Organised Sound (2025), **30**, 191–203
doi:10.1017/S1355771825100563

CAMBRIDGE UNIVERSITY PRESS

Article

OTTOsonics: Designing and implementing an open and affordable 3D spatial audio system

Manu Mitterhuber[1], Enrique Tomás[2] and Martin Kaltenbrunner[2]

[1]OTTO Kulturgenossenschaft, Ottensheim, Austria and [2]Tangible Music Lab, University of Arts Linz, Linz, Austria

ABSTRACT

OTTOsonics is an open hardware platform developed by a team of engineers and composers, designed to prioritise affordability and flexibility, addressing the needs of a broader community interested in spatial audio. At its core, the platform features a custom-designed, high-quality 4-inch speaker with a 3D-printed cabinet, and an affordable multichannel power amplifier. It also offers a comprehensive set of mounting accessories and a knowledge base for producing and presenting spatial audio using open-source software. Over the past three years, OTTOsonics has been adopted by multiple cultural initiatives, universities, and audio enthusiasts, enabling the production of new spatial audio works across genres such as electroacoustic, experimental, and pop music. This article outlines the key decisions made throughout the project and presents the technical and artistic outcomes after three years of operation. We discuss the key features of an open platform for spatial audio and how our designs address these needs, as well as future directions for further projects and initiatives.

(Received 13 September 2024; revised 7 April 2025; accepted 19 May 2025; First Published online 1 October 2025)

1. Introduction

Spatial audio has long been a fundamental component of sound-based entertainment, extensively employed in cinemas, theatres and immersive music systems to enhance audience experiences. Each year, cutting-edge spatial audio systems are increasingly installed in specialised venues and digital art spaces worldwide. On the consumer front, spatial audio is also becoming more accessible, with compatible audio hardware now within reach for individuals seeking to install home cinema systems. However, attending an immersive audio concert or a theatrical performance with multichannel audio remains a rare and exceptional experience. Moreover, the majority of music entertainment content continues to be produced in stereo format, and the listening experience in most nightclubs and discotheques rarely exceeds what can be described as an extended mono format.

For general-purpose venues aiming to acquire and install a high-quality multichannel system, such as an ambisonics system with a minimum of 24 audio channels and speakers, budget constraints and resource limitations are common deterrents. Currently, there is no official data, such as academic surveys, that documents the interest of multipurpose venues in adopting and utilising immersive audio formats and further research is needed in this area. What we have observed during the development of the past three years of our project is that many of these venues lack musical content referents to fully understand the possibilities of using immersive audio systems, as well as the technical expertise to, for example, diffuse stereo audio in multichannel formats during a live performance setting. As we will discuss in this article, the introduction of immersive audio formats into general-purpose venues presents a complex challenge encompassing social, economic, cultural, technical and educational factors. Through our project, OTTOsonics,[1] we have sought to partly address these issues by designing and developing an open, adaptable, lower cost and versatile hardware system that has been already adopted by several musical festivals, universities, summer schools and other cultural initiatives (Figure 1).

2. Design goals

Since the 1970s, stereo sound has been the dominant format in recorded music, broadcast radio, television, cinema and online media. While stereophonic systems utilise two independent audio channels to create the illusion of sound emanating from various directions between the two speakers, surround sound systems expand upon this spatial perception by incorporating additional audio channels around the listener (ITU-R BS.1116.1 2015). Although first introduced in the 1940s, surround sound has since been employed by numerous composers in electroacoustic and acousmatic music (Zvonar 2005). Spatialised audio can be created through various techniques, including Wave Field Synthesis (Berkhout et al. 1993), Ambisonics (Gerzon 1972), Vector and Direction Based Amplitude Panning (Pulkki 1997; Lossius et al. 2009), and surround recording techniques such as Virtual Microphone Control (Peters et al. 2008) and other diffusion techniques (Mooney 2005).

Surround sound rendering is currently performed primarily through software on computers, such as panning plugins within digital audio workstations and standalone applications. As many of

Corresponding author: Enrique Tomás; Email: enrique.tomas@kunstuni-linz.at

Cite this article: Mitterhuber M, Tomás E, Kaltenbrunner M. (2025). OTTOsonics: Designing and implementing an open and affordable 3D spatial audio system. *Organised Sound* **30**: 191–203, https://doi.org/10.1017/S1355771825100563

© The Author(s), 2025. Published by Cambridge University Press.

[1]The websites of OTTOsonics are www.ottosonics.com and https://tamlab.kunstuni-linz.at/projects/ottosonics/

https://doi.org/10.1017/S1355771825100563 Published online by Cambridge University Press



Figure 1. Left: a concert with the OTTOsonics platform at our main venue Alter Bauhof in Ottensheim, Austria, equipped with more than 100 channels. Right: the production studio of the Tangible Music Lab with a 23.1 system.

these programs are available as open-source software,[2] there has been a move towards standardising production and presentation workflows among artists, researchers and sound technicians. For instance, when using Ambisonics renderers, a venue only needs to provide information about the speaker layout and supply a simple JSON file with the specific Ambisonics decoder configuration that the performer can load into a plugin. This greatly facilitates the sharing and compatibility of music among venues.

To reproduce multichannel sound from a computer, particularly with setups involving many audio channels, several essential hardware components are required. These include an audio interface capable of handling the required number of channels, an equal number of active speakers (or passive speakers powered by multichannel amplifiers), speaker mounting systems, and usually hundreds of metres of audio cables. The acquisition of this hardware necessitates a substantial budget. Thus, while audio rendering can now be accomplished with open-source software, the cost of reproduction hardware remains the primary financial constraint in implementing surround sound systems.

In light of this context, we decided to fully investigate the factors that influence the feasibility and desirability of implementing surround sound systems across various community types. Based on our activities over the past three years, we identified several key groups with a vested interest in such systems: mid-size general-purpose music venues, theatres, music ensembles and bands, computer music departments at universities, independent artists and musicians, music festivals, and educators. Through direct communication with representatives from these groups during the last two years, we identified some key factors for the successful introduction of surround sound systems:

- *Economic constraints*: many venues face significant financial challenges, particularly in the aftermath of the COVID-19 pandemic, limiting their ability to make substantial investments.
- *Architectural limitations*: physical space within venues is often insufficient for the installation of large or heavy speakers, especially when a system requiring, for example, 24 speakers or more is proposed.
- *Technical expertise*: sound technicians and studio personnel typically lack training in specialised multichannel formats such as Ambisonics.
- *Content availability*: there is a scarcity of music produced in immersive audio formats, necessitating that venues either adapt existing content or produce new material specifically for these systems.
- *Dependency on commercial licences and formats*: the adoption of immersive audio hardware systems sometimes requires venues to rely on costly commercial licences and specific formats.
- *Accessibility*: spatial audio production is only available at a small number of very specialised studios such as universities and professional sound studios. They usually are expensive to rent or difficult to access for long periods of time.

Despite these challenges, we have observed that many cultural initiatives communicate a genuine interest in immersive audio formats. However, they often struggle to overcome these obstacles without resorting to expensive systems or hiring highly specialised personnel. Our conclusion was that further development was needed to make immersive audio more accessible to a broader range of venues. We decided to work around the possibility of a highly adaptable system, with a relatively low cost, but good sound quality, to investigate if venues would be interested in adopting such a system.

