

AMBISONIC DECODING WITH AND WITHOUT MODE-MATCHING: A CASE STUDY USING THE HEMISPHERE

Franz Zotter, Hannes Pomberger *

University of Music and Performing Arts
Institute of Electronic Music and Acoustics
Inffeldgasse 10, 8010, Graz, Austria
{zotter, pomberger}@iem.at

Markus Noisternig

IRCAM – CNRS UMR STMS
1 place Igor-Stravinsky, 75004, Paris, France
markus.noisternig@ircam.fr

ABSTRACT

Accurate decoding and most notably mode-matching has always been a matter of concern within the Ambisonics community; it can be further expected to play a major role in future discussion and research. Specifically, ambisonic decoding by mode-matching attempts to perfectly reconstruct incident sound fields using a surrounding spherical arrangement of loudspeakers. This reproduction is valid locally within a bounded central area, the *sweet-area*. Surprisingly, many (experienced) listeners have reported good reproduction quality and fair localization accuracy of real-world Ambisonics systems even outside this sweet area. Hence in practice, mode-matching decoding performs better than expected from theory, but it frequently poses numerical problems for incomplete or non-uniform spherical loudspeaker arrangements. To cope with these issues, this article presents, discusses, and compares several alternative Ambisonics decoding strategies with and without mode-matching in terms of simple quality criteria.

1. INTRODUCTION

Surround sound reproduction can be most accurately achieved by covering the entire angular domain around the listening area with as many loudspeakers as possible. This "very-high-order" Ambisonics approach theoretically provides perfect holophony in a dedicated central area of the reproduction room, the sweet-spot. This area is often referred to as sweet-area to express its spatial extent.

Real-world systems are often compromised by practical limitations, such as non-ideal loudspeaker positions and arrays covering only parts of a full sphere, limited number of independently driven loudspeaker channels, numerical instabilities of the decoder, and the acoustics of the reproduction room.

Referring to listening experiments and user experiences reported in literature, Ambisonics seems to work reasonably well even with non-ideal loudspeaker arrangements and with moderate reproduction orders. The sound reproduction quality of Ambisonics outside the sweet-area is usually not perceived as distracting or annoying. Some experiments show that the number of loudspeakers should not exceed by far those required by the reproduction order; too many loudspeakers result in audible artifacts due to phase distortions around the sweet area,

cf. [1, 2, 3, 4, 5]. In [1], loudspeaker signals have been time-aligned with reference to the center of the reproduction sphere. Listeners reported disturbing phasing effects and close-to-head localization of virtual sound-sources. Another experiment [6] shows that room reverberation can mask audible phase distortions.

Incomplete spherical loudspeaker arrangements further complicate the use of mode-matching for Ambisonics decoding, e.g. distance coding is no longer feasible and spherical harmonics become linearly dependent (numerically unstable).

This article discusses various decoding strategies for hemispherical loudspeaker arrays, as a particular case study for arrays that only cover parts of the sphere.

As a starting point, the holophonic mode-matching approach for a hemispherical loudspeaker setup is discussed in detail. Further decoder variants are proposed that explicitly avoid exact matching of the finite-order hemispherical modes. These variants include: order-weighting using an adapted Kaiser-window, regularized mode-matching decoders, Musil decoder, and a weighted direct-sampling decoder. Before introducing and evaluating the different decoders the mathematical/physical background is discussed.

Section two explains the theory of holophonic Ambisonics derived from the wave-equation and a continuous spherical distribution of point sources, cp. [7]. It furthermore explains how to obtain orthonormal band-limited modes of a partial-sphere or a hemisphere by regularization of the Gram-matrix of the spherical harmonics, cp. [8].

The **third section** describes mode-matching (or *modal source-strength matching*, cp. [7]), which reconstructs ambisonic signals on the whole sphere, i.e. on uniform spherical arrangements of discrete loudspeakers.

Section four summarizes practical findings and issues with this kind of sound-field reproduction.

In **section five**, we address practical issues by proposing possible improvements of Ambisonics decoding including spherical smoothing functions.

To evaluate and compare the different decoders, **section six** defines objective quality measures that cover three independent properties of Ambisonics: mode-matching quality, constant power, and spatial resolution.

Eventually, **section seven** presents *five* different decoding methods and techniques applicable to *hemispherical layouts of loudspeakers*. Most of them are based on the theory and improvements presented in sections 2, 3, and 5.

Our contribution concludes with a brief report and case

* This work is funded in part by the Austrian Research Promotion Agency (FFG), the Styrian Government and the Styrian Business Promotion Agency (SFG) under the COMET program.

study in **section eight**, applying the measures introduced in section 4; the results of preliminary listening sessions with a 24 channel hemispherical loudspeaker array are used to briefly discuss audible artifacts of different decoders.

2. CONTINUOUS AMBISONICS THEORY

Ambisonics can be regarded as discretization of a spherical distribution of sources at a given radius r_L driven by the angular amplitude distribution $f(\boldsymbol{\theta})$, cf. [9]. In the following, the Cartesian direction vector is defined as $\boldsymbol{\theta} = (\cos(\varphi) \sin(\vartheta), \sin(\varphi) \sin(\vartheta), \cos(\vartheta))^T$ and depends on the azimuth and zenith angle, φ and ϑ respectively. The Ambisonics approach can be derived from the nonhomogeneous Helmholtz-equation with the angular excitation $f(\boldsymbol{\theta})$

$$(\Delta + k^2)p = -f(\boldsymbol{\theta}) \frac{\delta(r - r_L)}{r^2}. \quad (1)$$

As discussed in [7, 10], the entire sound field inside a source-free volume can be reproduced by controlling $f(\boldsymbol{\theta})$ on the boundary surface. This is also referred to as *holophony*.

The modal solutions of this equation for the sound pressure are given as

$$p = \sum_{n=0}^{\infty} \sum_{m=-n}^n -ik j_n(kr) h_n(kr_L) Y_n^m(\boldsymbol{\theta}) \phi_{nm}, \quad (2)$$

wherein $j_n(kr)$ and $h_n(kr)$ are the spherical Bessel and Hankel functions, respectively, k is the wave-number, and $Y_n^m(\boldsymbol{\theta})$ a spherical harmonic, see Fig. 1. ϕ_{nm} are the coefficients of the spherical harmonics expansion

$$f(\boldsymbol{\theta}) = \sum_{n=0}^{\infty} \sum_{m=-n}^n Y_n^m(\boldsymbol{\theta}) \phi_{nm}, \quad (3)$$

The coefficients ϕ_{nm} are the discrete ambisonic signals in the frequency-domain $\phi_{nm} = \phi_{nm}(\omega) = \mathcal{DTFT}\{\phi_{nm}(t)\}$. Eq. (2) describes the expansion of the wave field for decoding ambisonic signals as discussed in the later sections.

2.1. Vector notation

It is common practice to combine the coefficients ϕ_{nm} and base-functions $Y_n^m(\boldsymbol{\theta})$ into vectors

$$\boldsymbol{\phi} = \text{vec} \{ \phi_{nm} \}_{n=0 \dots \infty}^{m=-n \dots n},$$

and $\mathbf{y}(\boldsymbol{\theta}) = \text{vec} \{ Y_{nm}(\boldsymbol{\theta}) \}_{n=0 \dots \infty}^{m=-n \dots n}$.

