# Spatial Composition – and What It Means for Immersive Audio Production

Julius Dobos

*The Creative Shop, University of Silicon Valley; San Jose, California*

Correspondence and requests for research data should be addressed to Julius Dobos (contact@thecreativeshop.com)
Preprint first available: November, 2024.
This revision of the preprint: September, 2025.

**ABSTRACT**

Immersive audio is an important concept in our listening experience as we enjoy films, games and other audio content. In musical works, while the concept of multichannel formats is not new, spatial audio technologies have been far underutilized in the inception of content; technologies such as ambisonics, wave field synthesis, object-based audio are, with an overwhelmingly wide margin, used in the recording, mixing and mastering stages of music production today. These spatial audio technologies are usually applied during music production; after the music creation but before the exhibition (performance or playback) phases, with spatial decisions typically made during the mixing process, often in absence of a spatial-musical concept. Here we propose the consideration of space as a compositional, rather than production element of music. We examine how its use in the earliest phase of conceptual design has a substantial effect on the composition itself.

Focusing on the creative possibilities offered by spatial audio design, what if decisions such as an instrument's location or its movement in space were not made during audio production by engineers, but rather, they were design choices made by composers prior to production, during the composition phase? Composers utilize melody, harmony, rhythm (and often timbre) to express a musical idea and evoke emotions – what if such ideas could be also shaped or even inspired by spatial location, area, position or movement? Would such space-derived decisions about notes, timbres and the roles of musical parts be the same if they were made at the time of composition, rather than during audio production? And, even more interestingly, would making spatial decisions during the composition process have a bearing on the musical message itself?

## 1 Spatial Music: Introduction

The concept of immersive audio is derived from our biology – the awareness and localization of sound sources in the environment around us in a 360-degree sphere. Edgard Varèse spoke of spatial location and movement ("sound projection") as the 4th dimension of music as early as 1936, and applied spatial design as a conceptual element in his work [1][2]. Using a custom playback system of over 350 speakers, the 1958 premiere of his *Poème Électronique* was a major milestone in the history of spatial music performance. Due to the lack of immersive audio technology, the performance utilized only three mono tape tracks [3]. Spatial performances did not catch on during the 1960's rock 'n' roll and Motown era – until in 1967 Pink Floyd started using their azimuth coordinator (a quad joystick controller), performing several quadrophonic shows [4]. Although the 1970's quadraphonic movement did not bear commercial success, the film industry's desire to evolve stereo playback systems into surround sound in cinemas, which would imitate our natural perception of sound, steadily lead to the development of the immersive audio systems we know today. [5]

Following the 1976 introduction of Dolby's analog 4-channel Dolby Stereo MP format (which featured a LRCS channel configuration), the 1980s brought about new concepts and inventions in immersive film sound. The playback systems necessary for delivering early immersive audio experiences for the public were first installed in movie theatres. The 4-channel Dolby Surround Pro Logic format was made available for home audio systems – still, with film audio playback in mind. Technically, these encoding-decoding formats were adequate for the reproduction of music content in four discreet channels (LRCS) [6] [7]. Yet, the recording industry did not see such developments as an opportunity to invest in and steer consumers to – and, without much standalone musical content produced for such systems, the music industry remained married to the stereo format.

The 1990s brought us digital surround sound for film, utilizing 5.1 discreet channels (Dolby Digital and DTS). Audiences experienced the benefits of digitally encoded surround sound in cinemas, including improved spatial specifications and a low-frequency effects channel. Despite the growing public awareness of the surround audio formats, still overwhelmingly used for film sound, surround music did not pique the interest of the general audience. The convenience of 2-channel systems, familiarity with the stereo format, the constant evolution of surround systems toward increasingly higher numbers of

surround channels and the higher price point of surround setups discouraged broad interest and investment in surround playback systems. In turn, the recording industry failed to develop a serious interest in releasing, let alone producing new, surround content for such narrow market.
By 1997, surround encoding plugins were available for major digital audio workstations, thus bringing multi-channel release capacity to music recording studios and artists. Starting with the first DTS music release (*Dave Grusin Presents West Side Story, 1997*) and over the next two decades, hundreds of 5.1-format music titles were released on DTS-CD, SACD and DVD-A [8][9]. However, the production of immersive music content was still slow; such releases were seen as niche products for a small audiophile market. Immersive, non-soundtrack music releases were exceedingly rare.

Dolby's Atmos platform was introduced in 2012 in the cinema and in the following year, Elton John's *Rocketman* was re-mixed and released in that format [10]. It was not until 2017, however, when the market demand for immersive music releases started to grow [11]. It is worth noting that at this point the Dolby Atmos platform had been out for five years, but it was still used overwhelmingly for film audio. By 2019, major labels, with exclusive access to the Atmos release pipeline, were re-releasing many of their music titles in that format. Most of these albums were originally conceived and produced in stereo and thus their re-releases underutilized the expansive potential of Atmos' object-based audio technology. Sony introduced its own object-based immersive format, 360 Reality Audio, in the same year.