Inspiration for this challenge was taken from the work done by previous researchers and audio enthusiasts who previously developed open and flexible 3D spatial audio systems. Lopez-Lezcano (2014) identifies the main goals for an open sound diffusion system: simplicity, transparency, versatility, use of commodity hardware, free software and minimum footprint of the equipment in the listening area. Faerber and Kocher (2010) built a mobile Ambisonics system to be 'much easier to handle, and drastically reduce the amount of time and work used for its installation'. Barbour designed a 3D spatial sound facility with the goal of unlimited access to researchers in a small domestic room, with a secondary design goal to be 'scalable to adapt to a larger space for performances and to an outdoor environment for public sound sculptures' (Barbour 2013). Smith and Cox (2016) propose a portable immersive intermedia system with the following design priorities: ease to transport, ease to set up, ease of tear down and ease of maintenance. Despite these successful developments, none of these sound diffusion systems was used by other initiatives than

[2]Two open-source suites of plugins are massively used. First the Ambisonics suite by the Institute of Electronic Music in Graz (https://plugins.iem.at/) and the plugins developed by researchers of the Aalto University (https://leomccormack.github.io/sparta-site/docs/plugins/sparta-suite/).

https://doi.org/10.1017/S1355771825100563 Published online by Cambridge University Press

by the authors. They were never evaluated in the context of the needs of other users. For this reason, we decided to orient our project towards an open and highly accessible design that others could easily replicate.

We propose the following key dimensions for the development of such suitable open systems:

- *Architectural adaptability*: the speakers should be as compact and lightweight as possible, allowing for easy installation and orientation in any direction.
- *Loudness*: the system must be capable of supporting loud concerts at a mid-size venue, with a typical listener distance capped at approximately 15 metres. This makes it suitable for electroacoustic music performances and mid-size pop/rock/techno venues.
- *Modularity*: while a typical standard configuration would consist of a minimum of 24 channels/speakers to assure immersive sound, the system should be scalable to support up to hundreds of speakers.
- *Cost*: the manufacturing cost for a typical 24-speaker system should be affordable.
- *Compatibility with commercial hardware and software*: musicians use all types of hardware and software solutions. The system must allow total compatibility.
- *Hackability*: both the software and the hardware architectures, along with as many components as possible, should be open-licensed, customisable, and hackable.
- *Fabricability*: the systems must be easy to produce by others without our direct involvement. These suggested the use of digital fabrication tools, such as 3D-printed enclosures and mounting systems, as well as amplifiers which could be produced in single layer PBCs.
- *Versatility*: our speaker system should not be designed for only one use or location.

Guided by these patterns, the first author, Manu Mitterhuber, explored the feasibility of designing and producing our own open hardware loudspeakers and multichannel amplifiers. The second and third authors, Enrique Tomás and Martin Kaltenbrunner, worked on defining suitable open network audio architectures, and on programming and exploring various open-source renderers solutions and presentation formats. The specifics of the OTTOsonics platform are detailed and discussed in the following section. As we described before in this article, the whole system has been released as open hardware and software: all plans and code can be downloaded from our affiliated websites.[3]

3. The OTTOsonics platform

3.1. Hardware: general design overview

The OTTOsonics loudspeaker (Figure 2) is built around a 4-inch professional driver manufactured by Faital. Based on the specific requirements of our project, we opted for 3D-printed speaker cabinets. While this approach has been explored by other researchers (Drack, Zotter and Barrett 2020), it has revealed challenges such as wall rigidity and material resonance frequencies. After several iterations, we designed a custom 3D-printed cabinet optimised for this driver. Our design delivers high-quality sound, and is both cost-effective and easy to produce. Detailed specifications of the loudspeaker are provided in the following subsection.

[3]For accessing all technical plans and files, please visit the links from our websites at https://tamlab.kunstuni-linz.at/projects/ottosonics/ and www.alterbauhof.at/ottosonics/.

The amplifier (Figure 2) is based on the TDA9294 integrated circuit, chosen for its simplicity and excellent sound as a Class A/B amplifier. Initially, we used a 12-channel system in a heavy 2U steel enclosure for early tests. After building 12 units, we switched to custom-designed enclosures from 1.5 mm aluminium sheets, cut with a waterjet, reducing weight and optimising the internal layout. This allowed us to create 16- and 24-channel amplifiers without increasing unit size. The fully analogue amplifiers feature balanced XLR inputs and Speakon connectors, chosen for their high power, ease of assembly and safety in time-sensitive setups. By integrating amplifiers, power supply and audio interfaces into a 5U rack, and using Euro containers for the speakers, we developed a portable 24-channel system that fits in a car and can be transported by train, bus, or plane. More details are provided in the following sections.

3.2. Loudspeaker design

The first decision to use passive speakers was primarily driven by the need for small, lightweight loudspeakers that are easy to install and transport. By using a single speaker cable for each speaker, we not only reduced costs – especially since setups often require more than 400 metres of cabling – but also gained flexibility, allowing us to easily shorten, extend and reuse cables. Removing all active components from the speakers further reduced their weight, enabling us to house multiple amplifiers within a single enclosure powered by a single source. The OTTOsonics speaker is shown in Figures 2 and 3.

Loudspeakers employ a range of different mechanisms for transforming electrical energy into sound pressure (e.g., dynamic vs. electrostatic transducers, ribbon transducers) and even the most commonly used dynamic loudspeakers provide a huge variety of electrical, mechanical and acoustic qualities that have to be considered and matched with the basic concept of a speaker. Again we set a few basic requirements to navigate through all those possibilities:

- *Frequency response* (*FR*): a linear FR is oftentimes achieved through a combination of different drivers for different frequency bands plus either electronically or digitally employed crossovers and filters. Since multichannel rendering is generally happening on computers, we simply skipped hardware filters entirely in order to implement any necessary filter later within a software. Furthermore while a 2- or 3-way speaker design – typical for studio monitor and coaxial speakers – with one or more crossovers can be beneficial for linearity and dispersion across frequency bands, it also creates problems in other areas such as phase alignment and scalability and makes the building of speakers more complicated and expensive.
- *Scalability*: one of the major advances in sound reinforcement was the introduction of line arrays, which offer a significant advantage over point sources by providing lower and adjustable energy loss over distance. This makes them the preferred choice for large-scale events. For a line array to function properly, all frequencies must be emitted consistently throughout the array, achievable through 1-way full-range speakers placed as close together as possible. Using coaxial speakers or 2-way monitors can cause phase cancellations and poor results. The reduced energy loss is especially useful for creating an immersive sound experience in multichannel systems with diameters exceeding 10 metres.

https://doi.org/10.1017/S1355771825100563 Published online by Cambridge University Press



Figure 2. Left: the OTTOsonics loudspeaker. Right: a rack with a 60-channel OTTOsonics amplifier.

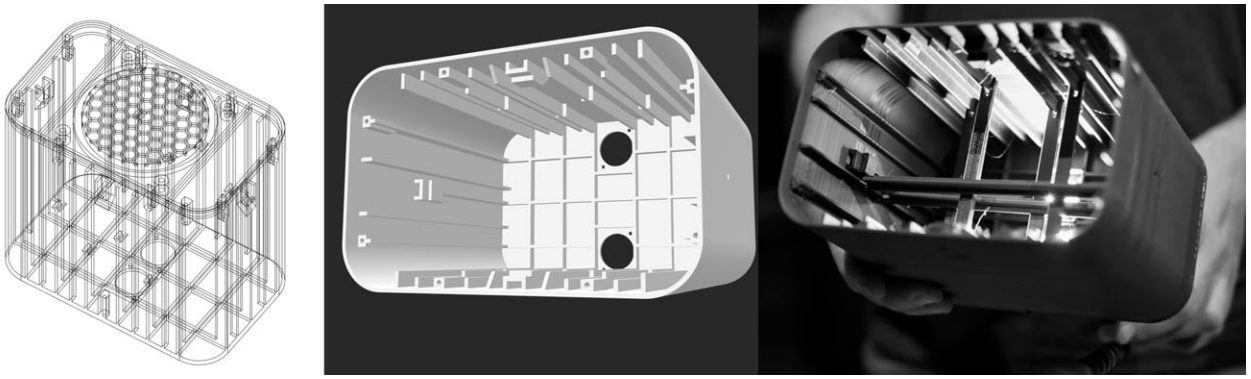


Figure 3. Left, middle: loudspeaker design. Right: 3D-printed result with struts.

