Analytic error control methods for rotation in Ambisonics

鄭任傑

Outline

- Motivation
- Ambisonics
- Rotation in Ambisonics
- Conclusion

Motivation

- Build a theory for spatial audio that is channel agnostic, homogeneous and coherent, but also has good localization with few channels.

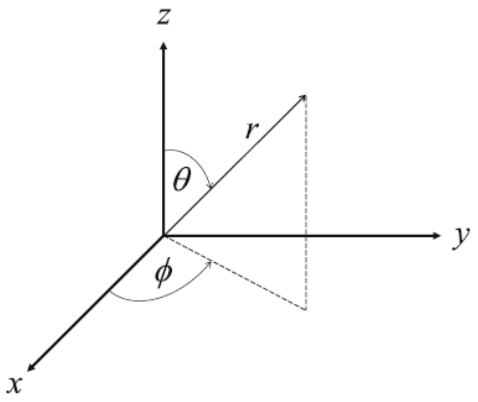
Spatial Audio Techniques

- Channel-based: the whole encoding/decoding and recording/reproduction is based on a specic channel layout, e.g. 2.0, 5.1, 7.1,..., Auro3D, Hamasaki 22.2.
- Layout-independent (channel-agnostic): the recording and encoding format is independent from the reproduction layout (includes sound field reconstruction methods and object-based formats). e.g. Ambisonics.

Ambisonics

- Ambisonics comprises both encoding, recording and reproduction (decoding) techniques that can be used live or in studio to present a 2-dimensional (planar, or horizontal-only) or 3-dimensional (periphonic, or full-sphere) sound field.
- Ambisonics encoding of the sound field whitch based on the acoustic wave equation and its accurate reconstruction in a point in space.

The Acoustic Wave Equation



$$\nabla^2 p(.,t) - \frac{1}{c}\frac{\partial^2}{\partial t^2} p(.,t) = 0$$

where ∇^2 is Laplacian in spherical coordinates $f(r,\theta,\phi)$

$$\nabla^2 f \equiv \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r} f\right) + \frac{1}{r^2 sin\theta}\frac{\partial}{\partial \theta}\left(sin\theta \frac{\partial}{\partial \theta} f\right) + \frac{1}{r^2 sin^2\theta}\frac{\partial^2}{\partial \phi^2} f$$

- The solution of the PDE is the spherical harmonics

$$Y_n^m(\theta,\phi) = \sqrt{\frac{2n+1}{4\pi}\frac{(n-m)!}{(n+m)!}} P_n^m(cos\theta) e^{im\phi}$$

where *n* is the order of the spherical harmonics

Spherical Harmonics

$n = 0$	$Y_0^0(\theta, \phi) = \sqrt{\frac{1}{4\pi}}$
$n = 1$	$Y_1^{-1}(\theta, \phi) = \sqrt{\frac{3}{8\pi}} \sin\theta e^{-i\phi}$
	$Y_1^0(\theta, \phi) = \sqrt{\frac{3}{4\pi}} \cos\theta$
	$Y_1^1(\theta, \phi) = -\sqrt{\frac{3}{8\pi}} \sin\theta e^{i\phi}$
$n = 2$	$Y_2^{-2}(\theta, \phi) = \sqrt{\frac{15}{32\pi}} \sin^2\theta e^{-2i\phi}$
	$Y_2^{-1}(\theta, \phi) = \sqrt{\frac{15}{8\pi}} \sin\theta \cos\theta e^{-i\phi}$
	$Y_2^0(\theta, \phi) = \sqrt{\frac{5}{16\pi}} (3\cos^2\theta - 1)$
	$Y_2^1(\theta, \phi) = -\sqrt{\frac{15}{8\pi}} \sin\theta \cos\theta e^{i\phi}$
	$Y_2^2(\theta, \phi) = \sqrt{\frac{15}{32\pi}} \sin^2\theta e^{2i\phi}$

