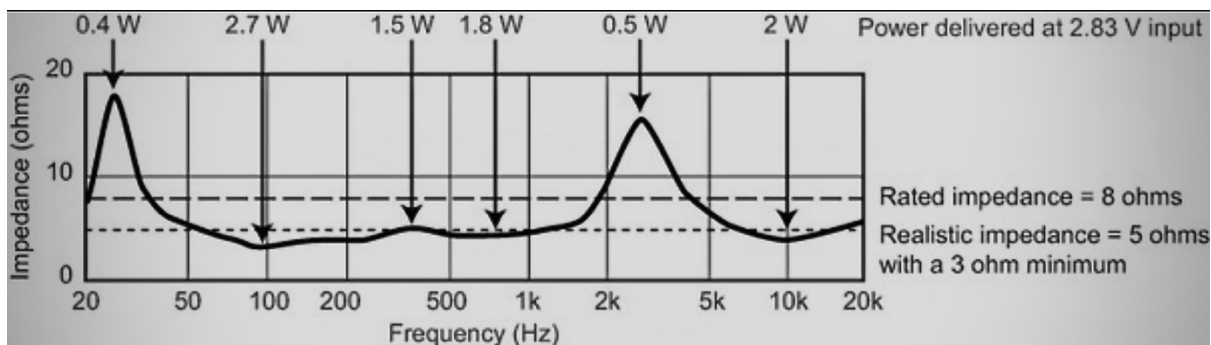
Experiences that developed my early cognitive image schema for music’s spatiality began after piano lessons ended at age ten due to disinterest. At age eleven, with renewed curiosity, I returned to the piano. I played our out-of-tune KARN upright piano daily and was spellbound when experimenting with the sustain pedal. The sustain pedal lifts the dampers so strings can ring openly and function like a physical reverberator. Single notes, chords, and arpeggios with variant dynamic attacks, created compelling decorrelated interaural reverberant sound worlds. At the time I could not explain any of this, but now I realise that I was fascinated with spectrum, energy, and time as they related to direct and indirect sound, intelligibility, intensity, pitch and height. Induction of these new-found relationships, perspectives, properties and attributes was forming my cognitive image schemata. I experimented with moving my head closer or further

from the piano and also turned my head sideways to differentiate between direct and indirect room sound. I soon realised that removal of the piano's upper and lower front panels projected more sound energy into the room. This represented the sound of an intimate proximity to the felt hammer-strikes on the strings. I perceived the extension of low and high frequency response, precise attack, expansive horizontal width, and more volume (energy) when my head was hyper-near to the strings. For the first time, gauging the correlation between volume change and proximity to an audio-object was revealed musically. Now, many years later, that awareness has helped me to imagine how my Dolby Atmos Hyper-Near Field Tiny Studio could deliver 85 dB of sound pressure level (SPL) within a hyper near (1.2-meter) speaker-to-listener distance.

#### 4.1.3 Hyper Near-Field Sound Energy

The human hearing system is measured with a logarithmic scale of successive amplitude and frequency points along an axis that are equal in distance. Sound pressure level is expressed in decibels (dB) and loudspeaker characteristics have been routinely measured at a ratio of energy to distance, usually 1 watt at 1 meter. Doubling energy delivers +3 dB, doubling distance yields -6 dB. Loudspeakers and amplifiers must be selected with matching impedances to insure the optimal transference of energy. Loudspeakers do not have the same effective impedance (opposition to varying electrical current) at all frequencies and Floyd Toole's graph (Figure 6) illustrates the non-linear results for rating loudspeaker sensitivity at a fixed 2.83 volts.

**Figure 6** shows the *actual power* of 2.7 watts at 100 Hz with a 3-ohm minimum impedance. It is different than the *actual power* of 0.4 watts at about 25 Hz with an 18-ohm impedance. Simply stated, the same voltage responds with different power outputs at dissimilar frequencies.



**Figure 6:** Impedance, wattage and frequency response (Toole, 2018).

*“Obviously, rating sensitivity according to power input does not work well. The domination of solid-state amplifiers really provided the solution. These amplifiers are essentially constant voltage sources with power rated according to what they can deliver into an 8-ohm resistor. If the load impedance drops to four ohms, the power will double. At 2-ohms, the power quadruples and so on until the amplifier can deliver no more current or exceeds some other limitation.”*

(Toole, 2018, p.437)

Power amplifiers must be able to drive low impedances but are usually rated at 8-ohms of resistance and the 1-watt standard voltage is 2.83 volts. Toole states that a distance of 2 meters is representative of the far-field for a small speaker. If the ear was located at 1 meter (half distance), 2.83 volts into 8-ohms would be 6 dB louder while a doubling of the distance decreases audio volume by 6 dB. Years ago, when I placed my ears closer to the piano, I was naively sensing logarithmic increase in loudness. Here, Toole not only pinpoints the physical theory at play in the design of speakers and amplifiers, but also how the intimate proximity to a sound source behaves and interacts with the (logarithmic) human hearing system.

With an 8-ohm speaker driven by at least 120 watts, placed at 1.2 meters to the subjective listener my Hyper Near-Field Tiny Studio can economically achieve Dolby's recommended speaker gamut SPL of 85 dB + 20 dB of headroom for True Peaks.

Those early in life experiments at the piano made me intuitively aware of the musical possibilities surrounding audio-object energy, intimate proximal listening, and the increased interaural sensation of extra-wide decorrelation in the horizontal left/right when introduced by the piano's sustain pedal. Ringing open strings created a physically tempered reverberance that was enveloping and immersive, in many ways similar to the perception of compelling Listener Envelopment (LEV) in acoustic concert hall sound. [See chapter 6.9](#) for more about LEV, which is one of my five Super Components involved for the creation of compelling 3D music spatiality.

## 4.2 Spatial Gamut, Attributes, Super Components, and the 9.1.6 Object-Bed

Compelling envelopment the Dolby Atmos 7.1.4<sup>50</sup> format involves twelve speakers, and the 9.1.6 gamut involves sixteen speakers by adding left and right wide speakers (Lw/Rw), as well as left and right top middle speakers (Ltm/Rtm). When compared to 2-channel stereo, the aggregate power of a large 12<sup>51</sup> or 16 speaker gamut can deliver dimensional separation with impact. The most common question I hear from emerging stereo practitioners who are curious to work in Atmos is “what do you do with all those speakers?”

---

<sup>50</sup> 7.1.4 requires 12 speakers in total, on three planes. The inferior (lower) plane locates the Low Frequency Effects (LFE) speaker. The standard mid-plane is approximately 120 centimeters above the floor and involves 7 speakers: Left, Center, Right, Left Side, Right side, Left side rear surround, Right side rear surround. The superior (height) plane involves 4 speakers: Top front right, Top front left, Top rear right, and Top rear left. Other music configurations are possible and range from 9.1.6 to 22.1.

<sup>51</sup> My Hyper Near-Field Tiny Studio now uses 2 6-channel Bryston BD-325 Class D power amps and a PSB Sub-Woofer amplifier delivering about 4000 Root Mean Squared (RMS) watts of power to 12 speakers at 8-ohms, about 5.5 horsepower.

My answer starts with what I propose to be the five Super Components of Auditory Spatial Impression (ASI) consisting of: 1) *Source Width* (SW); 2) *Apparent Source Width* (ASW); 3) *Ensemble Width* (EW), 4) *Apparent Ensemble Width* (AEW), and 5) *Listener Envelopment* (LEV). As noted by Antoniuk, Zieliński, and Lee, these five components are rooted in the acoustic design rationales for concert hall architecture and binaural spatial audio scene analysis (SASA). When viewed as Super Components, I propose that they provide effective conceptual zones for clean and unconflicted dimensional imaging in a 3D sound field. A fixed 9.1.6 object-bed (O-bed) is a metaphysical construct that routes audio-objects to locations in the virtual Atmos renderer sound-box. The O-bed was first invented by recording engineer Steve Genewick and Avid Technologies liked his idea so much that they started including pre-built object-bed presets in Pro Tools 2023 (Genewick 2024a).

Routing the five Super Components into a 9.1.6 object-bed<sup>52</sup> brings a compelling closed field sense of envelopment and utilizes the Atmos renderer to reliably re-render the large O-bed into a lower resolution 7.1.4 channel size and further down to stereo.

To test the Super Components, I used a live stereo jazz recording and assigned the natural stereo mix to the front stereo left and right speakers. I considered this to be the natural ensemble width (EW) which has a small amount of innate indirect sound radiosity “source bonded” (Smalley, 1997, 2007) to the direct sound, as captured by the recordist. Attempting to create virtual indirect spatiality, I cascaded two reverbs together to construct virtual Apparent Ensemble Width (AEW) and routed them to the left- and right-wide speaker zone in a 9.1.6 O-bed.

---

<sup>52</sup> A 9.1.6 O-bed involves speakers: Left, Center, Right, Left wide, Right wide, Left side surround, Right side surround, Left rear surround, right rear surround, Left top front, Right top front, Left middle top, Right middle top, Left rear top, Right rear top. This size of O-bed has a high level of spatial resolution.

Noticeably the center image was empty, so I created a mono sum of the original stereo recording and spectrally filtered it to feature the saxophone and vibes at center stage. This created a physical center of direct sound within the anterior horizontal image width involving the L, R, C, Lw, Rw speakers. The artificially derived center image was once-again split and routed to an 80-millisecond high pass filtered delay, routed back to the center speaker at a lower level and placed upstage (far from the listener) on the medial center line to delicately express anteriorly staged depth of field.

Listener envelopment (LEV) was generated using another stereo reverb with a pre-delay tuned in careful relation to the upstage center depth signal. These geometric spatial zones created a triangle of spatiality that emanated from the center and widened toward the rear surround and height speakers. Albeit it was a simple test, it suggests that a large O-bed—bigger than the available physical speaker array—provides a high-level of convincing dimensional zone resolution that is noticeably pleasing in speakers and binaural headphone playback.

I argue that the Super Components can facilitate fast and intuitive use of a 9.1.6 O-bed, giving emerging 3D practitioners a logical and reliable rationale for the design of Auditory Spatial Impressions when working with twelve or more speakers in the Dolby Atmos format.

## Chapter 5: Examine, Organize and Pinpoint the Components

---

### 5.1 Organize the Components

To prepare for the next episode, we must select and organize myriad cross-domain concepts, vocabulary, and terminologies that can describe the substance of popular music's 3D spatiality in the Dolby Atmos format. As we recall, Bosun Xie lists over 1000 bibliographical references regarding spatial sound and he states that the study is highly interdisciplinary but cross-domain literature is rare. With consideration of this vast canon of literature, William Moylan's organizational framework and semantic descriptions rang true for me as a composer, performer and music producer.

### 5.2 William Moylan's Organization of Spatial Perspectives, Properties, Variables, and Attributes

William Moylan calls musical spatiality an element of recording and describes it as illusory. He argues that "every record re-invents physics—the relationships and dimensions of our world ... space is refined precisely in every track to serve the artistic intentions (Moylan, 2020). His musicological encapsulations continue to shape my curiosity for the topic. They provide metaphoric ways of knowing about music's stereo spatiality in relationship to the listening experience and song-meaning and I believe they can extend effectively to three dimensions. When expanding his concepts from stereo to 3D, subjective scrutinization of spatial music began to reveal that the inherent 3D codes required new ways of knowing about the uniqueness of a song's spaces and places.

Moylan founds his concepts on facets of ecological psychoacoustics (Neuhoff, 2004) and ecological listening (Clarke, 2005) which facilitate spatial concepts. My focus on positional



relationships, cognitive perspectives, proportional properties and character attributes can be perceived as multidimensional layers of context found within the recorded music.

These layers form the spatial aggregate—a sum of all indirect sound variables that shape a song as much as any other element.

I share William Moylan’s purpose which is to define spatiality, explore how it appears, and then think extensively about how it can shape the recording, and be described.

Episode 2 is designed to pinpoint and describe the essential componentry. Hopefully it will bring a safe harbour to those who are adrift in their *lonely galaxy* and *lost in the space* of 3D music. I will propose a 3D spatial analysis synergy and the associated salient metaphysical components that build on the Ethos revealed in Episode One.

### 5.3 NASA: Lost in Space



**Figure 7:** Lonely Galaxy Lost in Space.

For a stereo (two speakers) creator, this image is a metaphor that suggests being oddly lost in the spatial sound field provided by twelve or more speakers in the Dolby Atmos Format.

Lonely Galaxy Lost in Space. NASA. ID GSFC\_20171208\_Archive-e0007110

Most galaxies are clumped together in groups or clusters. A neighboring galaxy is never far away. But this Galaxy, known as NGC 6503, has found itself in a lonely position, at the edge of a strangely empty patch of space called the Local Void. The Local Void is a huge stretch of space that is at least 150 million light-years across. It seems completely empty of stars or galaxies. The galaxy's odd location on the edge of this never-landed led stargazer Stephen James O'Meara to dub it the "Lost-in-Space Galaxy" in his 2007 book, *Hidden Treasures*. NGC 6503 is 18 million light-years away from us in the northern circumpolar constellation of Draco. NGC 6503 spans some 30,000 light years, about 1/3 of the size of the Milky Way.

NASA Permissions: Non-Commercial use agreement; News outlets, schools, and text-book authors. For educational or information purposes.

## EPISODE 2

### Chapter 6: 3D Spatial Music Theory and Semantic Analysis Synergy

---

#### 6.1 Overview

Setting forth from Episode 1, this episode proposes a musicological three-tool analysis synergy to posit the substance of the three-dimensional (3D) spatiality of music and its semantic language in the immersive Dolby Atmos format. “The natural listening environment can be defined as the acoustic space we occupy during our daily lives. A virtual auditory space is an acoustic environment created through the use of loudspeakers or headphones designed to replace or augment the natural listening environment” (Roginska, Geluso, 2018, p.1). Subjective listening in a hyper-calibrated 7.1.4 Dolby Atmos Hyper Near-Field Tiny Studio reveals that some music—*Peter Gabriel, i/o*—presents a highly compelling sensation of spatiality. This perception invites the following questions: 1) What is the substance of compelling spatiality in immersive 3D Dolby Atmos music recordings? and 2) How can we find it?

To answer these questions, this author proposes a theoretical 3-part analysis system called a Synergy.

**Part One** offers a semantic “Concept Map” (Seetharaman, Pardo, 2014, p.2), a thesaurus-like listing to look up 3D concepts, vocabulary, and terminology, organized with William Moylan's stereo framework of spatiality (Moylan, 2020, pp.287-356). The Concept Map guides the analyst to find their scope for perception of 3D spatial componentry involving positional relationships, cognitive perspectives, proportional properties and their character attributes.

**Part Two** presents a rudimentary Five-Facet Typology which distills salient concepts into five essential facets with simple descriptions of the metaphysical componentry involved.

**Part Three** proposes ear training exercises to improve the perception of 3D sound. This three-part system utilizes pre-existing “cognitive image schema” (Johnson, 1980; Lakoff, 1987; Leviton, 2006). It leverages intertextual cross-domain mapping of concepts, vocabulary, and terms that are borrowed from human anatomy, anthropology, psychoacoustics, acoustics, electroacoustics, musicology, and live concert/club/theatrical arts. Cross-domain mapping provides a “patterned way-of-knowing” [that is] “translatable from one sense-domain to another” (Moore, Schmidt, Dockwray, 2011, p.84). Human experiences of sound in spaces and places involve subject-listener emplacements with artful “acoustemological” attributes (Rice, 2018; Feld, 1994) that intend to convey subjective song meanings. Earmarked for composers, songwriters, mixing engineers, and pedagogues, this analysis synergy requires “ecological listening” (Clarke, 2005, p.105) and repeated practical application.

## 6.2 Episode 2 Structure

The 3D Concept Map and the associated distilled Five-Facet Typology will be revealed in an order that begins with William Moylan’s Concept Map for stereo music spatiality. It is the role model for an expansive 3D Concept Map. After that Moylan’s Concept Map is presented again but this time it contains this author’s Five-Facets within Moylan’s three categories of: A) Individual Sound Source; B) Composite texture; and C) Overall sound. Completing the reveal sequence is the Five-Facet Typology with component descriptions.

If the reader wishes to glean the [Concept Map](#) and [Five-Facet Typology](#) before reading on, please use the hyper-links provided above for quick access. Five examples of 3D spatial music analysis that utilize the author’s Synergy are presented at the end of this episode.

## 6.3 Introduction

The 3D Dolby Atmos music format can virtually immerse a listener in natural sound. Its spatial localization improves with training the listener's brain to perceive audio-objects located medially, anteriorly, vertically, laterally, and posteriorly. Eminent 3D music mixers enliven song meaning with staging and dimension to express music's environmental nature by broadening their roles to become immersive storytellers who spatiotemporally orchestrate audio-objects in 360-degrees around the listener. However, emerging immersive music creators often struggle to transition their stereo practice to 3D because Atmos pedagogy is underdeveloped. Many interviews with mixers exist online which provide insights into 3D music practice, but precise descriptions about the relationships, perspectives, properties and attributes of 3D music spatiality are conspicuously absent. However, “[i]n recent years, popular music scholars have increasingly turned their attention toward the virtual spaces of recorded music” (Gamble, 2017, p.105).

This episode posits the substance of implicit 3D spatiality heard in recorded music and introduces a synergy for 3D spatial analysis of music in the Dolby Atmos format.

### 6.3.1 Seven Principles for Perception of 3D Spatiality in Music

The following seven principles set the stage for 3D spatial analysis of recorded music in Atmos. 1) All auditory spatiality is a combination of direct and indirect sound; 2) Subjective listening in a closed field space involves a calibrated playback system with awareness of the critical listening distance. “[T]he *critical [listening] distance*<sup>53</sup> is where the direct sound from the

---

<sup>53</sup> The critical listening distance is also known as the reverberation radius of a closed field venue. In 3D spatial recording of indirect sound reverberation time, directionality of the sound source, and microphone pick-up pattern are essential components.

speaker is equal in magnitude to the sound of the reverberant field [or indirect sound]” (Serinus, 2022). 3) If the listener is within intimate hyper-near proximity to the speaker, room effect will become relatively innocuous; 4) Recorded *direct sound* is an audio-object containing little to no environmental spatial nature. 5) Recorded *indirect sound* is the reflected spatial nature set off by a direct sound, which is generated harmoniously in indoor/outdoor (closed- and free field) environments and can also be artificially generated during music production; 6) In nature, all direct sound begins as a monaural object and becomes dimensionally spatialized in 3D by environmental reflections, which are induced by the pinnae in humans and mammals. Low frequency infrasound is somatically induced through the body. 7) The human brain carries a lifetime of learned sonic understanding—a cognitive schematic—of 3D induced auditorial spatial relationships, perspectives, properties and attributes in association with visual imagery.

## 6.4 Analysis System Evolution

Musicological analysis systems are constantly evolving to better explain and systematize innate human creativity. With the rapid shift to 3D spatial music, the associated praxis, theory, and pedagogy is progressing quickly to hypothesize popular music’s 3D spatiality.

*“The history of music theory encompasses a broad range of thought and activities that have varied widely over time.” “... topics that at one point have been deemed fundamental to notions of music theory have often receded in importance, while other topics have emerged in the discipline that would not have been understood as belonging within the older paradigm.” “... [M]usic theory needs to be flexible, one in which the ever-changing subject matter and historically shifting boundaries of the discipline can be accommodated.”*

(Christenson, 2017)

This author's 3D Music Analysis Synergy builds on cognitive schemata and proposes the salient semantic language of 3D spatiality for popular music. Concepts, vocabulary, and terms map across domains to form specific intertextualities that clarify perceptions of 3D spatial relationships, perspectives, properties and attributes.

## 6.5 Relationships, Perspectives, Properties and Attributes

The theoretical spatial analysis frameworks are built on positional relationships, cognitive perspectives, proportional properties and character attributes which exist between the subjective-listener and the audio-object.

### 6.5.1 Relationships Determine Perspectives and the Circular Recording Method

There are three types of lateral and contralateral (clockwise/counterclockwise) positional relationships: 1) "subject-to-object" (Lungwitz, 1923, 1933; Bense, 1961); 2) object-to-subject; and 3) object-to-object. Subject-to-object, object-to-subject, and object-to-object relationships determine cognitive perspectives which are classed as "egocentric" or "allocentric" (Fernandez-Baizan et al, 2020). The laterally clockwise subject-to-object and contralaterally counterclockwise object-to-subject relationships involve an egocentric perspective, whereas the object-to-object relationships involve allocentric perspectives.

Circularly surrounding a performer, section, or full ensemble with an inward facing multi microphone array inverts the directionality of pickup but the relationships and perspectives remain. The allocentric perspectives of object-to-object section and ensemble staging can be captured naturally eliminating the need for panning during the final mix. A 3D microphone array (gamut) represents the subject-listener. The location and distance between microphones form a second layer of nested mic-to-mic allocentric perspectives; a virtual human with multiple ears.

The instrument/performer (section or ensemble) remains the audio-object and it contains nested allocentric staging perspectives. Throughout this phenomenon, the perspective of subject-listener (microphone gamut) to audio-object remains egocentric.

Inverting the directionality of a mic array to point inward at an audio-object can be classed as a *circular inward facing method*. It is a powerful 3D spatial recording technique that captures innate direct and indirect environmental sound in a critical balance. During playback the subjective listener uniquely experiences the audio-object circularly, while virtually situated at an (otherwise impossible) *inside* listening spot. This *circular inward facing method* is used in Justin Gray's 2025 ground-breaking Grammy Award nominated album entitled "Immersive."

#### 6.5.2 Properties Determine Attributes

Properties are the proportional objective grammar of 3D spatial art and craft. Properties can be measured subjectively or with metering. Properties determine attributes which describe song meaning, experiences, and resultant emotions. They could involve examinations of persona, fandom, proximity, community, scene, genre, social, culture, sub-culture, environmental spatial context, or any other context. A good example can be perceived from the properties of spectrum, delay, and reflection timings used to re-create the sonic attributes of a sports bar environment, as heard in Huey Lewis's "Heart of Rockn' Roll," mixed by Bob Clearmountain.

#### 6.5.3 The Subject-Listener

The subject-listener is the mixer, composer, musicologist, or general listener and the audio-object represents an instrumentalist, singer/performer, or audio sound source. The Dolby Atmos format positions the subject-listener in a locked egocentric perspective with audio-objects locating in three dimensions surrounding the listener.



In Dolby Atmos, relationships and perspectives are egocentrically perceived from the subject-listener's sweet spot. Allocentric object-to-object (sound-to-sound, musician-to-musician, or inward facing mic-to-mic) perspectives can describe associated onstage and offstage relationships in the 360-degree sound-box.

## 6.6 Human Induction and Perception

Auditory induction in humans is classified in two ways: 1) peripheral auditory processing, and 2) central auditory processing. Peripheral processing starts with the outer ear and ends at the cochlear nerve. "The central auditory system is composed of a number of nuclei and complex pathways that ascend within the brainstem" (Wenzal, Begault, Godfroy-Cooper, 2018, p.17).

"The parietal lobe [of the human brain] is part of the cerebral cortex lying between the occipital and front lobes" (Goldberg, 2001, pp.11051-11054). The parietal lobe is responsible for analysis of space using all sensory modalities including somatic induction of infrasound below about 18-20 Hz, which cannot be sensed by the ears. Natural acoustic immersion involves sound arriving from 720 degrees around the subject-listener, which is an inevitable consequence of human hearing in an acoustic space.

*"Auditory spatial perception refers to the ability to localize individual sound sources in 3D space even when multiple, simultaneous sources are present."*

(Wenzal, Begault, Godfroy-Cooper, 2018, p.10)

*"The localization of a sound source in the horizontal dimension (azimuth) results from the detection of left-right interaural differences in time arrival and interaural differences in intensity at the two ears (Middlebrooks & Green, 1991)."*

(ibid, p.11)

*“To localize a sound in the vertical dimension (elevation) and to resolve front-back confusions, the auditory system relies on detailed geometry of the pinnae, causing acoustic waves to diffract and undergo direction-dependent reflections” (Blauert, 1997; Hofman & Van Opstal, 2003).”*

(ibid)

## 6.7 Influential Literature and Concept Maps

Many researchers have created concept maps to semantically describe the components of spatiality in music. Popular musicology involves subjective intertextuality and much has been written about spatial relationships between material and musical forms. Please recall, Dr. Bosun Xie argues that “acoustics (physics), psychology, and physiology of hearing, electronics and signal processing, computers, and even the art of music” (Xie, 2023, p.xv) inform 3D spatial research.

For the analysis synergy, research considers: Brøvig-Hanssen and Danielsen (2013); Blauert (1983, 1997); Dockwray and Moore (2010); Meaux, E. and Marchand, S. (2021); Moore (1992); Holman, T. (2008); Patterson and Lee, (2021); Pawlak, Lee, Mäkivirta, and Lund (2023); Rumsey (2002, 2001); Denis Smalley (Smalley, D., 2007, 1997, 1986); and Xie (2023).

Attempting to centralize essential 3D concepts, semantic vocabulary, and terminologies, this author introduces a 3D Concept Map adapted from William Moylan’s stereo map of “spatial properties, attributes [and] variables, in three levels of perspective” (Moylan, 2020, p.292). Please note that Dr. Moylan’s “levels of perspective” differ from the cognitive mapping of environmental representation associated with *egocentric* and *allocentric* perspectives.

### 6.7.1 Cognitive Image Schemata and Metaphor

German philosopher Immanuel Kant describes schema as “a structure of imagination that connects concepts with percepts” (Johnson, 1989, p.116). “Image Schemata” (Johnson, 1987, 1989) and “Proxemics” (Hall, 1963) can unite, as is argued in *A Hermeneutics of Spatialization for Recorded Song* (Moore, Schmidt, Dockwray, 2011) to form a multi-faceted framework employing metaphoric “notions of ecological perception” (Clarke, 2005) in the sound-box.

*“Image schemata functions as the basis for rational thought by way of metaphorical extension serving as the pre conceptual ground for cross-domain mapping.”*

(Moore, Schmidt, Dockwray, 2011, p.85)

*“[O]ne of the significant analytical tools yielded by image-schemata-based analysis is the ability to perceive a subject position for both listener and [performer] persona...”*

(Moore, Schmidt, Dockwray, 2011, p.99)

### 6.7.2 Evaluating Proxemics

In 1963, the American anthropologist and cross-cultural researcher Edward Hall introduced Proxemics as four zones of space: intimate, personal, social, and public. “Hall's work is most useful to us as a means of theorizing the distance between the listener's perceptual position, which we can locate either within or outside the created spaces of recorded song, and the position of the song's persona (and the relationship it holds with its environment.)” (Moore, Schmidt, Dockwray, 2011, pp.98-99). “[T]he listener gauges the distance between her- or himself and [artist] persona by engaging in the overall scene of the track, which offers additional aural clues to spatial placement and perceived proxemic distance...” (Moore, Schmidt, Dockwray, 2011, p.101). Gauging distance relies on an innate sense of proxemic analysis.

Albert Bregman theorized this activity as Audio Scene Analysis [ASA]. “By analogy, auditory scene analysis is the process whereby all the auditory evidence that comes, over time, from a single perceptual environmental source is put together as a perceptual unit” (Bregman, 1993, p.11).

### 6.7.3 Evaluating Spatial Quality with Semantic Language (Rumsey)

In 2002 Francis Rumsey published his system of Spatial Quality Evaluation for Reproduced Sound. “Rumsey’s aim is the clarification of semantic language” (Moore, Schmidt, Dockwray, 2011, p.101) and his process is scene-based, involving “segmentation, segregation, and integration” (Bregman, 1993, p.23). “Rigor in the physical measurement of sound signals should be matched by equal rigor in semantics relating to subjective evaluation” (Rumsey, 2002, p.651). His approach involves terminology and meaning within an audio scene-based paradigm. He identifies the need for meaningful subjective evaluation of spatial quality in sound reproduction and differentiates *sound quality* from *sound character*. Sound *quality* involves preferential and emotive responses and sound *character* is purely descriptive.

Because numerically rated Mean Opinion Score (MOS) listening tests conflate sound *qualities* and sound *characteristics* into a single rating, typically from 1 (weak), to 5 (strong), the MOS rating system is not often used, prompting the question: What is the difference between an objective spatial sound *quality* and its subjective *character* attribute? William Moylan defines spatial properties as the sound we first hear and perceive, and then we “engage in observing their attributes” (Moylan 2020, p.287). It’s a blurry area and for clarity this author refers the reader back to [Chapter 6.5](#) for explicit interpretations of 3D positional relationships, cognitive perspectives, proportional properties, and character attributes.

Rumsey argues that “spatial attributes should be identified as meaningful to 1) individual subjects; 2) a well-defined group of expert subjects forming a listening panel that agrees upon a set of attributes to be graded; 3) expert listeners not associated with that listening panel; and 4) independent observers or readers of the results” (Rumsey, 2002, p.654). Rumsey highlights the work of Tobias Neher, a research student at the Institute of Sound Recording who created the distinction between “attributes of spaces” and “spatial attributes” (Neher, 2021) which then fostered extensive research carried out at IRCAM. Their Spatialisateur (SPAT) software and the MPEG-4 codec partially incorporate IRCAM’S resultant spatial audio scene language. **Tables 3 and 4** show IRCAM Concept Maps that use semantic language to clarify spatial attributes and properties for direct and indirect sound.

| <b><u>Group 1: (direct sound)</u></b> |                                                      |
|---------------------------------------|------------------------------------------------------|
| Source-related <u>attributes</u>      | Corresponding objective criteria, <u>properties</u>  |
| Source presence:                      | Energy of direct sound and early room effect         |
| Source warmth:                        | Variation of early energy with frequency             |
| Source brilliance:                    | Variation of early energy with frequency             |
| Room presence:                        | Energy of late room effect                           |
| Running reverberance:                 | Early decay time                                     |
| Envelopment:                          | Energy of early room effect relative to direct sound |

**Table 3:** Direct sound attributes, corresponding objective criteria, properties (Rumsey, 2002).

| <b><u>Group II: (indirect sound)</u></b> |                                                     |
|------------------------------------------|-----------------------------------------------------|
| Room-related <u>attributes</u>           | Corresponding objective criteria, <u>properties</u> |
| Late reverberance:                       | Late decay time                                     |
| Heaviness:                               | Variation of decay time with frequency              |
| Liveness:                                | Variation of decay time with frequency              |

**Table 4:** Indirect sound attributes, corresponding objective criteria, properties (ibid).

These Group 1 and II Concept Maps separate direct sound from indirect sound. This separation is essential in 3D ear training. In the early 1980s, Floyd Toole undertook subjective testing of loudspeakers, evaluating spatial quality using the Concept Map illustrated in **Table 5**.

|   | Category                                                    | Floyd Toole's elaboration                                                                        |
|---|-------------------------------------------------------------|--------------------------------------------------------------------------------------------------|
| 1 | Definition of sound images                                  | Stability, focus and source separation.                                                          |
| 2 | Continuity of sound stage                                   | A form of width homogeneity relating to the even distribution of sources across the sound stage. |
| 3 | Width of soundstage                                         | Related to the width between outer sources on the soundstage, not including reverberation.       |
| 4 | Impression of distance and depth                            | The definition suggests it is in fact depth that is meant                                        |
| 5 | Abnormal effects                                            | Unusual or unnatural spatial effects such as phasiness.                                          |
| 6 | Reproduction of ambiances, spaciousness, and reverberation. | (no elaboration from Toole)                                                                      |
| 7 | Perspective                                                 | Graded from you are there though they are here, to artificial versus contrived.                  |

**Table 5:** Floyd Toole's Concept Map for evaluation of spatial quality (Rumsey, 2002).

In **Table 6** Rumsey provides the IEC 60268 (1997), and EBU 562-3 (1990) Concept Maps for evaluating spatial quality.

|   | Category           | Attributes/Variables                                                                                                                                                                                                                                  |
|---|--------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | Image localization | Perceived spatial location of a reproduced sound source. The image may be well defined or blurred.                                                                                                                                                    |
| 2 | Image stability    | Perceived location of the reproduced sound source. May change with pitch, loudness, or timbre. It may also change as a function of listener position, head rotation, or other normal movements. If these effects are small, the image will be stable. |
| 3 | Width homogeneity  | Stereophonic image should be distributed uniformly between loudspeakers.                                                                                                                                                                              |

**Table 6:** IEC 60268, Concept Map for overall spatial quality.

An earlier IEC 268 Concept Map was a basic version of IEC 60268 which involved a proportional approach to properties of: 1) Spaciousness, (closed to spacious); 2) Distance, (distant to near); and 3) Location of sources, (unstable to stable). The EBU 562-3 Concept Map proposes attributes for multidimensional evaluation of spatiality involving multichannel sound for HDTV. It involved the following descriptors: 1) Apparent sound stage width; 2) Surround effect; 3) Apparent room size; 4) Horizontal and vertical localization; 5) Naturalness; 6) Sense of reality; 7) Agreeableness (Rumsey 2002, p.656).

## 6.8 Associating the Spatial Attributes of Concert Halls and Theatres

*“[S]patial quality has been studied in concert hall acoustics, and there is a certain amount that can be learned from these studies in relation to reproduced sound”, however, “reproduced sound and synthetic auditory scene creation can give rise to subjective attributes either not encountered or not considered relevant in natural acoustics.”*

(Rumsey, 2002, p.652)

Rumsey argues that understanding the natural spatial attributes in concert halls can benefit spatial analysis systems. Under the heading of Auditory Spatial Impression (ASI) he introduces two fundamental dimensions: 1) Apparent Source Width (ASW); and 2) Listener Envelopment (LEV). Heinrich Kuttruff also emphasises that ASI predominantly involves ASW and LEV, and he highlights experiments by Bradley and Soulodre that indicate, “angular distribution had the largest effect on spaciousness. The wider it was, the higher is the LEV” (Bradley, Soulodre, 1995). “The best objective predictor for LEV is the strength of the late lateral reflections” (Kuttruff, 2016, p.225-226). In 1981 Barron and Marshall found that the degree of spatial

impression (ASW and LEV) similar to a concert hall occurred with lateral delay times lasting between 5 to 80 milliseconds (Barron and Marshall, 1981, p. 230).