During the following years, the Atmos release pipeline opened up to all professionals. Under their Dolby Atmos Music program, Dolby developed tools specifically with music mixing and mastering in mind. The author of this paper was working on Atmos music mixes of independent electronic music in 2019 at Dolby HQ in San Francisco. Atmos music mixing was a still a curiosity at that time; neither set conventions nor recommended best practices were available for immersive music mixing. During the lockdown of the 2020-21 pandemic, the popularity of the format exploded. Dolby's rendering application was sold at the price of an average plugin, and by late 2023, the Atmos Renderer was integrated into Avid's Pro Tools as a no-cost feature. Now, virtually any engineer had convenient access to the 128-channel immersive audio production platform.

## 2 Spatial Music Happens in Production

Despite the easy access to spatial audio technologies, spatial music production tools are used, with an overwhelming majority, in the recording, mixing, and mastering stages of music production today. Popular phrases such as "Mixed in Atmos" or "Ambisonic Recording" reflect the nature of immersive audio production: the musical idea is initially created by the artist, composer, songwriter, producer and/or performed by musicians, then it is engineered in an immersive format by an audio engineer.

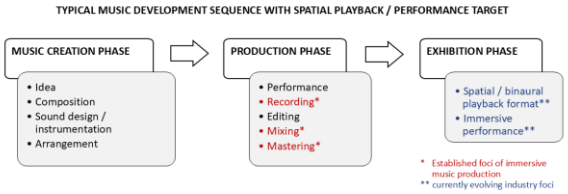


Figure 1. Music Development Sequence, highlighting the established areas of spatial music development (marked with asterisk) and currently evolving areas (double asterisk). No areas within the creation phase are marked, indicating lack of focus.

### 2a. The Recording Stage

Ambisonics is a technology primarily used for recording sounds within physical spaces, and has enjoyed enthusiastic application in VR/AR, game and 360-video productions. It is often the choice for recording orchestral performances, with the intent to document audio information in a given space. It is also used for recording impulse responses to capture the acoustic character of the space itself. Ambisonics is well-suited for documenting the three-dimensional perceptual aspects of a performance. Though it offers the option to isolate and reposition sound elements in space, post-processing of its recorded format is currently considered cumbersome with only a handful tools supporting such processes. Therefore, its application for recording abstract-space performances is impractical – i.e. for recordings not made in the same physical space as the final product is intended to convey, such as acousmatic works that involve isolated component recordings, direct input and virtual instrument sources.
Once recorded, spatial information is relatively difficult to modify in the mixing phase without incurring heavy encoding tasks and causing noticeable side effects (such as phasing) and compromised spatial images [12].

This is a preprint. Copyright ©, ℗ Julius Dobos / The Creative Shop.

**2b. The Mixing Stage**

When working with isolated sound sources using today's popular object-based spatial audio platforms, such as Dolby Atmos and 360 Reality Audio, decisions about spatial positioning and the perceived motion of musical parts within a 360-degree half- or full-sphere space are made during the mixing phase. These engineering and creative decisions are typically made by the mixing engineer, with or (often) without consultation with the artist. Engineers utilize the available spatial mixing technologies and playback systems with three distinctly different goals in mind:

1. Provide a set of satisfying solutions to manage mix conflicts. A mix engineer can increase clarity within a mix through spatial separation between acoustically conflicting elements in an arrangement.

2. Provide the listener with some degree of interactivity in the listening experience. This is possible when utilizing a playback system with a high number of actual audio channels distributed in a 360-degree sphere- or half-sphere environment, as opposed to inferred pseudo locations (such as the phantom center in stereo playback) and pseudo-spaces (such as the perceived size of an emulated space conveyed by reverb and other audio processes) customary in stereo mixing. Interactivity is accomplished when listeners can isolate musical components (objects) by "leaning closer" (mentally or physically) to a specific area of the playback environment. To provide such experiences, the engineer must decide during the mixing phase to isolate and precisely place a sound element – which is a fundamental concept in object-based mixing.

3. Improve spatial-dynamics in the three-dimensional space, thus adding new, creative qualities to the music mix. While such creative ideas certainly can make for a more dynamic listening experience, whether they represent authors' original intent is often questionable; they can potentially result in what listeners experience as "presentation gimmicks," taking the attention away from the musical message.

In other uses, the engineer may seemingly place the listener in the middle of the performance space ("stage") to provide an intimate listening experience.

