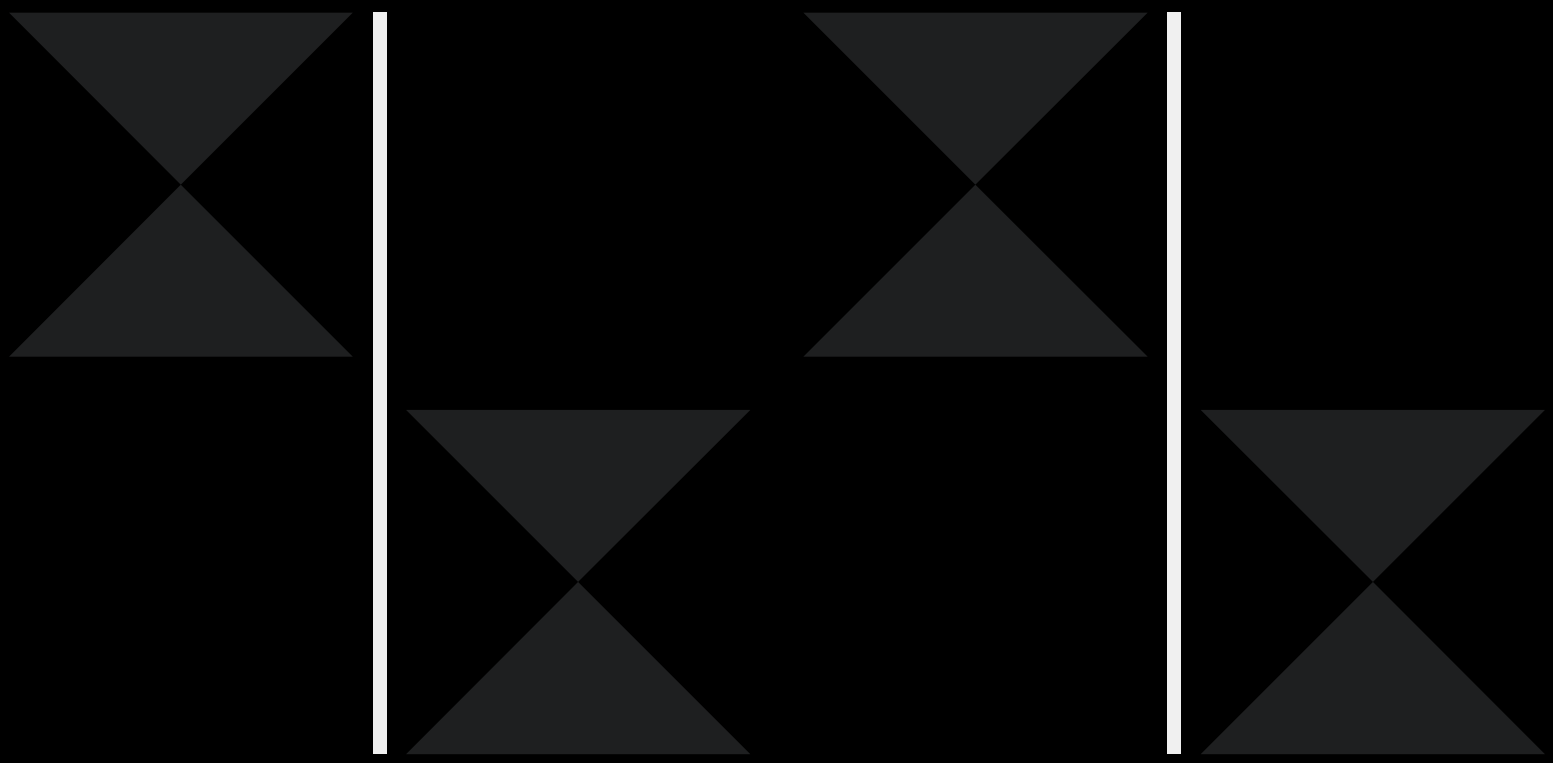
SoundID Reference

from Sonarworks



Immersive audio

eBook

Introduction

Immersive audio seems to be everywhere today. It's the buzzword of home entertainment systems, automobiles, smart speakers, games, and of course (most) music streaming platforms.

This ebook will demystify the technical and creative aspects of producing and listening to immersive music. Much like 5.1 and 7.1 audio, immersive audio faces an uphill battle for consumer acceptance due to the required playback equipment, physical space, and simply the technical understanding required to take advantage of all the variations in the delivery and playback environments. Nonetheless, immersive audio, building upon a century of triumphs and setbacks within the audio industry, has arrived at a time when music is more accessible than ever before and the technology exists to provide some sort of immersive experience in many different playback environments. These environments range from personal earbuds and living room smart speakers to elaborate multichannel listening rooms and theaters.

For those interested in creating immersive music and audio stories, we've included an interview with Richard Furch, a talented and extremely successful engineer and mixer, who will provide some insight into the creative approach and considerations when developing immersive mixes.

The field is so new that there are few rules and creators who put in the effort now will be the ones to define the future of the format. Read on and get inspired!

How To Listen to immersive → audio

In the beginning, there was mono, played from a single speaker with a limited sense of realism. Around 1940, two-channel recordings emerged, which produced the impression of a 3-dimensional space where the music seems to float between the left and right speakers in a panoramic perspective. Stereo became the de facto standard for music over the past 80 years. Multichannel sound for cinema, utilizing more than two speakers has been in place in one form or another since Disney's 1939 release of Fantasia.

Over the last three decades, music companies have attempted to market music in various surround formats, typically 5.1 or DTS Music. Surround music has never gained a serious foothold due to several factors, including large file sizes, complicated playback requirements, and mostly a lack of proper marketing and customer demand. However, home video and audio systems for TV, movies, and gaming have evolved to include surround-capable speaker systems, and audiences have become more accustomed to surround audio. This awareness, along with some recent technological developments has brought about the newest generation of multichannel or immersive audio. Let's examine some of the most common immersive audio formats and learn how we can enjoy our music this way.

Surround formats for film and music have historically been "channel-based" systems, where various audio elements can be routed to speakers in up to seven zones around the room.

For instance, dialog can be anchored in the center of the movie screen while a helicopter circles around the room. Modern immersive formats expand on the channel-based model and add audio "objects," which can move dynamically to virtually any specific location in a room, including above the listener. Moreover, the relative locations and movements of sounds will be the same in any compatible playback system, no matter if you're in a large theater or an automobile. For instance, sounds no longer come from "the right side of the room," but from an exact location in 3-D space around the listener. Immersive audio formats allow sounds to be individually placed in a channel-based bed or into object-based elements for extreme flexibility in creating a sense of spaciousness and dynamic movement inside a mix.

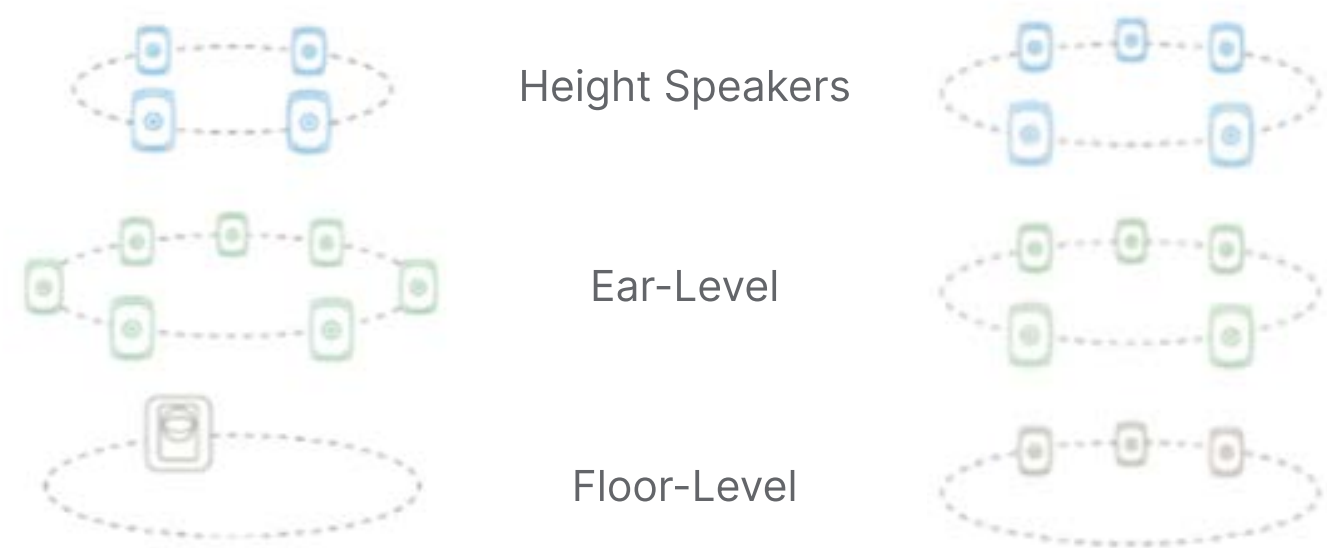
Music mixes that contain many layers of instrumentation and complex orchestration can sound great in stereo, but spreading individual sounds into an immersive soundfield brings out details that wouldn't be

noticed in a dense stereo mix. Sparse arrangements, like Herbie Hancock's "Watermelon Man" benefit from immersive mixing's ability to add dimension and depth beyond stereo mixes. Symphonic scores by Hans Zimmer, rock music by Alice In Chains, and chill tracks from Jack Johnson can all benefit from spreading out the music in creative ways that do not detract from the original intent of the song but instead seem to bring the listener physically into the mix. As the immersive music world matures, producers will create music with immersive playback in mind-creating parts and layers that are meant to be experienced in an immersive environment.

> The top two competing formats for immersive music are Dolby Atmos Music and Sony 360 Reality Audio, with other formats like Auro3D and DTS:X and THX Spatial Audio seeing some limited use.

These immersive formats can be enjoyed on multichannel speaker systems or on standard headphones and most streaming services support one or more of these formats. As an emerging technology, the methods for producing, distributing, and consuming these formats are

developing almost daily, and over the coming months and years, these formats will become commonplace, as will the ability to enjoy these formats on more and more playback systems.