Spherical Harmonics

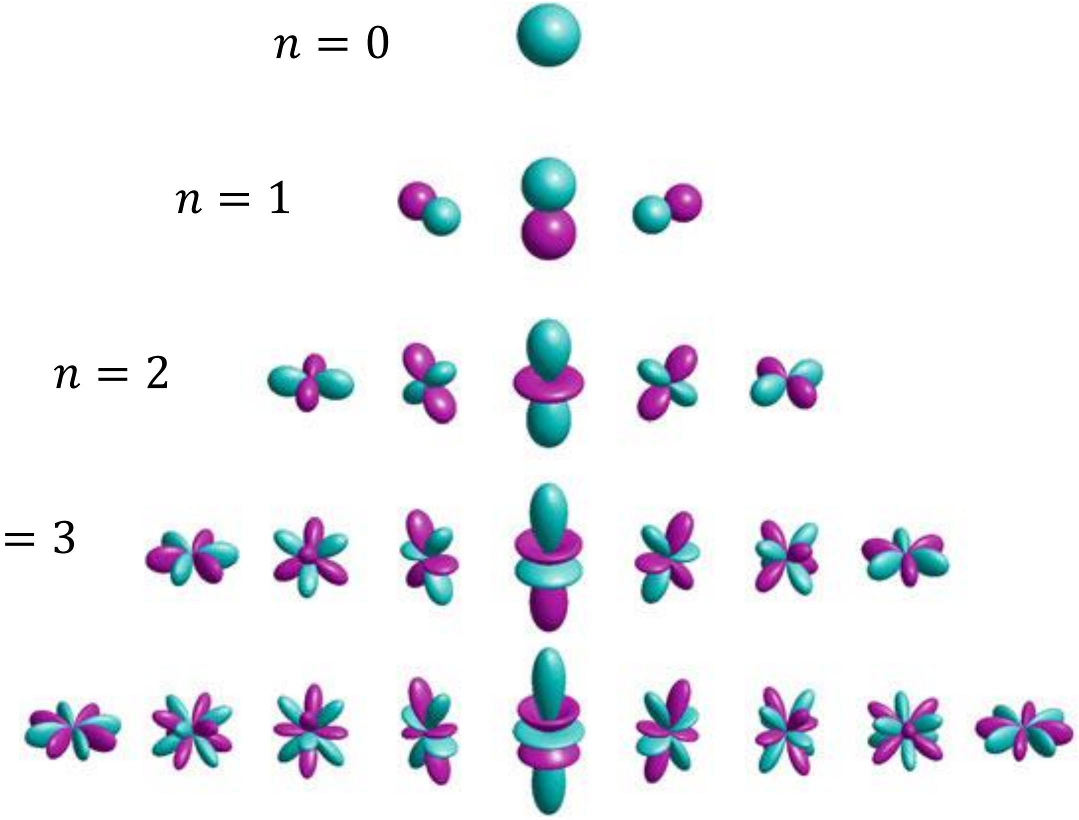
$n = 0$

$n = 1$

$n = 2$

$n = 3$

$n = 4$



Spherical Fourier Transform

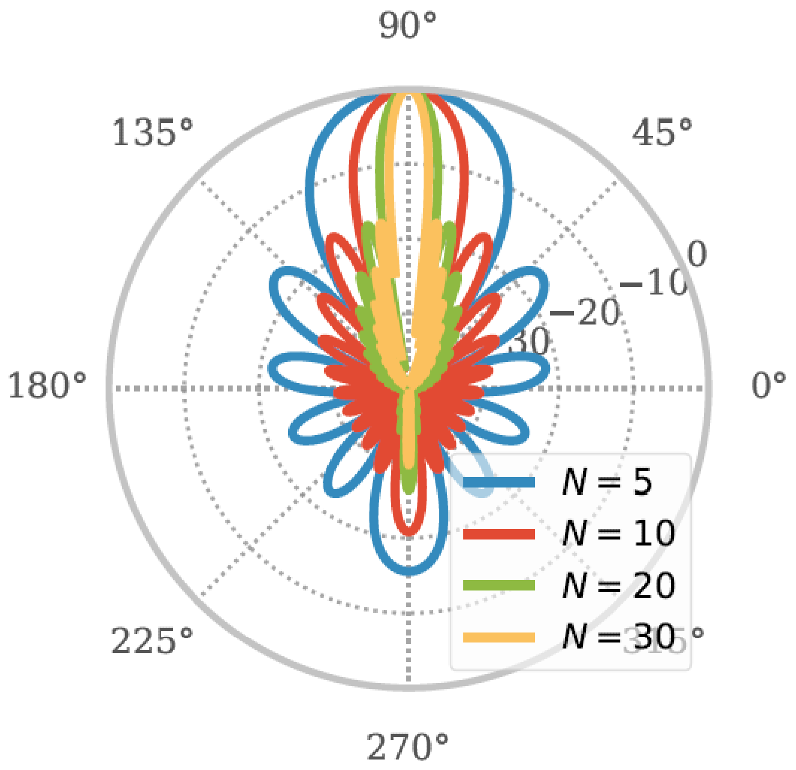
- The set of spherical harmonics $Y_n^m(\theta, \phi)$, for n ≥ 0 and −n ≤ m ≤ n , can be used to compose a wide range of functions on the sphere.
- It also be called Ambisonics representation or encoding higher order Ambisonics.

