

The Aural Cartographer's Dilemma

Austin Franklin
austin.franklin@mdu.se
Mälardalens University
Västerås, Sweden

Abstract

This paper introduces the aural cartographer's dilemma: contemporary digital tools for spatial audio excel at mapping sound objects within coordinate systems but conflate coordinate-based representations of space with the embodied, relational experience of sound. It traces this inheritance from stereo panning through MPEG-I, demonstrating that successive technical paradigms instead reproduce a composer-score-conductor model through ever sophisticated means. It argues this is not a technical limitation awaiting solution but a philosophical commitment requiring re-consideration.

Keywords

Computer Music History, Spatial Audio, Epistemology

1 Introduction

The aural cartographer's dilemma is a specific instance of a fundamental epistemological problem: Alfred Korzybski's axiom that "the map is not the territory" [19]. This principle warns against confusing a representation of reality with reality itself. In computer music, this fallacy is routine. Our most sophisticated tools, i.e., spatial audio systems, are intricate maps, abstractions that reduce sound to objects in a coordinate space. The field's prevailing effort has been to create ever more detailed and accurate versions of this map, refining the precision of sound-object placement. While Korzybski rightly observed that the map is not the territory, in spatial audio, the map is also not the experience.

Traditional cartography, for all its precision, remains ontologically committed to space as a container. This limitation finds its analogue in computer music, where tools for spatialisation often reproduce the very hierarchies they might disrupt. As Michel de Certeau observes in *The Practice of Everyday Life*, there is a fundamental difference between the "voyeur-god" perspective of the mapmaker and the embodied act of "walking in the city" [9]. That distinction persists in our digital musical tools, which privilege the composer's plan over the participant's experience.

The concert hall endures as our dominant metaphor. Its architecture of composer, conductor, and audience reappears in digital form through protocols that prioritise execution over emergence, plan over process. This persistence reflects less a technical limitation than a philosophical commitment to specific ideas about what music is and how it should be organised. The result is a peculiar paradox: technologies capable of extraordinary complexity deployed in service of centuries-old musical models. As Jaron Lanier notes, because software is "subject to an exceptionally rigid process of lock-in," the ideas it mediates have, in turn, "become more subject to lock-in than in previous eras" [22]. The

Western music tradition, with its emphasis on authorship and faithful reproduction, has profoundly shaped our imagination.

What follows is a critical genealogy intended to be diagnostic, not prescriptive. It traces the development and technical foundations of spatial audio's representational paradigms, articulates the precise nature of the dilemma, and demonstrates how each successive paradigm—stereo, multichannel, wave field synthesis, ambisonics, binaural, MPEG-I—reproduces a composer-score-conductor model. Section 3 discusses the turn towards relational sound studies, followed by speculations in Section 4 on what escaping the dilemma might require.

2 Representation in Spatial Audio

The aural cartographer's dilemma is rooted in the representational paradigms that have guided spatial audio practice for decades. This section traces the development of these paradigms to demonstrate a persistent continuity: each new technology refines the map but retains the same fundamental logic.

2.1 The Unit Circle and Stereo Field

The foundation of spatial audio representation is the unit circle, a mathematical abstraction that reduces the complex, three-dimensional experience of directional hearing to a two-dimensional continuum between left and right (Figure 1). This model formalises the practice of stereo panning, developed through Blumlein's pioneering patents in 1931 and commercialised with the advent of stereo recordings in the late 1950s [3]. The unit circle represents an idealised listening position at its center, with angular coordinates mapping to pan positions (typically -45° and $+45^\circ$ for left and right channels, respectively). Constant-power panning laws, formalised in the 1970s and 1980s, maintain perceived loudness across positions by applying trigonometric gains rather than linear crossfades [24]. This abstraction enables precise, repeatable positioning of sound sources but reduces the rich perceptual experience of directionality to a single parameter. The unit circle model prioritises compositional control and repeatability over perceptual fidelity or relational experience.

2.2 Multichannel Arrays

With the development of multichannel speaker arrays, the representational task becomes more complex but follows the same logic. Quadraphonic sound, introduced commercially around 1970, established the paradigm of four speakers surrounding a central listening position, though it remained a commercial failure [18]. The 5.1 surround standard was formalised as ITU-R BS.775 in 1992, cementing the geometry of five speakers plus a subwoofer [16]. More recent formats like Dolby Atmos, introduced in 2012, allow for "audio objects" with dynamic metadata (position, size, velocity) that renderers interpret in real time for different playback systems [10]. Speaker arrays are typically diagrammed as geometric arrangements around a central listening position (Figure 2). Tools for these systems, such as CueStation, allow sounds to be positioned within a three-dimensional coordinate grid. The composer's role is to populate this grid with sound



This work is licensed under a Creative Commons Attribution 4.0 International License.