*“The ‘sensation of spaciousness’ is primarily expressed from the lateral directions involving delay times of between 5 to 80ms”*

(Kuttruff, 2016, p.181)

This author agrees that ASW and LEV are two primary spatial impressions, but he proposes three more. He argues that five key spatial impressions can function as Super Components that empower a compelling sense of 3D spatiality. These were introduced in Episode 1 [Chapter 4.2](#) and will be further emphasized in the following Chapter 6.9. ASW and LEV primarily serve to create a compelling 3D listening experience which can also be described as a strong sense of spatiality. Placing audio-objects onstage or offstage is an effective way to conceive of abstracted spatiotemporal ensemble orchestration and the five spatial Super Component impressions help us to understand how closed field acoustic envelopment can be virtually constructed in the Dolby Atmos format. See **Figure 8**, Metaphoric Theatrical Object Staging Associations.

## 6.9 Auditory Spatial Impression (ASI) and Five Dimensional Super Components

Creating a compelling spatial impression in the Atmos format is aided with application of this author’s previously introduced five Super Components. “ASI” (Xie, 2023) houses the dimensions of, 1) Source Width (SW); 2) Apparent Source Width (ASW); which is also called “Audio Source Width” (ibid); 3) Ensemble Width (EW); 4) “Apparent Ensemble Width (AEW);” and 5) “Listener Envelopment (LEV),” (Rumsey, 2002; Kuttruff, 2016; Xie, 2023).



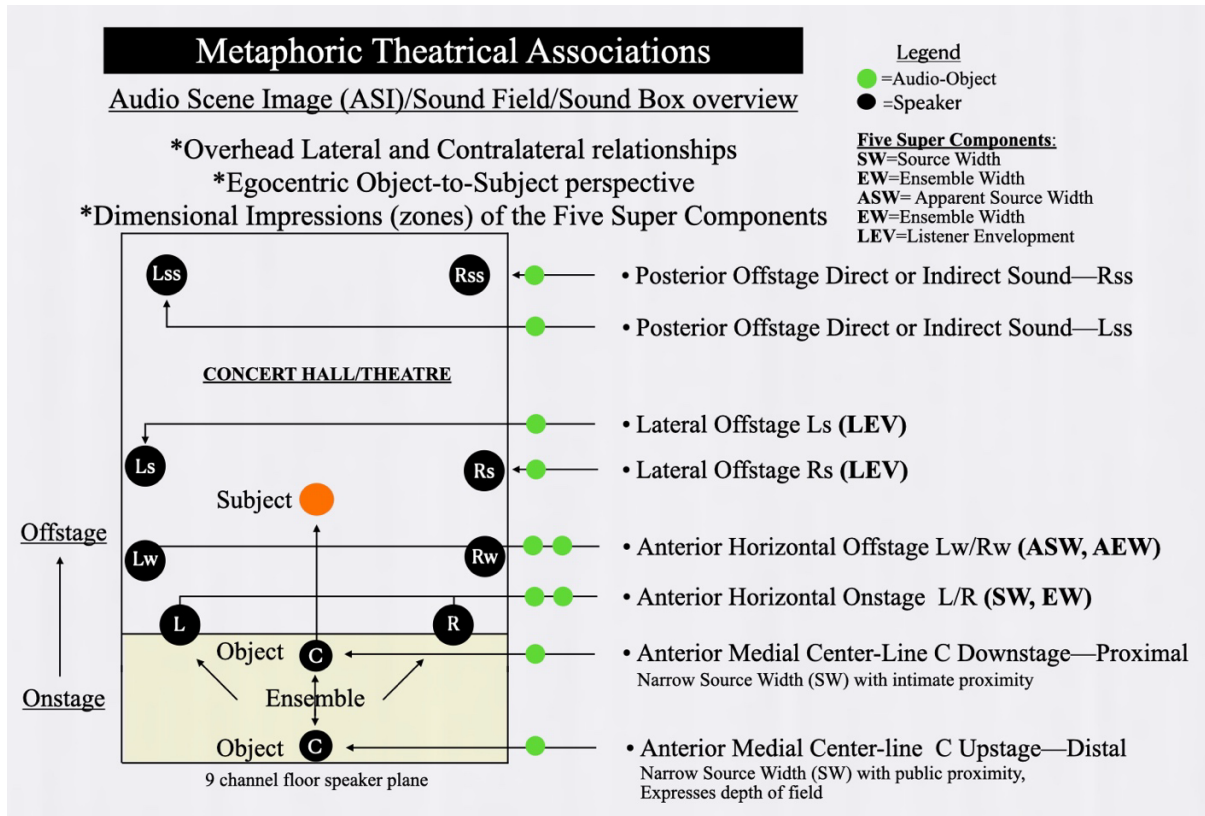
In the Atmos format SW, ASW, EW, and AEW are often perceived anteriorly onstage using the L, C, R speakers. Expanding or widening these components via offstage direct or indirect sound can be achieved with the Lw, Rw speakers in a 9.1.6 O-bed. See [Chapter 4.2](#). In 3D each of the Super Components can be expanded or dispersed using depth and height to fill holes in the sound field. This keeps audio-objects from sounding orphaned and bare.

In Atmos individually recorded performers are represented as audio-objects which can be positioned onstage or offstage in 360-degrees and in various proximities ranging from intimate to public. LEV can provide confident weighting of offstage indirect sound. It brings the essential substance of environmental width, which is predominant laterally, but also emanates posteriorly, and/or in the superior height.

The lateral side speakers (Ls, Rs) express LEV most effectively because they are on axis to the ears, meanwhile a portion of stereo indirect sound can also locate onstage in the L/R front speakers. This effectively portrays anterior stage width and depth. Monaural direct sound can also express depth of field in the C speaker by involving “source-bonded” (Smalley, 1997, 2007) natural indirect sound.

The properties of energy, spectrum, and timing work together to create the desired proxemic illusion (a character attribute) along the medial center stage line. Auditoriums and theatres may involve a proscenium arch at the front of the stage, which influences the radiation of onstage/offstage indirect and direct sound.

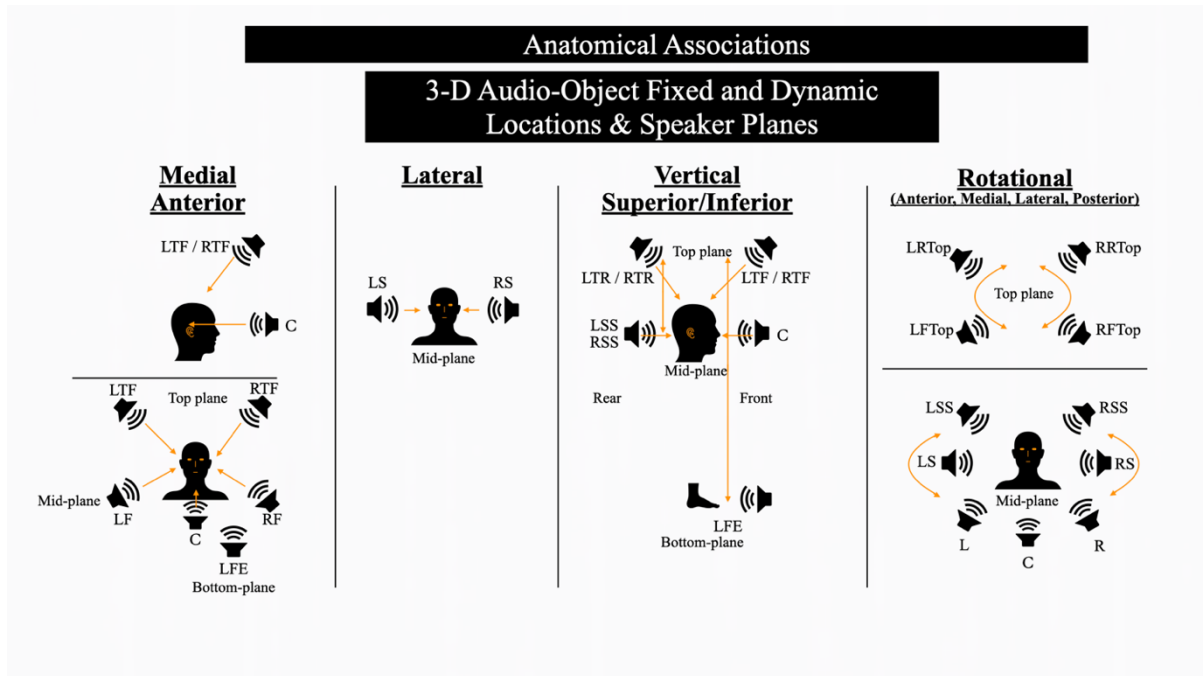
Please see **Figure 8**: Metaphoric theatrical associations.



**Figure 8:** Metaphoric theatrical associations.

## 6.10 Anatomical Associations

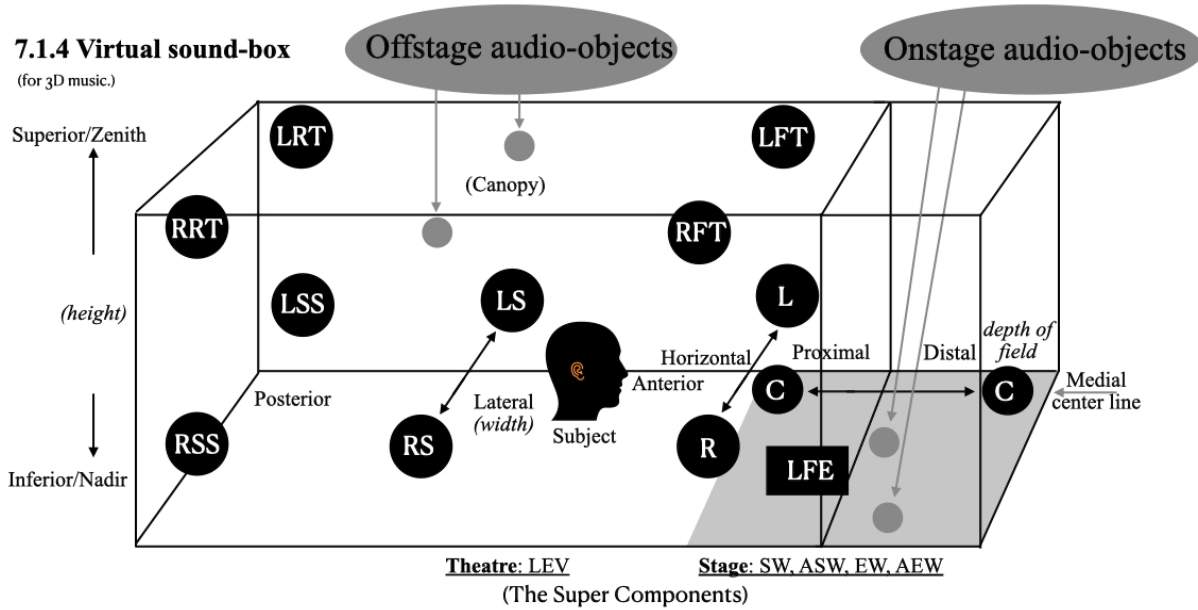
Leveraging terminology, concepts, and auditory experiences (stored as image schemata) from concert halls, music clubs, and theatrical venues can interoperate with common human anatomical associations to ingrain ready understandings of spatial relationships, perspectives, properties and attributes. The parietal lobe of the human brain analyses space with somatic induction using all sensory modalities. See **Figure 9:** Anatomical associations.



**Figure 9:** Anatomical associations.

Cognitive image schemas inform and extend notions of how to craft spatiotemporal orchestration. Audio-objects can be staged in new ways that may include 3D rotation, azimuthal tilt, or the re-recording of acoustic indirect sound with multi-mic arrays and ambisonic microphones. In 3D music production the precise design of relationships, perspectives, properties and attributes within the dimensions, SW, EW, ASW, AEW and LEV can strongly determine spatiotemporal orchestration. These concepts have inspired this author to reimagine the metaphoric 7.1.4 3D Atmos sound-box to reveal *dimensional zones* in the Dolby Atmos renderer.

It is specifically intended for music creators. **Figure 10** illustrates this author's music production Dolby Atmos sound-box for music recording and production.



**Figure 10:** This author’s reimagined virtual Dolby Atmos sound-box for music.

### 6.11 Natural Sound-Box Experience Ingrains a Palpable Sense of Spaciousness

The impression of naturalness and spaciousness was studied by Pfanzagl-Cardone and Höldrich in 2008. In their tests ( $r$ ) determines the measurement of spatial impression. Naturalness and spaciousness correlates to the proportional balance between indirect and direct sound energy balance ( $r=0.51$ ). Pfanzagl-Cardone finds that five attributes provide the highest correlation to listener preference. These are naturalness ( $r=0.84$ ), balance ( $r=0.69$ ), spatial impression ( $r=0.68$ ), sound colour ( $r=0.67$ ), and width ( $r=0.65$ ) (Pfanzagl-Cardone, 2023). He found that sound colour provides the strongest correlation to spaciousness, followed by localization and width. In a natural concert hall experience the subject-listener’s proxemic distance to the stage changes perception of the five Super Components. In association with the theatrical concepts of onstage, offstage, upstage, center stage, downstage, and medial center line,

Hall's theory of proxemic zones (intimate, personal, social, and public) can functionally guide the 3D creator toward creation of natural, virtual or abstract spatial environments.

## 6.12 William Moylan's Concept Map and Framework

This author assembles and organizes the components of 3D music spatiality with William Moylan's conceptual framework as a role model.

In Chapter 8 of Dr. Moylan's seminal book about recording analysis, William Moylan focuses on *The Illusion of Space as an Element of Recording*. He provides a framework of stereo concepts, vocabulary, and terminology, which this author uses to facilitate, adapt, and organize a new Concept Map for 3D music. "Spatial properties bring the track to life for the listener; the listener accepts the virtual reality as part of the context and expression of the track—part of what makes it unique" (Moylan, p.287). "Spatial properties are aesthetically central to the record's content and it's expression" (ibid, p.288). "The listener's perspective is used to calculate and define the spatial locations of sounds, and to understand the qualities of source host environments" (ibid, p.289). "This conceptual location is the listener's 'point of audition,' ... [a term used] to identify and locate an unchanging listener's position from which the track is observed" (ibid, pp.287-289). Moylan determines the following three categories of spatial properties: 1) Angular direction and width; 2) Span of distance location and depth; and 3) Dimensions of the environment. His stereo Concept Map organizes: 1) Levels of perspective; 2) Spatial properties; and 3) Variables/Attributes; to describe: A) Individual Sound Source; B) Composite Texture; and C) Overall Sound. See **Table 7**: Moylan's Concept Map. See **Table 8**: Adapted 3D Concept Map, [Chapter 6.13](#).

### 6.12.1 Moylan's Concept Map

| 1) Levels of Perspective           | 2) Spatial Properties                                                                                   | 3) Variables, Attributes                                                                                                                                                                                                                                                                                                                                                                               |
|------------------------------------|---------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <u>A) Individual Sound Source.</u> | <p>Lateral, Horizontal Location.</p> <p>Distance Location.</p> <p>Host Environment of Sound Source.</p> | <p>-Source position, image size, angular trajectory from listener, phantom image.</p> <p>-Source position, depth of image, distance from listener location, Aural image.</p> <p>-Size of enclosed space, patterns, and timings of reverberation characteristics (duration, density, dynamic contour), frequency content of the environment and reverberation, ambience, spaciousness, aural image.</p> |
| <u>B) Composite Texture.</u>       | <p>Sound Stage.</p> <p>Simultaneous environments, space within space.</p>                               | <p>-Lateral and distance positions in aggregate, Relationships and interrelationships of source positions, distance may include depth from source host environment.</p> <p>-Relationships between host environments of all sources, relationships of each host environment to the holistic environment.</p>                                                                                            |
| <u>C) Overall Sound.</u>           | <p>Sound stage dimensions.</p> <p>Holistic environment, space within space.</p>                         | <p>-Boundaries of source locations, left-to-right width, and front-to-back depth of an all-inclusive, overall staging area.</p> <p>-Overall environment (spatial 'place') of the record, relationship of the holistic environment to the aggregate of host environments of the sound stage.</p>                                                                                                        |

**Table 7:** Moylan's Concept Map.

## 6.13 The 3D Concept Map: A Thesaurus of Concepts, Vocabulary, and Terminologies

| 1-Levels of Perspective                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 2-Spatial Properties                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 3-Variables/Attributes                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <u>A—Individual sound source staging (direct or indirect)</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | <u>A—Individual sound source staging (direct or indirect)</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | <u>A—Individual sound source staging (direct or indirect)</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| <p><b>1) Direct sound &amp; indirect sound radiosity:</b></p> <p><b><u>Radiosity, Loudness delta:</u></b><br/>Perceived or measured difference between direct sound and environmental nature, involving proxemics, dynamic contour, and spectral contour, to indicate realism or surrealism.<br/>(The listening position in a 3D speaker sphere must be correlated.)</p> <p><u>Proxemics:</u> (Hall 1963)<br/>Distance</p> <p><b>Intimate zone</b></p> <ul style="list-style-type: none"> <li>• Very close to the listener (i.e. touching distance).</li> <li>• No intervening musical material.</li> </ul> | <p><b>1) Direct sound &amp; indirect sound radiosity:</b></p> <p><b><u>Dimension &amp; sound field:</u></b><br/>A—Mono<br/>B—Stereo<br/>C—Mid/side<br/>D—Multichannel<br/>E—Ambisonic (3D M/S)</p> <p>Source Width (SW)<br/>Apparent Source Width (ASW)<br/>Ensemble Width (EW)<br/>Apparent Ensemble Width (AEW)<br/>Listener Envelopment (LEV)</p> <p><u>Proxemics:</u><br/>Persona/environment</p> <p><b>Intimate zone</b></p> <ul style="list-style-type: none"> <li>• Persona set in front of environment.</li> <li>• Normally high degree of separation between persona and environment.</li> <li>• Vocal placed at front of sound-box and abutting the boundary of the sound-box.</li> </ul> | <p><b>1) Direct sound &amp; indirect sound radiosity:</b></p> <p><b><u>Spatial ethnography:</u></b><br/>Sonic environmental nature representing cognitive image schema indicating understanding of myriad indoor/outdoor sonic environments associated with nature, cultures, poetry, and mediated genres.</p> <p>Spectral and timing response characteristics, energy level, timbre, density, timing patterns in relation to music tempo.<br/>Bespoke spatial effect created by playing back music in a different environment from which it was originally recorded, and then separately recording a new spatial nature.</p> <p><u>Proxemics:</u><br/>Articulation of persona</p> <p><b>Intimate zone</b></p> <ul style="list-style-type: none"> <li>• Close range whisper.</li> <li>• Clarity of articulatory vocal sounds (coughs, breath intake).</li> <li>• Lyrical content suggests intimacy/potential physical contact and addresses interpersonal relationship between two people.</li> </ul> |

| <u>1-Levels of Perspective</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | <u>2-Spatial Properties</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | <u>3-Variables/Attributes</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>Personal zone</b></p> <ul style="list-style-type: none"> <li>• Close to listener (within arm's length).</li> <li>• Possibility of intervening musical material.</li> </ul> <p><b>Social zone</b></p> <ul style="list-style-type: none"> <li>• Medium distance from the listener.</li> <li>• Intervening musical material.</li> </ul> <p><b>Public zone</b></p> <ul style="list-style-type: none"> <li>• Sizable distance from listener.</li> <li>• High degree of intervening musical material.</li> </ul> | <p><b>Personal zone</b></p> <ul style="list-style-type: none"> <li>• Persona in front of environment.</li> <li>• Still a certain degree of separation but less than in [the] intimate zone.</li> </ul> <p><b>Social zone</b></p> <ul style="list-style-type: none"> <li>• Persona within the environment.</li> <li>• Little separation and more integration.</li> <li>• Vocal placed within the center of the sound-box.</li> </ul> <p><b>Public zone</b></p> <ul style="list-style-type: none"> <li>• Persona engulfed and towards rear boundaries of the environment.</li> <li>• High degree of integration.</li> <li>• Vocal towards rear of sound-box.</li> </ul> | <p><b>Personal zone</b></p> <ul style="list-style-type: none"> <li>• Soft to medium vocals.</li> <li>• Less clarity of vocal sounds.</li> <li>• Lyrical content addresses two to three people.</li> </ul> <p><b>Social zone</b></p> <ul style="list-style-type: none"> <li>• Medium to loud vocals.</li> <li>• Few, if any vocal sounds heard.</li> <li>• Lyrical content addresses small/medium group of people.</li> </ul> <p><b>Public zone</b></p> <ul style="list-style-type: none"> <li>• Full, loud vocals shout/semi-shout.</li> <li>• No vocal sounds heard.</li> <li>• Vocals address large group.</li> </ul> |
| <p><u>Microphone placement &amp; pattern:</u></p> <p>A—<b>Close</b>= dry, denatured</p> <p>B—<b>Near</b>=short &amp; long Hass effect, early reflections</p> <p>C—<b>Mid</b>=early reflection + reverberation</p> <p>D—<b>Far</b>=Pre-delay and long reverb duration</p>                                                                                                                                                                                                                                         | <p><u>Relationship, Perspective:</u></p> <p>Lateral or Contralateral</p> <p>Clockwise or counterclockwise</p> <p>Un flipped (↔) L-R</p> <p>Flipped (↑) (Front-back)</p> <p>Angle (0 to 360 degrees)</p> <p>Amount of spatialization captured with microphones.</p> <p>Circular inward facing</p>                                                                                                                                                                                                                                                                                                                                                                      | <p><u>Abstraction:</u> Low, Mid, High</p> <p><u>Characteristics:</u></p> <p>Variant combinations of small to large sizes of 3D cartesian (XYZ) enclosed to open-space patterns, timings, and directions of reflections, echo.</p> <p>Reverberation characteristics of duration, density, spectrum, and dynamic contour. Degree of spaciousness and natural decorrelation in an aural image created by balancing the direct source, staged, and recorded in a specific relationship with its natural environment. A recordist determines attributes of realism and/or surrealism with the</p>                            |



| <u>1-Levels of Perspective</u>                                                                                                                                                                                                                                                                                                                                                            | <u>2-Spatial Properties</u>                                                                                                                                                                                                                       | <u>3-Variables/Attributes</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>2) Impact, intensity, and intelligibility:</b></p> <p><i>Distance:</i> Energy density, spectrum, radiation, and plane.</p> <p>A Dry—(<i>ff</i>) full spectrum<br/> B Near—(<i>f</i>) full spectrum<br/> C Mid—(<i>mf</i>) filtered spectrum<br/> D Far—(<i>pp</i>) filtered spectrum</p> <p><i>Tonal Balance:</i><br/> A-Flat<br/> B-Smile<br/> C-Tilt<br/> D-Vintage/Era/Genre</p> | <p><b>2) Impact, intensity, and intelligibility:</b></p> <p><i>Distance:</i> wet vs. dry</p> <p>A—Dry<br/> B—Near<br/> C—Mid<br/> D—Far</p> <p><i>Tonal Balance:</i><br/> A-Flat<br/> B-Smile<br/> C-Tilt<br/> D-Vintage/Era/Genre</p>            | <p>selection of room, performer location, microphone, placement, pick-up pattern, and sound field.</p> <p><b>2) Impact, intensity, and intelligibility:</b></p> <p><i>Dolby Audio Renderer (DAR)-binaural options:</i></p> <p>A—Off<br/> B—Near<br/> C—Mid<br/> D—Far</p> <p><i>Localization:</i><br/> A—Single speaker<br/> B—Multiple speaker zone<br/> C—Multiple speaker array<br/> D—All speakers</p>                                                                                                                                                                                                       |
| <p><b>3) Pitch &amp; Height:</b><br/> (azimuth, spectral, vertical)</p> <p><i>Horizontal tilt</i></p> <hr/> <p><i>Spectral (Z) positioning:</i></p> <p>Height plane<br/> (See Smalley, Spatiomorphology)</p> <p>Standard plane</p> <p>Sub (LFE)</p>                                                                                                                                       | <p><b>3) Pitch &amp; Height:</b><br/> (azimuth, spectral, vertical)</p> <p><i>Nadir to Zenith</i></p> <hr/> <p><i>Spectral framing:</i></p> <p>Above 4K</p> <p>80 Hz to 4K</p> <p>20-80 Hz<br/> Combination of hearing and somatic induction.</p> | <p><b>3) Pitch &amp; Height:</b><br/> (azimuth, spectral, vertical)</p> <p><i>Music sphere: +/- 45-degrees</i></p> <hr/> <p><i>Spectral orchestration:</i></p> <p>High frequencies, instruments, harmonics, piccolo/flute, cymbals, strikes, plucks, transients, sibilance, special effects.</p> <p>Delay &amp; reverberant dimension and air of host environment bring palpability to the aural image. Size and body of sound sources and host environments.</p> <p>Low frequency effects, and bass managed low frequencies provide an omnidirectional correlated monophonic foundation to the aural image.</p> |

| <u>1-Levels of Perspective</u>                                                                                                                                                                                                                                                                                                                                                                                     | <u>2-Spatial Properties</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | <u>3-Variables/Attributes</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                                                                                                                                                                                                                                                                                                                                                                    | 0-20 Hz<br>Infrasound that is induced somatically.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| <u><b>B—Composite soundstage texture:</b></u><br><br><b>4) Spatial aggregate:</b><br><br>(Natural and virtual indirect contiguous sources)<br><br><i>Non-contiguous simultaneous environments, space within a space.</i><br><br>[See Smalley, Spatiomorphology]<br><br><br><u>Energy:</u><br>RT <sub>60</sub> , distance, energy of early reflections, spaciotemporal tonal character, colouration of reflections. | <u><b>B—Composite soundstage sonic ethnography:</b></u><br><br><b>4) Spatial aggregate:</b><br><br>(Natural and virtual indirect sources)<br><br><i>Natural and virtual instruments with their spaces</i><br><i>Simultaneous environments, space within a space.</i><br><br><i>Listener Envelopment (LEV)</i><br><br>[See Rumsey, Kuttruff]<br><br><br><u>Decorrelation vs Correlation:</u><br>Delay, flange, phase, rotation<br>Interaural cross feed character.<br>Size, damping, and liveliness, diffusion, absorption, colouration of reflections, clarity. | <u><b>B—Composite soundstage spatial image size:</b></u><br><br><b>4) Spatial aggregate:</b><br><br>(Natural and virtual indirect sources)<br><br><i>Channel variants</i><br>A—1.0 mono<br>B—2.0 Stereo<br>C—3.1 L, C, R, LFE<br>D—4 Quad<br>E—5.1 Surround Sound<br>F—7.1<br>G—7.1.2-4-6.<br>H—Variable “spatial aggregate,” image rotation, vertical ascension/descension.<br>J—Egocentric fixed point of audition.<br>K—Dynamic (unfixed) point of audition via headtracking. |
| <u><b>C—Overall sound character:</b></u><br><br><b>5) Holistic aggregate: Spatial Impression</b><br>(Kuttruff, 2016)<br><br>The aggregate characteristics of all direct sound sources combined with their indirect spatial attributes involving direct sound, direct sound staging, host environments.                                                                                                             | <u><b>C—Overall sound stage dimension:</b></u><br><br><b>5) Holistic aggregate: Spatial Impression.</b><br><br>Natural staging of ensemble involving all holistic environments and virtually cascaded spaces. Space within Space and Spaces within Spaces. Spatial zones may                                                                                                                                                                                                                                                                                    | <u><b>C—Overall positional boundaries:</b></u><br><br><b>5) Holistic aggregate: Spatial Impression.</b><br><br>Boundaries of source locations, left to right width, front to back depth, bottom to top height of an all-inclusive, staging area.<br>Overall ethnographic environment (spatial ‘place’) of                                                                                                                                                                        |

| <u>1-Levels of Perspective</u>                                                                                                                                                                                                                                                                                                                                                                          | <u>2-Spatial Properties</u>                                                                                                                                                                                                                                                                                         | <u>3-Variables/Attributes</u>                                                                                                                                                                                                                                                                                                              |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Recorded within direct sounds, virtual instruments, and their virtual environments.                                                                                                                                                                                                                                                                                                                     | involve myriad sub-ensembles staged within dimensional boundaries that can be fixed or dynamic and occur within the overall soundbox.                                                                                                                                                                               | the recording. Relationship of a source's holistic environment to the aggregate of all other natural and virtual environments in the soundbox.                                                                                                                                                                                             |
| <u>Audio-objects, reference frame:</u><br><i>Egocentric:</i> subject-object<br><i>Allocentric:</i> object-object<br><i>Cartesian:</i> XYZ<br><i>Planar:</i> Azimuth, Elevation, and Distance<br>(Commercial immersive music is exclusively an egocentric subject-object perspective)                                                                                                                    | <u>Location, reference frame:</u><br><i>Egocentric:</i> subject-object<br>Contralateral theatre perspective: object-subject<br><i>Allocentric:</i> object-object<br><i>Cartesian:</i> XYZ<br><i>Planar:</i> Azimuth, Elevation, and Distance                                                                        | <u>Static/dynamic positions:</u><br><br>Lateral, medial, vertical, rotational, anterior, posterior, superior, inferior, Static, Morph, Dynamic<br>Spatial multi-speaker zones<br>Lateral, egocentric subject-to-object, contralateral, egocentric object-to-subject.                                                                       |
| <i>Size, span, and scale:</i><br>A—Discrete=0%<br>B—Phantom center=50%<br>C—Periphonic=100%<br>D—Correlation & phase                                                                                                                                                                                                                                                                                    | <i>Size, span, and scale:</i><br>A—Discrete=0%<br>B—Phantom center=50%<br>C—Periphonic=100%<br>D—Correlation & phase                                                                                                                                                                                                | <i>Size and direction:</i><br>A—Single speaker<br>B—Between 2 speakers<br>C—Between all speakers<br>D—Correlation & phase                                                                                                                                                                                                                  |
| <u>Theatre perspectives:</u><br>Venue size<br><br>1) Arena-Stadium<br>2) Concert hall<br>3) Black-box theatre<br>4) <u>Proscenium or open stage</u><br><br>• Lecture hall<br>• Multi-purpose auditorium<br>• Shoe box concert hall<br>• Theatre<br>• Playhouse<br>• Worship space<br><br>5) Theatre in the round<br>6) Thrust<br>7) Traverse/Runway<br>8) Music club/bar<br>9) Home studio/concert room | <u>Theatrical staging zones:</u><br>Reverb time in seconds<br>1) Large: 2.2 — 10<br>1) Medium: 1.5 — 2.2<br>2) Small: .03 — 1.5<br>3) Infinite: 10 +<br><br><br>0.4 — 1.2<br>1.3 — 1.9<br>1.5 — 2.0<br>.03 — 1.5<br>.03 — 1.5<br>1.3 — 1.9<br><br><br>.03 — 1.5<br>1.5 — 2.0<br>.03 — 1.5<br>.03 — 1.5<br>.03 — .05 | <u>Static theatrical zones</u><br><br>▪ Medial centerline upstage (distal)<br>▪ Medial centerline, downstage (proximal)<br>▪ Horizontal stage left.<br>▪ Horizontal stage right.<br>▪ Offstage right rear.<br>▪ Offstage right rear.<br>▪ Offstage right front top.<br>▪ Offstage<br>▪ Offstage right rear top<br>▪ Offstage left rear top |

| <b><u>1-Levels of Perspective</u></b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | <b><u>2-Spatial Properties</u></b>                                                                                                                                                                                                                                                                                                                                                 | <b><u>3-Variables/Attributes</u></b>                                                                                                                                                                                                                                                                                                                                                                                              |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><u>Contralateral object-to-subject theatrical staging zones:</u></p> <ol style="list-style-type: none"> <li>1) <u>Upstage</u>: The distal area on the stage furthest from the audience.</li> <li>2) <u>Downstage</u>: The proximal area of the stage closest to the audience.</li> <li>3) <u>Stage Right</u>: The area of the stage to the performers left when facing the audience. Inverted Egocentric</li> <li>4) <u>Stage Left</u>: The area of the stage to the performers right when facing the audience. Inverted Egocentric.</li> <li>5) <u>Center line</u>: An imaginary reference line on the playing area that indicates the exact center of the stage, travelling from up to downstage.</li> <li>6) <u>Offstage</u>: The area surrounding the playing space not visible to the audience. Typically, this refers to spaces accessible to the performers but not the audience, such as wings, crossovers, and vomitorium pass ways. In Atmos Offstage includes rear surround and rear top.</li> <li>7) Onstage</li> </ol> <p><u>Virtual perspectives and surreal spatial distortions.</u></p> | <p><u>Indirect sound zone/stage:</u><br/>Natural or Virtual<br/>Speaker zone and radiation.</p> <ol style="list-style-type: none"> <li>1) C, RT, LT, L-R</li> <li>2) R, C, L</li> <li>3) R, RS, RT, C</li> <li>4) L, LS, LT, C</li> <li>5) C, LT, RT</li> <li>6) RS, RT, LS, LT</li> <li>7) L, C, R</li> </ol> <p><u>Virtual perspectives and surreal spatial distortions.</u></p> | <p><u>Dynamic, random, static: spatial zonal properties</u><br/>Volume, spectrum, zone boundary, radiation.</p> <p>Volume, Spectrum, Boundary</p> <p>Volume, Spectrum, Boundary</p> <p>Volume, Spectrum, Boundary</p> <p>Volume, Spectrum, Boundary</p> <p>Volume, Spectrum, Boundary</p> <p>Volume, Spectrum, Boundary</p> <p><u>Spatial zones:</u><br/>Medial, Lateral, Horizontal, Vertical, Rotational, onstage, offstage</p> |

| <b><u>1-Levels of Perspective</u></b>                                          | <b><u>2-Spatial Properties</u></b>                                                                                                                                                                                            | <b><u>3-Variables/Attributes</u></b>                                                                                                                                                                                                                                                                                                            |
|--------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <u>Energy and/or spectral split:</u><br>Transient location.<br>Tonal location. | <u>Delays:</u><br>Taps, pitched, filtered, distorted, feedback, reversed, and modulation shaped (LFO); triangle, square, sine, saw up, saw down, random.<br><br>Apparent Source Width (ASW)<br><br>Listener Envelopment (LEV) | <u>Virtual perspectives and spatial distortions.</u><br><br><u>Delay timing:</u><br>In relation to tempo, random, modulated.<br><br><u>Reverb characteristics:</u><br>Pre delay, reverse, radiosity, inversion, duration, highly combed, lightly combed, resonance and timbre, density, pitch modulated, filtered, mono, stereo, multi-channel. |

**Table 8:** The 3D Concept Map organized using William Moylan’s framework

The 3D Concept Map is an extensive work sheet and likely inaccessible for practical use. It was necessary to create and took months of head scratching. From it, five essential facets have arisen. The Five-Facet typology involves distilled topics with semantic descriptions for 3D spatial analysis. This typology can be reintroduced within William Moylan’s stereo concept map as way to extend his examination of music spatiality, but in 3D.

