Warping and Directional Loudness Manipulation Tools for Ambisonics

Matthias Kronlachner and Franz Zotter

The production and re-production of sound scenes in the Ambisonic domain offers flexibility regarding the loudspeaker placement around the listening area. Correct decoding should result in a spatial audio perspective that is independent of the loudspeaker configuration.
In case a modification of this perspective is needed, or directional alterations of the amplitude, applying such transformations in the Ambisonic domain is more challenging than directly changing the production or loudspeaker placement.
Nevertheless, for achieving flexibility during post-production and playback, finding such algorithms in the Ambisonic domain is feasible. This poster shows a simple way to describe any Ambisonic transformation by performing all manipulations in the angular domain instead of the spherical harmonics domain. For a practical implementation, the Ambisonic signals are sampled at sufficiently many discrete points in the angular domain, where the location of the sampling points or the weighting is manipulated. Re-expansion of the manipulated angular samples back into the spherical harmonics domain yields the Ambisonic transformation matrix.
The manipulations are ready to use and included in the *ambix* Ambisonic plug-in suite.

Describing any manipulation

A surround audio signal $f(\boldsymbol{\theta},t)$ from the direction

$$\boldsymbol{\theta} = \begin{pmatrix}\theta_x\\ \theta_y\\ \theta_z\end{pmatrix} = \begin{pmatrix}\cos\varphi\cos\vartheta\\ \sin\varphi\cos\vartheta\\ \sin\vartheta\end{pmatrix} \quad (1)$$

can be represented as

$$f(\boldsymbol{\theta},t) = \boldsymbol{y}^{\mathrm{T}}(\boldsymbol{\theta})\,\boldsymbol{\phi}(t). \quad (2)$$

with $\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{\theta})$ being the spherical harmonics and $\boldsymbol{\phi}(t)$ their expansion coefficients.
We want to find the matrix $\boldsymbol{T}$ which yields the transformed Ambisonic signals

$$\tilde{\boldsymbol{\phi}}(t) = \boldsymbol{T}\,\boldsymbol{\phi}(t). \quad (3)$$

Desirable transformations of the surround signal are (1) weighting by a direction-dependent gain $g(\boldsymbol{\theta})$ and (2) angular transformations $\mathcal{T}\{\boldsymbol{\theta}\}$ to modify the panorama

$$\tilde{f}(\boldsymbol{\theta},t) = g(\boldsymbol{\theta})\,f(\mathcal{T}^{-1}\{\boldsymbol{\theta}\},t). \quad (4)$$

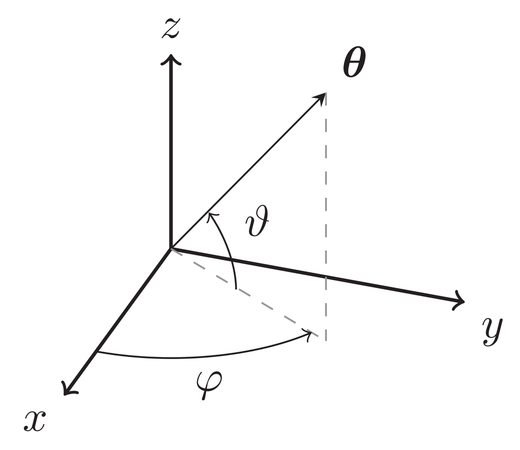


Figure 1: Cartesian and spherical coordinate system.

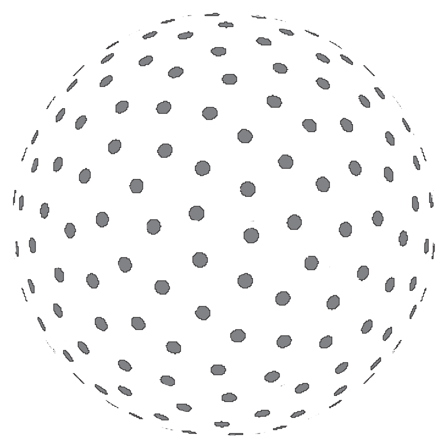


Figure 2: 240 sampling points 21-design allows order N=10.