NIME '26, June 23–26, 2026, London, UK

© 2026 Copyright held by the owner/author(s).

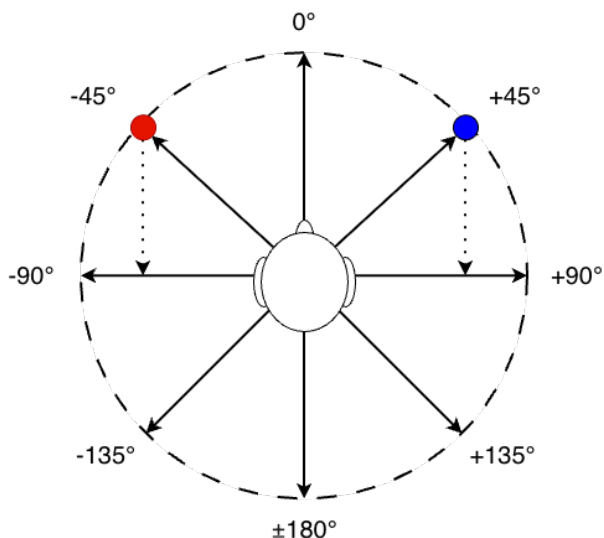


Figure 1: The Unit Circle and Stereo Field

objects, specifying their XYZ coordinates and sometimes trajectories. The underlying assumption is that if the sound object’s coordinates are correctly translated into speaker gains (using vector-based amplitude panning or more advanced algorithms), the intended spatial impression will be recreated for the listener at the "sweet spot." Despite increasing technical sophistication, the paradigm remains representational and object-centric. The composer specifies where sounds should be, and the system attempts to reproduce that specification. The listener’s experience is assumed to be a faithful decoding of this specification, not a context-dependent experience.

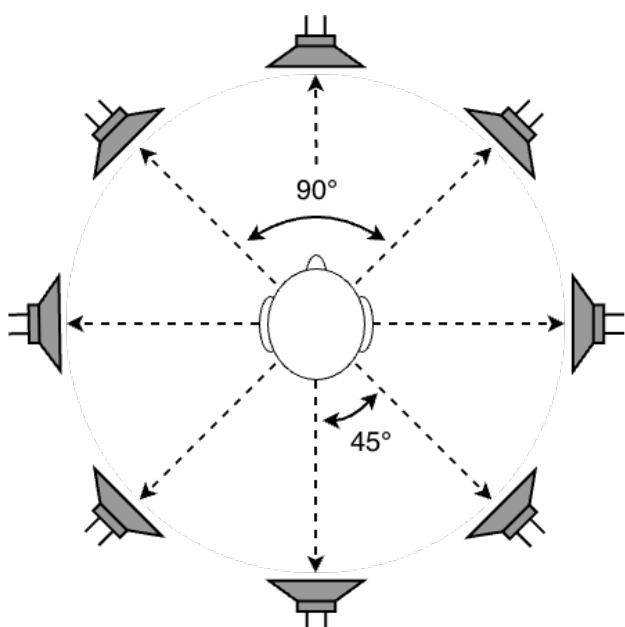


Figure 2: Multichannel Array

2.3 Wave Field Synthesis

Wave field synthesis represents a different, physics-based approach to spatial representation. The theoretical foundation was

established by Berkhout, de Vries, and Vogel in their foundational 1993 paper "Acoustic control by wave field synthesis" [2]. Rather than panning between discrete speakers, WFS uses large arrays of loudspeakers (often tens or hundreds) to physically recreate the wave fronts that would emanate from virtual sound sources (Figure 3). The mathematical representation is no longer a coordinate in a grid, but a set of driving functions for each speaker that synthesise the desired wave field.

While wave field synthesis can create remarkably stable and convincing spatial impressions over a large listening area (eliminating the sweet spot), it remains captured by the composer-score-conductor model. The system’s goal is to accurately reconstruct a pre-defined acoustic scene, an auditory "hologram" of virtual sources at specified locations, treating the speaker array as the concert stage and the field its audience. The representation has shifted from perceptual (panning curves) to physical (wave equations), but the fundamental relationship, composer specifies and conductor (i.e., system) reproduces, persists.

