An Investigation of Mastering Techniques for Immersive Audio

Aoife Dowd Bedidi

N00181074

Dún Laoghaire Institute of Art, Design and Technology

Creative Music Production/ DL838

Professional Project

Supervisor: Dr Ruth Moran and Dr Brian Carty

28th April 2023

Abstract

Mastering is the final stage of the audio production process, which involves balancing sonic elements and optimizing the track for playback on various systems and formats. This research aims to identify and evaluate new strategies for mastering audio within a rapidly and continuously evolving frontier. Immersive audio is a three-dimensional sound field created by combining lateral and overhead speakers. Dolby Atmos and Sony 360 RA are two formats that have played a pivotal role in making spatial audio available to consumers through the playback of binaural audio on headphones, creating a full 3D experience supported by online music streaming platforms. The concept of mastering for immersive audio differs from a traditional stereo workflow, and so requires new ways of working. This research aims to investigate the workflow for mastering content intended for this medium through interviews with industry professionals.

Table of Contents

Introduction

The final stage of the audio production process is known as "mastering." Once a song has been recorded and mixed, the mastering engineer's role is to balance the sonic elements and optimise the track for playback on various systems and formats. Mastering is crucial in ensuring that the final product sounds consistent across different playback systems and meets the technical criteria of the platform or medium for which it is intended to be distributed. This research constitutes a relatively new area that has emerged from the growing commercialisation of immersive audio and aims to investigate how an immersive audio track is mastered and prepared for distribution.

With immersive audio, music can be heard beyond the stereo left and right channels, offering full range to the music's height, depth, and width and delivering a richer, bolder reflection of the artist's true intentions.[1] Immersive audio formats are classified as being channel-based[2], scene-based[3] or object-based[4]. By definition, all immersive audio systems are multi-channel and must include speakers from above, i.e., height channels.

Since its existence, the mastering process has been purely channel-based, but the recent advent of object-based technologies demands new ways of working. The term "channel-based" implies that the master track was designed specifically for playback on discrete channels to specific speakers.[5] With object-based audio, rendering software can analyse the metadata[6] in the master file and route the sounds to the closest possible position to the original coordinates set in the mixing stage. In this way, the renderer is flexible enough to offer an adaptive listening experience across a wide range of playback speaker configurations.[7] The immersive format selected determines the exact location of these speakers. Several file formats have been created that can handle the task of object-based

[1] "What Is Dolby Atmos Music Mastering?" *Abbey Road*, Nov. 2022, www.abbeyroad.com/news/what-is-dolby-atmos-music-mastering-3273. Accessed 25 Apr. 2023.
[2] Channel-based immersive refers to surround formats beyond 5.1 that incorporate overhead speakers (height)
[3] Scene-based immersive formats present a single, complex data stream that describes a complete three-dimensional sound field
[4] Object-based immersive treats each sound source as an independent object with its own metadata
[5] Melchior, Frank, et al. "Spatial Mastering - a New Concept for Spatial Sound Design in Object-based Audio Scenes." *ResearchGate*, Jan. 2011, p. 1, www.researchgate.net/publication/267840563_Spatial_Mastering_-_a_new_concept_for_spatial_sound_design_in_object-based_audio_scenes. Accessed 25 Apr. 2023.
[6] Metadata is information about the object, such as location, volume and direction
[7] Herman, Gabe. "What Is Immersive Audio and How Does It Work?" *Sonarworks Blog*, 23 May 2022, www.scribbr.com/citation/generator/folders/1FhrFtAY7HPkBy4aEN3SP0/lists/5LrFnonaUexMQCLcIawvFJ/cite/blog-post. Accessed 25 Apr. 2023.

audio: Dolby ATMOS[8], Sony 360 Reality Audio (RA)[9], and Auro 3D[10] are three examples of commercial audio production and distribution formats that are competing within the market.[11]

Compared to channel-based mastering, the methodology of mastering audio objects is far less defined and understood. In the mixing stage of a stereo track, sounds are placed between two speakers, i.e., the left and right channels, and a panner is used to position the sound, i.e., to the left, right, or somewhere in between. In comparison, with object-based audio, the mix engineer has control over the lateral position (x), the depth position (y), and the vertical position (z) to assign the location of the sound. After the position and relative levels of the sounds are set, a master file is then created, capturing the audio and embedding its positioning information as metadata.

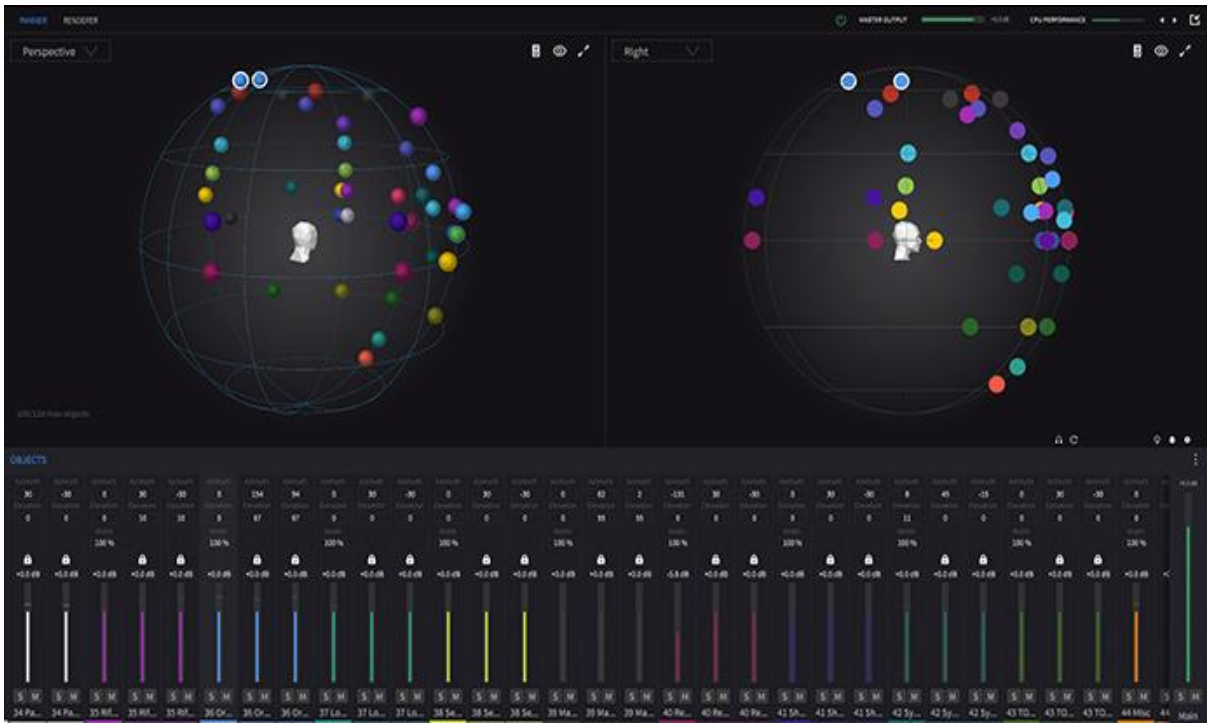

Fig.1 360 Reality Audio *"360 WalkMix Creator™ enables musicians and creators to produce music easily and creatively in an immersive spherical sound field using the 360 Reality Audio Music Format."* 2023

[8] Dolby Atmos is an object-based audio format
[9] Sony 360 Reality Audio is built on the MPEG-3D audio codec, and is a format that can contain objects, channel-based audio and ambisonic data.
[10] Auro Technologies' Auro 3D is a channel-based format
[11] Inglis, Sam. "An Introduction to Immersive Audio." Sound on Sound, Jan. 2022, www.soundonsound.com/techniques/introduction-immersive-audio. Accessed 27 Apr. 2023.

