Ambisonic data exchange formats

Data exchange formats for Ambisonics have undergone radical changes since the early days of four-track magnetic tape. Researchers working on very high-order systems found no straightforward way to extend the traditional formats to suit their needs. Furthermore, there was no widely accepted formulation of spherical harmonics for acoustics, so one was borrowed from chemistry, quantum mechanics, computer graphics, or other fields, each of which had subtly different conventions. This led to an unfortunate proliferation of mutually incompatible ad hoc formats and much head-scratching.

This page attempts to document the different existing formats, their rationales and history, for the terminally curious and those unfortunate enough to have to deal with them in detail.

Most modern applications use ACN and SN3D, although *traditional* first order is still common.

Spherical harmonics in Ambisonics

A common formulation for spherical harmonics in the context of Ambisonics is[1]

$$Y_\ell^m(\theta, \phi) = N_\ell^{|m|} P_\ell^{|m|}(\sin \phi) \cdot \begin{cases} \sin(-m\theta) & \text{if } m < 0, \\ \cos(m\theta) & \text{if } m \geq 0, \end{cases}$$

where Y denotes a spherical harmonic of degree ℓ and index m with a range of $-\ell \leq m \leq +\ell$.

(Note that if $m = 0$, then $\cos m\theta = \cos 0 = 1$.)

N is a normalisation factor (see below), and P_ℓ^m is the associated Legendre polynomial of degree ℓ and order m. The azimuth angle θ is zero straight ahead and increases counter-clockwise. The elevation angle ϕ is zero on the horizontal plane and positive in the upper hemisphere.

Unfortunately, the "Ambisonic order" ℓ is called the *degree* in mathematical parlance, which uses *order* for the "Ambisonic index" m.

Relationship of spherical harmonics and B-format signals

For a source signal S in direction (θ, ϕ), the Ambisonic components B_ℓ^m are given by

$$B_\ell^m = Y_\ell^m(\theta, \phi) \cdot S.$$

If we span a direction vector from the origin towards the source until it intersects the respective spherical harmonic, the length of this vector is the coefficient that gets multiplied with the source signal. Repeat for all spherical harmonics up to the desired Ambisonic order.

Prerequisites for successful data exchange

For successful exchange of Ambisonic material, some software requires the sender and receiver have to agree on the *ordering* of the components, their *normalisation* or *weighting*, and the relative *polarity* of the harmonics.

Since it is possible to omit parts of the spherical harmonic multipole expansion for content that has non-uniform, direction-dependent resolution (known as *mixed-order*), it might also be necessary to define how to deal with *missing components*.
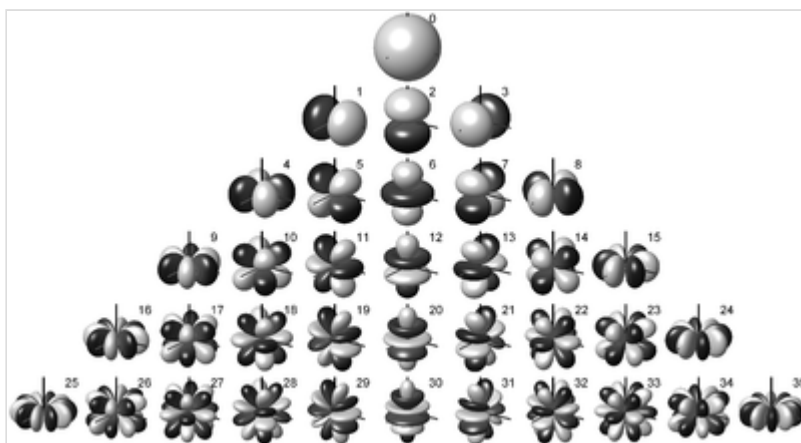
In the case of transmission "by wire", be it an actual digital multichannel link or any number of virtual patchcords within an audio processing engine, these properties must be explicitly matched on both ends, since there is usually no provision for metadata exchange and parameter negotiation. In the case of files, some flexibility might be possible, depending on the file format and the expressiveness of its metadata set.

However, in practice, just two formats are in widespread use. The first is *Furse-Malham higher-order format*, which is an extension of *traditional B-Format*, and the more modern *SN3D*, in *ACN* channel order. In neither case is there any ambiguity about ordering, normalisation, weighting or polarity and it is rare to see cases with missing components. A third format is in limited use: *N3D*, also in ACN channel order.

