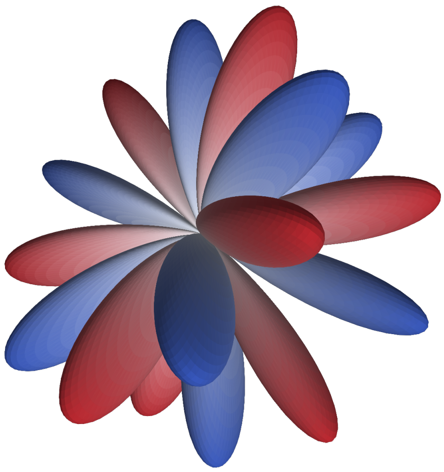
# Introduction to Ambisonics

Principles, Intuitions, and Artistic practices v0.2



# Contents

# Part I

## 1 Introduction

Ambisonics is a way of representing, transforming, and reproducing sound in space. It but one technique in the broader field of **spatial audio**, which is to organize sonic events not only in time, pitch, timbre, and intensity, but also in their position and behavior in space ( **movement**, **width**, **direction**, and **distance**). In this sense, Ambisonics is a tool for creating and manipulating sound in its three-dimensional aspect. Its position in space. A field that presents itself for artistic gestures.

In comparison to stereo productions, Ambisonics uses a variable and significantly larger number of loudspeakers surrounding the listening area, producing a greater sophistication in the resolution of sound positioning, immersion, and depth (Zotter & Frank, 2019). As well, surround formats such as 5. or 7.1 add more channels and spatial resolution but are constrained by a specific loudspeaker arrangement and a fixed mix on the audio on each channel. In contrast, Ambisonics works differently by **encoding the sound field itself** around the listener and boundaring with the loudspeakers, instead of encoding fixed signals into specific channel counts or loudspeaker layouts. Then, such sound field can be decoded for different playback systems: multichannel arrays, headphones, irregular installations, or other immersive systems. Notable drawbacks associated with Ambisonics are the presence of sweet spot and the perceived boundary surface created by the loudspeakr array (figure 1).

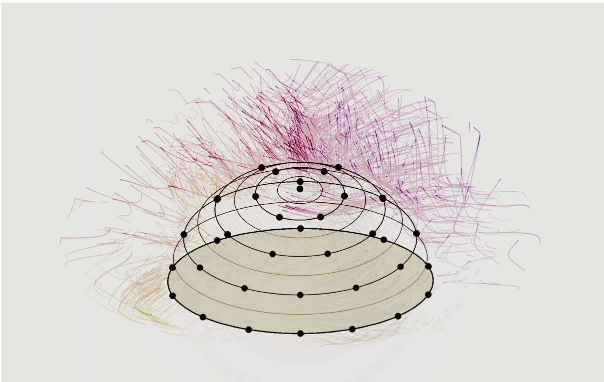


*Figure 1: A semi spheric sound system showing the boundary between the represented sound field, and the inner listening space.*

Ambisonics is used in live concert and acousmatic music, sound art, virtual reality, immersive installations, experimental electronic music, game audio, spatial research, and acoustic simulation. It is especially valuable when the artistic gesture is not simply to "pan" sounds around listeners, but to shape an entire soundscape: a world with sound objects, movement, environments, and relationships between virtual sounds, objects, and virtual spaces.

This text offers an introduction to Ambisonics: what it is, where it comes from, how it works, why to use it, and how it can be relevant artistically and musically.

## 1.1 Spatial Audio

Sound is inextricably spatial. We hear where a sound is coming from, how large the source seems, whether it is near or far, whether it is inside a room, outside in an open field, behind us, above us, or moving. Spatial listening is central to perception, orientation, memory, and emotional response. It helps to recognize environments, navigate danger, experience the acoustics, and distinguish between here and there.

Music and [sound] art have always workedn in this dimension, even before the development of digital spatial technologies: the architecture of a churches; choirs in Venetian medieval music; the distribution of instruments' groups; the use of space in opera; and the separation of sources in experimental music, to name a few, all show that space in music is not a recent concern. What changes with modern spatial audio systems is the degree of **control**, **precision**, and **processing** available to the composer, performer, and sound artist.

*Figure 2: A reverberant space. Kompresorinė, Vilnius, 2023. Lidar render, Ilme Vysniauskaite.*

For what concerns this text, Spatial Audio can be used artistically in several ways:

- It allows the listener to be surrounded by sound rather instead of in front of it (i.e. stereo).
- It offers Space as a musical gestural domain: where trajectories, densities, and movement.
- It makes the relation between sound, acoustics and architecture more explicit.
- It creates possibilities for soundscaping: immersion, disorientation, closeness, dramaturgy, real acoustic scenes and impossible sonic worlds.

In this context, Ambisonics establishes itself as one of the most important frameworks for working with spatial sound in a flexible, well adopted, knowledge and community rich, and conceptually coherent way.