Eq. (3) can be written in vector notation

$$f(\boldsymbol{\theta}) = \mathbf{y}^T(\boldsymbol{\theta}) \boldsymbol{\phi}. \quad (4)$$

Multiplying Eq. (4) by $\mathbf{y}(\boldsymbol{\theta})$ and integration over a spherical surface S^2 yields the transform of $f(\boldsymbol{\theta})$ into spherical harmonics

$$\int_{S^2} f(\boldsymbol{\theta}) \mathbf{y}(\boldsymbol{\theta}) d\boldsymbol{\theta} = \underbrace{\int_{S^2} \mathbf{y}(\boldsymbol{\theta}) \mathbf{y}^T(\boldsymbol{\theta}) d\boldsymbol{\theta}}_{:=\mathbf{G}} \boldsymbol{\phi}. \quad (5)$$

The matrix \mathbf{G} denotes the Gram-matrix of $\mathbf{y}(\boldsymbol{\theta})$. Integrating all pairs of spherical harmonics over the unit sphere $S^2 = \mathbb{S}^2$ yields

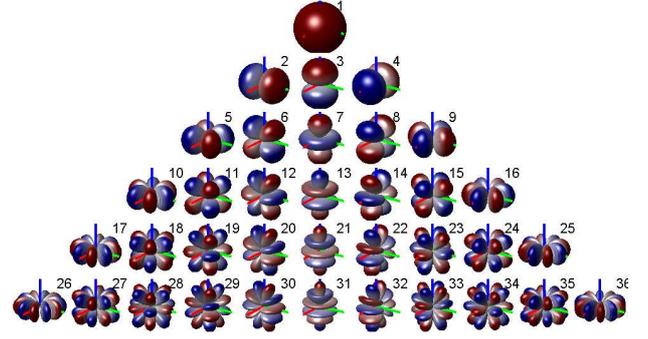


Figure 1: Balloon plots of real-valued spherical harmonics $Y_n^m(\boldsymbol{\theta})$ of orders $n \leq N = 5$ that show $|Y_n^m(\boldsymbol{\theta})|$ as radius and $\text{sign}\{Y_n^m(\boldsymbol{\theta})\}$ as color. The labels of the spherical harmonics correspond to $n^2 + n + m + 1$.

the matrix $\mathbf{G} = \mathbf{I}$, i.e. they are orthonormal. The transform reduces to

$$\boldsymbol{\phi} = \int_{S^2} f(\boldsymbol{\theta}) \mathbf{y}(\boldsymbol{\theta}) d\boldsymbol{\theta}. \quad (6)$$

In the following, a simplified notation is used for better readability

$$\text{diag}\{\text{vec}\{h_n(kr)\}_{n=0 \dots \infty}^{m=-n \dots n}\} = \text{diag}\{h_n(kr)\} \quad (7)$$

so that Eq. (2) reads as

$$p = -ik \mathbf{y}^T(\boldsymbol{\theta}) \text{diag}\{j_n(kr) h_n(kr_L)\} \boldsymbol{\phi}. \quad (8)$$

2.2. Encoding of point sources into Ambisonics

Point sources are used as a model of virtual sound sources and therefore represent (idealized) loudspeakers on the sphere; this corresponds to the discretization as mentioned above. A point source at the position $(\boldsymbol{\theta}_s, r_L)$ is mathematically described by an angular Dirac-delta function, which is non-zero at $\boldsymbol{\theta}_s$

$$f_s(\boldsymbol{\theta}) = \delta(\boldsymbol{\theta}_s^T \boldsymbol{\theta} - 1). \quad (9)$$

Inserting the above equation into Eq. (6), the spherical harmonics coefficient vector $\boldsymbol{\phi}_s$ of a point source located at $\boldsymbol{\theta}_s$ yields

$$\boldsymbol{\phi}_s = \int_{S^2} \delta(\boldsymbol{\theta}_s^T \boldsymbol{\theta} - 1) \mathbf{y}(\boldsymbol{\theta}) d\boldsymbol{\theta} = \mathbf{y}(\boldsymbol{\theta}_s). \quad (10)$$

2.3. Holophony of point-sources with distance coding

Ambisonics creates a perfect holophonic image of virtual point-sources at arbitrary positions r_s and $\boldsymbol{\theta}_s$ when distance coding is taken into account [11, 12]. The target modal source-strength distribution is inserted into Eq. (8) as

$$\boldsymbol{\phi}_s = \text{diag} \left\{ \frac{h_n(kr_s)}{h_n(kr_L)} \right\} \mathbf{y}(\boldsymbol{\theta}_s). \quad (11)$$

Obviously, in contrast to Eq. (10), the distance encoding of point sources is frequency-dependent.

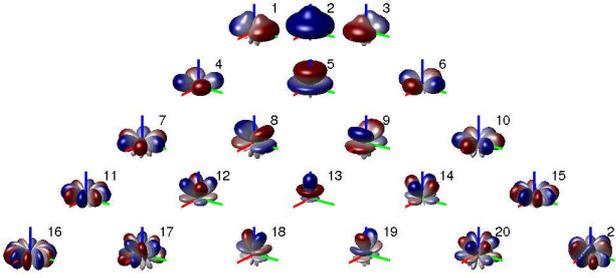


Figure 2: 21 orthogonal base-functions of limited order $n \leq 5$ on the hemisphere. These functions have been found according to [8], using an angle $\vartheta_{\max} = 104^\circ$ and a regularization threshold of $1/1.4$ from the largest eigenvalue.

2.4. Continuous Ambisonics on a bounded spherical domain

The existence of the spherical harmonics transform according to Eq. (5)

$$\phi = \mathbf{G}^{-1} \int_{S^2} f(\theta) \mathbf{y}(\theta) d\theta \quad (12)$$

is mandatory for encoding and decoding of Ambisonics but depends on the invertibility of the Gram-matrix \mathbf{G} . This inversion is often infeasible for an Ambisonics-layout on a partial-sphere $S^2 \subset S^2$.

As proposed in [8], inversion of the ill-conditioned Gram-matrix can be done in two steps. First, it is decomposed by eigen-decomposition $\mathbf{G} = \mathbf{V} \text{diag}\{\lambda\} \mathbf{V}^T$ into pairs of eigenvectors $\{\mathbf{v}_i \in \mathbf{V}\}$ and eigenvalues $\{\lambda_i \in \lambda\}$. Taking only eigenvalues above a certain threshold

$$\lambda_c \subset \lambda : \lambda_i > c \cdot \max(\lambda), \quad (13)$$

achieves regularization, where $c < 1$ denotes the regularization parameter. The regularized expression is inverted by $\mathbf{V}_c^T \text{diag}\{\lambda_c\}^{-1} \mathbf{V}_c$. The matrix \mathbf{V}_c contains the eigenvectors associated to the remaining eigenvalues.

In order to provide a better understanding, [8] outlines that the usage of a new set of base-functions corresponds to regularization

$$\tilde{\mathbf{y}}(\theta) = \mathbf{R} \mathbf{y}(\theta) \quad (14)$$

$$\text{with } \mathbf{R} = \text{diag}\{\lambda_c\}^{-\frac{1}{2}} \mathbf{V}_c^T, \quad (15)$$

which is an orthonormal basis on S^2 as $\tilde{\mathbf{G}} = \int_{S^2} \tilde{\mathbf{y}}(\theta) \tilde{\mathbf{y}}^T(\theta) d\theta = \mathbf{I}$.

The functions plotted in Fig. 2 have been found for the hemisphere; QR matrix triangulation was applied to decompose the matrix \mathbf{R} , as outlined in [8, 13], enabling to group the functions according to their similarities to spherical harmonics.