Dolby Atmos 7.1.4.
12 Channels

Sony 360 Reality Audio
13 Channels

If you've never heard immersive audio before, click here to experience an online headphone demo from Dolby that allows you to compare stereo mixes to Atmos mixes for a few different songs. This YouTube clip demonstrates the 3-D placement of sounds using Sony 360 Reality Audio over headphones.

Photo credit: SoundShepherds.Studio

How to enjoy immersive audio?

Immersive audio is produced and mixed on speaker systems that include ear-level surround speakers and additional height speakers above and even below the listening position.

A Dolby Atmos Music system, for instance, requires at least a 7.1.4 speaker system, while Sony 360 requires a 13-channel speaker system. Don’t fret, though, you can enjoy Atmos and 360 RA without purchasing a dozen speakers! Immersive formats are designed for consumers, so playback is supported on headphones and many consumer playback devices, such as soundbars and smart speakers

Immersive audio can be streamed on mobile devices, desktop computers, gaming consoles, and streaming devices like smart TVs, and the Apple TV 4K.

Keep in mind!

Immersive audio is an emerging technology, so there are some growing pains we must deal with in order to jump into listening to immersive audio.

Stream or Purchase?

Immersive audio can be streamed in Atmos, 360 RA, and Apple Spatial Audio formats, it is difficult to purchase immersive music.

The few sites that sell downloadable immersive music offer Dolby TrueHD, Auro3D, and Dolby Atmos Mp4, which are not true object-based immersive mixes, but the mixes have been downmixed 5.1 to 9.1 surround formats. If you wish to listen to these downloadable formats on speakers, you will need a player, like an Apple TV 4K or NVIDIA Shield TV Pro, plus an AVR home theater decoder to act as a monitor controller for your speaker system or headphones. For good or bad, most consumers will choose to stream immersive audio instead of purchasing it, so let's take a look at that scenario.

Streaming services such as Apple, Tidal, and Amazon offer immersive audio streaming in Atmos, 360 Reality Audio, or Apple Spatial Audio. Depending on which streaming service and which immersive format you wish to enjoy, your hardware requirements will differ. First, you must check with your streaming service to see which devices support their immersive audio stream. For instance, Tidal supports Atmos playback with a HiFi Plus subscription on mobile phones and Atmos-compatible streaming devices including the Apple TV 4K, Fire TV devices, and the NVIDIA Shield TV Pro. Tidal streams Sony 360 RA on any device that can run the Tidal app with headphones, but to play 360 RA over speakers, you will need to connect the app to speakers or an AVR with a Sony 360 RA chip built in.

Most listeners will use headphones to listen to immersive audio, so let's take a look at that case.

Immersive audio on headphones

Tidal and Amazon Music can stream Atmos and Sony 360 RA over any standard headphones from their app on an Apple or Android smartphone without any additional requirements.

Apple Music supports Atmos as Apple Spatial Audio on iOS devices and Apple silicon desktop Macs (M1 or M2) running the Mojave OS. Deezer supports 360 RA with its 360 mobile app. Nugs.net provides live concert streams encoded in 360 RA and PeerTracks.com provides streaming and blockchain distribution of virtual events and music encoded in 360 RA. Qobuz streams a limited number of titles in THX Spatial Audio, which is compatible with standard headphones and simulates a 7.1 surround environment.

Most streaming apps include a preference for stereo vs. immersive playback.

In Apple Music, for instance, simply set the preference for Dolby playback to Automatic or Always On and any song that has an Atmos version will play in Atmos, otherwise, it will play in stereo. Apple Music also supports dynamic head-tracking if you are using a current iOS and newer Apple Airpod or certain Beats headphones.

Headphone playback of immersive content provides a sense of ambience beyond standard stereo reproduction, but each streaming service may encode its immersive playback slightly differently and therefore the same music mix will sound different on different platforms. For instance, Apple utilizes custom binaural processing, instead of Atmos's processing, while other streaming services support Atmos's binaural processing.

This difference in encoding will affect the sense of space and tonality of a mix and you might not experience the mix exactly as the producer intended.

Hopefully, these differences in playback will disappear as the technology matures, but for right now, expect the immersive audio experience on headphones to differ slightly between different platforms.

Surround on speakers

Immersive audio can be enjoyed at home, but you must have a specific speaker setup and certain equipment to participate. For instance, you may be able to stream Atmos mixes via Apple Music or Tidal on your Apple TV 4K or other streaming appliance, but you must then connect that device via HDMI to an AVR (Audio/Video Receiver) with Atmos capability. The AVR can then route the audio to your speaker system. To hear full Atmos, you must have at least a 7.1.4 speaker system, but the AVR can also downmix a mix to work on smaller setups, including 7.1, 5.1, or stereo, albeit with less immersive quality. Sony 360 requires playback hardware, like an AVR or smart speakers, that contain a decoding chip specifically for 360 Reality Audio.

The Sony RA5000 wireless Smart Speaker contains 7 speakers to provide simulated Sony 360 RA from a single speaker. This tabletop speaker supports stereo and immersive audio playback.

Many soundbars and smart speakers are capable of playing back immersive formats. For instance, the Echo Studio smart speakers provide support for Atmos as do offerings from Sonos, Apple, Sony, Samsung, and many others. Sony and other manufacturers sell speakers that are compatible with 360 RA. . Smart speakers and soundbars are simple to set up and take up very little space in a room, but keep in mind that smart speakers and soundbars only produce an approximation of a full multichannel system. If you want to hear a true Atmos mix over speakers, you need to set up a proper 7.1.4 listening environment, while 360 RA is optimized for a 13-channel playback system, comprised of 5 surround speakers, 5 overhead speakers, and 3 floor-level speakers. Audiophiles and home theater enthusiasts will demand a proper immersive system, while the average consumer will likely be satisfied with the sound from a decent soundbar or set of smart speakers.

Immersive audio for creators

Music creators should mix immersive productions on properly calibrated multichannel monitor systems and also check the mixes on headphones.

Mixes should also be auditioned in stereo and binaural since we don't know how consumers will hear the mix. A mixer may have to coach the artist or client on the differences that will be heard between a full multichannel playback system, a soundbar, and the various streaming services with headphones. A mix that translates well to all platforms is possible, but there may also be a preference to sound best on a particular playback system or platform.

Immersive audio can be created in just about any modern DAW and tools are available to optimize the creation process. Dolby provides free panning plugins for DAWs as well as the paid Dolby Atmos Production Suite and Mastering Suite. Sony sells the 360 WalkMix Creator plugin software to author in 360 Reality Audio. Logic Audio (10.7 +) provides a fully integrated Atmos workflow and requires no additional software, so it may be the most cost-effective solution for Atmos music creation.

While immersive mixes should be checked on headphones, the mix should be created on a full speaker system. In addition to a DAW and speakers, a monitor controller that calibrates the speakers for level, EQ, and timing is necessary. This can be a hardware controller or a software solution.

Offerings from AVID, NTP, Grace, and MetricHalo provide hardware monitor and DSP control of an immersive speaker system, while software like SoundID Reference Multichannel provides these functions as an app on your computer.

To get started mixing in Atmos or 360 RA, I would advise spending some time with somebody who has expertise in the process.

I strongly recommend the in-depth information from Dolby on their free Dolby Atmos Music Training website. There are also many good Atmos mixing tutorials online from reputable sites, like Mix With The Masters and Puremix.net.

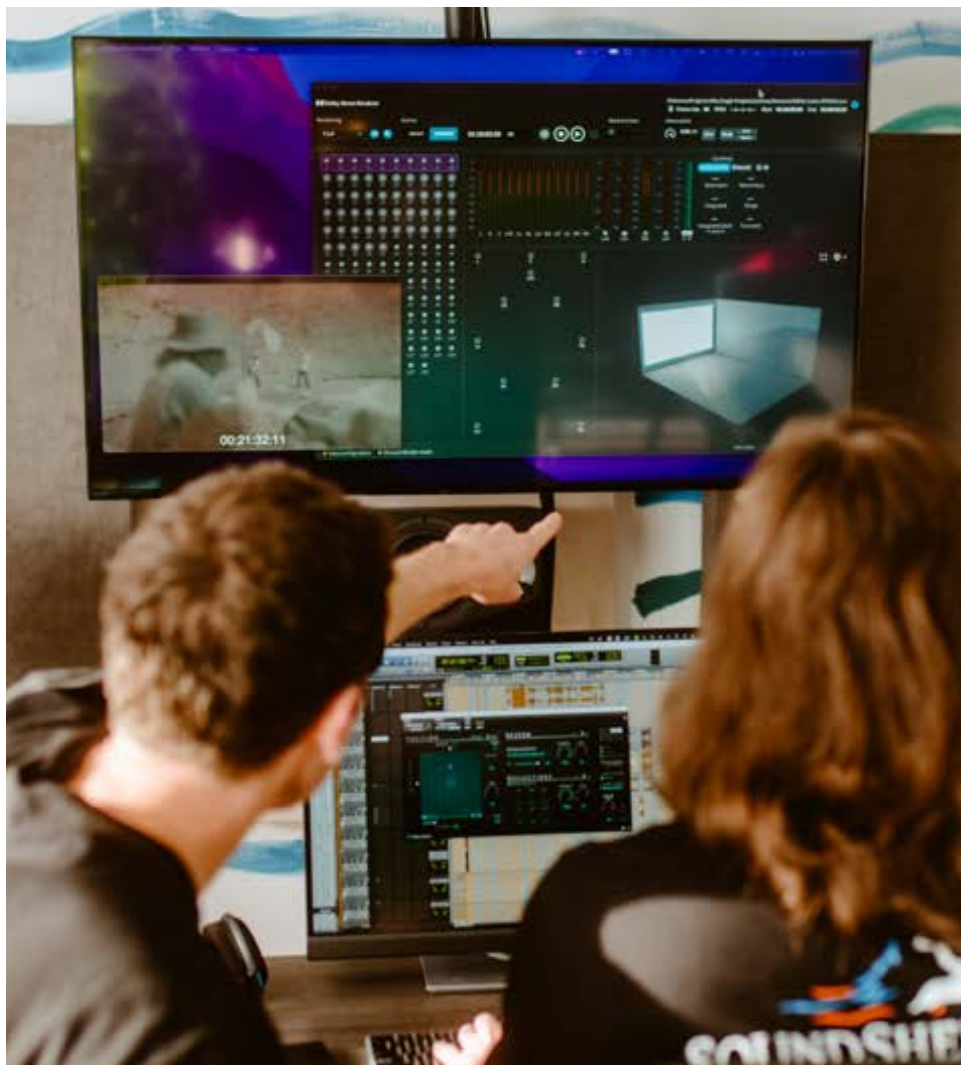


Photo credit: SoundShepherds.Studio

Mixing in immersive formats requires a different mindset and workflow than mixing in stereo and may take some time to master. If you want to become an expert in multichannel mixing, don't wait for your first project to come in before you start practicing!