## 1.2 Historical Context

Ambisonics emerged in the 1970s in the United Kingdom as part of a broader effort to develop a more convincing and more general system for sound reproduction than stereo. Its development is closely associated with the work of **Michael Gerzon**, alongside collaborators such as Peter Fellgett and others interested in psychoacoustics, signal theory, and immersive audio reproduction (Gerzon, 1973).

The original intention of Ambisonics was not simply to add more loudspeakers, but to create a system capable of reproducing a **sound scene** in a way that was not dependent on the playback format. In ordinary channel-based reproduction, there is a *hard* mix that corresponds to a fixed number of speakers: left, right, center, rear-left, rear-right, and so on. In Ambisonics, the spatial information of sources is encoded into an abstract intermediary representation, the so called *Soundfield domain* or *B-format* (also referred to as *Ambisonics domain*). Only later, during decoding, is that representation translated into actual loudspeaker signals, or *D - format* (Wiggins, n.d.), (Gerzon, 1985).

This means Ambisonics is a **scene-based** audio system rather than **channel-based**. Its aim is not simply to distribute sounds among fixed outputs but to preserve a more general structure of spatial relationships. Conversely, *oject -based* techniques, such as **Dolby Atmos** instead use separate sound sources accompanied by metadata describing their position and movement, which are then rendered to the specific playback system (Geier et al., 2010).

In the beginning, Ambisonics was developed as a 4 channel system, which encoded the sound field on the $x$, $y$, and $z$ axes, plus a chanel $w$, representing the overall sound pressure of the field. This is referred to as *First-Order Ambisonics.* The system then evolved to expand to higher channel counts, amounting to greater spatial resolution on the positions the virtual sound sources (VSS) can take. This is the so called *Higher-Order Ambisonics* (HOA).

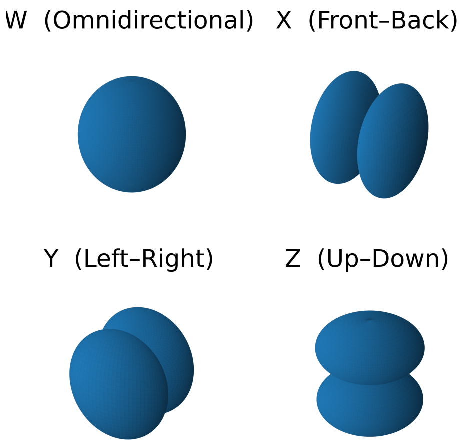



*Figure 3: The 4 channels used in First-Order Ambisonics.*

While Dolby Atmos is the industry standard for object-based audio in cinema and streaming, Ambisonics has carved a niche in music, academy, research, gaming and the arts. This is due to its open specification, its flexibility, and other factors or developments:

- modern computational power,
- flexible multichannel DAWs (i.e. Reaper),
- plugin ecosystems,
- interest in immersive concert formats,
- virtual reality, 360-degree videos and [interactive] gaming audio.

# Part II

## 2 What Is Ambisonics?

Ambisonics is a **mathematical framework and technique** for representing a three-dimensional sound field. Instead of assigning each sound directly to a specific speaker, Ambisonics encodes [virtual] sound sources as the signals that would be produced by a [virtual] multipole microphone which describe how energy is distributed acros multiple axes at the listening point.

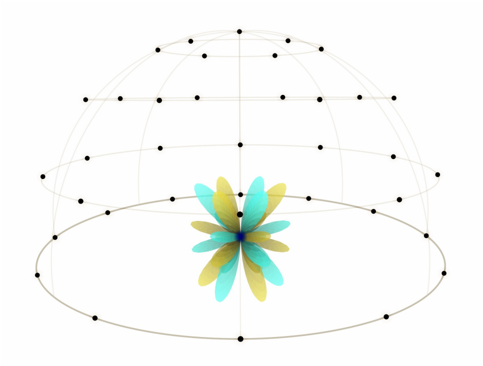


*Figure 4: Spherical harmonics in Ambisonics at the listening point.*

The fundamental idea is that sound can be decomposed into a set of *spherical harmonics* functions (the poles from before), which represents the invormation over the sphere around the listener. Later, this structured, multichannel information is interpreted (decoded) to the the signal that each loudspeaker would produce to contribue to the virtual soundfield.

This *encoding* of the spherical harmonics signal is often called **B-format** or an Ambisonic domain. A B-format multichannel signal is considered an intermediate representation, not intended for listening. However, processing and transformations can be aplied to a B-format signal. The enconded signal can experience various transformations of the signal distributed in space:

- rotation, translations, etc.,
- filtering and equalization,
- reverberation or other effects,
- addition, subtraction, convolution, etc.,

After processing, the B-format signal is *decoded* into a specific loudspeakers layout, to generate the signals to every loudspeaker would contribute to the sound field. This *speaker signals* is also referred to as **D-format**.

Because the B-format signals contain information about the sound field itself, and not about the hardware layout, it is said to be independent of the playback system or agnostic. This makes for a transportable format, lending itself to improved dissemination, distribution, archiving and sharing.