**2c. The Mastering Stage**

The goal of mastering in immersive audio production is to establish a consistent listening experience and provide deliverables based on target formats and specifications. Similarly to stereo music production, no musical content decisions nor choices concerning spatial location are made in the mastering stage.

## 3 Challenging the Status Quo of the Immersive Production Pipeline

Decisions about the location, space and movement of musical elements are typically made during the production phase, often by the audio engineer, regardless of the immersive technologies utilized. In fact, most spatial audio technologies, tools, and techniques available for immersive music production are developed for the audio production phase, rather than the content creation phase. Except for a few cases, such as interactive composition and improvised performance, the musical idea is first conceptualized, then the musical content is created via composition, arrangement, instrumentation, etc. No spatial considerations are typically made at this time.
Considering that composers utilize pitch, harmony, timing, and timbre to express a musical idea, could these elements be also be shaped or even inspired by space as a conceptual component? Would the result be the same if spatial decisions were made during conception and composition, rather than later during audio production? And, would making such spatial decisions during the composition process have a bearing on the resulting composition itself?

## 4 Space – a New Frontier

Prior to the 1900s, melody, harmony, and rhythm were considered the fundamental elements of music composition. Around the middle of the 20th Century, technologies to record and modify sound became available to composers, bringing a new element of music into focus: timbre. While some compositions as early as the late 1940s already emphasized timbre (some even excluding any melody, harmony, rhythm and exclusively focusing on timbre), it took decades for timbre to gain general acceptance as a critical element of composition.

Early composers of electroacoustic music – as most composers of acoustic music – considered space as an extrinsic attribute of music, typically tied to the location of a performance. Space was therefore not seen as a musical attribute that would influence music

This is a preprint. Copyright ©, ℗ Julius Dobos / The Creative Shop.

composition. Expressing the size of space around a sound source (e.g., an instrument), suggesting the distance between the source and the listener, or imparting mobility to a sound object by moving it around within the listening environment, were generally considered characteristics of recording and playback (documentation and imitation of the performance space), rather than attributes of the musical content.

In 1963, the Buchla 100-series modular synthesizer introduced quad-playback capability, evidencing their consideration of space as an intrinsic component of sound synthesis. Buchla later offered a more deliberately designed spatial modulation feature in their 200-series system. In 1969, the custom-made Azimuth Co-ordinator allowed Pink Floyd to move sounds around in the auditorium between six speakers during performance. However, with no spatial audio playback technology accessible by the general public, spatial music experimentation was largely confined to research labs and academia. While the evolution of music culture and the progress of music technology over the past 75 years have led to the complete acceptance of timbre as a fundamental element of music composition, space has not been accepted as an integral compositional component.

This paper proposes the re-evaluation and re-introduction of the spatial element as a fundamental musical constituent belonging to the composition phase. Conceptualizing space as intrinsic to the composition as a whole may have the potential to influence the creative process of composers today as much as new timbres did during the 1970s. The compositional use of spatial elements has become feasible as, 1) the latest technologies and accessible audio production tools provide satisfactory translation between the creation, production and playback stages, 2) the proliferation of release formats and distribution of spatial work do not depend on an exclusive release pipeline, and 3) a wide variety of consumer-level immersive audio playback systems, including binaural options have become affordable for a rapidly growing audience.
During the 20th Century, technology and access to new timbres led to a musical revolution. Are we at the brink of a similar musical revolution as spatial musical releases are becoming widespread? The film and game industries' push for immersive production and playback over the past decades would certainly indicate so; the audience's awareness of immersive audio content has grown exponentially.

In our Sonic Balls experiment, which is the basis of our upcoming paper, focus group listeners were able to match ten audio samples with ten written emotional descriptions, with greater than 85% confidence. The only difference among the ten audio samples was the spatial automation applied in each; the melodic, harmonic, rhythmic, timbral content, as well as loudness, dynamics and tempo were identical between all samples. Participants listened to the samples in a 7.1.4 monitoring system in random order and paired them with adjectives provided to them. This experiment proved that differences in spatial location and movement alone have the capacity to convey emotions. It also revealed that the spatial variable in music is capable of modifying emotional content, and with that, changing musical meaning.

Pitch, timbre, duration and loudness have long been established as the intrinsic qualities of musical notes. Unlike a single note's frequency, overtone content, duration and amplitude, its fixed spatial location is not an absolute, objectively measurable quality. However, we argue that *spatial movement* during the sustain phase of a musical note can be considered its intrinsic quality. Let's examine loudness for comparison. The constant, absolute loudness (SPL) of a single note, while measurable, is compositionally meaningless. At the same time, the note's dynamics – the relative changes in its loudness throughout the duration of the note – can lend the note musical expression and meaning (e.g. crescendo). Similarly, the spatial location of a musical note may not be musically meaningful, but its spatial dynamics, or spatial movement throughout a sustained note, can become an objectively measurable, inherent musical attribute. Moreover, considering the (fixed) positional relationship between two or more notes in a musical sequence, we assert that spatial location and movement indeed manifest as intrinsic musical qualities of a note sequence.

