A 2D Variable-Order, Variable-Decoder, Ambisonics based Music Composition and Production Tool for an Octagonal Speaker Layout

Martin J. Morrell[1] and Joshua D. Reiss[1] *

Centre for Digital Music, Queen Mary University of London, London, UK
martin.morrell@eecs.qmul.ac.uk

Abstract. This paper introduces a music production/composition tool for the spatialisation of sound sources played over an octagonal loudspeaker layout. The tool is based on Ambisonics theory, but does not produce any intermediary B-Format signals. The novel aspects of the tool is that it allows for variable-order and variable-decoder attributes on a per sound source basis. This allows creative control over the sounds' localisation sharpness. Distance including inside the speaker layout source placement and reverberation attributes can be assigned to each sound source to create a final spatial mix. The theory of variable-order, variable-decoder Ambisonics is discussed and the implementation aspects presented. The authors aim to bridge the gap between theory and usage of Ambisonics.

Keywords: Ambsionics, variable-order, variable-decoder, octagon, spatial audio, 2D, 3D.

1 Introduction

Ambisonics is a spatialisation technique for recording, panning and reproducing two and three-dimensional sound sources. Work on Ambisonics is often very much theoretical research and at other times is artists using Ambisonics ready tools to produce work. However the latter usually relies on using software such as Max/MSP to create custom software to then create artistic work or the use of plugins for digital audio workstations that are not well equipped for handling the amount of channels that Ambisonics can produce or the eventual speaker feeds needed. In this paper the authors present the theory behind the creation of a new tool that allows users to send audio from current digital audio workstation projects to be spatialised around an eight speaker octagonal layout. The tool offers some new novel features discussed in-depth in the paper to create variabe-order and variable-decoder based Ambisonics-esque signals. The spatialisation tool is controlled via standard midi protocol and the parameters stored in the same digital audio workstation project.

* This research was supported by the Engineering and Physical Sciences Research Council [grant number EP/P503426/1].

2 Ambsionics Background

Michael Gerzon led the original Ambsionics development team in the 1970s and wrote papers on the subject throughout his life [8–10]. Further work has been done to expand Ambisonics into Higher Order Ambisonics [2–4, 12] and to develop decoders, speaker layouts and evaluation of systems [1, 6, 7, 11, 13, 14, 18]. The basis of Ambisonics is to represent a three-dimensional auditory scene as a field representation that can later be reconstructed for any user loudspeaker layout. An Ambisonics representation is based on a fixed order that is linked to the localisation attributes of sound sources. Ambisonics theory is based on spherical harmonics calculated from legendre polynomials.

$$Y_{mn}^{\sigma(N2D)}(\theta,\phi) = \sqrt{2}\hat{P}_{mn}(\sin\phi) = \begin{cases} \cos n\theta n & \geq 0 \\ \sin n\theta & n < 0 \end{cases} . \tag{1}$$

$$\hat{P}_{mn}(\sin\phi) = \sqrt{(2-\phi_{0,n})\frac{(m-n)!}{(m+n)!}} P_{mn}(\sin\phi) . \tag{2}$$

The above equations use the N2D normalisation scheme. Several schemes exist for Ambisonics and affect the maximum gain of each spherical harmonic. When these are applied to a monaural sound source a sound field representation is created and is known as B-Format. The 2D representation is based only on the angular value θ as $\phi = 0$. The spherical harmonic expansion of the sound field is truncated to a finite representation known as the Ambisonic order M and each prior order m is included, $0 \leq m \leq M$. For each included order m the degrees calculated are $n = \pm m$. Where the total amount of harmonics in the sound field representation is $2M + 1$.