2.4.1. Encoding into new base-functions

For transcoding the spherical harmonics coefficients ϕ to the new base-functions, the orthogonality of the new functions can be exploited. Expanding given spherical harmonic coefficients

to $f(\theta) = \mathbf{y}^T(\theta) \phi$ and integrating them with $\tilde{\mathbf{y}}(\theta)$ over S^2 yields

$$\begin{aligned} \tilde{\phi} &= \int_{S^2} \mathbf{R} \mathbf{y}(\theta) \mathbf{y}^T(\theta) \phi d\theta \\ &= \mathbf{R} \mathbf{G} \phi = (\mathbf{R}^\dagger)^T \phi, \end{aligned} \quad (16)$$

with \mathbf{R}^\dagger denoting the right inverse of \mathbf{R} , i.e. $\mathbf{V}_c \text{diag}\{\lambda_c\}^{\frac{1}{2}}$.

2.4.2. Notes on distance coding on the hemisphere

It is reasonable that distance coding only works for loudspeaker arrangements on the whole sphere; it could not be applied to hemispherical setups. A horizontal loudspeaker array, i.e. a ring of loudspeakers, can be considered as limit case of a partial sphere. Distance coding is accessible to such arrays by reformulation of the problem in 2D and compensation of real-world 3D sources [14]. Nevertheless, distance coding for the hemisphere is not further regarded within this article.

3. AMBISONICS DECODING ON THE SPHERE: HOLOPHONY WITH LOUSPEAKERS

As stated above, to obtain holophony one has to control the angular amplitude distribution $f(\theta)$ of a continuous spherical source arrangement. Real-world Ambisonics playback facilities consist of a spherical arrangement of loudspeakers at discrete locations $\{\theta_l\}_{l=1 \dots L}$, which are controlled by their respective gains g_l . Assuming that the loudspeakers are point sources, cf. Eq. (9), the angular amplitude-distribution of the playback facility reads as

$$\hat{f}(\theta) = \sum_{l=1}^L g_l \delta(\theta_l^T \theta - 1), \quad (17)$$

or equivalently in the spherical harmonics/Ambisonics domain

$$\hat{\phi} = \sum_{l=1}^L g_l \mathbf{y}(\theta_l).$$

This linear combination of vectors is usually written in matrix-vector notation

$$\hat{\phi} = \mathbf{Y} \mathbf{g}, \quad (18)$$

$$\text{with } \mathbf{g} = [g_1, \dots, g_L]^T$$

$$\text{and } \mathbf{Y} = [\mathbf{y}(\theta_1), \dots, \mathbf{y}(\theta_L)].$$

The weight-vector \mathbf{g} is unknown and will be derived in the following paragraphs.

3.1. Mode-matching

To approximate the sound field of any continuous virtual source distribution $f_s(\theta)$ using Ambisonics with a loudspeaker distribution $\hat{f}(\theta)$, one could simply match the modal source-strength coefficients $\phi_s \stackrel{!}{=} \hat{\phi}$.

However, it is physically impossible to create an equivalent image of a point-source s located at θ_s without a loudspeaker at exactly the same angular position:

$$\delta(\theta_s^T \theta - 1) \neq \sum_{l=1}^L g_l \delta(\theta_l^T \theta - 1) \quad \forall \theta_s \notin \{\theta_l\}. \quad (19)$$

In this sense the matching conditions $\phi_s \stackrel{!}{=} \hat{\phi}$ are not applicable. In practice, a spatial band limitation, *i.e.* a limited angular resolution, is assumed. A band-limitation operator \mathcal{B}_N is applied in order to truncate the spherical harmonics expansion of $f(\theta)$ to $n \leq N$

$$\begin{aligned} \mathcal{B}_N \{f(\theta)\} &:= \mathbf{y}_N^T(\theta) \phi_N, & (20) \\ \text{with } \phi_N &= \text{vec} \{ \phi_{nm} \}_{n=0 \dots N}^{m=-n \dots n}, \\ \text{and } \mathbf{y}_N(\theta) &= \text{vec} \{ Y_{nm}(\theta) \}_{n=0 \dots N}^{m=-n \dots n}. \end{aligned}$$

Hence, the expansion coefficients and spherical harmonics of finite order are indicated by the subscript N .

This truncation "blurs" the given discrete directions and the matching condition results in

$$\mathcal{B}_N \{ \hat{f}(\theta) \} \stackrel{!}{=} \mathcal{B}_N \{ f_s(\theta) \}, \quad (21)$$

$$\Rightarrow \hat{\phi}_N \stackrel{!}{=} \phi_{s,N}. \quad (22)$$

With reference to Eqs. (10) and (18), the loudspeaker gains \mathbf{g} must fulfill the following matrix equation

$$\mathbf{Y}_N \mathbf{g} \stackrel{!}{=} \phi_{s,N}, \quad (23)$$

and the *decoding* matrix-equation that fulfills Eq. (23) writes as

$$\begin{aligned} \mathbf{g} &= \mathbf{D} \phi_{s,N}, & (24) \\ \text{with } \mathbf{D} &= \mathbf{Y}_N^T \mathbf{G}_d^{-1}, \\ \text{and } \mathbf{G}_d &= \mathbf{Y}_N \mathbf{Y}_N^T. \end{aligned}$$

\mathbf{D} denotes the right inverse of \mathbf{Y}_N , so that $\mathbf{Y}_N \mathbf{D} = \mathbf{I}$. The numerical stability of \mathbf{D} depends on the inversion of \mathbf{G}_d , the Gram-matrix of the discrete system \mathbf{Y}_N .

In contrast to the above-mentioned band-limitation, physical loudspeakers correspond to band-unlimited source-strengths $\hat{\phi}$. Despite holophony is achieved in the band-limited $n \leq N$ subspace, the uncontrolled higher-order components need to be addressed as well.

4. PRACTICAL ISSUES WITH AMBISONICS

The holophonic Ambisonics approach relies on mode-matching, but decoding is frequently a numerical challenge, especially for incomplete or non-uniform loudspeaker layouts on the sphere [8, 10].

Successful decoding of Ambisonics provides full control over the amplitude distribution at low orders. Nevertheless, it does not control the spherical harmonics at high orders $n > N$, which results in angular aliasing in $\hat{\phi}$.

4.1. Spatial aliasing and the sweet-area

A description of the sound field created by mode-matching is obtained by inserting the modal source-strengths of the loudspeakers Eq. (18) into the solution of the wave-equation Eq. (8). Decoders that fulfill Eq. (23) accurately reproduce the spherical modes of an incident sound field for $n \leq N$. The angular aliasing at higher orders $n > N$ creates artifacts (spatial aliasing) in the sound field.

One should consider that low order spherical modes can only describe *particular* sound fields. At high orders the spherical Bessel functions $j_n(kr)$ vanish for small arguments kr . Hence,

the aliased and band-limited components perfectly reconstruct the desired sound-field but only within the sweet area. The frequency-dependent radius of this holophonic sweet-area can be estimated by $r_{\max}/\lambda \approx N/6$, cp. [15, 16, 7].

Taking all the above-mentioned into account, it becomes clear that for $N < 20$ the sweet-area is smaller than the diameter of a listener's head for the high frequencies within the range of audibility. In practice, Ambisonics is generally used for mid to large scale concert venues, covering more than 100 listeners. Therefore, most – if not all – of the listeners are situated outside the sweet-area.

The following section discusses spatial aliasing and methods to reduce the perceivable artifacts.

4.2. Does the minimization of spatial aliasing result in better sound localization?

Many psychoacoustic evaluations of Ambisonics systems can be found in literature, *e.g.* for varying numbers of active loudspeakers [17, 18, 5, 6, 19] or fixed loudspeaker setups but different reproduction orders [1]. Some of them indicate that if the number of loudspeakers is higher than the one required by the reproduction order, audible artifacts occur in the vicinity of the central listening area. This further results in

- the impression of close-to-head sources and audible phasing artifacts (Frank [1], Malham [2])
- comb-filter effects near the sweet-area (Nettingsmeier [3]), and
- spectral impairment (Solvang [5]), as well as spectral unbalance (Daniel [4])

Outside the sweet-area, the sound field cannot be reproduced without artifacts. Considering the above-mentioned, the authors suggest that outside the sweet-area ambisonic synthesis sounds better with spatial aliasing than without; we propose the expression *friendly aliasing*.