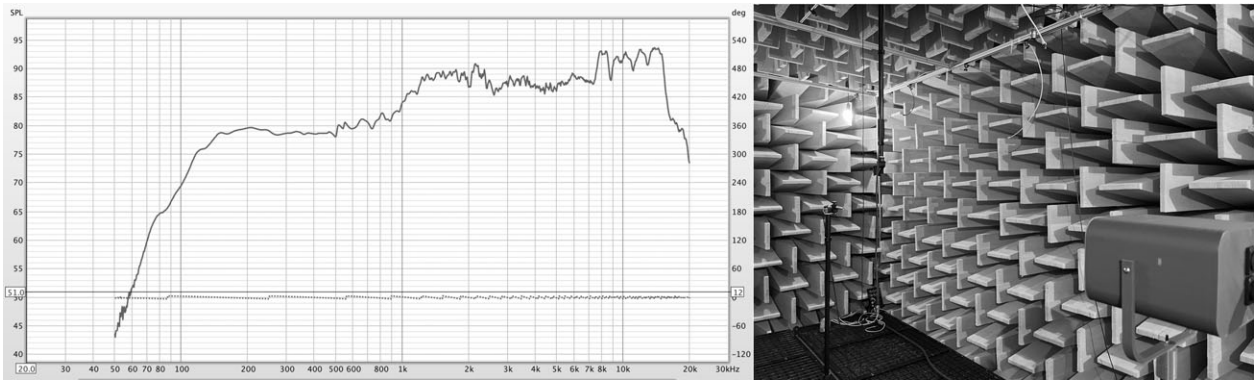


Figure 4. Frequency response at 0° measured at the anechoic chamber in MDW University Vienna.

The second advantage of scalable line sources is their ability to produce higher output than a single speaker of the same size. This extends their potential applications from studio setups and sound installations to live sound reinforcement for larger audiences, following the principle that doubling the number of speakers increases sound pressure level (SPL) by 6dB. Examples of our line arrays can be observed in Figure 4.

- *Lower cutoff frequency/crossover*: directional hearing is ineffective for low frequencies. Below 200 Hz, interaural level differences become imperceptible, though interaural time differences can still provide some cues about sound location. This effect fades entirely around 80 Hz. Therefore, speakers for spatial audio systems do not need to reach the low frequencies typical of studio monitors. This allows for a

https://doi.org/10.1017/S1355771825100563 Published online by Cambridge University Press

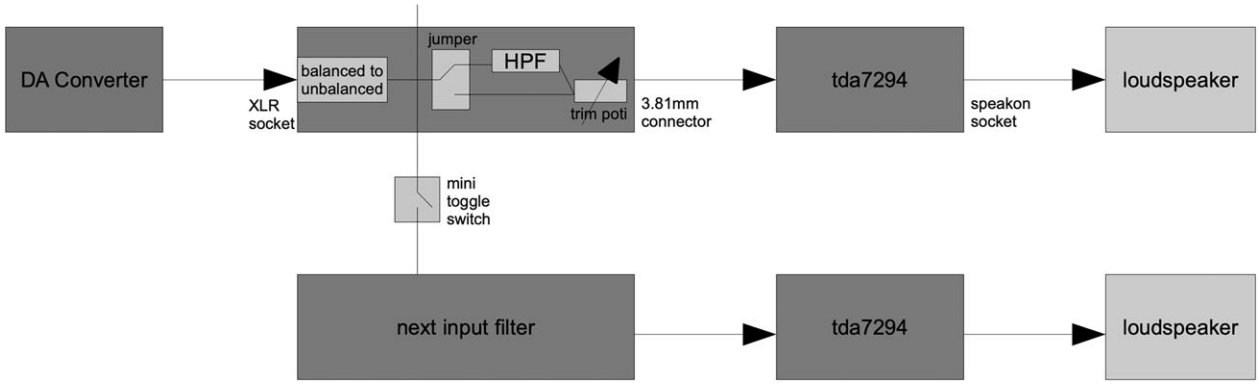


Figure 5. OTTOsonics amplifier: design and signal flow.

smaller, lighter design that conserves energy for higher frequencies. To determine the optimal speaker size and enclosure tuning, we simulated 3- and 4-inch drivers using Thiele-Small parameters[4] (Margolis and Small 1981) and Boxsim[5] software. We printed various enclosures, conducted acoustic measurements and performed extensive listening tests. Our results favoured 4-inch drivers due to their higher sensitivity and lower resonant frequency, yielding a linear frequency response down to 120 Hz. We selected the Faital 4FE32 16 Ohm driver, iteratively refining the design through simulation, printing, measurement and listening. Reducing the enclosure volume increased membrane damping due to air compression, raising the low-frequency rolloff. Smaller enclosures (1.5–2 litres) caused phase issues when used with 18-inch subwoofers in our venue, particularly around 160 Hz. Larger enclosures (2.8 litres) with a crossover frequency near 120 Hz significantly reduced these phase problems. Hence our OTTOsonics speakers are optimised for frequencies above 120 Hz and require an additional subwoofer for the non-directional lower frequencies. For economic reasons, the OTTOsonics project does not include an additional low-frequency speaker design, since subwoofers are commonly available at affordable prices.

- *Resonances*: speakers for high sound pressure levels have traditionally been made from low-resonance materials such as birch plywood or MDF. When using PLA, a common and affordable material for 3D printing, and opting for thin 2 mm walls to reduce weight and print time, we encountered significant and visible wall resonances. Since these movements were perpendicular to the walls, we added internal structures for more effective damping than simply thickening the walls. Additionally, we connected opposing walls, which naturally resonate out of phase, with hot-glued struts, reducing their movement through mutual cancellation (Figure 3). The ribs, struts and structures at the speaker's base were positioned asymmetrically to prevent higher resonant frequencies from aligning. The width, height and depth were also selected to avoid simple rational multiples, minimising matching room modes.
- *Overall dimensions and weight*: considering the preceding factors, the overall dimensions were chosen as a compromise to meet multiple requirements: avoiding room modes, minimising one dimension for use as line sources, achieving a volume of at least 2.8 litres, fitting the speakers into Eurobox containers, and maintaining a size that allows easy printing on most commercial 3D printers. The driver, with its neodymium magnet, weighs 270g, and combined with the printed parts and sockets, the total weight of 785g enables quick installations without requiring extensive safety precautions.
- *Connectivity*: each speaker is equipped with two parallel-connected Speakon sockets, allowing multiple speakers to be linked with short patch cables for line source setups. With a nominal impedance of 16 Ohms per speaker, connecting two, three, or four speakers reduces the impedance to 8, 5.3 and 4 Ohms, respectively. This remains manageable for most amplifiers, making the amplifier's maximum power output the limiting factor in our case.
- *Frequency correction*: the measured frequency response in an anechoic chamber at MDW University in Vienna yielded the result shown in Figure 5. For each channel, we correct the frequency response using a 2-band digital EQ[6] set at −8dB at 1700 Hz and −3dB at 5kH as a starting point, and usually further EQ according to the audio material and the room. For each possible type of array starting from two to eight or even more speakers the frequency correction has to be changed.
- *Horizontal dispersion*: owing to their size, 4-inch drivers exhibit increasingly narrow directivity at higher frequencies. To compensate for this, we use stacked speakers using a 4 mm nut at the top and bottom of the membrane. This lets us choose between a line array's long-throw effect or a wider horizontal dispersion. Our experience shows that, particularly at ground level, angling two speakers at 45° is effective due to the proximity of listeners. For larger distances, such as at outdoor concerts, a setup with multiple subwoofers in front of the stage and two slightly inclined line arrays per side (angled at 30° with eight speakers per array) provided ample sound pressure and frequency coverage. These stacked speakers can be observed in Figures 6 and 9.