Ambisonic B-format signals can also be decoded to headphones, using **Binaural audio** rendering, which includes simulating head, torso and ear acoustics.

# 3 Why Ambisonics

Ambisonics is selected in certain contexts because it offers a set of tools and knowledge to **capture, synthesise, transform**, and **reproduce multichannel representations of sound fields**, with the addition to relative **ease of transportation and playback**, independent of any single loudspeaker arrangement.

With that, Ambisonics is used because of characteristics and its posibilities to:

- **Immersion**: Ambisonics captures and/or synthesises convincing directionality of sound sources, and immersive impressions applied to films, music, or other artistic expressions. The spatial resolution depends on the Ambisonic order, loudspeaker count and layout, and room acoustcs.
- **Transformability**: The sound field can be manipulated in a number of ways(rotation, tilt, compression, warping, focusing, to name a few. ), which lends itself for creative expressions. Aside from music, this can be specially useful for VR navigations sptial illusion.
- **Soundscaping**: The previous points facilitate the use of Ambisonics for the creation of sonic worlds to envelop the listener. From a musical perspective to documentary.
- **Transportability**: Or agnosticism to playback system. Ambisonic signals do not corresopnd to speaker signals, rather, they represent an abstract space that is what is shared as multichannel audio files. Then these files are decoded accoring to a venues particular sound system.

## 3.1 Ambisonics in context: other spatial audio techniques

Ambisonics is not the only spatial audio approach:

- **Stereo**: Stereo is the standard practice, it uses 2 channels to encode sound with *front right* and *front left* speakers. Limited spatial experience.
- **Surround Sound**: Mostly 2D systems with fixed channel-based signals. They offer a more detailed spatial illusion, with signals hard-mixed for specific loudspeakers layouts. Examples are 5.1 or 7.1.
- **Object-Based Audio**: Detaching from fixed loudspeaakers setup, Object-based systems transmit position and movement metadata along the audio sources (objects). Dolby Atmos is the obvious example. In comparison with Ambisonics, Dolby works in a closed source and trasmits sound sources separate from their metadata, whereas Ambisonics encodes the spatial information in the B-format signals themselves.
- **Wave Field Synthesis**: Unlike other techniques so far, which create a virtual sound source (VVS)outside of the of the listening space delimited by the speakers array, WFS is capable of reconstructing wavefronts that form a VVS inside the listening space. It does so by using a large, dense array of speakers that synthesizes the sound filed at the speakers' frontier so that it would syncronise and form sound field patterns within the listening area (see Fig. 5).

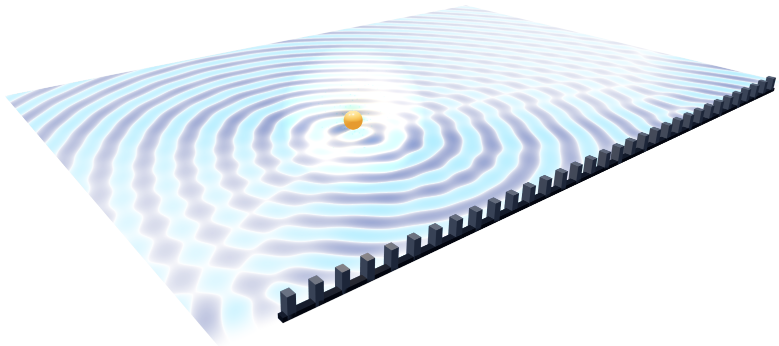


*Figure 5: Wave Field Synthesis. VSS on yellow.*

# 4 How Ambisonics?

In order to understand how Ambisonics works practically, it helps to think of the three main stages of the audio processing:

1. **Encoding**
2. **Processing / transformation**
3. **Decoding**

For the sake of demonstration, we will discuss these stages in the context of a DAW-based workflow (Reape in this case), but the same principles apply to other platforms (i.e. Max, which deserves a separate explanation), or raw processing of the audio for live performance, installation, or other contexts.

## 4.1 Encoding

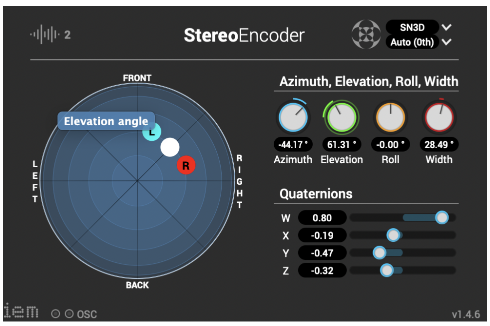



*Figure 6: IEM Stereo Encoder, vst3.*

In the encoding stage, a sound source is given a spatial position, or movement across a trajectory, along the surface of the *sphere* formed by the array of loudspeakers. An encoder process or plugin converts that source into an Ambisonic B-format signal by multiplying the source signal by factors that correspond to spherical harmonics for the given sound's *azimuth* (angle on the horizontal plane) and *elevation* (angle on the vertical plane) (Figure 7). The result is no longer a simple mono or stereo track but a multichannel representation of that source within a spatial field, or in other words, how it is captured by the virtual spherical harmonics at the listening point. As shown in figure 4, an encoder plugin will do that in an intuitive way.