The Five-Facets are:

- 1) Direct and indirect sound.
- 2) Impact, intensity and intelligibility.
- 3) Pitch & height.
- 4) The spatial aggregate.
- 5) The spatial impression of the holistic aggregate.

**Table 9**, Chapter 6.13.1 shows Moylan’s Concept Map with The Five-Facet Typology (shown in italics) inserted within category one, “Levels of Perspective” (Moylan, 2020).

### 6.13.1 Moylan's Concept Map Houses the Five-Facet Typology

| 1) Levels of Perspective                                                                                                                                                                                       | 2) Spatial Properties                                                                                   | 3) Variables, Attributes                                                                                                                                                                                                                                                                                                                                                                               |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>A) Individual Sound Source.</p> <p><u>Five-Facet Typology:</u></p> <p><u>1) Direct and Indirect sound:</u></p> <p><u>2 Impact, Intensity, and Intelligibility:</u></p> <p><u>3) Pitch &amp; Height:</u></p> | <p>Lateral, Horizontal Location.</p> <p>Distance Location.</p> <p>Host Environment of Sound Source.</p> | <p>-Source position, image size, angular trajectory from listener, phantom image.</p> <p>-Source position, depth of image, distance from listener location, Aural image.</p> <p>-Size of enclosed space, patterns, and timings of reverberation characteristics (duration, density, dynamic contour), frequency content of the environment and reverberation, ambience, spaciousness, aural image.</p> |
| <p>B) Composite Texture.</p> <p><u>4) Spatial Aggregate:</u></p>                                                                                                                                               | <p>Sound Stage.</p> <p>Simultaneous environments, space within space.</p>                               | <p>-Lateral and distance positions in aggregate, Relationships and interrelationships of source positions, distance may include depth from source host environment.</p> <p>-Relationships between host environments of all sources, relationships of each host environment to the holistic environment.</p>                                                                                            |
| <p>C) Overall Sound.</p> <p><u>5) The Spatial Impression of the Holistic Aggregate:</u></p>                                                                                                                    | <p>Sound stage dimensions.</p> <p>Holistic environment, space within space.</p>                         | <p>-Boundaries of source locations, left-to-right width, and front-to-back depth of an all-inclusive, overall staging area.</p> <p>-Overall environment (spatial 'place') of the record, relationship of the holistic environment to the aggregate of host environments of the sound stage.</p>                                                                                                        |

**Table 9:** Moylan's Concept Map housing the Five-Facet Typology.

**Table 10**, Chapter 6.13.2 illustrates the Five-Facet Typology with descriptions of the substantial components. The typology assembles essential concepts, semantic vocabulary and terminologies that can be utilized for discussion.

### 6.13.2 The Five-Facet Typology with Semantic Descriptions of the Substantial Components

|   | Facet                                                                   | Description of the substantial components                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|---|-------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | <a href="#"><u>Direct and Indirect sound</u></a>                        | Denatured direct sound; bonded radiosity between direct and indirect natural sound; spectral filtering; energy, and timing delta between direct and indirect sound to indicate proxemics; radiation into other channels; ecology of indoor environmental sound, and outdoor natural sound. Source Width (SW); Ensemble Width (EW); Apparent Source Width (ASW); Listener Envelopment (LEV); Sound field: mono, stereo, multichannel, ambisonic, and nulled channels, critical listening distance.                                                                                                                                                                                                               |
| 2 | <a href="#"><u>Impact, Intensity and Intelligibility</u></a>            | Transient vs tonal body dynamic range and energy delta; staging; proxemics; direct and indirect sound radiosity; inter-channel radiation; spectral filtering; spatial distortion.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 3 | <a href="#"><u>Pitch &amp; Height</u></a>                               | Staging; High frequency emphasis, direct, and indirect sound radiosity and radiation; frequency; tonal pitch; vertical object location; azimuth from nadir to zenith—limited application in the Dolby Atmos format.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 4 | <a href="#"><u>The Spatial Aggregate</u></a>                            | All reverberations, delays, and indirect sound. Listener envelopment (LEV)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 5 | <a href="#"><u>The Spatial Impression of the Holistic Aggregate</u></a> | Overall listening experience (OLE). All direct and indirect sound attributes focusing on envelopment and sensation of spaciousness; Abstraction, Ensemble staging; SW, EW, ASW, AEW, LEV; interchannel radiation; spatial zones and boundaries; nulled channels; spatial distortions; spectrum filtering; static vs dynamic object locations; inter-aural energy delta; timing; and spectrum difference; correlation versus decorrelation. Descriptions of sonic/acoustic ecology, culture, genre, ethnography, anthropology; music psychology; geometric architecture for indoor spaces and outside environments; persona and song meaning connoted by spatial proxemics in relation to the point of audition. |
|   |                                                                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |

**Table 10:** The Five-Facet Typology.

### 6.14 Improved Perception with Ear Training Exercises

Developing a heightened level of aural perception is aided by closing the eyes and ear training with the following exercises which can easily be made. The reader can access (via Apple Music or the 2L Music Store, see section 18, websites) recordist Morten Lindberg’s natively recorded *Voice Navigation 7.1.4 Immersive Audio* channel identification audio file.

On each channel, Morten Lindberg announces: L) “This is left front; R) This is right front; C) this is center,” etc., and he places a 100 Hz tone in the LFE channel.

This file can be adapted into three simple ear training exercises for personal use. Use Rogue Amoeba Loopback, or Audio Movers Omnibus to route 7.1.4 audio from the Apple Music desktop player or the Dolby Audio Renderer directly into 12 mono audio tracks in a DAW (Stagle, 2022) on an Apple computer with an M series processor. Then select, create, and mute regions to make the following three exercises.

Exercise 1: The Holistic Aggregate — (no muting needed)

Exercise 2: The Direct Sound — (mute all tracks except the voiced channel announcements)

Exercise 3: The Indirect Sound — (mute the direct sound voiced channel announcements)

## 6.15 The Analysis Process

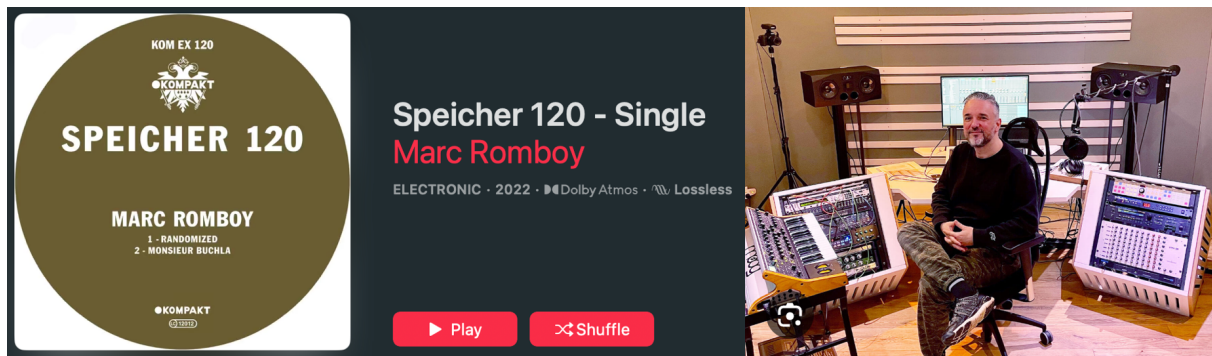
As previously stated in [Chapter 6.3.1](#), seven principles set the stage for 3D spatial analysis of recorded music. If necessary, please use the provided hyper link for review.

The analyst can follow the path of the Five-Facet Typology to yield valuable top-level subjective observations that evaluate: 1) Direct and Indirect sound; 2) Impact, Intensity and Intelligibility; 3) Pitch and Height; 4) The Spatial Aggregate; and 5) The Spatial Impression of the Holistic Aggregate. These analysis points can examine elements as 1) positional relationships, 2) cognitive perspectives, 3) proportional properties and 4) character attributes.

Channel-to-channel deconstruction is an essential practice for 3D spatial analysis, and it is aided with a high functioning center-desk monitor controller. This author uses Merging Technologies Anubis in his Hyper Near-Field Tiny Studio and his practice with the analysis synergy has garnered five examples to be presented in the following Chapter 6.16.



## 6.16 Analysis 1: “Randomized,” *Speicher 120*



**Figure 11:** Marc Romboy in his studio.

### 1) **Analysis Scope 1:** Scene by Scene Channel Deconstruction

Audio Scene One: 0:00 to 0:16 (ITPRA—Imagination phase-pre-outcome)

- A kick drum at 125 bpm, pulsing with the 8<sup>th</sup> note bass (E1), enters in the front L/R speakers. Only direct sound is heard. No LFE is used. The location is fixed.
- All other speakers contain low level indirect sound that is spectrally filtered to indicate a distant proximity from the direct audio. This tells the listener that the environmental context is in 360 degrees, even though the predominant kick drum motor (of the music) is staged anteriorly using the left and right speakers.

Audio Scene Two: 0:16 to 1:08 (ITPRA—Tension phase-pre-outcome)

- Direct and indirect sound build intensity together. The listener is arrested by the music and spatiality. The brain proceeds towards a flow state. A roto-tom (pitched to an A indicating E7sus4 harmony) is deeply spatialized (strong decorrelation) using a long stereo ping-pong delay and elite reverb that diminishes over 7 beats. Direct and indirect sound are premixed to about a 50-50 proportional blend, heard in the front top-height L/R speakers and the front mid-plane L, C, R speakers. Direct and indirect sound is a locationally unfixed spatial zone that rotates clockwise around the listener every two bars. The mind is occupied with spatiality in locked meter with the metronomic kick drum pulse.

Audio Scene Three: 1:08 to 1:59 (ITPRA—Predictive phase-post outcome)

- A deeply reverberant sonic spatial event strikes as the kick drum pulse pauses for one bar. A new 8<sup>th</sup> note, pulsing synth part begins on the dominant 7<sup>th</sup> of the chord. 16<sup>th</sup> note low drum toms accent the last two 16<sup>th</sup> notes of beat 4. They are

placed in the left-right side speakers and are lightly spatialized with early room reflections. Tension continues to build as new musical characters with differing properties and attributes indicate spaces within spaces. The listening place is virtual because what is perceived is not possible in a natural environment. This 3D music experience forms new cognitive image schemata in the brain of the first-time listener.

Audio Scene Four: 1:59 to 2:12 (ITPRA—Reaction response-post outcome)

- The last note of the 8<sup>th</sup> note pulsing synth is sent into a very long reverb which (twice) quickly rotates clockwise around the listener. This transitional moment is highly mesmerizing. It's an auditorial spectacle that transfixes the listener in a flow state. As the rotation stops and the reverberance decays, Roland 808 handclaps that have no spatial nature begin on beats 2 and 4 in the front left and right speakers while a foot-close hi-hat starts on the 8<sup>th</sup> note of each quarter note. It is located only in the center channel and has no spatial context.

Audio Scene Five: 2:13 to 2:35 (ITPRA—Appraisal of the past sonic event-post outcome)

- The bed layer of the holistic aggregate is now established and the track percolates in spatial stasis affording the listener time to process what has happened so far and then imagine what may come next.

Audio Scene Six: 2:35 to 2:50 (ITPRA—Tension phase)

- A breakdown begins; one by one kick and claps drop out leaving only the 8<sup>th</sup> note pulsing synth on the dominant 7<sup>th</sup> degree, with the bass note on the tonic.

Audio Scene Seven: 2:50 to 3:20 (ITPRA—Appraisal of the breakdown while Imagination restarts)

- Kick drum re-enters on ¼ notes and a more complex hi-hat pattern enters in the front left and right speakers. The foot-closed hi-hat re-enters in the center speaker.

Audio Scene Eight: 3:20 to 4:04 (ITPRA—Hypnotic dance flow state)

- A two-note spatialized motif appears as the melody. The quarter note motif starts on beat one. The notes are E3 and B3. This motif slowly rotates counterclockwise and is heard in the standard mid plane (ear-level) and top height speakers.

Audio Scene Nine: 4:04 to 4:38 (ITPRA—Imagination phase—tension phase—predictive phase)

- Another breakdown begins with the kick drum dropping out, tempo slows down. Listener envelopment increases on the lateral sides and in the rear speakers.

The primary motor of the music is replaced by a 3D spatial experience, suggesting the listener is only hearing the indirect sound of an indoor venue. Synth filters gradually close down via low pass filtering, Indirect reverberation dies down and the track dissolves almost to silence.

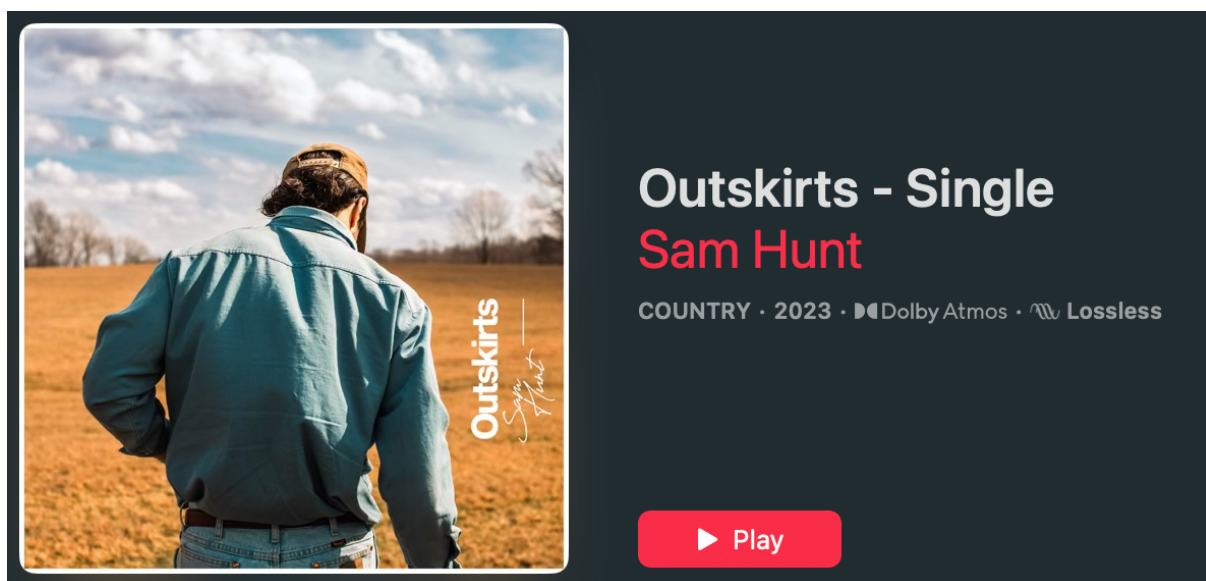
Audio Scene Ten: 4:38 to end (ITPRA—Imagination, Tension, Predictive, Reaction, Appraisal)

- The bass note restarts meanwhile the filter resonance opens up introducing distortion. The LFO speeds up to a fast pulse, and the note begins to rotate around the listener. The kick drum re-enters. A high-frequency analog and FM synth noise mashup occurs in all the height speakers (canopy). The rotation of space and place in this section is an extraordinary 3D spatial experience.

General Traits:

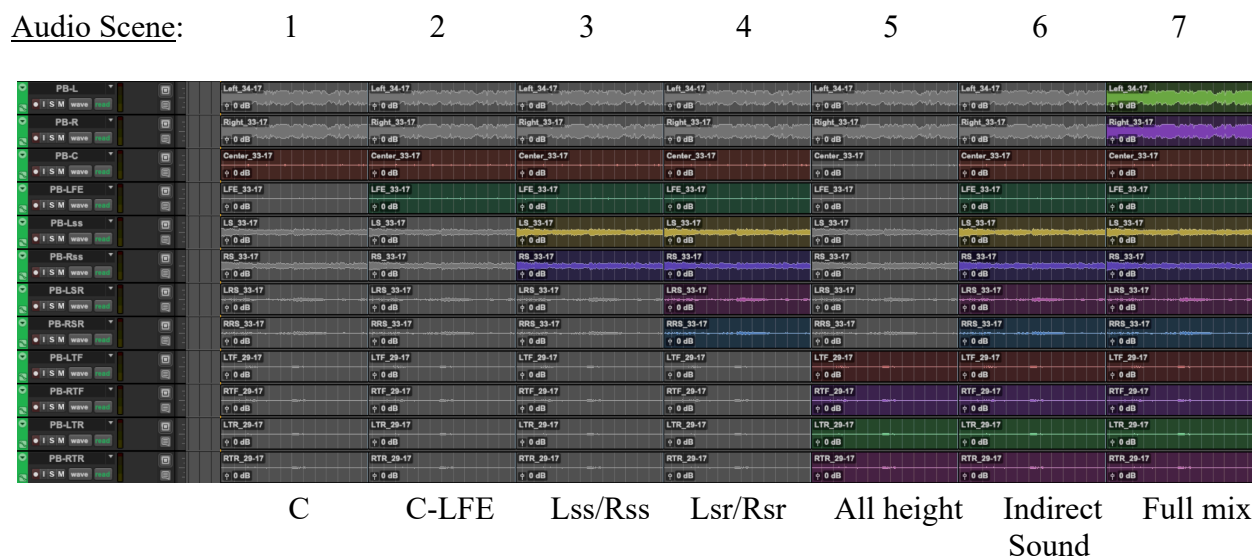
- There is excellent use of fixed and impactful direct sound heard in the anterior front. This is a counterpoint to the spatial distortions/abstractions involving spaces-in-spaces rotation. Distal and proximal audio-objects co-exist together. The sensation of space and listener envelopment is strong. Source width (SW) varies and audio-object sources involve direct and indirect sound created by cascaded multichannel delays and reverbs.

### 6.17 Analysis 2: “Outskirts,” Sam Hunt



**Figure 12:** “Outskirts,” Sam Hunt.

This repeating 4-bar phrase illustrates 7 audio scenes that focus on speaker zone function. It was looped for playback and using channel muting and unmute each phrase adds channels to exhibit the function of six speaker zones and full mix as named below the phrase. Creating a compelling sense of 3D spatiality in popular country music was well-achieved. This mix feels like a traditional stereo mix with lead vocals staged in the front L/R speakers and effected vocals in the Lss/Rss. The sensation of spatiality is strong with a confident use of the lateral side and rear speakers.



**Figure 13:** “Outskirts” channel breakdown.

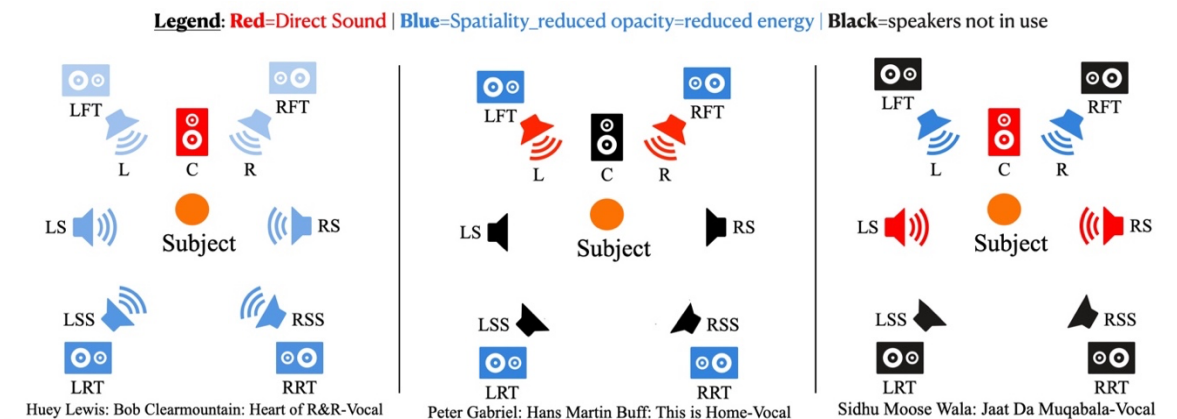
The center channel and LFE are used lightly. This may indicate that the mixer was uncertain about the opportunities afforded by these channels. It is also possible that the mix engineer was directed by the label or artist to use the C and LFE sparingly. In comparison to other Atmos mixes (Peter Gabriel, *i/o*, *In-Side Mix*, 2024) I think the center channel could have been more effectively used. In audio scene 3 the Lss/Rss provides the primary source of Listener Envelopment (LEV). In audio scene 4 the Lsr/Rsr add rear depth with reverb, delay, and offstage

instruments like electric piano and synths. In audio scene 5 all the height channels combine to form a quadraphonic height zone for the dobro, synths and electric guitars with an abundance of reverb and delays. In audio scene 6 the indirect spatial aggregate can be heard. The mixer places the music in an elite concert hall space.

## 6.18 Analysis 3: Vocal Staging, Three Comparisons



**Figure 14:** Vocal Staging, Huey Lewis, Peter Gabriel, Sidu Moose Wala.



**Figure 15:** Vocal staging: 1) Bob Clearmountain: 2) Hans Martin Buff; 3) mixer unknown.

### 6.18.1: “Heart of Rockn’ Roll,” Mixed by Bob Clearmountain

In “Heart of Rockn’ Roll,” the mixer Bob Clearmountain uses the center speaker to maximum advantage and places vocal direct sound there, along with a heartbeat, kick, snare, ride cymbal, picked bass, sequenced synthesizer, harmonica, and clean guitar. All of the sound in the

center channel is dry. As shown in **Figure 13**, indicated by shades of blue, other speakers are used for indirect sound, instrument offstaging, and special effects.

Clearmountain's rationale accounts for a wide range of listening positions in the speaker sphere. No matter where the subject-listener is located in the room the voice remains anchored to the center stage speaker as a *real* image. The source width is narrower than a Left/Right speaker *phantom* center image, but vocal intelligibility is good due to single center speaker "primacy of the first arrived sound," — "The Precedence Haas Effect," which some call "the law of the first wavefront" (Toole, 2018, pp.200-214). This vocal staging places the protagonist on the downstage medial center-line in a personal proximity to the subjective listener. All the other loudspeakers express the spatial environment of a low ceiling sports bar.

Specifically, the Lss/Rss speakers are used for a single repeat slap-back delay that creates an impression of room size, cued by bounced indirect sound from the rear wall.

Drawing upon personal cognitive image schemata from this author's memory, cognitive visualisation conjures a palpable feel for the room size. Sensation of smell also arises to recall the alcohol infused musty social environment of a sport bar. This subjective 3D music experience indicates that other modalities of visceral sensory perception can be triggered by compelling 3D music. Clearmountain creates an Overall Listening Experience (OLE) that is highly convincing. The sense of environmental spatiality is primarily communicated with laterally induced Listener Envelopment (LEV) by employing a delay length subjectively estimated to be about 60 to 80 milliseconds. This spatial property exemplifies exactly what Kuttruff asserts—previously cited in [Chapter 6.8](#).

#### 6.18.2: “This is Home,” Mixed by Hans Martin Buff

In this example Peter Gabriel is staged in the anterior, onstage front left and right speakers. If the listener moves to the left side the vocal emanates from there and the inverse occurs if the listener moves to the right. The lead vocal’s indirect reflected sound is not strong in level, but it is placed in the upper height canopy as indicated by the light blue shading in **Figure 15**. Music producers can observe that “multiple [early] reflections contribute to speech [vocal] intelligibility” (Toole, 2018, p.201). Placing Gabriel’s indirect sound in the canopy increased his vocal intelligibility.

#### 6.18.3: “Jaat Du Maqabla,” Sidu Moose Wala (Mix Engineer Unknown)

A highly unusual example of vocal Source Width (SW) staging is found in Sidu Moose Wala’s song *Jaat Du Maqabala*. Sidu’s direct vocal sound is placed downstage center via the C speaker, and laterally offstage via the Lss/Rss speakers to create a triangular 3-speaker zone. The result puts the singer’s direct sound in front and beside the listener.

Apparent Source Width (ASW) indirect sound fills in the triangular speaker zone with vocal reverb placed onstage in the front horizontal L/R speakers. *Jaat Du Maqabala* presents an effective egocentric listener-to-protagonist perspective using an extra-wide vocal source width via a geometric three-point speaker zone.

This creates a highly abstracted personal proxemic image that delivers strong intelligibility and impact. Artificially placing direct sound [periphonily] in 3 speakers C, Lss, Rss, *Jaat Du Maqabala* successfully breaks away from normalised 3D praxis for popular music lead vocal staging. This example of periphony does not compromise binaural playback.



## 6.19 [Analysis 4](#): “Rocket Man”, Mixed by Greg Penny and Bryan Pennington



**Figure 16:** “Rocket Man,” *Honky Château*, Elton John.

### iTunes Review

*Honky Château* is the pinnacle of Elton John and lyricist Bernie Taupin's obsession with all things American, not least because its themes are funny (“I Think I’m Gonna Kill Myself”) and personalized (“Amy,” a sly paean to young lust). Elton’s small group, augmented with some daredevil electric violin by Jean-Luc Ponty, adapts to everything from the New Orleans carousel music of “Honky Cat” to the Ray Charles-ready “Mellow.” Similarly riding varied currents, the man does some of his best singing here; he’s positively funky on the Band-informed “Susie (Dramas).”

**Figure 17:** *Honky Château*, iTunes review.

In [Chapter 1.6.2.1](#) Greg Penny’s approach to mixing “Rock Man” is described. This author also did an ITPRA analysis of the first two minutes of “Rocket Man”<sup>54</sup> which examines Greg Penny’s intuitive spatiotemporal decisions through the lenses of David Huron’s ITPRA theory. This exemplifies a connection between ITPRA theory, and Penny’s intuitive practice. Seeking an understanding of Greg’s answers to the elemental questions of *What*, *When*, *Where*, and *Why* is the purpose of this analysis. If necessary, please see the [glossary](#) to familiarize with 9.1.6 speaker sphere nomenclature. Please also see [Chapter 6.19.1](#), **Table 11**: ITPRA response analysis of “Rocket Man.”

---

<sup>54</sup> Beyond 2:00 minutes, Greg and Bryan’s spatiotemporal orchestration repeats. Since this analysis locates within the literature review [the first clause of this sentence needs to be re-written], brevity was a requisite.



### 6.19.1 ITPRA Response Analysis of “Rocket Man”

ITPRA is: Imagine; Tension; Prediction; Response; Appraisal.

| Response | When                                                                           | What, Where, and Why                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|----------|--------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| (I)      | 0:00 to 0:20<br>BPM—70.<br>Bar 1,<br><br>Introduction,<br>Verse 1              | <p>Direct vocal sound is in a “personal proximity” (Hall, 1963). It is staged in the center speaker. The horizontal front stereo left and right speakers are critically balanced with indirect sound. Stereo piano is placed in the stereo left and right.</p> <p>A subtle quad plate reverb (indirect sound) is placed in the <a href="#">Lss</a>, <a href="#">Rss</a>, <a href="#">Lrs</a>, <a href="#">Rrs</a>. Spatial staging remains traditional and pays homage to the original 1974 stereo mix. The quad (4-channel) plate reverb has about two seconds of decay. It suggests that the listener is seated in an elite concert hall in public proximity to the lead singer who is located at center stage. The height canopy speakers and LFE are not yet in use.</p>                                                                                                                                                          |
| (T)      | 0:20 to 0:23<br>Bar 5<br>Spatial event                                         | <p>Spatial ‘arousal’ intensifies with the sung word “high.” It teases with echo, creating a spatial sonic event. Located in the upper height canopy, this event introduces the third dimension (height) and triggers the anticipation of other events to come into that spatial zone. It constructs about 3 seconds of tension and release and forms a spatial motif. The echo effect suggests the disorienting effect of being “high.”</p> <p>(Greg Penny is gradually introducing spatial zones and connecting their function to the narrative and the singer’s performance. This exemplifies “how and when” context determines his creative spatiotemporal orchestration).</p>                                                                                                                                                                                                                                                     |
| (P) (T)  | 0:23 to 0:58<br><br><br><br><br><br><br><br>0:30 to :058<br>Bar 9<br>Verse 1.5 | <p>Traditional staging returns (resets) to the audio scene of Verse 1, bar 1. After the spatial sonic event finishes, a loss of sound in the height speaker canopy is sensed because the zone is unfilled. This emptiness creates tension that builds for 35 seconds during which the listener senses ‘intensification.’ It is a state of pre-outcome prediction while remaining in the post-outcome state with contemplation of the prior event meaning.</p> <p>Bass enters in the <a href="#">LFE</a>, <a href="#">Ls</a>, <a href="#">Rs</a>, <a href="#">Lss</a>, <a href="#">Rss</a>, <a href="#">Lrs</a>, <a href="#">Rrs</a>. In Atmos it is unusual to distribute the bass to multiple standard-plane speakers, but as done here, the bass provides a metaphor representing earth’s gravity. It is counterpoint boundary to the flyover of the outer space rocket ship, symbolized by the distorted electric slide guitar</p> |

|         |                                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|---------|--------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|         |                                      | placed in the top height speakers. Predictions about what is to come are forming as the spatial zones gradually get occupied.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| (P)     | 0:58 to 1:13<br>Chorus<br>Bar 17     | <p>Greg Penny refers to this musical moment as the “stun effect,” or perhaps the fulsome use of all spatial zones at once. His technique creates compelling decorrelated listener envelopment (LEV) with no holes in a clean and critically balanced sound-field.</p> <p>Distorted slide guitar is placed above—in the height canopy—with background vocals placed in the lateral and posterior locations. Doubled lead vocal locates in the center speaker and is shouldered into the Ls/Rs speakers. The anterior horizontal staging is wide and includes some indirect plate reverb sound. The primary reverb remains offstage, in the sound-box, beside and behind the listener.</p> <p>The background vocals lift out of the Ls/Rs and fix their location in the Lss/Rss, and Lsr, and Rsr.</p> |
|         | 1:13<br>Bar 21                       | The distorted slide guitar effect starts in the Ltf/Rtf and pans to the Ltr/Rtr speakers symbolizing a rocket ship flying over from front to back.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| (I) (T) | 1:16<br>Bar 22                       | Background singers support the melody by singing the words “Rocket Man,” heard in the Lss/Rss/Lsr/Rsr. This is an impactful moment.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| (R) (A) | 1:20 to 1:27<br>Bar 23-24            | Imagination and tension become active as the spatial zones deactivate and the chorus resolves. Reaction and appraisal stages are now active. The chorus repeats and Penny’s “stun” effect repeats. It has a satisfying impact because it’s a recapitulation. Tension is replaced with positive prediction and satisfying appraisal of the experience so far.                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|         | 1:26 to 2:15<br>Bar 25<br>Chorus two | The distorted slide guitar effect reverses direction and starts in the Ltr/Rtr and pans to the Ltf/Rtf speakers symbolizing a rocket ship returning on its journey in the third dimension of height.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|         | 1:41 to 1:56<br>Bar 29-33            | Repeat with variation.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|         |                                      | Etc.,                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |

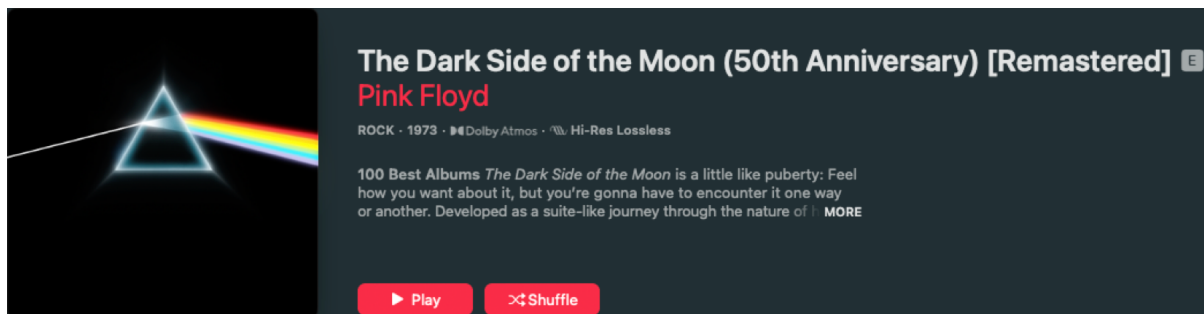
|                                   |                              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|-----------------------------------|------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                   | 1:56 to 4:35<br>Bar 33 to 80 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| “Rocket Man,”<br>Song<br>summary. | 0:00 to 4:35                 | Gregg Penny choreographs listener expectation and emotion using time-based spatial orchestration of X, Y, Z spatial zones. Informed by lyrical context he responds with sonic spatial metaphors that function as symbolic codes of meaning. The slide guitar mimics a spaceship journey that passes above from front to back and back to front. He places the bass in seven speakers to create dimensional separation. This gravity contrasts the slide guitar flyovers. The lead vocal is dry, placed in the center speaker but shouldered into the Ls/Rs to create the sense of extra Source Width (SW). Quad reverb placed in the Lss, Rss, Lrs, Rrs creates an offstage sound-box reverberance that envelopes the listener. Group vocals are placed in personal proximity to the listener using the Lss, Rss, Lrs, Rrs. There are three primary spatial zones to this mix: 1) L, C, R, 2) Lss, Rss, Lrs, Rrs, and 3) Lft, Rft, Lrt, Rrt. “Rocket Man” was inspired by Ray Bradbury’s short story titled “The Rocket Man,” published in his 1951 anthology entitled <i>The Illustrated Man</i> . |

**Table 11:** ITPRA response analysis of “Rocket Man.”

Greg Penny’s spatial recontextualization of “Rocket Man” is highly stimulating. For many people Elton John’s song “Rocket Man” is interpreted as a family song. This author never took the time to explore the songs’ themes of loneliness and separation from loved ones because he did not enjoy the sound of the original mix. To him it was harsh and dimensionally uninteresting. Greg Penny’s Atmos mix of “Rocket Man” changed all of that. This author now appreciates the song and its 3D spatial context. The 3D mix of “Rocket Man” faithfully puts the first time Atmos listener in a mind state of “flow” (Csikszentmihalyi, 2004).