Ensuring the successful translation of creative intent and technical features across all modes of listening is a critical task for mastering engineers. At a time when the way consumers experience music is ever-evolving, this proves to be an ever-present challenge. Maintaining seamless translation across systems is often influenced by the technical criteria of distribution channels. Streaming services such as Spotify require specific values related to the integrated loudness (LUFS)[12] of the master. Previously, there was an ongoing competitive increase in the loudness of commercially distributed music, referred to as the "loudness war"[13]. An underlying reason for this competitive increase in loudness was the illusion that louder recordings would be more effective at capturing the listener's attention and, in turn, sound better. This resulted in "squashed,"[14] hyper-compressed mixes with a much smaller use of dynamics. Concerns about loudness fatigue became more prevalent, and as streaming sites became increasingly popular, individual platforms began to utilize normalization as a way to have music play back at a consistent volume. Spotify, for example, sets a loudness target of -14 LUFS. If a track has a loudness level higher than this, the platform will decrease, or normalize, the volume to around -14 LUFS.

Due to the unique nature of object-based immersive audio, researchers suggest that many stereo mixing and mastering techniques must be relearned, particularly in relation to metering and headroom. For example, it is common practice for mix engineers to rely on master bus processing to shape the sound of a track. This practice makes use of the left-right channel, where all tracks are routed to and summed down to stereo. From there, the track is bounced to disk and ready for distribution. The master bus is the only point where the elements of a track are combined, and it is often utilised to apply global processing, such as compression, equalization (EQ)[15], or saturation, to polish the overall sound of a mix. However, in an object-based immersive format, there is no master bus. Even with channel-based surround audio, utilising master dynamics and EQ to unify the elements in a mix together is rarely applicable, as it can yield different results than in a stereo workflow. One use of master bus processing in stereo is to boost perceived loudness. The aim to not surpass 0 dBFS[16] on the master bus usually means mixes will be compatible with all stereo playback systems. With an immersive format, on the other hand, headroom is not only desired but also required. As the

[12] Loudness units full scale is a standard loudness measurement unit used for audio normalization
[13] The loudness war was a trend of increasing audio levels in recorded music, which reduced audio fidelity
[14] Squashed is a term used to describe audio that has little difference between the loud and soft, which can be fatiguing to years when listened to at high volume.
[15] EQ is the process of adjusting the volume of different frequency bands within an audio signal.
[16] dBFS refers to decibels relative to full scale

renderer determines what goes to each speaker in the listener's system during playback, there must be enough headroom to ensure that these calculations never result in distortion or clipping. Maintaining sufficient headroom is crucial to ensuring that a mix translates accurately across different playback systems. This requires careful attention to levels and dynamics throughout the mixing and mastering process.[17]

A good mix should not only sound great on high-end speakers but also on consumer-grade headphones and portable devices. To achieve this level of compatibility requires a deep understanding of both the creative and technical aspects of mastering. This research aims to provide insights into the challenges and opportunities that object-based audio presents to mastering engineers, a topic that underlines the need for continued innovation and development within this field, to ensure that audio engineers can fully leverage the potential of this exciting new medium. This study also aims to examine the accessibility of immersive audio. In an attempt to establish how such audio content is prepared for distribution, interviews with professional mixing and mastering engineers, coupled with observations made on existing research, will contribute to a clear and concise body of work that addresses the research question at hand.

[17] Inglis, Sam. "An Introduction to Immersive Audio." Sound on Sound, Jan. 2022, www.soundonsound.com/techniques/introduction-immersive-audio. Accessed 27 Apr. 2023.

Literature Review

This project responds to the work of audio professionals working with object-based audio, along with research projects that propose new methods of mastering for this medium. A selection of literature has been chosen, spanning an array of topics that must first be considered before approaching the research question proposed by this project. As well as establishing an understanding of immersive audio in the context of this study, it is important to also define the mastering stage. As previously explained, investigating object-based audio is of particular interest. Due to the rapidly evolving advances that come with researching this technology, it is important to consider that various problems and considerations proposed in these sources have evolved since the time of writing.

A work that can aid in rationalizing the evolution towards object-based mediums is Dennis Baxter's "Immersive Sound Production: A Practical Guide" (2022). This book introduces the essentials of immersive sound capture and creation. With real-world examples of microphones, mixing, and mastering, this book presents thorough explanations of production practices and the possibilities of immersive sound, and examines the technology that makes it possible.

Spatial audio is a continuing evolution. This term is used to describe audio that gives the impression of space beyond conventional stereo, allowing the user to pinpoint where a sound is coming from, whether above, below, or in a full 360-degree sound field. Understanding spatialization as a principle offers insight into the history of how users consume audio.
As explained by the author, spatialization refers to how sound is distributed in a specific environment and can be observed as the order, or degree, of separation. The first separation is from mono to stereo, where a sound can be clearly perceived and segregated into two channels. The second order of separation came when sound production evolved from stereo to surround sound. Even with this next step in the evolution of sound production, the audio is still only produced on a horizontal plane, whether in front of, to the sides of, or behind the listener. Adding height gives way to the third order of separation among sound components, which has been recognized as significantly expanding the apparent sound field. Placing a sound object anywhere in the 360-degree sound space around the listener with accurate localization in the horizontal plane is achieved by horizontal panning. Horizontal panning can

image a sound anywhere around the listener, including directly to a speaker or between two adjacent speakers, creating a phantom image between them.[18]

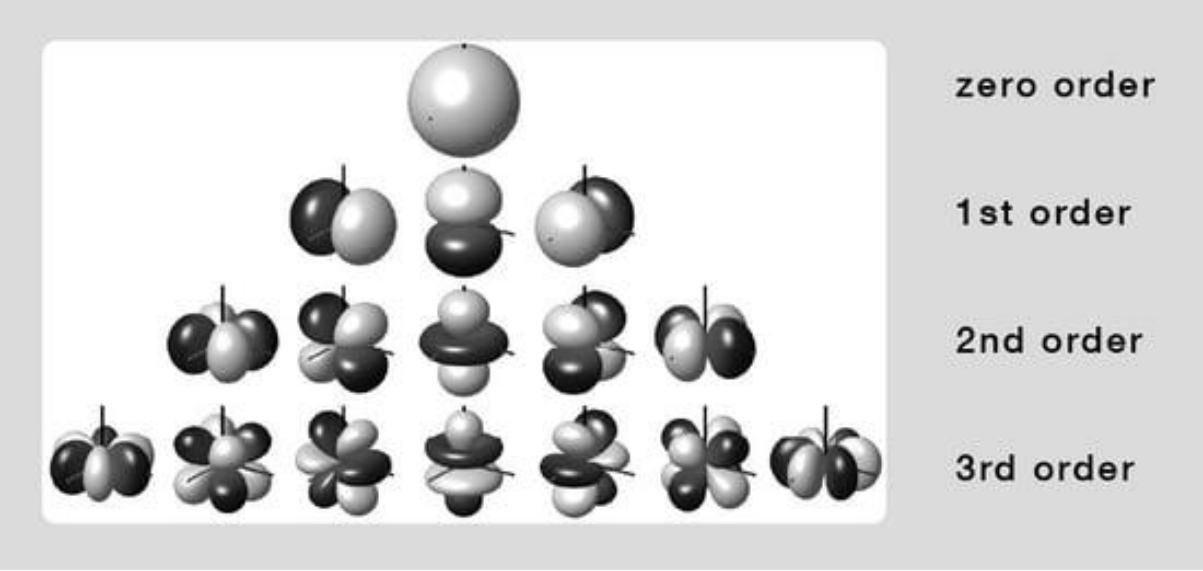

Fig. 1 Santos, Iolanda. *What is Ambisonics?* 2022

It is also suggested by the author that understanding how the brain perceives sound is imperative to rationalizing sound-shaping techniques based on psychoacoustic principles, that can be utilized during the mastering process. For example, it is suggested that spatial organisation and separation can significantly improve the clarity of a mix by unmasking audio elements. Frequency masking is an auditory phenomenon that occurs when the perception of one sound source is affected by the presence of another. Unmasking sound elements allows them to be separated and easily heard without increasing volume levels, as the ear can easily differentiate sounds that are only separated by a few degrees from each other. This book looks at equalization as a spatialization technique, i.e., rather than using elevation to achieve height, using equalization can also be effective in adding the impression of height. The concept of "added space" allows for sonic elements in a mix to be spread out across a number of speakers, rather than the convention two channels that stereo offers. This creates not only opportunities for the mastering engineer, but also raises potential problems that will need to be overcome.