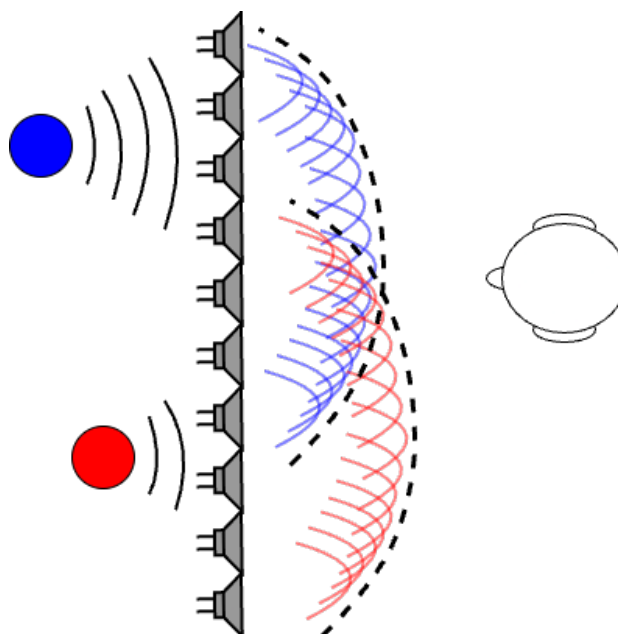


Figure 3: Wave Field Synthesis

2.4 Ambisonics

Ambisonics introduces another representational layer: the encoding of a sound field into spherical harmonics rather than speaker arrays. First-order ambisonics was developed by Gerzon et al. in the early 1970s, with patents filed in 1973 [13, 14]. Higher-order ambisonics, which uses more channels to capture finer spatial detail, was developed primarily in the 1990s and 2000s, notably in Daniel’s 2000 PhD thesis [8]. Ambisonics encodes a sound’s direction into channels representing an omnidirectional component and figure-eight components oriented along Cartesian axes. It is often shown as a spherical coordinate system, with sounds encoded as points or vectors on a sphere surrounding the listener (Figure 4).

Ambisonics is particularly interesting because it is an intermediate representation—a format-agnostic encoding of a sound field that can be decoded to various speaker layouts or binaurally for headphones. This flexibility could support relational listening,

where rendering adapts to listener position or movement. Yet in practice, ambisonics is used to reinforce the same cartographic logic: tools like Facebook's 360 Spatial Workstation or Reaper's Ambisonics Toolkit provide interfaces for placing sound objects in a sphere, treating the sound field as a container. The composer encodes; the decoder reproduces; the listener receives. The format's potential remains unrealised because the tools are built on assumptions inherited from the concert hall. Higher-order ambisonics merely refines the map's resolution; it does not question the map itself.

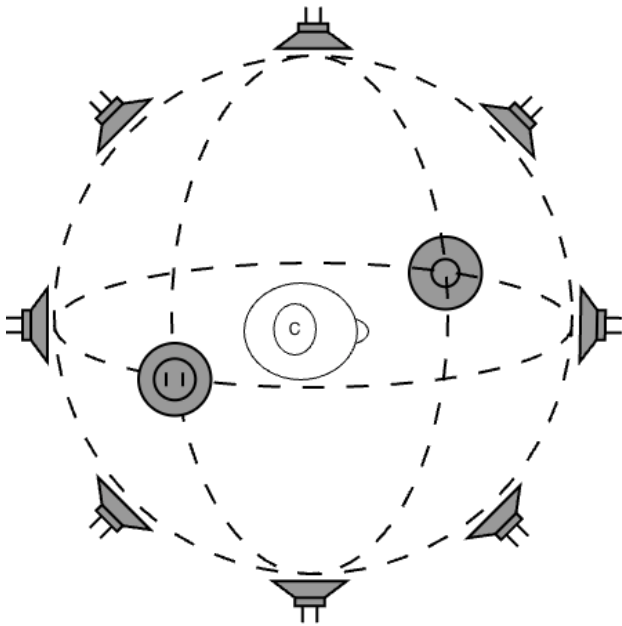


Figure 4: Ambisonics

2.5 HRTFs and Binaural Synthesis

Binaural synthesis attempts to bypass speaker-based reproduction altogether by modeling the acoustic filtering that occurs as sound reaches a listener's eardrums from different directions. This is achieved through Head-Related Transfer Functions (HRTFs)—directional filters measured from dummy heads or individuals. Dummy head measurements date to the 1970s, with the KU100 dummy head becoming a standard tool [25]. The CIPIC HRTF Database, published in 2001, provides a widely-used standardized dataset [1]. In this paradigm, space is represented as a database of HRTFs indexed by direction (azimuth and elevation). To spatialise a sound, the appropriate HRTF pair is convolved with the audio signal, and the result is played over headphones.