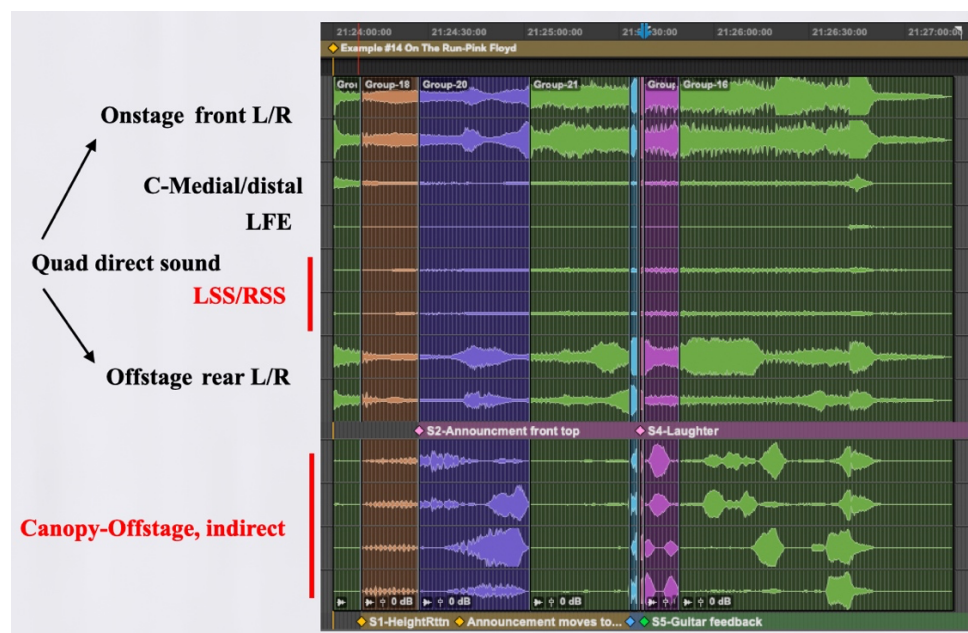
## 6.20 Analysis 5: “On The Run,” Pink Floyd, *The Dark Side of The Moon*

Mixed by James Guthrie (2023).



**Figure 18:** “On The Run”, *The Dark Side of The Moon*, 50<sup>th</sup> Anniversary, Remastered, Pink Floyd.

This analysis will look through three lenses, 1) Channel Usage; 2) Storytelling; and 3) Scene by Scene Activity.



**Figure 19:** “On The Run,” channel usage.

### 6.20.1 Channel Usage

James Guthrie preserved the quad direct sound effect from the original 1973 release by using the onstage front L/R and offstage rear L/R (Lsr/Rsr) speakers. The offstage Lss/Rss and height canopy are used for indirect sound and sound effects. The Center speaker is low in level while it contains the sequenced synthesizer and hi-hat pattern which functions as the compositional motor.

### 6.20.2 Storytelling

“On The Run” is an instrumental song created by David Gilmour and Roger Waters on the album *The Dark Side of The Moon*. The album’s central theme of madness is cultivated from an extensive touring schedule, and it suggests that the band were all starting to go a bit mad as a result.

“On The Run” is a story about the fear of flying and being in large public airport spaces. The 3D mix involves strategic and intuitive use of allocentric object-to-object relationships and egocentric subject-to-object relationships to express the auditory codes of space, place, human experience, human nature and societal class structure. These relationships root in the political economy of music industry fame and the perception of a luxurious “rich and famous” lifestyle, which contrasts the human need for a settled daily life experience to promote mental wellness.

Tension and release are central components in the composition, meaning David Huron’s ITPRA theory can be applied to examine the art and craft of James Guthrie’s spatiotemporal orchestration.

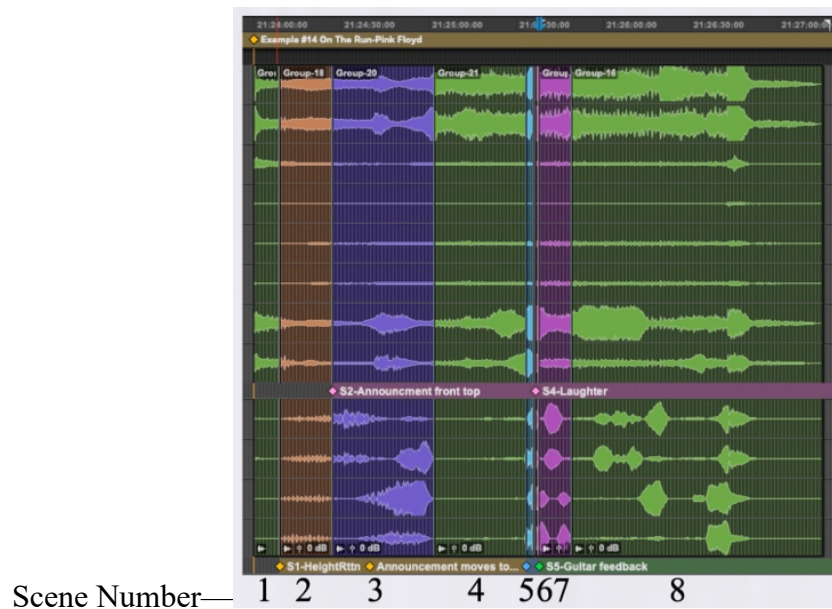
“On The Run” is a chase scene that unfolds in an airport and Guthrie creates a cinematic auditorial 3D environment that places the subjective listener as a disassociated witness to the events.

The listener is located in a subject-object relationship and egocentric perspective with the protagonist (the runner). Strong use of dynamically unfixed proxemic properties in varying proportions of intimate, personal, social and public zones make this audio scene a highly palpable and visceral audio experience.

The acoustic ecology of “On The Run” sounds like what a bystander would naturally experience. With the eyes closed, this virtual experience is compelling, and it reliably transports the Atmos listener to an immersive “flow” state, both in binaural headphone playback and a discrete 7.1.4 speaker system.

James Guthrie took great care to pay homage to the original quad version from 1973 by using the onstage front left and right speakers, and the offstage rear left and right speakers to recreate the quad speaker zone. The upper height canopy speaker zone is reserved for a rotational hoola-hoop effect and a surveillance helicopter. “On The Run” is movie music and sound with no visual cues necessary to communicate the materialities of the scene. Based on this 3D aural experience this author envisions the immense power and extended possibilities provided by the Dolby Atmos format for 3D audio book storytelling and podcasts.

### 6.20.3 Scene by Scene Activity Breakdown



**Figure 20:** Scene by scene activity breakdown.

#### **Scene by Scene Activity Breakdown.**

1. Establishes the 1974 vintage quad effect. A very low-level blending of sound is distributed into other channels to imitate natural radiation of a very large and public closed field space. This scenario begins the ITPRA (pre-outcome) phase of imagination.
2. A disorienting hoola-hoop sound effect in 360-degree rotation enters in the height speaker canopy. This scene audio scene stimulates the imagination phase.
3. Direct and indirect sound balanced as an environmental airport PA announcement is heard which intensifies the pre-outcome tension phase. Gradual dynamic movement from the front to rear signals that the subjective listener is advancing to a new location in the airport atrium. Guthrie's novel use of *dynamic audio-objects* (constantly in motion) pinpoints the allocentric relationship between the running protagonist and the overhead

public address announcement. This is Atmos art and craft at its finest. It is sonically equivalent to a cinematic close-up view of the protagonist as he runs through the congested airport terminal to escape a predatorial antagonist. The antagonist is portrayed by an ominous sounding synthesized helicopter that enters in the Left Top Rear, (Ltr) speaker zone with a gradual filtered-opening synth note, entering in the Right Front Top (Rft) speaker zone. It crosses over-top of the listener in a geometrical X pattern, employing doppler pitch change to signify changes in proximity and speed of movement.

4. A hi-hat and synth ostinato function as the musical motor at 160 bpm. Together they symbolize the protagonist's internal high degree of stress. The ominous synth helicopter suggests that aerial overhead surveillance is a persistent threat. Doppler effects and spectral filtering are routinely employed to indicate rate changes in localization and proximity.
5. The synth pulse accents are motivic sonic events that indicate a close call is being experienced. The protagonist's in-head narratives are voiced, "live for today, gone tomorrow, that's me."
6. Finally, laughter is heard. This author contemplates two explanations for this. Laughter could be from the protagonist who is deliriously mad, or perhaps it's a threatening expression from the gamified attacker as he enjoys the chase.
7. Backwards guitars indicate that the protagonist is psychologically disoriented. ITPRA tension and prediction phases are dynamic and intensely convincing.
8. The final explosion indicates a crashing aerial surveillance vehicle. The high-volume loudness of this sonic event expresses the zenith of the audio story and suggests that the



protagonist's intimate proximity to the crash was close. His final footsteps indicate the post outcome appraisal phase. He gets away safely, and his footsteps dynamically pan to the front right speaker zone, meanwhile the critical balance gradually shifts from direct to indirect sound, indicating an escape into the distance which provides resolution to the conflict.

## 6.21 Episode Review and Summary

Semantic descriptions of the substantial componentry found within 3D spatial music encoded for the Dolby Atmos format are complex, so they require precise but accessible analysis. This author's Three-Part analysis system is called a Synergy. The analyst is encouraged to use the Three-Part Synergy as they feel is best for them. Analyses [6.16](#) to [6.20](#) exemplify how this author applied the synergy. With practice the process can be customized for specific examinations.

**Part-One** proposes a 3D Concept Map for music spatiality that is built from the origins of myriad semantic concept map rationales. William Moylan's approach and framework was selected as a role model to adapt and form this author's 3D Spatial Music Concept Map. It functions as a thesaurus of scenarios, concepts, vocabulary, and terms for reference.

**Part-Two** proposes a Five-Facet Typology. It is distilled from the 3D Concept Map (thesaurus) and houses the seminal concepts, vocabulary, and terms essential for efficient analysis of "what matters most" (Nelson, 2022, p.2). Focused on first impressions of 3D spatial music it points the analyst to examine, 1) positional relationships, 2) cognitive perspectives, 3) proportional properties, and 4) character attributes of, 1) Direct and Indirect Sound, 2) Impact,

Intensity and intelligibility, 3) Pitch and Height, 4) The Spatial Aggregate, and 5) The Spatial Impression of the Holistic Aggregate.

**Part-Three** proposes ear training exercises that aim to tune the hearing system for improved perception of 3D spatial localization as a way to prepare the mind for induction of 3D music spatiality in the Dolby Atmos format.

There are seven precursory principles necessary for 3D music analysis to begin. Please see [Chapter 6.31](#) and [Chapter 6.15](#). Also worthwhile are David Byrne's and David Hurons theories about context and expectation. Please see [Chapter 3.1](#). The Synergy aims to be subjectively descriptive, not prescriptive or absolutist.

## 6.22 Episode Conclusion:

This author's analysis Synergy assembles seminal semantic language, scenarios, concepts, vocabulary and terminologies for 3D spatial analysis of 3D popular music recorded in the Dolby Atmos format. The associated concepts, vocabulary, and terms map across domains to form the necessary specific intertextualities that can clarify perceptions and pin-point understandings of substantial 3D spatial componentry examined in, 1) positional relationships; 2) cognitive perspectives; 3) proportional properties; and 4) character attributes. This research is aimed at composers, songwriters, mixing engineers, and pedagogues who teach the art and craft of 3D music composition and mixing for the Dolby Atmos format.

The next episode reveals a novel philosophy and benchmark for a cutting-edge, premium quality Dolby Atmos Hyper Near-Field Tiny Studio (HN-FTS). It is an eco-friendly, small-scale, large-scope control room/audio-lab designed for sole creators, professional mix engineers,

recording studios and educational institutions that require an economic solution for entry into the Dolby Atmos format.

### 6.23 NASA: Found in Space



**Figure 21:** Gemini-7 – Earth-Sky View – Limb – Outer Space.

This image is a metaphor for the eureka moment when metaphysical components take on agency to provide palpable immersions in 3D music spatiality. Eureka means “I have found it.” It was what Archimedes said when he identified the principle of buoyancy.

GEMINI-7 - EARTH-SKY VIEW - LIMB - OUTER SPACE

NASA ID: s65-65257

S65-65257 (4-18 Dec. 1965) --- Sunrise and Earth's limb, as photographed by astronauts Frank Borman and James A. Lovell Jr. during their Earth-orbital 14-day mission in the Gemini-7 (GT-7) spacecraft. Photo credit: NASA

Date Created:1965-12-07

NASA Permissions: Non-Commercial use agreement; News outlets, schools, and text-book authors. For educational or information purposes.

## EPISODE 3

### Chapter 7: The Hyper Near-Field Tiny Studio

---

#### 7.1 Overview

On May 17, 2021, Apple Music announced that it would bring Spatial Audio in the Dolby Atmos format, and Lossless Audio, to its entire catalogue at no additional cost to consumers. In addition, Apple brought Spatial Audio and “immersive music-authoring tools” to its digital audio workstation (DAW) Logic-Pro in late 2021. This gave music creators the software tools needed to publish music in these new formats.

The primary question for music makers worldwide became, “how do we effectively implement the associated physical studio infrastructure to support spatial and lossless mix and master operations via speakers and headphones for delivery to distributors and retail aggregates?”

The considerations for designing and building this small room are: 1) architecture, 2) acoustics, 3) electro-mechanics, 4) environment, 5) ergonomics and tactile large-scale operations in a small-scale, 6) reliability and transference, 8) economics, 9) hyper calibration, 10) meeting Dolby guidelines for sound pressure level (SPL), reverb time at -60 dB (RT<sub>60</sub>), and noise compliment at -25 dB (NC-25), and quality control (QC). Additionally myriad choices must be carefully weighed to solve small-room physical standing wave resonances.

The term “tiny studio” comes from the globally popular Tiny Home concept (Carrizosa, 2021) which is a form of architecture inspired by huts and tents built by indigenous peoples such as the Sioux, Inuit, and Samoans.

A tiny studio philosophy is made all the more necessary in the present era which prompts creators around the world to work remotely from their home spaces. This author's home-based Hyper Near-Field Tiny Studio (control room) is just 2m x 4.5m x 2m (X, Y, Z), and approximately 24m<sup>3</sup> cubic meters, which is 50% smaller than Dolby guidelines recommend for a Dolby listed near-field Atmos control room.

The physical limitations of this space created opportunities to bring unexpected and effective solutions that achieve an exceptionally flat aggregate room response curve, impressing veteran electro-acoustic engineers at Dolby, and a growing list of eminent Atmos music creators, engineers, and producers who have subjectively auditioned the Hyper Near Field Tiny Studio.

This episode pioneers a novel philosophy and benchmark for the Dolby Atmos 7.1.4 Hyper Near-Field Tiny Studio and encourages music creators to build their own premium 3D audio-music 3D control room on a comparatively modest budget.

## 7.2 Introduction

If you are considering expansion of a small stereo or 5.1 music production room to 7.1.4 Dolby Atmos, this episode will interest you. Although this author is not a professional control room designer, in Spring of 2020 he began research that aimed to define a small Atmos studio—a single operator 7.1.4 Hyper Near-Field Tiny Studio (HN-FTS) built for the purpose of audio mixing and mastering with reliable transference to mainstream professional and consumer-level playback systems. By Spring of 2022, creation of 3 dimensional (3D) immersive music began and on December 9<sup>th</sup>, 2022, this author released his inaugural Atmos album, *Summertime in Leith*, on which he played upright bass with Oscar Peterson protégé Robi Botos on piano. This release to digital streaming providers (DSPs) via The Orchard (distributor) stands as the first

empirical evidence to indicate that successful mix/master transference from this HN-FTS to the marketplace is achievable.

In spring of 2022, less than 200 Dolby Atmos music studios existed globally, and the HN-FTS category was undocumented. A HN-FTS is intended for independent professional and semi-professional music creators, institutions and commercial studios with a constrained space and budget. It utilizes applicable Dolby guidelines, control room design theory, and reaches beyond to exemplify how socio-economic Tiny Home philosophies of small-scale, large-scope, and buy-local can encourage innovation. Supporting research is gathered from diverse sources to convey essential understandings for a premium Atmos 7.1.4 control room that is approximately 50% smaller ( $24\text{m}^3$ ) than Dolby's minimum Near-Field guideline of  $>50\text{m}^3$ . Small spaces are acoustically problematic, while paradoxically, mainstream Atmos playback in the acoustically reverbless and high noise criterion automotive cabin is rapidly oncoming [Lucid, Mercedes Benz, Volvo]. As such this author argues that the low reverb time ( $\text{RT}_{60}$ ) HN-FTS is a valuable link to the hyper-small ( $8.5\text{m}^3$  to  $12\text{m}^3$ ) electric vehicle cabin Atmos playback space.

This HN-FTS philosophy is distinguished by premium sound quality, thrifty economics, and a single meticulously tuned mixing position. The benchmark design utilizes Dolby's Audio Room Design Tool (DARDT) to formulate criteria for music room function, geometry, speaker layout, amplifier wattage, sound pressure level (SPL), reverb time ( $\text{RT}_{60}$ ), and noise criterion (NC-25). Initially the inspiration for a HN-FTS came from a multichannel, full range small speaker design philosophy (Fielder, 2003; Sijan, 2008) that was pioneered at Audio Design Labs Blue Sky Studio Monitors, and extensively utilized at Skywalker Sound, a division of Lucasfilm.

This author has over 15 years of experience working with Blue Sky speakers in three Hyper Near-Field 5.1 control rooms that provided reliable transference to broadcasters,

streaming services, and Far-Field cinematic theatres. This HN-FTS philosophy presents an integrated path for small control room expansion to 7.1.4 Dolby Atmos and the project's research methodology relies on a blending of academic texts, specialist advisement, and practical experimentation; all intended for subjective interpretation and application in the creative domain of music and audio production. Texts from Philip Newell, Dr. Floyd Toole, Alton Everest, Kenneth Pohlmann, the AES-E-Library, Sound on Sound magazine, and Dolby Professional Support contribute essential science, while hands-on knowledge is gleaned from on-line video interviews with the most experienced Atmos practitioners. Contemporaneous discussion located in the Metaversity of Facebook groups (Atmos Music Mixing Professionals and Atmos Post Mixing Professionals) are also referenced. The research is synthesized to form a durable philosophy that encourages the independent designing and building of a 7.1.4 Hyper Near-Field Tiny Studio.

### 7.3 Tiny Studio First Steps

The Dolby Home Entertainment Studio Certification Guide (available from Dolby) provides foundational guidance for Atmos installations, but it generally describes larger Near-, Mid-, and Far-Field room sizes—specifically, small Atmos rooms (below 50m<sup>3</sup>) are not described or encouraged. This author's next step was to look outside of Dolby's guideline for Tiny Studio solutions, then only to discover Philip Newell's sobering observation:

*“The majority of people working in modern music recording studios have not had the necessary formal education in mathematics, acoustics, and electronics to make the standard textbooks appear as anything more than cold print, and direct contact with specialists is out of reach for most people.”*

(Newell, 2017, p.xxviii)

This author did not find Newell's book *Recording Studio Design* to be "cold print." It provided cogent information about the studio/control room system and its integration with the human hearing system. Floyd E. Toole's book, *Sound Reproduction: The Acoustics and Psychoacoustics of Loudspeakers and Rooms*, also yielded significant Tiny Studio determinants:

*"... rooms have a massive influence on what we hear at low frequencies"; "... there are no 'ideal' dimensions for rooms"; "Listeners have the ability to adapt to many aspects of room acoustics"*

(Toole, 2018, p.43)

Innately, humans do not all hear the same because each person's pinnae (outer ears) are unique. Newel supports this fact by arguing that hard objective science, ephemeral art, and adaptability must co-exist in the world of sound recording (Newell, 2017). Falling outside of the Dolby guideline, a Tiny Studio relies on the assembler's personal psychoacoustic perception to achieve its sweetest sound while utilizing science to avoid pitfalls. As will be shown, this Tiny Studio can meet many of the Dolby guidelines, but the biggest acoustic challenge is 1<sup>st</sup>-, 2<sup>nd</sup>-, and 3<sup>rd</sup>- order, low-frequency resonances called standing waves. Guided by DARDT, the Tiny Studio makes standing waves less problematic by hyper-focusing its correlated speaker system toward a single, hyper-tuned mix-position. When seasoned audio engineers learn the characteristics of a control room, psychoacoustic assessment of translational sonic relationships is standard practice, and they often combine the use of mastering-grade reference headphones with a finely tuned speaker system to average their results for reliable playback across external speaker systems in the consumer field. A Tiny Studio workflow may also utilize mastering-grade headphones and a tuned speaker layout but adds occasional "field trips" to listed Dolby Atmos control rooms to ensure reliable crafting of translatable Atmos mixes and masters (Dolby 2022c).



## 7.4 Why is a Tiny Studio Necessary

A Tiny Studio provides the necessary acoustic decorrelation and aural cross feed to generate an out-of-head (external) listening experience but with direct speaker sound hyper near (1.2m) to the mix-position. The economic advantages of a Tiny Studio are compelling, and they complement the affordable Atmos creator tools that are currently provided in digital audio workstations (DAWS)—Logic Pro, Pro Tools Ultimate, Nuendo, Pyramix, Ableton, and Davinci Resolve (Dolby, 2021d).

## 7.5 Psychoacoustics, Speaker, and Headphone Workflows

All audio work in 360-degree sound benefits from the practitioner understanding human psychoacoustic perception, auditory science, and physics. For example, when recording the native sound radiation of musical instruments, “to ensure that the interaural signal differences are captured, at least one microphone [placed] every 0.15 m is necessary, which roughly corresponds to the distance of a listener’s ears” (Ziemer and Bader, 2017, p.484). Furthermore, spatial hearing relies on interaural time difference (ITD), Interaural level difference (ILD), and Head Related Transfer Function (HRTF) (Rothermich, 2021). “HRTF involves frequency dependant interaural time differentiation that is achieved with pinna filtering” (Thomas, 2020), but headphones can facilitate a psychoacoustic perception of Atmos by synthesizing these as binaural spatial cues.

Dolby content creation specialist Margaret (Maggie) Tobin has developed the best practices for binaural mixing in Atmos on headphones and her master’s thesis investigates Atmos headphone-to-speaker translation (Tobin, 2021; Dolby 2020a). She works on a laptop with Sennheiser HD 600s and earbuds. She argues that hybrid headphone/speaker system

monitoring works, but “consumer headphones can be a bit bass heavy” and she warns, “low frequencies and the LFE channel can be difficult to judge”—because, in a 5.1 to 2.0 Dolby rendered downmix “The LFE channel is ignored ...” (Rothermich, 2021, p.95). Dolby Atmos supports autonomous headphone selection. Apple uses a proprietary playback renderer for unique binaural spatialization (with headtracking) when heard through AirPods. Some of the most experienced Atmos music mix engineers [Andrew Scheps, Greg Penny, Steve Genewick, Dave Way, and Fab Dupont] agree that you can get about eighty-five percent of the way towards a satisfactory Atmos mix using headphones. In a 2021 *Mix with the Masters* interview Steve Genewick (eminent Atmos mixing engineer since July 2017, Capitol Studios) says, “it’s not unusual for us to start [Atmos] mixes in binaural.” This author’s Tiny Studio praxis also utilizes mastering grade headphones and involves a binaural translator called Virtuoso—developed by the Applied Psychoacoustics Lab (APL) team at the University of Huddersfield, UK.

## 7.6 Atmos Economics

Building a Near-Field commercial Dolby Atmos control room is expensive and, as reported by practitioners in the *Mix with the Masters* interview, Atmos mixing often returns minimal economic compensation because record labels are frugal, but they realise that Atmos music mixes will future-proof their catalogue. The Tiny Studio philosophy and benchmark respond to this economic dilemma by providing music creators with an affordable control room investment when compared to premium Far-, Mid-, and Near-Field Atmos rooms. See section [7.24 Economic Comparison](#), **Figures 27, 28, 29 and 30**.

## 7.7 Audio History, Speaker Location, and Format Evolution

Atmos stands on the shoulders of late 19<sup>th</sup>-century monaural sound reproduction, Alan Blumlein's patented binaural stereo sound (Fox, 1982) and 5.1 surround sound pioneered by Walter Murch in 1979 on Dolby 70mm film for the movie *Apocalypse Now* (Ondaatje, 2002).

Atmos expands the 5.1 speaker layout to add height—5.1.2-, 7.1.4-, 9.1.6-, ...up to 64 speakers for cinema, and 22 speakers for music. Audio tracks can be assigned to a channel-bed or 128 audio-objects that utilize spatial location metadata which is rendered as coded (grouped) spatial elements/clusters (Rothermich, 2021; Nair, 2022) to reduce data size during streaming playback. Objects can be monaural or stereo and Blumlein's binaural stereo format is used for headphone playback. Atmos speaker systems are located on two planes, and the subwoofer is placed on the floor. The ear-level speakers (approximately 1.2m above the floor) are positioned on the standard mid-plane, and the top height-plane speakers must be correctly angled and hung safely from the ceiling. The Atmos format is evolving rapidly with a community made up of pioneers who have been creating music in Atmos since 2017. In 2019, Amazon HD+ and Tidal introduced Atmos music on their respective streaming platforms (Dolby, 2020b).

## 7.8 Atmos Basic Theory and Operation

7.1.4 Dolby Atmos uses 12 channels of audio. Their accumulative energy is expressed as a K-weighted, full-scale loudness measurement (LKFS). An Atmos renderer (software or hardware) does this by folding down all channels and objects into a 5.1 configuration for measurement of total loudness, targeting -18 LKFS (integrated) with a maximum peak of -1 dB to provide 17 dB of headroom. Music mixers routinely create spatial decorrelation by cascading horizontally panned stereo reverbs. Binaural playback rendering via either Dolby Atmos or

Apple Spatial Audio renderers introduce virtual binaural spatiality in positions of close, near, and far to simulate the psychoacoustic signals needed for out-of-head audio immersion when listening on stereo headphones.

Atmos affords different workflows for film/tv audio post-production and music production. In film/tv teams of mixers (dialogue, music, and effects) work together at once, whereas music typically involves one mix engineer. The workflow type and stage of work—pre-mix, mix, or master—to be done affects the control room design as is exemplified by Steve Genewick’s COVID-19 “guerrilla Atmos music installation” in his living room (*Mix with the Masters*, 2021). He used only two top-height speakers but claimed that it allowed him to create mix basics for music. In addition, he took “field trips” to other Dolby Atmos control rooms for quality assurance. Dolby strongly encourages this practice and to support it they provide a global listing of approved Atmos studios (Dolby, 2024a).

## 7.9 A, B, and QC Chains

The A and B chain, and what this author calls the QC chain (quality control) make up the primary system chains for an Atmos studio. Atmos relies on a renderer, which the DAW connects to through 128 channels of input/output (i/o). The A chain involves everything used before the renderer—computer hardware and software; The B chain involves everything used after the renderer—audio interface, speaker controller, room tuning hardware/software, physical absorption, diffusion, and speakers. The QC chain is used to analyse commercially released Atmos mixes through the discrete Atmos speaker playback system or headphones. Typically, it involves an audio-visual receiver (AVR), or Arvus H2-4D decoder, or Apple TV or a Blu-Ray player. A significant breakthrough saw macOS Monterey (12.4) provide discrete channel Atmos

monitoring via 16 channels of core audio. This discovery was introduced to the Dolby Atmos Post Mix Professional Facebook group in April 2022 by David Stagle and was tested by group practitioners. It works with Apple Music and Apple TV, and instructions are published publicly on [pro-tools-expert.com](https://pro-tools-expert.com) (Stagle, 2022). In macOS Sequoia (15) Atmos is now channelized directly from the Apple Music and TV desktop applications to a 7.1.4 speaker gamut.

### 7.10 Design with DARDT and Do More with Less

Available from [Dolby.com](https://dolby.com), DARDT is an interactive spreadsheet tool that uses advanced macros to pre-build the Atmos control room virtually. Upon registration with Dolby, control room design begins with the self-explanatory DARDT Quick Start Video Series. Beyond DARDT, this HN-FTS design strives to do more with less. For example, the global Tiny Home movement inspired economical, large-scope, small-scale hardware solutions such as the eco-friendly Anubis, Hapi MKII audio hardware and Canadian designed PSB speakers.

### 7.11 Decorrelated Sound from a Correlated Speaker Layout

In the parlance of physics and philosophy, the *Oxford Dictionary* depicts the adjective [\*spatiotemporal\*](#) as a descriptor of an object that belongs to both space and time, or to space-time.

Discrete immersive multichannel, locationally and temporally dynamic *object-based audio*, and [\*Ambisonic audio\*](#) is perceived by the human auditory system as virtual decorrelated sound that emanates from a 3D speaker gamut. Increasing the number of speaker channels present a higher accuracy of auditorial location.

Cinematic Far-Field Atmos theatres use speaker arrays for each audio channel with numerous speakers for audio-objects, to disperse direct sound into a large room at cinematic sound pressure level (SPL).

In a 7.1.4 Tiny Studio one speaker per audio channel focuses direct sound to the sole mix-position to function as a well-tuned (correlated) Hyper Near-Field speaker playback system. A highly decorrelated Atmos music mix presents very effective differences in timing, spectre, location, and volume between channels and speakers to construct an immersive out-of-head audio experience—the opposite of “In-head localization” (Toole, 2018, p.182).

A correlated mix presents the same sound from adjacent speakers to create a non-immersive “periphonic” (Gerzon, 1972, p.1) in-head sensation. When a multichannel mix is not supported with interchannel decorrelation, periphony results, potentially leading to variant degrees of phase cancelation and comb filtering when folded down to binaural audio.

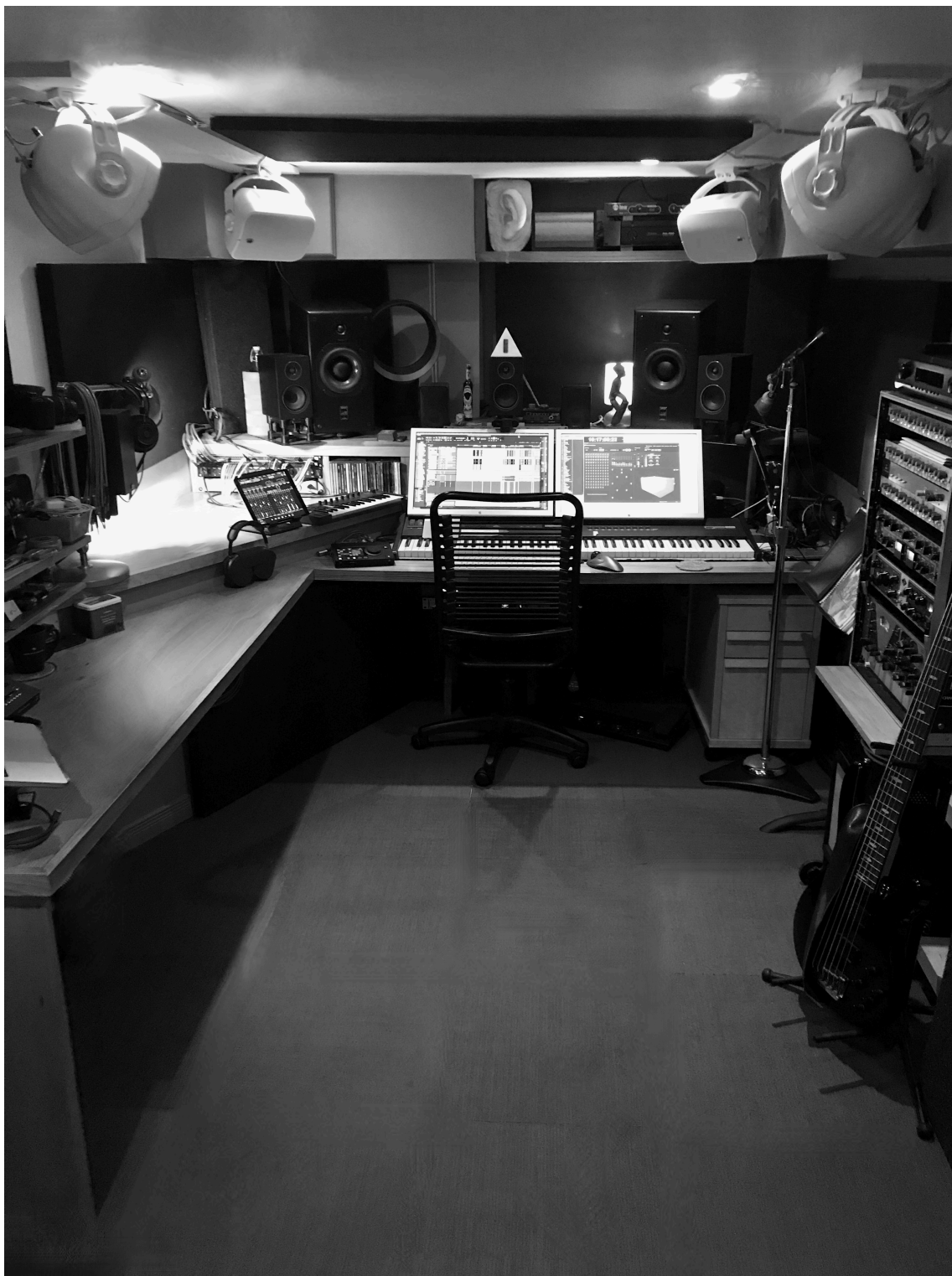
This author argues that for music, custom-prepared stereo (Left only/Right only/Lo/Ro) fold-downs remain superior to a static fold-down pre-set (Novotny, 2017), however the Atmos renderer can create acceptable stereo (static) fold-downs. When creating mixes on an untuned (decorrelated) playback system the result will be unpredictable mix behaviour when heard on a calibrated (correlated) playback system.

