

Springer Topics in Signal Processing

Volume 19

Series Editors

Jacob Benesty, INRS-EMT, University of Quebec, Montreal, QC, Canada

Walter Kellermann, Erlangen-Nürnberg, Friedrich-Alexander-Universität,
Erlangen, Germany

The aim of the Springer Topics in Signal Processing series is to publish very high quality theoretical works, new developments, and advances in the field of signal processing research. Important applications of signal processing will be covered as well. Within the scope of the series are textbooks, monographs, and edited books.

More information about this series at <http://www.springer.com/series/8109>

Franz Zotter · Matthias Frank

Ambisonics

A Practical 3D Audio Theory for Recording,
Studio Production, Sound Reinforcement,
and Virtual Reality

Franz Zotter
Institute of Electronic Music and Acoustics
University of Music and Performing Arts
Graz, Austria

Matthias Frank
Institute of Electronic Music and Acoustics
University of Music and Performing Arts
Graz, Austria



ISSN 1866-2609

ISSN 1866-2617 (electronic)

Springer Topics in Signal Processing

ISBN 978-3-030-17206-0

ISBN 978-3-030-17207-7 (eBook)

<https://doi.org/10.1007/978-3-030-17207-7>

© The Editor(s) (if applicable) and The Author(s) 2019. This book is an open access publication.

Open Access This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The intention of this textbook is to provide a concise explanation of fundamentals and background of the surround sound recording and playback technology Ambisonics.

Despite the Ambisonic technology has been practiced in the academic world for quite some time, it is happening now that the recent ITU,¹ MPEG-H,² and ETSI³ standards firmly fix it into the production and media broadcasting world.

What is more, Internet giants Google/YouTube recently recommended to use tools that have been well adopted from what the academic world is currently using.^{4,5}

Last but most importantly, the boost given to the Ambisonic technology by recent advancements has been in usability: Ways to obtain safe Ambisonic decoders,^{6,7} the availability of higher-order Ambisonic main microphone arrays (Eigenmike,⁸ Zylia⁹) and their filter-design theory, and above all: the usability increased by plugins integrating higher-order Ambisonic production in digital audio workstations or mixers.^{7,10,11,12,13,14,15} And this progress was a great motivation to write a book about the basics.

¹ <https://www.itu.int/rec/R-REC-BS.2076/en>.

² <https://www.iso.org/standard/69561.html>.

³ https://www.techstreet.com/standards/etsi-ts-103-491?product_id=1987449.

⁴ <https://support.google.com/jump/answer/6399746?hl=en>.

⁵ <https://developers.google.com/vr/concepts/spatial-audio>.

⁶ <https://bitbucket.org/ambidecodertoolbox/adt.git>.

⁷ <https://plugins.iem.at/>.

⁸ <https://mhacoustics.com/products>.

⁹ <https://www.zylia.co>.

¹⁰ <http://www.matthiaskronlachner.com/?p=2015>.

¹¹ <http://www.blueripplesound.com/product-listings/pro-audio>.

¹² <https://b-com.com/en/bcom-spatial-audio-toolbox-render-plugins>.

¹³ <https://harpex.net/>.

¹⁴ <http://forumnet.ircam.fr/product/panoramix-en/>.

¹⁵ http://research.spa.aalto.fi/projects/sparta_vsts/.

The book is dedicated to provide a deeper understanding of Ambisonic technologies, especially for but not limited to readers who are scientists, audio-system engineers, and audio recording engineers. As, from time to time, the underlying maths would get too long for practical readability, the book comes with a comprehensive appendix with the beautiful mathematical details.

For a common understanding, the introductory section spans a perspective on Ambisonics from its origins in coincident recordings from the 1930s, to the Ambisonic concepts from the 1970s, and to classical ways of applying Ambisonics in first-order coincident sound scene recording and reproduction that have been practiced from the 1980s on.

In its main contents, this book intends to provide all psychoacoustical, signal processing, acoustical, and mathematical knowledge needed to understand the inner workings of modern processing utilities, special equipment for recording, manipulation, and reproduction in the higher-order Ambisonic format. As advanced outcomes, the aim of the book is to explain higher-order Ambisonic decoding, 3D audio effects, and higher-order Ambisonic recording with microphones or main microphone arrays. Those techniques are shown to be suitable to supply audience areas ranging from studio-sized to hundreds of listeners, or headphone-based playback, regardless whether it is live, interactive, or studio-produced 3D audio material.