4.3. Must it be clean holophony?

Panning, *i.e.* suitably weighted playback of a signal $s(t)$ on a loudspeaker setup $\mathbf{x}(t) = \mathbf{g} s(t)$ can produce the impression of one single sound source. The perceived location of this phantom source can be adjusted by the weights \mathbf{g} . Phantom sources appearing between pairs (stereophony, surround) and triplets of loudspeakers (VBAP) are well-described in literature [20, 21, 22]. It might be convenient to violate the requirements of holophonic wave-field reconstruction with Ambisonics in order to obtain a panning law that enforces stable phantom sources.

Santala [6] states that room reverberation is able to mask the phasing effect. Spatial aliasing is further reported to be capable of masking the side-effects of ideal order truncation, which normally occur at the border of the sweet-area.

Frank [1] has shown for the IEM-CUBE that digital time-alignment of loudspeaker signals deteriorates the perceived quality of the sound reproduction for low order decoding near the sweet-spot. In-phase decoding further decreases the effective order [4] and performs inferior to max-rE or basic decoding using time-aligned signals.

Compressive sampling [23] is a non-mode-matching optimization technique for computing the decoder for given loudspeaker setups. It makes use of the L1-norm optimization, which

yields driving signals for the smallest possible set of active loudspeakers. Nevertheless, it seems to be a promising technique delivering good perceptual results.

For simplification, the influence of room acoustic conditions and loudspeaker radiation patterns are usually neglected in most practical implementations; most recently this has been studied for WFS systems [24] and Ambisonics [25].

Ambisonics for binaural sound reproduction over headphones (Daniel [4], Noisternig *et al.* [26], Kan [27], etc.) allows to perfectly center the listener's head. Therefore, it provides a more accurate implementation of holophony than most loudspeaker setups.

3D spatialization with a low number of loudspeakers and a moderate acoustical treatment of the room raises the question: *Must it be clean holophony, or does the interpretation as a panning-law serve our needs?* In [28, 29, 30] Ambisonics is regarded as a panning function; this approach can be further generalized by applying weighting functions for spherical convolution, which is shown in the following section.

4.4. Erroneous localization near the loudspeakers

A virtual point-source in Ambisonics corresponds to an ideally truncated spherical Dirac-delta distribution. For circular horizontal Ambisonics systems this function corresponds to a periodic sinc-function, and for spherical Ambisonics systems it corresponds to a rotationally symmetric, sinc-like function. This spatially/angularly band-limited distribution results in distracting side lobes, which often evoke localization errors and front-back confusion for the listeners outside the sweet-area, see also Fig. 3a.

To diminish these distracting side lobes angular smoothing functions have been applied to higher-order Ambisonics [4], cp. Fig. 3b. In the spherical-harmonics domain, angular smoothing is accomplished by attenuating the higher order components, *i.e.* *spherical convolution* [31].

A source at the north pole $\theta_s = (0, 0, 1)^T$ has a normalized, rotationally symmetric, angular amplitude-distribution

$$f_s(\vartheta) = (N+1)^{-2} \sum_{n=0}^N (2n+1) P_n(\cos(\vartheta)) \quad (25)$$

that depends on the zenith angle $\vartheta = 0 \dots \pi$ and the Legendre polynomials $P_n(\cos(\vartheta))$. Applying the order-weights a_n , this distribution becomes

$$f_s(\vartheta) = \left[\sum_{n'=0}^N a_{n'} (2n'+1) \right]^{-1} \sum_{n=0}^N a_n (2n+1) P_n(\cos(\vartheta)). \quad (26)$$

Known types of weighting functions a_n in Ambisonics literature are

- *In-phase*: full side lobe suppression, biggest main lobe,
- *max-rE*: best angular power concentration, and
- *max-rV*, narrowest source in the amplitude domain,

and frequency-dependent combinations thereof. These smoothing functions can be found, *e.g.*, in Jérôme Daniel's thesis [4]. However, these smoothing-functions do not provide a freely adjustable the side-lobe rejection.

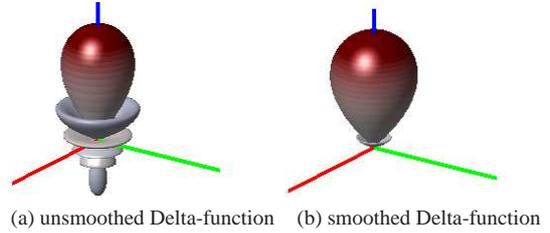


Figure 3: Angular amplitude-distribution for a virtual point-source of the order $N = 5$ at $\theta_s = (0, 0, 1)^T$ plotted as balloon diagram $|f_s(\vartheta)|$, without and with smoothing.

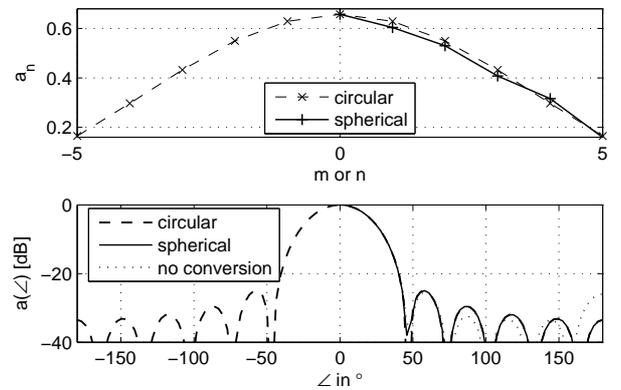


Figure 4: Conversion of a Kaiser-window to the spherical harmonics (Legendre) domain is necessary to obtain the corresponding circular and spherical angular functions.

5. PROPOSED IMPROVEMENTS

5.1. Spherical Kaiser smoothing-function

The Kaiser-window provides a parametrization of the side-lobe attenuation. In the following, a Kaiser-window with the parameter $\beta = 2.75$ is employed, resulting in a side-lobe attenuation of 23dB . Given the original window coefficients in a vector \tilde{a} , a matrix \mathbf{W} achieves conversion from the circular to the spherical harmonics domain [32],

$$\mathbf{a} = \mathbf{W} \tilde{\mathbf{a}}. \quad (27)$$

First-order Ambisonics decoders often apply frequency-dependent smoothing functions. Referring to literature [1, 18], frequency-dependent smoothing have not been found to improve higher-order Ambisonics (HOA) decoding.

6. QUALITY MEASURES FOR AMBISONICS

The evaluation of Ambisonics decoders requires objective quality measures, which have been chosen to be optimal for a

- constant amplitude in the direction of the virtual source (band-limited case, condition met by mode-matching),
- constant decoded overall-power for any virtual source direction,

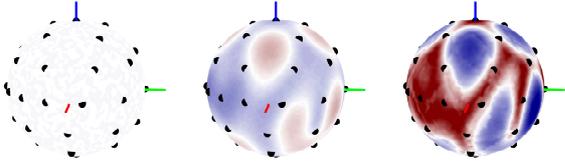


Figure 5: Quality measures for mode-matching on the entire sphere.

- constant concentration of the decoded power towards loudspeakers near the virtual source.

All three measures express the stability of loudness for a virtual source with varying direction. The first quality measure estimates the loudness inside, the second one the loudness outside the sweet-area, and the third measure roughly describes the stability of the discrete angular loudspeaker driving signals around the virtual source direction.