Supported by our Sonic Balls experiment data, we propose that space is to be recognized as the 5th intrinsic quality of musical notes and a compositional design – as opposed to production design - element.

## 5 Definition of Purpose and Challenges

The primary purpose of this study was to uncover - through music composition and focus testing - whether making spatial location and spatial movement choices during the composition process influence the resulting composition. As the qualitative evaluation of a composition can be highly subjective, focus audiences auditioned alternate

This is a preprint. Copyright ©, ℗ Julius Dobos / The Creative Shop.

versions of a sample composition. One version of an original theme was composed in a stereo monitoring environment and with stereo playback in mind, while the other version of the identical theme was composed in an immersive monitoring environment and for immersive playback.

The secondary purpose of the experiment was to find out in what way, if any, spatial composition in and for an immersive environment influences a creator's use of the fundamental musical elements, compared with the choices they make when composing in and for a stereo environment. The differences in melodic, harmonic, rhythmic, phrasing, timbral, instrumentation and arrangement choices would indicate no, some, or significant influence.

The tertiary goal of the experiment was to compare the narrative and stylistic qualities of a musical composition created in one versus another playback system. Listeners were asked to compare purely musical contents, thus removing the variable of the exhibition format during blind A-B testing.

To achieve maximum objectivity, we formed focus groups consisting of listeners with various critical listening skill levels and listening habits. The musical material had to satisfy these preconditions: 1) it must be based on an original musical idea, to avoid conscious or subconscious influences of preexisting work, 2) it should use a mixed instrumentation approach: be acousmatic in nature but include acoustic instrument samples, in order to consider the expectations of traditional sound element positions (e.g. a violin section on the left), yet allow for the potential disruption of such conventions. The composer should consider the final realization and playback of the music; however, they must implement spatial ideas in the form of compositional choices, rather than via production treatment in the later mixing stage. 3) The compositions must be created by the same composer to rule out the consequences of differences between multiple composers' personal expressions, skill levels, stylistic preferences and former experiences.

There were several challenges to satisfy the above requirements. The two compositions must be created along the same timeline, as much in parallel as possible. Utilizing the same amount of time during the same part of the same day(s) minimizes the variation in the effects of the ever-changing environment (e.g. temperature, humidity, ambient light, ambient sounds) on the two compositions, as well as the physiological and psychological effects (e.g., tiredness, inspiration, mood). I.e., the composer must create the two versions in parallel, using the same tools. Audio monitoring during composition must be done using the playback system targeted by the given composition: one version should be composed using stereo monitoring, the other must be monitored in an immersive system. Furthermore, the arrangement and instrumentation choices in the two different compositions should also be kept identical, except for differences inspired by space. Ideas related to spatial location and movement of musical parts should inform the direction of the composition; the composer must consciously reject any new ideas during the writing process that are not derived directly or indirectly from the spatial element of music - or - implement non-spatial ideas equally in both compositions. An example of the direct influence of a spatial idea on the composition is "writing a part to fill the top right rear corner of the space". This is a conceptual choice. In contrast, a melody inspiring the notes of a new counter-melody is not a conceptual choice, and should be either ignored or implemented into both compositions equally, as should be any new melodic, harmonic or rhythmic idea not directly inspired by space.

## 6 Methodology

The methodology used to satisfy the above requirements was established along the following priorities:

1st priority: the composition choices about the stereo and spatial versions of the same core theme should be informed exclusively by the real-time experience of the composer. The spatio-musical ideas originating during each of these two isolated composition phases will therefore derive from monitoring the process in real time in a stereo and in an immersive environment, respectively.

2nd priority: the production processes (recording, mixing, rendering) following the composition phase should have a minimal or no influence on the outcome and should not affect the musical content. The mixing engineer must strive to create two identical mixes in terms of presentation, impact and perceived loudness.

3rd priority: while the two compositions should be monitored only in their native playback format during the music creation and production phases, the final renders should be made available for feedback and content analysis by the focus groups in a common format, i.e. both stereo or both immersive. This

This is a preprint. Copyright ©, ℗ Julius Dobos / The Creative Shop.

requires the conversion of both mix versions to a different, non-native exhibition format: the "up-mix" of the stereo composition to a spatial audio format, and a "fold-down" stereo render of the immersive mix. As some differences might result from the up-mixing process, and minor artifacts (e.g., phasing) might occur during the fold-down conversion, adjustments should be made to minimize such artifacts. However, any adjustment made must not alter the musical content in any way.