Using (2) and (4) yields

$$\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{\theta})\,\tilde{\boldsymbol{\phi}}(t) = g(\boldsymbol{\theta})\,\boldsymbol{y}^{\mathrm{T}}(\mathcal{T}^{-1}\{\boldsymbol{\theta}\})\,\boldsymbol{\phi}(t) \quad (5)$$

and we use orthogonality $\int_{\mathbb{S}^2}\boldsymbol{y}(\boldsymbol{\theta})\,\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} = \mathrm{diag}\{\boldsymbol{a}\}$ to remove $\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{\theta})$

$$\tilde{\boldsymbol{\phi}}(t) = \underbrace{\mathrm{diag}\{\boldsymbol{a}\}^{-1}\int_{\mathbb{S}^2}\boldsymbol{y}(\boldsymbol{\theta})\,g(\boldsymbol{\theta})\,\boldsymbol{y}^{\mathrm{T}}(\mathcal{T}^{-1}\{\boldsymbol{\theta}\})\,\mathrm{d}\boldsymbol{\theta}}_{:=\boldsymbol{T}}\;\boldsymbol{\phi}(t), \quad (6)$$

$$\text{with } \boldsymbol{a} = \left[\tfrac{4\pi}{2n_{ACN}+1}\right]_{ACN}.$$

We recognize this integral as spherical harmonics transform $\boldsymbol{T} = \mathcal{SHT}\{g(\boldsymbol{\theta})\,\boldsymbol{y}^{\mathrm{T}}(\mathcal{T}^{-1}\{\boldsymbol{\theta}\})\}$ which we can perform as discreet spherical harmonics transform using a suitable distribution of L directions $\boldsymbol{\Theta} = [\theta_1, \ldots, \theta_{\mathrm{L}}]^T$.

$$\boldsymbol{T} = \mathcal{DSHT}\{\mathrm{diag}\{\boldsymbol{g}(\boldsymbol{\Theta})\}\,\boldsymbol{Y}(\mathcal{T}^{-1}\{\boldsymbol{\Theta}\})\} = \boldsymbol{Y}^{\dagger}(\boldsymbol{\Theta})\,\mathrm{diag}\{\boldsymbol{g}(\boldsymbol{\Theta})\}\,\boldsymbol{Y}(\mathcal{T}^{-1}\{\boldsymbol{\Theta}\}). \quad (7)$$

$\boldsymbol{T}$ is constant as long as the angular transformation $\mathcal{T}\{\boldsymbol{\theta}\}$ and the weighting function $g(\boldsymbol{\theta})$ are not changing. Angle-distortion or directional-loudness-weighting requires the manipulated Ambisonic signal to be of higher order which can be found in [1, Tab 1].

Efficient evaluation of $\boldsymbol{T}$ by t-designs

The number of sampling points must be at least $\mathrm{L} \geq (\mathrm{N}+1)^2$. For a transform of the Ambisonic order N, we would need a spherical t-design of $t \geq 2\mathrm{N}$. Hardin and Sloan [2] provided coordinates $\boldsymbol{\Theta}_t$ for various t-designs. A more recent collection of t-designs up to degree 1000 was provided by Gräf and Potts [3].
The 21-design with 240 points [Fig. 2] allows the use up to Ambisonic order $\mathrm{N}=10$, and it allows a $\mathcal{DSHT}$ without any pseudo-inversion:

$$\boldsymbol{T} = \mathrm{diag}\{\boldsymbol{b}\}\,\boldsymbol{Y}^{\mathrm{T}}(\boldsymbol{\Theta}_t)\,\mathrm{diag}\{\boldsymbol{g}(\boldsymbol{\Theta}_t)\}\,\boldsymbol{Y}(\mathcal{T}^{-1}\{\boldsymbol{\Theta}_t\}), \quad \text{with } \boldsymbol{b} = \left[\tfrac{2n_{ACN}+1}{\mathrm{L}}\right]_{ACN}. \quad (8)$$

Directional loudness modification

Modifying the loudness of specific directions is especially useful for post production of microphone array recordings.
To perform loudness modifications in the angular domain, we consider a cap function to crop out a part of the surround sound scene (Fig. 3). For this purpose, we use Eq.8 with a neutral angular transformation $\mathcal{T}\{\boldsymbol{\theta}\} = \boldsymbol{\theta}$ and a gain function

$$g(\boldsymbol{\theta}) = g_1\,u(\boldsymbol{\theta}_c^{\mathrm{T}}\boldsymbol{\theta} - \cos\tfrac{\gamma_c}{2}) + g_2\,u(\cos\tfrac{\gamma_c}{2} - \boldsymbol{\theta}_c^{\mathrm{T}}\boldsymbol{\theta}). \quad (9)$$