[18] Baxter, Dennis. Immersive Sound Production: A Practical Guide. Routledge, 2022. 29-46

Along with understanding the evolution of sound towards immersive formats, it is also beneficial to the comprehension of this study to define key terms that relate to the mastering process. The Audio Engineering Society's technical council presented "Recommendations for Loudness of Internet Streaming and On-Demand Distribution" in 2021, which represents the collective knowledge of a group of experts and lists recommendations for establishing and implementing an effective distribution loudness for streaming. Previously, loudness targets were referred to within the context of the competitive increase in loudness that existed for many years in the music industry. For a mastering engineer, loudness and dynamics are two key considerations. Ensuring a song is loud enough to compete among commercial tracks in the market without compromising the dynamic nature of a performance can be a crucial task.

Measured by the unit LUFS, or LKFS. integrated loudness is defined as the average loudness between two points in time, measured electrically using the BS.1770 loudness algorithm and gating.[19] In other words, it demonstrates the average loudness measured over a whole track. The acronyms LUFS (loudness units with respect to digital full scale) and LKFS (k-weighted loudness with respect to digital full scale) are interchangeable and represent identical values. For the rest of this research paper, LKFS will be used. As noted previously, streaming services often have specifications regarding integrated loudness levels, and may also have a slight tolerance specified. In the event that they don't meet these specifications, the tracks are normalized. Loudness normalization applies uniform attenuation or gain in order to match the loudness of content to the desired distribution loudness. The intention of this was noted in this document as a way to "reduce annoying loudness jumps and the incentive to produce loud content." In other words, it is used to maintain loudness consistency and discourage mastering tracks to an unnecessarily loud level, as was often the case with the "loudness war."

[19] "Dolby Atmos Music Master Delivery Specification: Version 2022.07." Sales Force, 15 July 2022, dolby.my.salesforce.com/sfc/p/#700000009YuG/a/4u000000lFJ9/IVkxI54tPvDbGymtjQ6tyIjEvkaA7xPa_Byq6.vH_dA. Accessed 22 Apr. 2023.

Content		Distribution Loudness (LUFS)	Upper Tolerance (LU)	Loudness Measurement Method
Assorted [1]	Speech is measurable [2]	-18	+1	Dialog Integrated Loudness [3]
	Speech is not measurable (See Section 5B)	-18	+2 Format-specific see Table 2	Integrated loudness
Music [4]	Track-normalized [5]	-16	+0.2	Integrated loudness
	Album-loudest track (e.g., on-demand music services, See Section 5D)	-14 [6]	+0.2	Integrated loudness
Interstitial [7]		-18	+0.2	Integrated loudness
Virtual Assistant		-18	n/a [8]	Integrated loudness of assistant's voice preceding volume control

Fig. 2 AES, *Recommendations for Loudness of Internet Streaming and On-Demand Distribution.* 2021

When analysing how a source's loudness control impacts the loudness of a downmix[20] to fewer channels (for example, from stereo to mono), it's vital to grasp the distinction between acoustic and electrical summation of program elements in stereo or multi-channel mixes. For example, an element that is equally present in both the left and right channels of a stereo recording will sound about 3 dB louder in stereo than in either channel alone. However, when the stereo channels are electrically summed to mono, the elements add arithmetically, so in this example, the element will be 6 dB louder in the mono mix, and the relative balance of the elements in the mono sum will differ from the stereo; the balance will change depending on the element's placement in the stereo panorama. This means that when heard in mono, loudness control applied to a stereo mix can have errors as large as 3 LU[21]. This concept raises questions when thought of in the context of immersive audio. For example, just as a stereo mix (two channels) can be downmixed to mono (one channel), a key principle of the Dolby Atmos format is that it can be downmixed to other speaker configurations, i.e., if the playback system has fewer speakers available, the renderer will output the audio signals to

[20] Downmixing refers to mixing of a number of distinct audio channels to produce a lower number of channels
[21] LU is an abbreviation for the term Loudness Unit. 1 LU is equal to 1 dB

suit the system. It is important for the mastering engineer to understand how this may influence the level of a track

This document also details how peak limiters are used in audio codecs to prevent clipping. An audio codec is used for the compression or decompression of digital audio data.[22] However, each limiter design has its own sound, which can make the resulting audio quality unpredictable. Mobile devices, for example, have built-in peak limiters that prevent power amplifiers from clipping. Content with a high peak-to-average ratio may activate these limiters, with unpredictable, subjective results. As presented in the journal article "MPEG-H Audio—The New Standard for Universal Spatial/3D Audio Coding," audio codecs such as MPEG-H, allow for the universal carriage of encoded 3D sound from channel-based, object-based, and HOA-based[23] input.[24] This codec is used by the Sony 360 RA format, and it is important to understand its capabilities and performance, particularly in relation to concepts shared with the mastering process.

In this research article published by the Journal of the Audio Engineering Society (AES), Juergen Herre et al. address two concepts that are common to the workflow of mastering in the context of this audio codec: loudness normalization and dynamic range control. Within MPEG-H 3D audio, comprehensive loudness related measures, according to ITU-R BS.1770-3, are embedded into the stream for loudness normalization. The audio signal is normalized by the decoder to map the program's loudness to the desired target loudness for playback. As discovered in the AES recommendations, downmixing and dynamic range control can alter the loudness of the signal. As such, a dedicated program loudness metadata can be included in the MPEG-H bitstream[25] to ensure correct loudness normalization. Dynamic range control (DRC)[26] is stated as a vital framework to control the dynamic range and is also noted to have improved clipping prevention and peak limiting. This concept required further investigation

[22] Techopedia. "What Is an Audio Codec? - Definition From Techopedia." Techopedia, 25 Jan. 2017, www.techopedia.com/definition/26302/audio-codec. Accessed 28 Apr. 2023.
[23] HOA refers to higher-order ambisonics, a technology that allows for accurate capture and delivery of 3D audio on any device
[24] Herre, Juergen, et al. "MPEG-H audio-The New Standard for Universal Spatial/3D Audio Coding." Journal of the Audio Engineering Society, Dec. 2014, www.researchgate.net/publication/286938585_MPEG-H_audio-The_New_standard_for_universal_spatial3D_audio_coding.
[25] Bitstream is a sound conversion system that translates the analog output audio signal into digital bits, transferring the information to a receiver
[26] Dynamic range control (DRC) is the adaptive adjustment of the dynamic range of a signal. The dynamic range of a signal is the logarithmic ratio of maximum to minimum signal amplitude specified in db.

to grasp the relationship between loudness normalization and dynamic range control as technical aspects of the audio codec and how they are managed creatively during mastering process.