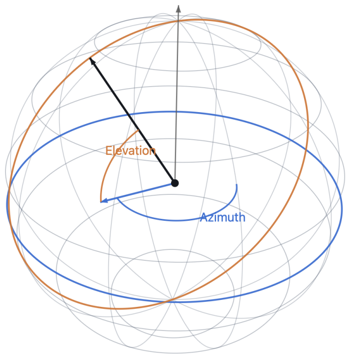



*Figure 7: Azimuth and Elevation on the listening space.*

Staying in the context of a DAW, to encode more than one source, there are various strategies that can be followed: a sources per track with their own encoder; several sources summed into the same encoderl; or several sources from a multichannel audio file encoded to different positions using one single encoder plugin (i.e. IEM multiencoder plugin). If several sources are encoded, they can be summed together into a common Ambisonic *bus*. This bus contains the combined sound field.

## 4.2 Processing and Transformation

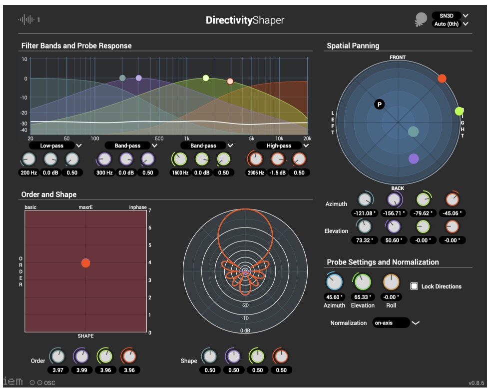



*Figure 8: IEM Directivity Shaper, vst3*

Once in the Ambisonic domain, that is, after the Encoding process (or plug-in, i.e. vst3), the multichannel signal can be all processed affecting all channels, which results on *transformations* on the *representations* of the sound field as a whole. Spatially localised transfromations are also possible affecting only a section of the sound field. Types of transformations include: rotation, warping, compression, reverb, delay, directivity shapers (figure 8), among others (Kronlachner, 2014), (Zotter & Frank, 2019).

Importantly, these transformations can affect the whole scene without separately processing each source, but also can be applied to individual sources and its resulting B-format signal summed at the *Ambisonic bus*.

### 4.3 Decoding

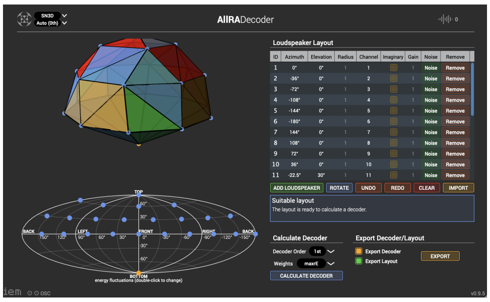



*Figure 9: Decoder plugin by IEM suite.*

After all the spatial processing, an Ambisonic signal is mapped onto the position of physical loudspeakers accoring to their azimuth and elevefation angles. This mapping is the *decoding* of the signal, an opposite operation to the encoding stage. The resulting multichannel signal is also referred to as **D-format** (Wiggins, n.d.), where each channel is the signal corresponding to a specific loudspeaker in the array, so that the sound produced by the loudspeakers would reproduce the sound field intended by the B-format signal.

In this stage, the signal can be decoded into a symmetrical optimal loudspeaker setup (rarely), a non conventional speaker layout, a virtual listening environement, or into a pair headphones (2 speakers), using **Binaural decoding.**

### 4.4 Binaural Ambisonics - Headphones

A sound field encoded as Ambisonics can also be decoded into a pair of headphones by convolving the Ambisonic signal with *head-related transfer functions (HRTFs)* for both ears. These HRTFs are impulse responses that, when applied on the Ambisonic signals, produce the sound as heard in each hear, given a particular position in space and filtering and interaural time-level differences (ITD and ILD).

This type of decoding is specially useful because of its use for headphone only delivery of content, such as intearctive media, games, and other artistic applications, offering a soundfield that can rotate and suffer other transformations. One such transformations is rotation by *head-tracking*, which rotates the sound following the direction of the head by some means (ie. sensors on the phone or the headphones). This has expanded Ambisonics beyond specialized concert halls and studios.

However, binaural listening is not identical to loudspeaker listening. The result may differ: spatial clarity, externalization, and elevation perception depend on many variables, including the HRTF model and the listener's own perceptual adaptation. It is therefore often wise to treat binaural and loudspeaker presentations as related but not interchangeable experiences, although, there is of course always room for creativity.

*Figure 10: Binaural rendering for headphones representing an Ambisonic sound field.*

## 5 Orders of Ambisonics

As mentioned before, the concept of Ambisonics is to represent the sound field around a listening point by sampling a sound source as it would be captured by virtual microphones with increasing spatial resolution as the order increases. This is done by encoding the sound field into a set of spherical harmonics.

