World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:17, No:11, 2023

A Guide to the Implementation of Ambisonics Super Stereo

Alessio Mastrorillo, Giuseppe Silvi, Francesco Scagliola

Abstract—This paper explores the decoding of Ambisonics material into 2-channel mixing formats, addressing challenges related to stereo speakers and headphones. We present the Universal HJ (UHJ) format as a solution, enabling the preservation of the entire horizontal plane and offering versatile spatial audio experiences. Our paper presents a UHJ format decoder, explaining its design, computational aspects, and empirical optimization. We discuss the advantages of UHJ decoding, potential applications, and its significance in music composition. Additionally, we highlight the integration of this decoder within the Envelop for Live (E4L) suite.

Keywords—Ambisonics, UHJ, quadrature filter, virtual reality, Gerzon, decoder, stereo, binaural, biquad.

I. Introduction

AMBISONICS, a powerful spatial audio format, has gained prominence in audio production and immersive sound experiences. However, decoding Ambisonics content for 2-channel systems can be challenging. Stereo decoding excludes information from various directions, while binaural decoding is limited to headphone use. In this context, the UHJ format offers an intriguing alternative. It has the capability to convey information from the entire horizontal plane of the Ambisonics sphere while remaining compatible with both headphones and traditional speakers. Furthermore, the information contained in the two Left and Right channels can be decoded back into the E-format, resulting in a reconstructed B-format that retains information from the entire horizontal plane. This unique combination of features places the UHJ format as a promising solution, addressing some of the limitations associated with both stereo and binaural decoding techniques.

II. The UHJ Format

Ambisonics C-format, also known as UHJ (from *Universal UD-4*, *Matrix H* and *System 45J*), is an evolution of Ambisonics designed to be compatible with both mono and stereo (2-channel) sound reproduction systems. It achieves stereo compatibility without the need for additional decoding, delivering a notably wider stereo image compared to traditional stereo systems. Ambisonics C-format effectively conveys audio information from First-Order Ambisonics, and as Gerzon noted:

A. Mastrorillo is Professor of Computer Music, New Technologies and Musical Languages Department, State Conservatory ”Piccinni”, Italy (e-mail: mastrolex@gmail.com).

G. Silvi is a Full Professor of Electroacoustic, New Technologies and Musical Languages Department, State Conservatory ”Piccinni”, Italy.

F. Scagliola is a Full Professor of Electroacoustic Music Composition, New Technologies and Musical Languages Department, State Conservatory ”Piccinni”, Italy.

. . . The front-stage material is reproduced with sharply defined images, occupying virtually the whole of the stereo stage, and some sound positions can appear marginally beyond the loudspeakers. Rear-stage sounds appear with rather less well-defined images between the stereo loudspeakers [1].

Furthermore, Ambisonics UHJ can be seamlessly mixed with conventional stereo material. Notably, UHJ-encoded stereo material can be found in various commercial album releases [2].

The UHJ format takes the four *WXYZ* components and derives four *LRTQ* components capable of recreating the entire sonic sphere. Here, *LR* represents two signals suitable for stereo reproduction, while *TQ* enhances localization accuracy and provides full-height surround sound. However, for the proposed algorithm, only the first three components (*WXY*) will be used to convey information from the horizontal plane.

III. The Envelop for Live Suite

Envelop for Live (E4L) is an open-source audio production framework designed for spatial audio composition and performance. It seamlessly integrates with the music production environment of Ableton Live 10+ and Max for Live. E4L's core design philosophy revolves around modularity and flexibility, empowering artists to compose and perform spatial audio while also enabling developers to create innovative audio effects within the Ambisonics domain [3].

Before the author's contribution, E4L did not feature a UHJ decoder. However, through collaborative efforts with the Envelop Sound community of developers, a brand-new UHJ decoder was created and subsequently incorporated into the suite [4]. This decoder has found practical application in the author's electroacoustic compositions, including *Trippin' on the edge of time* and *Emptiness of the hanging*. These compositions have garnered recognition and been featured in events hosted by Electric Sound, SMC, ICMC, eviMus, Arte e Scienza, and Sin[x]Thésis.