6.1. Peak-amplitude of the virtual source (within the sweet-area)

An ideal mode-matching decoder produces the identity $\mathbf{Y}_N \mathbf{D} = \mathbf{I}$, cp. Eq. (24). Therefore, decoding/re-encoding of an Ambisonics setup should ideally reproduce the analytical spherical-harmonic patterns when regarding a band-limited subspace, *i.e.* $n \leq N$. Within this subspace and considering the order-weights \mathbf{a} , cf. Eq. (27), a virtual source at the direction $\boldsymbol{\theta}_s$ should be represented by the ideal angular amplitude distribution $f(\boldsymbol{\theta}) = \mathbf{y}_N^T(\boldsymbol{\theta}) \text{diag}\{\mathbf{a}\} \mathbf{y}_N(\boldsymbol{\theta}_s)$, which has a constant peak-amplitude

$$q_1^{(ideal)} = \sum_{n=0}^N \frac{2n+1}{4\pi} a_n. \quad (28)$$

In practice, accurate mode-matching will often fail so that the peak-amplitude differs from the one given in Eq. (28). Therefore, the ratio between the actual peak-value for every virtual source angle $\boldsymbol{\theta}_s$ and its ideal value $q_1^{(ideal)}$ is considered as a quality criterion. Given the decoder \mathbf{D} and the order weights \mathbf{a} , this first measure writes as

$$q_1(\boldsymbol{\theta}_s) = \mathbf{y}_N^T(\boldsymbol{\theta}_s) \mathbf{Y}_N \mathbf{D} \text{diag}\{\mathbf{a}\} \mathbf{y}_N(\boldsymbol{\theta}_s) / q_1^{(ideal)}. \quad (29)$$

It varies locally depending on the decoders found by regularization, approximation, or for decoders without mode-matching. Orthogonal angular loudspeaker layouts are a special case: *t-design*, or *quadrature* using the decoder $\mathbf{D} = \text{diag}\{\mathbf{w}\} \mathbf{Y}_N^T$ always perform optimally [10].

6.2. Overall power of decoded signals (outside the sweet-area)

Decoding a virtual source usually does not – but ideally should – yield a constant sum of squared loudspeaker signals, independent of the source position $\boldsymbol{\theta}_s$

$$q_2^{(ideal)} = \sum_{n=0}^N \frac{2n+1}{4\pi} a_n^2. \quad (30)$$

For mode-matching decoders, the actual sum tends to deviate from the ideal values. Given the decoder matrix \mathbf{D} and the order weights \mathbf{a} , the second quality measure becomes

$$q_2(\boldsymbol{\theta}_s) = \|\mathbf{D} \text{diag}\{\mathbf{a}\} \mathbf{y}_N(\boldsymbol{\theta}_s)\|^2 / q_2^{(ideal)}. \quad (31)$$

Orthogonal angular layouts of loudspeakers are again an ideal exception, cf. *t-design*, or *quadrature* with decoder $\mathbf{D} = \text{diag}\{\mathbf{w}\}^{\frac{1}{2}} \mathbf{Y}_N^T$, [10].

6.3. Angular power-distribution of decoded signal

In order to prevent largely erroneous localization, the power of the discrete signals should concentrate within the angular proximity to a virtual source. Introducing the angle ϑ_p that discriminates between proximal and distant angles, we obtain a constant ratio in the ideal analytic case

$$q_3^{(ideal)} = \frac{\int_{\cos(\vartheta_p)}^1 |\sum_{n=0}^N a_n (2n+1) P_n(\mu)|^2 d\mu}{\int_{-1}^{\cos(\vartheta_p)} |\sum_{n=0}^N a_n (2n+1) P_n(\mu)|^2 d\mu}. \quad (32)$$

For the given angular function in Fig. 4 and $\vartheta_p = 40^\circ$ we get $q_3^{(ideal)} \approx -17\text{dB}$.

Given the decoder matrix \mathbf{D} and the order weights \mathbf{a} , the ratio between the power of the loudspeaker signals in angular proximity to the virtual source and of those far from it yields

$$q_3(\boldsymbol{\theta}_s) = \frac{\|\text{diag}\{\boldsymbol{\theta}_s^T \boldsymbol{\Theta}_L \geq \cos(\vartheta_p)\} \mathbf{D} \text{diag}\{\mathbf{a}\} \mathbf{y}_N(\boldsymbol{\theta}_s)\|^2}{\|\text{diag}\{\boldsymbol{\theta}_s^T \boldsymbol{\Theta}_L < \cos(\vartheta_p)\} \mathbf{D} \text{diag}\{\mathbf{a}\} \mathbf{y}_N(\boldsymbol{\theta}_s)\|^2}, \quad (33)$$

with $\boldsymbol{\Theta}_L = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_L)$; the inner product $\boldsymbol{\theta}_s^T \boldsymbol{\Theta}_L$ expresses the cosine of the angular distance between the loudspeakers and the virtual source.

The angle ϑ_p has been chosen to enclose 3-5 loudspeakers around any virtual source location.

6.4. Example

Fig. 5 shows three different quality measures in dB for a mode-matching decoder on the entire sphere, cf. Eq. (24). The amplitude levels are constant but the other measures do not perform optimally.

7. HEMISPHERICAL AMBISONICS DECODING

Why should decoding to a hemispherical loudspeaker arrangements be challenging, e.g. regarding the IEM-CUBE, a small concert room depicted in Fig. 6? The following sections discuss the difficulties of decoding by mode-matching and present various approaches to improve the results for hemispherical loudspeaker arrangements. Moreover, a non-mode-matching approach is given.

7.1. Mode-matching based on hemispherical base-functions

As described in [8, 13], a new orthogonal set of functions can be calculated for the hemisphere by truncated eigendecomposition. For the spherical domain restricted to zenith angles $\vartheta = 0 \dots \vartheta_{\max} < \pi$ with $\vartheta_{\max} = 104^\circ$ we obtain orthonormal functions as depicted in Fig. 2.



Figure 6: The IEM-CUBE is a 24 channel hemispherical Ambisonics system.

Encoding the loudspeakers with this new subset of band-limited spherical harmonics according to Eq. (16) yield

$$\tilde{\mathbf{Y}}_N = (\mathbf{R}^\dagger)^\top \mathbf{Y}_N. \quad (34)$$

The matching-condition using these new base-functions is similar to Eq. (23)

$$\tilde{\mathbf{Y}}_N \mathbf{g} \stackrel{!}{=} \tilde{\phi}_{s,N}$$

and yields

$$\begin{aligned} \mathbf{g} &= \tilde{\mathbf{Y}}_N^\top \tilde{\mathbf{G}}_d^{-1} \tilde{\phi}_{s,N}, \\ \text{with } \tilde{\mathbf{G}}_d &= \tilde{\mathbf{Y}}_N \tilde{\mathbf{Y}}_N^\top. \end{aligned} \quad (35)$$

The resulting decoding equation with ordinary spherical harmonics encoding from Eq. (16) is

$$\begin{aligned} \mathbf{g} &= \mathbf{D} \phi_{s,N}, \\ \text{with } \mathbf{D} &= \tilde{\mathbf{Y}}_N^\top \tilde{\mathbf{G}}_d^{-1} (\mathbf{R}^\dagger)^\top. \end{aligned} \quad (36)$$

7.2. Mode-matching based on even spherical harmonics

Another variant of hemispherical decoding [8] excludes the odd spherical harmonics with respect to z ; Eq. (23) becomes invertible. The matrix \mathbf{R} contains ones and zeros in order to extract the $(N+1)(N+2)/2$ even-symmetric harmonics from the entire $(N+1)^2$ set of base functions, see Fig. 7. With $\mathbf{R} = (\mathbf{R}^\dagger)^\top$, the encoded loudspeaker positions become

$$\tilde{\mathbf{Y}}_N = \mathbf{R} \mathbf{Y}_N$$

resulting in the Gram-matrix $\tilde{\mathbf{G}}_d = \tilde{\mathbf{Y}}_N \tilde{\mathbf{Y}}_N^\top$ of the discrete system $\tilde{\mathbf{Y}}_N$. For the hemisphere, this achieves $\tilde{\mathbf{G}} = \mathbf{R}^\top \mathbf{G} \mathbf{R} = \mathbf{I}$ by discarding the obvious linear dependencies.