The book comes with various practical examples based on free software tools and open scientific data for reproducible research.

Our Ambisonic events experience: In the past years, we have contributed to organizing Symposia on Ambisonics (Ambisonics Symposium 2009 in Graz, 2010 in Paris, 2011 in Lexington, 2012 in York, 2014 in Berlin), demonstrated and brought the technology to various winter/summer schools and conferences (EAA Winter School Merano 2013, EAA Symposium Berlin 2014, workshops and Ambisonic music repertory demonstration at Darmstädter Ferienkurse für Neue Musik in 2014, ICAD workshop in Graz 2015, ICSA workshop 2015 in Graz with PURE Ambisonics night, summer school at ICSA 2017 in Graz, a course at Kraków film music festival 2015, mAmbA demo facility DAGA in Aachen 2016, Al Di Meola's live 3D audio concert hosted in Graz in June 2016, and AES Convention Milano 2018.

In 2017 (ICSA Graz) and 2018 (TMT Cologne), we initiated and organized *Europe's First and Second Student 3D Audio Production Competition* together with Markus Zaunschirm and Daniel Rudrich.

Graz, Austria
February 2019

Franz Zotter
Matthias Frank

Acknowledgements

To our lab and colleagues: Traditionally at the Institute of Electronic Music and Acoustics (IEM), there had been a lot of activity in developing and applying Ambisonics by Robert Höldrich, Alois Sontacchi, Markus Noisternig, Thomas Musil, Johannes Zmöltnig, Winfried Ritsch, even before our active time of research. Most of the developments were done in pure-data, e.g., with [iem_ambi], [iem_bin_ambi], [iem_matrix], CUBEmixer. Dear colleagues deserve to be mentioned who contributed a lot of skill to improve the usability of Ambisonics: Hannes Pomberger and his mathematical talent, Matthias Kronlachner who developed the ambix and mcfx VST plugin suites in 2014, and Daniel Rudrich, who developed the IEM Plugin Suite, that also involves technology that was elaborated together with our colleagues Markus Zaunschirm, Christian Schörkhuber, Sebastian Grill.

We thank you all for your support; it's the best environment to work in!

First readers: We thank Archontis Politis (Aalto Univ., Espoo and Tampere Univ., Finland), Nicolas Epain (b<>com, France), and Matthias Kronlachner (Harman, Germany/US), to be our first critical readers, supplying us with valuable comments.

Open Access Funding: We are grateful about funding from our local government of Styria (Steiermark) Section 8, Office for Science and Research (Wissenschaft und Forschung) that covers roughly half the open access publishing costs. We gratefully thank our University (University of Music and Performing Arts, “Kunstuni”, Graz) for the other half, *transition to open access* library, and vice rectorate for research.



Outline

First-order Ambisonics is nowadays strongly revived by internet technology supported by Google/YouTube, Facebook 360°, 360° audio and video recording and rendering, as well as VR in games. This renaissance lies in its benefits of (i) its compact main microphone arrays capturing the entire surrounding sound scene in only four audio channels (e.g., Zoom H3-VR, Oktava A-Format Microphone, Røde NT-SF1, Sennheiser AMBEO VR Mic.), and (ii) it easily permits rotation of the sound scene, allowing to render surround audio scenes, e.g., on head-tracked headphones, head-mounted AR/VR sets, or mobile devices, as described in Chap. 1.

Auditory events and vector-base panning: Chapter 2 of this book is dedicated to conveying a comprehensive understanding of the localization impressions in multi-loudspeaker playback and its models, followed by Chap. 3 that outlines the essentials of practical vector panning models and their extensions by downmix from imaginary loudspeakers, which are both fundamental to contemporary Ambisonics.

Harmonic functions, Ambisonic encoding and decoding: Based on the ideals of accurate localization with panning-invariant loudness and perceived width, Chap. 4 provides a profound mathematical derivation of higher-order Ambisonic panning functions in 2D and 3D in terms of angular harmonics. These idealized functions can be maximized in their directional focus ($\max\text{-}\mathbf{r}_E$) and they are strictly limited in their directional resolution. This resolution limit entails perfectly well-defined constraints on loudspeaker layouts that make us reach ideal measures for accurate localization as well as panning-invariant loudness and width. And what is highly relevant for practical decoding: All-Round Ambisonic decoding to loudspeakers and TAC/MagLS decoders for headphones are explained in Chap. 4.