The order of Ambisonics then refers to the degree of spatial resolution of the spatial position of virtual sound sources. The higher the order the more channels are used to represent the sound field.

The number of channels in an Ambisonic system is given by:

$$(N+1)^2$$

where $N$ is the order.

Thus:

- 0th order = 1 channel
- 1st order = 4 channels
- 2nd order = 9 channels
- 3rd order = 16 channels
- 4th order = 25 channels

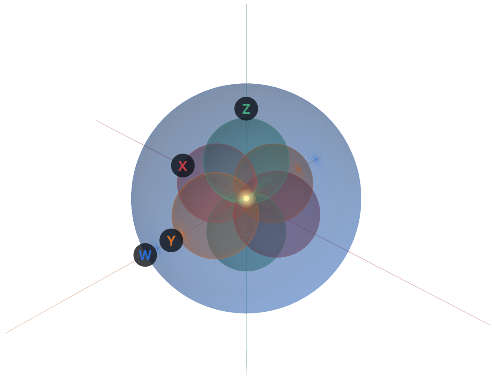



*Figure 11: W, X, Y and Z polar patterns as in FOA.*

## 5.1 First-Order Ambisonics

First-Order Ambisonics (FOA) uses four channels, which represent the sound captured as if by an omnidirectional microphone W, and three figure of eight microphones X, Y, and Z, along those axes. FOA can be produced with 4 coincident microphones with the pattern as previously described. But it can also be produced synthetically, by encoding a mono source into this spherical harmonics representation.

Even though FOA has an obvious limited spatial resolution, this format is widely used because of its ease of transportation, and found in virtual environments, game audio, commercial players and streaming platforms.

## 5.2 Higher-Order Ambisonics

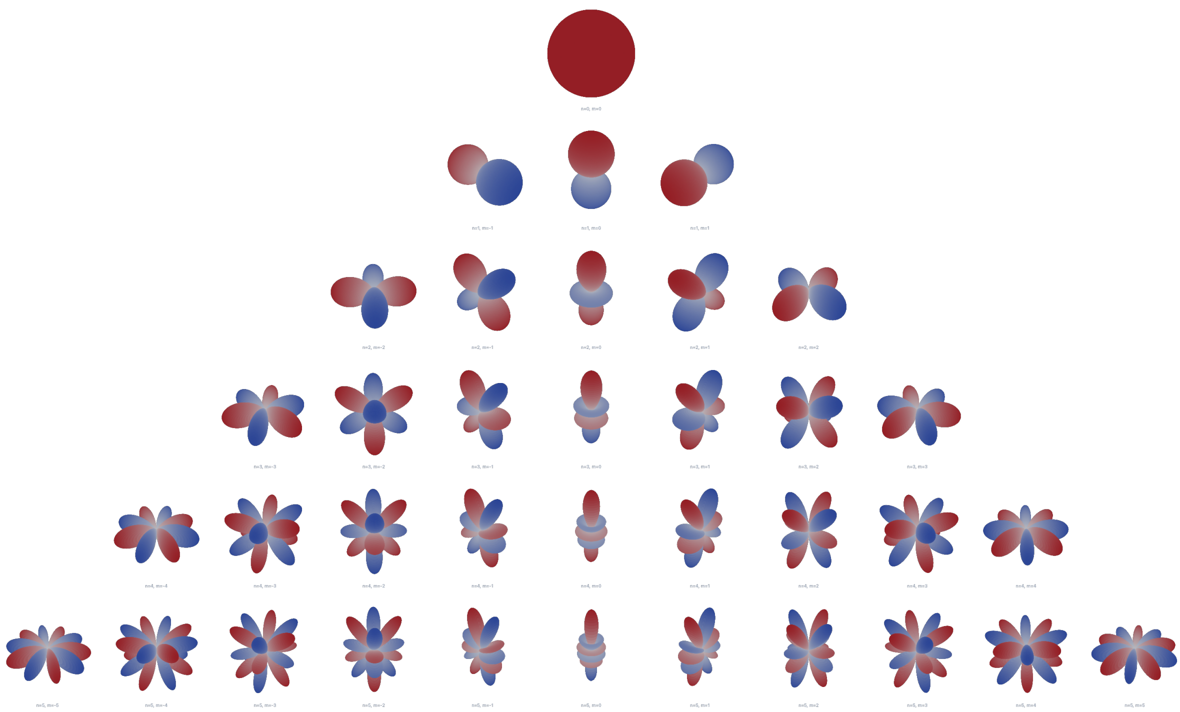


*Figure 12: Spatial Harmonic in Ambisonic encoding up to 5th order.*

Higher-Order Ambisonics (HOA) is the extension of the basic FOA format to encompass greater degrees of spatial resolution. in other words, virtual sound sources can be represented more precisely in their position in space thanks to the use of increasingly complex directivity patterns on the virtual microphones as imagined in the center of the listening space. These increment in directivity patterns is expressed in an increase in the number of channels.