Before continuing towards the practise of mastering object-based audio, it is first imperative to address the purpose of the mastering stage. As previously noted, mastering is the final stage in the record making process. It is the last creative step in the audio production process, and serves as a bridge between mixing and replication, via distribution.[27] In his book, "Mastering Audio: The Art and the Science," Bob Katz describes it as "the last opportunity to enhance sound or repair problems within an acoustically treated room." It is a craft that requires meticulous attention to detail as well as technical and musical knowledge. This book acts as a complete guide to both the technical and artistic principles of mastering audio.

The mastering engineer's workflow includes critical auditioning, editing, level processing, and output to the final medium. At the mastering stage, the mix is done. The mastering engineer may make modest EQ corrections, but nothing that would alter the mix. During mastering, the final adjustments are made to ensure that the audio is consistent and balanced across all playback systems. This may involve enhancing the overall sound quality by adding equalization, compression, and limiting. The ability to notice audio dropouts, phasing, frequency ranges of sibilance, electrical noises, such as ticks, clicks, and pops, and bad edits are some of the things an experienced mastering engineer should be able to recognize. The author also stresses the importance of ear training, stating that it should be "a lifelong activity, and no one can become an expert overnight."

Within his comprehensive and practical book, Katz details clear and comprehensive explanations of the mastering process. This philosophy of mastering, which has been referenced by many, has proved to be a valuable resource, and the ideas expressed continue to be relevant within the ever-evolving subject of immersive audio. Many ideas expressed throughout this resource are seen to be expressed by the subjects interviewed for the purpose of this study, which supports the idea that research into this topic is not for the purpose of redefining what it means to master audio, but rather identifying how the goals and intentions

[27] Katz, Bob. *Mastering Audio: The Art and the Science*. 2nd ed., Focal Press, 2007, p.19-83

that have always existed as formats developed and changed can be respected and upheld within this new frontier.

It is also noted by the author that what defines the master has become vague, as projects may have authoring, i.e., metadata, added at some stage down the line.[28] At the time of publication, this resource had only gone as far as mastering techniques for stereo and surround sound. Further analysis is required to understand how adding metadata can affect the master. One principle that has remained the same throughout the evolution of audio formats is “the mastering engineer's responsibility to ensure the audio quality that leaves the studio will be the same quality that will be represented on the final medium.” The goal of presenting the mix in the best possible way is still at the core principle of the mastering stage. In this book, Katz suggests that it is best to avoid mixing and mastering at the same time, as this can distract from the goal of mastering. This is an interesting point, and previous research undertaken for the purpose of this study calls into question its relevance in the practice of mastering for immersive audio. It has been observed that as the mastering workflow is still somewhat undefined for immersive audio, some engineers tend to overlap the mixing and mastering stages. Reliance on the format chosen to render a mix to different playback modes can result in issues if time isn’t taken to learn the nuances of mastering for immersive formats.

Moving on, it is important to consider the research carried out, which directly relates to the objective of this research. In their 2018 paper "An Approach for Mastering Audio Objects” Hestermann et al. couples the art and science of the mastering process with the evolution towards immersive audio formats, as they present ideas for object-based spatial audio mastering. As well as outlining the main limitation that relates to mastering objects-based audio, a concept introduced in this paper is universally applicable to both metadata and audio signals. Despite a growing interest in object-based audio, both commercially and as a research topic, there have been few production tools entering the market that are designed to work with this type of audio. This paper presents new object based spatial mastering techniques, showing different interaction approaches for mastering using “mastering objects,” which were realized in a prototype to evaluate the concept as part of an interactive system.

[28] Katz, Bob. *Mastering Audio: The Art and the Science*. 2nd ed., Focal Press, 2007, p.19-83

Audio content creation can be traditionally divided into three stages: recording, mixing, and mastering. Although all three steps are well defined for channel-based audio, this research once again notes that object-based audio requires new workflows. The mixing process only differs slightly, as the sound engineer no longer pans signals to dedicated loudspeakers. Instead, all positions of audio objects are usually set in a spatial authoring tool, which allows for the definition of the metadata component of every audio object. A complete mastering process for audio objects has not been developed so far, although it is observed in this paper that such a process would have benefits similar to those in channel-based workflows. As another practise of mastering engineers is to increase the overall perceived loudness of a mix to conform to distribution standards, engineers need to ensure satisfying sound quality across different audio codecs and listening scenarios. Therefore, various formats and bitrates[29] are monitored during the mastering stage of this study in order to ensure the best quality product.

One explicit challenge of mastering object-based audio scenes lies in the existence of both metadata and audio data, as opposed to only audio material in channel-based mastering. Although adjustments to the audio component of audio objects are likely to be more important, tools for mastering metadata are also necessary to optimise spatial scenes. This paper was successful in proposing a tool that would optimize object-based mixes for optimal sound quality, and while the mastering process remains undefined after publication, the authors are successful in demonstrating why it is a worthwhile subject to be investigated as they explain how the translation of audio objects spatialized by a renderer creates room for many unexplored issues. In his paper "Mastering Object-Based Music with an Emphasis on Philosophy and Proper Techniques for Streaming Platforms" (2020), Michael Matakis provides an evaluation of cutting edge object-based mastering techniques, along with a philosophy for mastering audio-objects in any scenario.

This study states that measuring the subjective quality of object-based masters requires the use of specific parameters unique to the medium's traits. Within his work, Matsakis references a previous study in which object-based masters made in Dolby Atmos were played through a 5.1[30] setup as well as a 5.1 channel-based master of the same material. They

[29] Bitrate is the amount of data being transferred into audio
[30] 5.1 refers to a surround sound system that uses five audio channels and one LFE channel

defined five measurable aspects of the sound as spatial immersion, localization, dynamics, audio quality, and format preference.

A key focus of this work was identifying and creating techniques for mastering audio objects for streaming platforms, while also exploring and noting the effects of experiencing the same object-based master across several playback systems. To achieve this, four professional mastering engineers who were currently working with immersive audio content were interviewed for this study. The interviews were centered around each engineer's techniques and philosophies for working within the medium. The techniques proposed by these mastering engineers were then put into practice, in order to master an immersive piece of music, and detail the effects of listening to the same object-based master across several playback systems. The strengths and weaknesses were noted, and using this information, a new workflow was established. These masters were then put forward to an expert panel as a means of further, unbiased review. The conclusion was that each of the techniques was viable for mastering audio objects, and while mastering object-based audio is possible with the current technology, it requires a deep understanding of the software's technical limitations as well as a concrete philosophy for thinking about immersive audio itself. Matakis noted that each of the subjects made it a point to stress that engineers who are used to working in channel-based environments "must *first* relearn how they think about the positioning of sound sources in the space of a mix".

This research was successful in demonstrating different workflows for mastering engineers working with immersive audio, and so a similar approach was adopted for this study. Conversations with both mixing and mastering engineers proved to be a viable way of learning about the mastering process for this medium, as well as understanding the differences between different industry-standard formats. There are several immersive audio formats that exist in the marketplace. One format that is often referenced is Dolby Atmos (Atmos). Originally, it was marketed as an evolution of cinema sound, bringing 3D immersion where once there was just stereo. Today, it is one of the leading immersive audio formats, adopted by Apple for subscribers of Apple Music to experience spatial audio, and integrated into Logic Pro[31] for creators to begin creating in immersive. In version 2022.07 of the "Dolby Atmos Music Master Delivery Specification," technical requirements for

[31] Logic Pro is a DAW developed by Apple Inc.

deliverables are set out, and make reference to synchronization and alignment, loudness, mix and approval environment, and the formats specific master parameters. At the time of writing, this was the most up to date version of the document, and observations can be made regarding what a mastering engineer must keep in mind when delivery for this format.