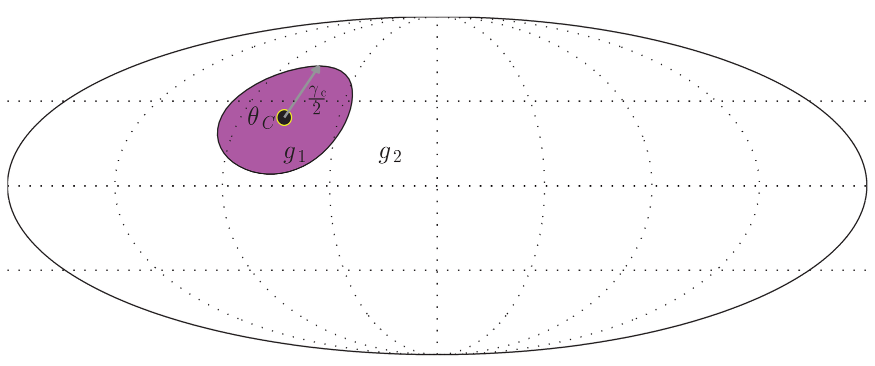


Figure 3: Spherical cap with center $\boldsymbol{\theta}_c$, size $\frac{\gamma_c}{2}$, gain factor g_1 inside the cap and g_2 outside the cap.

Warping the surround image

Warping is used to stretch a certain region of the surround image while squeezing it in other regions to prevent overlap. We are using an angular transformation $\mathcal{T}\{\boldsymbol{\theta}\}$ to warp the elevation angle (Fig. 5).
Warping with regard to other directions are accessible through pre- and post rotation.
As example we outline warping towards the northpole (Fig. 4a, *Gerzon's dominance effect*):

$$\tilde{\mu} = \frac{\mu+\alpha}{1+\alpha\mu} \qquad \text{with} \quad \mu = \sin(\vartheta) \quad (10)$$

An enlarged virtual source of constant amplitude activates more loudspeakers. To prevent the hereby invoked loudness increase, we use the de-emphasis

$$g(\mu) = \frac{\sqrt{1-\alpha^2}}{1+\alpha\mu} \quad (11)$$

to attenuate enlarged regions.