Recommended immersive → listening

Some songs work extremely well in immersive formats, either because of the arrangement or simply the attention and care put into the immersive mix. Here are a few recommended songs to inspire you and provide an idea of how effective immersive mixes can be.

“Drivers License” Olivia Rodrigo

This simple pop arrangement spreads out from left to right and front to back to enhance its emotional impact. The ambience doesn’t detract from the song, but feels natural, even when the chorus vocals come from the sides and back.

“Bet” Snarky Puppy

This room-filling jazz/rock fusion track shows off the clarity of a dense mix and also the low-end clarity that is available in immersive mixing. This track lets the listener sit in the room with the band.

“Boom” Tiesto and Sevenn

This club track provides the atmosphere of a pounding dance club when listened to in Atmos. The bass and kick fill the room and the percussive elements float around the room.

“Dream of Arrakis” Hans Zimmer

This atmospheric score of layered percussion, eerie synths, and processed vocals is best experienced in an immersive format.

“Kill or Be Killed” Muse

This nu-metal guitar-laced anthemic track shows off the depth and clarity available for dense, high-energy rock songs. Muse is known for their sonic approach and this track really shines in Atmos.

“Meet the Moonlight” by Jack Johnson

This cozy laid-back track fills the room with just two wide-panned guitars and yet the mix can still grow to include layered vocals, more guitars, and synth.

“Bury a Friend” Billie Eilish

This clubby track spreads out in Atmos even beyond the stereo version, making room for the squeaks, vocal harmonies, and crackles that create the spooky mood of this song. This song has become a showpiece in Atmos rooms.

“Dil Bechara” A.R. Rahman

This Sony 360 RA mix provides the warmth and room-shaking bottom of a classic reggae record and provides space for the percussion, vocals, and electronic elements to fill the room around you.

“Rocket Man” Elton John

This groundbreaking Atmos mix by Greg Penny shows off the effec-tiveness of surrounding the listener with the music. In this mix, you’ll notice the background vocals move from the front to the upper rear and the slide guitar travels over your head, like a rocket. This track has become a standard listening test for Atmos speakers setup. It is almost as if this 1972 track was produced to be heard in Atmos.

“Watermelon Man” Herbie Hancock

This simple classic fusion record from 1962 sounds amazing in Sony 360 Reality Audio, spreading out the flutes and vocals while keeping the presence and funk of the bass and drums. This mix pays homage to the original mix while adding just enough dimension to feel fresh and immersive.

Spend some time cruising your favorite streaming service and search under Atmos or Sony 360 for staff recommendations, curated playlists, and immersive mixes from artists you are familiar with. Once you develop a taste for immersive mixes you may never wish to listen in stereo again.

The immersive audio → family tree

How do we begin developing immersive audio skills? Surprise! You probably possess some already. Although immersive seems like a totally new animal, immersive audio is only the most recent branch on the audio family tree. Immersive audio didn't drop out of the sky. Multi-channel immersive arrays have organically evolved from earlier formats in a natural progression.

This article provides a historical introduction to formats that shaped the technical and philosophical elements of modern immersive audio. You will observe how elements from earlier systems are still evident in the architecture of immersive audio systems today. The insights gained from these past formats can be effectively applied to our approach in a modern immersive environment. As the cliché goes, "The more things change, the more they stay the same."

Mono

Consider the nascent audio industry when everything came through one speaker (Ca.1877). The quality of recordings was limited by the quality of the engineers' equipment. Audio recording was a technological marvel — a miracle of science! Inventors of early recording machines like phonographs, wire recorders, and gramophones were all vying for the title of most life-like machine. But lifelike wasn't yet possible, so a greater value was placed on composition and performance. This axiom still holds true; there's more benefit to a poor recording of a great performance than a great recording of a poor performance.

Early recording engineers learned that loud marches could cut through the noise of the recording medium. Since there was no mixing to be done after the recording process, everything had to be perfectly blended on the way in by carefully placing musicians around a mic in a good-sounding room. The foundation of creativity, emotions, and passion should always be the most salient elements of a recording. The best historic and modern recordings showcase performances that connect emotionally with the audience.

Even as the number of audio channels increased from one (mono) to two (stereo), and on to many (multichannel), the mono presentation of music remains an important consideration, as many listeners will still experience a mono, or mono-sounding version of a song out in the world.

Stereo

By the 1930s, the novelty of recording had begun to lose its luster, and inventors began to research using more than one channel. British audio scientist, Alan Blumlein, became fascinated with the idea that recordings could capture sound in a similar fashion to how we hear with our ears. Beginning around 1933, Blumlein tinkered with microphone placement methods that could mimic human binaural perception. Although originally

Record labels embraced the idea of immersive sound and partnered with manufacturers to produce stereo playback systems. By the 1970s, just about every new vinyl release was issued in stereo, and mono was relegated to television and AM radio broadcast. Blumlein demonstrated that the way we hear music matters. The more lifelike a recording seems, the easier it becomes for the listener to transport themselves into the music.

Today's engineers create transcendent stereo mixes. Just as with mono, the left and right stereo speakers orient sound in front of the listener, but with a wide soundstage and phantom center image.

Quadraphonic (4.0)

One limitation of stereo is that the sound stage is always in front of the listener. Envelopment is possible, but it doesn't quite match the sonic environment of sitting in a room with a live band. One proposed solution to this problem was to add a set of stereo speakers directly behind the listening position. These speakers would "surround" the listener with ambient cues to simulate the experience of a venue. This format was called Quadraphonic Sound ("Quad"), and it was the first commercially available multi-channel surround sound format. Quad derived its name from its delivery method, which consisted of four discreet channels

1. Left
2. Right
3. Left Surround
4. Right Surround

In 1971, CBS introduced Quad in both vinyl and 8-track formats. Specially pressed vinyl records had four different etchings in each groove. Turntables for quad needed special styluses with four needles. Each channel carried the sound for one of the dedicated speakers. 8-track cartridges offered the same quad experience on analog tape.

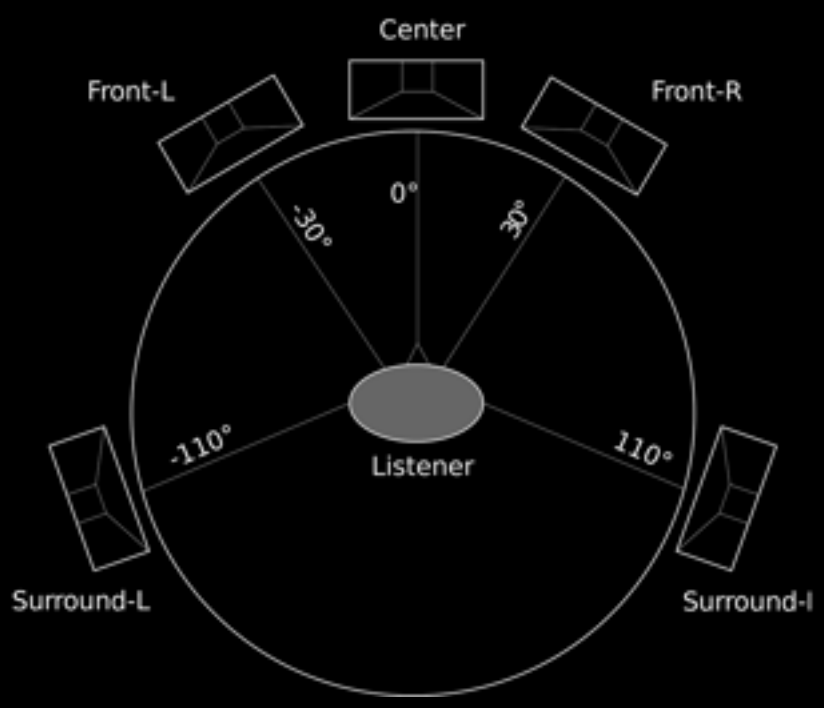
Consumers had difficulties accepting quad as a format. The quality and output level of quad recordings suffered as compared to stereo on both vinyl and 8-track. In addition, quad turntables and commercially released records were expensive and incompatible with standard stereo and mono systems. To experience quad sound, the listener had to remain in the center of the speaker setup. Because of these issues, quad flopped in the commercial market, but there were significant lessons learned from the experience that set the groundwork for subsequent surround formats.

5.1

Despite the apparent failure of quad, surround sound remained an alluring idea for audio research labs like Dolby. Digital audio offered improvements to frequency response, dynamic range, and noise floor, which created an opening for surround to make a comeback. The film industry rushed to digital and worked closely with Dolby to develop a discreet six-channel format that consisted of a quadrophonic setup with the addition of a center speaker, and a low-frequency effects channel (LFE). This format was called 5.1 surround; the 5 defines the number of speakers at ear level, and the .1 is the dedicated LFE channel.

With 5.1, center-channel dialog would remain centered on a large movie screen, while music and ambient sounds had a home in the left-right channels Ambience, effects, and music inhabit the rear channels and sounds could dynamically move around the audience to create special effects. Subsequent versions of surround included more channels to improve coverage and increased the size of the listening sweet spot in large theaters (7.1 or more speakers).

Batman Returns (1992) was the first blockbuster movie to be mixed in 5.1, with Jurassic Park following shortly. The compelling sound design and exciting surround mix contribute to the success of these films. This surround sound ushered in a new, exciting, and sometimes terrifying cinematic experience. Theater sound soon came into people's living rooms with "Home Theater Systems," where movies on DVD-A or Blu-ray provided surround to go along with your big screen TV. The future looked bright for the format, and the music industry began issuing releases in 5.1 on SuperAudio-CD and DVD-A.