Once encoded, Ambisonics material can be played back over various different loudspeaker layouts using a suitable decoder. The minimum number of loudspeakers to correctly reproduce 2D Ambisonics is $2M+2$ [17]. For a regular layout, i.e. one that has the loudspeakers equally spaced, the angular separation is simply $360^o/L$ where L is the number of loudspeakers for 2D reproduction. For a regular layout the decoder matrix can be calculated by using the Moore-Penrose pseudo-inverse matrix of the spherical harmonics (equal to the source material and appropriate for the amount of loudspeakers) at each loudspeaker position.

$$pinv \begin{pmatrix} Y_{(0,0)}(spk1) & Y_{(1,-1)}(spk1) & Y_{(1,1)}(spk1) & \ldots Y_{(M,m)}(spk1) \\ Y_{(0,0)}(spk2) & Y_{(1,-1)}(spk2) & Y_{(1,1)}(spk2) & \ldots Y_{(M,m)}(spk2) \\ Y_{(0,0)}(spk3) & Y_{(1,-1)}(spk3) & Y_{(1,1)}(spk3) & \ldots Y_{(M,m)}(spk3) \\ \vdots & \vdots & \vdots & \ddots \vdots \\ Y_{(0,0)}(spkN) & Y_{(1,-1)}(spkN) & Y_{(1,1)}(spkN) & \ldots Y_{(M,m)}(spkN) \end{pmatrix} . \tag{3}$$

The given pseudo-inverse decoder results in the standard, rV, decoder matrix. Gerzon specified criteria for low and high frequencies reproduction known as rV and rE vectors [8, 9, 11]. To create a decoder that maximises the rE vector the decoder is then multiplied with gains $g'm$ based on each component's order and the system order.

$$g'm = P_m(\text{largest root of } P_{M+1}) \ . \tag{4}$$

Furthermore the decoding can be changed to what is known as in-phase decoding so that there are no negative gains used to create the sound's directionality.

$$g'm = \frac{M!}{(M+m)!(M-m)!} \ . \tag{5}$$

Ambisonics can be seen as creating a polar pattern of M^{th} order in the direction of the sound source where the polar pattern is sampled by discrete loudspeaker positions. By increasing the amount of loudspeakers the resolution of the polar pattern is increased. In turn, by increasing the order, the directionality is increased and by using different decoders as described above, the rear-lobe is altered.

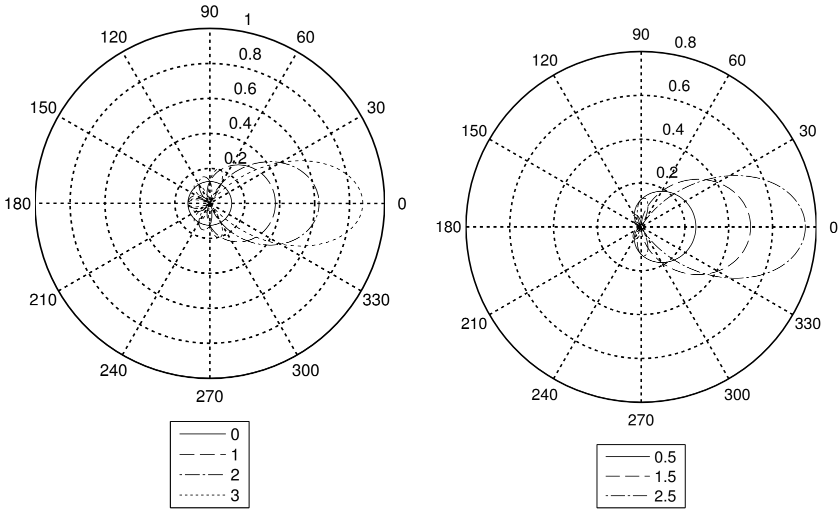
3 Variable-Order and Variable-Decoder Concept

In this section the authors present the novel idea of variable-order and variable-decoder Ambisonics. This concept allows for varying the reproduced polar pattern, and therefore the sharpness of localisation, by setting the order used to a non-integer value. Further to this, the idea of a variable-decoder is discussed that can alter the amount of rear lobe of the sampled polar pattern. The two variables are linked but not interchangeable. The order alters the width of the main lobe, whilst altering the amount of and gain of, the rear lobes. The decoder alters the gain of rear lobes whilst consequently altering the width and gain of the main lobe.