$$f(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} f_{n,m} \, Y_n^m(\theta, \phi)$$

- inverse spherical Fourier transform or decoding higher order Ambisonics.

$$f_{n,m} = \int f(\theta, \phi) [Y_n^m(\theta, \phi)]^* \, d\Omega$$

SH Representation Plot



Head rotation in SH domain

- rotation matrix

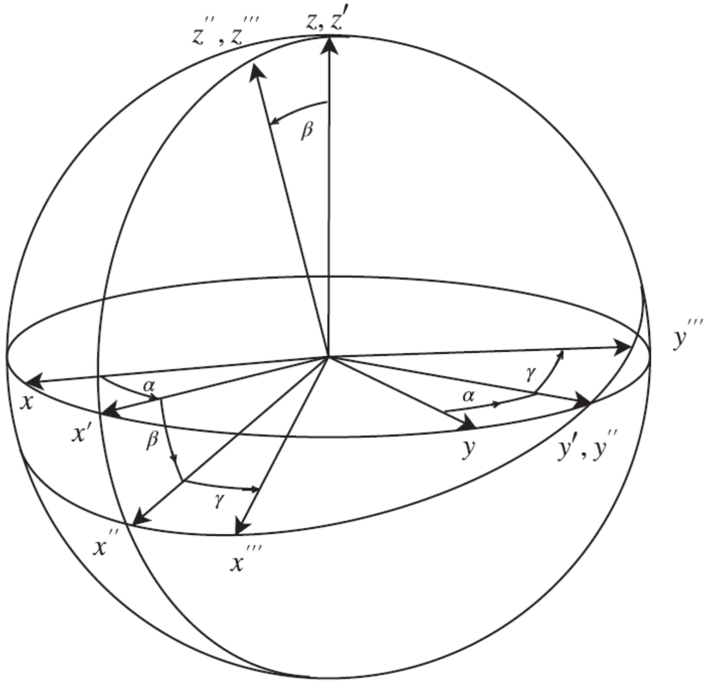
$$\hat{g} = \hat{u}(\alpha)\hat{a}(\beta)\hat{u}(\gamma) = \begin{bmatrix} cos\alpha & sin\alpha & 0 \\ -sin\alpha & cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} cos\beta & 0 & -sin\beta \\ 0 & 1 & 0 \\ sin\beta & 0 & cos\beta \end{bmatrix} \begin{bmatrix} cos\gamma & sin\gamma & 0 \\ -sin\gamma & cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$f'_{n,m} = \sum_{m\prime=-n}^{n} f_{n,m'} \mathcal{D}^{n}_{m',m}$$