The film industry made the most of surround sound, but the music industry was less successful. To get the best experience, the listening position must be a fixed point, and the speakers must properly surround the listener This was just not practical for the dimensions of the average living room, and consumers took a very lax approach to adhering to placement guidelines. The result was a complete breakdown of quality control and ultimately, a disappointing listening experience.

It should be noted that around the same time (ca. 2005), Apple introduced the revolutionary iPod, which allowed listeners to carry their entire music collection in their pockets. Goodbye, beloved CD wallet! iPods played stereo music only and spurred the mp3 format that quickly became the de-facto consumer playback format.

Immersive formats

Only 15 years after the iPod, we welcome the modern immersive audio formats, which expand upon the basic setup of a 5.1 or 7.1 surround system. Traditional surround formats generate a 2-D sound field where sounds can move laterally around the listening position in a circle. Immersive systems add the third dimension–height channels. Now sound can truly surround and envelope the listener. Additionally, immersive formats, like Atmos and Sony 360 Reality Audio encode audio channels with metadata that instructs the playback system where to place sound in a particular room. The soundfield is therefore scalable from small rooms (even cars) up to large cinemas.

The LFE channel is used in an immersive mix as in earlier surround applications—only for low-frequency effects. Some speaker systems may use the subwoofer channel as an extension of the main speakers, and this is referred to as a bass-managed system. The LFE channel typically uses its own subwoofer and only contains special effects low-frequency content.

> The minimum setup:
> For Dolby Atmos playback on speakers is a 7.1.4 system, comprised of 7 ear-level speakers, 4 overhead speakers, and one subwoofer. Sony 360 RA requires three layers, 5 ear-height, 5 overhead, and 3 floor-level.

MPEG-H Audio is quickly becoming the immersive format for interactive virtual and augmented reality content, such as gaming and live events. The Auro-3D, DTS:X, and Pure Research immersive formats are used in some TV, film, and multimedia presentations.

Immersive audio can be enjoyed with soundbars or smart speakers, which may contain several speakers to emulate a multi-point playback system. Additionally, immersive mixes can translate acceptably well to headphones, and Atmos and Sony 360 RA contain binaural information

(first studied by Alan Blumlein) that provides a sense of ambience when multichannel audio is replayed on headphones. Binaural information simulates how human ears perceive sound in real spaces and works for headphone listening.

For the creator

The architects of immersive formats are students of history. Immersive audio builds on the successes (and failures) of previous platforms and addresses unresolved issues. We have learned that multichannel audio from speakers is impractical for most consumers, but what works for the film industry should also work for the music industry. Most importantly, we have learned that music needs to be portable and able to travel with the consumer.

Artists, labels and even mix engineers haven't come to a consensus about how to best use the center, rear, overhead, and LFE channels. Many engineers prefer not to use the center channel for lead instruments because the tonality seems strange, while other engineers fully utilize the center. Some mixers utilize the rear and side channels for effects and ambience, while others feel comfortable placing important elements away from the front channels. Many engineers avoid using the LFE, while some embrace it. One thing that can be agreed on is that mono sounds lend themselves to movement around the soundfield more effectively than stereo sounds.

The first step for creators is choosing what immersive audio format they'd like to create. It really comes down to Dolby Atmos Music or Sony 360 Reality Audio as these are the two formats supported by the big streaming services. As of June 2023, Spotify does not support any type of immersive audio or high-resolution audio, for that matter.

The key to mixing immersive music is to maintain the essence of the song while taking advantage of the expanded soundfield.

The real-life problem comes with the realization that some people will hear the mix on speakers while the vast majority will hear your mix on headphones. Headphones will play a binaural rendering of the Atmos or Sony 360 RA mix, but different streaming services utilize slightly different render engines so there are a variety of versions that will be heard across platforms and headphone brands.

The best we can do right now is keep on top of the technology and try to target the playback systems that we expect most of our listeners will use.

The evolution

A lesson we can take from all of this is that history lives among us in the present. Immersive audio is an evolution of mono, stereo, quad, and surround; with height channels and objects to add a new dimension. Audio engineers should never stop learning new skills, which even means practicing mixing in all formats. Immersive mixing contains mono, stereo, and surround elements, so every immersive mix is a history lesson, and each mix can teach us something new. Oh, if Alan Blumlein could see us now...

Multichannel audio setup 1.0

Anyone paying attention to music or movies has surely noticed the buzzwords “immersive,” “spatial,” “3D,” or just plain “surround” attached to any and all streaming media. As content producers, we need to understand the principles of creating immersive audio, generically called “multichannel” audio. We have spent years developing the craft of stereo (and mono) mixing, perfecting our listening skills, and carefully building highly accurate environments for our monitor systems — our stereo monitor system. All of a sudden we have to expand that monitor system to twelve or more speakers! This article will help demystify the setup and calibration of a multichannel system necessary for achieving mixes that successfully translate to the rest of the world, on playback systems from headphones to smart speakers to home theaters.

This article will focus on multichannel music production, and the same principles and technologies generally apply to film, television, and even audio for games.

The most common multichannel audio formats include 5.1 and 7.1 surround, Dolby Atmos, Sony 360 Reality Audio, MPEG-H, and Auro 3D. Atmos leads the pack for music distribution and streaming, although 360 Reality Audio also makes a strong presence. We will have to wait and see how things shake out for all the competing formats, but for our purposes, we will build on the standards set by Atmos, and these principles will translate to the other formats.

Multichannel speaker setups

Stereo audio uses two speakers to present an arc of sound that reaches from one speaker to the other, while multichannel audio provides several speakers that encircle the listener at ear height, from above, and in some cases from below.

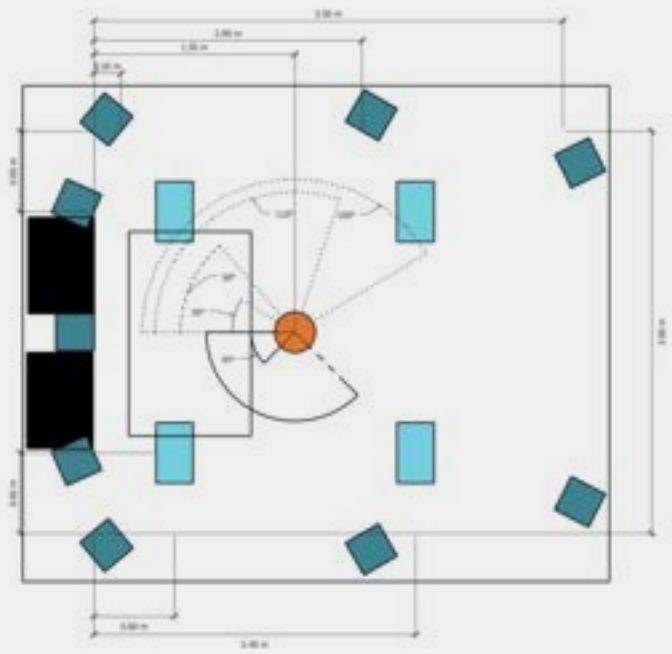
A traditional 5.1 speaker system contains five speakers at the listener's ear height along with a subwoofer (the .1) for low-frequency effects. The speakers are labeled left, center, right, left surround, right surround, and LFE. 7.1 systems expand this setup to include left and right "side" channels. In a large room or theater, there may be several side speakers to cover the entire length of a room, so any sound panned to the side doesn't have a pinpoint location, but rather a "zone" of space that it comes from. This arrangement is referred to as "channel-based" mixing. Modern multichannel setups are named like 7.1.4, which means seven surround speakers, one subwoofer, and four overhead speakers. Again, the ".1" subwoofer is dedicated to the LFE channel.

Room setup

Dolby publishes guidelines for Atmos Music studio setups. These best practices for mixing music in Atmos are thorough but fairly simple. The speakers (7.1.4 minimum, 9.1.4 or above recommended) should all be capable of reproducing full-range audio (40 Hz to 18 kHz, ±3 dB) and each speaker should be capable of generating 85 dB SPL at the listening position with >20 dB of headroom. The subwoofer for the LFE channel should be able to handle 31.5 Hz up to 150 Hz. For commercial music rooms, Dolby recommends a dedicated LFE subwoofer along with any additional subs that are needed to supplement the bass from the main speakers. When a subwoofer is used as part of the main speakers, the system is referred to as “bass-managed,” and many monitor controllers offer crossover control and routing options for bass-managed and full-range speaker systems. Most production rooms, and even home studios, should be large enough for Atmos setups, as the recommended speaker layout dimensions range from 3.5m x 3m x 2.4m (L x W x H) up to 6.4m x 5.5m x 3m (11.5’ x 9.8’ x 7.9’ up to 21’ x 18’ x 9.8’).

> Note!
> The speaker layout uses the face (baffle) of the speakers as the measurement point, not the surrounding walls or ceiling.