3.1 Variable-Order

The result of encoding a monaural sound source to Ambiosnics B-Format and then decoding it for a loudspeaker layout is equivalent to applying a gain to the monaural sound and sending it to each loudspeaker. Therefore in the authors' approach the audio signal is not converted to B-Format. Instead the gains are calculated numerically and applied based on the octagon layout.

The variable-order is created by calculating the decoders, of same type, for each order. Since we are dealing with an octagonal layout the orders used are 0 through 3. The spherical harmonic values are calculated for all orders for the sound source location θ and speaker gains obtained. By using linear algebra the variable-order can be created by a mixture of 0^{th} and 1^{st}, 1^{st} and 2^{nd}, and 2^{nd} and 3^{rd} speaker gains. Figure 1 a) shows the sampled polar pattern for the whole orders. Figure 1 b) shows the half orders using the variable-order approach. As



(a) Orders 0, 1, 2 and 3

(b) Orders 0.5, 1.5 and 2.5

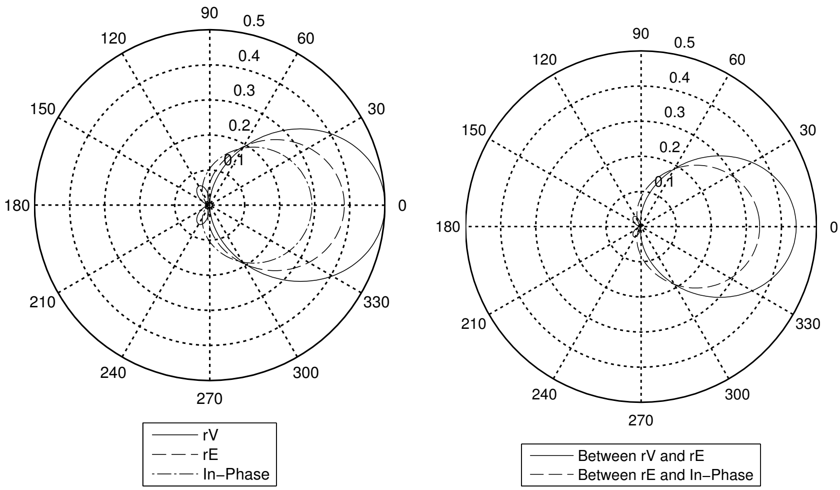
Fig. 1: The reproduced polar pattern of a sound source at $\theta = 0^o$ for Ambisonic orders 0 through 3 are shown in a). The half orders of 0.5, 1.5 and 2.5 are shown in b).

can be expected the polar pattern of half orders is directly between the whole orders. The variable-order approach can be used to create the polar pattern of any decimal value order representation. For an Ambisonics representation the gain of all loudspeakers must equal 1. This has been calculated to be true, but from simple algebra if the two speaker feeds both equal one and the weighting applied equal one, then so must the resultant equal one. This fact is important so that a sound source does not experience an overall gain boost when the variable-order is used as a creative feature.

3.2 Variable-Decoder

Three types of Ambisonics decoders have been presented in section 2 and each is used for a specific purpose. However these decoders offer an aspect of creativity over being able to manipulate the rear lobe of the polar pattern, thus altering the shape of the sound sources' polar pattern.

As for the variable-order concept, the variable-decoder can be calculated in the same manner. By using a weighted ratio that equals 1 of two types of decoder, a variable pattern can be created. The weighting is done between rV and rE decoders and the rE and in-phase decoders. This is because the rE polar pattern lies between the basic and in-phase patterns.



(a) rV, rE and In-Phase decoders (b) Variable-decoders

Fig. 2: The three standard decoder types for order 1.5 are shown on the left and the intermediate decoders on the right.