Component ordering

The *traditional B-format* ($WXYZ$) only concerned itself with zeroth and first Ambisonic order. Because of a strong correspondence between the spherical harmonics and microphone polar patterns, and the fact that those polar patterns have clearly defined directions, it seemed natural to order and name the components in the same way as the axes of a right-hand coordinate system.

Spherical Harmonics up to *Ambisonic order* 5 as commonly displayed, sorted by increasing Ambisonic Channel Number (ACN), aligned for symmetry.

For higher orders, this precedent becomes awkward, because spherical harmonics are most intuitively arranged in symmetric fashion around the single z-rotationally symmetric member *m=0* of each order, with the horizontal sine terms *m<0* to the left, and the cosine terms *m>0* to the right (see illustration).

Furse-Malham

			W_0			
		Y_2	Z_3	X_1		
	V_8	T_6	R_4	S_5	U_7	
Q_{15}	O_{13}	M_{11}	K_9	L_{10}	N_{12}	P_{14}

In *Furse-Malham higher-order format*, an extension of traditional B-format up to third order,[2] orders 2 ($RSTUV$) and 3 ($KLMNOPQ$) begin with their z-rotationally symmetric member and then jump outward right and left (see table), with the horizontal components at the end. Higher-order extensions are trivially defined, but are not used.[2]

SID

			0			
		2	3	1		
	5	7	8	6	4	
10	12	14	15	13	11	9

In his seminal 2001 thesis,[3] Daniel used a three-index nomenclature Y_{mn}^{σ} for the spherical harmonics, which corresponds to $Y_{\ell|m|}^{sgn(m)}$ in the notation used here.[note 1] He implied yet another channel ordering, subsequently developed into an explicit proposal called *SID* for *Single Index Designation*[4] which was adopted by a number of researchers. This scheme is compatible with first-order B-format, and continues to traverse the higher spherical harmonics in the same fashion, with the z-rotationally symmetric component at the end, going through the horizontal components first. It is, however, incompatible with Furse-Malham. SID ordering is not in widespread use.

ACN

			0			
		1	2	3		
	4	5	6	7	8	
9	10	11	12	13	14	15

For future higher-order systems, adoption of the *Ambisonic Channel Number (ACN)*[5] has reached wide consensus. It is determined algorithmically as $ACN = \ell^2 + \ell + m$.

ACN is used widely with SN3D and N3D, below.

More simply:

FuMa = WXYZ | RSTUV | KLMNOPQ

ACN = WYZX | VTRSU | QOMKLNP

SID = WXYZ | UVSTR | PQNOLMK

SID used in iem_ambi in Puredata.

Normalisation

For successful reconstruction of the sound field, it is important to agree on a normalisation method for the spherical harmonic components. The following approaches are common:

maxN

The maxN scheme by Daniel normalizes each single component to never exceed a gain of 1.0 for a panned monophonic source. Malham states that *"[w]hilst this approach is not rigorously "correct" in mathematical terms, it has significant engineering advantages in that it restricts the maximum levels a panned mono source will generate in some of the higher-order channels."*[2] This property is particularly interesting for fixed-point digital interfaces. The maxN weights may be determined by visual inspection up to the third order; above this value the maxima of each polynomial need to be determined explicitly.[2]

MaxN is used in the Furse-Malham format (with the exception of a -3dB correction factor for W, which makes it directly compatible with traditional B-Format). Otherwise, it is not in widespread use.

SN3D

SN3D stands for Schmidt semi-normalisation and is commonly used in geology and magnetics. The weighting coefficients are[6]

$$N_{\ell,m}^{\text{SN3D}} = \sqrt{(2 - \delta_m)\frac{(\ell - |m|)!}{(\ell + |m|)!}}, \delta_m \begin{cases} 1 & \text{if } m = 0 \\ 0 & \text{if } m \neq 0. \end{cases}$$

Originally introduced into Ambisonic use by Daniel, he notes: *"High degree of generality - the encoding coefficients are recursively computable, and the first-order components are unity vectors in their respective directions of incidence".*[7]

With SN3D, unlike N3D, no component will ever exceed the peak value of the 0th order component for single point sources.[1] This scheme has been adopted by the proposed AmbiX format.

SN3D (in the ACN channel order) is in widespread use and a common choice in new software development.

In the Ambix specification paper the term $(2 - \delta_m)$ is replaced with $\dfrac{(2 - \delta_m)}{4\pi}$.