IV. The Gerzon's Math

In Gerzon's work, ”Ambisonics in Multichannel Broadcasting and Video” [1], the equations for converting B-format into UHJ format involve a 90° phase-shifted j component, as expressed below:

$$\Sigma = 0.9397W + 0.1856X \tag{1}$$

$$\Delta = j(-0.3420W + 0.5099X) + 0.6555Y \tag{2}$$

The left and right signals can then be derived using the following equations:

$$L = 1/2(\Sigma + \Delta) \quad (3)$$

$$R = 1/2(\Sigma - \Delta) \quad (4)$$

These equations take as input the three Ambisonic B-format components (*WXY*) and produce the intermediate components Σ and Δ, from which the stereo left and right signals can be obtained. Fig. 1 illustrates the polar patterns of the Σ and Δ intermediate components.

The *j* component is generated using a quadrature technique. A sinusoid can be represented as the sum of a sine function with a phase of zero and a cosine function with a phase of $\pi/2$. If the sine part is considered the in-phase component, the cosine part can be referred to as the phase-quadrature component. In general, phase-quadrature implies being '90° out of phase' [5].

The mathematical process for computing phase quadrature is known as the Hilbert Transform [6].

When you calculate such quadrature, the output is decomposed into two components with a relative phase separation of 90° between them.

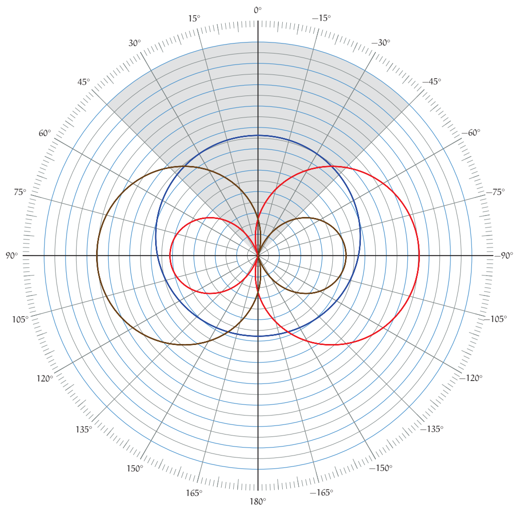

Fig. 1 Σ in blue in the centre, Δ and $-\Delta$ in red and brown in the sides

V. Implementing the Decoder and Quadrature Filter

The decoder is implemented in *gen~* codebox, which is an extension of the Max/MSP patching environment. This code runs at the sample-rate, operating per-sample for minimal latency and precise timing. The coefficients used in this *gen~* algorithm are calculated with double-precision for accuracy.

The filter responsible for signal quadrature is constructed using a cascade of three Biquad all-pass filters, which follow the equations:

$$y[n] = b0w[n] + b1w[n-1] + b2w[n-2] \quad (5)$$

$$w[n] = x[n] - a1w[n-1] - a2w[n-2] \quad (6)$$

These equations are applied separately for the real and imaginary parts of the signal, resulting in a sixth-order filter. This filter configuration can be customized, and the order can be increased by adding more Biquads to the cascade.

The input signal is processed through two filters, each consisting of three Biquads in cascade, following the path shown in Fig. 2. The coefficients for these filters are determined based on the values of control frequency, quality factor, and sample rate. The output of these two filters yields the real and imaginary components of the input signal.

The three Ambisonic *WXY* components are then filtered to produce the real parts *Wr, Xr,* and *Yr* and the imaginary parts *Wi* and *Xi*. In (1) and (2), the real parts are used in both Σ and Δ, while the imaginary part is employed in the *j* arguments within Δ only.

It İs worth noting that the E4L suite internally employs the AmbiX format with ACN channel order and SN3D weighting. Therefore, a $-3dB$ gain adjustment is applied to the W component to convert the content to Fu-Ma format, as mandated by (1) and (2).