Binaural audio is best depicted as a sphere around a schematic head, showing the direction of arrival for the virtual source (Figure 5). While binaural synthesis can create compelling illusions of externalised sound, it reduces spatial hearing to a linear, time-invariant filtering process. The experience is predetermined by the HRTF measurement; individual anatomical variations mean the "map" (the HRTF set) may not match the "territory" (the listener's own hearing), leading to front-back confusions or lack of externalisation. The composer specifies direction; the system applies the corresponding filter; the listener receives the result. That individual ears differ from the dummy head is treated as a system limitation rather than an equally valid experience.

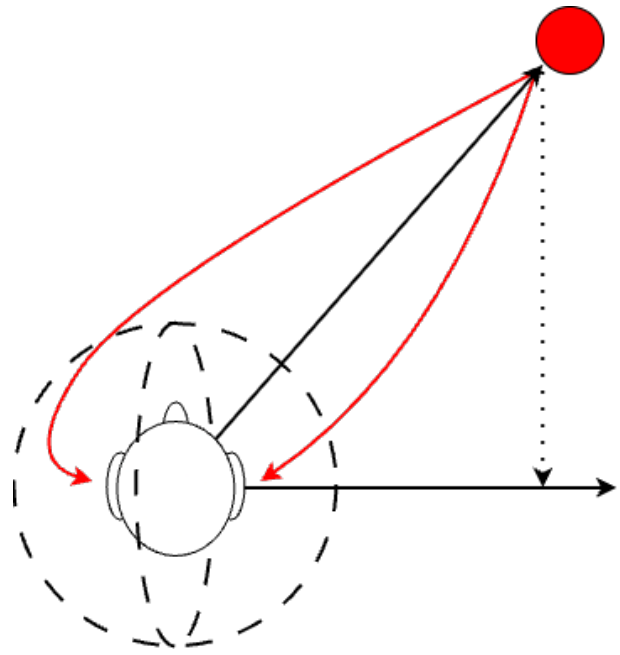


Figure 5: HRTFs and Binaural

2.6 The MPEG-I Standard

The MPEG-I (Immersive) standard, formally ISO/IEC 23090-4, represents the state of the art in standardised spatial audio representation for virtual and augmented reality. Building on the MPEG-H 3D Audio standard released in 2015, MPEG-I was developed between 2017 and 2021 [17]. Its architecture and coding mechanisms are designed to render audio in environments where a user possesses six degrees of freedom, moving and rotating freely in space.

MPEG-I defines a comprehensive pipeline for creating, encoding, transmitting, and rendering immersive audio scenes. Its architecture is object-based and scene-graph oriented. The Encoder Input Format (EIF) is an XML schema that serves as the definitive, pre-authored "map" of an audio scene (Figure 6). It is a declarative description specifying all sound sources (both continuous and interactive event sounds), their static or dynamic positions, acoustic properties, room acoustics, occlusions, and the logical updates that can be triggered by user interaction.

Proponent encoders compress the audio signals and, crucially, encode the spatial metadata from the EIF into a compact, serialised bitstream. This process translates the declarative scene description into a temporal sequence of data packets, containing both audio payloads and the parametric instructions for their spatial presentation. The decoding and spatialisation are performed by renderer modules. The renderer ingests the bitstream and, in real time, decodes the audio signals and associated metadata. Using HRTF sets, it synthesises a binaural output tailored to the user's continuously tracked position and orientation.

MPEG-I is the latest iteration of this representational tradition. The EIF is a map; the bitstream is a conductor's baton; the renderer is an obedient orchestra. The standard encodes a composer-score-conductor model at the level of data structure. It is, in this sense, less a technical innovation than the formalisation of assumptions that have guided spatial audio development for nearly a century.

