Stereo and Ambisonics: A reflection over parallel spatialization techniques in the piece *Scie Luminosissime*

By Giuseppe Pisano

As a composer of acousmatic music, I have often presented my work on a variety of different loudspeaker systems whese can be subdivided in two macro-families: those intended for the playback of multichannel pieces – often symmetric systems, in the shape of a circle, a dome, or a regular polygon – and those intended for acousmatic diffusions – the loudspeaker orchestras, mainly employed for the diffusion of stereo pieces, where the performer spatializes sound in real time. In this practice, it is not uncommon to listen to re-adaptations of pieces, presented in different formats to suit different types of systems, and I have always found the Ambisonics workflow a very successful approach to scale my pieces on systems belonging to the first macro-family. However, when I first faced the need of presenting my multichannel pieces on systems meant for stereo diffusion, I was very unhappy with the results I achieved through the typical methods employed to adapt Ambisonics

to stereo, such as rendering my pieces in Ambisonics UHJ format, or using virtual microphones to record the sound field in stereo. Based on this experience, I decided to compose a piece in two versions in parallel, for stereo-setting and Ambisonics.

This process initiated a reflection process on how I would manage sonic materials differently according to the format to emphasize their characteristics and has become an acquired strategy in my compositional practice.

The first piece I composed in parallel as a stereo- and Ambisonics-version was *Scie Luminosissime* (2018-2020)[1]. In this paper I outline the compositional process of this piece, and highlight some conceptual differences that determined my choices in the use of the materials and the approach to spatialization.

Scie Luminosissime was inspired by the album *Fratello Mare* by Mike Cooper and the poem *Sii dolce con me. Sii gentile* by Mariangela Gualtieri. This inspiration is rather indirect but provides the setting in which I wanted to compose.

The materials used in *Scie Luminosissime* include field recordings ob-

tained during the summers of 2016 and 2017 – mostly fragments of human agencies in different reverberating spaces – in dialogue with concrete and synthesized sounds – the *scie*, usually distorted and/or convoluted trails of synthesized sound tones – and a few bits of sound sampled from different guitar pieces by Mike Cooper. My goal was to generate momentary and elusive visualizations of human interactions in an otherwise diffused blur of gestures and fragments.

The time-structure of the piece can be divided in two parts, each subdivided in two sections.

The first section is mainly characterized by simple tones (1) to which other concrete elements are gradually added, occupying wider portions of the spectrum (2). The cumulation ends with a gesture (3) that opens a new section and reiterates later (4) with a different organization of the materials. The idea was to create a new perspective, as if the same picture was seen from a different angle. After this, the space dissipates fading back into luminous trails again.

In his article "Space-form and the Acousmatic image", Denis Smalley
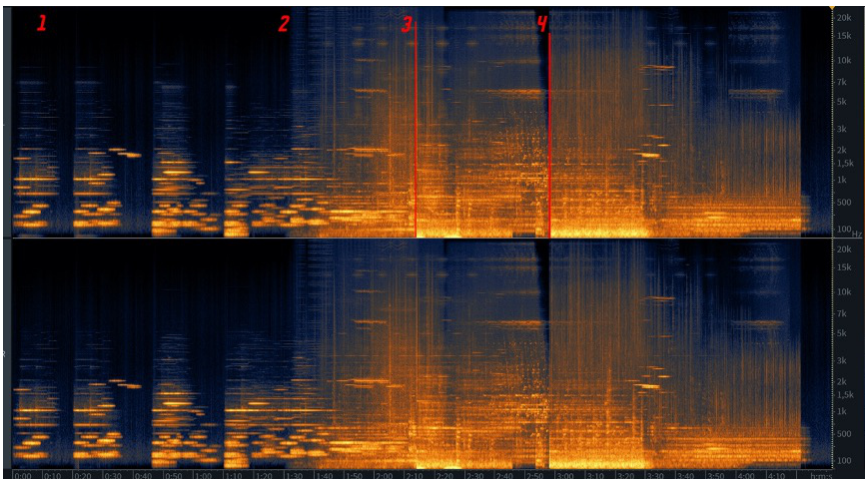

Figure 1: spectrogram of *Scie Luminosissime*, the numbers highlight the four sections of the piece.

(2007) describes a soundscape he listens to sitting by his window. He is inside, and the whole world is outside, beyond the fourth wall. The idea of being a spectator of sound from a vantage point reinforces once again the parallel between acousmatic music and cinema which has been structural in the discourse originated in the French literary and musical school (Chion 1990; Normandeau 1992). This connection is even more important now that the expansion of sound from stereo to more immersive formats is becoming a widely available technology, especially through the film industry and other audio-visual media, and demonstrates how Smalley's front-based approach to the description of the soundscape in the article is in reality much more techno-centric than openly intended.