Dolby’s Technical Guideline (Dolby, 2021b) provides standards for correlating speaker systems in equidistant or orthogonal layouts that vary in size from near- to far-field. In comparison to larger multiple mix-position Atmos rooms, a primary HN-FTS differentiation is the hyper tuned sole mix-position. This blended equi-orthogonal Tiny Studio layout is highly correlated with speakers on the standard mid-plane placed away from the wall with left/center/right (LCR) aimed toward the sole operator.

The front top surround speakers are placed slightly inward of the L, C, R field about 50° toward the mix-position and angled at 51° to aim just over the operator's shoulders. The LFE is laterally positioned on the floor between the centre and right-front speakers, and direct loudspeaker sound is minimally affected by reflections from room geometry. Low frequency standing waves along with mid-to-high frequency diffusion influence the perception of mix decorrelation and spectral colouration. Expert Atmos mixers bake in their unique mix decorrelations which should then translate reliably when heard on other correlated speaker systems that are precisely positioned, timed, and voiced for the room. This Tiny Studio speaker layout involves speaker distances that range between 102 to 106 centimetres (please see **Table 12**) to mix position, resulting in a room that looks (please see **Figure 22** and **Figure 23**), feels, and sounds like big headphones with acoustic cross feed. Electronic timing adjustment in the Anubis speaker controller corrects speaker timings that are within a 24-centimetre range of difference.

| Speaker | Distance to Mix-Position |
|---------|--------------------------|
| Left    | 120 cm                   |
| Right   | 120 cm                   |
| Center  | 120 cm                   |
| Sub     | 110 cm                   |
| LS      | 102 cm                   |
| RS      | 102 cm                   |
| LRS     | 126 cm                   |
| RRS     | 126 cm                   |
| LTF     | 106 cm                   |
| RTF     | 107cm                    |
| LTR     | 111 cm                   |
| RTR     | 111 cm                   |

**Table 12:** Speaker distance to mix-position.



**Figure 22:** Hyper Near-Field Tiny Studio (HNFTS), photography by Paul Novotny, 2022.





**Figure 23:** Hyper Near-Field Tiny Studio, sole mix-position, photography by Paul Novotny, 2022.

## 7.12 The Components of a Correlated Playback System

Playback system correlation can be seen as a series of components: acoustic room tuning to absorb resonances below the Schroeder frequency and diffuse soundwaves above (Brown, 2021); uniform electronic system polarization; minimal phase distortion; amplitude correction; precise system clock; speaker-delay timing; and speaker equalization with bass management that correlates frequencies below 80 Hz by sending them to a single monaural subwoofer. DARDT indicates specifications and measurements for a correlated system. Fine tuning is accomplished

with real-time-analysis (RTA) and psychoacoustic perception. Playback system correlation relies on a combination of hard science and aural perception.

### 7.13 Small Control Room Acoustics and the “My Room Principle.”

Small control room dimensions invoke problematic low-frequency energy pressure zones—standing waves that must be suppressed. DARDT calculates their location and frequency. They occur in first, second, and third order. Floyd Toole describes what causes standing waves—also called eigenfrequencies, eigentones, eigenmodes, modes, or resonances:

*“Sounds of all frequencies and wavelengths reflect around rooms in a somewhat chaotic manner, but when the wavelengths are of the same order as the dimensions of a room some special events occur. Between parallel surfaces that are half a wavelength apart, or multiples of half-wavelengths apart, the sound travelling in both directions add constructively, reinforcing each other, and a standing wave develops. This is the basis of a room mode or resonance.”*

(Toole, 2018, p.100)

Finding an effective method to render them innocuous in a sound field is a significant challenge because small rooms do not have enough space for low-frequency absorption material. In his AES preprint, “A Controlled-Reflection Listening Room for Multichannel Sound,” BBC research engineer R. Walker observes:

*“The three main components of the sound field in the vicinity of the listener are the direct sound, the early reflections and the later reflections which merge to form the reverberant field. All these components are functions of time and frequency.”*

(Walker, R. 1998, p.25)

Walker references the *International Telecommunications Union* and *European Broadcast Union* recommendations—ITU-R BS.1116 [1] and EBU Rec.R22 (tech.Doc.3276 [3]) as “closely-controlled parameters for listening rooms for critical listening.” As expected, this Tiny Studio breaks these rules prompting this author to investigate the novel “MyRoom principle.”

Bogic Petrovic and Zorica Davidovic describe their “MyRoom principle” in a 2010 AES paper. “In general, rooms smaller than 80m<sup>3</sup> are not advised to be used for a control room for stereo and surround production and reproduction.” (Petrovic, B., Davidovic, Z., 2010). They concluded that their small-room designs (Studio 1 at 76m<sup>3</sup>, and Studio 2 at 33.30m<sup>3</sup>) could not rely on existing designs of Live-End Dead-End (LEDE), Non-Environment (N-E), Reflection Free Zone (RFZ) or Andrew Parry’s design called “Early Sound Scattering” (ESS). They set out to create a new system they describe as an “acoustic treatment in which all surround speakers would be equally treated, while not compromising stereo monitoring.”

*“Every surface had to be hybrid, containing a reflective and absorptive part, providing absorption in low to mid frequency range, and diffuse reflection in mid to high range, with a crossover region in which both are present.”*

(Petrovic, Davidovic, 2010, p.3)

They conclude that their hybrid surface “MyRoom principle” is intended for a home studio located in an apartment/house. They propose the technical names of Non-Coherent (N-C) acoustic design or Early Reflections Delay Encryption (ERDE). Their design (please see Figure 24) is unique and the Studio 2 room size (33.30m<sup>3</sup>) is comparable to this author’s Tiny Studio which is (24m<sup>3</sup>).



**Figure 24:** “MyRoom Principle”—studio 1 with hybrid surfaces.

Everest and Pohlmann (2022) observe that a frequency range of 20 Hz to 20 kHz spans 10 octaves and represents wavelengths ranging from 57 ft to 0.6 inches. Newell tells us that absorbent material needs about “one quarter of a wavelength” (Newel, 2017, p.18) to suppress the constructive effect of standing waves—approximately one meter of thickness is needed at 100 Hz. Everest and Pohlmann (2013) argue that perfect diffusion is difficult to achieve in small rooms due to less absorption, which leads to variant short reverb times. They also warn about irregular room geometry.

*“... modes are easy to calculate in a six-sided rectangular room with parallel surfaces. A room with different geometry, for example, splayed walls, also supports modes but the physical placement of the pressure response of the modes is distorted by the nonparallel surfaces, splaying walls does not solve mode problems; it only makes them harder to predict.”*

(Everest, Pohlmann, 2013, p.143)

Bryan Pennington (Dolby Senior Field Applications Engineer) adds his recommendations for what he terms “a well-behaved Dolby Atmos room”.

*“You want to make sure it’s not too big or too small ... but I can’t stress enough that bigger is not always better, you want to make sure that a room is not too big to turn out a good mix. You want to make sure that you’re probably more in a rectangular room, less in a square room because you can load yourself up with acoustic problems.”*

(Pennington, Dolby, 2022c, 8:35-10:10)

If the room is too dead the mixer might over-reverberate. However, this author argues that low  $RT_{60}$  (please see **Table 13**) makes a Tiny Studio valuable for judging music playback in the hyper small Atmos car cabin. Dolby says  $RT_{60}$  times are best when they are between .5 to .8 milliseconds (*ibid*), but this Tiny Studio  $RT_{60}$  is considerably less. Atmos specification for noise criteria (NC-25) means the room will not have an inherent noise level that masks low-level reverb tails at -60 dB when audio is played at 85 dBC RMS. Paradoxically, mainstream Atmos playback in high-noise, low- $RT_{60}$  automotive cabins is oncoming (Dolby 2022a; Dolby 2023; Dolby 2024b).

A 2016 Tesla Model XP90D has a cabin noise level ranging from 34.4 dB at idle, to 67.3 dB at 140 km/h (auto-decibel-db.com, 2023). This indicates that at a cinematic playback level of 85 dBC, approximately 10 dB of reverb tail will be masked at idle, and at 140 km/h, approximately 42 dB of reverb tail will be masked by cabin/road noise.

| Measured RT <sub>60</sub> [s] |              |              |              |             |             |             |             | Notes             |
|-------------------------------|--------------|--------------|--------------|-------------|-------------|-------------|-------------|-------------------|
|                               | <b>125Hz</b> | <b>250Hz</b> | <b>500Hz</b> | <b>1kHz</b> | <b>2kHz</b> | <b>4kHz</b> | <b>8kHz</b> | <b>Average of</b> |
| Measurement 1                 | 0.07         | 0.07         | 0.09         | 0.05        | 0.07        | 0.06        | 0.14        | 3 samples         |
| Measurement 2                 | 0.07         | 0.16         | 0.14         | 0.17        | 0.07        | 0.21        | 0.27        | 5 samples         |
| Measurement 3                 | 0.07         | 0.08         | 0.07         | 0.07        | 0.09        | 0.11        | 0.06        | 3 samples         |
| Measurement 4                 | 0.11         | 0.12         | 0.16         | 0.11        | 0.16        | 0.15        | 0.22        | 7 samples         |
| Measurement 5                 | 0.07         | 0.09         | 0.08         | 0.08        | 0.08        | 0.08        | 0.03        | 4 samples         |
| Measurement 6                 | 0.12         | 0.08         | 0.07         | 0.44        | 0.41        | 0.06        | 0.14        | 5 samples         |
| Measurement 7                 | 0.10         | 0.10         | 0.09         | 0.09        | 0.09        | 0.09        | 0.07        | 1 sample          |
| Measurement 8                 | 0.17         | 0.09         | 0.08         | 0.06        | 0.07        | 0.07        | 0.08        | 36 samples        |
| Measurement 9                 | 0.12         | 0.13         | 0.10         | 0.12        | 0.12        | 0.07        | 0.08        | 1 sample          |
| Measurement 10                | 0.10         | 0.11         | 0.10         | 0.10        | 0.08        | 0.11        | 0.13        | 3 samples         |
| <b>HN-FTS Average</b>         | <b>0.10</b>  | <b>0.10</b>  | <b>0.10</b>  | <b>0.13</b> | <b>0.12</b> | <b>0.12</b> | <b>0.12</b> |                   |

**Table 13:** DARDT-RT<sub>60</sub> times for this author’s Hyper Near-Field Tiny Studio (HN-FTS).

## 7.14 Hyper Near-Field Room Geometry and Room Volume

Room geometry and speaker layout size are two different measurements. Dolby’s smallest room category is Near-Field with a room volume of fifty cubic meters (50m<sup>3</sup>). But Dolby’s Bryan Pennington states that “there is no hard and fast minimum room size” (Pennington, 2022; Dolby 2021e). As such, this Hyper Near-Field benchmark precisely defines the Tiny Studio’s room geometry, room volume, and speaker layout size. **Table 14** compares the Dolby specifications (Dolby, 2021b) of a Near-Field room to this author’s Hyper Near-Field Tiny Studio.

| <b>Comparison Table</b>          | <b>Near-Field</b>     | <b>Hyper Near-Field</b> |
|----------------------------------|-----------------------|-------------------------|
| <u>Dimension</u>                 | <u>Specification</u>  | <u>Tiny Studio</u>      |
| Minimum speaker layout height    | 2.4m                  | 1.75m                   |
| Minimum speaker layout width     | 3m                    | 2.0m                    |
| Minimum speaker layout length    | 3.5 m                 | 2.4m                    |
| Recommended room volume          | >50m <sup>3</sup>     | 24m <sup>3</sup>        |
| Speaker distance to mix-position | ≤5m (<4m recommended) | 1.2m                    |

**Table 14:** Comparison of the Near-Field to the Hyper Near-Field.

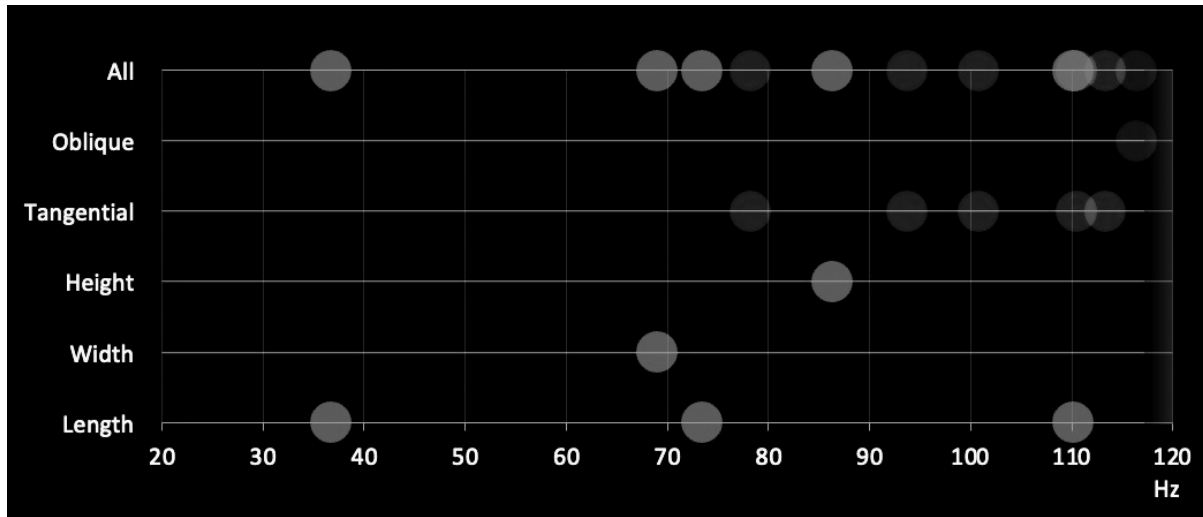
This author's Tiny Studio room geometry is rectangular. The mix-position is located at the long end. The rear wall has a 1.96m-high, 70mm-wide door, which functions as a bass trap when left open. The Room length is 4.7m, the Width is 2.57m. The basement ceiling height is low at 1.98m, regardless, this author perceives no compromise in dimensional separation between the standard mid-plane and the height plane (canopy) speakers. The equi-orthogonal speaker sphere is approximately 4.2m<sup>3</sup> and speaker to mix-position measurements are taken from the center point of each speaker grill. In a small room with low height, orienting standard mid-plane speakers in a tweeter-on-top vertical stance reduces side-to-side smear, improves perception of spectral height from the L/C/R speakers, and tweeters with waveguides will help to control high frequency splay and speaker crossover. Speaker and mix-position placements should locate outside of the eigenmode peaks and valleys.

*“It is very difficult to make a good quality studio, free of problematical compromises, in spaces with inadequate height. Control rooms require height because of the need to avoid acoustic problems between the parallel floor and ceiling.”*

(Newell, 2017, p.9)

## 7.15 Standing Waves and the Subwoofer Position

DARDT calculated an optimal speaker layout with a subwoofer position that minimized standing waves in the room dimensions (**Figure 25**). The 1<sup>st</sup> mode was at 35 Hz, 2<sup>nd</sup> at 69 Hz, 3<sup>rd</sup> at 104 Hz, and 4<sup>th</sup> at 138 Hz and axial modes were at 67 Hz, 134 Hz, 201 Hz, and 268 Hz.



**Figure 25:** Hyper Near-Field Tiny Studio, dimension of standing waves.

## 7.16 Dolby Room Curve

The Dolby modified X curve emanates from cinematic practices that date back to the 1970s. This music-oriented Tiny Studio does not utilize the X curve or the Dolby music curve, although the subwoofer is aligned to Dolby specification (pink noise plays at approx. +5 dBC above the center speaker) to balance bass managed energy (86 Hz crossover) and low-frequency effect (LFE) with the satellite speakers. (Please see Ioan Allen's 2006 SMPTE journal paper for more information on the Dolby X curve.)



### 7.17 Timbrally Neutral Canadian Loudspeakers

Paul Barton (Toronto, PSB founder and Chief Engineer for Lenbrook Industries Group of Companies which include PSB Speakers, NAD Electronics, and Bluesound) provided this author with generous advice and shared a compelling Canadian story involving Barton's mentor, Dr. Floyd E. Toole.

Moncton, New Brunswick native Floyd Toole and Paul Barton are central figures in international loudspeaker design, and their collaboration benefitted this Tiny Studio. In 1965, Floyd Toole graduated with a Ph.D. in electrical engineering from the Imperial College of Science and Technology, University of London. His multi-disciplinary research involved binaural hearing and sound localization from signal generation. After graduation Dr. Toole became a research scientist at the National Research Council of Canada (NRCC) where he designed an anechoic chamber. For PSB speakers, Toole did the research and Barton implemented the science. In 1991, after 26 years of research, Dr. Toole was hired as Vice President of Acoustical Engineering at Harman International Industries, and now in 2023 PSB is a highly respected international speaker brand possessing a long history with Dolby. Floyd Toole maintains that: "If we are to set an objective for loudspeaker performance, it should be timbral neutrality" (Toole, 2018, p.16). "The most common features of small low-cost loudspeakers are lack of bass, and an inability to play loud" (Toole, 2018, p.360).

Bryan Pennington (Dolby, 10:10, 2022c) advises that: "Atmos works with all speakers and price does not denote quality" and he describes two uses of corrective speaker EQ for Atmos rooms—to make speakers sound more like each other, and to help mixes translate to other environments.

For a Tiny Studio philosophy these understandings are economically liberating since the largest expense of a larger Atmos studio is numerous high-powered speakers that are capable of meeting Dolby's SPL requirements. Paul Barton agreed with Toole and Pennington. He suggested that using seven passive PSB Alpha P3 bass reflex speakers with a rear firing port which can be bunged to fully close the cabinet (like Blue Sky speakers), or be partially bunged, or left fully open would be helpful when voicing the room. Fully open ports ended up sounding the best.

This equi-orthogonal Tiny Studio speaker layout provides a 1.2m distance from speaker to mix-position and DARDT calculates that these PSB speakers will play at a common (RMS) level of 84 dBC with a true peak of 104.7 dB. The on-axis frequency range is 57-21,000 Hz with a low frequency cut-off of -10 dB at 43 Hz. Originally this Tiny Studio powered satellite speakers economically with multiple ART SLA 1 (2 X 100 watts) stereo amplifiers, but now Canadian made Bryston BD 225 power amps are utilized providing 225 watts per channel. Paul Barton recommended using copper 12-gauge speaker wire with the shortest lengths possible.

The 2020 TEC Award winning PSB 10" CSIR passive subwoofer is in a slim enclosure, merely 10.2cm deep—Paul Barton's design places the magnet in front of the speaker cone. DARDT indicated that the CSIR subwoofer should be located on the floor against the wall between the centre and right speakers. The gracile enclosure afforded a position that helped to suppress standing waves. Barton's CSIR 10" and the PSB CS 500W Architectural Subwoofer Amplifier are well suited to a Tiny Studio and DARDT calculates that this 500-watt amp/speaker combo will peak at 103 dB, giving the subwoofer 83 dBC of common level. Now with the Bryston BD 225 amplifiers this Tiny Studio has 2975 watts of total power in a 4.2m<sup>3</sup> speaker playback sphere and easily surpasses Dolby's cinematic SPL specification of 85 dBC.

Bryan Pennington (Dolby, 8:35, 2022c) recommends that all Atmos studios will benefit from a single speaker manufacturer because uniform voicing crosses all models. Paul Barton agreed and suggested the PSB CS500 Universal In-Outdoor speakers for the top plane. They provide a safe, simple, and elegant mount-aim-dismount solution.

Passive speakers reduce control room heat and noise because power amps can be placed outside of the control room and standard wiring is preferable when repair is needed. This speaker system meets Toole's requirement for timbral neutrality and wide dynamic range while being cost-effective. Additionally, this Tiny Studio uses ATC-SCM20-ASL PRO MKII reference monitors for critical stereo listening and when the ATCs are paired with the PSB subwoofer (2.1), the system plays to 24 Hz. The ATCs are inverted so the PSB/ATC drivers match location.

*“‘Small rooms’ are notorious for bad bass.”*

(Toole, 2018, p.215)

The human ear has difficulty locating sound below 80 Hz, while “below 50 Hz, localisation is all but impossible” (Newell, 2017, p.583). Based on our auditory biology this author suggests that “eleven speakers that play to 40 Hz” (Dolby recommended) are intended for far-field cinematic rooms and perhaps not necessary for music. Frequencies below 80 Hz radiate in an omnidirectional pattern and by directing them to a bass managed single monaural subwoofer, system correlation is improved. Reducing satellite speaker frequencies below 80 Hz allows them to play with less distortion, more power, and wider dynamic range, while for music, single subwoofer bass management evokes the monaural bass sound associated with vinyl. Blue Sky speaker design philosophy (Fielder, 2003; Sijan, 2008) affirms the need for bass

management and Philip Newell references Floyd E. Toole, encouraging monaural bass management when in less than “first class” acoustic conditions.

*“... if first class acoustic conditions are neither available nor achievable for practical reasons, then fast, flat, mono bass below 80 Hz may be greatly preferable to non-flat resonant bass in stereo.”*

(Newell, 2018, p.586)

## 7.18 Speaker Isolation

Canadian made ISO Acoustic pucks decouple speaker energy transmission from a speaker mantle or speaker stands. They secure speaker footings during powerful compression and rarefaction cycles, and mitigate external energy feedback, preventing woofer blur. Canadian acoustic design engineer Dave Morrison conceived of the PSB CSIR subwoofer isolation system and as perceived by this author, Morrison’s ISO pucks cleared up this Tiny Studio’s L, C, R low and low-mid frequencies while improving horizontal imaging.

## 7.19 Measurement Tools

A measuring tape or laser measuring tool and an SPL meter able to measure in A and C weights with fast and slow responses are necessary for level matching speakers. C-weighting is used for speakers with similar spectral response down to 40 Hz, and A-weighting is used for dissimilar speakers.

## 7.20 Sound Pressure Level (SPL) Specifications

Based on this Tiny Studio’s room geometry, an equi-orthogonal layout was needed. There is no option for a blended equi-orthogonal layout, however it works very well in this Studio.

An Atmos room must be calibrated to operate at the SPL of 79 dBC, 82 dBC, or 85 dBC. These are common root mean square (RMS) energy levels requiring an additional peak range of +20 dB. 85 dBC is used for cinematic QC, while approximately 77 dBC is common for daily music mixing.

## 7.21 Audio Over Internet Protocol

Atmos control rooms often utilize the Audio Over Internet Protocol (AOIP) using ethernet (CAT 6) cabling to connect audio hardware to computers. AES67-Ravenna and DANTE are the current AOIP specifications and computer audio hardware functions within their ecosystems. 128 inputs/outputs or more are routinely used for Atmos audio.

## 7.22 Room Tuning and Calibration

Metaphorically speaking, when a large-scale visual image is reduced to a smaller scale the image tightens up. This author's Hyper Near-Field Tiny Study does the same thing to a creatively decorrelated audio image. Room tuning involves an order of operations that combine to express a flat, wide dynamic range, with a uniform and fast impact response. Mixing engineers adopt their preferred sound pressure levels for daily work and that level is used to hyper-tune the room. This author's preferred working level is 77 dBC. For the calibration Dolby pink noise was played at 77dB and measured using a C weighting with a slow response time. The Hyper Near-Field Tiny studio philosophy requires only one calibration location, rather averaging multiple calibration points. Larger Atmos rooms require averaging multiple points because the sweet spot needs to accommodate several mixers. The SPL meter and omni calibration microphone must be placed precisely at mix position, and at ear height which is often about 48" above the floor.

A room's Schroeder frequency (Schroeder, 1975) guides neutralisation of control room colouration by indicating the transition point between low frequency absorption and high frequency diffusion. This Tiny Studio's Schroeder frequency is 317 Hz. Absorption involves 1.5" (3.7mm) rockwool panels behind the L, C, R, side walls, ceiling cloud, and carpet tile floor. As expected, uneven bass response does appear below about 400 Hz and it must be dealt with.

This author's studio requires the LS, RS and LSS speakers to be placed close to the wall. The first stage of tuning involved a speaker port bung test to see if bunging would reduce low frequency energy that bounces off the wall surface directly behind the LS and RS speakers. Partial bunging of the PSB speaker ports with a small piece of foam lowered speaker sensitivity by about .05 dB, but when the ports were fully open, as intended by Paul Barton, the speaker gamut felt faster, more dynamic, and had better tonal balance.

The second stage involved hyper refinement of the speaker-to-mix position. Measured from the center grill to the SPL meter microphone head, the hyper small scale of the speaker sphere benefits from a +/- 1 millimetre tolerance. Each speaker distance was documented and then entered in a custom-made speaker timing calculator in MS Excel to create pre-timed speaker gamut which Sonarworks would then improve. Please see **Table 15**.

| Atmos HN-FTS Speaker Delay Calculator                                                |                        |                                      |                       |                            |                            |               |               |  |
|--------------------------------------------------------------------------------------|------------------------|--------------------------------------|-----------------------|----------------------------|----------------------------|---------------|---------------|--|
| Legend                                                                               |                        |                                      |                       |                            |                            |               |               |  |
| Base Distance/Base Time ~ Speed of Sound = 1,125ft/sec ~ 1.125 ft/ms                 |                        |                                      |                       | Anubis Firmware: 1.5.11b55 |                            |               |               |  |
| 30.5 cm = 1ms ~ 1/10th of a ms = .03 microseconds ~ 35 micro seconds = 1 centimeter. |                        |                                      |                       | Anubis Buffer: 1.33ms (64) |                            |               |               |  |
| Speaker                                                                              | ATC 2.0                | PSB 7.1.4-64KB Anubis Buffer=1.33 ms | Distance to mix in cm | Difference in cm           | Difference in Microseconds | Speaker Delay | Delay Setting |  |
| LS                                                                                   | No correction required | Room Timing Correction               | 122                   | 4                          | 0.013114754                | 0.013114754   | 1.428         |  |
| RS                                                                                   | "                      | Room Timing Correction               | 122                   | 4                          | 0.013114754                | 0.013114754   | 1.428         |  |
| C                                                                                    | "                      | Room Timing Correction               | 122                   | 4                          | 0.013114754                | 0.013114754   | 1.428         |  |
| LFE                                                                                  | "                      | Room Timing Correction               | 110                   | 16                         | 0.052459016                | 0.052459016   | 1.441         |  |
| SL                                                                                   | "                      | Room Timing Correction               | 102                   | 24                         | 0.078688525                | 0.078688525   | 1.487         |  |
| SR                                                                                   | "                      | Room Timing Correction               | 102                   | 24                         | 0.078688525                | 0.078688525   | 1.487         |  |
| Left Side-Rear                                                                       | "                      | Base Distance                        | 126                   | 0                          | 0                          | 0             | 1.33          |  |
| Right Side-Rear                                                                      | "                      | Base Distance                        | 126                   | 0                          | 0                          | 0             | 1.33          |  |
| TL                                                                                   | "                      | Room Timing Correction               | 106                   | 20                         | 0.06557377                 | 0.06557377    | 1.487         |  |
| TR                                                                                   | "                      | Room Timing Correction               | 107                   | 19                         | 0.062295082                | 0.062295082   | 1.48          |  |
| TRL                                                                                  | "                      | Room Timing Correction               | 113                   | 13                         | 0.042622951                | 0.042622951   | 1.457         |  |
| TRR                                                                                  | "                      | Room Timing Correction               | 111                   | 15                         | 0.049180328                | 0.049180328   | 1.457         |  |

**Table 15:** Hyper Near-Field Tiny Studio speaker timing delays.

Sonarworks Multichannel Sound ID Reference implements Finite Impulse Response (FIR) technology and supports sample rates up to 192kHz (Sonarworks, 2023). Operating it requires attention to the tutorial and this author ran the measurement test 5 times to learn how consistent the software was. It requires very precise placement of the calibration microphone to garner repeatable results.

For final verification this author invited established the well-established Dolby Atmos room tuning electro-acoustic engineer Edwin Segeren (Los Angeles, New York City, Toronto) to verify the HN-FTS results. He subjectively auditioned and scientifically analysed the SoundID Reference room response using Rational Acoustics SMAART software. This author's Hyper Near-Field Tiny Studio presented astonishingly good results.

The simple concept for control room tuning involves achieving as much physical correlation as is possible. This forms a rigid framework for electronic tuning to add a gentle touch of electronic software-based refinement. A highly correlated Atmos speaker gamut requires minimal intervention from electronic tuning.

This Tiny Studio benefits from Sonarworks SoundID Reference multichannel calibration which operates elegantly within the Anubis monitor controller. Quick and easy comparison between calibrated and uncalibrated monitoring indicates a significant improvement in room response at middle to low frequencies. A custom equalization curve is used that only tunes the frequencies below 600 Hz and no Dolby music curve is used. See **Figure 26: HN-FTS Aggregate Room Response, 12-19-23**, measured by Edwin Segeren.



**Figure 26:** HN-FTS Aggregate room response, 12\_19\_23, measured by Edwin Segeren.

### 7.23 Equipment List (HN-FTS)

| #  | Item                   | Description and Use                                      |
|----|------------------------|----------------------------------------------------------|
| 7  | PSB Alpha P3 speakers  | Use: standard mid-plane                                  |
| 4  | PSB CS500 speakers     | Use: upper height speaker canopy                         |
| 6  | ART SL 1 Power amps    | Use: 11 channels (100 watts per channel)                 |
| 1  | PSB 10" CSIR (speaker) | Use: Low Frequency Effects (LFE) and bass management     |
| 1  | PSB CSW 500 (amp)      | Use: Low Frequency Effects (LFE) and bass management     |
| 18 | ISO Acoustic isolators | Use: Decouple speakers from the room                     |
| 3  | Merging T.—Anubis      | Use: Center desk monitor control, (AD-DA-12 in X 24 out) |
| 1  | Merging T.—Hapi MK II  | Use: AD-DA—16 in X 18 out                                |
| 2  | ATC-SCM 20-ASL Mk II   | Use: Critical stereo mixing and mastering                |
| 1  | Grado RS II Reference  | Use: Critical headphones mix and master                  |
| 1  | Apple AirPods Max      | Use: Critical Apple Spatial binaural reference           |
| 1  | Virtuoso               | Use: Binaural rendering customized headphone responses   |
| 1  | Dolby Atmos Renderer   | Use: Processing of Atmos mixes                           |
| 1  | Atmos Album Assembler  | Use: Finalizer for Atmos album delivery                  |
| 1  | Logic Pro              | Use: Digital Audio Workstation (DAW)                     |



|   |                            |                                                            |
|---|----------------------------|------------------------------------------------------------|
| 1 | Reason (DAW & Rack)        | Use: Digital Audio Workstation (DAW)                       |
| 1 | Pro Tools Ultimate         | Use: Digital Audio Workstation (DAW)                       |
| 2 | Bryston BD 225 (installed) | Use: Replacement amps for the SLA 1 amps (increased power) |

**Table 16:** Hyper Near-Field Tiny Studio equipment list (10\_05\_25).

The list is dated 04\_02\_25 and shows new purchases that have taken place since the initial studio build. Two more Anubis' have been added to the Audio Over Internet (AOI) network which provide additional headphone monitoring and custom mixing at performance/recording stations.