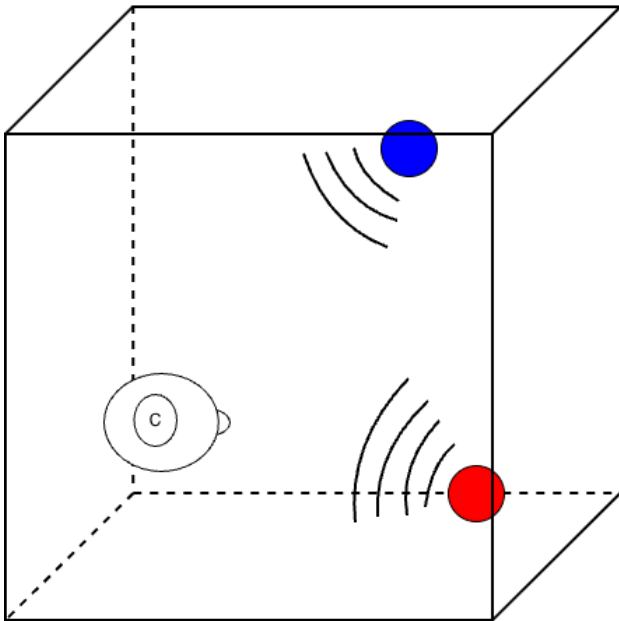


Figure 6: MPEG-I virtual scene

3 Towards Relational Sound Studies

The aural cartographer's dilemma arises precisely because of a composer-score-conductor model inherent to these practices. The word "practices" here aligns with Bourdieu's concept of "durable, transposable dispositions," where repetition builds habits into technical protocols [5]. Software that assumes ideal use cases or workflows actively re-teaches habits when it becomes widely distributed. In spatial audio, this inheritance was not accidental; it was actively cultivated. The spatial music of the 1950s–70s established the composer as a spatial architect, though through markedly different approaches. Stockhausen's *Gesang der Jünglinge* (1955–56) and *Kontakte* (1958–60) fixed spatial trajectories to tape, treating space as a parameter [12]. Xenakis's *Terretektorh* (1966) distributed musicians throughout the audience, creating a relational space whose experience varied with listening position, a more open but still authored architecture.

Cage's *Variations VII* (1966), by contrast, abandoned authorial control entirely, using multiple sound sources to create conditions for sonic emergence rather than specifying outcomes. What these works share is not a unified method but a collective expansion of music's spatial possibilities. Yet it was the Stockhausen model—space as specifiable, repeatable, and authorially controlled—that became codified in software. IRCAM's *Spat*, developed in the 1990s, rendered this logic executable, translating the concert hall's hierarchies into digital form [6]. Xenakis's relational architectures and Cage's anti-authorial gestures remained, largely, roads not taken.

Yet the Stockhausen model, though dominant, was never total. A parallel tradition emerged from the World Soundscape Project (WSP) at Simon Fraser University in the early 1970s, directed by R. Murray Schafer. The WSP proposed a listener-based approach using "ear cleaning" and "soundwalks" to counter habituated non-listening [27]. Barry Truax, an original WSP member, extended this subjective concept into an information-based model of acoustic communication [31]. In this framework, sound results in meaning based on two types of listener knowledge: properties

of the sound itself (spectral and temporal patterns) and environmental, social, and cultural context. Listening occurs at different levels of attention, from foreground analytical to background habitual. For Truax and the WSP tradition, space is an acoustic community where "acoustic information plays a pervasive role in the lives of the inhabitants" [31]. This model is relational from the outset: the listener's position, movement, and contextual knowledge are constitutive of spatial experience, not corrections to an ideal. The WSP approach thus offers an alternative genealogy for spatial audio rather than imposing a pre-authored Euclidean geometry.

Within computer music, scholars have questioned the field's unexamined commitments. Magnusson argues that "the computer is the great impersonator," too often replicating older musical traditions instead of enabling new ones [23]. Born critiques the focus on a Euclidean perspective in electroacoustic music that becomes limiting [4]. Likewise, speaker agnosticism defined by Catena and Frisk rejects fixed speaker layouts in favor of perceptual categories like movement, placement, and occupancy: "it is more important that a sound moves, rather than defining a precise spatial trajectory that is difficult to locate and perceive, perhaps distorting the real compositional intent" [7]. Finally, Sterne questions the ideology of fidelity that equates technical accuracy with aesthetic value [29], while Voegelin confronts Western knowledge and its "reliance on visual categories, straight lines, and universalising principles that pretend objectivity and a distant view" [32]. Yet these critiques, while essential, remain largely within the technological frame. A broader critical tradition in sound studies illuminates what is at stake in the aural cartographer's dilemma by reframing sound not as object but as relation.

LaBelle's concept of "acoustic territories" provides a crucial vocabulary. For LaBelle, sound is a "performative dynamic" that constitutes and contests space. His project traces auditory life to uncover "the embedded tensions and potentiality inherent to sound as it exists in the everyday spaces around us" [21]. Eidsheim extends this line of thinking by arguing that sound is not a fixed object but a vibrational practice, produced through the act of listening itself [11]. Together, these concepts reveal what coordinate-based representations systematically obscure: that listening is always located, always negotiated, always implicated in larger cultural histories and ideologies. From this perspective, our current representational paradigms are an acoustic politics of space that position listeners in relation to an ideal rather than honouring the multiplicity of actual listening experiences.

4 An Escape Attempt

The aural cartographer's dilemma describes the persistent philosophical commitment embedded within spatial audio technologies: the reproduction of a composer-score-conductor model across successive paradigms. In order to plot an escape, we must extend the diagnosis to three additional dimensions: the politics of listening position, the experience of time, and the social life of audio. This section extends the analysis to consider these dimensions while speculating what an escape might require.