3.3. Amplifier design

As a first step, we locked down a few basic criteria for choosing a suitable amplifier circuit: very low distortion, very low noise floor

[4]Thiele-Small parameters (commonly abbreviated T/S parameters, or TSP) are a set of electromechanical parameters that define the specified low-frequency performance of a loudspeaker driver. These parameters are published in specification sheets by driver manufacturers so that designers have a guide in selecting off-the-shelf drivers for loudspeaker designs.

[5]Boxsim is a software created by Visaton for the simulation of hi-fi speakers: www.visaton.de/en/literature-software/software.

[6]For this EQ we use the IEM open source plugin MultiEQ which allows 64 channels per instance.

https://doi.org/10.1017/S1355771825100563 Published online by Cambridge University Press

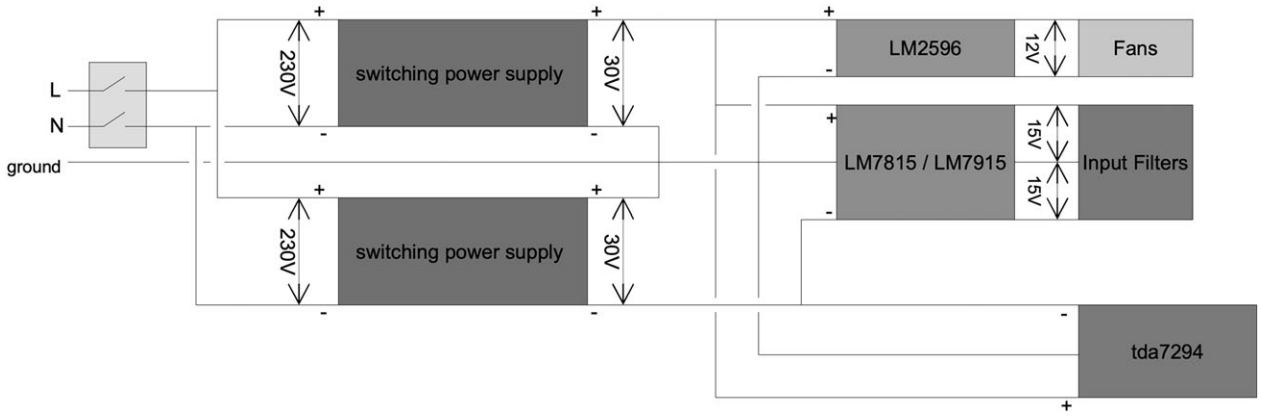


Figure 6. OTTOsonics power supply design.

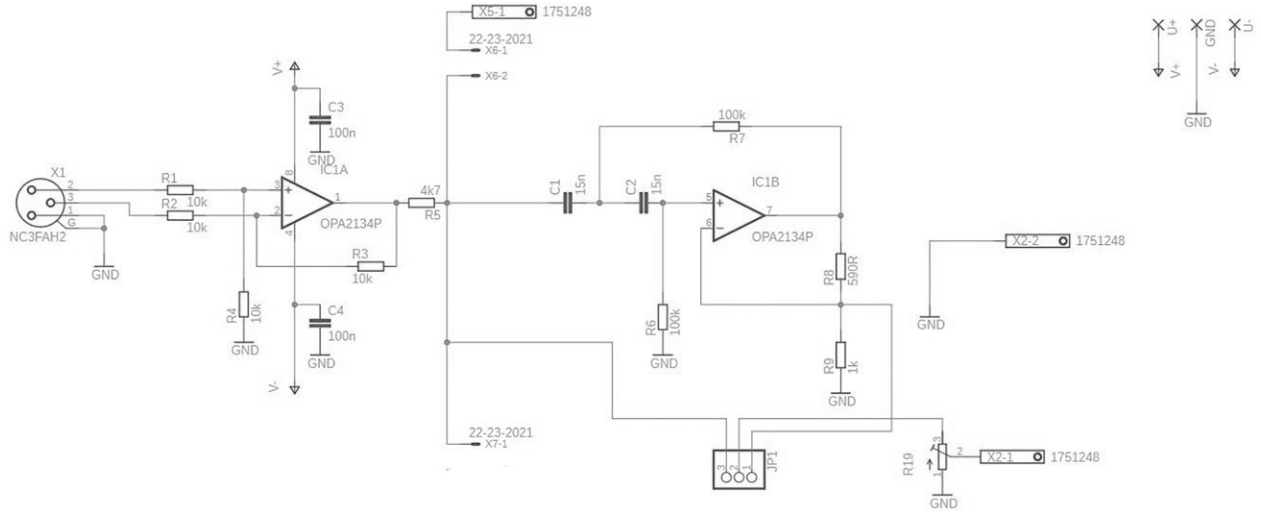


Figure 7. Input filter design.

which otherwise would be especially annoying in multichannel setups, high operating voltage and high output power enabling it to power one or more speakers at the same time, and in general highest possible sound quality at an affordable price. For all these reasons we chose the integrated circuit TDA7294[7] by ST, a class A/B analogue amplifier that we already knew sounded excellent and fulfilled our needs. It can be supplied with up to ± 40VDC delivering a maximum of 100W. Given the maximum power handling of the Faital 4FE32 of 30W RMS in combination with 16 Ohm nominal impedance, we tuned the power supply to ± 30VDC resulting in a maximum of

$$P = U_{eff}{}^2/R = (30/sqrt(2))^2/16 = 28.125W$$

This also allowed for a usage of two or three speakers connected in parallel, in which case the impedance would be reduced to 8 or 5.3 Ohm resulting in 56W or 85W provided by the amp and making the system more modular and flexible. The block diagram of the amplifier is shown in Figure 7.

[7]TDA7294 datasheet: www.st.com/resource/en/datasheet/tda7294.pdf (accessed 15 September 2024).

Providing power for up to 24 TDA7294 chips required a special solution since both positive and negative voltage were needed. While some boards could accept AC from heavy toroidal transformers due to built-in diodes and capacitors, smaller boards lacked this feature. Additionally, toroidal transformers are bulky and costly, so we opted to test lightweight, affordable switching power supplies capable of high currents. Concerns about HF electromagnetic noise and unstable power were mitigated by isolating the power supplies in a separate metal enclosure, which proved effective. We connected two 1200W power supplies in series with the middle contact grounded, creating a −30, 0 and +30 VDC supply, providing a total of 2400W (Figure 8). Factoring in the manufacturer's recommendation to keep the load under 80%, and the system's efficiency, we estimated this setup could power up to 36 speakers (30W each) via a 12-channel amp (three speakers/channel), a 16-channel amp (two speakers/channel), or a 24-channel amp (one speaker/channel).

After testing and evaluating all components, we finalised the amplifier layout (Figure 5). Given that Class A/B amplifiers operate at around 60 per cent efficiency, proper heat dissipation was necessary. U-shaped aluminium heat sinks with 50 mm silent fans offered an affordable and effective solution while also allowing the TDA7294 boards to be mounted directly through the integrated circuit.

https://doi.org/10.1017/S1355771825100563 Published online by Cambridge University Press



Figure 8. Different mounting options with the OTTOsonics speakers.