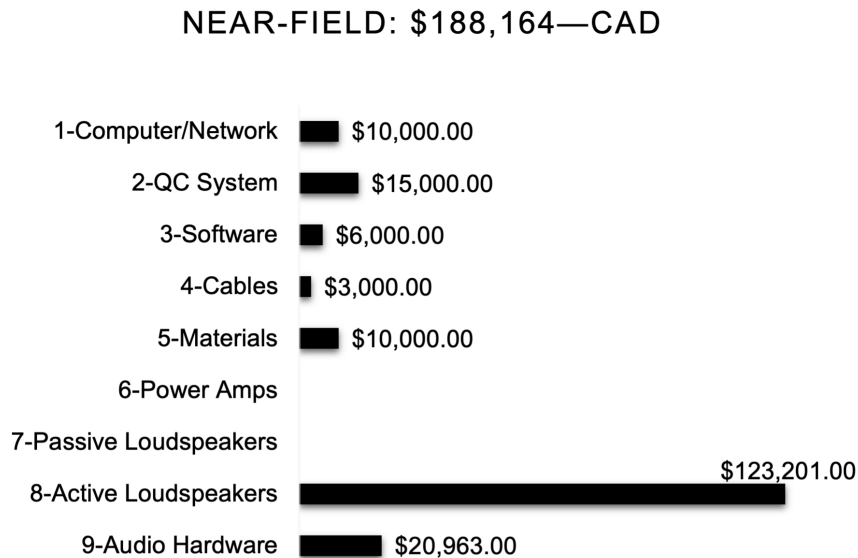
The ART SLA 1 amplifiers were suitable for the initial Hyper Near-Field Tiny Studio build, but the new Bryston BD 225 Class D amplifier has better performance specifications.

Each 5.25-inch Bryston chassis houses 6 amplifiers with a frequency response of 1 Hz to 45 KHz. The power rating for each amp is 225 watts and it's available to each channel with 8 ohms of speaker load impedance. Continuous RMS wattage available to all channels is 1250 watts and dynamic wattage to all channels is 3750 watts with 100ms of latency. According to Ian Calhoun, CEO at Bryston, this amplifier design generates almost no heat and sips electricity making it a perfect fit for the ecological considerations of the Hyper Near-Field Tiny Studio philosophy. The economics of these high-end Bryston power amps do increase the base cost of this author's studio by \$8,976.00 bringing the cost to \$62,582.00. These amps bring the studio to an *esoteric audiophile level*. The *premium level* HN-FTS economic comparison in Chapter 7.24 does not include these next generation Class D Bryston power amplifiers.

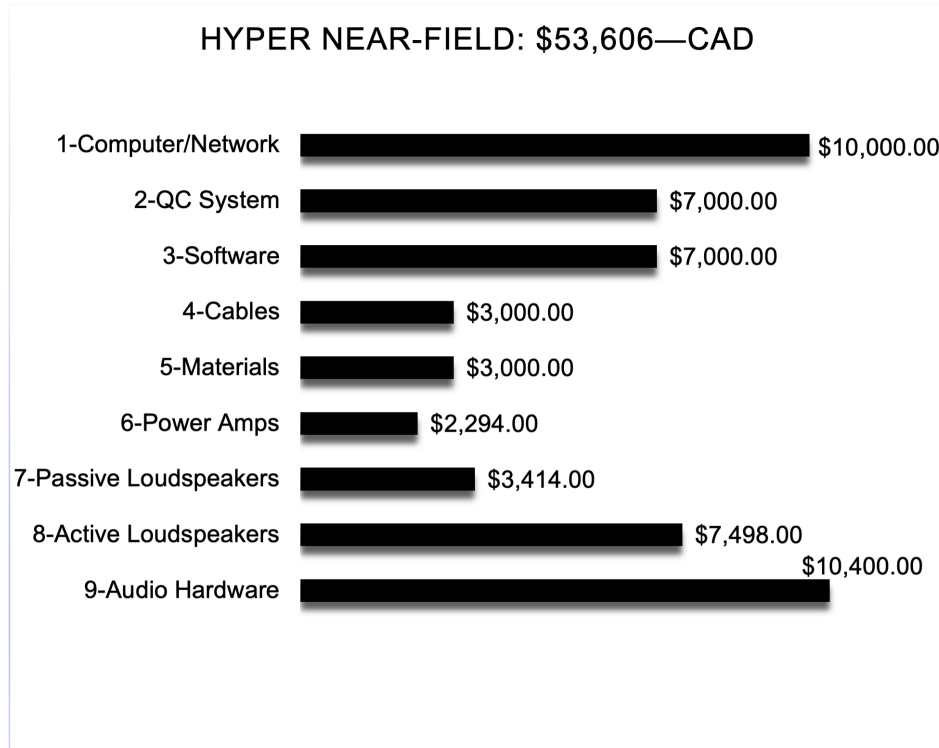
## 7.24 Economic Comparison

This Hyper Near-Field Tiny Studio costs about 35% of a virtually designed (DARDT) Near-Field Atmos control room. Near-Field loudspeakers use 65% of the budget, while this Hyper Near-Field loudspeaker system uses about 20%. Loudspeakers are the biggest cost advantage of a Tiny Studio because the sole mix-position is hyper near, offering affordable timbral neutrality at a high SPL. Premium analogue-digital-analogue (ADA) hardware is expensive for both control rooms, but this Tiny Studio economizes by prioritizing mastering grade i/o—16X18 Hapi MkII, 4X8 i/o Anubis, all costing approximately 50% less than Near-Field 32X32 i/o. The B chain is the most expensive chain in both rooms.

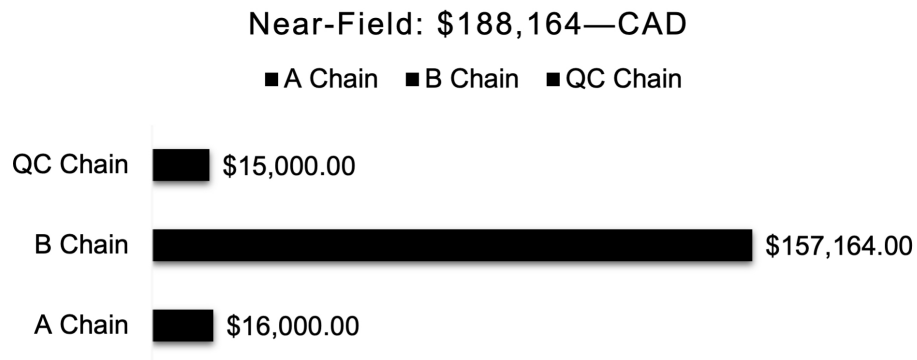
The following bar-graphs (**Figures 27, 28, 29, and 30**) provide estimations for economic comparison between Near-Field and Hyper Near-Field room types. Pricing was done in 2023.



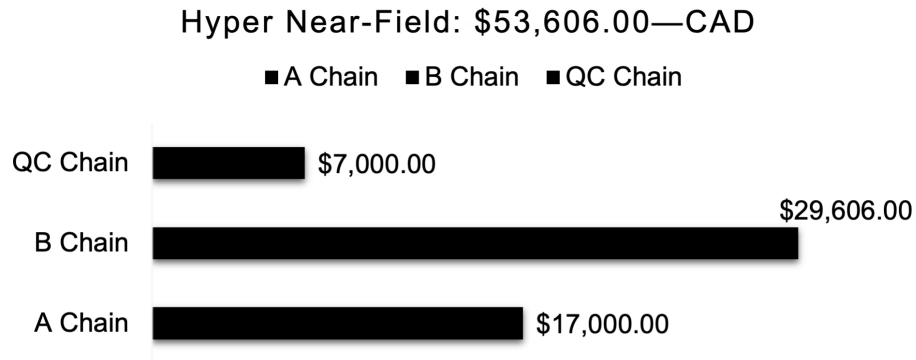
**Figure 27:** Near-Field economic estimate.



**Figure 28:** Hyper Near-Field economic comparison.



**Figure 29:** Near-Field economic breakdown by chain.



**Figure 30:** Hyper Near-Field economic breakdown by chain.

### 7.25 Episode Summary and Conclusion

This HN-FTS philosophy establishes a premium level small control room benchmark which is aimed toward independent professional, semi-pro, higher music education institutions, and commercial facilities. This HN-FTS control room is approximately 50% smaller than Dolby's minimum Near-Field guideline. Below is an eleven-point summary of attributes.

1. Meets Dolby's premium level music SPL specification of 82 dBC +20 dB TP. (85 dBC +20 dB TP is possible and requires 898 watts for the sub and 126 watts per L/R front speaker). The esoteric level Canadian-made Bryston BD 225 class D Atmos amplifiers provide 12 channels, each with 225 watts of continuous power, affording the HN-FTS ample ability to easily surpass Dolby's 85 dBC +20 SPL specification. The total HN-FTS combined power of the Bryston and PSB amplifiers are 2975 watts—about 4 mechanical horsepower.
2. Single operator, 4m<sup>3</sup> speaker sphere, timbrally neutral speakers, low distortion, wide dynamic range.
3. Provides a highly correlated, hyper-tuned and voiced playback system.
4. Compared to Near-Field, reduces control room cost by approximately 75%.
5. Can achieve a noise compliment equal to or below Dolby guideline of NC-25.
6. Provides essential room decorrelation and crosstalk for externalised control room monitoring.

7. Subjectively, this Tiny Studio sounds *tight*, like a big set of 3D headphones.
  8. Has a low  $RT_{60}$  time which is valuable for judging playback in an Atmos automotive cabin and makes electronic room tuning highly effective.
  9. With careful subwoofer selection, placement, electronic tuning, and a rear door bass trap in a rectangular room, this HN-FTS can render standing waves suitably innocuous at the sole mix-position.
  10. Workflow utilizes headphones, discrete speaker playback, and comparative field trips.
  11. Offers low-energy consumption, quiet and cool studio operation.
  12. Encourages individual solutions for large-scope, small-scale, affordability.
- Increased sustainability and upgraded sound (to an esoteric audiophile level) is within reach when using the Bryston BD 225 amplifiers.

A Tiny Studio may not meet “first class acoustics” (Newell, 2018), it won’t be listed in the Dolby.com database of approved Atmos studios. Currently, Dolby does not approve Atmos music studios.

This control room does provide first class control room sound. It works extremely well and has been auditioned by eminent Dolby Atmos 3D music creators Richard Chycki, Justin Gray, Ron Searles, Micheal Francis Nunan, Ryan Mc Cambridge and others. It is the smallest Atmos control room imaginable. It provides a cozy aesthetic with a scientifically tuned playback system that translates reliably to outside larger Atmos playback environments. In an onboarding email with Dolby, Ceri Thomas responded,

*“We do have rooms which are, as you describe, Hyper Near-Field, as they do not meet our best practices, they would not be eligible for listing on Dolby.com. That doesn’t mean you can’t still personally create in them.”*

(Dolby Labs/Thomas, 2022)

In conclusion this author advocates for further research and development of the HN-FTS philosophy and benchmark, inviting Pro Audio manufacturers to recognize the Atmos Hyper Near-Field Tiny Studio as a viable category for specific development. This author argues that independent music creators can confidently assemble their own 3D Atmos audio HN-FTS which is distinguished by the levels of premium or esoteric audiophile sound quality, thrifty economics, and a well-tuned sole mixing position. This HN-FTS benchmark introduces sturdy and adaptive solutions that consider essential music business and socio-economic context which routinely require the independent music creator, studio owner or institution to do more with less.

## Chapter 8: *Found in Space*, Review and Summary

---

### 8.1 Organization of the Appeals, Summary of The Metaphysical Components

The appeals of *Found in Space* are complex due to their interdisciplinary nature. The episodic format breaks complex discussions into three manageable themed narratives and invites the reader to order their reading as they please, but a top to bottom read is recommended.

The episodic format engenders three consequential journeys that accumulate knowledge about 3D popular music spatiality in the Dolby Atmos format for composers, songwriters, recordists, mixers and pedagogues. This chapter will review and summarize *Found in Space* with a diagrammatic map of the substantial metaphysical components.

In Episode 1 this author's appeals form an ethos for the approach toward 3D spatial music composition. Episodes Two and Three contain the supporting logos as reasoned rationales which provide separate examinations of salient metaphysical componentry , along with a studio design philosophy and benchmark. The Pathos of *Found in Space* is provided as music examples with subjective evaluations derived from this author's use of his 3D Spatial Analysis Synergy. See [Chapters 6.16 to 6.20](#). Please hear the music on this author's [Apple Music Playlist](#).

The general approach and organization of associated scenarios, concepts, vocabulary and terminology build on William Moylan's framework, whilst semantic concept maps previously published by Francis Rumsey and others have also been informative.

### 8.2 The Objective

The objective of *Found in Space* is to expose theoretical rationales that support the substantial components of compelling spatiality in 3D popular recorded music. It does this by

gathering research from a wide body of interdisciplinary literature. Using metaphor as a way to know and then locating fundamental concepts, vocabulary, and terms that can adjacently crossover and extend to epistemic intertextual knowledge, this author seeks to present his theories by gathering the components into a diagrammatic one-page summary. [See Chapter 8.9.1.](#)

This author's approach is multifaceted and amalgamates specifically chosen aspects of cognitive image schemata, spatial context, listening experience, mind state, emotional outcome, psychoacoustics, playback system correlation, creative decorrelation, and biological induction of sound. As such, an admixture of citation styles is utilized.

Rationales by David Byrne and David Huron are central to the *Found in Space* ethos. Byrne unpacks artistic music context, spatial context and creativity while Huron theorizes cognitive temporal musical expectation, experience, and their resultant emotions. This author argues that involving these rationales helps to guide effective spatiotemporal orchestration of audio-objects in the Dolby Atmos format.

*Found in Space* assembles the salient semantic descriptors for the 3D spatiality of music which one day soon may form metacontextual descriptors for immersive consumer media and devices.

### 8.3 The Need

Due to the nascent stage of 3D spatial popular musicology and its theories, this author believes that established stereo music creators and students need durable and precise pedagogy to guide their transition of praxis toward 3D popular music. Many stereo music creators are lost in the expanded range of possibilities and opportunities offered by 3D spatiality, meanwhile



eminent 3D music mixers like Hans Martin Buff,<sup>55</sup> Steven Wilson, Morten Lindberg, Justin Gray and others are flourishing in the Atmos format. Hans Martin Buff enlivens the meaning of songs with his compelling 3D staging and spatial dimension, and his binaural headphone mixes translate well via mainstream devices like mobile iPhones, iPads, and computers. To date seasoned 3D music creators intuitively worked out their approaches but now *Found in Space* shares decades of this author's stereo, surround sound, and 3D experience in association with his scholarly research to reveal the compelling substance of 3D music spatiality.

## 8.4 Principles for Spatial Analysis (Review)

Experience has led this author to pinpoint 7 principles as a way to set the stage for 3D spatial analysis of recorded music. Please [See Chapter 6.3](#).

## 8.5 The Acoustemological Components

*Found in Space* organizes the substantial componentry of 3D spatiality into the two primary categories of Acoustemology and Techoustemology. For purposes of musicological 3D music spatial theory, the domains of acoustics, acoustic ecology, induction, psychoacoustics/cognition, popular musicology, ethnomusicology, music technology and consumer audio visual media and devices locate well between Acoustemology and Techoustemology. Context, spatial context, spatial recontextualization, and spatiotemporal orchestration are essential concepts for the creation of compelling 3D popular music.

---

<sup>55</sup> Hans Martin Buff won a 2025 Grammy Award in the category of Best Immersive Album for his mix of Peter Gabriel's 2024 release entitled *I/O (In-side Mix)*.

Four questions are central to the contextual determination of spatial substance: 1) *What*; 2) *When*; 3) *Where*, and 4) *Why*. Answers to those questions can be guided by David Huron's ITPRA theory: 1) Imagination response; 2) Tension Response; 3) Prediction Response; 4) Reaction Response, 5) Appraisal Response.

Eight acoustemological components are substantial to 3D spatial music theory; 1) contexts; 2) listener expectation; 3) listener experience; 4) listener emotions; 5) positional relationships; 6) cognitive perspectives; 7) proportional properties; and 8) character attributes.

## 8.6 The Techoustemological Components

To induce and perceive the acoustemoligical components we also must ask *how* to do it. For this, an Atmos studio with a highly correlated speaker gamut is essential. For this six components are substantial: 1) acoustic room tuning to absorb resonances below the Schroeder frequency and diffuse soundwaves above, resulting in an innocuous influence of room character; 2) uniform electronic system polarization; 3) precise system clock with minimal drift to eliminate phase distortion; 4) speaker gamut amplitude correction; 5) speaker-delay timing; and 6) speaker equalization with bass management that correlates frequencies below about 80 Hz by sending them to a single monaural subwoofer. Once these conditions are met the question of *how* looks at creator praxis, a topic that this dissertation does not delve into in detail. However, this author is aware of other Ph.D. researchers like Marcella Rada at the University of Surrey who is focused on “guidelines, workflows, methodologies and practices” for spatial audio.

## 8.7 The Super Components—SW, ASW, ES, AEW, LEV

Under the heading of Auditory Spatial Impression (ASI) this author proposes that five Super Components can be utilized to efficiently create compelling 3D spatial music. The rationale for these dimensional Super Component impressions is found in durable concert hall acoustic design principles. With well-considered application into a 16-channel 9.1.6 Atmos O-bed, which resolves into 7.1.4 via the Atmos renderer, a clean sound field can emerge efficiently.

The Super Components are: 1) Source Width (SW); 2) Apparent Source Width (ASW); 3) Ensemble Width (EW); 4) Apparent Ensemble Width (AEW); and 5) Listener Envelopment (LEV). This author suggests that a rudimentary method for students to learn how the Super Components can benefit compelling spatiality involves placing direct sound—the SW and EW—in the onstage L, C, R, speaker zone. Then, cascaded stereo reverbs can add separate impressions of ASW and AEW which are highly effective when placed in the Left- Right-wide (anteriorly horizontal) speaker zone. As Heinrich Kuttruff argues—(see [Chapter 6.8](#) and Damaske, 1967/68; Baron, 1981, 1981, Kuttruff, 2016)—compelling LEV is best achieved with decorrelated indirect sound, created by using a delay between 0.5 ms and about 80 ms, and/or reverb and placed in the Left- Right-side (lateral) speaker zone. This concept adds compelling and tasteful immersive weight when applied in a critical balance with direct sound. (Please see [Chapter 6.9](#), **Figure 8**).

For the emerging 3D music creator precise design of their relationships, perspectives, properties and attributes within the dimensions, SW, EW, ASW, AEW and LEV quickly form durable impressions of 3D spatiality resulting with a strong sensation of spatial envelopment, dimensional separation, depth, impact, and intelligibility. The Super Components strongly suggest *what to do with all those speakers*.

When mixing in 3D, the audio-object's Source Width is a hefty determinant of direct sound impact and intelligibility, but it must be done carefully when spreading across multiple speakers to avoid undesirable energy buildup from excessive correlation. The L, C, R, front speakers will often be suitable for SW, and additionally the front height speakers can extend SW into Source Height (SH). An example of extended pitch/height is heard when elevating drum kit cymbals into the upper front top canopy.

#### 8.7.1 A Highly Unusual Use of Vocal Source Width (SW)

Worthy of highlighting in this review is an unusual example of compelling vocal Source Width staging in Sidu Moose Wala's song "Jaat Du Maqabala." Sidu's direct vocal sound is placed at center stage via the C speaker, and offstage in the lateral Ls/Rs speakers to create a triangular 3-speaker Source Width zone. This puts the direct vocal sound in front and beside the listener. Placed onstage via routing to the front horizontal L/R speakers, the vocal reverb (indirect sound) creates the impression of Apparent Source Width (ASW). This spatiotemporal orchestration creates a highly personal proxemic lead vocal zone and successfully breaks away from normalised 3D praxis for vocal staging in popular 3D music. This author suggests that Atmos song mixers should listen closely to this mix. It can be found in this author's [3D music playlist on Apple Music](#).

#### 8.8 The Order of Operation for 3D Music Analysis and Praxis

3D music spatiality is complex, so this author argues that the associated pedagogy can be reduced to a 3-step order of operation. 1) Train the brain to perceive and understand the [five-facets](#) in mono, stereo and then 3D; 2) Understand [1\) Relationships; 2\) Perspectives; 3\)](#)

[Properties and 4\) Attributes](#). Practice 3D spatial creation with the [Super Components using a 9.1.6 O-bed](#); and 3) Analyse the results with the Five-Facet Typology in 3D. An ITPRA analysis can be revealing because it focuses on the linear development of spatiotemporal orchestration. Please see [Chapter 6.19](#), ITRPA analysis of “Rocket Man”.

3D mix training can begin in mono, and advance to stereo (2.0), 2.1 (L, C, sub), 3.1 (L, C, R, sub), 5.1, 7.1 and finally arrive at mixing in 7.1.4.

A 9.1.6 object-bed provides the highest practical resolution of 3D spatiality for mainstream popular music, and the Dolby Atmos renderer will resolve the 16 channel O-bed into the 7.1.4 speaker gamut. Channel muting or soloing is an essential way to understand spatiotemporal orchestration. Analysis examines speaker zone architecture and dimensional sound field cohesion between the standard mid-plane and the upper height canopy speakers.

### 8.8.1 Cross-Domain Rationales

Musicological research and discourse about 3D popular music production sits at the intersection of musicology and technology. Natural acoustical concepts associate closely with virtual 3D spatiality and cognitive awareness of environment. These associations form an ethos of cross-domain rationales that underpin research, studio design, philosophy, benchmark, ear training, subjective analysis, and praxis.

### 8.9 3D Diagrammatic Spatial Substance Summary Map, (Spider Sheet)

**Figure 31** is a diagrammatic Spatial Substance Map (Spider Sheet) that aggregates the substantial constituents to form a DNA equivalent for 3D music spatiality. The Spider Sheet can be used to stimulate new branches of inspection that seek other undiscovered theories and

answers to the central questions: 1) What is the substance of compelling spatiality in immersive 3D Dolby Atmos music recordings? and 2) How can we find it?

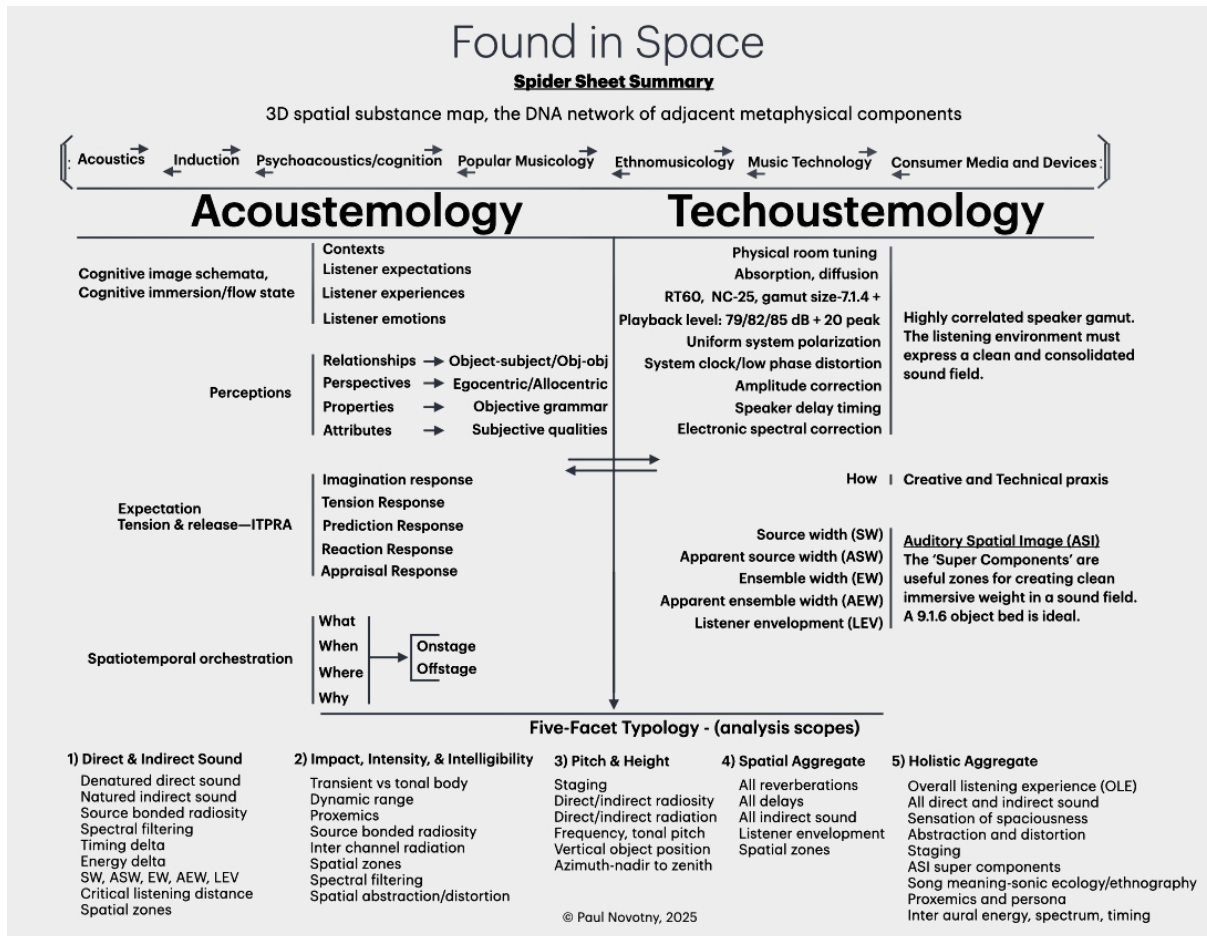
It can function as an idea generator, with a process that serially joins or combines selected textual components. It urges the creator to imagine, theorize, and describe bespoke subjective 3D spatial examinations, conceptualizations, contexts, and perceptions of their own 3D music mix.

A primary take-away from these journeys is that the substantial metaphysical components of 3D spatiality are abstract and reside in the mind. They are illusory metaphors that can be described with intertextual rationales, drawn from the domains of well-established research literature to form new ways of knowing about music's 3D spatiality. Some examples of serial component combinations are:

- Examining Contexts combined with Source Width (SW).
- Examining Listener experience combined with Listener Envelopment (LEV).
- Examining Attributes combined with Proxemics
- Examining What combined with Ensemble Staging
- Examining When combined with dynamic audio-object location such as Pitch & Height, azimuth-nadir to zenith, or spectral abstractions and distortions.
- Examining Attributes combined with Direct and Indirect Sound

[See Figure 31.](#)

### 8.9.1 The 3D Diagrammatic Spatial Substance Map



**Figure 31:** 3D diagrammatic spatial substance map, (spider sheet).

### 8.10 The Hyper Near-Field Tiny Studio

Metaphorically speaking, travelling to the third and fourth dimensions (X, Y, Z, T = *time*) of music spatiality requires a robust space vehicle. Designing and building the Hyper Near-Field Tiny Studio is a form of technological luthiership that results with a durable music-spacecraft.

It is a hyper-tuned instrument that took this author two-years to research, design and build in order to conceive of the research contained in *Found in Space*. Once the vehicle was operational, learning to perceive the metaphysical components described in **Episodes 1** and **2**

became possible. **Episode 3** articulates his philosophy and benchmark for this Dolby Atmos

7.1.4 Hyper Near-Field Tiny Studio which has now been verified by specialists as a unique and accurate music studio control room.

The substantial considerations for this small room involve: 1) architecture; 2) acoustics; 3) electro-mechanics; 4) environment; 5) ergonomics and tactile large-scope operation in a small-scale; 6) reliability, and transference; 8) economics; 9) hyper calibration; 10) meeting Dolby guidelines for sound pressure level (SPL); and noise compliment at -25 dB (NC-25); and quality control (QC).

### 8.11 The Economics of the HN-FTS

Affordable accessibility to the Atmos format is challenging for the next generation of 3D music creators. In response, this author's Hyper Near-Field Tiny Studio costs about 35% of a virtually designed<sup>56</sup> Near-Field Atmos control room. Near-Field loudspeakers use 65% of the budget, while this Hyper Near-Field loudspeaker system uses about 20%. Loudspeakers are the biggest cost advantage of a Tiny Studio because the sole mix-position is hyper near (1.2m), offering affordable timbral neutrality at Dolby's specified cinematic playback SPL of 85 dBC.

### 8.12 General Discussion, Impact, and Contribution of Knowledge

During this five-year research journey, many aspects of the work have been presented at international conferences and shared with some of the world's leading experts in the Dolby Atmos professional and academic communities. The empirical result of this author's Hyper

---

<sup>56</sup> Virtual Design is done with the Dolby Room Design Tool (DARDT).



Near-Field Tiny Studio indicates that music playback in 2-channel stereo to 12-channel Atmos sounds as good as the verified aggregate response curve indicates. [See Chapter 7.22.](#)

Over time this author hopes his 3D analysis synergy may contribute viable pedagogy for 3D music analysis, creation and production. This author hopes that *Found in Space* will find its place amongst the existing body of musicological, technological and music composition literature. It is uniquely situated as cross-domain, intertextual, and interdisciplinary research about 3D music spatiality. Bosun Xie notes, “During the past 40 years, especially since 1990, spatial sound has been developed greatly. Current spatial sound techniques differ considerably from those in 1970 in many aspects of basic physical, auditory principles, and technical means. Therefore, a book on the principles and applications of spatial sound should be rewritten” (Xie, 2023, p.xvi).

## Chapter 9: Conclusion

---

### 9.1 The Skeptics

Currently, mainstream 3D Music listening experiences are not well understood. Now is a time akin to the early 60s when recorded music was transitioning from mono to stereo.

*“So, The Beatles spend 7-8 months on Sergeant Pepper, including sitting in on the mixes, which is all mono, and then they “---- off” on vacation because they're done and the production guy and the engineers sit down for three hours and make the whole thing in stereo, in some way, just because they have to. Fast forward 15 years, the mono thing is out, and that really half-assed stereo mix that nobody was there for is the main thing, and people think that's what it [the format] is. So that's why I'm saying ... even if you hate it [Atmos], and you think [in]stereo, you're [we're] done with the stereo, think about it, ... in 10 years this may be the version that people are going to listen to ... so we might as well put some love into it [our Atmos mixes], or you know, at least support it [the format], or kind of get your head around it [Atmos], see how it works”*

(Wendl, Buff, 2024)

This author has heard arguments from Atmos skeptics asserting that humans only have two ears, and twelve speakers are just not necessary. Of course, this argument is not true and perhaps *Found in Space* may also help to reconfigure misunderstandings about 3D music and the Dolby Atmos format for music.

*Found in Space* is intended for the next generation of 3D music creator. Professionals and students eagerly yearn for deeper understanding of theories about 3D popular music spatiality to avoid problems with their art and craft and maximize myriad opportunities that can captivate audiences with compelling immersive music experiences. Atmos format progression is rapid and soon the general public will have even easier access to 3D music in the Dolby Atmos enabled car cabin—possibly a new 3D main-stage for the next generation of 3D music creators.

## 9.2 Ending the Regressive Listening Trend

In about 1550 to 1585, during the High Renaissance, the *maestro di cappella* of St. Mark's Cathedral in Venice, Flemish composer Adrian Willaert dared to physically separate vocalists. He “used a chorus in left and right parts for antiphonal singing, matching the two organ chambers oriented to either side of the alter” (Holman, 2008, p.2). Adrian Willaert and Giovanni Gabrieli were the first composers who spatially separated choral parts, and it is only now that music listeners who travel on a bus, subway, plane, or drive in their Atmos enabled car cabin can experience spaces and places as separated audio-objects, spatiotemporally orchestrated in 360 degrees, perhaps similarly to how Willaert and Giovanni staged their choirs.

Despite the Atmos format requiring a costly proprietary licence for device manufacturers,<sup>57</sup> the format's palpability when heard through speakers or binaurally on headphones appears capable of reversing *technologically regressive listening trends* brought about by weak file formats like the low resolution .mp3. Corporate heavy hitters like Netflix, Amazon, Apple, Tidal, and others are committed to the Dolby Atmos format. The Apple Spatial Audio Format (ASAF) which includes Dolby Atmos, uses the Apple Positional Audio Codec (APAC) so Apple devices can encode head-tracked, multi-source audio with an iPhone 16. Dolby Atmos and Apple Spatial appear to have a bright future in comparison to previous surround sound formats.

---

<sup>57</sup> Manufacturers must purchase a licence to place the Atmos renderer in their devices.

### 9.3 The Development of Open-Source 3D Audio

Recently Samsung and Google have announced Eclipsa, based on the Immersive Audio Model Formats, (IAMF). Eclipsa is a royalty free open media format endorsed by the Alliance for Open Media (AOM). Eclipsa arrived on Samsung TVs and Soundbars in 2025 and it will support up to 28 fixed or dynamic channels which are rendered to output speakers, or headphones for binaural playback. No matter what renderers come into the marketplace, the 7.1.4 audio format has been adopted by the Audio Engineering Society (AES) and the Society of Motion Picture and Television Engineers (SMPTE) as the new standard for 3D spatial audio archival. 3D music seems to be here to stay.

### 9.4 Preparing the Next Generation of 3D Music Creator

Perhaps the most noble endeavour of *Found in Space* may be to assist the next aspiring Hans Martin Buff when realising compelling 3D music in their own Dolby Atmos 3D Hyper Near-Field Tiny Studio, and as Hans says, “put some love into it” (*ibid*).

Some recent Atmos releases (2025) this author recommends as tour-de-force art and craft would be Steven Wilson’s *The Overview*, Grammy nominated Trio Mediaeval, *Yule*, Justin Gray’s Grammy nominated release entitled *Immersive*, Max Richter’s minimalist (Classical Crossover) release entitled *In a Landscape*, and Peter Gabriel’s 2024 Grammy winning *I/O (In-Side Mix)*.

It is my hope that future 3D music researchers will build upon *Found in Space* and embark on new journeys of their own that endeavour to discover other approaches and theories that have not yet been found.

In the near future, mature software and hardware solutions that make the creation and distribution of Atmos music easier will become more prominent. The mainstream consumer requires reliable and easy ways to experience spatial music in the home using speakers. Samsung's HW-Q990F 11.1.4 3D soundbar system with rear speakers and a subwoofer is a good example of what is currently possible for externalized home theatre 3D music playback.