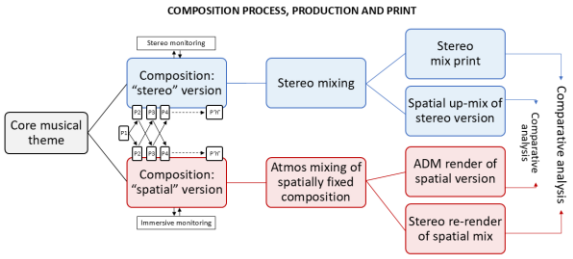


Figure 2. Composition & Evaluation Methodology

As seen in the above diagram, the core musical theme is produced in two versions by the same composer in parallel: a stereo and a spatial version. The musical parts are written in a chronologically alternating pattern, as inspired or influenced by the space and experienced via stereo vs immersive (7.1.4 or binaural) monitoring. The order of composing new parts for the two variants is reversed at times (spatial vs stereo first) to randomize the process. For evaluation, the two compositions are monitored in a common, shared format; stereo mixes of both compositions, or both compositions listened to on an immersive playback system, in order to compare exclusively their musical contents.

## 7 Production Process

The comparative methodology underpinning this paper is dependent on the accurate real-time monitoring and audio rendering technology used during the composition and production phases. The following hardware and software platforms were chosen: Avid Pro Tools, Dolby Atmos Renderer, iyO VAD Pro audio device and renderer, a 7.1.4 studio monitoring setup, a 64-speaker dome environment.

The individual musical parts of both stereo and spatial variants of the composition were rendered as audio stems in their respective Pro Tools sessions. This included audio content and final spatial positions (spatial location and automation data). Hence, the mixing process started with all spatial variables already implemented. The mixing phase employed traditional audio processing (i.e., equalization, compression, reverb) using the same tools, identical effect plugin settings and near-identical non-spatial automation (mostly volume automation). The minor processing differences between the mixes of the two versions were necessitated by differences in content, rather than being subjective choices of the engineer.

The final renderings yielded a stereo file and a Dolby Atmos ADM file. Additionally, an opposite-format render of each composition, as explained above, were created. The stereo fold-down of the spatial version was created using the 2.0 direct re-render algorithm of the Dolby Atmos software. The up-mix of the stereo version was created manually by distributing its content to a 7.1.2 bed channel and to numerous objects, mirroring the mix structure and spatial treatment of the spatial piece's immersive mix, and by importing pan automation data from it. Similarly, the same musical parts of the stereo composition were selected for height channel placement, at identical position and mix-to-height channel level ratio, as the ones inherently featured in the height channels of the spatial composition's native immersive mix. The overall impact of both versions was compared on monitoring systems featuring 2.0, 5.1, 7.1, 7.1.4 and 63.1 speaker configurations.

## 8 Results: The Effects of Compositional Design

Composer's notes on the effects of alternating monitoring formats and experiential differences were analysed and the results documented in a Composition Experience Chart. This document reveals that the compositional space (and playback format) had no direct effect on 28 (74%) of the final piece's 38 instrumental parts. Of the 10 differing instrumental parts (26% of the total number of parts), 6 parts were significantly different (or absent in one or the other version), and 4 parts were somewhat different between the two compositions. Compositional space had the largest influence on melodic phrasing and rhythmic patterns, effecting a total of 9 parts, as well as on melody-writing, which was influenced by space in 7 musical parts – 5 of the parts showing a significant melodic difference. Most of the timbral differences between the two compositions were the combined consequences of melodic and phrasing differences as well as their spatial locations and movement. Not surprisingly, timbral differences are most noticeable in synthesized (as opposed to sampled acoustic) sound sources, one case showing a significant difference, and two musical parts showing small timbral differences.

This is a preprint. Copyright ©, ℗ Julius Dobos / The Creative Shop.

Harmonic progression was the only fundamental element of composition that remained unaffected by the change of space. However, the differences in melodies occurring in the same harmonic context created variances in the momentary manifesting harmonies. Though in our experiment these variations remained relatively small, the extent of such differences in any given composition have the potential to alter harmonies as a natural consequence of melodic changes.

## 9 Evaluation of Stereo vs. Spatially Composed Content

In accordance with the evaluation method above, we conducted listening sessions with 125 individuals assigned into eight cohorts. Listeners participated in 11 group listening sessions, for a total of 159 individual impressions over four months. Some sessions included only one exhibition format, while others included both stereo and immersive playback.

Focus groups noted their preference exclusively for musical content. The playback order of compositions and mix format pairs were reversed to mitigate first impression bias and recency bias. The summary of focus group evaluations from all eight sessions follows below.