The Ambisonic signal processing chain and effects are described in Chap. 5. It illustrates the signal flow from source encoding through Ambisonic bus to decoding and where input-specific or general insert and auxiliary Ambisonic effects are located. In particular, the chapter describes the working principles behind frequency-independent manipulation effects that are either mirroring/rotating/re-mapping, warping, or directionally weighting, or such effects that are frequency-dependent. Frequency-dependent effects can introduce widening, depth or diffuseness, convolution reverb, or feedback-delay-network (FDN)-based diffuse

reverberation. Directional resolution enhancements are outlined in terms of SDM/SIRR pre-processing of recorded reverberation and in terms of available tools such as HARPEX, DirAC, and COMPASS for recorded signals.

Compact higher-order Ambisonic microphones rely on the solutions of the Helmholtz equation, and their processing uses a frequency-independent decomposition of the spherical array signals into spherical harmonics and the frequency-dependent radial-focusing filtering associated with each spherical harmonic order, which yield the Ambisonic signals. The critical part is to handle the properties of radial-focusing filters in the processing of higher-order Ambisonic microphone arrays (e.g., the Eigenmike). To keep the noise level and the sidelobes in the recordings low and a balanced frequency response, a careful way for radial filter design is outlined in Chap. 6.

Compact higher-order loudspeaker arrays oppose the otherwise inwards-oriented Ambisonic surround playback, as described in Chap. 7. This outlooking last chapter discusses IKO and loudspeaker cubes as compact spherical loudspeaker arrays with Ambisonically controlled radiation patterns. In natural environments with acoustic reflections, such directivity-controlled arrays have their own sound-projecting and distance-changing effects, and they can be used to simulate sources of specific directivity patterns.

Contents

1	XY, MS, and First-Order Ambisonics	1
1.1	Blumlein Pair: XY Recording and Playback	2
1.2	MS Recording and Playback	3
1.3	First-Order Ambisonics (FOA)	5
1.3.1	2D First-Order Ambisonic Recording and Playback	6
1.3.2	3D First-Order Ambisonic Recording and Playback	9
1.4	Practical Free-Software Examples	13
1.4.1	Pd with Iemmatrix, Iemlib, and Zexy	13
1.4.2	Ambix VST Plugins	13
1.5	Motivation of Higher-Order Ambisonics	18
	References	21
2	Auditory Events of Multi-loudspeaker Playback	23
2.1	Loudness	24
2.2	Direction	25
2.2.1	Time Differences on Frontal, Horizontal Loudspeaker Pair	25
2.2.2	Level Differences on Frontal, Horizontal Loudspeaker Pair	25
2.2.3	Level Differences on Horizontally Surrounding Pairs	27
2.2.4	Level Differences on Frontal, Horizontal to Vertical Pairs	28
2.2.5	Vector Models for Horizontal Loudspeaker Pairs	28
2.2.6	Level Differences on Frontal Loudspeaker Triangles	31
2.2.7	Level Differences on Frontal Loudspeaker Rectangles	31
2.2.8	Vector Model for More than 2 Loudspeakers	32
2.2.9	Vector Model for Off-Center Listening Positions	32