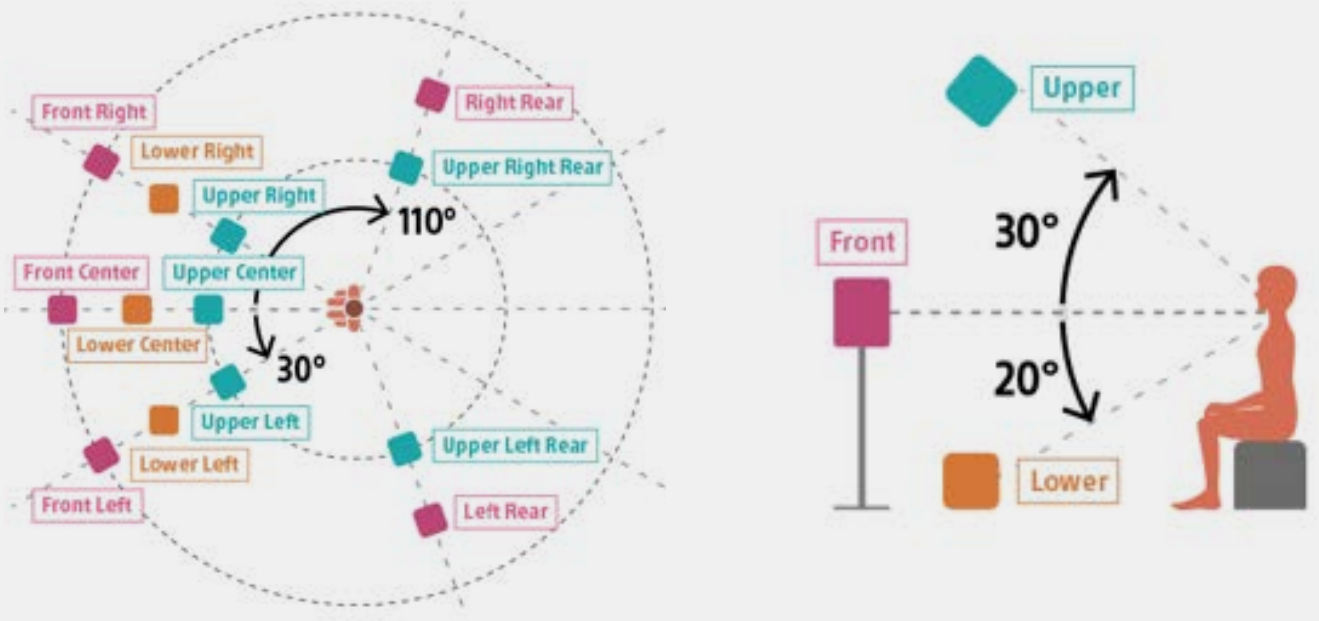
The speaker geometry and distance from the mix position can be based on equidistant (circular) or orthogonal (rectangular) speaker layouts. As always, the rooms should be set up as symmetrically as possible, but the actual distance between the listener and the front or rear speakers can vary somewhat as the monitor system will be calibrated to the specific geometry of the room. Remember that, in mixing rooms, there is typically only one best listening position, which would be the same location for stereo and multichannel setups. The position of the ear-height speakers and their angles relative to the listening position is shown in the image below. The height speakers are placed 45 degrees



Orthogonal room layout for a Dolby Atmos 9.1.4 music setup. Image from Dolby's technical guideline.

to the outside and 45 degrees forward and behind the listening position. Atmos follows the ITU-R BS.775.3 specification, which allows for slight variations in speaker placement for practical reasons, such as doors, or consoles.

The Atmos recommended speaker setup is compatible with many other surround formats. Auro 3D, however, adds a top speaker layer (above the height layer), referred to to as the "voice of God," and 360 Reality Audio adds a lower height layer. If you are designing a room for Atmos, spend some time reading the Dolby Atmos Music Room Best Practices and try out the Home Entertainment Dolby Audio Room Design Tool, a spreadsheet utility that allows you to build a specific room layout using common brands of monitors or custom configurations. The full recommendations for Atmos music studios can be found here. For comparisons of the many available multichannel and immersive audio formats, take a look at Genelec's Immersive Hub.



Top and side views of the recommended speaker setup for mixing Sony 360 Reality Audio, which requires a minimum of 13 speakers, including 3 lower-height speakers. Like Atmos, Sony 360 RA is compatible with headphones. Image from Sony Corporation

Now that we have some guidelines for the types of speakers needed and the placement of the monitors, it should be noted that the speakers must be calibrated for loudness, frequency response, and timing relative to the mix position. We also need a convenient method for simultaneously controlling the volume of all the speakers.

Monitor control

Multichannel setups for small music production rooms require monitor control for at least 12 speakers, plus perhaps headphones and alternate speakers. Larger rooms may require monitor control for up to 64 speakers. The options for monitor control range from software-only solutions to sophisticated hardware units, but even the simplest Atmos setup requires an audio interface with at least 12 outputs. If you feel comfortable or have a limited budget, you can control the speaker volume/DIM/mute from the Dolby Renderer app or even a monitor volume fader in your DAW. Cubase and Nuendo are especially well set up for multichannel monitor control. The volume fader in SoundID Reference for Multichannel could even be used as your monitor volume controller. Ginger Audio's Sphere software provides monitor control and DSP processing for up to 9.1.6 channels and also provides the ability to run AU plugins, like SoundID Reference Multichannel, on monitor sources and outputs.

Dolby Atmos Renderer application. The top right yellow box highlights the speaker control section, and the top left yellow box indicates the available speaker layout selections.

Some people are wary of software-based volume control and feel more comfortable with a hardware monitor controller. As of now, only a handful of monitor controllers are designed specifically for multichannel setups, including options from AVID, Focusrite, Genelec, Grace, JBL, Merging Technologies, NTP, and Trinnov. These brands provide speaker control plus DSP processing for calibrating the monitors. Other devices,

such as Merging Technologies, Metric Halo, and MOTU interfaces can be user-configured to control multiple speaker outputs from a single volume knob and provide varying levels of monitor calibration. Expect to pay at least a few thousand dollars for a multichannel hardware monitor controller, especially if you prefer integration with dedicated mix controllers, like a EUCON controller.

Speaker calibration

Multichannel speaker setups require precise calibration for frequency response, volume, and time alignment for each speaker. If the timing (phase) or frequency response of all the speakers do not match, sounds that play from more than one speaker could suffer from comb filtering and will sound unnatural, inaccurate, or have poor imaging at the listening position. If the loudness of each speaker is not calibrated, then your spatial mix will not translate well to other systems.

For loudness calibration, Dolby recommends that every speaker except the LFE is calibrated to generate the same level at the listening position. An SPL meter along with a pink noise generator or measurement software along with sweep tones may be used to measure each speaker's output. For large mix rooms, pink noise played at -20dBFS should generate 85 dB SPL (C-weighted, slow) from every speaker. A subwoofer used for LFE should have an output of 89 to 91.5 dB SPL for its frequency range. A subwoofer used for bass management should play at 79 – 81.5 dB SPL with its full-range speaker muted.

For smaller home studios and production rooms, you may wish to decrease all the SPL recommendations by 6 dB. Immersive mixes tend to feel comfortable at slightly louder monitor levels than stereo mixes.

To calibrate the frequency response of each speaker, a sweep tone along with software that can create a measurement plot is recommended

although some calibration is possible using pink noise and an RTA. Software like SoundID Reference Measure, Room EQ Wizard, and Smaart can be used as complete solutions to measure the loudness, frequency response, and timing of every speaker. SoundID, for example, measures each speaker from several locations around the mix position and generates a multichannel correction profile that is then imported into the SoundID plugin. Both SoundID and Room EQ Wizard allow you to port their measurement results to certain monitor controllers with built-in DSP. A selection of hardware monitor controllers, including those from Genelec JBL, MiniDSP, and Trinnov, provide some amount of built-in measurement and room correction tools.

Playing and capturing impulses or short clicks is necessary to calibrate the timing of each speaker. This process could be performed manually by playing a click from a DAW and recording the click with a microphone located at the listening position. This must be done individually for each speaker and then the delay times can be measured in the DAW. These delays need to be somehow applied to each speaker output, probably through the monitor controller's DSP or via the Dolby Atmos Renderer or Mastering Suite software. SoundID Reference measures speaker timing (actually phase response) and includes timing correction as part of the multichannel calibration profile it generates. Some hardware monitor controllers can perform time alignment as part of their room correction DSP.

While the Dolby software and a few hardware monitor controllers provide some room correction capabilities, SoundID Reference for Multichannel measures loudness, frequency response, and timing for systems from stereo up to 9.1.6. The measurements cover a comforta-

ble sweet spot around the listening position and a calibration profile is generated. This calibration profile can be saved and imported into the SoundID DAW plugin. or it can be ported from SoundID to several hardware monitor controllers, including DaDman interfaces with the SPQ card, AVID's MTRX Studio, Merging Technology interfaces, Wayne Jones studio monitors, and the Dolby Atmos Renderer Application. So there are several calibration options available, depending on how much money you wish to spend and what type of monitor control integrates with your particular workflow.

Acoustic treatment for multichannel rooms

Electronic room correction is necessary, but should only be performed after room acoustics have been optimized with proper treatments. We should all be familiar with acoustic treatments for stereo production rooms, but multichannel rooms present a slightly more complex challenge because sound comes from many places in the rooms and more than two speakers (often facing each other) can emit the same sounds. Time alignment, loudness calibration, and level matching go a long way toward dialing in an accurate monitor system, but let's take a look at some of the primary acoustic treatment considerations.

A major concern for speakers placed near walls or the ceiling is speaker boundary interference (SBIR).

SBIR occurs when the low frequencies from a speaker reflect off the nearest wall/ceiling and create a cancellation of a specific and narrow range of low frequencies emitted from the front of the cabinet.

With stereo monitors, SBIR can be mitigated in one of three ways: (1) soffit mounting speakers, (2) placing the speakers close to the front wall while treating the front wall to absorb low frequencies down to about 150Hz, (3) placing the speakers 2 meters or more from any wall to lower the SBIR frequency to a very low frequency. One or more of these solutions can and should be applied, even in a 3.5m x 3m x 2.4m room. As an example of SBIR, speakers mounted very close to the walls and ceiling will require at least 10cm to 15cm (4” – 6”) of acoustic treatment behind them to absorb a reasonable amount of energy down to 125Hz. Without acoustic treatment, a speaker with its face 60cm (24”) from a wall or ceiling will create a problematic dip at just about 140Hz. Moving that speaker so its face is only 30cm (12”) from the wall raises the frequency cancellation to 280Hz — a much easier problem to treat. In small rooms, it’s therefore recommended to treat the walls with 10cm (4”) of absorption and place the speakers as close to the wall as is practical.

Immersive speakers are often mounted directly to the walls or ceiling, and their bass output may also be exaggerated due to acoustic space loading. This phenomenon can effectively be cured by adjusting the frequency response of the speakers with EQ. Space loading is one of the few frequency response problems of a speaker in a room that can be easily fixed with a simple shelving EQ and many powered speakers feature a 3-position switch to correct the bass output for various placements.

Time alignment should ensure that sounds that emanate from more than one speaker arrive in phase at the listening position. Further, it is recommended to utilize the same family of speakers for all the speakers in the system. For the most cohesive sonic impression, the left, center, and right speakers should be identical, as should all the surrounds, as should all the overheads. SoundID Reference will properly align even mismatched speakers, but using the same model for all the speakers will produce the most cohesive soundfield.