N3D

N3D or full three-D normalisation is the most obvious approach to normalisation. Daniel describes it as follows: *"Orthonormal basis for 3D decomposition. Simple relationship to SN3D [..]. Ensures equal power of the encoded components in the case of a perfectly diffuse 3D field. [..] Obvious significance for solving decoding problems [..] (3D reconstruction)."*[8]

The relation to SN3D is[9]

$$N_{\ell,m}^{\text{N3D}} = N_{\ell,m}^{\text{SN3D}} \sqrt{2\ell + 1}.$$

This normalisation is standard in physics and mathematics and is supported by some Ambisonic software packages. It is used in MPEG-H. However, SN3D is now much more common.

As N3D and SN3D differ only by scaling factors, care is needed when working with both, as it may not be obvious on first listening if an error has been made, particularly on a system with a small number of speakers.

N2D / SN2D

Additionally, two schemes exist which consider only the horizontal components. This has practical advantages for fixed-point media in the common situation where sources are concentrated on the horizontal plane, but the normalisation is somewhat arbitrary and its assumptions do not hold for strongly diffuse soundfields and sound scenes with strong elevated sources. Since Ambisonics is meant to be isotropic and the 2D schemes definitely are not, their use is discouraged.

Polarity

A third complication arises from the quantum mechanical formulation of spherical harmonics, which was adopted by some Ambisonics researchers. It includes a factor of $(-1)^m$, a convention called *Condon–Shortley phase*, which will invert the relative *polarity* of every other component within a given Ambisonic order. The term can be folded both into the formulation of the associated Legendre polynomials or the normalisation coefficient, so it may not always be obvious.

MATLAB and GNU Octave both **include** Condon–Shortley phase in its `legendre(ℓ,X)` functions, but **undo** it by applying the factor again in the Schmidt semi-normalized form `legendre(ℓ,X,'sch')`.[10][11]

Wolfram Language also **includes** C-S phase in its `legendreP(ℓ,X)` implementation,[12] and **retains** it in `SphericalHarmonicY[ℓ,m,θ,φ]`, which is fully normalized.[13] Note that this function returns complex values and uses the physics convention for spherical coordinates where θ is the zenith angle (angle from the positive Z-axis) and ϕ is the azimuth (counter-clockwise angle around the positive Z-axis).

The presence of Condon–Shortley phase in parts of the signal chain usually manifests itself in erratic panning behaviour and increasing apparent source width when going to higher orders, which can be somewhat difficult to diagnose and much harder to eliminate. Hence, its use is ***strongly discouraged*** in the context of Ambisonics.

None of the ambisonic exchange formats described above use Condon–Shortley phase. Polarity is generally only a concern when trying to reconcile theoretical formulations of the spherical harmonics from other academic disciplines.

Reference table of layouts and normalisations

The following table gives an overview of all Ambisonic formats published so far.

- For Furse-Malham (and traditional B-format), sort by FuMa column and multiply the spherical harmonic by the maxN*[note 2] factor.
- For SID according to Daniel, sort by SID and apply the relevant normalisation factor as specified in the file metadata.
- For basic AmbiX, sort by ACN and use the SN3D factor.

Conversion factors can be applied either to the Ambisonic components B or the spherical harmonics Y.

The data is taken from Chapman (2008).[14]

__Please do not rely on this table until it has been thoroughly checked and the "Under construction" notice has been removed.__