Figure 4: Warping of the elevation

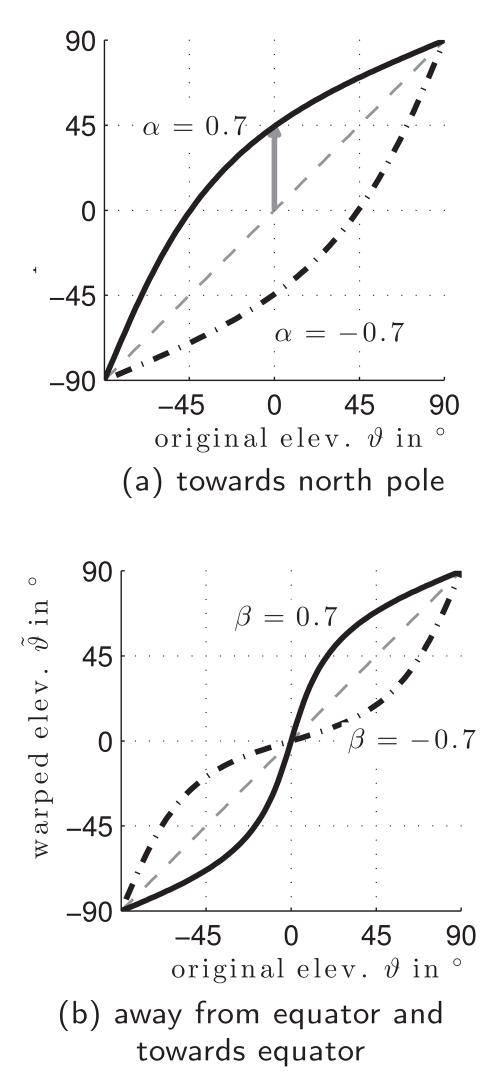
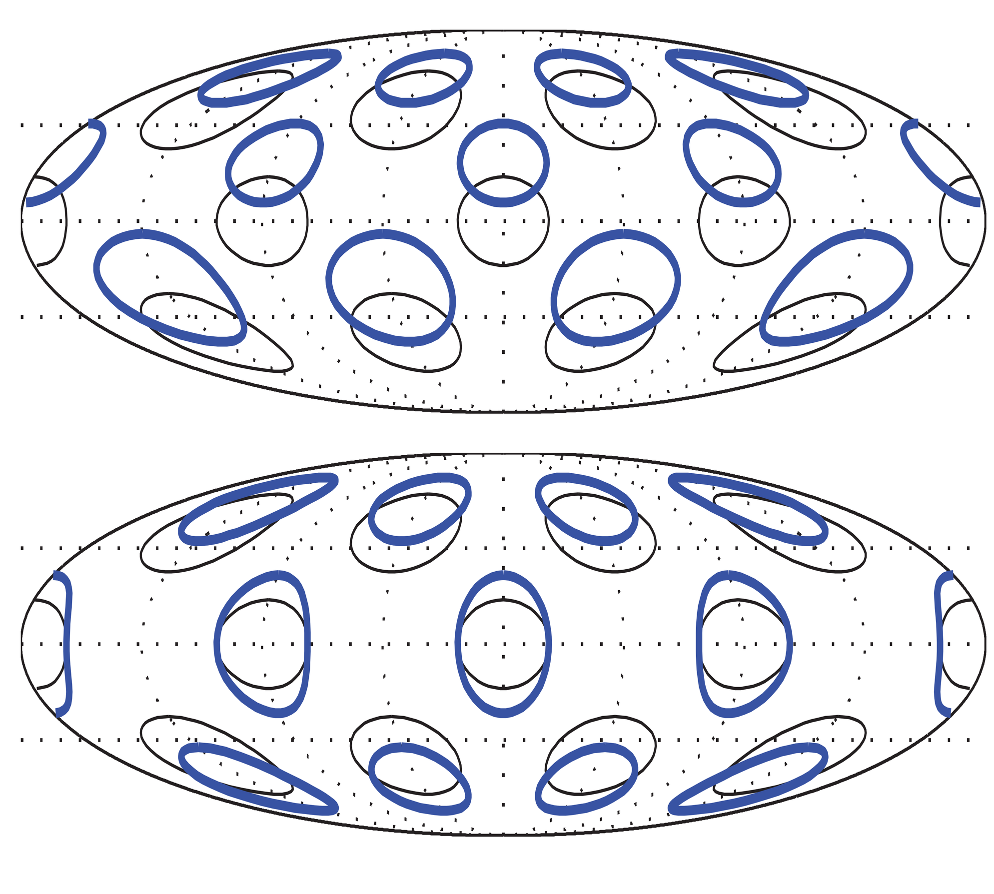


Figure 5: Warping scheme, thin lines indicates unmodified surround image, warping towards the north-pole and warping away from the equator.



Rotation

While rotations of spherical harmonics around the z-axis are fairly easy to implement, the rotation around the x- and y-axis is mathematically more demanding.
To maintain mathematical simplicity we can perform the rotation of the spherical harmonics in the angular domain using Eq. 8 with the angular transformation

$$\mathcal{T}\{\boldsymbol{\theta}\} = \mathbf{R}(\phi,\theta,\psi)\,\boldsymbol{\theta}, \quad (12)$$

where

$$\mathbf{R}(\phi,\theta,\psi) = \underbrace{\begin{pmatrix}1 & 0 & 0\\ 0 & \cos\phi & -\sin\phi\\ 0 & \sin\phi & \cos\phi\end{pmatrix}}_{\mathbf{x-axis-rotation(roll)}} \cdot \underbrace{\begin{pmatrix}\cos\theta & 0 & \sin\theta\\ 0 & 1 & 0\\ -\sin\theta & 0 & \cos\theta\end{pmatrix}}_{\mathbf{y-axis-rotation(pitch)}} \cdot \underbrace{\begin{pmatrix}\cos\psi & -\sin\psi & 0\\ \sin\psi & \cos\psi & 0\\ 0 & 0 & 1\end{pmatrix}}_{\mathbf{z-axis-rotation(yaw)}}. \quad (13)$$