Photo credit: SoundShepherds.Studio

In immersive playback rooms, just like in stereo, the decay (T20) should be even across all frequencies and generally between 150ms and 250ms. Broadband absorption should be evenly distributed around the room for multichannel systems since speakers face all directions. Low-frequency absorbers may be necessary behind each speaker location to minimize SBIR cancellations. Diffusers are unnecessary in all but the largest multichannel rooms as there are sound sources all around the

room and the soundfield should only be that of the playback system. A final consideration for multichannel rooms would be to carpet the floor as much as is practical, since the overhead speakers fire at the floor, and side speakers may also have floor-bounce to consider.

Subwoofer placement for stereo setups often requires some trial and error to find practical locations that constructively match the sub with the main speakers without creating destructive standing waves. With multichannel setups, we may need one or two subs to supplement our horizontal plane of speakers and another sub (or two) to supplement the overhead speakers.

Due to the practicalities of mounting overhead speakers, these speakers probably produce the most band-limited audio, and a subwoofer is often recommended to supplement their output. This sub will most likely be located at least 2.5 meters from the ceiling and at a different distance from each of the four overhead speakers. How do we align one or two subs to play well with the overheads? Subs that crossover at 80Hz or lower probably can’t be localized, but they may still have noticeably different phase interactions with the different overhead speakers. My recommendation is to avoid using subs associated with overhead speakers and instead find a way to utilize full-range overheads.

As with stereo mixing and mastering environments, the key to properly understanding a multichannel mix and ensuring translation to other playback systems is to treat your mix room to have an even frequency response and similar decay times for all frequency ranges. Treating bass modes may be even more critical in multichannel rooms as more speakers are placed against walls and near corners, where modal excitement is sure to occur.

Part of the excitement of multichannel mixes is that they produce greater dynamics, an enhanced sense of transients, and more apparent bass presence due to the number of full-range speakers in the room. However, since multichannel rooms provide such a wide soundstage, dynamic range, and clarity, it is important to proof a mix on headphones.

A large majority of listeners will undoubtedly hear the mix via a streaming service on headphones, so it behooves us to verify that our mix translates well to headphone playback both with and without binaural rendering.

Flat sound and target curves

Room-to-room translation of mixes is always a concern and Dolby has pretty much nailed this subject in regards to cinema playback. Mixing stages for films for theatrical release need to be calibrated with the “X-Curve” target curve. This X-Curve was developed to take into account the size and acoustics of movie theaters and the distances that listeners sit from the speakers. For stereo music, mixers and mastering engineers tend to prefer a neutral or flat frequency response that translates well to typical living room home stereo systems. The Dolby Atmos Music Studio Best Practices document recommends the “Dolby Atmos Music Target Curve” for commercial music mixed in Atmos. This target curve takes into account nearfield mixing environments and how the mixes will translate to typical home stereo environments.

Dolby recommends strict adherence to the target curve, although many mixers feel the Dolby Atmos Music curve is not significantly different from their customary target curve. SoundID Reference provides a target curve that has been verified in professional Atmos mix rooms to deliver the Dolby Atmos Music curve. Sony 360, on the other hand, suggests a flat frequency response for music mixing. Because of all of these standards, SoundID Reference allows the user to quickly switch their monitor's target curve between Dolby's target and a flat target.

The Dolby Atmos Music target curve compared to the flat target displayed in SoundID Reference. The user can easily and quickly toggle between a flat target and the Dolby target My with one mouse click.

Dolby recommends averaging room measurements around the listening area and SoundID performs 37 spatially averaged measurements to arrive at its calibration profile. SoundID Reference for Multichannel allows you to easily switch between a flat target curve, the Dolby Atmos music curve, or any other custom curve you wish to use.

If you port a SoundID correction curve to a hardware device, like the AVID MTRX Studio, you can choose to port whichever curve you prefer.

Conclusion*

Setting up a multichannel audio system may seem a daunting task, but by taking into account proper acoustic treatment, speaker alignment, and calibration you can achieve an exceptional listening environment that will translate well to other multichannel playback systems. Multichannel music for consumers, even more than stereo, will likely be listened to on a system very different from the system in a typical mixing environment. It is therefore important to develop an accurate and consistent mixing environment and, even then, to assess your mixes outside of your studio on many different systems. This process will provide important feedback on producing mixes that sound great in a calibrated room and also translate to consumer playback systems, streaming platforms, and typical home environments.

By Adam Kagan

Photo credit: SoundShepherds.Studio

What is immersive audio and how does it work?

The hottest topic in audio right now is immersive and spatial audio. Just about every audio trade magazine, trade show, and conference has highlighted presentations using the word “immersive.”

The Spatial Audio, Atmos and immersive categories on streaming services like Apple Music, Tidal, Netflix, and Amazon HD have piqued listeners' curiosity with the promise of a novel listening experience for new songs and well as old favorites. Luxury automotive companies offer immersive audio systems that give audiophiles goosebumps. Record companies, streaming services, equipment manufacturers, and, of course, artists and content creators are all excited about this new world. You might expect such a shiny new thing like immersive audio to be incredibly complicated, but it really isn't as difficult to grasp as most imagine. Undoubtedly, there are fresh elements to comprehend when delving into the immersive realm, which may require some time to fully digest, but adept engineers are consistently ready to learn new technology and rapidly adapt to changes in the market. As always, engineers who embrace immersive audio early will be the first to reap the rewards.

Why is now a good time to take the leap into immersive?

Major labels like Capitol (UMG) began in 2019 to remix or adapt their entire stereo catalog into immersive formats—financial evidence that labels take immersive seriously. Many popular artists are releasing immersive mixes, giving their fans a new way to enjoy their art. There are technical standards in place to help you build an immersive system with accuracy, but the creative rules haven't yet been written. This early stage of evolution affords an opportune moment for creators, producers, and engineers to learn about, and experiment with immersive audio.

With this article, engineers who are familiar with how to mix in stereo and are curious about the world of immersive audio will gain a better understanding of immersive and spatial audio, and how to conceptualize the workflow. Readers are encouraged to think of this as a prep/pep-talk for a larger conversation about how to integrate immersive and spatial audio into professional practice. It's not as complicated as it seems, and it's a lot of fun!

By Adam Kagan

What is immersive audio?

The goal of immersive audio is to create an experience that envelops the listener from every angle in three dimensions. This may be to recreate the experience of an event such as a basketball game, or a concert in an orchestra hall. It may be used in a movie to draw the viewer into a scene with ambient cues from above and behind. It can be used to imagine new sonic landscapes where instruments seem to defy the laws of physics or present fantastic reimaginations of music that was made a long time ago. To envelop the listener, whether it be with precise realism or whimsical fantasy, one must first understand how a listener perceives the sounds around them.

Our hearing apparatus is already immersive. With our eyes closed we understand sound cues that arrive from any angle around our heads. We take advantage of this ability when we create immersive listening environments. Sounds that come out of speakers behind and above us sound to our brains as if they are behind or above us because they are actually behind or above us. There is no special computer processing needed for our brains to be able to decipher directionality.

We can perform this trick of locating sounds in space, even with our eyes closed, because our two ears and brain work together as a team. Our vision is binocular, which means we have one dominant eye and a second eye that is compared to the first. This comparative math is what we use to generate our sense of depth. Want a quick demonstration? Have a friend toss you a ball from six feet away and catch it with one hand. Easy. Now, repeat the experiment, but cover one of your eyes. With one eye closed, it is harder to discern distance and you'll notice that it's difficult to gauge where the ball is in space as it comes towards you. We've modified your binocular vision into monocular vision. Fun!

The brain/ear team works in a similar fashion. When sound arrives at both of (left and right), the brain calculates time and amplitude differences between the two ears to understand where the sound source is located. A sound directly in front of you will arrive at both ears at the same time and amplitude. As the sound moves to the left or right, sound will arrive at one ear slightly sooner and louder. The brain does quick comparative math to understand the exact difference in time and amplitude and translates that data into an approximation of the location of the source without the help of the eyes. This works for vertical sounds as well. We call this system of hearing "binaural" (bi = two, aural = ear).

So, what is “immersive” audio, and how does it work?

Sometimes referred to as “ambisonics,” immersive audio is a 3-D sound field created by a combination of lateral and overhead speakers. All immersive audio systems are by definition multi-channel and must include speakers from above (height channels). There are several speakers pointing at you from locations around you at ear level, and additional speakers pointing at you from around you but above your head. In immersive, all speakers are focused directly at one fixed point called the “listening position”. The exact placement of these speakers is directed by the immersive format you choose. Popular formats include Dolby Atmos, Sony 360 Reality Audio, and Auro 3D.

Each format has strict standards for the number of speakers and where the speakers should be placed in a room for optimized listening or content creation.

All immersive formats use dedicated software to define the 3D environment as X,Y,Z coordinates. When the software is set up in a particular room, the user defines how many speakers are connected to the system, and their relative position and orientation in the room. Once calibrated, the software becomes a bridge between the virtual (DAW panning) and the physical listening space. The mix engineer can use a standard DAW, along with the software to place sound elements in 3-D space. The software determines what combination of speakers in the room will best recreate a sound’s position in the physical listening environment.

The audio channels contain metadata that tell any playback system where to place sounds in that particular room, so the system is scalable from small rooms up to large theaters.

> The software is referred to as a “renderer” that takes raw data like panning information and renders it into a sonic experience for the listener.

The renderer also generates the master audio file which contains all of the audio and metadata necessary for a consumer to hear the immersive mix the way the mix engineer intended, regardless of whether they are listening on a full speaker system, a soundbar, or headphones.

> This concept is quite different from a traditional stereo workflow. When performing a mix in stereo, sounds are placed between two speakers; “left channel” and “right channel”.