B			$Y_\ell^m (\equiv Y_{ACN})$			Conversion factors	
ACN	FuMa	SID	ℓ	m	Spherical harmonic in N3D	to SN3D	to maxN*[note 2]
0	W_0	0	0	0	1	1	$\dfrac{1}{\sqrt{2}}$
1	Y_2	2	1	-1	$\sqrt{3}\sin(\theta)\cos(\phi)$	$\dfrac{1}{\sqrt{3}}$	$\dfrac{1}{\sqrt{3}}$
2	Z_3	3	1	0	$\sqrt{3}\sin(\phi)$	$\dfrac{1}{\sqrt{3}}$	$\dfrac{1}{\sqrt{3}}$
3	X_1	1	1	1	$\sqrt{3}\cos(\theta)\cos(\phi)$	$\dfrac{1}{\sqrt{3}}$	$\dfrac{1}{\sqrt{3}}$
4	V_8	5	2	-2	$\dfrac{\sqrt{15}}{2}\sin(2\theta)\cos^2(\phi)$	$\dfrac{1}{\sqrt{5}}$	$\dfrac{2}{\sqrt{15}}$
5	T_6	7	2	-1	$\dfrac{\sqrt{15}}{2}\sin(\theta)\sin(2\phi)$	$\dfrac{1}{\sqrt{5}}$	$\dfrac{2}{\sqrt{15}}$
6	R_4	8	2	0	$\dfrac{\sqrt{5}}{2}(3\sin^2(\phi)-1)$	$\dfrac{1}{\sqrt{5}}$	$\dfrac{1}{\sqrt{5}}$
7	S_5	6	2	1	$\dfrac{\sqrt{15}}{2}\cos(\theta)\sin(2\phi)$	$\dfrac{1}{\sqrt{5}}$	$\dfrac{2}{\sqrt{15}}$
8	U_7	4	2	2	$\dfrac{\sqrt{15}}{2}\cos(2\theta)\cos^2(\phi)$	$\dfrac{1}{\sqrt{5}}$	$\dfrac{2}{\sqrt{15}}$
9	Q_{15}	10	3	-3	$\sqrt{\dfrac{35}{8}}\sin(3\theta)\cos^3(\phi)$	$\dfrac{1}{\sqrt{7}}$	$\sqrt{\dfrac{8}{35}}$
10	O_{13}	12	3	-2	$\dfrac{\sqrt{105}}{2}\sin(2\theta)\sin(\phi)\cos^2(\phi)$	$\dfrac{1}{\sqrt{7}}$	$\dfrac{3}{\sqrt{35}}$
11	M_{11}	14	3	-1	$\sqrt{\dfrac{21}{8}}\sin(\theta)(5\sin^2(\phi)-1)\cos(\phi)$	$\dfrac{1}{\sqrt{7}}$	$\sqrt{\dfrac{45}{224}}$
12	K_9	15	3	0	$\dfrac{\sqrt{7}}{2}\sin(\phi)(5\sin^2(\phi)-3)$	$\dfrac{1}{\sqrt{7}}$	$\dfrac{1}{\sqrt{7}}$
13	L_{10}	13	3	1	$\sqrt{\dfrac{21}{8}}\cos(\theta)(5\sin^2(\phi)-1)\cos(\phi)$	$\dfrac{1}{\sqrt{7}}$	$\sqrt{\dfrac{45}{224}}$
14	N_{12}	11	3	2	$\dfrac{\sqrt{105}}{2}\cos(2\theta)\sin(\phi)\cos^2(\phi)$	$\dfrac{1}{\sqrt{7}}$	$\dfrac{3}{\sqrt{35}}$
15	P_{14}	9	3	3	$\sqrt{\dfrac{35}{8}}\cos(3\theta)\cos^3(\phi)$	$\dfrac{1}{\sqrt{7}}$	$\sqrt{\dfrac{8}{35}}$
16	ø	17	4	-4	$\dfrac{3}{8}\sqrt{35}\sin(4\theta)\cos^4(\phi)$	$\dfrac{1}{3}$	ø
17	ø	19	4	-3	$\dfrac{3}{2}\sqrt{\dfrac{35}{2}}\sin(3\theta)\sin(\phi)\cos^3(\phi)$	$\dfrac{1}{3}$	ø
18	ø	21	4	-2	$\dfrac{3}{4}\sqrt{5}\sin(2\theta)(7\sin^2(\phi)-1)\cos^2(\phi)$	$\dfrac{1}{3}$	ø

19	ø	23	4	-1	$\dfrac{3}{4}\sqrt{\dfrac{5}{2}}\sin(\theta)\sin(2\phi)(7\sin^2(\phi)-3)$	$\dfrac{1}{3}$	ø
20	ø	24	4	0	$\dfrac{3}{8}(35\sin^4(\phi)-30\sin^2(\phi)+3)$	$\dfrac{1}{3}$	ø
21	ø	22	4	1	$\dfrac{3}{4}\sqrt{\dfrac{5}{2}}\cos(\theta)\sin(2\phi)(7\sin^2(\phi)-3)$	$\dfrac{1}{3}$	ø
22	ø	20	4	2	$\dfrac{3}{4}\sqrt{5}\cos(2\theta)(7\sin^2(\phi)-1)\cos^2(\phi)$	$\dfrac{1}{3}$	ø
23	ø	18	4	3	$\dfrac{3}{2}\sqrt{\dfrac{35}{2}}\cos(3\theta)\sin(\phi)\cos^3(\phi)$	$\dfrac{1}{3}$	ø
24	ø	16	4	-4	$\dfrac{3}{8}\sqrt{35}\cos(4\theta)\cos^4(\phi)$	$\dfrac{1}{3}$	ø

However, please note that only the Furse-Malham and SN3D/ACN encodings are in wide use. (Traditional B-Format is a subset of Furse-Malham.) For both of these encodings, the equations can be expressed directly, without separate normalisation or conversion factors, and there is no ambiguity around ordering.