Results of Spatial Composition

Focus Group Sessions

| Group / Session | Version playback order | Playback format | Preference # of listeners | |
|---|---|---|---|---|
| | | | Stereo comp. | Spatial comp. |
| Grp #1 S #1 | SP, ST | BIN | 8 | 12 |
| Grp #2 S #2 | ST, SP | BIN | 16 | 8 |
| Grp #3 S #3a | SP, ST | 2.0 | 0 | 10 |
| Grp#3 S #3b | | 5.1 | | |
| Grp #4 S #4a | SP, ST | 7.1.4 | 6 | 10 |
| Grp #4 S #4b | | 2.0 | | |
| Grp #5 S #5 | ST, SP | 2.0 | 8 | 10 |
| Grp #6 S #6 | SP, ST | 2.0 | 8 | 12 |
| Grp #7 S #7a | ST, SP | 2.0 | 6 | 2 |
| Grp #7 S #7b | | 7.1.4 | | |
| Grp #8 S #8 | SP, ST | 7.1 | 4 | 5 |
| 125 individual listeners<br>159 listening test impressions | | | **68**<br>**42.8%** | **91**<br>**57.2%** |

## 10 Discussion

Various theories on the synergistic relationships between melody, rhythm, harmony and space have been proposed, agreeing on the importance of compositional intent. [13] Our comparative spatial composition experiment provides data and evidence, that using space as a compositional element has a significant effect on the resulting musical content. Feedback notes from focus group members detail the individual preferences between the two compositions, in many cases identifying the effects of spatio-musical design, regardless of the mix and monitoring format.

Focus group evaluation showed an overall preference for the spatial composition by 14.5 percentage points. Interestingly, while auditioning in the stereo playback environment, listeners preferred the musical content of the spatial composition by a smaller margin, than in the immersive environment. In the immersive listening environment, their preference for the musical content of the spatial composition was more pronounced. This confirms our expectation, that an immersive playback system provides an advantage for conveying the spatio-musical details that were concepted during the immersive composition phase. The stereo composition did not have such advantage. However, the narrative meaning of the spatial composition's spatio-musical details diminished when the piece was auditioned in stereo.

An unexpected finding from our focus groups was that most listeners had clear personal preferences for musical content, regardless of the listening format and environment. This confirms that the influential differences were rooted in musical content design and were not the effects of differing presentations. Listeners provided detailed reasons for their preferences, thus, giving us high quality data and the confidence to rule out listener indifference and random choices. Also, a broad demographic diversity, as well as the wide spread in musical tastes and critical listening skill levels among focus group members ensured that the results did not carry an inherent ability- or cultural bias.

In conclusion, without the existence of two different spaces which respectively inspired two different compositions, one of the compositions (which was preferred by 33.8% more listeners), would not have existed. Had the writing work happened only in a stereo environment, the resulting composition would have missed details that made our spatial version more satisfactory for 57.2% of our listeners, regardless of the mix and playback format. Therefore,

This is a preprint. Copyright ©, ℗ Julius Dobos / The Creative Shop.

while using space as a compositional element might not result in a composition objectively inferior or superior relative to one created without any spatial consideration, we conclude that the use of space as a musical component can lead to significant enough content differences to cause some listeners to prefer the result – in our case, the majority of listeners.

More importantly, notes taken independently by focus group participants and the composer revealed similar differences between the compositions in terms of emotional effect, and individuals used much the same descriptors for these experiential differences. We collected and analysed over 1,000 data points from our focus group sessions. Keywords describing the stereo composition include "melody, melodic, beautiful, pretty, emotional, connected, familiar, realism, expected, cohesive, classic, warm, full, celeste". Keywords describing the spatial composition were quite different: "fresh, impact, action, modern, interesting, unique, contrasting, exciting, fragmented, movement, intriguing, playful, fun, adventurous, journey, bass, strings".

One listener who preferred the spatial version summarized their experience as: "The spatial version is fresh, modern, exciting" while another listener wrote "the stereo version tried to 'go big' in a more cliche way, whereas the spatial version was more powerful in a smarter way". Listeners who preferred the stereo composition remarked that "[it] felt more like the storytelling arch of a character" and "the spatial version was amazing but the stereo composition had my emotions activated the entire time".