With ordinary spherical harmonics encoding, the above can be re-written according to [8], as a special case of Eq. (16)

$$\begin{aligned} \mathbf{g} &= \mathbf{D} \phi_{s,N}, \\ \text{with } \mathbf{D} &= \tilde{\mathbf{Y}}_N^\top \tilde{\mathbf{G}}_d^{-1} \mathbf{R}. \end{aligned}$$

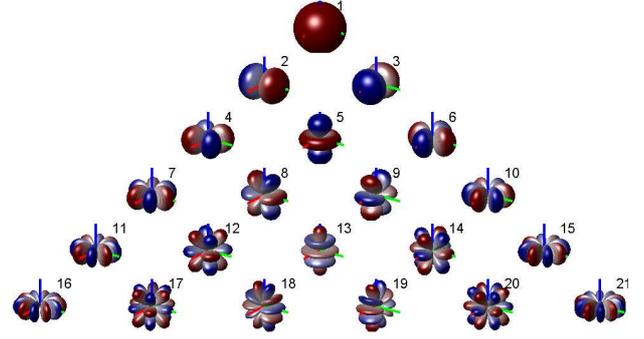


Figure 7: Selection of the 21 even-symmetric spherical harmonics $n \leq 5$ with respect to the z -axis.

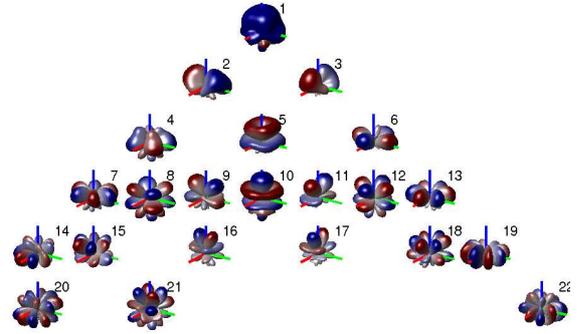


Figure 8: Base-functions that are orthonormal for the hemispherical loudspeaker-layout of the IEM-CUBE.

7.3. Mode-matching regularizing the discrete Gram-matrix

Another feasible way of decoding to a hemispherical loudspeaker layout is regularization of Eq. (24), cp. [28, 33, 30].

As in the analytic case, Sec. 2.4, regularization results in new base-functions that are orthonormal for a given loudspeaker layout. This is achieved by truncated eigendecomposition of the Gram-matrix for the given loudspeaker system \mathbf{Y}_N

$$\begin{aligned} \mathbf{G}_d &= \mathbf{Y}_N \mathbf{Y}_N^\top = \mathbf{V} \text{diag}\{\boldsymbol{\lambda}\} \mathbf{V}^\top, \\ \mathbf{R} &= \text{diag}\{\boldsymbol{\lambda}_c\}^{-\frac{1}{2}} \mathbf{V}_c^\top, \\ \tilde{\mathbf{Y}}_N &= (\mathbf{R}^\dagger)^\top \mathbf{Y}_N. \end{aligned}$$

The new Gram-matrix is orthonormal, $\tilde{\mathbf{G}}_d = \tilde{\mathbf{Y}}_N \tilde{\mathbf{Y}}_N^\top = \mathbf{I}$ and therefore simplifies the right-inverse as given above. With Eq. (16), the corresponding mode-matching decoder becomes

$$\begin{aligned} \mathbf{g} &= \mathbf{D} \phi_{s,N}, \\ \text{with } \mathbf{D} &= \tilde{\mathbf{Y}}_N^\top (\mathbf{R}^\dagger)^\top. \end{aligned}$$

Multipole-matched rendering [33] is similar but additionally takes the regularization of frequency-dependent acoustic near-fields into account.

7.4. Mode-matching based on virtual and phantom loudspeakers – the Musil decoder-design

Alternatively to designing new restricted-domain orthonormal or regularized base-functions in the modal domain, modification can be done in the angular domain as well. For this purpose, a given hemispherical layout is completed to a full sphere by the virtual loudspeakers $l > L$, also referred to as phantom loudspeakers. The decoder is computed for this full set of loudspeakers

$$\mathbf{Y}_N = \begin{pmatrix} Y_0^0(\theta_1) & \dots & Y_0^0(\theta_L) & Y_0^0(\theta_{L+1}) & \dots \\ \vdots & \ddots & \vdots & \vdots & \ddots \\ Y_N^N(\theta_1) & \dots & Y_N^N(\theta_L) & Y_N^N(\theta_{L+1}) & \dots \end{pmatrix}. \quad (37)$$

The driving signals for the virtual loudspeakers can be (a) omitted, or (b) mapped to their nearest neighboring loudspeakers, in order to preserve the power of their signals. The matching condition uses gains for both physical and phantom loudspeakers

$$\mathbf{Y}_N \begin{pmatrix} g_1 \\ \vdots \\ g_L \\ g_{L+1} \\ \vdots \end{pmatrix} \stackrel{!}{=} \phi_{s,N}. \quad (38)$$

All the gains are calculated by right-inversion $\mathbf{Y}_N^T \mathbf{G}_d^{-1}$ and take into account linear combinations to build phantom loudspeakers

$$\mathbf{g} = \mathbf{D} \phi_{s,N}, \quad (39)$$

$$\text{with } \mathbf{D} = \begin{pmatrix} 1 & 0 & \dots & 0 & g_{L+1,1} & \dots \\ 0 & 1 & 0 & \vdots & g_{L+1,2} & \dots \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots \\ 0 & \dots & \ddots & 1 & g_{L+1,L} & \dots \end{pmatrix} \mathbf{Y}_N^T \mathbf{G}_d^{-1}. \quad (40)$$

with $\mathbf{G}_d = \mathbf{Y}_N \mathbf{Y}_N^T$.

This decoder is currently implemented at the IEM-CUBE, see Fig. 9. Virtual loudspeakers are placed in weakly sampled regions, *i.e.* the “missing speaker” at the north-pole, and the southern hemisphere. Anywhere virtual loudspeakers are located near the real loudspeakers, their signal is mapped to the latter by suitable weighting functions (projection). This decoder-design approach allows to obtain perceptually good results but requires a lot of experience.

7.5. Direct sampling of the spherical harmonics using Voronoi weights

For fully spherical layouts, the Gram-matrix of \mathbf{Y}_N can be improved by multiplying suitable weights w_l to the L loudspeaker nodes. This changes the Gram-matrix $\tilde{\mathbf{G}}_d$ to:

$$\tilde{\mathbf{G}}_d = \mathbf{Y}_N \text{diag}\{\mathbf{w}\} \mathbf{Y}_N^T. \quad (41)$$

In general, weights \mathbf{w} that diagonalize $\tilde{\mathbf{G}}_d$ might not exist. However, if the given loudspeaker nodes and weights w_l form a *quadrature rule*, the Gram-matrix equals identity $\tilde{\mathbf{G}}_d = \mathbf{I}$; this only holds for special sets of nodes and weights, cf. [7].

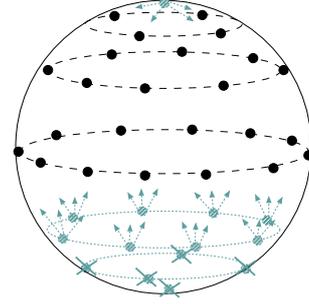


Figure 9: Hemispherical Ambisonics arrangement with the decoding approach of Thomas Musil, using virtual phantom loudspeakers.