Figure 9. OTTOsonics line arrays with six and eight speakers.

Typical signals came from audio interfaces with balanced line-level outputs. We needed to phase-invert one signal and combine it with the other for unbalancing. Additionally, we aimed to protect the 4-inch drivers from low frequencies and ensure low impedance before the power amp stage. We selected an OPA2134PA operational amplifier: one side handled the balanced-to-unbalanced conversion, while the other provided a high-pass filter fixed at 110 Hz (Figure 9). We added features such as a filter bypass via a jumper, a bridge function to route one input to multiple power amps, and a precision potentiometer to adjust the output level for optimal signal-to-noise ratio and distortion-free input. A PCB was designed and manufactured for two audio channels, incorporating two XLR sockets for easy mounting into the amp enclosure.

3.4. Mounting accessories

Positioning speakers in 3D space based on predefined coordinates, directed at a specific point, typically requires multiple hanging points and expensive flying accessories. Taking advantage of the reduced weight of our speakers, we developed a set of 3D-printed hangers and clamps (Figure 8). The hangers consist of a U-shaped piece and a top piece, connected via M4 nuts inserted and glued at the speaker's centre of mass, allowing two-axis rotation. The top piece, with two holes, enables the speaker to hang between two points, offering rotational stability and flexible positioning for different spaces and installations. The top piece can also be swapped for a clamp to attach the speaker to a 50 mm pipe or truss.

For speakers at ear height or in multipurpose rooms where removal is necessary, we designed an adapter for microphone stands, with a sturdier version available for two speakers. A different solution was needed for hanging line arrays of three or more speakers (Figure 9). Their increased weight is supported by a 3D-printed bar at the array's base. Each speaker is connected via headless 14 mm M4 threads, secured with an M4 bolt to the bar and a top piece with multiple holes for adjustable inclination. A 4 mm nylon paracord runs through the top and bottom pieces, with 3D-printed tensioners at both ends for quick height adjustments.

3.5. Audio network architecture and interface connectivity

The typical OTTOsonics spatial audio system features a hemisphere with 24 or more symmetrically arranged speakers in three to four concentric rings. For systems with 48+ speakers, networked audio solutions using Ethernet cables are preferred, saving time and costs. While various protocols exist, such as Cobranet and Dante, Audio Video Bridging (AVB) has become our platform's standard. AVB, defined by IEEE standards, ensures synchronised Ethernet communication with low latency, making it ideal for

https://doi.org/10.1017/S1355771825100563 Published online by Cambridge University Press

professional audio. Although AVB is still emerging and commercial products are limited, it offers the potential for broader adoption with affordable networking devices.

Most AVB-compatible products are network switches or audio-specific endpoints. We integrated MOTU AVB interfaces but identified their use as a limitation for our open hardware system. Thus, we are developing a custom multichannel AVB digital-to-analogue converter to integrate with our amplifiers, allowing computers to connect directly to AVB endpoints via Ethernet without relying on costly commercial hardware. A typical setup requires digital audio playback and live audio capture from both analogue and digital sources, with flexible routing across the network. Our standard audio network includes computers for sound playback and rendering, AVB analogue input/output interfaces, AVB network switches, professional digital mixers, volume controllers and WLAN gateways. Computers connect via USB or Ethernet, and we use modular 19-inch rack modules containing MOTU 24AO interfaces and 24-channel amplifiers. This architecture was tested in various musical contexts, implementing up to 120 channels during live diffusion of analogue audio from acoustic instruments at music festivals. The system exhibited no perceivable latency, and the overall architecture proved to be both technically reliable and robust.

3.6. Spatial audio rendering software

The OTTOsonics project relies heavily on two open-source plugin suites, IEM and SPARTA (introduced in section 2) for Ambisonics and VBAP audio rendering. Key plugins used include multichannel encoders, Ambisonics decoders, B-format converters, impulse response encoders, and Ambisonic effects (delays, compression, reverb, granular spatialisation). These tools run on the Reaper DAW, chosen for its support of 128 channels per track, enabling Ambisonics encoding up to the 10th order, though we typically use 7th order to save CPU and storage.

An alternative fully open-source setup was employed in Pure Data (Pd), using the *snake~* extension to host VST plugins and connect them modularly with a single multichannel output, offering a fully open-source solution without the need for Reaper.

3.7. Fabrication process and production cost

Detailed instructions for producing OTTOsonics speakers, amplifiers, and power supplies are available on our website under the 'knowledge base' menu. This includes photos, parts lists, and files for fabrication:

- STL files for 3D printing speakers and accessories
- Eagle files for PCB production
- DXF files for CNC cutting metal enclosures.

We use standard PLA with a 1 mm nozzle to reduce printing time, adjusting slicer settings as needed. The speaker enclosure is designed for a Faital 4FE32 4-inch driver, though swapping drivers may produce unpredictable results. PCBs can be ordered online, and CNC cutting can be done via waterjet or laser cutting, with files usable for various machines. Alternatively, metal enclosures can be drilled manually for input/output sockets. Essential tools include a drill, soldering iron, multimeter, hot glue gun, M3 and M4 thread cutters, and metal drills. Basic soldering and electronics knowledge is required to troubleshoot and ensure safe assembly.

The approximate material cost of a 24-channel OTTOsonics system in Europe in 2024 (without including subwoofer, audio interface and computer) varies based on supplier and market conditions. This is a rough estimate in Euros: 24 speaker drivers (800); 24 channels amplifier (500); power supply (200); aluminium enclosures (150); cables and plugs (150). The total cost is 1,800 Euro.

4. Implementation

In this section we describe the different application areas of our project. At each of these case studies we discuss the specificities of implementing spatial audio systems, the feedback from the stakeholders involved, and what we have learned about each of the case studies.

4.1. For multichannel electroacoustic and acousmatic music

Initial tests of OTTOsonics were conducted during artist residencies with electroacoustic music composers. In May 2022, Mariam Gviniashvili, who usually works at the Ambisonics studio at NOTAM in Oslo, spent a week with a hemisphere of 30 speakers. Following her, other composers, including Åke Parmerud, Ida Hiršenfelder, and Amélie Nilles, also used the system in 2022, each concluding their residencies with a concert (Figure 10).

Over two years, more than 50 composers have presented music with OTTOsonics at workshops, concerts, workshops, and summer schools. The system was also featured at the Ars Electronica Festival in Linz, Austria, in 2022, 2023 and 2024. From our experience, OTTOsonics's wide dynamic range and precise sound positioning suit electroacoustic music needs. For venues of about 100 people within a 15-metre speaker hemisphere, the system offers sufficient loudness. In rectangular setups, we use two or three speakers in line arrays to address differences in perceived sound pressure.

4.2. For sound reinforcement of rock, pop, electronica bands

The OTTOsonics platform has been used for sound reinforcement in over 40 concerts featuring rock, pop, hip hop and electronic bands, held in indoor venues with capacities ranging from 100 to 500 people. Ambisonics rendering through a multichannel system offers more accurate sound localisation over a larger sweet spot compared with traditional stereo PA systems, enhancing the connection between visual and auditory elements for live performances.

Using 15 line arrays with six speakers each (Figure 9), surrounding the audience, plus 20 overhead speakers, we delivered sufficient sound pressure for various genres (e.g., rock, pop, electronica, DJ sets). In smaller setups, we combined a powerful stereo PA for mono signals (e.g., bass drum, snare, bass, vocals) with a multichannel system for guitars, synthesisers and effects. Ambisonics also achieves better audio separation and a widening effect, especially useful for effects such as delays and reverb, as well as instruments such as guitars and keyboards that are typically panned hard left or right in studio productions.