In Smalley's narration, the window is the point from which the sound radiates in the listening space, it is both an access to the external sonic environment and a filter that facilitates the localization of the different sound sources.

When we listen to music in stereo, loudspeakers behave in a very similar way as this window. They define a front from which sounds radiates in our listening space, a two-dimensional surface on which a three-dimensional illusion is represented as in a trompe-l'oeil (more a trompe-l'oreille in our case) that uses perspective to generate depth. This changes radically when we are surrounded by loudspeakers. We do not sit by a window, we sit under a "glass dome", we can still have a frontal perspective, but our perception responds to stimuli coming from every direction. In this situation, music is then not merely projected in the listening space, but it becomes the listening space. The listener is in the sound field, immersed, and the sound is an architecture.

Materials in an immersive sound-field context are hierarchically organized in a figure/ground relationship. This gestalt notion, that has been applied before to many other fields such as soundscape analysis (Porteous and Mastin 1985) and acousmatic music (Santarcangelo and Wanke 2020), is of course still relevant for immersive formats where music-space and listening space overlap. A good example might be the nomenclature used to

distinguish between the two types of sound sources in Dolby Atmos: Beds and Objects. Beds are multi-channel sources, meant to establish a scene, while Objects are mono sources that would benefit from a more precise positioning and from panning.[2] This distinction is not just practical or operational, but reflects a difference in the perceptual and musical roles of the sound materials, which can also be applied to other immersive formats. When composing in Ambisonics, authors can adopt a similar approach. Some sounds can be treated as structural pillars of a scene, establishing the environment. Other materials can have more prominent roles, not only in terms of levels and clarity, but also in their spatial behavior with faster trajectories, swift movements, dramatic panning. However, the absence of a structural distinction between sources and their roles allows for more nuanced musical solutions and alterations in the roles of the materials.

An example of how I apply this paradigm is at the beginning of *Scie Luminosissime*, where, in the beginning, every single tone is a fixed position, like in a constellation, where

every star has its own place. These become spatial references, describing and defining the sonic environment with their presence. As the piece develops, more resonances add up together with similar elements of different timbres, moving slightly or appearing in opposite positions to the sound they originate from. These new materials generate a movement around the original structure, and every new tone is introduced with similar copies of itself, each presenting slightly different characteristics. The original instances are still fixed in space, as solid statements of presence. The other ones emerge from the background, creating a shapeshifting environment.

In the stereo version of the piece, this effect was achieved by determining a position for each original source on the horizontal plane, choosing to leave low frequency sounds in the middle and to enlarge the picture progressively at higher frequencies. Resonances were obtained with a stereo reverb, inverting channels after a low pass filter. In the Ambisonics version, I used a different technique encoding the original mono sources with different

instances of IEM MultiEncoder,[3] creating new copies of some of them to fill the entire space.

For the additional sounds emerging, I decided to create small alterations in their position modulating the values for the coordinates with a random LFO at a very small amplitude, which generated an organic liveliness.

Lower frequency sounds and resonances which I wanted to sound bigger and further away were encoded using the AmbiX suite by Matthias Kronlachner (2013). This software presents an additional parameter called size that can be used to adjust spatial directivity, adjusting the ratio between the levels of the omnidirectional signal and those of the other spherical harmonic components. At higher values of size, the sound sources are spatialized as more omnidirectional, equally distributing the sound to all loudspeaker, making them appear bigger, but also less directional. Adjusting this parameter, I was able to spatialize sounds with a weaker sense of directionality which played a different role.

Gestural sounds which, in the stereo version of the piece, were spatialized using traditional linear panning, in Ambisonics expand their movement to the second and third dimensions, swooshing from one side to the other of the sphere. These movements are achieved using crossfades between two encoded monophonic sources, and are presented over multiple diameters, to increase the sense of tridimensionality.

Another important difference between 3D Ambisonics and stereo is the introduction of verticality. However, while real height can only be found in three-dimensional sound formats, different techniques can be employed in stereo to generate a perceptual illusion of height. Denis Smalley provides some good examples for this when introducing the concept of gravitation, referring to the work of French composers such as François Bayle, François Delalande, and in particular to Bernard Parmegiani and his piece *Geologie Sonore* (Kronlachner 2013). In his writings, Smalley addresses height from a spectromorphological standpoint, claiming that the illusion of vertical movement is based on the

quality and the behavior of the sonic materials. A psychoacoustic approach to the production of vertical panning illusions is provided by Curtis Roads in the book *Computer Music Tutorial* where he stresses the role of the pinnae in sound localization on the vertical plane. Showing the Head Related Transfer Function spectra of sounds coming from above and below ear level, he describes the functioning of vertical panning (Roads 1996).