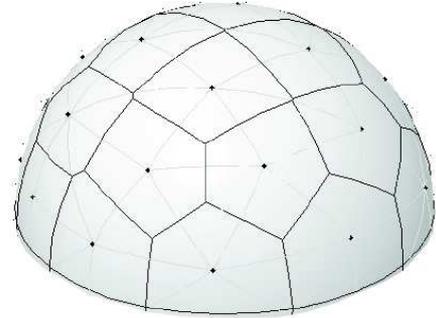


Figure 10: IEM-CUBE loudspeaker layout with Voronoi cells determining the coverage area of each loudspeaker.

Nevertheless, $\tilde{\mathbf{G}}_d \approx \mathbf{I}$ can be approximated by weighting the l^{th} loudspeaker signal with the discrete surface w_l it covers. The Voronoi-algorithm STRIPACK [34, 35] for the sphere calculates these weights for a given arrangements of loudspeakers. For distributions on a full sphere, the weights obtained by this algorithm deliver results of reasonable playback quality with the decoder

$$\mathbf{D} = \text{diag}\{\mathbf{w}\} \mathbf{Y}_N^T. \quad (42)$$

This method provides a relatively easy way to determine the decoder coefficients for incomplete spherical layouts. However, cells without enough neighbors have to be manually limited, *e.g.* for hemispherical layout the loudspeakers on the equator, see Fig. 10.

8. CASE STUDY

Seven decoder examples have been evaluated at the IEM-CUBE using the loudspeaker-locations provided [36]. For all the examples, spherical Kaiser-smoothing has been applied with the parameter $\beta = 2.75$. In particular, these decoders have been studied:

- even symmetrical mode-matching,
- hemispherical-base mode-matching, regularization parameter $c = 1.4$, and $\vartheta_{\max} = 104^\circ$,

- regularized Gram-matrix decoder, regularization parameter $c = 2.7$,
- Voronoi-weighted,
- Musil-decoder (phantom loudspeakers),

The objective quality measures evaluated for the test cases are shown in Fig. 11. The figures illustrate the upper hemisphere, using the azimuth φ as polar angle and the zenith θ as radial coordinate. Black dots indicate the loudspeaker positions.

Obviously the mode-matching decoders in Figs. 11a, 11b perform quite well for the q_1 criterion, except for the “hole” at the north-pole.

8.1. Perceptual evaluation – informal listening sessions

The above-mentioned objective quality measures might provoke misinterpretations regarding the perceptual quality. This section briefly discusses informal listening sessions in the IEM-CUBE. One should note that this must not be understood as evidence as the presented impressions are not a result of formally objective psychoacoustic listening experiments.

Informal listening sessions including the authors and Thomas Musil indicate that the q_2 criterion is perceptually much more important than q_1 . It well represents audible loudness variations, *i.e.* the decoders in Figs. 11a and 11b were perceived to have a bad loudness balance, especially because of the loudness-boost of virtual sources positioned near the north-pole.

The bad performance of the q_3 criterion for mode-matching decoders could not be verified in the informal listening sessions.

The subjects of informal listening sessions attributed a good performance to the decoders depicted in Figs. 11c, 11d, 11e. The decoder shown in Fig. 11c was in favor by all of the subjects, even if a loudness loss was noticeable for virtual sources close to the equator. Decoders in Figs. 11c, 11d, 11e have been subjectively rated on the second, third, and fourth place, respectively.

9. CONCLUSIONS

This article discusses the theoretical background of Ambisonics on the sphere, the hemisphere, and in more general the mode-matching approach. Three objective quality criteria were presented that are applicable for evaluation and comparison of different decoders. We have proposed an improvement of Ambisonics decoding by introducing the spherical Kaiser smoothing filter. Most notably, we have described five different methods/techniques for Ambisonics decoding to a hemispherical arrangement of loudspeakers and evaluated their performance objectively, also providing some informal information about their subjective performance. The methods and techniques presented show different ways to solve the decoding problem on the hemisphere or other partial spheres.

Many further aspects of Ambisonics decoding remain subject to future research, as for instance, the geometric distortion and/or optimization of loudspeaker positions on the sphere. Investigation of additional constraints for improving the loudness balance of decoders seems promising as well.

Moreover, hybrid approaches mapping the supplementary virtual loudspeakers (Musil decoder) by applying VBAP should be regarded more closely. Extensive listening test are required to evaluate different decoding strategies and to improve the objective models of subjective qualities.

10. ACKNOWLEDGMENT

The authors are grateful to Thomas Musil for his help with decoder design, implementation, and for attending the listening sessions. We would like to thank IEM for supporting this research as well as IRCAM and the partial support of the EACEA (Education Audiovisual and Culture Executive Agency), Projet Culture 2007-2013 of the European Commission for providing the infrastructure for a successful demonstration at the symposium.

11. REFERENCES

- [1] M. Frank and F. Zotter, “Localization experiments using different 2D Ambisonics decoders,” in *25. Tonmeister-tagung, Leipzig*, 2008.
- [2] D. Malham, “Re: [Sursound] November 2008 paper on 2D decoders (Chris Travis),” Sursound – Surround Sound discussion group, 2009. [Online]. Available: <http://mail.music.vt.edu/mailman/listinfo/sursound>
- [3] J. Nettingsmeier, “[Sursound] more anecdotal evidence for and against delay compensation,” Sursound – Surround Sound discussion group, 2009. [Online]. Available: <http://mail.music.vt.edu/mailman/listinfo/sursound>
- [4] J. Daniel, “Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia,” PhD Thesis, Université Paris 6, 2001.
- [5] A. Solvang, “Spectral impairment of two-dimensional higher order Ambisonics,” *AES Journal*, 2008.
- [6] O. Santala, H. Vertanen, J. Pekonen, J. Oksanen, and V. Pulkki, “Effect of listening room on audio quality in Ambisonics reproduction,” in *126th AES Convention, Munich*, 2009.
- [7] F. Zotter, H. Pomberger, and M. Frank, “An alternative Ambisonics formulation: Modal source-strength matching and the effect of spatial aliasing,” in *126th AES Convention, Munich*, 2009.
- [8] H. Pomberger, F. Zotter, and A. Sontacchi, “An ambisonic format for flexible playback layouts,” in *Proc. 1st Ambisonics Symposium, Graz*, 2009.
- [9] Y. J. Wu and T. D. Abhayapala, “Soundfield reproduction using theoretical continuous loudspeaker,” in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008.
- [10] F. Zotter, “Sampling strategies for holography/holophony on the sphere,” in *NAG-DAGA*, Rotterdam, 2009.
- [11] J. Daniel, “Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format,” in *Proceedings of the AES 23rd International Conference, Copenhagen*, 23rd May 2003.
- [12] S. Moreau, “Étude et réalisation d’outils avancés d’encodage spatial pour la technique de spatialisation sonore Higher Order Ambisonics: microphone 3d et contrôle de distance,” PhD Thesis, Université du Maine, 2006.
- [13] R. Pail, G. Plank, and W.-D. Schuh, “Spatially restricted data distributions on the sphere: the method of orthonormalized functions and applications,” *Journal of Geodesy*, vol. 75, pp. 44–56, 2001.