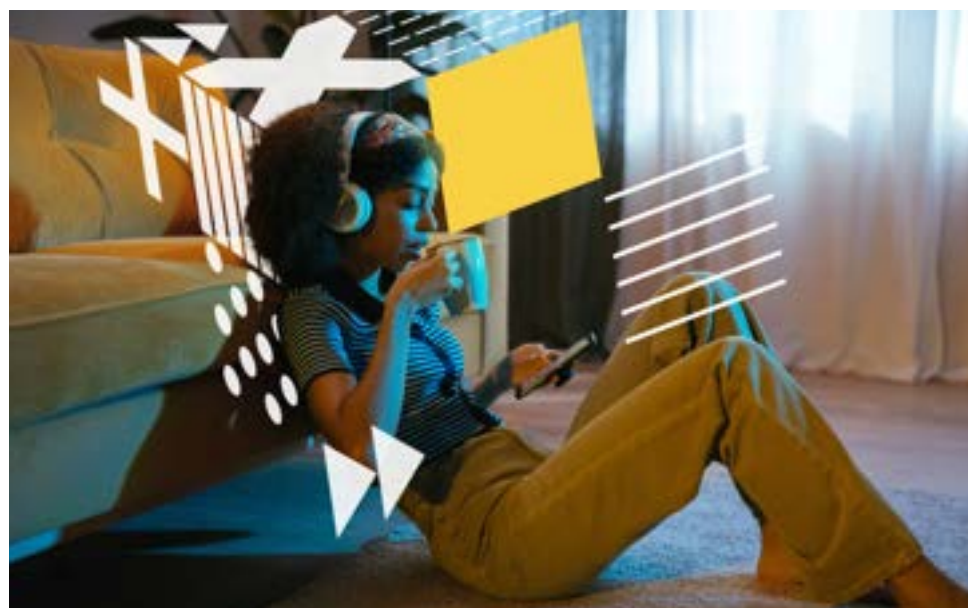
A panner is used to send each track to the left, the right, or somewhere in between. A stereo mix only requires two channels to properly store and replay the mix. When mixing in immersive software, the user operates a 3-D panner that represents the lateral (x, side-to-side) position, the depth position (y, font-to-back), and the vertical position (z, up-down) of the sound. The software renderer then computes the most accurate speaker or combination of speakers necessary to play the sound in the room. The renderer encodes up to 128 audio channels, which can be individual sounds or subgroups, with information that defines its X-Y-Z coordinates. An immersive master for a three-minute song may be a few gigabytes, as it contains up to 128 channels of audio and metadata.

There are a number of ways that consumers can experience an immersive audio file, but it always requires a receiving device (phone, computer, streaming hardware, or AVR) equipped with a decoder that matches the immersive format being streamed.

When a consumer plays an immersive file, a re-rendering app analyzes the metadata in the master file and manages the playback experience. The consumer's receiver is aware of the total number of speakers connected to the system and their relative placement. The re-renderer routes the sound to a position that matches the original coordinates. The re-renderer is flexible enough to generate a faithful listening experience even if the listener has fewer or more speakers than the mix studio had.

The ability of the renderer to adapt to its host environment is remarkable. Smart speakers and sound bars equipped for immersive playback can recreate a version of an immersive mix without the operator having any knowledge of speaker placement or calibration. Consumer devices such as Amazon's Echo Studio support Dolby ATMOS or Sony 360 RA and automatically calibrate for an immersive playback experience. Consumers will be using these devices to listen to your mixes, so you might consider using one as a mix reference in your studio.

How does immersive audio translate to headphones?



The short answer is that it doesn’t quite translate, but you can still experience a version of immersive audio on regular headphones with either binaural or Spatial audio.

Binaural encoding is the method that Dolby and Sony use to create a convincing 3-D space in stereo headphones that mimics the soundfield of an Atmos or Sony 360 mix, though with only two speakers.

Spatial Audio is Apple’s proprietary conversion of Atmos into headphones that changes some of the binaural information that is encoded in the Atmos master file.

While Apple has trademarked Spatial Audio as their headphone playback codec for immersive sound, you may also see the word spatial to generically represent immersive audio over headphones, which may be one of many re-renders of the immersive master. All this means that, as creators, we don’t know exactly how our audience will hear our music!

Conclusion*

Immersive is an exciting and fresh way to experience sound. It may seem complicated, but human evolution has already achieved the miraculous part. Thanks to our binaural perception, we enjoy a rich world full of engaging soundscapes. When you work in immersive audio, you can recreate these events in life-like detail, generate new experiences that spark the imagination, and bring new opportunities for artists and content creators. It’s a very rewarding format. Learning a new workflow is fun, and now is a great time to get in on the ground floor of this nascent technology. As you mull all of this over, I encourage those who haven’t done so to seek out a facility equipped with an immersive playback system and ask to take a test drive. Dive into spatial audio on your favorite streaming app and find out what’s out there.

“Believe what you see with your eyes, trust what you hear with your ears, know what you feel with your flesh”

— Brian Staveley

By Katrina Allikas

Immersive audio and mixing in Atmos with Richard Furch

Martins Popelis

Richard Furch

Our CPO and co-founder Martins sat down with 6-times Grammy-awarded mixing engineer Richard Furch to discuss music, technology, and mixing in stereo and Atmos. In his 20-year career, Richard has worked with Prince, Frank Ocean, Jay-Z, Rick Ross, Snoop Dogg, The Game, and The Weeknd to mention a few.

Martins Popelis: How did you get into music, what brought you into it, and how did your career develop from there?

Richard Furch: I'm a jazz pianist basically by trade. I was pretty good as a pianist, but I learned also that unless you're Herbie Hancock, being a jazz pianist is quite one travel down the road. Along that line, I decided my side is on the other side of the glass. I went to SAE and Berkeley for audio engineering and music, and then ended up in New York. I got a job as a runner, and I made it through the ranks, and finally here in LA. I've been mixing, and engineering a lot of R&B and hip-hop that you heard of, and keep going on the daily.

MP: In your career, what are some of the highlights, some of the artists, and cool projects that you've worked with that stand out in your memories?

RF: I have to start with the coolest, which I was Prince's engineer for about a year. We did a couple of albums together. I also worked with Frank Ocean, Jimmy Jam and Terry Lewis, and Tobi Lou. Like I said, a bunch of great R&B: Usher, Chaka Khan, some hip-hop, Jay-Z, Outkast, that kind of stuff. It came from a start at the New York Studios, the big studios. Shout out to Sound On Sound Studios. That's where it started to blossom for me. Once you have a hit in one style, more of that work appears. I've been fortunate with a very, very good range of artists and projects.

There's an unspoken rule in the music industry: Success chooses you.

MP: Do you remember your first studio?

RF: That was in my bedroom in Berlin. I used to have the E-MU E64 samplers, and we were on cakewalk for sequencers at the time. Then many iterations of that in Boston and New York followed. My first studio where I was like, this is actually a real room, was probably here in LA. Now, we're in this version. The Mix House is probably Version 4 or so. This is now a full-blown professional facility.

MP: This is not only stereo. This is Atmos, right?

RF: It is Atmos. It has been always 5.1 as well before Atmos became a bigger part. It's full-blown. There are five stereo systems here, a big system, a Phantom Focus System made by Carl Tatz, and then a couple of other smaller systems, and now the Atmos experience.

MP: Awesome. You've been working in a lot of rooms and had the opportunity to work with a lot of great talent. All the rooms sound different as far as I've seen them. Can you work in any room or do you prefer to work in rooms that have accurate sound?

RF: You start as an engineer, leaving the nest of your recording studio, which, for me, was Sound On Sound Studios at the time. You start going around and try to give yourself a leg up. In my case, that was bringing these speakers around, the ProAc Studio 100s. I think they're fantastic speakers and I brought them around to every session. The interesting part is that really fast you realize that even though you're bringing your own speakers, none of it sounds the same. It's like your best chance, but still, the low end is fully different. Just the way they fit together with whatever studio you're in is not the same at all.

Over time, I realized I can't really, really rely on that. Mostly, I relied on headphones like Beyer DT 770s, et cetera, just to double-check stuff. Over time, the only real answer was to build my own room. In this room, this version of the Mix House, I've been here for 12 years now. I know this room inside and out, so every record that comes out here doesn't need to be car checked. If you know you have a compromised room, a car check might be a great thing for you, but once you get to a room that you actually trust, a car check is not really that valuable anymore. After all this time in this room and the way I know how this room translates. The studio becomes the instrument, and I don't really want to work anywhere else for mixing purposes. Now, if I need to record somewhere else, I'll travel to all the great rooms in town. That's fun, but for the final mix, I'll end up here mostly.

MP: That makes sense. When you were working in the other rooms and you were learning that they all sound different. How did you deal with that?

RF: Cry into my pillow! No, it's hard. I want to say that translation is the hardest thing to understand. The fact that you could work on something for eight hours, and then you go into another environment and it sounds totally different, is heartbreaking. Especially at the beginning of your career when you're trying to figure it all out. Is it me? Is it because I'm not good, or is it because I don't hear it enough? That's a very confusing situation to be in.

MP: I can understand that. Let's talk about Sonarworks in your room. You have this great room. You're used to it. Now, you've had some experience with our latest multi-channel version of the software. How would you describe your experience so far?

RF: In my situation, I had to combine a new Atmos system with a very, very dialed-in stereo system that I already had. I knew I couldn't start over. I also knew I couldn't really use the same speakers, partly because it would be insanely expensive, and partly because my speakers are so tuned and so dialed-in that they translate fantastically. Whatever I send out to mastering comes back untouched. I really trust these speakers.

I decided, okay, to get my feet wet – I already had a 5.1 system with extra JBL speakers. These are the 308s for that matter. I think Chuck Ainlay uses the same ones in Nashville. I decided, let me just buy more of those, so basically add to the 5.1 system with more speakers of the same kind. At least we're moving forward with them and we're not starting over.

We did a bunch of stuff with time aligning, measuring and recording clicks, shifting speakers, and putting little delays in. It worked pretty well for that matter. Then you told me about the multichannel version of the Sonarworks SoundID Reference plugin, I was like, "Let's try that."

We were shooting for something that I already know what it should sound like because I'm happy with the stereo side. I was like, we need to clone these speakers 11 times or something. I was really interested in getting the surround speakers and the height speakers to be at the same listening precision. They should be very similar in overall frequency response. That's exactly what the [SoundID Reference Multichannel] software did. All of a sudden, everything came into focus, so to speak. The good things about the left and right stayed the way they were. Actually, they got a little better, because now you have a little bit more correlation between left and right. Then the other speakers joined the pack and became focused very much without overlapping frequencies or masking. Everything became a little cleaner, tighter, and more accurate. The overall tuning was still what I knew so well from this studio.

MP: Let's talk about Atmos a little bit. How did you first get into it? How do you feel about it? Where do you think it's going?