2.3	Width	34
2.3.1	Model of the Perceived Width	35
2.4	Coloration	36
2.5	Open Listening Experiment Data	38
	References	38
3	Amplitude Panning Using Vector Bases	41
3.1	Vector-Base Amplitude Panning (VBAP)	42
3.2	Multiple-Direction Amplitude Panning (MDAP)	44
3.3	Challenges in 3D Triangulation: Imaginary Loudspeaker Insertion and Downmix	48
3.4	Practical Free-Software Examples	50
3.4.1	VBAP/MDAP Object for Pd	50
3.4.2	SPARTA Panner Plugin	51
	References	52
4	Ambisonic Amplitude Panning and Decoding in Higher Orders	53
4.1	Direction Spread in First-Order 2D Ambisonics	54
4.2	Higher-Order Polynomials and Harmonics	57
4.3	Angular/Directional Harmonics in 2D and 3D	58
4.4	Panning with Circular Harmonics in 2D	58
4.5	Ambisonics Encoding and Optimal Decoding in 2D	61
4.6	Listening Experiments on 2D Ambisonics	61
4.7	Panning with Spherical Harmonics in 3D	67
4.8	Ambisonic Encoding and Optimal Decoding in 3D	71
4.9	Ambisonic Decoding to Loudspeakers	72
4.9.1	Sampling Ambisonic Decoder (SAD)	72
4.9.2	Mode Matching Decoder (MAD)	73
4.9.3	Energy Preservation on Optimal Layouts	73
4.9.4	Loudness Deficiencies on Sub-optimal Layouts	74
4.9.5	Energy-Preserving Ambisonic Decoder (EPAD)	74
4.9.6	All-Round Ambisonic Decoding (AllRAD)	75
4.9.7	EPAD and AllRAD on Sub-optimal Layouts	77
4.9.8	Decoding to Hemispherical 3D Loudspeaker Layouts	77
4.10	Practical Studio/Sound Reinforcement Application Examples	81
4.11	Ambisonic Decoding to Headphones	85
4.11.1	High-Frequency Time-Aligned Binaural Decoding (TAC)	88
4.11.2	Magnitude Least Squares (MagLS)	89
4.11.3	Diffuse-Field Covariance Constraint	90

4.12	Practical Free-Software Examples	91
4.12.1	Pd and Circular/Spherical Harmonics	91
4.12.2	Ambix Encoder, IEM MultiEncoder, and IEM AllRADecoder	92
4.12.3	Reaper, IEM RoomEncoder, and IEM BinauralDecoder	95
	References	96
5	Signal Flow and Effects in Ambisonic Productions	99
5.1	Embedding of Channel-Based, Spot-Microphone, and First-Order Recordings	101
5.2	Frequency-Independent Ambisonic Effects	103
5.2.1	Mirror	104
5.2.2	3D Rotation	106
5.2.3	Directional Level Modification/Windowing	107
5.2.4	Warping	109
5.3	Parametric Equalization	110
5.4	Dynamic Processing/Compression	111
5.5	Widening (Distance/Diffuseness/Early Lateral Reflections)	112
5.6	Feedback Delay Networks for Diffuse Reverberation	114
5.7	Reverberation by Measured Room Impulse Responses and Spatial Decomposition Method in Ambisonics	116
5.8	Resolution Enhancement: DirAC, HARPEX, COMPASS	119
5.9	Practical Free-Software Examples	120
5.9.1	IEM, ambix, and mcfx Plug-In Suites	120
5.9.2	Aalto SPARTA	125
5.9.3	Røde	126
	References	127
6	Higher-Order Ambisonic Microphones and the Wave Equation (Linear, Lossless)	131
6.1	Equation of Compression	132
6.2	Equation of Motion	132
6.3	Wave Equation	133
6.3.1	Elementary Inhomogeneous Solution: Green's Function (Free Field)	133
6.4	Basis Solutions in Spherical Coordinates	135
6.5	Scattering by Rigid Higher-Order Microphone Surface	137
6.6	Higher-Order Microphone Array Encoding	139
6.7	Discrete Sound Pressure Samples in Spherical Harmonics	141
6.8	Regularizing Filter Bank for Radial Filters	142
6.9	Loudness-Normalized Sub-band Side-Lobe Suppression	145
6.10	Influence of Gain Matching, Noise, Side-Lobe Suppression	146

6.11	Practical Free-Software Examples	148
6.11.1	Eigenmike Em32 Encoding Using Mcfx and IEM Plug-In Suites	148
6.11.2	SPARTA Array2SH	150
	References	151
7	Compact Spherical Loudspeaker Arrays	153
7.1	Auditory Events of Ambisonically Controlled Directivity	154
7.1.1	Perceived Distance	154
7.1.2	Perceived Direction	154
7.2	First-Order Compact Loudspeaker Arrays and Cubes	155
7.3	Higher-Order Compact Spherical Loudspeaker Arrays and IKO	157
7.3.1	Directivity Control	159
7.3.2	Control System and Verification Based on Measurements	162
7.4	Auditory Objects of the IKO	164
7.4.1	Static Auditory Objects	164
7.4.2	Moving Auditory Objects	165
7.5	Practical Free-Software Examples	166
7.5.1	IEM Room Encoder and Directivity Shaper	166
7.5.2	IEM Cubes 5.1 Player and Surround with Depth	167
7.5.3	IKO	169
	References	169
	Appendix	171