Figure 2 shows the three decoders for order 1.5 on the left and the decoders half way between the rV snd rE decoders and the rE and in-phase decoders. The variable-decoder lies at the given ratio between the standard decoders.

3.3 Observations

The proposed methodology creates a set of variable-order, varidable-decoder loudspeakers signals for an octagon arrangement of loudspeakers. The end result is sampling at regular intervals of a third oder polar pattern [5]. The resultant gain g_L for loudspeaker at position θ_L can be calculated by equation 6. The sum of the gain of each order must equal one.

$$g_L = a_0 + a_1 \cos(\theta + \theta_L) + a_2 \cos(2(\theta + \theta_L)) + a_3 \cos(3(\theta + \theta_L)) \ . \qquad (6)$$

Therefore the variable-order is equivalent to increasing the next order gain whilst the ratio of the prior orders' gains remains the same. The variable-decoder, is like altering the ratio between the a_0 and a_1 gain coefficients thus changing the base polar pattern, as well as altering the ratio between higher orders.

4 Composition/Production Tool Implementation

The tool to use the variable-order, variable-decoder methodology has been implemented in the Max/MSP 5 software environment for Mac OSX. The tool is designed to receive audio signals from digital audio workstations, e.g. via Jack or Soundflower, for a total of 16 monaural and 4 stereophonic signals. The controls for each channel are sent via midi commands which can be stored in a digital audio workstation project. The authors built User Control Panels for this function for the Cubase/Nuendo environment, but VSTs, AUs or other midi capable software can be used to control the settings for each sound source. The premise for this is that no extra saved data is needed that cannot be stored in a common audio project.

Figure 3 shows the user interface for the tool. The only user definable parameters on the interface are On/Off, midi driver, audio driver and where to save a recorded file. The interface has eight LED style meters for monitoring the signal level going to each loudspeaker so that distortion can be avoided. Since users may not always have an eight speaker layout available, a binaural (over headphones) mix is simultaneously available.

Fig. 3: The user interface for the variable-order, variable-decoder spatialisation tool.

4.1 Tool Features

The novel features in this tool have already been presented in this paper, however, there are some other features that are note worthy. This includes the handling of distance and reverberation.

Distance Distance is a user definable parameter and is accomplished by gain manipulation only. No delay has been included since for music purposes pitch shifting of sound sources will affect the overall tonal effect of the work, alter the speed and therefore ensemble timing of the music and finally can include zipper noise. The $1/r$ inverse law is used to implement the gain change at sources greater than 1.0 where the maximum value is 10. Since the roll off of $1/r$ simulates anechoic conditions, the feature is given as for creative not real-world application. For sources that are placed inside the speaker layout the distance calculation changes to $1 + \cos(90^o r)$ so that infinite gain is not reached. The maximum gain at the central position is 2.0, or approximately +6dB.

Inside Panning Sound sources that have a distance less than 1.0 are placed inside the loudspeaker array. This is done by altering the polar patterns. If the order of reproduction is 1 then this is the same as cancelling out the 1st order spherical harmonics and doubling the zeroth order spherical harmonics [15]. For the case of third order 2D Ambisonics, the maximum allowed in this tool, the inside panning function is expanded. The result is that even orders are doubled and odd orders are cancelled out. This again is all done as numerical and not audio calculations. Figure 4 shows the polar pattern change going from 1.0 to 0.0. The result is strong lobes from opposite poles giving the psychoacoustic illusion of being in the centre of the array.

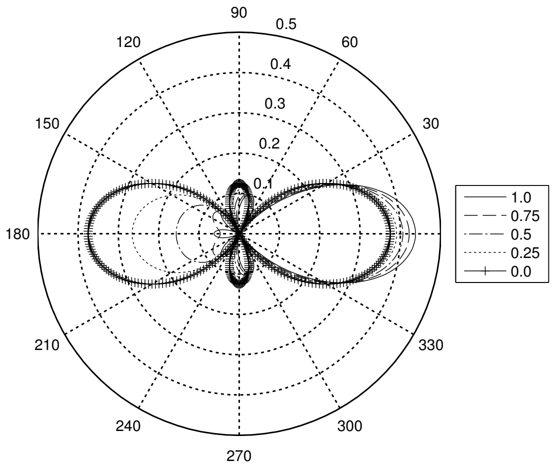


Fig. 4: The change in polar pattern exerted by a third order sound source as it is moved from a distance of 1.0 to 0.0 to be placed in the middle of the speaker array.