In practical terms, this is to say that with higher orders of Ambisonics, the localisation of a virtual sound source will be shaper, accomplishing greater detail on movement and localisation of individual, separate ssources. On the downside, this escalation of order brings increased complexity in the processing of more chanells, memory use and routing of multichanel signals.

One thing to bear in mind is that an Ambisonic signal is not a multichannel audio file that is simply sent to hardware speakers, but instead they represent the sphereical harmonics components of the perceived or synthesized sound field (see figure 12). These channels are to be decoded in the right order or the spatial field is corrupted. However, for creative purposes this corruption and mix of channels can be intentional and explored (Tveit, 2025).

## 5.3 Conventions: Channel order and normalisation

As mentioned before, the sound field in teh sapce is encoded or embedded into spherical harmonics components, and the order of this componets matters for the processing agreement in between systems, algorithms, plugins, etc. Several conventions exist for ordering and scaling channels. The most common terms encountered are:

- **ACN** – Ambisonic Channel Numbering. Used in modern plugin suites like Ambix, IEM and Sparta. In this format the channels follow the position of the spherical harmonics pyramid in consecutive fashion:0, 1,2,3,4, etc.
- **FuMa** – Furse-Malham format follows the position of the spherical harmonics pyramid like so: 0, 3, 1, 2, 6, 7, 5, 8, 4, etc.

And for scaling, the most commonly found are:

- **SN3D** – Schmidt semi-normalization, which reduces the gain of the higher harmonic patterns in order to avoid excesive loudnesss from their contribution. Also used in most modern plugins.
- **N3D** – full normalization instead normalises at equal loudness between lower harmonic and higher spherical harmonics components.

If different parts of a workflow use incompatible conventions, the result may be spatially incorrect. In many contemporary workflows, especially plugin-based ones, **ACN** + **SN3D** is the most common practical standard. Older materials may use FuMa.

# Part III

## 6 Ambisonics in Art and Music

In an artistic context, Ambisonics provides for another tool for creative gestures, narratives, dramaturgy, and sonic aesthetics. For such cases, the technical discussions of orders and channel numbering take on another sense, since they can be thought of not as absolute factors of quality, but as variables for creative expresion and experimentation. Finer control of variables (i.e. higher order, spatial localization) does not guarantee a better experience o improved ideation of the sound works and vice versa.

### 6.1 Uses of Ambisoncis

Ambisonics is used in a wide range of artistic, technical, and research contexts.

- **Electroacoustic Music and Sound Art**: Traditional and specially suited to acousmatic music, Ambisonics is also used in live electronics, installation art, and experimental composition.
- **Immersive Music Performance**: Beyond live electronicmusic, Ambisonics is used for concerts with spatial audio surrounding the audience. There are different strategies for amplifying instruments, performers or virtual ensembles.
- **Immersive sound for Cinema**: Ambisonics is also used for the final mix or soundtrack of films,which have to be watched in a venue with an adecquate sound system.
- **VR, XR and 360-Degree Media**: *Headtracking* panning or rotating makes the use of Ambisonics highly useful for VR, XR and 360-degree videos.
- **Game Audio and Interactive Media**: Similar to the last point, Ambisonics contributes in the cration of sonic sceneries in virtual and interactive environemtns.
- **Ambisonics field recordings**: Besides synthesis, the sound field can be recorded with multichannel microphones and encoded into B-format signals (most comonly in FOA).
- **Research and Simulation**: Ambisonics is also used in psychoacoustics, room acoustics, perceptual research, and the study of auditory scene reproduction.

### 6.2 Artistic gestures with Ambisoncis (and spatial audio in general)

Within the mentioned applications, spatial audio, and Ambisonics in particular, present opportunities for artistic expressions, which variables can include but are not limited to:

- **Composing the space**: The space as compositional parameter. A recurrent topic on the break between music and sound art. Whereas the firs is traditionally concerned with the arrangement of sound in time, the second tends to work with the orgransing of sound in space. **Space itself becomes compositional material**.

- **Movement as gesture**: In relation to the previous point, movement can also convey musical, artistic and symbolic expression: slow, sudden, converting, diverging, layering, circular or shapped motions. Thus, movement is not an add-on. It can be rhythm, phrasing, tension, release, orientation, or narrative.
- **Engulfing the listener:** With the audience at the centre of the sound system, listeners are embedded *within* the sound field, which makes for potential compelling, engulfing, and some times disorienting experiences. Ambisonics is particularly suited for atmospheric sound, soundscaping, and donic narratives.
- **Realism and Abstraction**: The soundscapes created with Ambisonics can range from figurative to impossible scnearios. Transporting the audience to an alternate reality with sound only is not a trivial task.

### 6.3 Artistic Challenges

When composing or designing with Ambisonics, it is useful to ask artistic, perceptual questions, such as:

- The *sweet spot*, is small.
- Effective illustion of VSS position and movement (Kendall & Cabrera, 2011) (Schumacher et al., 2021)
- Is the audience seating, standing or moving? where and how are they situated?
- What is the spatialisation strategy? live, automation, what interfaces are used?
- How is space thought? as symbol, structurally, compositionally, theatrically, figuratively?
- What is the form of the sound events? functional, expressive, ritual, environmental, destabilizing?
- What is the aesthetic of the sound? pointilistic, textural, soundscapist, etc?