File formats and metadata

For file-based storage and transmission, additional properties need to be defined, such as the base file format and, if desired, accompanying metadata.

AMB

The `.amb` file format was proposed and defined by Richard Dobson in 2001,[15] based on Microsoft's WAVE_FORMAT_EXTENSIBLE (http://msdn.microsoft.com/en-us/windows/hardware/gg463006.aspx) amendment to the WAV audio file format. It mandates the use of Furse-Malham encoding.

From its parent, it inherits a maximum file size of 4GB, which is a serious limitation for live recording in higher orders.

`.amb` Files are distinguished from other multichannel content by their suffix and by setting the file subtype Globally Unique Identifier in their header data to either of the following values:

- `SUBTYPE_AMBISONIC_B_FORMAT_PCM {00000001-0721-11d3-8644-C8C1CA000000}` for integer samples, or
- `SUBTYPE_AMBISONIC_B_FORMAT_IEEE_FLOAT {00000003-0721-11d3-8644-C8C1CA000000}` for floating point.

The definition mandates that the WAVE_EX `dwChannelMask` must be set to zero. Furthermore, it recommends that the file should contain a `PEAK` chunk, containing the value and position of the highest sample in each channel.

The channels within an `.amb` file are interleaved, and any unused channels are omitted. This makes it possible to identify traditional #H#P mixed-order content by the number of channels present, as per the following table:[15]

.amb to B-format channel mapping

No of chns	Order		Description	Layout of components
	horiz.	vert.		
1	0	0	mono	W
2	½[note 3]	0	M/S stereo	WY
3	1	0	first-order horizontal	WXY
4	1	1	first-order full-sphere	$WXYZ$
5	2	0	second-order horizontal	$WXYUV$
6	2	1	second-order horizontal, first-order full-sphere	$WXYZUV$
7	3	0	third-order horizontal	$WXYUVPQ$
8	3	1	third-order horizontal, first-order full sphere	$WXYZVUPQ$
9	2	2	second-order full-sphere	$WXYZRSTUV$
11	3	2	third-order horizontal, second-order full sphere	$WXYZRSTUVPQ$
16	3	3	third-order full sphere	$WXYZRSTUVKLMNOPQ$

The free and open source C library libsndfile has included `.amb` support since 2007.

Dobson's format has been instrumental in making native Ambisonic content easily accessible to enthusiasts, and to pave the way for research and deployment of Higher-order Ambisonics. While it cannot scale any further than third order and does not accommodate #H#V mixed order sets, its capabilities are more than sufficient for most Ambisonic content in existence today, and backwards-compatibility to `.amb` is an important feature of any real-world Ambisonic workflow.

AmbiX

AmbiX[1] adopts Apple's Core Audio Format or `.caf`. It scales to arbitrarily high orders and has no practically relevant limitation of file size. AmbiX files contain linear PCM data with word lengths of 16, 24, or 32 bit fixed point, or 32 or 64 bit float, at any sample rate valid for `.caf`. It uses ACN channel ordering with SN3D normalisation.

The *basic format* of AmbiX mandates a complete full-sphere signal set, the order of which can be uniquely and trivially deduced from the number of channels. Only the minimum header information required by the `.caf` specification are present and no other metadata is included.

The *extended format* is marked by the presence of a User-Defined Chunk with the UUID

`1AD318C3-00E5-5576-BE2D-0DCA2460BC89`.

(The original specifications used `49454D2E-4154-2F41-4D42-49582F584D4C`, which is an invalid UUID [16]). Additionally, the header now contains an *adaptor matrix* of coefficients, which needs to be applied to the data streams before they can be played back. This matrix provides a generic way of mapping payloads in any previous format and any mix of orders to canonical periphony, ACN ordering and SN3D normalisation. Theoretically, it can even accommodate sound fields that span only subsets of the sphere.

AmbiX was originally proposed at the Ambisonic Symposium 2011, building upon previous work by Travis[17] and Chapman et al.[5]

Notes

1. sgn(x) is the Sign function.
2. MaxN* (starred) denotes maxN normalisation with the additional -3dB correction factor for W.
3. Dobson (2001) uses "1", which would imply a complete set of horizontal components WXY.