4.1 The Politics of Listening Position

The optimal listening position at the center of a stereo field or multichannel array, or the "sweet spot," is a spatial politics. The "sweet spot" refers to the specific location at the geometric center of a speaker array where a listener perceives the intended spatial image with minimal distortion [15]. As Thompson documents,

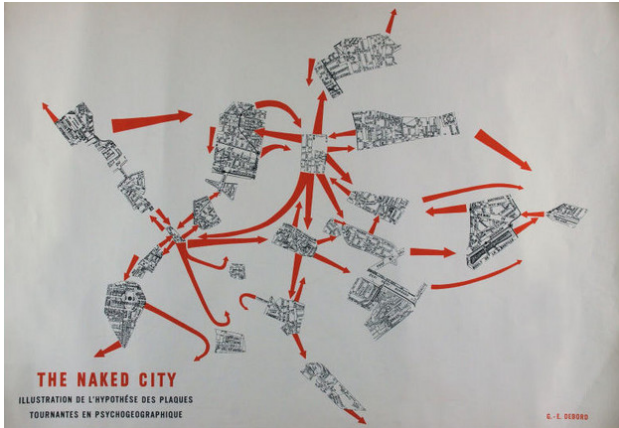


Figure 7: The Naked City by Guy Debord

this ideal listening position is not acoustically inevitable but historically produced, emerging alongside modern architecture and recording technologies [30]. It enforces a single, stationary, and privileged listening position. The more central the position the more expensive the ticket, positioning some listeners as ideal and others as marginal. Even wave field synthesis, which eliminates the sweet spot by reconstructing a pre-authored wave field for an entire area, merely substitutes one form of spatial discipline for another: every position becomes equally good at receiving the composer's specification, but the specification itself remains untouched.

This political reading resonates with the Situationist practice of psychogeography, which sought to map the emotional and atmospheric currents of the city over its bureaucratic grid. Where traditional cartography documented physical infrastructure, psychogeography charted zones of affective intensity and relation [33]. Guy Debord's *The Naked City* (1957) exemplifies this approach, presenting a map fragmented into "unities of ambiance" connected by directional arrows, suggesting flows of pedestrian attention that escape the grid of streets (Figure 7). These were static illustrations, but they articulated a vision of space as lived and relational rather than abstract and geometric.

A genuinely relational spatial audio might abandon the sweet spot entirely, not by making every position equally good at receiving the author's specification, but by making each position produce a different but equally valid experience. This might require distributed systems that respond to listener location and generate unique audio for each participant. However, this raises a crucial question: does the composer's authorship enable or restrict the participant's ability to *dérive*—to drift without intent? The distinction between abandoning the sweet spot versus multiplying it is not a trivial one.

4.2 The Experience of Time

The aforementioned commitment to spatial representation has obscured an equally fundamental dimension: time. A composer-score-conductor model assumes a linear temporality: the work has a fixed duration, events occur at specified times, the listener follows along. This temporal logic is embedded in the structure of audio objects and their metadata, which specify not only where sounds should be but when they should occur. Even interactive

systems that respond to user input typically do so within a predetermined temporal framework: events are triggered, parameters updated, but the overall temporal arc remains authored.

Yet spatial experience is not temporally linear. We dwell in spaces, return to them, move through them at variable paces, attend to different features at different times. A room experienced for the first time differs from that same room after years of habitation; a sound heard once differs from the same sound heard repeatedly, accumulating associations and expectations. Janet Cardiff's audio walks, such as *The Missing Voice* (1999), demonstrate one alternative. The listener hears recorded voices that refer to the present moment while recalling past moments and anticipating future ones, creating a dense temporal layering that exceeds linear chronology. Teri Rueb's *Core Sample* (2007) similarly uses GPS-triggered audio to create a collage of geographical and historical sounds at a single site. These works treat time as multiple and layered rather than linear and sequential, suggesting possibilities for spatial audio systems that operate with similar temporal complexity.

What would spatial audio look like that honoured this temporal complexity? Perhaps systems whose behaviour changes with repeated listening, accumulating memory of previous encounters, or systems that respond to the listener's duration of occupancy, becoming more detailed or more abstract as time passes. A truly relational spatial audio would need to engage time as actively as it engages space.

4.3 The Social Life of Audio

Spatial audio increasingly mediates social listening. We listen together in cinemas, concert halls, and living rooms. We share headphones. We participate in multi-user VR environments. These social contexts complicate the individualist assumptions of a composer-score-conductor model. This resonates with Small's concept of *Musicking*, which argues for a holistic view of music as an activity including composing, performing, listening, and dancing rather than music as a fixed object [28].