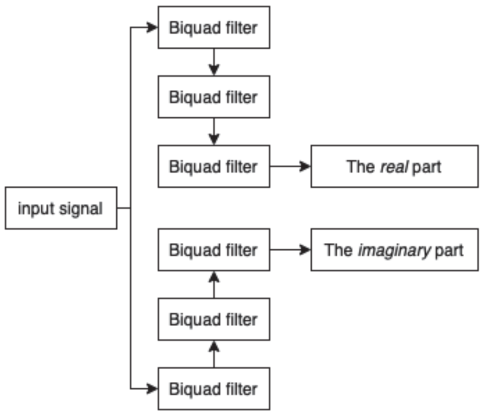

Fig. 2 The quadrature filter

A. Empirical Optimization

Configuring the filter settings is a critical aspect of UHJ decoding. Our empirical approach involved the use of 31 simultaneous oscillators and filters to observe filter behavior across various frequencies concurrently.

We optimized control frequency and quality factor values to achieve quadrature, ensuring precise phase relationships.

Finding the appropriate values for control frequency and quality factor for each biquad in the quadrature filter involves a systematic test procedure. This procedure is designed to optimize the phase relationship between the real and imaginary parts of the signal. Here's a summary of the test procedure:

1) Test Signal Generation: A test signal, denoted as *x*, is generated. This test signal consists of 31 oscillators, each being a sinusoidal waveform with an amplitude of 1. The frequencies of these oscillators are evenly distributed by one-third of an octave, spanning the range from 20 Hz to 20 kHz.
2) Filtering with Biquads: The test signal *x* is simultaneously passed through 31 filters, with each

filter using the same set of control frequency and quality factor values.

3) Monitoring Filter Outputs: The output of each filter is closely monitored during this process.
4) Parameter Search: The goal is to find the optimal values for control frequency and quality factor. These values are determined by observing the outputs of all filters simultaneously. The complex equation below is used for this purpose:

$$Re(x)^2 + Im(x)^2 = 1 \tag{7}$$

In this equation, *Re(x)* represents the real part of the signal *x*, and *Im(x)* represents the imaginary part of the same signal. The values for control frequency and quality factor are adjusted until this equation is satisfied, ensuring the desired phase relationship between the real and imaginary components.

The values discovered through this systematic procedure are documented in Table I.

TABLE I
THE VALUES FOR CONTROL FREQUENCY AND QUALITY FACTOR

frequency Hz	**Q factor**
for the *real* part	
374.1	0.1093
666.8	0.4210
17551	0.5750
for the *imaginary* part	
35.61	0.2571
3723	0.3464
6786	0.1200

B. Setting up the Filter

The implementation of the Biquad filters and the incorporation of Gerzon's equations are characterized as being straightforward, stable, and computationally efficient. However, configuring the filter settings was a time-intensive process, requiring significant effort.

Using 31 simultaneous oscillators and filters can be resource intensive in terms of memory and CPU time, particularly when testing filter responses at high sample rates. Despite the computational cost, this approach is essential because it allows for the observation of filter behavior across different frequencies simultaneously. This simultaneous observation is crucial for achieving accurate filter settings.

In the development of our quadrature filter, we focused on achieving precise phase relationships and spatial characteristics across the frequency spectrum. To reach these goals, we employed an empirical approach involving rigorous testing and optimization efforts. This approach was integral to the configuration of the filter settings, ensuring their accuracy and stability in practice.

Our testing procedure made use of a set of 31 simultaneous oscillators and filters, enabling us to observe the behavior of the filters across various frequencies concurrently. This comprehensive assessment was essential for fine-tune the filter parameters.

The results of our empirical approach are presented in Table II. Notably, these values represent the culmination of our efforts and highlight the significance of the empirical approach in our research.

TABLE II
RESULT OF THE TEST WITH (7)

Frequency Hz	**Min**	**Max**
10	1.021	0.981
20	1.013	0.989
25	1.008	0.993
31.5	1.004	0.995
40	1.002	0.999
50	1.003	0.996
63	1.007	0.997
80	1.010	0.990
100	1.002	0.996
125	1.016	0.984
160	1.003	0.990
200	1.021	0.996
250	1.022	0.982
315	1.008	0.996
400	1.006	0.987
500	1.001	0.994
630	1.006	0.996
800	1.015	0.971
1000	1.016	0.993
1250	1.001	0.998
1600	1.013	0.998
2000	1.005	0.990
2500	1.014	0.998
3150	1.001	0.997
4000	1.011	0.986
5000	1.008	0.979
6300	1.007	0.990
8000	1.012	0.985
10000	1.025	0.964
12500	1.011	0.992
16000	1.148	0.901
20000	1.244	0.719