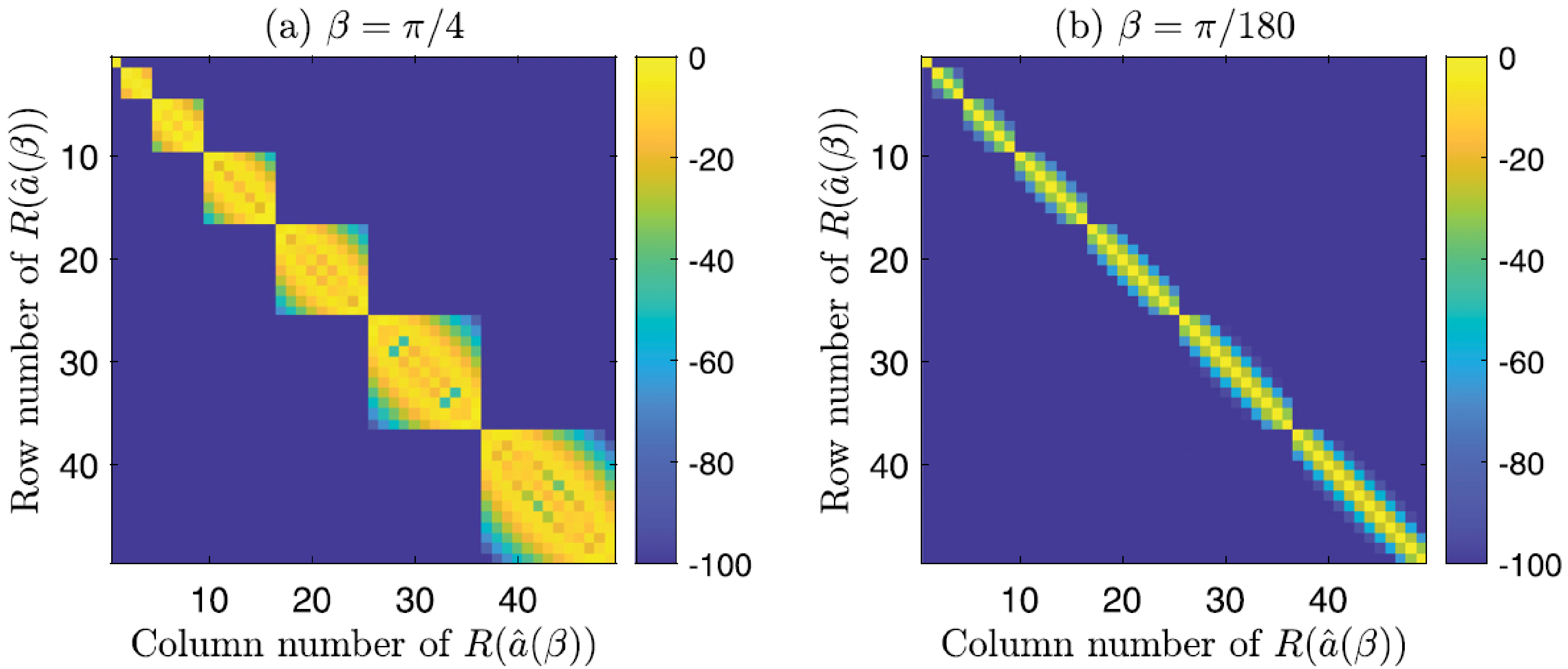
where $\mathcal{D}^{n}_{m',m}$ is the element of Wigner D-matrix.