Ambisonics and other spatial audio methods can easily become a demonstration of technology, instead of carring a deeper artistic sense. Therefore it becomes important to evaluate the role and relevance of the spatialisation in the development of pieces tu support the artistic expression (perceptual, stylistic, narrative, decorative, musical)

## 7 Ambisonics in Practice

Ambisonics, the same as other spatial audio techniques, can be used in a variety of contexts, from fixed media to live concerst and from symmetrical optimal setups to irregular or sound art installation layouts. Likewise, the tools that are available to the composer and sound artist range from a number of DAWs, pluginns, microphones to technical and aesthetic decisions that can be made in the process of composition, recording, production or performance.

*Figure 13: Spatial audio concert setup, with live spatialisation and hardware instruments.*

### 7.1 Ambisonic Synthesis and Playback

The most common way to work with Ambisonics is through a DAW, which allows for the composition, processing, and mixing of multichannel audio. Reaper is a popular choice due to its flexible routing and support for multichannel tracks. Plugins such as IEM's Ambisonic suite, SPARTA, or AMBIX provide tools for encoding, processing, and decoding Ambisonic signals. Max/MSP is often used for real-time and composed spatialization, and for the power of its flexibility. For installations, custom software or hardware solutions may be developed to manage the sound field.

The workflow typically involves encoding mono or stereo sources into B-format to give it a *virtual location*, or in other words to create a *virtual sound source*. Then, aggregating several sources in a bus, and applying spatial or multichannel transformations before or after the aggregation, and finally decoding the Ambisonic signal bus into a specific loudspeaker layout or binaural render for headphones. In the case of arbitrary sound systems positions, techniques or tools, as is often the case in sound art installations, the multichannel signal can be Ambisonics as much as an arbitrary design for signals in each speaker, and audio files hard-mixed for the specific layout.

Typical DAWs used in an Ambisonic or spatial audio workflow include Reaper, Ableton Live, Max/MSP, and Logic Pro. Each has its strengths and weaknesses in terms of multichannel support, routing flexibility, and plugin compatibility (see Table 1).

| DAW | Multichannel Track Support | Routing | Commonly Used With | License |
|---|---|---|---|---|
| Reaper | YES | Arbitrary | IEM, AMBIX, SPARTA | Low-cost |
| Ableton | Limited | Restricted | ENVELOPE | Paid |
| Max MSP | YES | Arbitrary | SPAT | Paid |
| Logic | Limited | Restricted | Atmos | Paid |

*Table 1: Comparison of Digital Audio Workstations (DAWs) in the context of Ambisonics workflows.*

### 7.2 Ambisonic Recording

Another worksflow is to record Ambisonics directly. This is equivalent to obtaining an already recorded B-format signal, which can be processed and decoded. This captures the sound field of a real acoustic environment, in order to be played back as is, or after applying transformations.

The following Table 2is a range of first-order and higher-order Ambisonic microphones used in spatial audio production:

| Name | Manufacturer | Ambi. Order | Ch. Count | Incl. Soft. | Price |
|---|---|---|---|---|---|
| Zoom H3-VR | Zoom | FOA | 4 | Yes | €300–400 |
| Zoom H8 + VRH-8 Capsule | Zoom | FOA | 4 | Yes | €450–600 |
| RØDE NT-SF1 | RØDE | FOA | 4 | Yes | €900–1100 |
| Sennheiser AMBEO VR Mic | Sennheiser | FOA | 4 | Yes | €1200–1600 |
| Core Sound TetraMic | Core Sound | FOA | 4 | Yes | €1400–1800 |
| SPCMIC | SPC | FOA | 4 | Partial | €1500–2500 |
| Audio-Technica BP3600 | Audio-Technica | FOA (enh.) | 8 | No | €3000–3500 |
| ZYLIA ZM-1 | Zylia | HOA (3rd) | 19 | Yes | €1200–1500 |
| mh acoustics Eigenmike EM32 | mh acoustics | HOA (4th) | 32 | Yes | €20,000–30,000 |
| Voyage Audio Spatial Mic | Voyage Audio | FOA | 4 | Yes | €1000–1300 |

*Table 2: Comparison of Ambisonic microphones including supported order, channel count, and software integration.*

### 7.3 Ambisonics Live

For the case of Live performances, a number of tools and challenges appear. Mostly, the problem to solve is the input of signals into the Ambisonic domain, real-time processing and output into the speaker layout. This can be done with a number of software and hardware solutions, which are listed in Table 3.