Immersive audio has become more accessible in recent years, and this format allows for the creation of compelling immersive audio that does not require knowledge of the target speaker configuration. In addition to being an object-based format, an Atmos mix can contain one or more 'beds." Each bed is essentially a conventional channel-based stream in up to 7.1.2 surround (i.e. seven horizontal channels, one LFE channel, and two height channels). The idea behind this format is for the Atmos playback tools to intelligently downmix a track, when playing back on a system with fewer speakers. As well as distributing audio signals to available speakers during playback, it calculates the levels required to best recreate the location originally chosen by the audio mixer. Hence, a proposed benefit of this would be that the audio would only need to be mastered once, and it would no longer prove necessary to do separate 5.1 and 7.1 mixes. From a mastering engineer's perspective, it is imperative to become familiar with how this may affect the audio quality of a finished track, as again, one of the main objectives is to ensure accurate translation across all formats.

This document also suggests that all deliverables should be monitored in a room with the 5.1.X[32] speaker configuration. Furthermore, it recommends that deliverables should be monitored in room approved for home listening, with a speaker layout of at least 7.1.4-ch speakers. It is also suggested that all deliverables should be monitored over headphones with the headphone renderer mode set to "binaural." There are strict guidelines when it comes to synchronization and alignment, as well as loudness. If a master file does not conform to these specifications, it will be rejected by streaming platforms, i.e., Apple Music. It is stated that all deliverables must be synced to the corresponding stereo deliverable. This means that there should not be any silence at the start or end of the deliverable when compared to the corresponding stereo deliverable.

[32] "X" refers to the number of overhead channels

In terms of loudness, an important consideration for a mastering engineer, this document states that loudness should be measured using the Dolby Atmos Renderer[33]. With this format, individual tracks must achieve an integrated Dolby Atmos loudness that does not exceed minus 18 LKFS based on ITU recommendations BS.1770-4. It also illustrates that an integrated loudness lower than - 18 LKFS is permissible to enable wider dynamic range content. A tolerance of ± 0.1 LKF is permitted due to differences that can be introduced by the measurement process for albums. However, loudness must be measured on an individual track basis rather than the whole album. The loudest track on an album should not exceed -18 LKFS, and again, there is a tolerance of ± 0.1 LKFS due to differences that may be introduced by the measurement process.

The true peak[34] specification states this level is to be no more than –1 dB for the Dolby Atmos loudness measurement. A tolerance of ± 0.1 DB is again permitted. This loudness target is said to ensure translation across different playback modes without modification; it also allows for a wider dynamic range. As the music is spread out around to the listener, there is greater space in the mix, which allows the mix engineer to make modest EQ and compression adjustments, allowing for subtle harmonics and timbres to exist freely in a mix rather than turning down louder sonic elements to prevent masking. Understanding Dolby Atmos file requirements is essential for mastering for this format. Further Dolby Atmos music master parameters are outlined in this document, which state that all deliverables must use 24 bits PCM[35] resolution and a sampling rate[36] of 48 kHz. For projects created at 96 kHz DAMF (Dolby Atmos Master File) should be used if delivering for archiving. This means that while projects can be created at the higher resolution of 96 kHz, the deliverable should be 48 kHz.

On episode 81 of his podcast, "The Mastering Show," Ian Shepherd offers an engaging insight into the practical work that goes into making an immersive track release ready. Through conversations with two immersive mastering engineers, this discussion puts into context the limitations and practicalities of immersive mastering, with a focus on the Dolby

[33] A Dolby Atmos application that generates metadata, allowing for accurate spatial translation
[34] True Peak refers to the maximum level that a signal reaches
[35] PCM refers to pulse-code modulation. 24 bit PCM is the number of bits of information in each sample, and it directly corresponds to the resolution of each sample
[36] Sample rate is the number of samples per second that are taken of a waveform to create a discrete digital signal, measured in kHz

Atmos format. While examining the ways in which mastering for Atmos is different than traditional stereo mastering, they detail how it works in practice.

This podcast further explains the Dolby Atmos specification discussed previously, which states that the loudest song must not exceed -18 LKFUS integrated. Here, it is explained that this value is stated to allow for the necessary processing that needs to happen if the mix is to be played back on different systems. This specification also allows for a more dynamic mix. With the target loudness of Atmos to be no more than –18 LKFS, the mastering engineers describe it as "freeing," as they decide how much limiting and compression to use based purely on the sonics of a mix, rather than trying to achieve extra loudness.

One observation offered is that there will be demands for mastering Atmos mixes as it may be needed to fix issues in the mixing due to unfamiliarity with the format itself, i.e., the deliverables and loudness specifications. During this discussion, Shepherd speculates about the future of mastering this format. He puts asks to the interviewees whether they believe a small, select group of mixing engineers will setup their studios with a full Atmos rig, and mastering-grade monitoring, resulting in tracks requiring only light touch mastering, or rather the more affordable option of Atmos mixing using only headphones, and keeping the mastering process separate to check how the mix then translates on a full speaker setup. This idea relates to the accessibility of the format, particularly for music creators, and is of interest for this study to explore further. The opinions offered conclude that having extensive experience working on speakers is necessary before working from headphones, as it's important to understand how the binaural engine responds to the material. Having a good workflow holds potential for getting the most out of a headphone-based workflow.

The conversational style of this resource allowed for further explanation of topics touched upon throughout this literature review and proved valuable as a way of coupling research with the practical applications of immersive mastering.

Conclusion

In conclusion, various resources have been examined in this chapter, each providing valuable insight into some of the key aspects pertaining to both mastering and immersive audio, and how they coexist. The concept of ensuring the overall balance, level, and consistency of a song is stated as some of the main focuses of the mastering engineer in the resources

examined. It is clear that establishing a method for mastering immersive audio requires a deep understanding of a broad scope of topics, such as those touched upon in this review. It is evident from the above resources that there are mastering engineers that have found ways to adapt to the object-based medium, and it is of benefit to the progress of this research to investigate this further.

Methodology

The question remains: how is an immersive audio track mastered and prepared for distribution? Before moving on to the research design, it is important to keep in mind the objectives of this research and to investigate how mastering engineers are currently working with this content. In this chapter, this research objective will be addressed in more detail in order to inform the chosen methodological approach undertaken as well as address the limitations faced.

i. Research Design

There is limited information available online on what it means to master an immersive audio track. Despite this, streaming platforms such as Apple Music, TIDAL, and Amazon Music support immersive audio.[37] Without any clear answers readily available online, it became clear that an effective way of researching this topic would be to include interviews and conversations with various mix and mastering engineers who are working with labels as well as independent artists to create immersive mixes for release, in order to establish the work flow of professionals. Approaching such discussions first required a solid understanding of the topic, informed by the literature review, in order to inform what questions needed to be asked. In order to gain clarity on where the gaps in research fall, it is imperative to include an up-to-date evaluation of techniques utilized by professionals using modern technology. These conversations will serve as a foundation to addressing research question at hand, in order to progress this research further.

ii. Interviewing Process

Two mastering engineers were chosen to be interviewed, along with one mix engineer. Chris Le Monde and Darcy Proper are both mastering engineers who are currently working on mastering immersive projects. To provide comparison, mixing engineer Ruadhri Cushnan gave a deeper insight into the mixing process and gave his opinion on why an immersive track needs to be mastered. The interviews consisted of a semi-structured approach where subjects were asked several questions, informed by previous research, listed in appendix. A. The eligibility criteria desired when selecting subjects to interview were mixing and

[37] Scarrott, Becky. "Hi-res Music Streaming Services Compared: Which Should You Sign up For?" *What hifi*, May 2022, www.whathifi.com/advice/hi-res-music-streaming-services-compared. Accessed April 2023.

mastering engineers with experience with stereo who have evolved their workflow to include immersive audio, and it was of benefit to the research project to gather their opinions on the future of the format. Although mastering is the main focus of this study, the decision was made to briefly investigate the mixing stage, as it is a crucial step in the audio production process and lays the foundation for the mastering stage. Understanding the relationship between these two stages and identifying how they may overlap can lead to a more efficient and effective workflow in immersive audio production.