Reverberation Reverberation is produced in the tool by transforming the sound source into B-Format and processing it through either the Wigware VST reverberation plugin [18] based on the freeverb algorithm or using a convolution plugin using B-Format impulse responses, such as those freely available [16].

5 Conclusion

The authors have presented a novel approach to implementing the theory of Ambisonics that does not use the intermediary B-Format representations for a fixed octagonal loudspeaker layout. These conditions however mean that composers/artists do not need to worry about designing speaker layouts. Furthermore by fixing the speaker layout of the tool, calculations are done numerically, and a variable-order and decoder is created for each sound source. The result is being able to mix orders to create a composer defined rather than technologically defined sound field that the listener hears. Work is being undertaken to use this tool for an original composition.

References

1. Benjamin, E.: Ambisonic loudspeaker arrays. In: Audio Engineering Society Convention 125 (10 2008)
2. Daniel, J.: Reprèsentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimèdia. Ph.D. thesis, l'Universite Paris 6 (2000)
3. Daniel, J.: Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format. In: Audio Engineering Society Conference: 23rd International Conference: Signal Processing in Audio Recording and Reproduction (5 2003)
4. Daniel, J., Moreau, S.: Further study of sound field coding with higher order ambisonics. In: Audio Engineering Society Convention 116 (5 2004)
5. Eargle, J.: The Microphone Book. Focal Press (2001)
6. Furness, R.K.: Ambisonics-an overview. In: Audio Engineering Society Conference: 8th International Conference: The Sound of Audio (5 1990)
7. Furse, R.W.: Building an openal implementation using ambisonics. In: Audio Engineering Society Conference: 35th International Conference: Audio for Games (2 2009)
8. Gerzon, M.A.: Periphony: With-height sound reproduction. J. Audio Eng. Soc 21(1), 2–10 (1973)
9. Gerzon, M.A.: Practical periphony: The reproduction of full-sphere sound. In: Audio Engineering Society Convention 65 (2 1980)
10. Gerzon, M.A., Barton, G.J.: Ambisonic decoders for hdtv. In: Audio Engineering Society Convention 92 (3 1992)
11. Heller, A., Lee, R., Benjamin, E.: Is my decoder ambisonic? In: Audio Engineering Society Convention 125 (10 2008)
12. Käsbach, J., Favrot, S.: Evaluation of a mixed-order planar and periphonic ambisonics playback implementation. In: Forum Acusticum 2011 (2011)
13. Menzies, D.: W-panning and o-format, tools for object spatialization. In: 22nd International Conference on Virtual, Synthetic and Entertainment Audio (June 2002)
14. Moore, D., Wakefield, J.: The design of ambisonic decoders for the itu 5.1 layout with even performance characteristics. In: Audio Engineering Society Convention 124 (May 2008)
15. Morrell, M.J., Reiss, J.D.: A comparative approach to sound localization within a 3-d sound field. In: Audio Engineering Society Convention 126 (5 2009)
16. Murphy, D.T., Shelley, S.: Openair: An interactive auralization web resource and database. In: Audio Engineering Society Convention 129 (November 2010)
17. Poletti, M.A.: Three-dimensional surround sound systems based on spherical harmonics. J. Audio Eng. Soc 53(11), 1004–1025 (2005), `http://www.aes.org/e-lib/browse.cfm?elib=13396`
18. Wiggins, B.: Has ambisonics come of age? In: Proceedings of the Institute of Acoustics. vol. 30. Pt. 6 (2008)