Another topic to consider is the strategy for spatialisation in live performance. This can be done by the performers themselves, an engineer or an automation algorithm. The choice of gestures, interfaces and spatialisation decisions can carry the artistic expression.

| Name | Type | Max Ch. | Latency | Network | Use Case |
|---|---|---|---|---|---|
| JackTrip | Software | 2–128+ | Very Low | Internet/LAN | Networked performance |
| Soundjack | Software | 2–64 | Very Low | Internet/LAN | Remote ensemble playing |
| SonoBus | Software | 2–32 | Low | Internet | Flexible remote collaboration |
| Dante | HW/SW | 512+ | Very Low | LAN | Live sound, installations |
| AVB | Protocol | 100+ | Very Low | LAN | Time-synchronized audio networks |
| OSC | Protocol | N/A | Low | Internet/LAN | Control and metadata (no audio) |
| NDI Audio | Software | 16–64 | Low | LAN/Internet | Audio-visual streaming |
| ReaStream | Plugin | 2–64 | Low | LAN/Internet | DAW-to-DAW routing |
| Max/MSP | Software | Flexible | Low | Internet/LAN | Custom spatial systems |
| SuperCollider | Software | Flexible | Low | Internet/LAN | Algorithmic spatial audio |
| MADI | Hardware | 64–128 | Very Low | Coax/Fiber | Stage and studio transport |
| AES67 | Protocol | 100+ | Very Low | LAN | Interoperable AoIP systems |

*Table 3: Tools, protocols, and systems for multichannel audio exchange in spatial audio performance and ensemble contexts.*

## 7.4 Technical Challenges

Along with the artistic challenges, there are technical ones that affect the oucome of an Ambisonic or spatial audio work:

- Loudspeaker system may be imperfect, irregular, or arbitrary. This can be a challenge for an orthodox Ambisonic setup, but an opportunity for site specific designs.
- Sound systems have to be calibrated for gain, equalization and delay compensations.
- Architectural spaces can be acoustically beneficial or detrimental. This gives way to use the space as a commpositional element.
- Sound systems may result in fairly expensive or complicated configurations and signal flow. Also refer to the previous table 3 for live performance.
- Artist expression should trump over technical demonstration.
- Postion and movement of the VSS, beyond its perceptual challenges, can be difficult to operate in the DAW, whether for live or fixed media.
- Spatialisation strategies and interfaces can pose a challenge. Choices include MIDI controllers, OSC protocol, custom interfaces, or manual control of the parameters in the DAW.

A strong Ambisonic, or spatial audio work is therefore not just technologically elaborate, but also robust enough to survive real space, speakers, listeners, and acoustic imperfections.

## 8 Conclusion

Ambisonics is but one of many spatial audio techniques. Its relevance and importance in contemporary practices is due to its flexibility, transportability, the wealth of tools, knowledge, research and community around it, and its adoption in key contexts, such as virtual reality, immersive music, and sound art.

Even thought, it is not the only nor the best spatial audio technique, and it is not without issues and limitations, it is a powerful tool for artistic expression that also allows an entry point to the diverse context of spattial and multichannel audio and music. Therefore, conceptual understanding of its principles, uses, technicalities and artistic possibilities is dealt with in this text for the purpose of providing a comprehensive introduction to the topic, and to encourage further exploration and experimentation with Ambisonics in artistic contexts.

## 9 References


Geier, M., Ahrens, J., & Spors, S. (2010). Object-based audio reproduction and the audio scene description format. *Organised Sound*, *15*(3), 219–227. https://doi.org/10.1017/S1355771810000324

Gerzon, M. A. (1973). Periphony: With-height sound reproduction. *Journal of the Audio Engineering Society*, *21*(1), 2–10.

Gerzon, M. A. (1985). Ambisonics in Multichannel Broadcasting and Video. *Journal of the Audio Engineering Society*, *33*(11), 859–871.

Kendall, G., & Cabrera, D. (2011). Why things don't work: What you need to know about spatial audio. *Proceedings of the International Computer Music Conference*, 37–40.

Kronlachner, M. (2014, June). *Spatial transformations for the alteration of ambisonic recordings* [Master's thesis, University of Music, Performing Arts Graz, and Graz University of Technology] [Master's thesis]. https://www.matthiaskronlachner.com/wp-content/uploads/2013/01/Kronlachner_Master_Spatial_Transformations_Mobile.pdf

Schumacher, F., Espinoza, V., Mardones, F., Vergara, R., Aránguiz, A., & Aguilera, V. (2021). Perceptual recognition of sound trajectories in space. *Computer Music Journal*, *45*(1), 39–54. https://doi.org/10.1162/comj_a_00593

Tveit, A. (2025). *Spatial sound transformation toolkit for max* (Version latest) [Accessed: 2026-03-26]. https://github.com/Andertve/Spatial-Sound-Transformation-toolkit-for-Max

Wiggins, B. (n.d.). Ambisonic channels [Ambisonic.info, accessed 9 March 2026].

Zotter, F., & Frank, M. (2019). *Ambisonics: A practical 3d audio theory for recording, studio production, sound reinforcement, and virtual reality*. Springer. https://doi.org/10.1007/978-3-030-17207-7