These impressions are consistent with the composer's notes describing the effect of space on the composition process. Such experiences seem to be caused by the diminished focus on spatial concepts in the stereo composition, in which the composer relied more heavily on melody, harmony and longer melodic phrases, resulting in a composition driven by one or more of the traditional elements of music. Conversely, increased focus on spatial expression during the spatial composition diminished the focus on melody, harmony or rhythm as primary tools of expression, resulting in potentially shorter, rather fragmented melodic phrases, yet more unconventional ways of expressing the musical concept. The majority of listeners and composer have shared the view that the stereo piece has stronger ("more beautiful") melodies and better flow, while the spatial piece is more thrilling, dynamic ("unique and fun") in unconventional ways. It is worth noting that the spatial composition prompted nearly twice as many adjectives and longer feedback notes from all focus groups – without participants knowing which version they were describing.

The potency of space as a compositional tool and conceptual element of artistic expression cannot be understated; its use during composition can have a complex and profoundly influential impact on the resulting musical piece.

## 11 Future Implications

### 11a. Space as an element of composition

While we acknowledge the limited scope of our experiment, its outcomes suggest that using space as a conceptual compositional element is a potent technique to influence emotions and the message. Focus group data also confirm that the compositional use of space exerts a far more significant influence on the other, traditional elements of music, and on the listener's experience, than the spatial techniques that can be applied during audio production. Spatio-musical design may help the composer to express concepts and inspire new compositional ideas that would not surface without its deliberate use during the writing process. In this regard, space may become an influential element in the music creation process – as timbre came to be some 75 years ago.

### 11b. Spatial concepts in production – What it means for audio engineers

Spatial ideas generated and implemented during the audio production process by the audio engineer, whether or not involving the composer, may complement or may conflict with the musical message. A lack of connection or contrast between composers' preconceived narrative ideas and the spatial narrative subjectively interpreted and expressed by the engineer during production may lead to inconsistent recording and mixing decisions, conceptual conflicts and even disconnection from the narrative concept. Moreover, spatio-musical details may not be possible to implement to their full potential during the audio production phase; as mixing engineers typically work with previously recorded content, spatial decisions applied during the mixing phase would not influence melodic writing, phrasing, harmony, rhythm or timbre. This, again, points to the need for composers and engineers to coordinate and preserve the musical concept during production, to ensure that preexisting spatio-musical

This is a preprint. Copyright ©, ℗ Julius Dobos / The Creative Shop.

ideas will be expressed through the mixing process when targeting immersive playback and performance formats. It is the composer, not the engineer, who should be expected to take a lead in communicating such spatial design concepts, as an integral part of the musical message.

How would the role of the mixing engineer change when it comes to making mixing decisions about spatial location, image and movement? How should the spatial concept be preserved and transferred from composer to engineer? Using the model of the current music production pipeline, we can presume that specific spatial information can be conveyed via one of the following three practices:

a) The composer makes spatial decisions and records them into a DAW session in the form of either spatially-positioned, rendered audio recordings (mock-up or final), and/or via spatial automation, before handing over the session to a mixing engineer. This works well for generated sound sources that are under the composer's control, but requires an advanced technical competency from composers.
b) The composer and mixing engineer work together in real time and verbally coordinate the execution of spatial concepts. This requires clear and precise communication, as well as significant time, which may translate to increased costs. While this method is the least efficient and requires the composer to be present during production session(s), it is the most effective in cases when recordings are made post-composition in the recording studio, and the production schedule allows for spatial experimentation and compositional changes.
c) The composer utilizes sheet music featuring a notation system that conveys detailed spatial location, imaging and movement of musical parts and events. This requires the development and standardization of a new spatial notation language. A component of such system, a spatial notation symbol, Locus, was proposed by Luís Zanforlin in 2018. [14] While Zanforlin's symbols mark location and movement in three-dimensional space, they do not scale up to complex musical compositions and potentially high accuracy of placement in the listening environment (e.g., precise location, object size, etc.). However, the Locus symbols provide a solid starting point for future universal spatial notation systems that contain detailed performance instructions for a large number of musical parts. A significant advantage of this method is the potential integration of a standard visual notation system across all stages of music composition, production, and performance by all participants of a comprehensive pipeline, extending to live sound and possibly even to lighting design.

## 11. Expressing Spatial Concepts in Live Performance

Although there are more examples of the conceptual use of constructed spaces in acousmatic music, today's technology allows for real-time signal processing of live-performed musical parts (including acoustic sources), which would allow the real-time performance of pre-defined spatial concepts. A preconceived spatial narrative could be conveyed in any of the following ways:

a) The performance can be conducted to a prepared tempo map while various instruments or instrument sections are being spatially positioned in real time. In this case, pre-determined pan automation can be playing back on input / aux channels, in sync with the live performance. This method allows for less spontaneity, and may require a high level of timing precision from the performers. However, the precision and complexity of such spatial automation can achieve complex, otherwise unattainable results in real time.
b) A lively solution would involve spatial scores featuring spatial notation, which would guide dedicated spatial mixing engineers to perform live spatial mixing, following the spatial score in sync with the performers (or conductor) in real-time. This method would require the development and practice of a new skill from the engineer and a comprehensive, standardized spatial notation system. However, such coordinated work can make the performance spatially dynamic, lively, and turn it into a non-repeatable experience for listeners.
c) Further evolving the real-time spatial mixing engineer concept, spatial conductors, (secondary to a traditional conductor, if any) could join performers on stage. Referencing a spatial score and using 3-dimensional gyroscope-equipped wireless conductor batons, spatial conductors could directly control the placements of various instruments in real time from the stage, selecting sources and placing them by pointing at areas of the physical performance space with their baton.