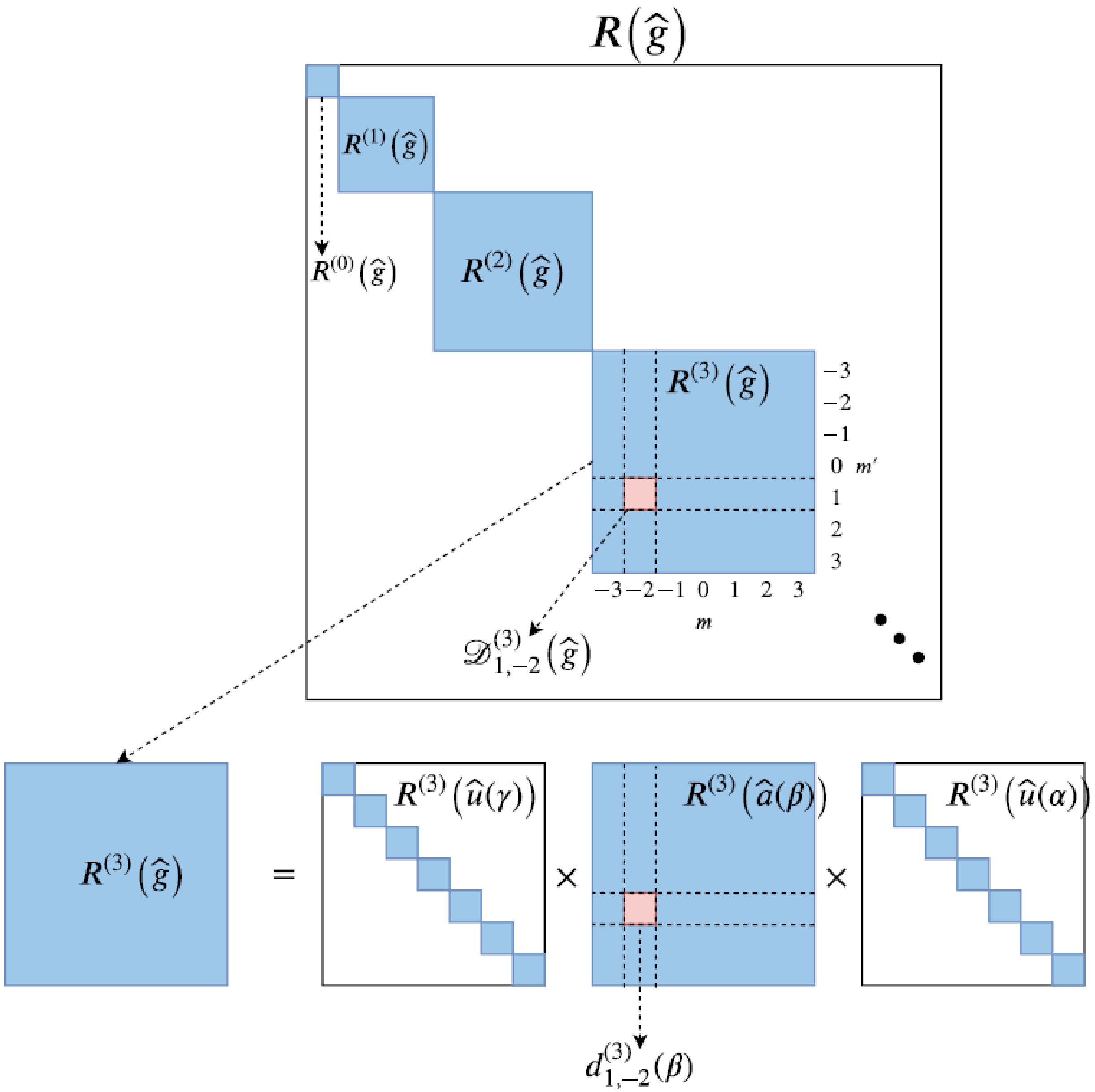
$$\mathcal{D}^{n}_{m',m} = e^{im'\gamma} d^{n}_{m',m}(\beta) e^{im\alpha}$$

where $d^{n}_{m',m}$ is the element of Wigner's (small) d-matrix.