Recently, applications like *Randonautica* and *Outside Lies Magic* emphasise the collaborative aspects of space within a user's local environment [20]. These platforms treat space as something to be explored collectively, with outcomes emerging from the intersection of algorithmic generation and social interaction. Similarly, the *Greenwich Emotion Map* (2005) by Christian Nold equips participants with GPS and galvanic skin response sensors to create maps where topography is overlaid with data of emotional arousal (Figure 8). This project attempts to map the internal experience of moving through a space, making visible the affective dimension of spatial practice [26]. While it remains analytical, documenting experience rather than generating it, it points toward a social, embodied understanding of spatial encounter that could inform generative systems.

Taken together, these examples suggest models for spatial audio that are not merely responsive to individuals but co-constituted through shared activity. A social spatial audio would need to handle multiple listeners with different positions, orientations, and attentional states. It would need to negotiate between them, perhaps generating different renderings for different listeners while maintaining some shared reference. It might enable listeners to point out sounds to each other, to coordinate attention. It might treat listening as a collective activity rather than individual reception.

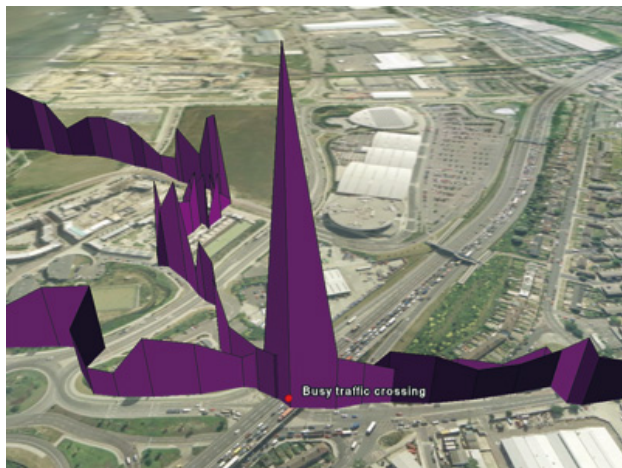


Figure 8: The Greenwich Emotion Map by Christian Nold

4.4 The Old in the Way of the New

What unites the three dimensions above is that they do not demand new technologies. Ambisonics can be decoded relationally, rendering sound fields that shift with listener movement rather than assuming a fixed center. Binaural synthesis can accommodate individual anatomical variation through personalised HRTFs or adaptive filtering. Object-based formats like MPEG-I can support dynamic, multi-user scenarios where rendering differs per listener. Networked audio systems can enable collective listening experiences. The technical capacity for relational, temporal, and social spatial audio is already here. What prevents its realisation is a commitment to the models they were designed to serve. A composer-score-conductor model is embedded in the assumptions of those who build the interfaces, write the standards, and design the workflows. The aural cartographer's dilemma is therefore a problem of imagination.

5 Conclusion

From the unit circle to MPEG-I, spatial audio has refined the map while mistaking it for the territory. Each new paradigm has advanced technical precision while preserving the fundamental relationship: composer specifies and conductor reproduces. What genuinely relational spatial audio systems would require remains an open question. This paper's contribution is diagnostic, not prescriptive. However, the aural cartographer's dilemma will not be solved by better panning algorithms, higher-order ambisonics, or more sophisticated scene graphs. It requires, instead, a reconsideration of what we believe music to be and whom we believe it is for.

6 Ethics Statement

This research adheres to the ethical guidelines established by the NIME community. This paper is theoretical and genealogical in nature, involving no human participants, user studies, or data collection. As such, no ethical approval was required. All sources cited are publicly available academic publications, technical standards, and historical documents. The arguments advanced are based on critical analysis of existing technologies and their representational paradigms, not on empirical investigation involving human subjects. The author has no conflicts of interest to declare.

Acknowledgments

Information Retrieval in Embedded Systems for Audiovisual Artistic Processes (IRESAP) is supported by The Knowledge Foundation (KKS).