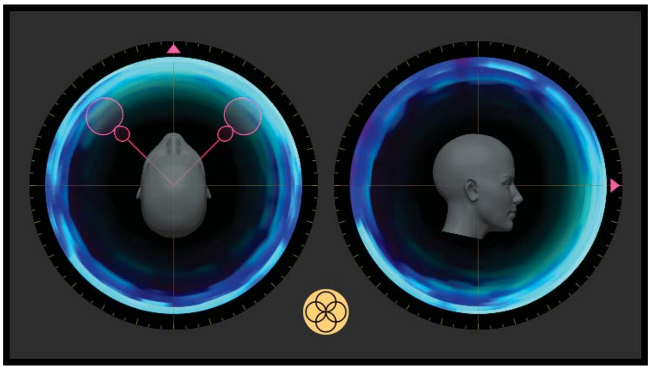

Fig. 3 The soundscape of the First Order Ambisonics

VI. RESULTS AND APPLICATIONS

The results of our empirical approach are presented in Table II: here, the Min and Max values represent the minimum and maximum values observed in the output oscillation, respectively. These values represent the most precise and stable configuration among various iterations. The UHJ decoder has been tested with Ambisonics recordings at sample rates up to 192 kHz, demonstrating stable behavior and rich sonic detail.

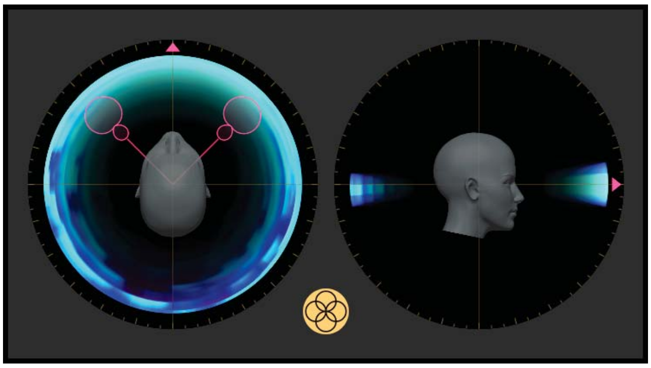

Fig. 4 The reconstructed E-format after conversion to UHJ

A smaller deviation from 1 suggests a smaller error in achieving quadrature, which is desirable for precise phase relationships between the real and imaginary parts. In cases where the value is greater, such as at 10 Hz or above 16,000 Hz, it implies that the phase shift between the imaginary and real parts is less than 90. This can result in a reduction in stereo width at the lowest and highest frequencies. Moreover, this does not cause destructive summing in the signal. These observations are crucial for fine-tuning the quadrature filter parameters to achieve the desired response across the audio spectrum.

VII. Conclusion

Decoding Ambisonics signals into a 2-channel mixing format presents challenges that depend on the intended reproduction system, whether it be stereo speakers or headphones. However, the UHJ format offers a solution that transcends this choice, preserving the entire horizontal plane and offering unique advantages such as head rotation implementation and controls for balancing the omnidirectional W and directional XY components.

The stereo image achieved with UHJ decoding is characterized by clarity, width, and depth, making it an attractive choice for listening. Ambisonics recordings, including those captured with Soundfield microphones and others, can be decoded using UHJ for stereo playback, delivering a wider stereo image without the limitations of binaural techniques. As the UHJ format covers the entire horizontal plane, it is even possible to reconstruct the horizontal Ambisonics plane using the equations provided by Gerzon.

Fig. 3 illustrates the soundscape of a First Order Ambisonics recording before conversion to UHJ stereo, showcasing information spanning a 360° sphere. In Fig. 4, we present the reconstructed E-format, emphasizing that, while some height information may be lost, the horizontal plane is effectively reconstructed. To fully restore the original spherical sound field, complete UHJ decoding utilizes all four *WXYZ* components and returns *LRTQ*. Furthermore, the UHJ format demonstrates potential applications in streaming, where the first two channels are readily available for stereo reproduction, while the two *TQ* components can be employed for decoding surround formats if needed, such as quadriphony or 5.1.

Our proposed algorithm has been rigorously tested with Ambisonics recordings, even at sample rates up to 192 kHz, demonstrating stable performance and delivering rich detail and sonic quality in the output.