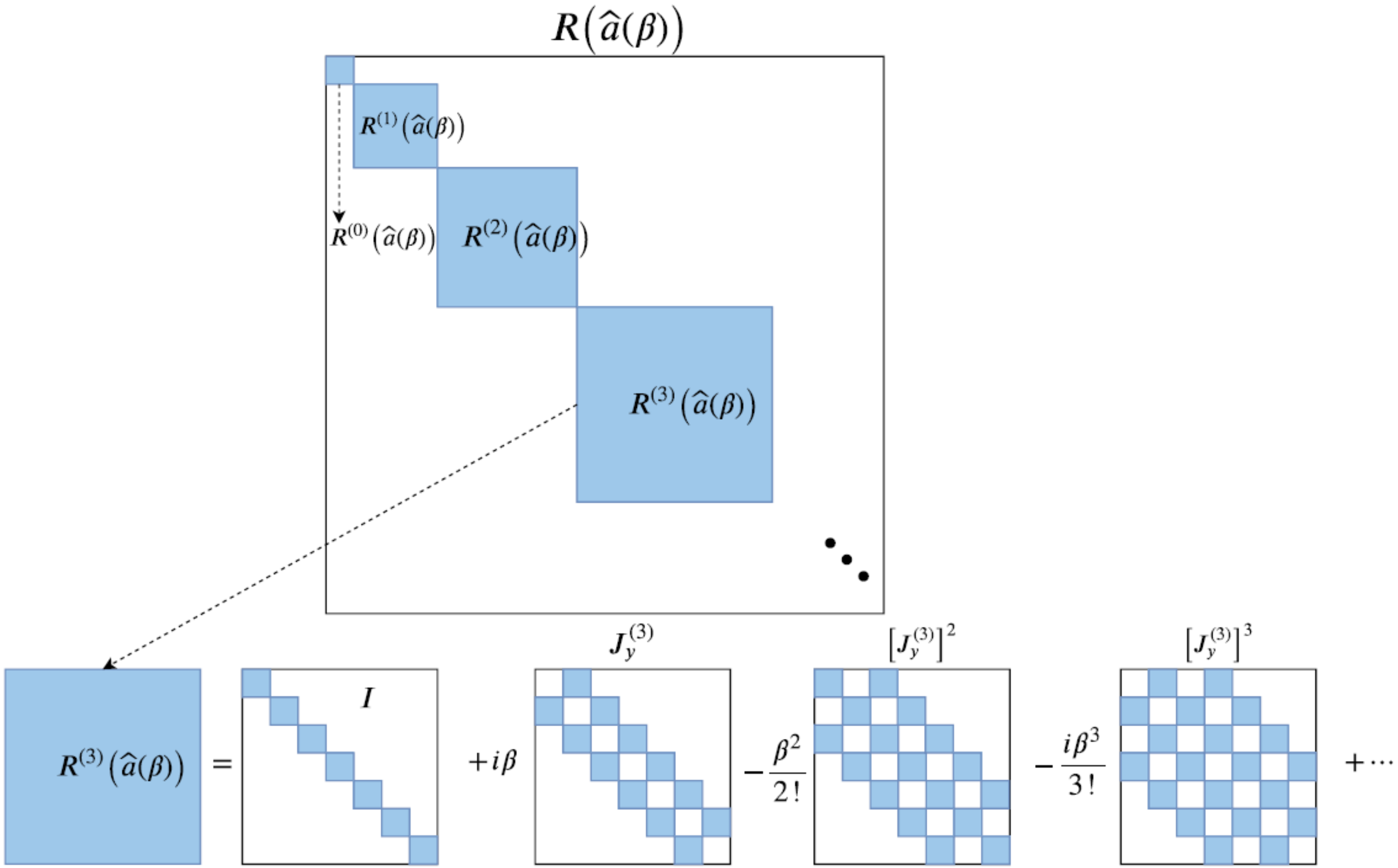


Magnitudes of WdMs





$R(\hat{g})$
$R^{(0)}(\hat{g})$
$R^{(1)}(\hat{g})$
$R^{(2)}(\hat{g})$
$R^{(3)}(\hat{g})$
−3 −2 −1 0 1 2 3 m'
−3 −2 −1 0 1 2 3
m
$\mathscr{D}^{(3)}_{1,-2}(\hat{g})$
$R^{(3)}(\hat{g})$ = $R^{(3)}(\hat{u}(\gamma))$ × $R^{(3)}(\hat{a}(\beta))$ × $R^{(3)}(\hat{u}(\alpha))$
$d^{(3)}_{1,-2}(\beta)$



$R(\hat{a}(\beta))$
$R^{(1)}(\hat{a}(\beta))$
$R^{(0)}(\hat{a}(\beta))$
$R^{(2)}(\hat{a}(\beta))$
$R^{(3)}(\hat{a}(\beta))$
$R^{(3)}(\hat{a}(\beta)) = I + i\beta J_y^{(3)} - \frac{\beta^2}{2!}[J_y^{(3)}]^2 - \frac{i\beta^3}{3!}[J_y^{(3)}]^3 + \cdots$

Conclusion

Ambisonics provide

- Nice physical formulation
- Reproduce the sound field
- Getting a lot of application in Augmented/Virtual Reality (AR/VR)

Reference

1. T. Magariyachi and Y. Mitsufuji, “Analytic error control methods for efficient rotation in dynamic binaural rendering of Ambisonics,” *The Journal of the Acoustical Society of America*, vol. 147, no. 1, pp. 218–230, 2020.
2. C. Hold, H. Gamper, V. Pulkki, N. Raghuvanshi, and I. J. Tashev, “Improving Binaural Ambisonics Decoding by Spherical Harmonics Domain Tapering and Coloration Compensation,” *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019.
3. B. Rafaely, “Analysis and design of spherical microphone arrays,” *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 1, pp. 135–143, 2005.
4. B. Rafaely, *Fundamentals of Spherical Array Processing*. Cham: Springer International Publishing, 2019.