For musicologists, I hope that many more “Journeys to the Third Dimension of Music Spatiality in the Dolby Atmos format,” and other spatial audio formats lie ahead. Bon voyage!

## BIBLIOGRAPHY · Episode 1 (Pilot)

- Adorno, T., (1991). *The Culture Industry: Selected Essays on Mass Culture*, “On the Fetish Character in Music and the Regression of Listening,” London, Rutledge.
- Adorno, T., Eisler, H., (2003). *Movie Music: The Film Reader*, “Prejudices and Bad Habits,” Kay Dickenson, ed., Routledge, London.
- Adorno, T., Eisler, H., (1947, 2019). *Composing for the Films*, Continuum, ISBN-10 0826499023, ISBN-13 978-0826499028.
- Ahrens, J., (2012). *Analytic Methods of Sound Field Science*, Springer-Verlag, Berlin Germany.
- Ambrose, C., (2003). *Frequency Range of Human Hearing*, *The Physics Fact Book, An Encyclopedia of Scientific Essays*, [Online], Available: <https://hypertextbook.com/facts/2003/ChrisDAmbrose.shtml> [Accessed 10/24/24].
- Antoniuk, P., Zieliński, S., Lee, H., (2024). *Ensemble Width Estimation in HRTF-Convolved Binaural Music Recordings using an Auditory Model and a Gradient-Boosted Decision Trees Regressor*, EURASIP Journal on Audio, Speech, and Music Processing. Springer Open.  
<https://doi.org/10.1186/s13636-024-00374-2>
- Barron, M., (1971). The Subjective Effects of First Reflections in Concert Halls, J Sound Vibr; 15:475.
- Barron, M., Marshall, A.H., (1981). *Spatial Impression Due to Early Lateral Reflections in Concert Halls*, The Derivation of Physical Measure. J. Sound Vib. 77(2),211-232,  
[https://doi.org/10.1016/S0022-460X\(81\)80020-X](https://doi.org/10.1016/S0022-460X(81)80020-X)
- Beard, D., Gloag, K., (2016). *Musicology, The Key Concepts*, Second Edition, Routledge, Taylor & Francis New York.
- Begault, D., R., (1994). *3D Sound for Virtual Reality and Multimedia*, Academic Press Professional, Cambridge, MA.
- Bie, O., (1899). *History of the Pianoforte and Pianoforte Players*, J.M. Dent & Company, London.  
[online website available at <https://www.gutenberg.org/ebooks/71333>], accessed 11, 26, 2024.
- Blauert, J., (1983, 1997). *Spatial Hearing, The Psychophysics of Human Sound Localization*, The MIT Press, Harvard MA.
- Blumlein, A.D. (1931). “British Patent Specification 394, 325,” reprinted in the Journal of the AES, (1958) 6, p. 91 and in the JAES anthology Stereophonic Techniques.
- Bregman, A., S., (1993). *Auditory Scene Analysis, Hearing in Complex Environments*, in Thinking in Sound: The Cognitive Psychology of Human Audition, ed. Stephen Mc Adams and Emmanuel Bigand, Clarendon Press, 10-36, Oxford.

- Brøvig-Hanssen, R., Danielsen, A., (2013). *The Naturalized and the Surreal: Changes in the Perception of Popular Music Sound*, University of Oslo, Department of Musicology, Box 1017, NO-0315, Oslo, Norway.
- Byrne, D., (2012). *How Music Works*, Mc Sweeney's, San Francisco.
- Carrizosa, P., (2021). *Could Tiny Homes be the Adorable, Affordable and Sustainable Housing that our Planet Needs?* (Website), available online from, <https://ideas.ted.com/impact-of-tiny-homes-on-the-environment-and-affordable-housing/>.
- Cobussen, N., Nielsen, M., (2016). *Music and Ethics*, Ashgate, Routledge, Taylor Francis Group, London and New York.
- Cox, C., Warner, D., (2004). *Audio Culture Readings in Modern Music*, Chapter 14, Chapter 7, Continuum, New York, London.
- Csikszentmihalyi, M. (1990). *Flow, The Psychology of Optimal Experience*, Harper Perennial, New York.
- Csikszentmihalyi, M. (2004). *Pursuit of Happiness* Ted Talk\_Flow [online website accessed 11,13,24], <https://www.pursuit-of-happiness.org/history-of-happiness/mihaly-csikszentmihalyi/#:~:text=Mihaly%20Csikszentmihalyi%20discovered%20that%20people,which%20involves%20their%20creative%20abilities.>
- Csikszentmihalyi, M. (2014). *Flow and the Foundations of Positive Psychology*, The Collected Works of Mihaly Csikszentmihalyi, Springer, ISBN set: 978-94-017-9083-3.
- Damaske, P., (1967/68). *Subjective Investigations of Sound Fields* (in German), *Acoustica*; 19:199
- Dayal, G., Ferringo, E., (2012). *Electronic Dance Music (EDM)*, Oxford Music [Online], Available, [http://www.oxfordmusiconline.com.ezproxy.york.ac.uk/subscriber/article/grove/music/A2224259type=article&search=quick&q=synthesiser+bass&pos=16&\\_start=1#firsthit](http://www.oxfordmusiconline.com.ezproxy.york.ac.uk/subscriber/article/grove/music/A2224259type=article&search=quick&q=synthesiser+bass&pos=16&_start=1#firsthit) [Accessed 11,18,24]
- Dientsfry, E. (2016). *The Myth of The Speakers: A Critical Re-examination of Dolby History*, Film History, Indiana University Press, 28, 1, pp. 167-193, [www.jstor.org/stable/10.2979/filmhistory.28.1.06](http://www.jstor.org/stable/10.2979/filmhistory.28.1.06)
- Dolby, (2020C). *Greg Penny on Mixing “Rocket Man”* <https://www.youtube.com/watch?v=ggzfcUKDqdo> [website accessed 11,13,24]
- Dolby, (2022D). *Dolby Atmos Music Creation: Studio Setup and Tuning*, Pennington, B., 10:10, <https://www.youtube.com/watch?v=pPDY6mXBmgs> [Website accessed 05, 26, 2022]
- Dolby, (2024A). *The New, Electric Volvo EX90 SUV to Feature Dolby Atmos*, [website accessed 11,11,24] <https://professional.dolby.com/music/dolby-atmos-for-cars/volvo/#gref>

- Dolby, (2024B). *Rush in Atmos – Reimagining the Classics*, #DolbyCreatorTalks [website accessed 11,11,24] <https://www.youtube.com/watch?v=DMjwsdzdxIQ>
- Faconnier, G., (1997). *Mappings in Thought and Language*, Cambridge: Cambridge University Press.
- Fiebig, A., Popper, A. N., Fay, R. R., Schulte-Fortkamp, B., & Sisneros, J. A. (2023). Preface 1992. *Soundscapes: Humans and Their Acoustic Environment* (Vol. 76). Springer International Publishing AG.
- Feld, S., (1994). “*From Ethnomusicology to Echo-Muse-Ecology: Reading R. Murray Schafer in the Papua New Guinea Rainforest.*” *Soundscape Newsletter* 8: 9-13.
- Gerzon, M.A. (1992). *Applications of Blumlein Shuffling to Stereo Microphone Techniques*, *Journal of the AES*, #3448
- Graven, N. S., Browne, V. J., (2008). *Auditory Development in the Fetus and Infant*, Department of Community and Family Health, College of Public Health, University of South Florida, Tampa, FL; University of Colorado Denver School of Medicine and The Children’s Hospital, Department of Pediatrics, JFK Partners, Aurora, CO., Elsevier Inc., 1527-3369/08/0804-0278\$34.00/0, doi:10.1053/j.nainr.2008.010.
- Hall, T. E., (1963). *A System for Notation of Proxemic Behaviour*, *American Anthropologist*, 65/5, 1003-26.
- Hawkins, S. (2016). *Queerness in Pop Music: Aesthetics, Gender Norms, and Temporality*, Routledge, New York.
- Holman, T., (2008). *Surround Sound: Up and Running*, Second edition, Boston, MA: Focal Press, Elsevier.
- Huron, D., (2006). *Sweet Anticipation, Music and the Psychology of Expectation*, A Bradford Book, The MIT Press, Cambridge, Massachusetts, London, England. (iBooks)
- Husserl, E., (1931, 2013). *Ideas, General Introduction to Pure Phenomenology*, George Allen and Unwin, Routledge, Taylor and Francis Group, Muirhead Library of Philosophy.
- Johnson, M., (1987). *The Body in the Mind*, Chicago University Press.
- Käsbach, J., Hahmann, M., May, T. Dau, T., (2016). *Assessing and Modeling Apparent Source Width Perception*, In Proc. of DAGA (Deutsche Gesellschaft für Akustik), Aachen, Germany.
- Kim. Y., H., Choi, J., W., (2013). *Sound Visualization and Manipulation*, John Wiley & Sons, Singapore.
- Klapholz, J. (1991). *Fantasia—Innovations in Sound*, *Journal of the AES*, #5994, 39
- Kotler, S., Mannino, M., Keslo, S., Huskey, R., (2022). *First few seconds for flow: A Comprehensive Proposal of the Neurobiology and Neurodynamics of State Onset*, *Neuroscience and Biobehavioral Reviews*, 143, 104956, Elsevier.



- Kauffman, S., (2013). *What is life, and can we create it?* BioScience, August/Vol.63 No 8
- Kauffman, S., (2020). *Answering Schrödinger's "What is Life?"* Entropy, 22, 815; doi:10.3390/e22080815
- Kuttruff, H., (2016). *Room Acoustics, Sixth Edition, Institute of Technical Acoustics*, Aachen University, Germany, Routledge, CRC Press, Taylor & Francis group, Boca Raton, London, New York. (Kindle)
- Lee, H., Millns, (2017). *Microphone array impulse response (MAIR)*, library for spatial audio research, (AES, NewYork, NY, USA) Convention e-Brief 356, In *PROC. of the 143<sup>rd</sup> AES Convention*.
- Lee, H., (2024). *Echo-project*, [website] visited on 12\_15\_24, available at, [https://aplhud.com/echo?fbclid=IwY2xjawHNlQNleHRuA2FlbQIxMAABHTM8akTDESI0m3Btx\\_F2MK5uXWMdiAM1JMW3Dwt\\_gc7UqZZZYbP4WQzXPg\\_aem\\_3bF4r3Pf-ZIAJt0LUYawNw](https://aplhud.com/echo?fbclid=IwY2xjawHNlQNleHRuA2FlbQIxMAABHTM8akTDESI0m3Btx_F2MK5uXWMdiAM1JMW3Dwt_gc7UqZZZYbP4WQzXPg_aem_3bF4r3Pf-ZIAJt0LUYawNw)
- Lee, H., (2024). *The Echo-Project, Projects, Publications*, <https://apl-hud.com/projects/>, <https://aplhud.com/publications/>[https://aplhud.com/echo?fbclid=IwY2xjawHNlQNleHRuA2FlbQIxMAABHTM8akTDESI0m3Btx\\_F2MK5uXWMdiAM1JMW3Dwt\\_gc7UqZZZYbP4WQzXPg\\_aem\\_3bF4r3Pf-ZIAJt0LUYawNw](https://aplhud.com/echo?fbclid=IwY2xjawHNlQNleHRuA2FlbQIxMAABHTM8akTDESI0m3Btx_F2MK5uXWMdiAM1JMW3Dwt_gc7UqZZZYbP4WQzXPg_aem_3bF4r3Pf-ZIAJt0LUYawNw)
- Lomax, A., (1968). *Folk Song Style and Culture*, Transaction Publishers, Republished in 2017 by Routledge, Taylor Francis Group.
- Lutzke, J., Henggl, M., F., (2009). *The Rhetorical Triangle, Understanding the use of Logos, Ethos, and Pathos*, School of Liberal Arts, Writing Center, "Because writers need readers," Cavanaugh Hall, University Library 2125, Indiana University, [www.iupui.edu/~uwc](http://www.iupui.edu/~uwc).
- Meaux, E., Marchand, S., (2021). *Synthetic Transaural Audio Rendering (STAR): A Perceptive 3D Audio Spatialization Method*, J. Audio Eng. Soc., vol. 69, no.7/8, pp. 497-505, (July/August), DOI:<https://doi.org/10.17743/jaes.2021.0015>
- Merchel, S., (2014). *Auditory-Tactile Music Perception* (ISBN: 978-2-8440-3161-4), Technische Universität Dresden, Shaker Verlag, Germany.
- Meyer, L., (1956, 1961). *Emotion and the Meaning of Music*, The Chicago University Press, Chicago, London.
- Middleton, R., (1990). *Studying Popular Music*, Buckingham, UK, Open University Press.
- Moore, A.F., Schmidt, P., Dockwray, R., (2011). *A Hermeneutics of Spatialization for Recorded Song*, Twentieth-Century Music, 6/1, 83-114, Cambridge University Press, doi:10.1017/S1478572210000071.
- Moore, A., (1992). *Rock: The Primary Text, Developing a Musicology of Rock*, Open University Press, Buckingham England.
- Moylan, W., (2020). *Recording Analysis: How the Record Shapes the Song*, Chapter 8, The Illusion of Space as an Element of Recording, Focal Press, Routledge Academic, Taylor & Francis Group.

- Nelson, R., (2022). *A History of Inspiration through Metaphors of Learning, The Height of Teaching*, Routledge, Taylor Francis Group, London and New York.
- New Oxford American Dictionary (2018). *Mac OS*, Apple Inc.,
- Patterson, J., Lee, H., (2021). *3D Audio, Perspectives on Music Production*, Focal Press, Routledge Academic, Taylor & Francis Group.
- Pawlak, A., Lee, L., Mäkiparta, Lund T., (2023). *Spatial Analysis and Synthesis Methods: Subjective and Objective Evaluations Using Various Microphone Arrays in the Auralization of a Critical Listening Room*, IEEE/ACM Transactions on Audio, Speech, and Language Processing, in peer review.
- Porcello, T., (2004). "Afterword." In *Wired for Sound: Engineering and Technologies in Sonic Cultures*, edited by Thomas Porcello and Paul D. Greene, 269-81. Middletown, CT: Wesleyan University Press.
- Rambarran, S., (2021). *Virtual Music, Sound Music and Image in the Digital Era*, Bloomsbury Academic, New York, London, Oxford, New Delhi, Sydney.
- Read, G., (1953). *Thesaurus of Orchestral Devices*, Sir ISAAC Pitman & Sons Ltd., The Pitman Press, Bath, Melbourne, Johannesburg, London, Toronto, Praeger Publishers, Westport CT., Greenwood Publishing Group Inc.
- Reardon, G., Calle, J.S., Genovese, A., Zalles, G., Olko, M., Jerez, C., Flanagan, P., Roginska, A., (2017). *Evaluation of Binaural Renderers: A Methodology*, AES, Convention e-Brief 359, 143<sup>rd</sup> convention, 10\_18-21\_17, New York.
- Rice, T., (2018). *Acoustemology*, University of Exciter, Researchgate, DOI: 10.1002/9781118924396
- Robinson, S., M., (2013). *Chamber Music in Alternative Venues In the 21<sup>st</sup> Century U.S.: Investigating the Effect of New Venues on Concert Culture, Programming and the Business of Classical Music*, University of South Carolina, <https://scholarcommons.sc.edu/etd/1655> , [visited on 10/16/24]
- Roginska, A., Geluso, P., (2018). *Immersive Sound, The Art and Science of Binaural and Multichannel Audio*, Audio Engineering Society presents, Routledge Focal Press. (Kindle)
- Rumsey, F., (2001). *Spatial Audio*, Focal Press, Taylor & Francis Group. Oxford, Auckland, Boston, Johannesburg, Melbourne, New Delhi.
- Rumsey, F., (2002). *Spatial Quality Evaluation for Reproduced Sound: Terminology, Meaning, and a Scene-Based Paradigm*, Institute of Sound Recording, University of Surrey, UK, and School of Music in Piteå, Luleå University of Technology, Sweden, JAES, Vol 50, No 9, 2002, September.
- Sato, S., Ando, Y., (2002). *Apparent Source Width (ASW) of Complex Noises in Relation to the Interaural Cross-Correlation Function*, J. Temporal Des. Arch. Environ. 2 (1), 29-32.
- Schafer, R., M., (1977). *The Tuning of The World*, McClelland & Stewart, First Canadian Edition.

- Schafer, R., M., (1994). *The Soundscape: Our Sonic Environment and the Tuning of the World*, Destiny Books, Rochester, Vermont. ISBN, 9780892814558, 9781594776687. [Online, York Univ. Libraries, EBSCOhost, accessed 11,13,24]
- Schafer, R., M., (2004). *The Music of the Environment*, Chapter 7, Editors-Cox. C., Warner, D., *Audio Culture Readings in Modern Music*, Continuum, New York, London.
- Schaeffer, P. (1966). *Traité des objets musicaux: Essai interdisciplines*, Éditions du Seuil, Paris, ISBN 978-2-02-133541-5.
- Schaeffer, P. (2004). *Acousmatics*, Chapter 14, Editors-Cox. C., Warner, D., *Audio Culture Readings in Modern Music*, Continuum, New York, London.
- Schaeffer, P., (2017). *Treatise on Musical Objects: An Essay Across Disciplines*, University of California Press, Oakland, California.
- Schieffelin, E., [1976] (2005). *The Sorrow of the Lonely and Burning of the Dancers*, New York, Palgrave Macmillan.
- Shapira, E., Pujol, R., Plaksin, M., Kimmel, E., (2016). *Sound-Induced Motility of Outer Hair Cells explained by Stochastic Resonance in Nanometric Sensors in the Lateral Wall*. Vol 2, December 2016, Pages 1-11, Physics in Medicine, Science Direct, Elsevier.
- Sijen, P., (2008). *Design Philosophy: Blue Sky Studio Monitors*, AudioDesignLabs, (website) available from, <http://audiodesignlabs.com> [accessed 28 Aug. 2021].
- Simon, G., Olive, S., Welti, T., (2009). *The Effect of Whole-Body Vibrations on Preferred Bass Equalization of Automotive Audio Systems*, presented at the 127<sup>th</sup> Conv. Audio Engineering Society, New York.
- Smalley, D., (1997). *Spectromorphology: Explaining Sound-Shapes*, Organized Sound 12 (1): 35-58.
- Smalley, D., (2007). *Space-Form and the Acousmatic Image*, Organised Sound, 2(2): 107-26.
- Steele, G., (2008). *Auditory Technologies and the Regression of Listening: The iPod, and the Rhetoric of the Soundtrack*, The University of Rhode Island, The Journal of Technology, Knowledge, and Society, Vol. 4, No 3.
- Théberge, P., (2001). *Plugged In, Technology and Popular Music*, S. Frith, W. Straw, & J. Street, (editors), Cambridge Companion to Pop and Rock (pp.3-25), Cambridge, Cambridge University Press.
- Toole, E. F., (2018). *Sound Reproduction: The Acoustics and Psychoacoustics of Loudspeakers and Rooms*, Third Edition, Audio Engineering Society Presents (AES), Routledge, Taylor & Francis Group.
- Torick, E. (1998). *Highlights in the History of Multichannel Sound*, Journal of the AES, 46, 1/2 January/February, pp. 27-31

Walker, K., Martens, W., L., (2006). *Perceptions of Audio-Generated and Custom Motion Programs in Multimedia Display of Action-Oriented DVD Films, Haptic and Audio Interaction Design*, Springer, Berlin, Germany.

Wenzal, M, E., Begault, R. D., Godfroy-Cooper, M., (2018). "Perception of Spatial Sound," Chapter 1, *Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio*, Audio Engineering Society, Edited by Agnieszka Roginska and Paul Geluso, Routledge, Taylor & Francis, New York and London. (Kindle)

Whiteley, S., Bennett, A., Hawkins, S., (2004). *Music, Space and Place: Popular Music and Cultural Identity*, (editors), Ashgate, Routledge Taylor & Francis Group.

Xie, B., (2013). *Spatial Sound, Principles and Applications*, CRC Press, Taylor & Francis Group. (Kindle)

Zentener, M., Grandjean, D., Scherer, K, R., (2008). *Emotions Evoked by the Sound of Music: Characterization, Classification, and Measurement*, University of Geneva and University of York, American Psychological Association, 1528-3542/08/\$12.00, DOI: 10.1037/1528-3542.8.4.494,

Zotter, F., Frank, M., (2012). *All Around Ambisonic Panning and decoding*, JAES, Soc., 60 (10), 807-820.

## BIBLIOGRAPHY · Episode 2

- Barron, M., Marshall, A.H., (1981). *Spatial Impression Due to Early Lateral Reflections in Concert Halls*, The Derivation of Physical Measure. J. Sound Vib. 77(2),211-232, [https://doi.org/10.1016/S0022-460X\(81\)80020-X](https://doi.org/10.1016/S0022-460X(81)80020-X)
- Bense, M., (1961). *Bewusstseinstheorie, [Theory of Consciousness]*, Grundlagenstud. Aus Kybern. U Geisteweiss. 2, 65-73.
- Blauert, J., (1983, 1997). *Spatial Hearing, The Psychophysics of Human Sound Localization*, The MIT Press, Harvard MA.
- Bradley, J., S, Soulodre, G., A., (1995). *Objective Measures of Listener Envelopment*. J. Acoust Soc Am; 98:250.
- Bregman, A., S., (1993). *Auditory Scene Analysis, Hearing in Complex Environments*’, in Thinking in Sound: The Cognitive Psychology of Human Audition, ed. Stephen Mc Adams and Emmanuel Bigand. Oxford: Clarendon Press, 10-36.
- Brøvig-Hanssen, R., Danielsen, A., (2013). *The Naturalized and the Surreal: Changes in the Perception of Popular Music Sound*, University of Oslo, Department of Musicology, Box 1017, NO-0315, Oslo, Norway.
- Christenson, T., (2017). *History of Music Theory, Oxford Bibliographies*, <https://www.oxfordbibliographies.com/view/document/obo-9780199757824/obo-9780199757824-0199.xml> (accessed 9 Mar. 2024).
- Clarke, E., (2005). *Ways of listening: An Ecological Approach to the Perception of Musical Meaning*, Oxford University Press.
- Damaske, P., (1967/68). *Subjective Investigations of Sound Fields* (in German), Acoustica; 19:199
- Dockwray, R., Moore, A., F., (2010). *Configuring the Sound-Box 1965-1972*, Popular Music, Vol. 29/2, Cambridge University Press, pp.181-197, doi:10.1017/S0261143010000024
- Dockwray, R., Moore, A.F., and Schmidt, P., (2009). *A Hermeneutics of Spatialization for Recorded Song*, Twentieth-Century Music, vol 6, no 01, pp. 83-114
- EBU REC. 562-3, (1990). *Subjective Assessment of Sound Quality*, European Broadcast Union, Geneva, Switzerland.
- Faconnier, G., (1997). *Mappings in Thought and Language*, Cambridge University Press, Cambridge.
- Fernandez-Baizan et al, (2020). *Egocentric and Allocentric Spatial Memory in Typically Developed Children*, Brain and Behaviour, Wiley.
- Frith, S., (2007). *Taking Popular Music Seriously, Chapter 18, Music and Identity*, Routledge, 1st Edition, eBook ISBN, 9781315087467, Taylor and Francis eBooks, [Accessed July 1st, 2024]

- Gamble, S., (2017). *Listening to Virtual Space in Recorded Music*, Art of Record Production 2017, Proceedings, J-O Gullö (ed.), Stockholm, Sweden.
- Goldberg, M. E., (2001). *Parietal Lobe*, International Encyclopedia of the Social and Behavioural Sciences, ISBN:0-08-043076-7, Elsevier Science Ltd.
- Hall, T. E., (1963). *A System for Notation of Proxemic Behaviour*, American Anthropologist, 65/5, 1003-26.
- Hofman, P., & Van Opstal, A. (2003). *Binaural Weighting of Pinna Cues in Human Sound Localization*, Experimental Brain Research, 148(4), 458-470.
- Holman, T., (2008). *Surround Sound: Up and Running*, Second edition, Boston, MA: Focal Press, Elsevier.
- IEC 60268, (1997). [Draft]. *Sound System Equipment—Part 13: Listening Tests on Loudspeakers*, International Electrotechnical Commission, Geneva, Switzerland.
- ITU-R BS.116, (1994). *Methods for Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems*, International Telecommunications Union, Geneva, Switzerland.
- Johnson, M., (1987). *The Body in the Mind*, Chicago University Press.
- Johnson, M., (1989). *Image-Schematic Bases of Meaning*, Association Canadienne de sémiotique/Canadian Semiotic Association, Vol. 9.
- Kuttruff, H., (2016). *Room Acoustics*, Sixth Edition, Institute of Technical Acoustics, Aachen University, Germany, Routledge, CRC Press, Taylor & Francis group, Boca Raton, London, New York.
- Lakoff, G., (1987). *Women, Fire, and Dangerous Things*, Chicago: Chicago University Press.
- Lakoff, G., Johnson, M., (1980). *Metaphors We Live By*, Chicago University Press.
- Leviton, D., (2006). *This is Your Brain on Music*, Plume, Penguin Group.
- Linberg, M., (2022). *Channel Identification audio for Dolby Atmos 7.1.4, 2L*.
- Lungwitz, H., (1923). *DieEntdeckung der Seel-Allg. Psychobiol*, [The Discovery of the Soul—General Psychobiology]. Brücke Verlag Kurt Schmiersow, Kirchhain N.L.
- Lungwitz, H., (1933). *Lokalisation der akustischen Gegenstände*, [The Localization of Acoustical Objects], In: Lehrbuch der Psychobiologie, vol. 2, Walter de Gruyter, Berlin.
- Meaux, E., Marchand, S., (2021). *Synthetic Transaural Audio Rendering (STAR): A Perceptive 3D Audio Spatialization Method*, J. Audio Eng. Soc., vol. 69, no.7/8, pp. 497-505, (July/August), DOI:<https://doi.org/10.17743/jaes.2021.0015>)

- Middlebrooks, J. C., & Green, D. M. (1991). *Sound Localization by Human Listeners*, Annual review of Psychology, 42 (1), 135-159.
- Moore, A.F., Schmidt, P., Dockwray, R., (2011). *A Hermeneutics of Spatialization for Recorded Song*, Twentieth-Century Music, 6/1, 83-114, Cambridge University Press, doi:10.1017/S1478572210000071.
- Moore, A., (1992). *Rock: The Primary Text, Developing a Musicology of Rock*, Open University Press, Buckingham England.
- Moylan, W., (2020). *Recording Analysis: How the Record Shapes the Song*, Chapter 8, The Illusion of Space as an Element of Recording, Focal Press, Routledge Academic, Taylor & Francis Group.
- Neher, T., (2001). *Stimulus Manipulation for 3D Audio Evaluation*, Int. Rep., Institute of Sound Recording, University of Surrey, UK.
- Neuhoff, J., (editor) (2004). *Ecological Psychoacoustics*, Sandiego: Elsevier
- Patterson, J., Lee, H., (2021). *3D Audio, Perspectives on Music Production*, Focal Press, Routledge Academic, Taylor & Francis Group.
- Pawlak, A., Lee, L., Mäkiparta, Lund T., (2023). *Spatial Analysis and Synthesis Methods: Subjective and Objective Evaluations Using Various Microphone Arrays in the Auralization of a Critical Listening Room*, IEEE/ACM Transactions on Audio, Speech, and Language Processing, in peer review.
- Pfanzagle-Cardone, E., (2023). *The Art and Science of 3D Audio Recording*, Springer.
- Pfanzagl-Cardone E., Höldrich R., (2008). *Frequency-dependent signal-correlation in surround- and stereo microphone systems and the Blumlein-Pfanzagl-Triple (BPT)*. Paper 7476 presented at the 124<sup>th</sup> Audio Engineering Society Convention, Amsterdam.
- Rice, T., (2018). *Acoustemology*, University of Exciter, Researchgate, DOI: 10.1002/9781118924396
- Roginska, A., Geluso, P., (2018). *Immersive Sound, The Art and Science of Binaural and Multi-channel Audio*, Audio Engineering Society, Edited by Agnieszka Roginska and Paul Geluso, Routledge, Taylor & Francis, New York and London. (Kindle)
- Rumsey, F., (2002). *Spatial Quality Evaluation for Reproduced Sound: Terminology, Meaning, and a Scene-Based Paradigm*, Institute of Sound Recording, University of Surrey, UK, and School of Music in Piteå, Luleå University of Technology, Sweden, JAES, Vol 50, No 9, 2002, September.
- Schaeffer, P. (1966). *Traité des objets musicaux: Essai interdisciplines*, Éditions du Seuil, Paris, ISBN 978-2-02-133541-5.
- Seetharaman, P., Pardo, B., (2014). *Crowdsourcing a Reverberation Descriptor Map*, Northwestern University, EECS Department, MM'14, ACM 978-1-4503-3063-3/14/11, November 3-7, 2014, Orlando, Florida, USA.

- Serinus, J., V., (2022). *NWAA Labs: Measurement Beyond The Atomic Level*, Reference, Stereophile, (Website), Available online from <https://www.stereophile.com/content/nwaa-labs-measurement-beyond-atomic-level> [accessed 18, Jan, 2025]
- Smalley, D., (1997). *Spectromorphology: Explaining Sound-Shapes*, Organized Sound 12 (1): 35-58.
- Smalley, D., (2007). *Space-Form and the Acousmatic Image*, Organised Sound, 2(2): 107-26.
- Stagle, D., (2022). *How To Listen to Dolby Atmos Without Costly Hardware*, (Website), Available online from <https://www.pro-tools-expert.com/production-expert-1/how-to-listen-to-dolby-atmos-without-costly-hardware?fbclid=IwAR0Qr32l9UPcmJpXQxf-KJr7cEyDcCRFrXrGQX-2DP1FtCulZRjegQosBxA> [accessed 25, Aug. 2022].
- Toole, F., E., (1985). *Subjective Measurements of Loud-speaker Sound Quality and Listener Performance*, J. Audio Eng. Soc., vol 33, pp.2-32 (Jan-Feb)
- Wenzal, M, E., Begault, R. D., Godfroy-Cooper, M., (2018). "Perception of Spatial Sound," Chapter 1, *Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio*, Audio Engineering Society, Edited by Agnieszka Roginska and Paul Geluso, Routledge, Taylor & Francis, New York and London. (Kindle)
- Xie, B., (2013). *Spatial Sound, Principles and Applications*, CRC Press, Taylor & Francis Group. (Kindle)
- Yost, W. A., & Dye, R. H. (1997). *Fundamentals of Directional Hearing*, Seminars in Hearing, 18(4), 321-344. New York: Thieme Medical Publishers.
- Zagorski-Thomas, S., (2014). *The Musicology of Record Production*, Cambridge University Press, UK.



### BIBLIOGRAPHY · Episode 3

- AES, (2022). *AES International Conference on Automotive Audio*, AES.org, (Website), available online from, <https://aes2.org/events-calendar/aes-2022-international-automotive-audio-conference/>, [Accessed 15 Jun. 2022].
- Allen, I., (2006). *The X-Curve: Its Origins and History*, (Website), available online from, [https://www.associationdesmixeurs.fr/wp-content/uploads/2015/10/Dolby\\_The-X-Curve\\_SMPTE-Journal.pdf](https://www.associationdesmixeurs.fr/wp-content/uploads/2015/10/Dolby_The-X-Curve_SMPTE-Journal.pdf) [Accessed 24 Jul. 2022].
- Alton Everest, F., Pohlmann, C., K., (2013). *Handbook of Sound Studio Construction, Rooms for Recording and Listening*, McGraw-Hill, Kindle-iPad edition. (p.20, 27, 111, 56, 143).
- Alton Everest, F., Pohlmann, C., K., (2022). *Master Handbook of Acoustics, Seventh Edition*, McGraw-Hill, Kindle-iPad edition. (p.489).
- Auto-decibel-db, (2023). *The Car Interior Noise Level Comparison site – Auto innengeräusch vergleich*. (Website), available online from, [https://www.auto-decibel-db.com/index\\_kmh.html](https://www.auto-decibel-db.com/index_kmh.html) [Accessed 06 Apr. 2023].
- Barton, P., (2021). *PSB Speakers, Designer, Engineer, Business Owner at Lenbrook Industries*, I had numerous telephone discussions with Paul between September and December 2021, (Website store), available online from, <https://www.psbSpeakers.com/product-category/series/alpha/>.
- Brown, P., (2021). *Divide and Conquer – The Schroeder Frequency*, (Website), available online from, <https://www.prosoundtraining.com/2021/10/14/divide-and-conquer-the-schroeder-frequency/> [Accessed 24 Jul. 2022].
- Carrizosa, P., (2021). *Could Tiny Homes be the Adorable, Affordable and Sustainable Housing that our Planet Needs?* (Website), available online from, <https://ideas.ted.com/impact-of-tiny-homes-on-the-environment-and-affordable-housing/>.
- Dolby, (2019). *Dolby Atmos Home Entertainment Studio Certification, Guide*, 08.07.19, professional.dolby.com, (link is no longer available, Dolby no longer certifies rooms), [Accessed 25 Jun. 2021].
- Dolby, (2020a). *Dolby Atmos Music via Headphones*, Webinar 3: How to create Dolby Atmos Music on Headphones, Ceri Thomas, (Website), available online from, <https://professional.dolby.com/events/dolby-atmos-music-via-headphones/>
- Dolby, (2020b). *Why Dolby Atmos, DLB\_Atms\_WP\_v04\_200807*, (Website), available online from, <https://professional.dolby.com/siteassets/content-creation/dolby-atmos/why-dolby-atmos-white-paper.pdf>, [Accessed 15 June 2022]—[Accessed 28 Jul. 2022].
- Dolby, (2021a). *The DARDT Quick Start Video Series*, (Website), available online from, [https://professionalsupport.dolby.com/s/article/HE-DARDT-Quick-Start-Video-Series?language=en\\_US](https://professionalsupport.dolby.com/s/article/HE-DARDT-Quick-Start-Video-Series?language=en_US), [Accessed 7 Jul. 2021].