This is a preprint. Copyright ©, ℗ Julius Dobos / The Creative Shop.

By wearing augmented reality headsets, the spatial conductor could work in sync with the performers on the physical stage while viewing a panning container overlay that represents the venue and the position of sound objects in it (akin to the Dolby Atmos 3D visualizer). In this setup, spatial conductors would select, position and move sound sources as objects in the virtual environment, which would be replicated in the actual venue's playback system in real time. While offering a solution to theories which suggest that, in lieu of being part of an audiovisual work, performative spatio-musical expression needs visual representation, [15] this approach would produce a truly organic performance and dynamic live experience.

In any of the above scenarios, the goal of spatial notation is to perform the spatial expressions previously designed by the composer, as opposed to adding a new spatial layer to the live performance.

## About the Author

Julius Dobos is an award-winning, platinum-album European-American composer, sound synthesist, music director and mastering engineer. He is best known for his 11 electronic- and electro-orchestral albums, music produced for film, television, installations, major events and international brands. His composition credits span over 30 years and include You Don't Mess with the Zohan, Paul Blart: Mall Cop, The Zookeeper, the DragonBall Z series, NCIS, to name a few. His production work includes Grammy®-awarded artists Martha Sebestyen, Erykah Badu and a colorful list of music & audio projects.

Under the pseudonym forgotten future, Dobos crafts one-of-a-kind, dynamic electronic music with a distinct style and sonic character. His live music performances have attracted audiences worldwide, the 2019 Los Angeles Human Intelligence Control show featured a novel musical experience through digitally-tracked interaction with the audience.

Dobos is a Distinguished Adjunct Professor at the University of Silicon Valley, a published author and speaker on electronic music and production culture, a long-time synthesizer- and audio technology beta-tester. He is a BMI and Artisjus writer, publisher, and the Vice Chair of the Audio Engineering Society San Francisco section.

## Acknowledgements and Founding

The author would like to express his gratitude to Dr. Timothy W. Duncan, DMA for his review of this paper. The author would like to acknowledge David Bowles for his consultation on ambisonics. Dolby Laboratories, Inc. provided the Dolby Atmos Renderer software and iYO, Inc. provided the VAD Pro hardware and anechoic chamber session time at no cost to author. Author received no compensation from the above companies related to this study and has no financial interest in their technologies.

## References

[1] Staff Writer, *Varese Envisions 'Space' Symphonies,* The New York Times (1936)

[2] E. Varèse and C. Wen-chung, *The Liberation of Sound,* Perspectives of New Music (1966)

[3] V. Lombardo, et al, *A Virtual-Reality Reconstruction of Poème électronique Based on Philological Research,* Computer Music Journal, 33, p.24--47 (2009)

[4] M. Calore, *May 12, 1967: Pink Floyd Astounds With 'Sound in the Round',* Wired (2009)

[5] L. Robinson, *Does Surround Sound Have a Future?,* MidiMagic (2022)

[6] Dolby, Inc., *Company History – A Formula for Film Sound Success,* dolby.com (2004)

[7] *A Chronology of Dolby Laboratories May 1965-May 1998,* dolby.com

[8] D. Zych, *Dave Grusin: Inside Story,* JazzTimes (July 2024)

[9] M. Anderson, *DTS CD Discography,* The Surround Discography, Version 2.2a (2023)

[10] B. Schultz, *Classic Hits go Immersive,* Mix Online (2020)

[11] C. Chagas, *Spatial Talks w/ Greg Penny, S*ound Particles (2025)

[12] F. Grond, PhD, *Higher-order Ambisonics in Dolby Atmos Productions,* Zylia (2022)

[13] M. A.J. Baalman, *Spatial Composition Techniques and Sound Spatialisation Technologies,* Cambridge University Press (2010)

[14] L. Zanforlin, *Locus: Spatial Music Notation,* luiszanforlin.com (2018)

[15] E. Bates, *The Composition & Performance of Spatial Music,* Enda Bates - Trinity College Dublin (2009)

This is a preprint. Copyright ©, ℗ Julius Dobos / The Creative Shop.