References

- [1] V. Ralph Algazi, Richard O. Duda, Dennis M. Thompson, and Carlos Avendano. 2001. The CIPIC HRTF database. In *2001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. 99–102. <https://doi.org/10.1109/ASPA.2001.969552>
- [2] A. J. Berkhout, D. de Vries, and P. Vogel. 1993. Acoustic control by wave field synthesis. *The Journal of the Acoustical Society of America* 93, 5 (1993), 2764–2778. <https://doi.org/10.1121/1.405859>
- [3] Alan Dower Blumlein. 1933. Improvements in and relating to Sound-transmission, Sound-recording and Sound-reproducing Systems. Filed 14 December 1931, accepted 14 June 1933.
- [4] Georgina Born. 2013. *Music, sound and space: Transformations of public and private experience*. Cambridge University Press.
- [5] Pierre Bourdieu. 2020. Outline of a Theory of Practice. In *The new social theory reader*. Routledge, 80–86.
- [6] Thibaut Carpentier, Markus Noisternig, and Olivier Warusfel. 2015. Twenty years of Ircam Spat: looking back, looking forward. In *41st International Computer Music Conference (ICMC)*. 270–277.
- [7] Stefano Catena and Henrik Frisk. 2024. A speaker agnostic approach to spatialisation in electroacoustic music. In *SMC 2024*.
- [8] Jérôme Daniel. 2000. *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia*. Ph. D. Dissertation. Université Paris 6, Paris, France.
- [9] Michel De Certeau. 1984. *The Practice of Everyday Life*. 1984. *Trans. Steven Rendall*. Berkeley: U of California P (1984).
- [10] Dolby Laboratories. 2012. Dolby Atmos White Paper. <http://www.dolby.com/uploadedFiles/Assets/US/Doc/Professional/Dolby-Atmos-Next-Generation-Audio-for-Cinema.pdf>. Accessed: 2026-02-16.
- [11] Nina Sun Eidsheim. 2015. *Sensing sound: Singing and listening as vibrational practice*. Duke University Press.
- [12] Jonathan Fielder. 2016. A history of the development of multichannel speaker arrays for the presentation and diffusion acousmatic music. *Austin, TX* (2016).
- [13] Michael A. Gerzon. 1973. Periphony: With-height sound reproduction. *Journal of the Audio Engineering Society* 21, 1 (1973), 2–10.
- [14] Michael A. Gerzon. 1975. Ambisonics. Part Two: Studio Techniques. *Studio Sound* 17 (1975), 24–30.
- [15] Tony Grajeda. 2015. The 'Sweet Spot': The Technology of Stereo and the Field of Auditorship. *Living Stereo: Histories and Cultures of Multichannel Sound* (2015), 37–64.
- [16] Tomlinson Holman. 2007. *Surround Sound: Up and Running* (2nd ed.). Focal Press, Amsterdam; London. Previous edition published as 5.1 Surround Sound, 2000.
- [17] International Organization for Standardization. 2021. ISO/IEC 23090-4:2021 Information technology — Coded representation of immersive media — Part 4: Immersive audio.
- [18] International Telecommunication Union. 1992. *Recommendation ITU-R BS.775-1: Multichannel stereophonic sound system with and without accompanying picture*. Technical Report. ITU Radiocommunication Assembly, Geneva.
- [19] Alfred Korzybski. 1931. *A non-Aristotelian system and its necessity for rigour in mathematics and physics*.
- [20] Andrew Kuka. 2024. *Outside Lies Magic: Finding the Extraordinary in Everyday Places*. (2024).
- [21] Brandon LaBelle. 2010. *Acoustic Territories: Sound Culture and Everyday Life*. Continuum, New York.
- [22] Jaron Lanier. 2010. *You are not a gadget*. Vintage.
- [23] Thor Magnusson. 2019. *Sonic writing*. (2019).
- [24] Brian C. J. Moore. 1997. *An Introduction to the Psychology of Hearing* (4th ed.). Academic Press, San Diego, CA.
- [25] Henrik Møller, Michael Friis Sørensen, Dorte Hammershøi, and Clemen Boje Jensen. 1995. Head-related transfer functions of human subjects. *Journal of the Audio Engineering Society* 43, 5 (1995), 300–321.
- [26] Christian Nold. 2006. Greenwich emotion map. (2006).
- [27] R Murray Schafer. 1977. *The tuning of the world*. (No Title) (1977).
- [28] Christopher Small. 1998. *Musicking: The meanings of performing and listening*. Wesleyan University Press.
- [29] Jonathan Sterne. 2020. *MP3: The meaning of a format*. Duke University Press.
- [30] Emily Thompson. 2003. *Soundscape of Modernity: Architectural Acoustics and the Culture of Listening in America, 1900-1933*.
- [31] Barry Truax and Gary W Barrett. 2011. Soundscape in a context of acoustic and landscape ecology. *Landscape Ecology* 26, 9 (2011), 1201–1207.
- [32] Salomé Voegelin. 2024. Sonic epistemologies: confrontations with the invisible. *Open Philosophy* 7, 1 (2024), 20240002.
- [33] Peter Wollen. 1989. The situationist international. *New Left Review* 174, 1 (1989), 1.