We also experimented with eliminating the stage entirely, allowing artists to perform within the audience without monitors, assuming they could hear themselves throughout the room. The resulting sound felt as though it was coming from all directions within the hemisphere, creating a novel listening experience. Informal audience feedback emphasised the comfort and immersion of the sound, which, while loud overall, was not overwhelming due to the even distribution across numerous speakers. Many found the

https://doi.org/10.1017/S1355771825100563 Published online by Cambridge University Press



Figure 10. Left: electroacoustic music concert at the end of the Phonon Ambisonics Summer School in July 2023. Right: concert for album release of the band Laikka at WEST Space Vienna in June 2023.



Figure 11. Measuring the impulse response of Brucknerhaus Linz, and implementation of virtual acoustics for a string quartet concert at our venue Alter Bauhof.

sensation of sound coming from above and around them refreshing compared with the typical frontal setup.

Beyond the technical aspects of sound reinforcement, our artistic goals in these concerts focused on enhancing expressivity through spatialisation. Since most of the musicians and bands had no prior experience with spatial techniques, we integrated spatial trajectories into the musical structures and emotional moments of the performances. Given the typically limited soundcheck time, we developed a suite of ready-to-use spatial effects, including signal fading between positions, dynamic sound movement at variable speeds, and built-in effects such as reverb and delay, all controllable via MIDI interfaces. During longer soundchecks, we collaborated with the artists to design specific *spatial presets* and *spatial trajectories* for key moments in the concert, such as abrupt changes in sound source positions and variations in trajectory speed.

4.3. For virtual acoustics rendering at classical music concerts

Virtual acoustics rendering uses real-time room impulse responses to simulate different acoustic environments (Wefers et al. 2014). In our project, this technology was primarily used to improve the dry acoustics of our venue, making it more suitable for brass ensembles and string quartets, and also as an expressive audio effect during live performances. For these, the audience sat beneath a speaker hemisphere, with the ensemble nearby (Figure 11).

We used the open-source MCFX plugin suite by Matthias Kronlachner (2014), particularly the *mcfx_convolver*, which encodes mono sources into ambisonic format using the Ambisonics impulse response. We also developed a programming framework for measuring and calculating multichannel impulse responses with Ambisonics microphones, now available on our GitHub.[8] This framework has been applied in four performances: one brass quintet, two string quartets, and one string quintet. During these, we altered the impulse response and its intensity in real time, changing the level of reverberation and spatial reflections, while also applying sound filtering based on the properties of the spaces where the impulse responses were recorded. In one anecdotical instance, a string quartet initially believed the enhanced acoustics were natural and were eventually surprised how different the venue actually sounded when the system was turned off.

[8]The python code of this framework together with a tutorial in form of Jupyter notebook can be downloaded from https://github.com/tamlablinz/Multichannel-IR-with-Jupyter.

https://doi.org/10.1017/S1355771825100563 Published online by Cambridge University Press

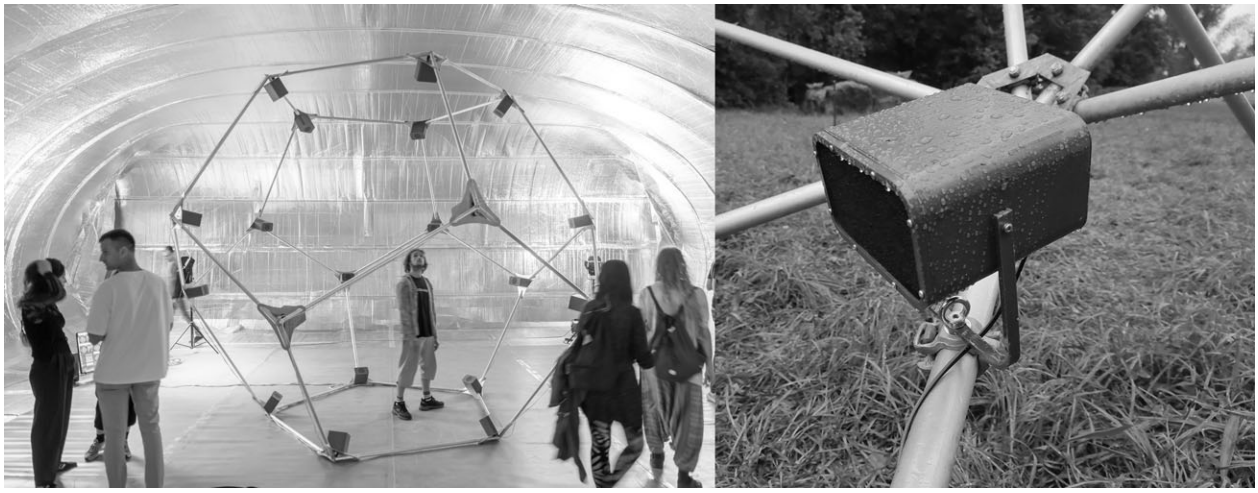


Figure 12. Left: dodekaOTTO (photo credit: Ales Rosa). Right: waterproof OTTOsonics speaker.

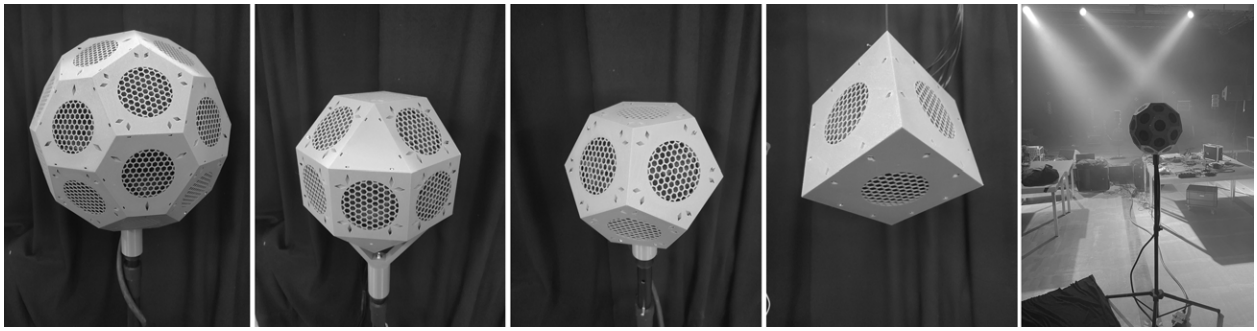


Figure 13. Speaker arrays created with OTTOsonics with 31, 18, 11 and 6 drivers. Right setup for performance during Klangfestival 2024.

4.4. For creating public space listening sculptures

After gaining experience with our system indoors, we explored designing a 3D spatial audio system for outdoor public spaces. We chose a dodecahedron for its sculptural aesthetic and created the dodekaOTTO, a spatial audio sculpture made of aluminium. The dodekaOTTO features 20 OTTOsonics speakers arranged at the CNC-cut corners in four rings of five speakers each. Measuring 1.40 metres per side and standing just over 4 metres tall, it uses the same OTTOsonics amplifier and subwoofer as our indoor setups. For heavy weather environments we created a waterproof version of the speaker too (Figure 12).

Since spring 2023 we built two dodekaOTTOs with the design support of Ben Wesch, with three more constructed in Slovenia, France and Greece. Plans for building your own dodekaOTTO structure are available on our website as open design.[9] The system excels in reproducing 3D spatial audio in outdoor settings, accommodating about 10 people inside the sculpture with speakers positioned close to the audience. In July 2024, we held a hybrid soundscape composition workshop at the Summer Academy of the Ionian University in Corfu, Greece, where students created works to blend with the site's existing soundscape. Additionally, the Slovenian dodekaOTTO was used by Mauricio Valdés of PINA for workshops and electroacoustic music presentations at the Ars Electronica Festival 2024 in Linz.