References

1. Christian Nachbar; Franz Zotter; Etienne Deleflie; Alois Sontacchi (June 2–3, 2011). *AmbiX – A Suggested Ambisonics Format* (https://iaem.at/ambisonics/proceedings-of-the-ambisonics-symposium-2011/ambix-a-suggested-ambisonics-format). Ambisonics Symposium 2011. Lexington (KY).
2. Malham, David (April 2003). "Higher order Ambisonic systems" (http://decoy.iki.fi/dsound/ambisonic/motherlode/source/higher_order_ambisonics.pdf) (PDF). *Space in Music – Music in Space (Mphil thesis)*. University of York. pp. 2–3. Retrieved 2 November 2007. "Max-Normalisation (MaxN) by Daniel [...] The factors for this can be obtained by inspection up to about third order but over this point it becomes more difficult and requires the maxima of each polynomial to be determined (either mathematically or numerically) explicitly and then inverted."
3. Jérôme Daniel, *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia* (http://gyronymo.free.fr/audio3D/download_Thesis_PwPt.html), Paris 2001, p.151
4. Jérôme Daniel, *Spatial Sound Encoding Including Near Field Effect: Introducing Distance Coding Filters and a Viable, New Ambisonic Format* (http://gyronymo.free.fr/audio3D/publications/AES23%20NFC%20HOA.pdf), 23rd AES Conference, Copenhagen 2003, p. 13
5. Michael Chapman et al., *A standard for interchange of Ambisonic signal sets* (http://iem.kug.ac.at/fileadmin/media/iem/projects/2009/ambixchange09.pdf), Ambisonics Symposium, Graz 2009
6. Nachbar, Zotter, Deleflie, and Sontacchi (2011) lc, p.3, eq(3)
7. Daniel (2001), lc, p.156, translated from French "Grande généricité: calcul récursif des coefficients d'encodage, les composantes d'ordre 1 étant celles du vecteur incidence (unitaire) $\vec{u}$."
8. Daniel (2001) lc, p.156, translated from French "Base orthonormée pour la décomposition 3D. Relation simple á SN3D (facteur $\sqrt{2m+1}$). Assure une puissance égale des composantes encodées dans le cas d'un champ parfaitement diffus 3D (intérêt dans le domaine analogique). Intérêt évident pour la résolution (en 3.3) des problèmes de décodage (restitution 3D)."

9. Daniel (2001), lc, p.150 eq(3.9)
10. MathWorks documentation: legendre (http://nl.mathworks.com/help/matlab/ref/legendre.html)
11. GNU Octave documentation: legendre (http://octave.sourceforge.net/octave/function/legendre.html)
12. Wolfram language documentation: LegendreP (http://reference.wolfram.com/language/ref/LegendreP.html)
13. Wolfram language documentation: SphericalHarmonicY (http://reference.wolfram.com/language/ref/SphericalHarmonicY.html)
14. Michael Chapman, *Ambisonic channel sequence (proposed standard)* (http://ambisonics.ch/standards/channels/) Archived (https://web.archive.org/web/20120930011635/http://ambisonics.ch/standards/channels/) 2012-09-30 at the Wayback Machine
15. Richard Dobson *The AMB Ambisonic File Format* (http://people.bath.ac.uk/masrwd/bformat.html) Archived (https://web.archive.org/web/20140422234520/http://people.bath.ac.uk/masrwd/bformat.html) 2014-04-22 at the Wayback Machine
16. IEM, *AmbiX reference implementation (API documentation)* (https://iem-projects.github.io/ambix/apiref/format.html)
17. Travis, Chris, *A new mixed-order scheme for Ambisonic signals* (http://ambisonics.iem.at/symposium2009/proceedings/ambisym09-travis-newmixedorder.pdf/at_download/file) Archived (https://web.archive.org/web/20091004015535/http://ambisonics.iem.at/symposium2009/proceedings/ambisym09-travis-newmixedorder.pdf/at_download/file) 2009-10-04 at the Wayback Machine, Ambisonics Symposium, Graz 2009

External links

- Notes on Basic Ideas of Spherical Harmonics (http://www.regonaudio.com/SphericalHarmonics.pdf), an introductory text by Robert E. Greene

Retrieved from "https://en.wikipedia.org/w/index.php?title=Ambisonic_data_exchange_formats&oldid=1278451473"