Results

On completion, the interviews were transcribed verbatim and thematically analysed. One practical advantage of this methodology is that highlights hidden information. Sometimes, an engineer would share information unrelated to the question at hand but relevant to a previous or forthcoming question. Therefore, the method allowed for a more comprehensive analysis of the data and a better understanding of the context in which the information was presented. Additionally, it could potentially lead to new insights and avenues for further research.

There were certain limitations to this method of investigation. Limitations to the study design include time and access to appropriate software and studio setups, which may have been useful as a method to practice object-based mastering. Questions were designed to compensate for such limitations by asking what is needed to begin working with immersive audio. A qualitative method was argued for as the most effective methodological approach based on the limited relevant data and contextual appropriateness. The qualitative method also allowed for in-depth exploration of the participants' experiences and perspectives, which would not have been possible with a quantitative approach.

The topic of immersive audio is extremely broad, and there are many factors to consider, even when investigated in the context of mastering. As observed from the literature review chapter, current research in the area of immersive audio is continuously evolving, and a mastering workflow is still largely undefined. Having the resources to put into practice what was learned throughout the interviews may prove an effective way to confirm the workability of a workflow and hold the potential for future research. The next sections detail the most significant findings on mastering object-based content from the interviews, as well as the specific mastering signal flows from each interviewed engineer.

Analysis

This chapter presents the analysis work carried out throughout this thesis and focuses on how effective the chosen methodological approach has been in relation to the project objectives. The chapter will also highlight the limitations of the study and suggest areas for future research. Overall, this analysis provides valuable insights into the effectiveness of the chosen methodology and its potential implications for similar projects in the future.

i. Analysis of Methodology

Due to the exploratory nature of this research, the qualitative methods utilized proved efficient and effective. Approaching a subject with the goal of understanding a workflow that can be used by prospective master engineers meant paying particular attention to the opinions of those with experience in this field. It was observed that there is a large number of engineers working worldwide who are excited by the evolutions in immersive audio and eager to continuously learn and experiment with new ways of working. At the heart of these conversations was always the delivery of a final mix that remained true to the intentions of the artists.

ii. Analysis of Interviews

The aim of this research was to draw up a series of guidelines that could help those who want to adapt their master workflows to immersive music. Meaningful learning was derived from conversations with mix and master engineers working with immersive audio. To achieve this, important analysis of interviews was used to identify common factors among various workflows and responses to similar questions. Two of the interviews took place during visits to respective studios set up for immersive audio, and one took place over zoom. The perspective offered from direct sources was an integral part of understanding how an immersive audio track becomes release-ready.

iii. Interview Results

When asked about the workflow for mastering, subjects tended to agree that mastering for immersive is a relatively new concept, and that it hasn't exactly been figured out yet. The way in which EQ and limiting play a role in traditional mastering has yet to be defined in the context of immersive audio. The subjects interviewed each made a point of stressing the importance of understanding the delivery specifications of different formats, i.e., loudness specifications, the number of channels on the output, and naming conventions.

Subjects were asked what they receive at the beginning of a project. For context, it is typical that a mastering engineer would receive a finished mixed mix and begin the mastering process. Alternatively, they may receive the individual stems of a finished mix, in order to have more control over individual elements. This is known as "stem mastering."
Le Monde began by explaining that a typical workflow would begin with receiving all the stems of a finished track along with the mastered stereo file that is used as the reference track. In order to create a 3-dimensional mix that reflects the creative intentions of the song, the stereo master is used as a reference track. In a scenario such as this, he is both mixing and mastering the immersive version of the track. Working from a song that has already been mixed in stereo, the goal is to replicate the stereo mix as closely as possible in the 3D surround space. The ability to have creative control over the individual elements at this stage may be imperative to achieve this, and as well draws a comparison to the practise of stem mastering.

As put forth by Cushnan, it is important that the immersive mix match the stereo mix as closely as possible; therefore, it is important to be aware of a balance shift that may happen when you spread a sound source out in an immersive mix. For example, it is important to be aware of the impact of spreading a sound source out into five speakers rather than the usual two-channels with stereo, and how this affects the tonality of that sound in the mix.

A potential issue brought up by subjects was that mistakes in the stereo mix may be revealed in an immersive mix. The theory of frequency masking,[38] as referred to previously, may cover up an element in a stereo mix that now sticks out in an immersive mix. Another common finding related to the dynamic range of a track. As there is a definite loudness specification for the Dolby Atmos file delivery, each engineer stressed that if it is not met, it will result in a track being rejected from streaming services.

When asked how one learns to master immersive content, subjects tended to respond that the place to start is learning how to mix immersive audio, in order to learn the nuances of formats. A variety of perspectives were expressed among the interviewees. The consensus

[38] Frequency masking is an auditory phenomenon that occurs when two similar sounds play at the same time, or in the same general location. One masks the other, confusing ones perception of either sound.

was that the end format is an important consideration from the start. Due to the differences between immersive audio formats, there is no one method for mastering all immersive audio content. Having the knowledge and experience to recognise what might cause issues is an asset. Knowing the studio experience, learning the speaker array, and how they translate other devices is beneficial, and a common recommendation observed in all interviews was to visit other studios and listen to as many speaker setups as possible.

Conclusion of Analysis

Having discussed the topic at length, it is clear that compiling a set of guidelines for mastering immersive audio would prove more difficult than once anticipated due to the complexities of working with multiple formats. Furthermore, it has been established that this is a rapidly evolving topic, meaning that since the time of writing there may have been new advancements or ways of working. As a result, the understanding of immersive audio as a blanket term to many different formats has significantly improved. It is crucial to keep up with the latest developments in immersive audio to fully comprehend its potential and limitations.

Secondly, when looking at the use of interviews throughout this project, it can be seen that they were executed effectively and with precision. Despite the difficulties in understanding a common practise, the interviews elicited exactly what was needed for the development of the project, and provided valuable insights into the diversity of approaches used by professionals. The findings from the interviews can be used to inform future research and practise in the field. Overall, the interviews proved to be a successful method for gathering data on this complex topic.

In the following chapter of this research paper, important points made will be discussed and explained in further detail, which will be useful to progressing the research, as it may be useful to eventually compiling a list of guidelines. The discussion will also shed light on the challenges faced during the project and how they were overcome, providing insights for future similar projects. These guidelines can serve as a starting point for other researchers in the field of mastering, and may also help establish a more standardized approach to the process.

Discussion

The previous analyses showed that while this thesis has been successful in exploring both mastering and immersive audio, there remains room for learning. The discussion chapter of this thesis will demonstrate the learning gained from undertaking the project, and highlight interesting points that warrant further discussion. Furthermore, the successes of the project will be detailed and discussed in comparison to its original objectives.

Before attempting to establish an evolved mastering workflow, it was first helpful to ask why an immersive audio track needs to be mastered. As is the case with Dolby Atmos, one of the formats' main purposes is that a file can be translated across many different playback methods. Similarly, the MPEG-H codec used by Sony 360 RA contains information that prevents clipping and normalizes the music to ensure safe translation across playback systems.[39] Cushnan suggested that if the mix has been done by an engineer with enough experience working with immersive and who understands the nuances of the format for which they are delivering, the mastering process may consist of simply ensuring there are no errors, and everything lines up the way that it should. For engineers who may not have access the necessary speaker setup at home, it is plausible that many will work on a headphone-based workflow. In this case, having the option of sending the mix to a mastering engineer who can listen to it on the appropriate speaker set up may be imperative to ensuring that everything sounds as the artist intended. Overall, the consensus was that quality control checks were critical at this stage to ensure that a finished track is not get rejected by streaming platforms.