RF: Actually, it was by accident, because at the very beginning, a friend of mine said, "You need to have an Atmos room." I'm like, "Okay, let's do it." Then the price tag was 60 grand, and then please open your ceilings, which is all – sorry, they're all covered with fabric, etc. I'm like, no, thank you—not doing that. Then, over time, Dolby loosened their restrictions and basically said you can do this on one computer and you can actually use any speakers you want. Okay, so now we can explore it. Then all of a sudden, a couple of projects came in. They're like, "Can you do the Atmos thing?" I'm like, "No. Crap." All of a sudden, you felt like there was an actual need. That's when I was like, okay, let's get this happening.

I would say the whole transition was from I can't do Atmos to, okay, we got this – it was about six weeks to two months. Then just like every new technology, it's a huge learning curve of what are we going to do creatively. Even though I am very, very familiar with the 5.1 surround system that I had, it's quite a dimension, literally, in height. It's quite different to handle.

MP: Have you noticed any new creative ideas or dimensions come up because of that ability to mix in space rather than just left-right?

RF: Yes, I would say maybe the number one difference or change or advantage is – part of stereo mixing is making sure that everything that's in the record actually fits between the two speakers, which obviously requires a bunch of techniques and care for that matter. In an Atmos environment, the beautiful part is I can put these layers into a different 3D environment and just have more space for it all, that job becomes easier. Now, once that job becomes easy, it also means now we have space to add other things. What I like, for instance, in EDM, you come into the drop, and there might be something like a snare roll, and it ends up in a cymbal in a little explosion. Let's make that a big explosion! What is a cool move is to have it come from the back and then the cymbal comes in the front

maybe. Then the tail of that splashes over the whole field back into the back. These are cool movements that you can do that add to the actual instrumentation that is already there. That's how creativity can get a little bit more elevated there.

It was very clear that everybody really wants the stereo mix, just bigger— it should be Atmos, but please don't change it too much. If you follow my masterclasses, how I describe my mixing style is I take your rough mix, which sounds like this, and I explode it and make it wider and bigger and hopefully clearer and more impactful. All the good things that you're trying to do in mixing. Then I realized to myself, that's exactly what Atmos is. Now we're just doing it to the stereo mix and making it even bigger. All of a sudden, I felt, oh, there's a purpose. I've been doing that for a very, very long time, but now we're doing it on more and more speakers. All of a sudden, creatively, it was very, very exciting.

MP: Sounds very inspiring. Now you have mixed in Atmos for quite some time. Are there any Atmos-specific mixing tips that you have discovered that you think might help somebody who is just starting in Atmos?

RF: One of the interesting parts that you have to be very careful about is – I mentioned earlier we're looking for a bigger, more expansive version of this stereo mix, at least in many cases. These might be delivered as stereo stems from the original mix. Now wouldn't it be beautiful if you just align those stems and pan them into space and everything would be done? You just print and you're good. Unfortunately, it's not that easy.

What you will figure out pretty quickly is even if you make no changes to the actual stems, the actual resulting Atmos mix will sound quite different than your original stereo mix. Basically, in order to do quality control, you have to figure out how far I have to change the stems in order for them to appear as if we were never there. It's a very long process to figure out how to do that. It is really hard to take a stereo stem of any kind, so let's say drums, and just put it into

an Atmos mix and for it to come out the same. Every single time you change a binaural setting, it changes where it appears in the Atmos system.

I notice very often that the vocals, especially if they come in on a stereo stem, always seem to be a little softer in the Atmos mix if I just don't do anything. I tend to boost them a little bit. I tend to put some more ambiance on it into the high channels. I might even put the dry part more into the height and also more into the center – a little bit of the center channel. Obviously, that adds a little bit of level to them officially, if they come out of different channels.

Also, even if you set the binaural setting to off, it is also not quite the same where it lands. Learning how that happens, learning that the important parts of the record, vocals, drums have to be in the same place where they were before so that you have then the opportunity to take, let's say, guitars or choirs or pads and move them around a little bit more. Always double-checking. I'll print one, listen to it on the Apple headphones, etc, because that can be only listened to offline, not in real-time, unfortunately. That is the process. That is the learning curve. That's why it's a little bit more complicated than just a stereo mix.

MP: How do you feel about where Atmos is right now? Do you feel it's here to stay?

RF: I think, first of all, on the positive side, it's growing. I'm starting to hear quality and ideas that actually goes far beyond the stereo mix. All of a sudden there are records where I'll go back between stereo and Atmos, and I'm like, "The Atmos is definitely better. I don't know exactly why or what's going on, but it's actually better."

We're all practicing. All the engineers I talk to, we're all going like, "Have you heard this?" or, "How did you do that?" or, "I did this mix yesterday, and it did this really weird thing. Are you doing the same thing?" We're trying to talk. We're all learning and trying to make it better. On the positive side, that is beautiful.

MP: Studios are moving to 9.1.6 and people are moving to mono.

RF: On the more complicated side, of course, most people don't even have a stereo speaker anymore. We can record a 192 kHz and everybody listens to mp3s. Those are funny things, but it's true. I have 11 speakers here, and I'm just going to tell you that a normal mortal person that is just a music fan has absolutely zero chance to set up an array like this and to really experience an actual Atmos mix in speakers, where I think it is the most powerful. When you sit a client down in the middle of the speakers – and these clients are like, "Why are we doing this?". You just press play and they're like, "Okay, I get it now. That's cool." On headphones, that experience is much more limited. It's not bad or anything.

MP: With seeing all the big-weight tech companies getting behind this – as you said, six months ago it wasn't sounding maybe as good as it does now. Six months from now, it's probably going to be even better.

RF: I really hope so. Obviously, I have high hopes for that. The interesting part is that the codecs that are being delivered on phones, etc, are slightly different from the Dolby codec that we have

here. Even if I pan something to the back, and it comes clearly out of my back surround speakers, you cannot hear that on headphones. You hear the sound changing a little bit. There is a binaural component. That is true, but you can't really say, hey, I heard that coming from the back. It doesn't really work like that yet. It's a hard task to make that all translate again, but I think we're moving in the right direction there.

I know they're working on it with the HRTF profiles, with the Sony 360s, but we're still probably a generation of hardware, like specific headphones, away from that being really impactful. That doesn't mean though that we should stop. It just means we should get better at it.

There's a beauty in listening to stuff in surround, and the beauty of Atmos is that it is scalable from a headphone system to a whatever, 11-point-something system, while the other systems, 5.1, 7.1, etc, were pretty much fixed. You either have that kind of system at home or you didn't. That's it. Considering we have a format now that could live in all these places, I have high hopes that it will propagate better and that the music listener, the music fan at one point goes, "You know what? Music sounds better than it did in 2015." I'm not knocking 2015, but our job is to make stuff better over time. That's what I hope will happen.

MP: What's the best Atmos mix you have heard so far?

RF: That's a very good question. Musically speaking, there's some really great stuff on the Olivia Rodrigo albums, or album, because there's only one – but on the songs. Then like I just said, the first one that actually really impressed me quality-wise was probably the Harry Styles record. I think that's one where I like the Atmos mix better than the stereo mix. There's something, not just the creative part of what they did with it, but also the clarity. I think somehow they solved something there. It really translated really well. I was impressed. That one I actually heard first on the AirPods Pros, not the Maxes.

MP: For somebody who is just starting in music, what would be your advice in terms of how they should think about their career to be successful or what they should or shouldn't do? From your experience, what do you think is the key?

RF: An engineer in 2022, I think you have a lot of really exciting opportunities in the way that, for instance, most software that you need to do your job as I do is certainly very cheap considering if you compare it to the time when I came up. The great news is you can literally – with a laptop, with the software that's available, and with time, you literally have the same setup that I have. You can practice things that were hard to do at the time. You had to sneak into studios to get studio time. Now, you can actually practice at home, so it's all about connections. What a beautiful time to get started. As long as you're patient and realize that this won't happen in two months, but it could happen, very, very likely in two years or five years, then that's a decent outlook, I think. What more chance do you want?

Just try to be positive, like, how do I add to the community around me? How do I help a musician be a little bit more successful? How do I help another engineer who could be your competitor? Once you move that thing out of your mind and go – how about that could be your best friend, and maybe you could do this together—I think that's a better way to insert yourself into this music scene. You're going to be the most loved and valuable part of this community. That will come back in spades.

I was lucky. I walked in on a Jay-Z session, and all of a sudden I was part of one of the most successful hip-hop albums of all time. That's lucky. It was the preparation meets opportunity. I worked really hard on getting a job and getting that particular gig meant being there, in the studio. These kinds of things are important, and they're a little harder right now because certainly, you're not going to make all of that happen in your bedroom. Jay-Z or Gunna is not going to walk into your bedroom tomorrow. That is pretty sure.

MP: Thank you. I've been enjoying this conversation on my part and hope you had fun as well.

RF: Yes, absolutely. You are helping me make better records. That's what it's all about.

Sonarworks tech is loved and used by 140,000+ studios & 55+ Grammy-Awarded-Engineers worldwide

"SoundID Reference for Multichannel is a game changer! It gives me added confidence in knowing that what I'm hearing out of my monitors is correct, which allows me to spend more time on the creative process and finishing my mixes faster."

— Aaron Mattes

Grammy Nominated and RIAA Platinum Certified Mixing Engineer - Dolby Atmos

NLE Choppa, 50 Cent, Young Thug, Big Sean, Gucci Mane, Lil Baby, Polo G, Flo Rida, Macklemore, Cheif Keef, Travis Baker, Lotto, Moneybagg Yo, Ro James

“Sonarworks software provided the missing puzzle piece I was searching for: How to rein in my entire ATMOS speaker set and make them sound cohesive and uniform. The combination of measurements and target curves provided focus and clarity and really tied the room together. Translation is excellent.”

— Richard Furch

6 Grammy Winning Albums out of 23 Nominations Diamond/Multi-Platinum Selling, Engineer, Mixer, Pianist

Prince, Jam and Lewis, Tyrese, Frank Ocean, Usher, India.Arie, OutKast, 林俊傑 JJ Lin 鄧紫棋 G.E.M. 蔡健雅 Tanya Chua 蔡依林 Jolin Tsai, Wang Leehom, Boyz II Men, Macy Grey, Jay-Z

Try Out SoundID Reference 21-day free

→ No credit card required, no hidden conditions.
→ A quick and guided setup.
→ Try out on any of supported headphones and speakers for up to 9.1.6. setups.

Learn More →

SoundID Reference