[9]This technical information is accessible at: https://tamlab.kunstuni-linz.at/projects/dodekaotto/.

4.5. For creating speaker arrays

Inspired by commercial icosahedral speaker arrays such as the IKO (Zotter et al. 2017) and other smaller portable versions (Drack et al. 2020), we decided to combine our speakers to create spatial audio systems capable of diffusing sound from positions that do not necessarily surround the audience. Our goal was to 3D print structures that could house the same Faital 4-inch loudspeakers. After a few iterations, we have successfully created arrays consisting of 6, 11, 18 and 31 speakers, as illustrated in Figure 13.

These speaker arrays have been used for three primary purposes. First, for reproducing sweep tones during Ambisonics impulse response recordings. Second, as stage monitors for musicians. Third, drawing inspiration from the work of Gerriet K. Sharma (2016), for spatialising audio in acoustically complex environments where reflections from walls and other irregular surfaces generate intriguing sound effects by altering the trajectories of the reflected waves. Composer Enrique Tomás successfully spatialised the live performance of the Postdigital Lutherie Ensemble using the 31-speaker array during Klang Festival 2024 in Gallneukirchen. For this, we encoded the audio from each ensemble member in different

https://doi.org/10.1017/S1355771825100563 Published online by Cambridge University Press

spatial directions and developed various physical model algorithms to rotate and attract the virtual sound positions, enhancing the spatial experience of the performance. Although further scientific and artistic exploration of the acoustic properties of these speaker arrays is necessary, we view our designs as a promising alternative for constructing frameworks to facilitate research on speaker arrays.

4.6 Resume of initiatives equipped with OTTOsonics

By September 2024, OTTOsonics has been adopted by the following projects and venues as a permanent facility:

- Alter Bauhof Ottensheim: 96 channels (ch) concert room, 10 ch studio, 1 × 16 ch, 2 × 24 ch, and 2 × 32 ch mobile systems, 20 ch dodekaOTTO
- Tangible Music Lab, University of Arts Linz: 24 ch studio, 48 ch auditorium, 20 dodekaOTTO
- Kulturpool Gallneukirchen: 32 ch concert room
- Theater phönix, Linz: 24 ch sound installation, 24 ch theatre room
- Applied Science University Upper Austria: 24 ch studio
- Black Page Orchestra Vienna: 24 ch studio
- PINA Cultural Association Koper, Slovenia: 24 ch system, 20 ch dodekaOTTO
- St Joseph University Macau, China: 32 ch studio
- Ionian University Corfu, Greece: 24 ch. system, dodekaOTTO
- Britt Hatzius Theater Company Heidelberg & Brussels: 24 ch system
- Phonon Cultural Association Usti nad Labem, Czech Republic: 16 ch system
- Kulturverein Kapu, Linz: 16 ch system
- Musrara University Jerusalem, Israel: 12 ch system.

We expect that two more systems will be fabricated at universities in the UK and Hungary before the end of 2024.

5. Discussion

5.1. Openness and affordability as main features

The market for loudspeakers and audio amplifiers offers a vast range of options, from low-cost models to high-end, professional systems. However, OTTOsonics has garnered significant interest from prominent university departments, popular music venues and renowned composers. In our view, within this saturated market, OTTOsonics stands out as an open hardware alternative developed by a small community of enthusiasts, addressing the needs of a broader community interested in spatial audio. The project has never been driven by commercial interests but rather by a personal commitment to creating a platform that first and foremost serves our own needs for producing and presenting 3D audio in our venues. This motivation led us to design a system that prioritises certain features, particularly affordability and flexibility.

While the OTTOsonics community of musicians acknowledges that there are more precise and technically more developed systems available, they also recognise that the cost of acquiring a full set of 24 or more professional speakers and corresponding amplifiers is beyond their financial means. Our evaluations of OTTOsonics in various implementations show that the system performs well – exceeding expectations to the point of attracting widespread interest. The availability of a 24-channel system has enabled venues, music ensembles, and universities to begin producing 3D audio music, something that had previously seemed out of reach for many.

The socioeconomic context addressed by OTTOsonics is central to its appeal. Lowering the per-channel-cost to the bare minimum while maintaining a high-quality sound standard has allowed us to liberate multichannel sound systems from the academic ivory tower. This not only makes multichannel sound more affordable for cultural institutions but may eventually generate new artistic and musical approaches inspired by the modular and scalable potential of this platform. By fully opening the project and making available every schematic, 3D printing file, code and tutorial online, we empower users to take ownership of the system. They can maintain and repair it, customise it, and even copy, fork or clone our designs. This open-source design approach is not coincidental – the authors of this paper have been dedicated to open-source development for the past two decades.

5.2. Redefining the listening experience

OTTOsonics systems, with configurations of 24, 32, 48, or more speakers, have been used across genres such as hip hop, electropop, electroacoustic music, rock, classical, theatre, children's music and techno. Traditionally, these genres are experienced through front-facing setups with two main speaker arrays near the stage. However, our festival implementations, which feature audiences surrounded by speaker hemispheres, have altered the traditional listening experience. Such a condition of course requires the additional intervention of an intermediate spatialisation artist, who takes care of the spatial distribution of such a live performance.

From those experiences we can assert that spatialising music originally composed in mono or stereo through multichannel setups has enriched both musicians and audiences. However, objectively comparing the listening experiences between conventional stereo systems and immersive audio setups with smaller speaker hemispheres poses a challenge, as digital music is typically produced for traditional formats rather than 3D audio. The auditory imagination of musicians is shaped by the formats and media they use, which makes direct comparisons rather difficult. To address this, we plan to conduct user studies to better understand audience experiences with traditional and newer immersive formats, as current data are insufficient.

We envision that adopting Ambisonics sound reinforcement could mirror the shift in the 1970s when stereo systems became widespread. We do not advocate for a single best approach but observe that the presentation format of digital music is evolving. We believe that increasing the availability of multichannel sound venues could encourage more musicians to explore and create more 3D audio content, enhancing immersive musical experiences.

5.3. Spatial audio training

Over the past three years with OTTOsonics, we have encountered many initiatives interested in spatial audio. A common issue is the widespread misunderstanding of spatial audio capabilities. Many specialists mistakenly believe that proprietary or licensed formats, such as Dolby Atmos, are the only options. However, spatial audio should not be limited to licensed formats or specific workflows. For example, Dolby Atmos, designed for film, may not meet the needs of every electroacoustic music composer.

We frequently educate and advise venues and music production specialists on the diverse tools and multichannel formats available.

https://doi.org/10.1017/S1355771825100563 Published online by Cambridge University Press

Some universities, having invested heavily in licensed systems, find that students rarely use them due to their complexity and limitations and they are shifting their opinion towards open platforms.

To address this, we advocate for investing in open training and support in spatial audio, targeting in particular the following key groups of people:

- *Sound technicians*: professionals who work with musicians but often lack training in spatial audio, especially live spatialisation and spatial audio reinforcement.
- *Venue managers*: those who organise spatial audio projects but may not understand the technical possibilities or integration into their venues.
- *Music students*: many universities have computer music studios that are not easily accessible to students not deeply involved in electroacoustic music, limiting the production of other genres of multichannel music.

We believe that expanding open access to training and resources for these key groups will foster a broader adoption and exploration of spatial audio in diverse musical and performance contexts.