In Ambisonics, I found it very useful to spatialize sounds over the vertical plane trying to imitate or to refer the behavior and spectral characteristics of sounds that I am most likely to perceive above or below ear level in real life. For example, there are sounds of birds or sounds that "behave like" birds, there is the sound of rain and a sound that "behaves like" rain. Exploiting the vertical plain to spatialize sounds can provide an extra layer of realism to unreal sounds. Similar effects are much more difficult to achieve in stereo. As pointed out by Roads (1996, 10), realistic vertical panning illusions depend on too many factors such as speaker placement and listener's position, and are therefore extremely situational and very fragile. But Roads also describes that sound localization on the vertical plane is stronger for sounds above 6 kHz. Keeping this in mind, I found it useful to split sounds with very complex spectra over different frequency bands using a set of EQs, and spatialize these bands individually. Increasing the width of high frequency components did not provide any clear illusion of a vertical panning movement, but it widened the stereo image, giving more air and a stronger sense of openness to the sound image. I found it also useful to combine this frequency-based approach to the material- and spectromorphology-based approach described by Smalley, and to see them as two sides of the same coin which together contribute to create an illusion of height in a stereo mix. Spatializing filtered components of sound sources is a technique that was used extensively in *Scie Luminosissime*, especially for granular materials, where multiple monophonic sources are processed using the GRM Space Grain[4] a multi-channel granulator whose resulting signals are then filtered and encoded using multiple IEM Multi-Encoder with some personal presets.

Scie Luminosissime was the first piece I composed using this parallel method, which I later adopted for other works, too. I found this compositional process of creating two versions of the same piece very useful to thoroughly exploit the strengths of both formats. The versions share the same materials, the same structure and the same duration, but manifest themselves animated by very different spirits. While the stereo version appears more compact, perhaps with a stronger body, more punch and a concrete foundation, the Ambisonics version is more enveloping and relaxed: gestures are perhaps less dramatic, but more realistic. When working in Ambisonics, my feeling was that certain sounds did not require the same dynamic presence that I was using in stereo. Instead, they could interpret their role in the narrative development of the piece much more successfully with just a different organization of the space. Another relevant aspect is how the relationship between full and empty space was managed differently in the two versions of the piece. Filling space in Ambisonics simply requires "more sound", and adding multiple voices of a sound material, if correctly positioned, will rarely results in the space feeling crowded. It was therefore necessary to double some sound sources in the Ambisonics version or to cut out few others in the stereo version to balance the sense of fullness and emptiness in different moments of the piece to make sure that the same impression was achieved.

In conclusion I think that my approach to the composition of this piece has perfectly filled the purpose of making the two versions perfectly functional to the two different types of loudspeaker systems and performance styles. Expanding this workflow to other pieces has made possible to present the entirety of my work in all possible contexts.

Notes

[1] www.giuseppepisano.com/fixed-media/ (last access Feb. 16, 2023).

[2] Dolby Atmos Renderer v3.0. Dolby Lab. Inc, 2 August 2018, 116.

[3] IEM – Plug-in description "multiencoder", 2022. https://plugins.iem.at/docs/plugin descriptions/#multiencoder (last access Feb. 16, 2023).

[4] InaGRM – Spaces, 2022. https://
inagrm.com/en/store/product/15/
spaces (last access Feb. 16, 2023).

References

Smalley, D. (2007). "Space-Form and
the Acousmatic Image." *Organised
sound* 12(1): 35-58.

Chion, M. (1990). *L'audio-Vision. Son
Et Image Au Cinéma*. Nathan, Paris.

Normandeau, R. (1992). "Un Cinema
Pour L'oreille." Doctoral dissertation,
Montreal: Universite de Montreal.

Porteous, J. D., and Mastin, J. F.
(1985). "Soundscape", *Journal of
Architectural and Planning Research:*
169-86.

Santarcangelo, V. and Wanke, R.
(2020). "The Early Stage of
Perception of Contemporary Art
Music: A Matter of Time." *Organised
Sound* 25 (2): 130-41.

Kronlachner, M. (2013). "Ambisonics
Plug-in Suite for Production and
Performance Usage." Paper
presented at the Linux Audio
Conference.

Roads, C. (1996). "Sound
Spatialization and Reverberation". In
idem, *The Computer Music Tutorial.*
MIT press: 466-469.