There are strict specifications for the delivery of finished mixes to streaming services. It is specified during each interview that, where possible, the stereo and immersive versions should line up in terms of the file length. Matching where musical events fall inside that file is imperative. Lining up to the stereo in this sense is very important, as it should allow the listener to switch back and forth between the immersive and the stereo version of a song. However, while they should line up in relation to time and loudness, it is recommended to

[39] Herre, Juergen, et al. "MPEG-H audio-The New Standard for Universal Spatial/3D Audio Coding." Journal of the Audio Engineering Society, Dec. 2014, www.researchgate.net/publication/286938585_MPEG-H_audio-The_New_standard_for_universal_spatial3D_audio_coding.

make use of the immersive space, as instructed by the artist or label, and not just make another stereo master.

Exceeding the stated loudness specification was highlighted by Le Mode and Cushnan as another common reason why a track may be rejected by a streaming platform. As previously noted, with Doby Atmos, there is a loudness specification for the loudest song on an album to be -18LKFS. This allows the mastering engineer to preserve dynamic mixes. Traditionally, in mastering, there has been a push for things to be loud, which can result in a squashed sound limit the dynamic range of a track. With the added space offered by immersive, there is a freedom for engineers to make decisions regarding limiting and compression based purely on the sonic elements rather than pushing for extra loudness where it is not needed. In order to ensure a good balance, it was recommended by all subjects to listen and ensure all elements are sitting nicely before beginning with the loudness process.

When further explaining the loudness specification for Dolby Atmos, immersive mastering engineer Darcy Proper explains that -18 LKFS is measured at the 5.1 downmix of the Atmos mix, and that gets further folded down for the binaural render. Therefore, if the immersive version is too loud, once it gets translated for the binaural headphones, it is probable that it will clip and cause distortion. Hence, -18 LKFS is a number that is generally safe. This means that audio engineers should aim to mix the immersive version at or below -18 LKFS to prevent distortion and ensure a high-quality listening experience for users. Additionally, this standard helps maintain consistency across different platforms and devices and provides a safe listening experience for consumers.

It was pointed out during each interview that in comparison to Dolby Atmos, Sony 360 RA has not stated a particular loudness specification. Sony 360 RA relies on the user's discretion to avoid potential harm from excessive volume levels. While they have not stated a particular loudness level, they do provide a limiter as a safety feature to prevent distortion[40] and maintain sound quality when a mix is folded down into the binaural headphone experience. It is once again worth noting that this limiter acts as a safety feature of sorts, and is not used as a creative tool, unlike a limiter that may be included in the mastering engineer's mastering chain. This means that the limiter is not meant to enhance or alter the sound in any way but

[40] Distortion refers to the altering or deformation of an audio signal's original waveform

rather to ensure that it remains consistent and distortion-free. Overall, this demonstrates a commitment to providing a high-quality listening experience for users of binaural headphone technology.

An objective of this research has been to investigate the accessibility of this spatial audio. It is no longer necessary for consumers to have elaborate speakers set up in order to listen to sound created for immersion. Streaming services now offer both new and existing tracks in spatial audio, which can be listened to over headphones. As it is becoming increasingly accessible for consumers to listen to immersive audio in a variety of ways, the most popular being headphones, the question arises if advancements in head-related transfer function (HRTF)[41] tracking and binaural rendering merit the use of headphones when mastering immersive content rather than with a speaker array. This idea was put forth to interviewees, and a variety of perspectives were expressed.

As the translation of audio objects spatialized by a renderer creates room for many unexplored issues, each subject leaned towards using a full speaker setup, rather than mixing and mastering solely on headphones. However, practise of cross-referencing their work on several fold-downs was recommended to ensure translations and compatibility across different systems. The binaural downmix of immersive content can be somewhat unreliable, as the technology for personalized HRTFs is still evolving. As such, a master may sound different to different people if the same HRTFs are assumed for every listener. Therefore, using the binaural render as a reference to make slight adjustments, after they have finished making larger processing decisions was recommended. Cushnan offered the observation that there may be low-frequency content that is hidden when spread across multiple speakers, but is overpowering the mix when listened to on headphones, and certain frequencies may have been attenuated when downmixed into two channels.

In the case that engineers choose to mix and master immersive content using headphones, for the reasons outlined above, it may be necessary to send the finished track to a mastering engineer with a full speaker set-up for the final seal of approval. Knowing what happens when objects are placed in certain areas and understanding the tonal differences from the binaural engine usually result from spending extensive time getting familiar with all the

[41] HRTF is a phenomenon that describes how an ear receives sound from a sound source

intricacies of each format. Proper offered, it is important to have enough listening experience to be sonically aware of potential timing and phase[42] issues, distortion, balance issues, and other kinds of problems that can crop up. Being able to recognize these types of problems and figure out if they are inherent in the material received is an important skill. If was inherent in the material received, it's important to consider if there is anything that can be done to fix it. On the other hand, if it was induced during the mastering process, it is crucial to identify the cause and rectify it before proceeding with the final mix. Proper went on to explain that, particularly in the case of album work, it's imperative to have a good sense of how the material should fit together to create a coherent listening experience.

Subjects were asked what they believed to be the necessary experience to begin working with immersive audio content. Trends observed included the importance of being well versed in routing and signal flow, as well as audio fundamentals such as healthy gain structure and calibration. If incorporating an analogue workflow, it's imperative to ensure that everything is calibrated. Without proper calibration, the analogue workflow may introduce inaccuracies and inconsistencies that could negatively impact the final output. Therefore, it is crucial to have a thorough understanding of the calibration process and ensure that all equipment is properly calibrated. Additionally, audio networking knowledge is essential to seamlessly connecting different machines and devices for a smooth workflow. Most importantly, it's recommended to have sensitivity and respect for artists and their music. Being capable and willing to approach a project with the goal of doing as much or as little as necessary to effectively deliver the artist's message to the end listener. Being careful not to take artistic liberty was a point made by each subject.

When asked about the future of this technology, subjects were optimistic about the evolution of a defined workflow. Proper offered an insightful observation relating to high-resolution pictures, stating that it is often picture that drives the audio industry. There is a continued need for sound that matches the visual experience of a high-resolution picture, as in the film and television industries, the audio quality can greatly enhance or detract from the overall viewing experience. As a result, there has been a push for advancements in audio technology to keep up with ever-improving visual quality.

[42] Phase refers to the position of a sound wave in time

As time goes on, it appears that the streaming industry is becoming increasingly eager to support object-based music content. However, the medium is *currently* struggling to match the predictability of channel-based audio. While immersive audio can boost listener engagement and help refresh or revitalize an older catalogue, there is also an opportunity to consider new implications of this technology and further research potential.

Overall, the insights gathered throughout this research have proven to answer the question put forth at the start of this study. Understanding that there are many formats that fall under the term “immersive audio” helped to conclude that there are no rules or instructions that will work for everything. Continuing to be curious about the format and gaining practical experience are key to establishing an individual workflow for mastering immersive audio content. While the professional experience of the engineers interviewed for this study has allowed them to adapt to new ways of working as audio delivery formats have evolved, it was universally agreed that there is no established mastering workflow as of yet. Despite this, it can be observed that when mastering immersive content, the stereo master should be used as a reference, both for an album and an individual track. Continuously checking that timing, loudness, and artistic intent lines up with the stereo file is imperative to delivering a good master.