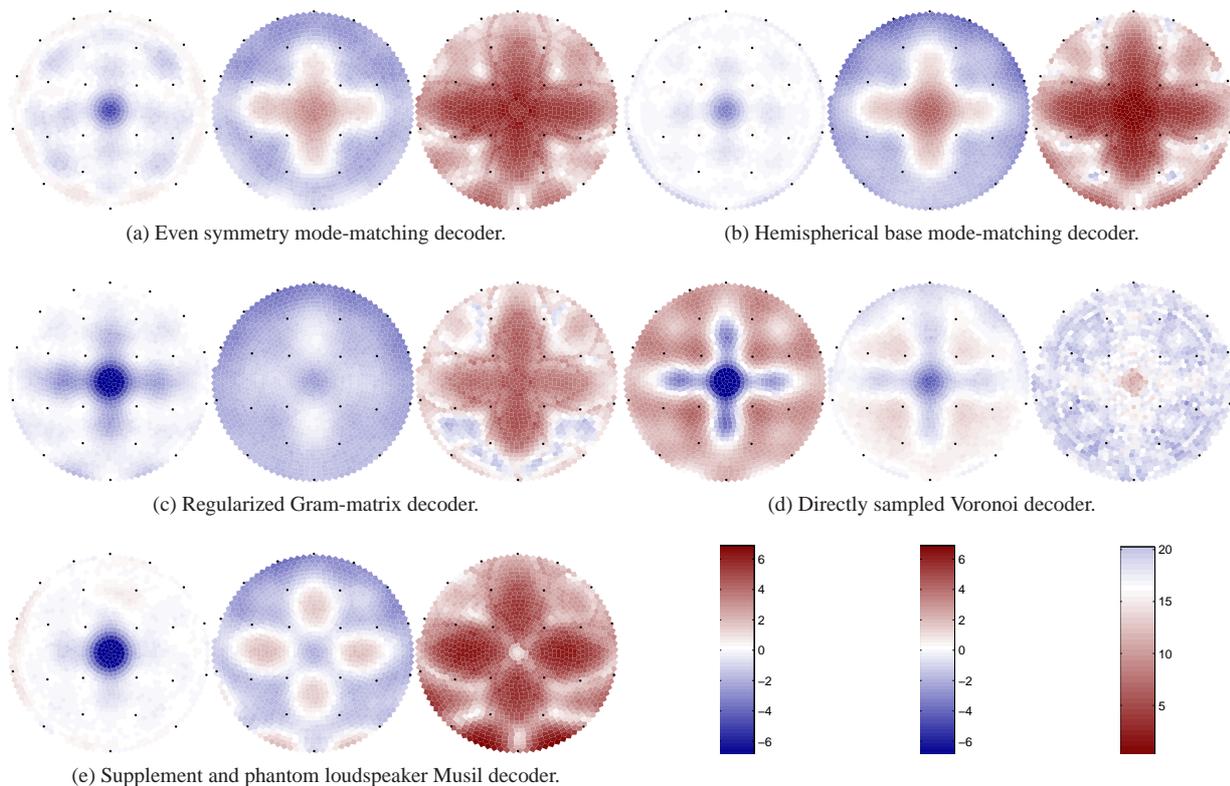


Figure 11: (a) even, (b) hemi-base, (c) regularized, (d) Voronoi, (e) Musil. Kaiser-parameter 2.75, $N=5$, $\text{cond} \approx 1.4-1.5$ for even, hemibase, regularization. The three color bars refer to the three columns of each diagram: colors first (virtual source amplitude) and second column (total loudspeaker signal power): $-7 \dots 0 \dots 7$ dB corresponds to blue \dots white \dots red, third column: red \dots white \dots corresponds to $0 \dots 17$ dB (power concentration within $\pm 40^\circ$).

- [14] J. Ahrens and S. Spors, "An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions," *Acta Acustica united with Acustica*, 2008.
- [15] D. Ward and T. Abhayapala, "Reproduction of a plane-wave sound field using an array of loudspeakers," *IEEE Transactions on Speech and Audio Processing*, 2001.
- [16] M. A. Poletti, "Three-dimensional surround sound systems based on spherical harmonics," *AES Journal*, 2005.
- [17] M. Kratschmer and R. Rabenstein, "Implementing Ambisonics on a 48 channel circular loudspeaker array," in *Proc. 1st Ambisonics Symposium, Graz*, 2009.
- [18] S. Bertet, J. Daniel, E. Parizet, and O. Warusfel, "Investigation on the restitution system influence over perceived higher order Ambisonics sound field: a subjective evaluation involving from first to fourth order systems," in *Proc. Acoustics-08, Joint ASA/EAA meeting, Paris*, 2008.
- [19] M. Frank, "Perzeptiver Vergleich von Schallfeldreproduktionsverfahren unterschiedlicher räumlicher Bandbreite," University of Music and Performing Arts, Institute of Electronic Music and Acoustics, Graz, and Deutsche Telekom Labs, Berlin, 2009.
- [20] J. Blauert, *Spatial Hearing*. MIT Press, 1983.
- [21] G. Theile, "Über die Lokalisation im überlagerten Schallfeld," Ph.D. dissertation, Technische Universität Berlin, 1980.
- [22] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *Journal of the Audio Engineering Society*, vol. 45, no. 6, pp. 456–466, 1997.
- [23] N. Epain, C. Jin, and A. van Schaik, "The application of compressive sampling to the analysis and synthesis of spatial sound fields," in *127th AES Convention, New York*, 2009.
- [24] S. Spors, H. Buchner, and R. Rabenstein, "A novel approach to active listening room compensation for wave field synthesis using wave-domain adaptive filtering," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. Citeseer, 2004.
- [25] J. Ahrens and S. Spors, "Sound field reproduction employing non-omnidirectional loudspeakers," in *126th Conv. of the AES, Munich*, 2009.
- [26] M. Noisternig, A. Sontacchi, T. Musil, and R. Höldrich, "A 3D ambisonic based binaural sound reproduction system," in *24th international AES Conference: Multichannel Audio, Banff*, 2003.
- [27] A. Kan, C. Jin, and A. van Schaik, "Distance variation function for simulation of near-field virtual auditory

- space,” in *2006 IEEE International Conference on Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings*, vol. 5, 2006.
- [28] M. Poletti, “Robust two-dimensional surround sound reproduction for nonuniform loudspeaker layouts,” *Journal of the Audio Engineering Society*, vol. 55, no. 7/8, p. 598, 2007.
- [29] M. Neukom and J. Schacher, “Ambisonics equivalent panning,” in *Proceedings of the 2008 International Computer Music Conference, Belfast, UK*, 2008.
- [30] J.-M. Batke and F. Keiler, “Investigation of robust panning functions for 3d loudspeaker setups,” in *128th AES Convention, London*, 2010.
- [31] J. R. Driscoll and D. M. J. Healy, “Computing fourier transforms and convolutions on the 2-sphere,” *Advances in Applied Mathematics*, vol. 15, 1994.
- [32] F. Zotter, “Analysis and Synthesis of Sound-Radiation with Spherical Arrays,” PhD thesis, University of Music and Performing Arts, Graz, 2009.
- [33] J. Hannemann, “Broadband Multipole-Matched Rendering: The Current State of Affairs,” in *1st Ambisonics Symposium, Graz*, 2009.
- [34] J. Keiner, “CSTRIPACK. C translation of Voronoi algorithm on the sphere,” 2007. [Online]. Available: <http://www.math.uni-luebeck.de/keiner/sonstiges.shtml>
- [35] R. Renka, “Algorithm 772: STRIPACK: Delaunay triangulation and Voronoi diagram on the surface of a sphere,” *ACM Transactions on Mathematical Software (TOMS)*, vol. 23, no. 3, p. 434, 1997.
- [36] K. Macheiner, “Bestimmung lokaler Koordinaten eines Lautsprecherfeldes im Akustiklabor Inffeldgasse 10, Institut für Elektrische Musik und Akustik (IEM),” 2008. [Online]. Available: <http://ambisonics.iem.at/symposium2009/audio-infrastructure>