6. Future work

Our future plans involve several key initiatives. First, we aim to develop an open hardware AVB amplifier with an integrated 24-channel DAC, which would directly convert multichannel AVB content into analogue signals to drive the loudspeakers. Second, we plan to rethink the manufacturing process of the loudspeakers, exploring more industrial and yet cheaper production methods such as injection moulding and the use of sustainable materials. Those industrial designs will allow to scale our open-source production methods and further reduce the total cost of an OTTOsonics system. The design specifications are suitable for both DIY and industrial means of production, and thus ensure openess and scalability.

Finally, we intend to conduct more formal research on audience perception of sound reinforcement using immersive audio systems, to better understand the impact of such formats on the listening experience.

7. Conclusion

In this article, we have outlined the rationale behind the design of the OTTOsonics open hardware sound diffusion system and highlighted its key design features. After extensive testing and implementation across various musical genres and venue types, we conclude that the system performs well, and its core attributes – affordability and flexibility – have garnered interest from numerous cultural initiatives, many of which have independently created their own OTTOsonics systems. The growing number of universities, theatres and music venues adopting the system suggests that it has been positively received by these communities.

Our primary hypothesis was whether introducing an open, affordable and flexible sound diffusion system could foster greater production of immersive audio. After three years of activity, we have come to realise that the challenges are not solely technical or economic but also sociocultural. A key obstacle lies in the general unfamiliarity with multichannel audio formats and the scarcity of available multichannel content. Nonetheless, we believe that the OTTOsonics project has made a meaningful contribution by enabling the creation of numerous multichannel musical works in a short period and by disseminating knowledge about immersive audio. Consequently, we assert that this system has significantly helped various communities across Europe, Latin America and Asia to engage with spatial audio. This success indicates that the development of more alternative, open sound diffusion systems by audio enthusiasts and engineers worldwide could inspire more musicians to explore immersive audio formats.

In this article, we have also detailed the technical specifications of our speakers, amplifiers and audio architecture, as well as acknowledged the system's technical limitations. While the system's sound quality has exceeded our expectations, there remains a need for the development of new speaker and amplifier designs that are more efficient, neutral in frequency response, louder, more sustainable and even more cost-effective. Significant work remains to be done in this area.

Finally, we believe that OTTOsonics has contributed valuable insights into the critical design patterns necessary for an effective open spatial diffusion system. These patterns have been tested across a broad range of musical venues, university departments, theatres, composers, and spatial audio enthusiasts. We hope that future developers and researchers will build upon the knowledge we have gathered during this initial three-year phase.

Acknowledgements. OTTOsonics was funded in 2022 by the Austrian *Neustart Kultur* programme, supported by the Federal Ministry for Arts and Culture, and the Upper Austrian Department of Culture. Additionally, part of the research equipment was provided by the University of Arts Linz through the Tangible Music Lab. We extend our gratitude to our community and all those who evaluated our system within their respective initiatives, offering invaluable insights and feedback.

References

Barbour, J. L. 2013. The Equidome, a Personal Spatial Reproduction Array. *Proceedings of the 39th International Computer Music Conference, ICMC, 2013*, Perth, Australia, 12–16 August. Ann Arbor, MI: Michigan Publishing. http://hdl.handle.net/2027/spo.bbp2372.2013.062.

Berkhout, A. J., de Vries, D. and Vogel, P. 1993. Acoustic Control by Wave Field Synthesis. *Journal of Acoustic Society of America 93*(5): 2764–78.

Drack, V., Zotter, F. and Barrett, N. 2020. A Personal, 3D Printable Compact Spherical Loudspeaker Array. *Audio Engineering Society Convention 148. Audio Engineering Society. 2*–5 June. Online Convention.

Faerber, P. and Kocher, P. 2010. The Mobile Ambisonics Equipment of the ICST. *Proceedings of the 2010 International Computer Music Conference*, New York, 24–9 August. Ann Arbor, MI: Michigan Publishing. http://hdl.handle.net/2027/spo.bbp2372.2010.040.

Gerzon, M. A. 1972. Periphony: With-Height Sound Reproduction. *Journal of the Audio Engineering Society 21*(1): 2–10.

ITU-R BS.1116.1. 2015. Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems. www.itu.int/rec/R-REC-BS.1116 (accessed 15 September 2024).

Kronlachner, M. 2014. Plug-in Suite for Mastering the Production and Playback in Surround Sound and Ambisonics, Gold Award at AES Student Design Competition, Berlin, April. www.matthiaskronlachner.com/wp-content/uploads/2013/01/kronlachner_aes_studentdesigncompetition_2014.pdf (accessed 5 September 2025).

Lopez-Lezcano, F. 2014. Towards Open 3D Sound Diffusion Systems. *Music Technology meets Philosophy – From Digital Echos to Virtual Ethos. Joint Proceedings of the 40th International Computer Music Conference, ICMC 2014, and the 11th Sound and Music Computing Conference, SMC 2014.* Athens, Greece, September 14–20. Ann Arbor, MI: Michigan Publishing. http://hdl.handle.net/2027/spo.bbp2372.2014.136.

https://doi.org/10.1017/S1355771825100563 Published online by Cambridge University Press

Lossius, T., Baltazar, P. and de la Hogue, T. 2009. DBAP-Distance-Based Amplitude Panning. *Proceedings of the 2009 International Computer Music Conference*. Montreal, 16–21 August. http://hdl.handle.net/2027/spo.bbp2372.2009.111.

Margolis, G. and Small, R. H. 1981. Personal Calculator Programs for Approximate Vented-Box and Closed-Box Loudspeaker System Design. *Journal of the Audio Engineering Society 29* (June): 421–41.

Mooney, J. R. 2005. Sound Diffusion Systems for the Live Performance of Electroacoustic Music. PhD dissertation, University of Sheffield.

Peters, T., Matthews, T., Braasch, J. and McAdams, S. 2008. Spatial Sound Rendering in Max/Msp with Vimic. *Proceedings of the 2008 International Computer Music Conference*, Belfast, Ireland, 24–9 August. http://hdl.handle.net/2027/spo.bbp2372.2008.021.

Pulkki, V. 1997. Virtual Sound Source Positioning Using Vector Base Amplitude Panning. *Journal of the Audio Engineering Society 45*(6): 456–66.

Sharma, G. K. 2016. Composing with Sculptural Sound Phenomena in Computer Music. PhD dissertation, Institute of Electronic Music and Acoustics (IEM) at the University of Music and Performing Arts Graz.

Smith, B. D. and Cox, R. 2016. Big Tent: A Portable Immersive Intermedia Environment. *Proceedings of the 39th International Computer Music Conference, ICMC, 2016*, Utrecht, The Netherlands, 12–16 August. Ann Arbor, MI: Michigan Publishing. http://hdl.handle.net/2027/spo.bbp2372.2016.066.

Wefers, F., Stienen, J., Pelzer, S. and Vorlaender, M. 2014. Interactive Acoustic Virtual Environments Using Distributed Room Acoustic Simulations. *Proceedings of the EAA Joint Symposium on Auralization and Ambisonics*, Berlin, Germany. https://doi.org/10.14279/depositonce-9.

Zotter, F., Zaunschirm, M., Frank, M. and Kronlachner M. 2017. A Beamformer to Play with Wall Reflections: The Icosahedral Loudspeaker. *Computer Music Journal 41*(3): 50–68. https://doi.org/10.1162/comj_a_00429.

Zvonar, R. 2005. A History of Spatial Music. *eContact!* 7.4. https://econtact.ca/7_4/zvonar_spatialmusic.html (accessed 15 September 2024).

https://doi.org/10.1017/S1355771825100563 Published online by Cambridge University Press