Conclusion

This research set out to establish how an immersive audio track is mastered and prepared for release. Extensive research and analysis were carried out, and as a result, this methodological process has yielded significant learning outcomes that could be used to provide a foundation for mastering object-based audio. The reality of working with audio in the creative technology industry is that there is no single approach that would be effective for every situation. The techniques and workflows investigated in this research should generally be taken as starting points for beginning to work with object-based formats.

Although it is reasonable to desire clear and concise results, the potential to progress further comes from analysing shortcomings of a project. If this research were to be further investigated, experimenting with the workflows uncovered during this study would better establish the capabilities of individual formats and their compatibility with industry-standard DAWs and would provide more evaluation of popular techniques used in the professional world. Furthermore, efforts to understand the listener's perception of object-based music content and their preferences towards it could also aid in the adoption and growth of this medium. It is important to consider both technical and consumer aspects when evaluating the potential success of object-based music content. For example, if the audio will primarily be listened to in a car, it may be necessary to adjust the mix accordingly to ensure optimal sound quality.

It is important to understand the potential impact of immersive audio on the music industry and how it may shape the way music is produced and consumed in the future. For example, Dolby has collaborated with car makers Mercedes-Benz to deliver a Dolby Atmos listening experience to a broad range of their vehicles, bringing “the next generation of in-car entertainment to the consumer.”[43] This illustrates a move towards immersive listening experiences, and highlights the need for audio professionals to consider it as part of their workflow. In conclusion, this project design proved appropriate for answering the research question as it successfully used real world sources to uncover best professional practice.

[43] “Dolby Atmos for Cars - Dolby Professionals.” Dolby Professional, professional.dolby.com/dolby-atmos-for-cars. Accessed 28 Apr. 2023.

Bibliography

360 WalkMix Creator™. "Home - 360 WalkMix Creator." *360 WalkMix Creator - Plugin Software Compatible With Popular Digital Audio Workstations*, 2 Mar. 2023, 360ra.com.

"Audio Bit Depth and Sample Rate Explained." *iZotope*, 5 May 2020, www.izotope.com/en/learn/digital-audio-basics-sample-rate-and-bit-depth.html. Accessed 28 Apr. 2023.

Baxter, Dennis. *Immersive Sound Production: A Practical Guide*. Routledge, 2022.

"Definition of 5.1 Channel." *PCMAG*, www.pcmag.com/encyclopedia/term/51-channel. Accessed 28 Apr. 2023.

"Definition of Immersive Audio." *PCMAG*, www.pcmag.com/encyclopedia/term/immersive-audio. Accessed 28 Apr. 2023.

"Dolby Atmos for Cars - Dolby Professionals." *Dolby Professional*, professional.dolby.com/dolby-atmos-for-cars. Accessed 28 Apr. 2023.

Dolby Atmos Mastering. episode 81, 1 Mar. 2022. *Apple Podcasts*, podcasts.apple.com/ie/podcast/81-dolby-atmos-mastering/id1095931813?i=1000552610353. Accessed 22 Apr. 2023.

"Dolby Atmos Music Master Delivery Specification: Version 2022.07." *Sales Force*, 15 July 2022, dolby.my.salesforce.com/sfc/p/#700000009YuG/a/4u000000lFJ9/IVkxI54tPvDbGymtjQ6tyIjEvkaA7xPa_Byq6.vH_dA. Accessed 22 Apr. 2023.

Vickers, Earl. "The Loudness War Do Louder, Hypercompressed Recordings Sell Better?" *AES*, June 2011. *AES E-Library*, www.aes.org/e-lib/browse.cfm?elib=15934.

Herman, Gabe. “What Is Immersive Audio and How Does It Work?” *Sonarworks Blog*, 23 May 2022, https://www.sonarworks.com/blog/learn/immersive-audio. Accessed 22 Apr. 2023.

Herre, Juergen, et al. “MPEG-H audio-The New Standard for Universal Spatial/3D Audio Coding.” *Journal of the Audio Engineering Society*, Dec. 2014, www.researchgate.net/publication/286938585_MPEG-H_audio-The_New_standard_for_universal_spatial3D_audio_coding.

Hestermann, Simon, et al. “An Approach for Mastering Audio Objects.” *Journal of the Audio Engineering Society*, vol. P10-2, July 2018. *AES E-Library*, www.aes.org/e-lib/browse.cfm?elib=19631.

Inglis, Sam. “An Introduction to Immersive Audio.” *Sound on Sound*, Jan. 2022, www.soundonsound.com/techniques/introduction-immersive-audio. Accessed 27 Apr. 2023.

Katz, Bob. *Mastering Audio: The Art and the Science*. 2nd ed., Focal Press, 2007.

Melchior, Frank, et al. “Spatial Mastering - a New Concept for Spatial Sound Design in Object-based Audio Scenes.” *ResearchGate*, Jan. 2011, p. 1, www.researchgate.net/publication/267840563_Spatial_Mastering_-_a_new_concept_for_spatial_sound_design_in_object-based_audio_scenes.

Michael, Matsakis. *Mastering Object-Based Music With an Emphasis on Philosophy and Proper Techniques for Streaming Platforms*. New York University, 2021.

Olivieri, Ferdinando et al. “Scene-Based Audio and Higher Order Ambisonics: A Technology Overview and Application to Next-Generation Audio, VR and 360° Video.” *European Broadcasting Union*, Nov. 2019, tech.ebu.ch/docs/techreview/trev_2019-Q4_SBA_HOA_Technology_Overview.pdf. Accessed 28 Apr. 2023.

"Recommendations for Loudness of Internet Audio Streaming and On-Demand Distribution." *AES E-Library*, 24 Sept. 2021, www.aes.org/technical/documentDownloads.cfm?docID=731. Accessed 28 Apr. 2023.

Santos, Iolanda. "What Is Ambisonics?" *Sound Particles*, 7 Oct. 2022, blog.soundparticles.com/what-is-ambisonics. Accessed 28 Apr. 2023.

Scarrott, Becky. "Hi-res Music Streaming Services Compared: Which Should You Sign up For?" *Whathifi*, May 2022, www.whathifi.com/advice/hi-res-music-streaming-services-compared.

Techopedia. "What Is an Audio Codec? - Definition From Techopedia." *Techopedia*, 25 Jan. 2017, www.techopedia.com/definition/26302/audio-codec. Accessed 28 Apr. 2023.

"What Is Dolby Atmos Music Mastering?" *Abbey Road*, Nov. 2022, www.abbeyroad.com/news/what-is-dolby-atmos-music-mastering-3273. Accessed 25 Apr. 2023.

"What Is HRTF?" *3Dio*, 19 June 2019, 3diosound.com/blogs/learn-about-sound/what-is-hrtf. Accessed 28 Apr. 2023.

Appendix A.

Template for Interview Questions

1. How have your mastering strategies evolved when working with sounds beyond the stereo format?

2. What does your typical workflow look like, and how do traditional stereo mastering techniques fit into mastering an immersive track?

3. With added space, have you noticed any changes in regard to the dynamic range and loudness specs of immersive tracks, or how the timbre of individual sounds changes due to special trajectory?

4. Do you use the stereo master file as your reference track?

5. What is your checklist for deciding an immersive track that you have mastered is finished/ release ready?

6. What are the audio deliverables? Do you complete separate masters for different formats (i.e dolby atmos, sony 360 etc.)

7. In your opinion, would advancements in HRTF tracking and binaural rendering merit the use of headphones when mastering immersive content, rather than with a speaker array?

8. What would you see as the potential issues of opting to master immersive audio on headphones?

9. What experience do you need to begin mastering immersive audio?

10. What advice would you give to mastering engineers looking to adapt their workflows to include immersive mastering?

11. How do you expect the role of a mastering engineer to evolve in the future?