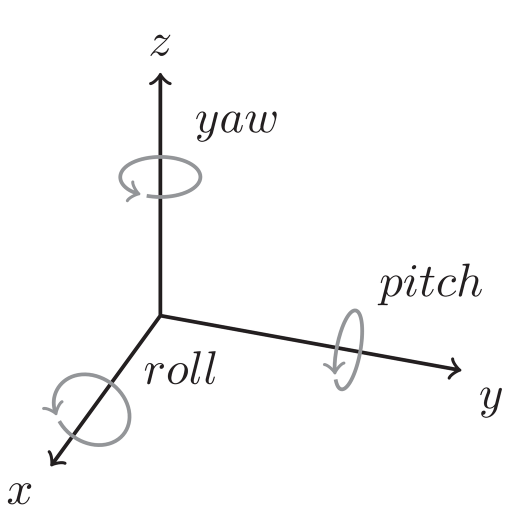
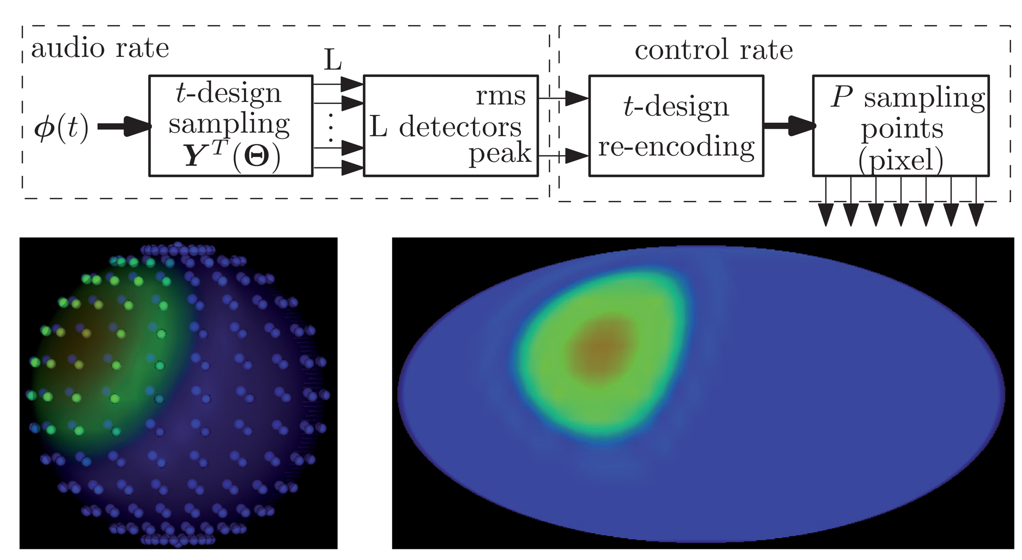


Figure 6: Rotation around xyz-axis

Metering

To verify the described manipulations an Ambisonic metering tool was developed. The rms and peak level is measured at discrete points on the sphere. A texture is generated using the rms levels and mapped back onto a sphere for 3D visualization. The peak values are indicated by balls that are distributed on the sphere (Fig. 7).
Additionally the Mollweide projection is used to create a two-dimensional representation of the directional Ambisonic loudness levels.
Currently the visualization is implemented in Pure Data. It is feasible to transfer all computations to the GPU which is usually not busy with audio applications.

Figure 7: Ambisonic metering by sampling the surround scene, measuring audio levels and interpolating the measurement points with spherical harmonics to generate the texture.



Conclusion

A pragmatic approach for calculating Ambisonic transformation matrices has been presented. These transformations can be used to attenuate or boost certain directions in Ambisonic recordings, to rotate, and to warp the spatial image in certain directions. The algorithms have been implemented as ready-to-use audio plug-ins applicable to production and postproduction of Ambisonic recordings. Additionally the transformations can be used to adapt Ambisonic recordings to certain playback situations. For all that, a new way of metering the Ambisonic surround production is required and was successfully presented.

Acknowledgment

The authors would like to thank the Music Innovation Studies Centre Vilnius where parts of the development and tests where carried out.

References

[1] F. Zotter and H. Pomberger, "Warping of the Recording Angle in Ambisonics", in 1st International Conference on Spatial Audio, Detmold, 2011.

[2] R. H. Hardin and N. J. A. Sloane, "McLaren's Improved Snub Cube and Other New Spherical Designs in Three Dimensions", in Discrete Computational Geometry, vol. 15, pp. 429-441, 1996.

[3] M. Gräf and D. Potts, "On the computation of spherical designs by a new optimization approach based on fast spherical Fourier transforms", Numer. Math. 119, 699 - 724, 2011.

[4] M. Kronlachner and F. Zotter, "Spatial transformations for the enhancement of Ambisonic recordings", in 2nd International Conference on Spatial Audio, Erlangen, 2014.

EAA Joint Symposium on Auralization and Ambisonics 2014
TU Berlin, Germany, April 3-5, 2014.

Matthias Kronlachner and Franz Zotter

Institute of Electronic Music and Acoustics
Inffeldgasse 10 / III, 8010 Graz, http://iem.at