The inclusion of this decoder in the E4L suite may lead to further development, including the exploration of higher orders and a broader frequency response for the quadrature filter.

Beyond its technical merits, Ambisonics UHJ has also proven to be a powerful tool in music composition. The production and performance of first author's compositions, such as *Trippin' on the edge of time* and *Emptiness of the hanging*, have provided valuable insights into the entire creative process, from production to the concert hall.

In conclusion, Ambisonics UHJ, when coupled with our decoder, represents a versatile and powerful solution for spatial audio reproduction, offering advantages for both technical applications and artistic endeavors.

Acknowledgment

We wish to acknowledge the Conservatorio Niccolò Piccinni in Bari, Italy, for the support throughout the entirety of this project. This technical and artistic contribution have been pivotal in shaping the outcome of this work.

References

[1] M. Gerzon, "Ambisonics in multichannel broadcasting and video," *Journal of the Audio Engineering Society*, p. 31, 1983.
[2] "Surround Discography website," www.surrounddiscography.com, accessed: 2023-01-24.
[3] "Envelop for Live webpage," https://www.envelop.us/software, accessed: 2023-01-24.
[4] "UHJ decoder for E4L," https://github.com/EnvelopSound, accessed: 2023-01-24.
[5] Julius O. Smith III, *Mathematics of the Discrete Fourier Transform (DFT), with Audio Applications - Second Edition*. W3K Publishing, 2007.
[6] C. B. Gold, A. V. Oppenheim, "Theory and implementation of the discrete Hilbert Transform," 1969.

Alessio Mastrorillo Alessio 'Al Mastro' Mastrorillo (1987) is an electroacoustic music composer and researcher, selected by SMC, ICMC, eviMus, Sin[x]thèsis, Arte e Scienza, UrtiCanti and Electric Sound. He has spent an artistic residency in Tempo Reale (Florence) in 2021. He is graduated Summa Cum Laude in Electroacoustic Music with thesis on Ambisonics and Music Composition at State Conservatory "Piccinni" (Italy). He contributed to the 48° National Congress of Associazione Italiana di Acustica (Matera 2022) and XXIII Colloquio di Informatica Musicale (Ancona 2022). He has been touring as drums player, producing electronic music, short-film scores and engineering live music. His mainly focus is algorithmic composition, Ambisonics and live performances.

Giuseppe Silvi Giuseppe Silvi (1981) is a professor of Electroacoustics at State Conservatory "Piccinni" (Italy). He trained at the S. Cecilia Conservatory in Rome where he studied Electronic Music with Giorgio Nottoli, Nicola Bernardini and Michelangelo Lupone, graduating in 2015. He attended courses with Alvise Vidolin at the Accademia Chigiana and Michelangelo Lupone at the Accademia S. Cecilia. Fundamental to his path were meetings with Piero Schiavoni, Enzo Filippetti, Mario Bertoncini, Walter Branchi, and Giorgio Netti and participation in EMUFest (International Electronic Music Festival of Rome). He collaborates with the Centro Ricerche Musicali in Rome. He is a member of the improvisation group METADIAPASON. He develops a spherical omnidirectional listening system called S.T.ONE. Founder of the LEAP (Laboratorio Elettroacustico Permanente, Rome) space and the SEAM (Sustained ElectroAcoustic Music) project, he curates interpretations of the Roman electroacoustic repertoire to which he tries to look at with perspective by writing music to listen to.

Francesco Scagliola Francesco Scagliola (1971) is a Composer and Sound Artist who graduated with honours in Piano, Composition, and Electroacoustic Music. He specialized in Composition at the National Academy of Santa Cecilia (Rome) and Computer Music at Tempo Reale (Florence). He is a Full Professor of Electroacoustic Music Composition and Coordinator of Computer Music at the School of Electronic Music - State Conservatory "Piccinni" (Italy). He is focused on music production, instrumental and electronic music, and multimedia arts. His research fields are Music Theory and Music Composition (including Algorithmic Composition). Founder of the research group Sin[x]Thésis, he has performed with the Arditti Quartet, L. Bacalov, A. Corghi, S. Reich, and the Swingle Singers.