- Dolby, (2021b). *Dolby Atmos Home Entertainment Studio, Technical Guidelines*, 06.05.21, professional.dolby.com, (Website), available online from, <https://dolby.my.salesforce.com/sfc/p/#700000009YuG/a/4u0000001FHc/UYA0IZeD632SUXVmEPmUcr.wIuhpHp6Q7bVSl4LrbUQ>, [Accessed 12 Sept. 2021].
- Dolby, (2021c). *Dolby Atmos Post Production Learning*, (Website), available online from <https://learning.dolby.com/hc/en-us/sections/360010367372Dolby-Atmos-Post-and-Music-Production-Learning> [Accessed 12 Jan. 2022].
- Dolby, (2021d). *Dolby Professional Support, What DAWs Support Dolby Atmos*, (website), available online from, [https://professionalsupport.dolby.com/s/article/What-DAWs-support-Dolby-Atmos?language=en\\_US](https://professionalsupport.dolby.com/s/article/What-DAWs-support-Dolby-Atmos?language=en_US) [Accessed 20 Sept.2022]
- Dolby, (2021e). *What room dimensions are needed to mix content in Dolby Atmos for Home Entertainment?*, (Website), available online from, [https://professionalsupport.dolby.com/s/article/What-room-dimensions-are-needed-to-mix-content-in-Dolby-Atmos-for-Home-Entertainment?language=en\\_US](https://professionalsupport.dolby.com/s/article/What-room-dimensions-are-needed-to-mix-content-in-Dolby-Atmos-for-Home-Entertainment?language=en_US) [Accessed 06 Jul. 2022].
- Dolby, (2022a). *Experience the future of in-car entertainment with Dolby at the AES 2022 Conference*, (Website), available online from, <https://professional.dolby.com/music/dolby-atmos-for-cars/#gref>, [Accessed 15 Jun. 2022].
- Dolby, (2022b). Thomas, C., Brooks, E., Kunau, M., *Dolby Atmos onboarding email*, Feb. 5<sup>th</sup>.
- Dolby, (2022c). *Dolby Atmos Music Creation: Studio Setup and Tuning*, Dolby, 10:10, <https://www.youtube.com/watch?v=pPDY6mXBmgs>
- Dolby, Developer, (2022d). *Dolby Atmos is an immersive audio format that can be delivered via multiple audio codecs ...* (Website), available online from, [https://developer.dolby.com/technology/dolby-atmos/overview/#:~:text=Dolby%20Atmos%C2%AE%20is%20an,\(but%20NOT%20Dolby%20Digital\).](https://developer.dolby.com/technology/dolby-atmos/overview/#:~:text=Dolby%20Atmos%C2%AE%20is%20an,(but%20NOT%20Dolby%20Digital).) [Accessed, 25 Aug. 2022].
- Dolby, (2023). *Dolby Atmos for cars: Volvo Cars*, (Website), available online from, <https://professional.dolby.com/music/dolby-atmos-for-cars/volvo/>, [Accessed 06, Jan. 2023].
- Dolby, (2024a). *Dolby Atmos Studio Directory*, (website), available online from <https://professional.dolby.com/music/dolby-atmos-music-studios/#gref>
- Dolby, (2024b). *The New Electric Volvo EX90 SUV to Feature Dolby Atmos*, (Website). Available online from <https://professional.dolby.com/music/dolby-atmos-for-cars/volvo/#gref>, [Accessed 25, Nov., 2024].
- Fazenda, B., Angus, J., Cox, T., (2015). *The SOS Guide to Control Room Design, Making Space*, Mixing / Production, (Website), available online from, <https://www.soundonsound.com/techniques/sos-guide-control-room-design> [Accessed 12 July 2022].
- Fielder, L., (2003). *Dolby Labs, Analysis of Traditional and Reverberation-Reducing Methods of Room Equalization*, JAES-2249, Vol.51, No. ½ 2003, January/February.

- Fox, B., (1982). *Early Stereo Recording (Alan Blumlein)*, Studio Sound and Broadcast Engineering, May 1982.
- Genewick, S., (2021). *White Lightning, Staci Greisbach*, (Website), available online from, <https://mixwiththemasters.com/videos/mixing-in-atmos-2-staci-greisbach-white-lightning/part/1>, [Accessed 04 Jul. 2022].
- Gerzon, M.A., (1972). *Width-Height Sound Reproduction*, (Periphony), Journal of the Audio Engineering Society, JAES, January/February 1973, Volume 21, Number 1.
- Kunst, A. (2019). *Purposes headphones are used for in the United States 2017* (What do you use your headphones for?), STATISTA (website), available online from <https://www.statista.com/statistics/696862/uses-of-headphones-in-the-us/> [accessed 15 May 2022].
- Lucidmotors, (2021). *Lucid Air is the World's First Vehicle to Integrate Dolby Atmos*, (website). Available online from <https://www.lucidmotors.com/media-room/lucid-air-is-first-vehicle-to-integrate-dolby-atmos>, [accessed 27 Aug. 2022].
- Murch, W., (2012). *Walter Murch on the Digital Revolution*, (Website), available online from, <http://www.filmdetail.com/2012/03/16/walter-murch-on-the-digital-revolution/> [accessed 6 Jul. 2022].
- Nair, S., (2022). *Clustering, An expanded version of my Avid Blog Post*, Atmos Post Mixing Professionals, Facebook group, (website), available online from, <http://facebook.com> [accessed 10 Aug. 2022]
- Newell, P., (2017). *Recording Studio Design, Fourth Edition*, (9781138936072), Audio Engineering Society Presents, (AES), Routledge, Taylor Francis Group, Kindle-iPad edition.
- Newell, P., (2018). *Philip Newell Acoustic Design*, (Website), available online from, <http://philipnewell.net/index9.html> [accessed 20 Jul. 2022].
- Novotny, P., (2017). *Creating an Immersive Fold-Out – Look Ahead*, Proceedings of the 12<sup>th</sup> Art of Record Production Conference, Mono: Stereo: Multi — Stockholm 2017, KMH.
- Ondaatje, M., (2002). *The Conversations, Walter Murch and the Art of Editing Film*, Humble Sounds, Vintage Canada Edition, Randomhouse.ca., iBooks version, pg.23.
- Pennington, B., (2022). *Dolby Atmos Music Creation 101: Studio Setup and Tuning*, (website), available online from <https://www.youtube.com/watch?v=pPDY6mXBmgs> The interview starts at 8:35 in the transcript. [accessed 15 Apr. 2022].
- Petrovic, B., Davidovic, Z., (2010). *Acoustical Design of Control Room for Stereo and Multichannel Production and Reproduction*, a Novel Approach, MyRoom principle, AES paper #8295.
- Petrovic, B., Davidovic, Z., (2016). *A Novel Approach of Multichannel and Stereo Control Room Acoustic Treatment*, Second Edition, AES paper #9526.

- Rothermich, E., (2022). *Logic Pro 10.7.3, Monitor Dolby Atmos through Apple's Spatial Audio Engine, Music Tech Explained*, The Visual Approach, (website), available online from, [https://www.youtube.com/results?search\\_query=music+tech+explained](https://www.youtube.com/results?search_query=music+tech+explained) [accessed 28 Jul. 2022].
- Rothermich, E., (2021). *Mixing in Dolby Atmos, #1 How it works*, Kindle e-book, [www.dingdingmusic.com](http://www.dingdingmusic.com).
- Scheps, A., (2019). *Andrew Scheps talks to Awesome People*, YouTube—PureMix (website), available online from <https://www.youtube.com/watch?v=cnX4pVrxqq0> [accessed 12 Apr. 2022].
- Schroeder, R., M., (1975). *Diffuse Sound Reflected by Maximum Length Sequences*, Journal of the Acoustical Society of America, 57, Pg.149-150.
- Schroeder, R., M., Gerlach, R., (1976). *Response to "Comments on 'Diffuse Sound Reflection by Maximum Length Sequences'"*, Journal of the Acoustical Society of America, 60, Pg. 268-268.
- Sijen, P., (2008). *Design Philosophy: Blue Sky Studio Monitors*, AudioDesignLabs, (website) available from, <http://audiodesignlabs.com> [originally accessed 28 Aug. 2021].  
<http://audiodesignlabs.com/wordpress/2008/05/design-philosophy-blue-sky-studio-monitors/#more-16> [accessed 03 Oct. 2025] (posted by Pascal Sijen, May 17, 2008)
- Sonarworks Studio Reference, (2023). *White Paper*, (Website) Available online from [https://www.sonarworks.com/whitepaper#:~:text=Sonarworks%20solution%20implements%20cutting%20edge,Response%20\(FIR\)%20filter%20technology](https://www.sonarworks.com/whitepaper#:~:text=Sonarworks%20solution%20implements%20cutting%20edge,Response%20(FIR)%20filter%20technology) [accessed 12, Nov. 2023]
- Stagle, D., (2022). *How To Listen to Dolby Atmos Without Costly Hardware*, (Website), Available online from <https://www.pro-tools-expert.com/production-expert-1/how-to-listen-to-dolby-atmos-without-costly-hardware?fbclid=IwAR0Qr32l9UPcmJpXQxf-KJr7cEyDcCRFrXrGQX-2DP1FtCuLZRjegQosBxA> [accessed 25 Aug. 2022].
- Tobin, M., (2021). *Mixing for Headphones, Dolby Atmos Creation 101*, (Website), available online from, <https://www.youtube.com/watch?v=qs1tffnAjPE> [accessed 28 Jul. 2022].
- Toole, E. F., (2018). *Sound Reproduction: The Acoustics and Psychoacoustics of Loudspeakers and Rooms, Third Edition*, Audio Engineering Society Presents (AES), Routledge, Taylor & Francis Group, Pg.16 Kindle-iPad version. pgs. xvii, xix, pg.321. (Kindle)
- Walker, R., (1996). *Optimum Dimension Ratios for Small Rooms*, 100<sup>th</sup> convention of the Audio Engineering Society, Preprint 4191 (H-4).
- Walker, R., (1998). *A Controlled-Reflection Listening Room for Multichannel Sound*, R&D Development Department, British Broadcasting Corporation, Proceedings of the Institute of Acoustics, Vol 20, Part 5.
- Zeimer, T., Bader, R., (2017). *Psychoacoustic Sound Field Synthesis for Musical Instrument Radiation Characteristics*, JAES, Vol. 65, No. 6, pp.482-496.

## WEBSITES VISITED · Episode 1

- Ascendo, <https://www.aia-cinema.com/products/subwoofers.html>
- Ascendo, *Subwoofer Cinematic Installation*, <https://www.youtube.com/watch?v=ELTlx2pE8oI>
- Applied Psychoacoustics Lab, Huddersfield UK, *Projects*, <https://apl-hud.com/projects/>
- Applied Psychoacoustics Lab, Huddersfield UK, *Publications*, <https://apl-hud.com/publications/>
- Applied Psychoacoustics Lab, Huddersfield UK, *Resources*, <https://apl-hud.com/resources/>
- Behrman, S. (2022). *The Political Economy of Classical Music*, <https://jacobin.com/2022/03/political-economy-classical-european-music-history-revolution-elitism>
- Brie, O., (2023). *Public domain version*, <https://www.gutenberg.org/ebooks/71333>
- Dolby-A, (2024). <https://professional.dolby.com/music/dolby-atmos-for-cars/volvo/#gref>
- Dolby-B, (2024). <https://www.youtube.com/watch?v=DMjwsdzdxIQ>
- Dolby (2020). <https://www.youtube.com/watch?v=ggzfcUKDqdo>
- Fairchild, J., (2020-204). *Nature Walks with Judy*, Blog, Charleston, South Carolina, BBC Earth, YouTube.  
<https://naturewalkswithjudy.com/2021/05/24/alligators-communicate-with-infrasonic-bellows/#:~:text=Alligators%20communicate%20with%20Infrasonic%20Bellows,%2D%20Nature%20Walks%20with%20Judy>
- Huberman, A., *Enhance Your Learning Speed & Health Using Neuroscience Based Protocols*, | Dr. Poppy Crum, September 29, 2025. [https://www.hubermanlab.com/episode/enhance-learning-speed-neuroscience-poppy-crum?utm\\_source=linkedin&utm\\_medium=social](https://www.hubermanlab.com/episode/enhance-learning-speed-neuroscience-poppy-crum?utm_source=linkedin&utm_medium=social)
- Lee, H., *Echo-project* (2024). [https://apl-hud.com/echo?fbclid=IwY2xjawHNLQNIeHRuA2FlbQIXMAABHTM8akTDESIOM3Btx\\_F2MK5uXWMdiAM1JMW3Dwt\\_gc7UqZZZYbP4WQzXPg\\_aem\\_3bF4r3Pf-ZIAJt0LUYawNw](https://apl-hud.com/echo?fbclid=IwY2xjawHNLQNIeHRuA2FlbQIXMAABHTM8akTDESIOM3Btx_F2MK5uXWMdiAM1JMW3Dwt_gc7UqZZZYbP4WQzXPg_aem_3bF4r3Pf-ZIAJt0LUYawNw)
- Genewick, S., (2024a). <https://stevegenewick.com/dolby-atmos>
- Genewick, S., (2024b). *Apple Music playlist*, <https://music.apple.com/ca/playlist/steve-genewick-atmos-mixes/pl.u-76oNkrMF1JKbp>
- Ministry of Sound (1991). <https://www.ministryofsound.com/>
- Oxford English Dictionary (2024). York University Library, [visited on 09,24,24] <https://www.oed.com/search/dictionary/?scope=Entries&q=context>

Saunders, S., (2023). *AI and the Adjacent Possible*, Art & Science, <https://www.artscience.ca/insights/ai-and-the-adjacent-possible/>

Sen, D., (2025). *Meet Apple Spatial Audio Format and APAC | Meet with Apple*, Apple Developer, [https://www.youtube.com/watch?v=DpvR\\_74V4T4](https://www.youtube.com/watch?v=DpvR_74V4T4), [Visited on 12,05,25]

Zentner, M., (2021). *9 Great Composers Explained in 9 Emotions*, TEDx Talks, 9 Dimensions of the GEMS Model, <https://www.youtube.com/watch?v=WPOsrRQfkk>

## **WEBSITES VISITED · Episode 2**

<https://music.apple.com/ca/album/voice-navigation-7-1-4-immersive-audio/1693419553?i=1693419891>

[https://shop.2l.no/products/immersive-channel-identification-7-1-4-test-files?variant=41851094696119&\\_pos=1&\\_psq=atmos+adm&\\_ss=e&\\_v=1.0](https://shop.2l.no/products/immersive-channel-identification-7-1-4-test-files?variant=41851094696119&_pos=1&_psq=atmos+adm&_ss=e&_v=1.0)

<https://www.pro-tools-expert.com/production-expert-1/how-to-listen-to-dolby-atmos-without-costly-hardware?fbclid=IwAR0Qr32l9UPcmJpXQxf-KJr7cEyDcCRFrXrGQX-2DP1FtCuIZRjegQosBxA>

<https://irepod.com/podcast/pro-tools-expert-podcast/the-science-of-listening-with-professor-hyunkook-l>

<https://rogueamoeba.com/loopback/>

<https://audiomovers.com/omnibus/>

## **WEBSITES VISITED · Episode 3**

<https://www.pbspeakers.com/product-category/series/alpha/>

[https://www.auto-decibel-db.com/index\\_kmh.html](https://www.auto-decibel-db.com/index_kmh.html)

<https://ideas.ted.com/impact-of-tiny-homes-on-the-environment-and-affordable-housing/>

<https://professional.dolby.com/music/dolby-atmos-for-cars/#gref>

<https://mixwiththemasters.com/videos/mixing-in-atmos-2-staci-griesbach-white-lightning/part/1>

<https://www.statista.com/statistics/696862/uses-of-headphones-in-the-us/>

<https://www.pro-tools-expert.com/production-expert-1/how-to-listen-to-dolby-atmos-without-costlyhardware?fbclid=IwAR0Qr32l9UPcmJpXQxf-KJr7cEyDcCRFrXrGQX-2DP1FtCuIZRjegQosBxA>



## WEBSITES VISITED · Conclusion

Wendle, M., (2024). *Studio Manager* || *Immersive Studio Manager* || APL || Dolby Atmos Mix Engineer || Senior Technician, University of Huddersfield, UK, Linkedin.

[https://www.linkedin.com/posts/markwendl\\_lets-keep-talking-hans-martin-buff-activity-7257285107206574081-wf26?utm\\_source=share&utm\\_medium=member\\_desktop](https://www.linkedin.com/posts/markwendl_lets-keep-talking-hans-martin-buff-activity-7257285107206574081-wf26?utm_source=share&utm_medium=member_desktop)

NASA: *Images and Media Usage Guidelines*, <https://www.nasa.gov/nasa-brand-center/images-and-media/>

## DISCOGRAPHY

Gabriel, P, (2024). *i/o (In-side Mix)*, Apple Music.

Gray, Justin, (2025). *Immersed*, Apple Music, Pure Audio, Blu Ray Disc, Auro 3D, CD, IAN records.

Hunt, S., (2023). “Outskirts” Apple Music.

John, Elton (2017). “Rocket Man”, [Dolby Atmos version], Apple Music.

Lewis, H., (2023). “Heart of Rockn’ Roll,” Apple Music.

Moose, W., S., (2023). “Jaat Du Maqabala,” Apple Music.

Novotny, P, (2015). “Now’s The Time,” *Look Ahead* featuring Robi Botos, TR10019, Triplet Records, Toronto, Canada, Apple Music.

Novotny, P, (2025). Spatially Compelling 3D Music—Paul Novotny-Apple Music Playlist, <https://music.apple.com/ca/playlist/spatially-compelling-3d-music-paul-novotny/pl.u-jV890pDFdjPzgp>

Novotny, Paul (2022). *Summertime in Leith, In Concert at the Historic Leith Church*, featuring Robi Botos, TR10026, Triplet Records, Toronto, Canada, Apple Music.

Pink Floyd, (2023). “On The Run,” *The Dark Side of The Moon*, Apple Music.

Richter, Max, (2024). *In a Landscape*, Apple Music.

Romboy, M., (2022) “Randomized,” *Speicher 120*, Apple Music.

Rush, (2022). “Tom Sawyer,” *Moving Pictures, 40th Anniversary*, Ume/Mercury and Anthem Records, 04/15/2022, Apple Music.

Wilson, Steven, (2025). *The Overview*, Apple Music.

## GLOSSARY

### Acoustics, Conceptual, Physics:

|                               |                                                                                                                                                                                                                           |
|-------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Anechoic Chamber              | A room that is free from echo and used for measurement and design of loudspeakers and microphones. Direct sound of an audio-object is all that is heard in an Anechoic Chamber.                                           |
| Apparent Ensemble Width (AEW) | Direct and indirect sound width of an ensemble. Can be effectively deployed in the Lw, and Rw speakers. An important component for wide spatial resolution in binaural listening.                                         |
| Apparent Source Width (ASW)   | Direct and indirect sound width of an audio-object source. Usually deployed within L, C, R speakers. associated with impact, intelligibility, and intensity.                                                              |
| ASA                           | Audio scene analysis.                                                                                                                                                                                                     |
| Correlated                    | A mutual relationship between components, essential for Dolby Atmos studio calibration.                                                                                                                                   |
| Crosstalk                     | Essential transference of signal between left and right headphone channels to re-create the aural decorrelations of listening in a room. Electronic crosstalk is undesirable when designing multichannel audio equipment. |
| dB                            | Decibel: A unit used to measure the intensity of a sound or the power level of an electrical signal by comparing it with a logarithmic scale.                                                                             |
| Decorrelated                  | A separate relationship between components, essential for a compelling spatial experience. Decorrelation makes the sound appear outside of the head.                                                                      |
| Direct sound                  | The “denatured” (Schaeffer, 1966) original audio object with little to no indirect sound or environment.                                                                                                                  |



|                            |                                                                                                                                                                                                                                                                                |
|----------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Direct sound source depth  | Distance perspectives described as proximal or distal.                                                                                                                                                                                                                         |
| Distal                     | A far depth of field.                                                                                                                                                                                                                                                          |
| Ensemble Depth (ED)        | The perceived distance from front to back of the ensemble.                                                                                                                                                                                                                     |
| Ensemble Width (EW)        | Direct sound ensemble sound width.                                                                                                                                                                                                                                             |
| HRTF                       | HRTF is a term that explains interaural induction of sound using the ears. It involves interaural timing differences (ITD), interaural level differences, (ILD), and interaural cross-correlation (IACC), which are frequency spectrum based.                                  |
| IACC                       | Interaural cross-correlation.                                                                                                                                                                                                                                                  |
| ILD                        | Interaural level difference.                                                                                                                                                                                                                                                   |
| Indirect sound             | The reflected environmental nature that echoes. direct sound. It can be created virtually or recorded natively.                                                                                                                                                                |
| Infrasound                 | Sound wave frequencies below the lower limit of human audibility.                                                                                                                                                                                                              |
| ITD                        | Interaural timing difference.                                                                                                                                                                                                                                                  |
| Listener envelopment (LEV) | The sensation of spatial sound envelopment. Created using the lateral side speakers and involves delay times of between .05ms to 80ms. (Damaske, 1967/68; Baron, 1981, 1981, Kuttruff, 2016). LEV or “sense of spatiality” was investigated thoroughly by Barron, M., in 1971. |
| MIR                        | Music information retrieval associated with textual codification of music properties and qualities. MIR involves genre classification, recommendation, music source separation, instrument recognition, automatic music transcription, and music generation.                   |

|                       |                                                                                                                                                                                                                     |
|-----------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Object bed (O-bed)    | An artificial speaker sphere that can bigger than the physical speaker sphere. Mixing into an O-bed creates higher spatial resolution in binaural playback.                                                         |
| Periphonic, periphony | “Sound reproduction in both vertical and horizontal directions around a listener.” (Gerzon, M, A, 1972, p.1) This is a correlated expression of equal sound from all speakers.                                      |
| Proximal              | A close depth of field and the opposite of distal.                                                                                                                                                                  |
| Radiation             | The propagation of sound in a particular direction through a medium, originating from an audio-object, loudspeaker, or loudspeaker array.                                                                           |
| Radiosity             | The amount of “source-bonded” (Smalley, 1997,2007) indirect sound.                                                                                                                                                  |
| RT <sub>60</sub>      | A measurement of the duration needed for indirect reverberant sound to decay or reduce by 60 decibels (dB) after the direct sound source stops.                                                                     |
| SASC                  | Spatial audio scene characterization. Based on Rumsey’s scene-based paradigm spatial content of music could be characterized at three abstraction levels: low, mid, and high. (Antonouk, Zieliński, Lee, 2024, p.2) |
| Source width (SW)     | Direct sound width of an audio-object source.                                                                                                                                                                       |
| X, Y, Z               | Cartesian planar coordinate system that locates a point from an origin (0,0,0) where three orthogonal axes cross to indicate the coordinate positions of width, height, and depth.                                  |
| <b>Anatomy:</b>       |                                                                                                                                                                                                                     |
| Anterior              | Situated in the front of the body.                                                                                                                                                                                  |
| Azimuth (horizontal)  | The horizontal angle of tilt.                                                                                                                                                                                       |

|                   |                                       |
|-------------------|---------------------------------------|
| Horizontal        | Parallel to the plane of the horizon. |
| Inferior (Nadir)  | Situated below.                       |
| Lateral           | Situated on the sides of the body.    |
| Posterior         | Situated behind the body.             |
| Superior (Zenith) | Situated above.                       |

### Cognitive Science:

|                |                                                                                                                                                                                                                                     |
|----------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Allocentric    | Audio-object to audio-object (musician to musician) staging relationships. Allocentrism is a psychological dimension which corresponds to cultural collectivism.                                                                    |
| Egocentric     | Describes subjective listener to audio-object head-centered relationships. Egocentrism refers to self-centered relationships. The egocentric relationship can invert to the musician stage perspective looking out at the audience. |
| Image Schemata | “Functions as the basis for rational thought by way of metaphorical extension, serving as the pre conceptual ground for cross-domain mapping (Moore, Schmidt, Dockwray, 2011, p.85).                                                |
| Clairaudience  | The supposed faculty of perceiving, as if by hearing, what is inaudible.                                                                                                                                                            |

### Proxemic Zones:

|          |                        |
|----------|------------------------|
| Intimate | Touching distance.     |
| Personal | Arm’s length distance. |
| Social   | Medium distance.       |
| Public   | Sizeable distance.     |

### Musicological:

|               |                                                                                                                              |
|---------------|------------------------------------------------------------------------------------------------------------------------------|
| Acoustemology | “A portmanteau that combines acoustic and epistemology to foreground sonic experience as a way of knowing” (Rice, 2018, p1). |
|---------------|------------------------------------------------------------------------------------------------------------------------------|

|                 |                                                                                                                                                                               |
|-----------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Instrumentator  | A term used by Gardener Read to describe the occupation of a music orchestrator.                                                                                              |
| ITPRA           | A theoretical response system designed by David Huron to measure experience and emotion in five stages: 1) Imagination, 2) Tension, 3) Prediction, 4) Reaction, 5) Appraisal. |
| Sound-box       | A term coined by musicologist Allan Moore that describes a closed field listening space.                                                                                      |
| Techoustemology | “A complimentary term to acoustemology rooted in the social construction of technology and technoculture” (Porcello, 2004, p.270; Rice 2018, p.5).                            |

#### Speaker Sphere: (9.1.6)

|                 |                                                                                                                                                                                                                                                                                                                                                                                                           |
|-----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Ambisonic audio | A 3D full-sphere technology that uses a microphone with a four-capsule tetrahedral array. W represents total sound energy, X represents back to front sound, Y represents left and right sound, and Z represents sound from above and below. These relationships, perspectives and properties are captured as B-format information which can be decoded for binaural headphones or a multi-speaker gamut. |
|-----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

#### (Multichannel)

|                         |                 |    |
|-------------------------|-----------------|----|
| Left                    | L               | 1  |
| Center                  | C               | 2  |
| Right                   | R               | 3  |
| Low Frequency Effects   | LFE (subwoofer) | 4  |
| Left wide (9.1.6 only)  | Lw              | 5  |
| Right wide (9.1.6 only) | Rw              | 6  |
| Left side surround      | Lss             | 7  |
| Right side surround     | Rss             | 8  |
| Left rear surround      | Lrs             | 9  |
| Right rear surround     | Rrs             | 10 |
| Left top front          | Ltf             | 11 |
| Right top front         | Rtf             | 12 |

|                               |     |    |
|-------------------------------|-----|----|
| Left top middle (9.1.6 only)  | Ltm | 13 |
| Right top middle (9.1.6 only) | Rtm | 14 |
| Left top rear                 | Ltr | 15 |
| Right top rear                | Rtr | 16 |

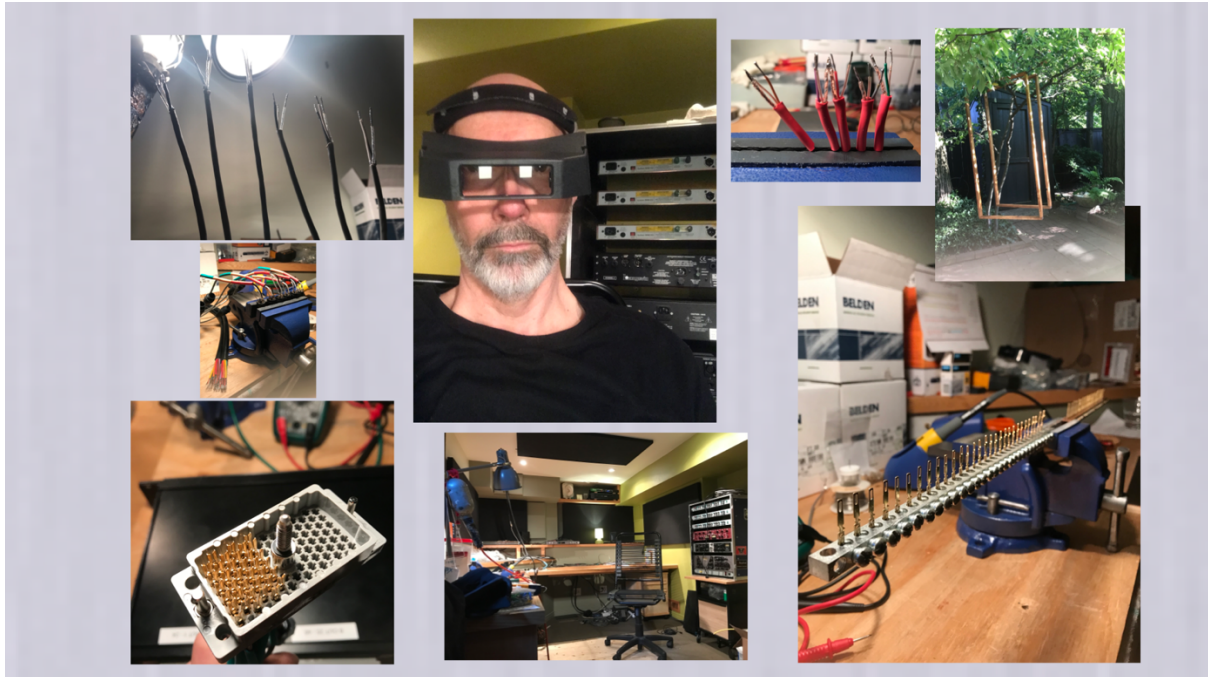
### Technological:

|         |                                                                                                                                                                                                                                                                                                                                                                           |
|---------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Reverb: | A lingering auditory sustain resulting from direct-sound that is reflected off of surfaces in closed- or open-field spaces, and places. Reverb can be virtually simulated or created during music production by hardware, software, or re-micing of audio-objects. 3D re-micing is an effective technique for capturing the genuine characteristics of spaces and places. |
|---------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

### Theatrical and Stage:

|                    |                                                                                                                                                                                                                                           |
|--------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Downstage          | The closest stage proximity to the subject-listener.                                                                                                                                                                                      |
| Center stage       | The horizontal middle of the stage.                                                                                                                                                                                                       |
| Medial center line | The geometric center line of a stage which in Atmos can be used to anchor the stereo image using the center speaker. Medial center line proximity and distance add highly effective dimensions of depth to the anterior horizontal image. |
| Onstage            | An audio-object placed on the stage of a concert, club, or theatre.                                                                                                                                                                       |
| Offstage           | An audio-object placed off the stage of a concert, club, or theatre environment.                                                                                                                                                          |
| Stage right/left   | Expresses the subject horizontal stage location of an audio-object in an audio-object to subjective listener relationship. It can invert contralaterally to be expressed from the audience location                                       |

## IMAGE GALLERY



**Figure 32:** Building the HN-FTS, wiring EDAC connectors.



**Figure 33:** Building the HN-FTS, patch bay with connectors.

## MEDIA LINKS

Novotny, Paul (2023). *Understanding Dolby Atmos with Producer Paul Novotny*, Billboard, Canada, <https://ca.billboard.com/fyi/interview-producer-paul-novotny-helps-understand-dolby-atmos>

Novotny, Paul (2023). *Paul Novotny | The Tiny Studio Concept*, Headliner Radio, USA. <https://www.everand.com/podcast/733149255/Headliner-USA-E147-Paul-Novotny-The-Tiny-Studio-Concept>

Novotny, Paul (2023). *Paul Novotny | The Tiny Studio Concept*, Headliner USA E147, Podcast, [https://www.buzzsprout.com/450409/13654318-headliner-usa-e147-paul-novotny-the-tiny-studio-concept?client\\_source=small\\_player&iFrame=true&referrer=https://www.buzzsprout.com/450409/13654318-headliner-usa-e147-paul-novotny-the-tiny-studio-concept.js?container\\_id=buzzsprout-player-13654318&player=small](https://www.buzzsprout.com/450409/13654318-headliner-usa-e147-paul-novotny-the-tiny-studio-concept?client_source=small_player&iFrame=true&referrer=https://www.buzzsprout.com/450409/13654318-headliner-usa-e147-paul-novotny-the-tiny-studio-concept.js?container_id=buzzsprout-player-13654318&player=small)

Novotny, Paul, (2023). *The Many Lives of Paul Novotny*, Pro Audio Spotlight, <https://www.proaudiospotlight.com/the-many-lives-of-paul-novotny-2/>

Novotny, Paul (2023). *Paul Novotny, Tiny Studio*, Toronto, CA, Merging Technologies, [https://www.merging.com/news/use\\_cases/paul-novotny--tiny-studio-toronto--ca](https://www.merging.com/news/use_cases/paul-novotny--tiny-studio-toronto--ca)

Novotny, Paul, (2023). *The Audio Producers' Owner Paul Novotny on Bringing Immersive Audio to Tiny Studios*, Headliner UK, <https://headlinerhub.com/audio-producers-paul-novotny-immersive-audio-tiny-studios.html>

Novotny, Paul, (2023). *Building a Hyper Near-Field Tiny Studio Using Dolby Atmos*, Webinar, <https://www.youtube.com/watch?v=22aMYubZNfI>

Novotny, Paul (2022-current). *Apple Spatial Music: Compelling 3D Music*, Apple Music Playlist, <https://music.apple.com/ca/playlist/spatial-compelling-3-d-music-paul-novotny/pl.u-jV890pDFdjPzgp>

## FOUND IN SPACE

Journey to the Third Dimension of  
Music Spatiality in the Dolby Atmos Format



## THE END

© PAUL ANTHONY NOVOTNY 2025  
All Rights Reserved