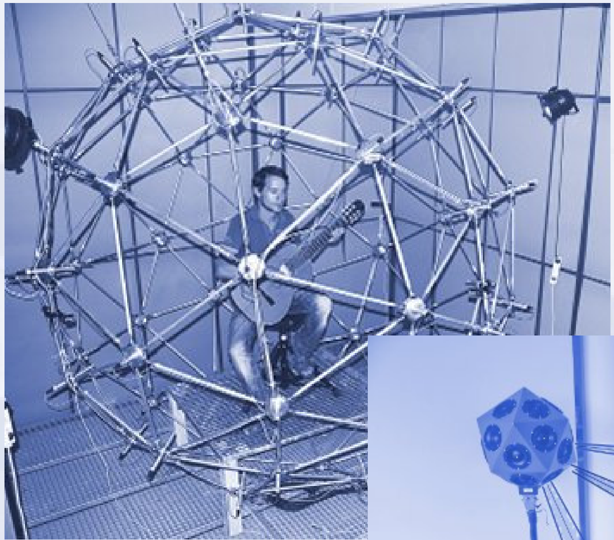
Techniques and Considerations on Sound Field Recording and Reproduction Using Spherical Harmonics



Franz Zotter

Hannes Pomberger

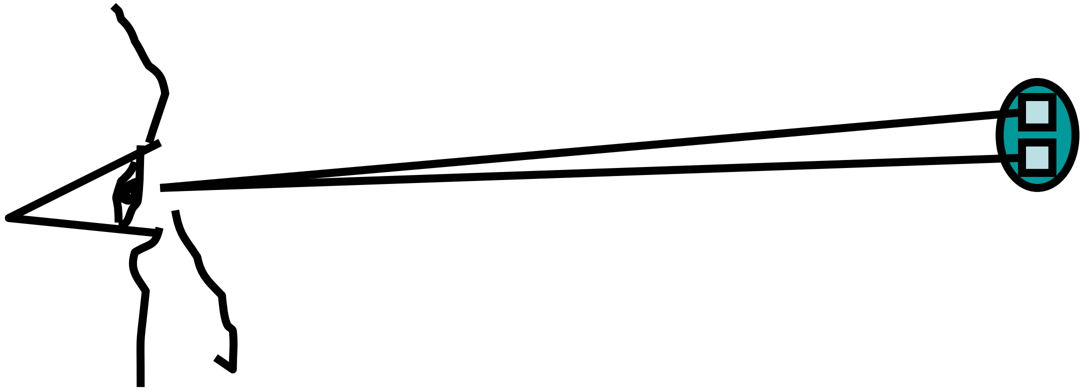
Matthias Frank

Georgios Marentakis

Basics of directional resoultion

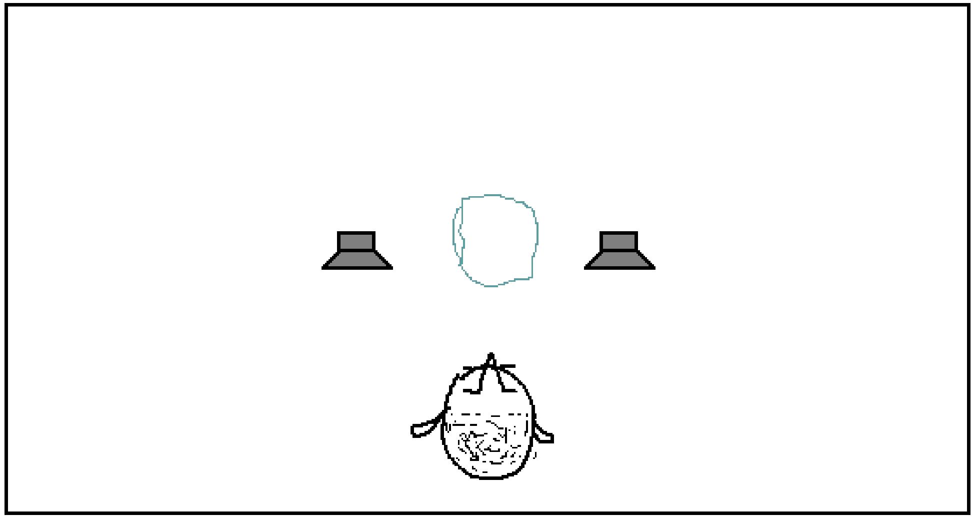
Resolution of the eye:

Pixels start to mix for angles < 1/40 °



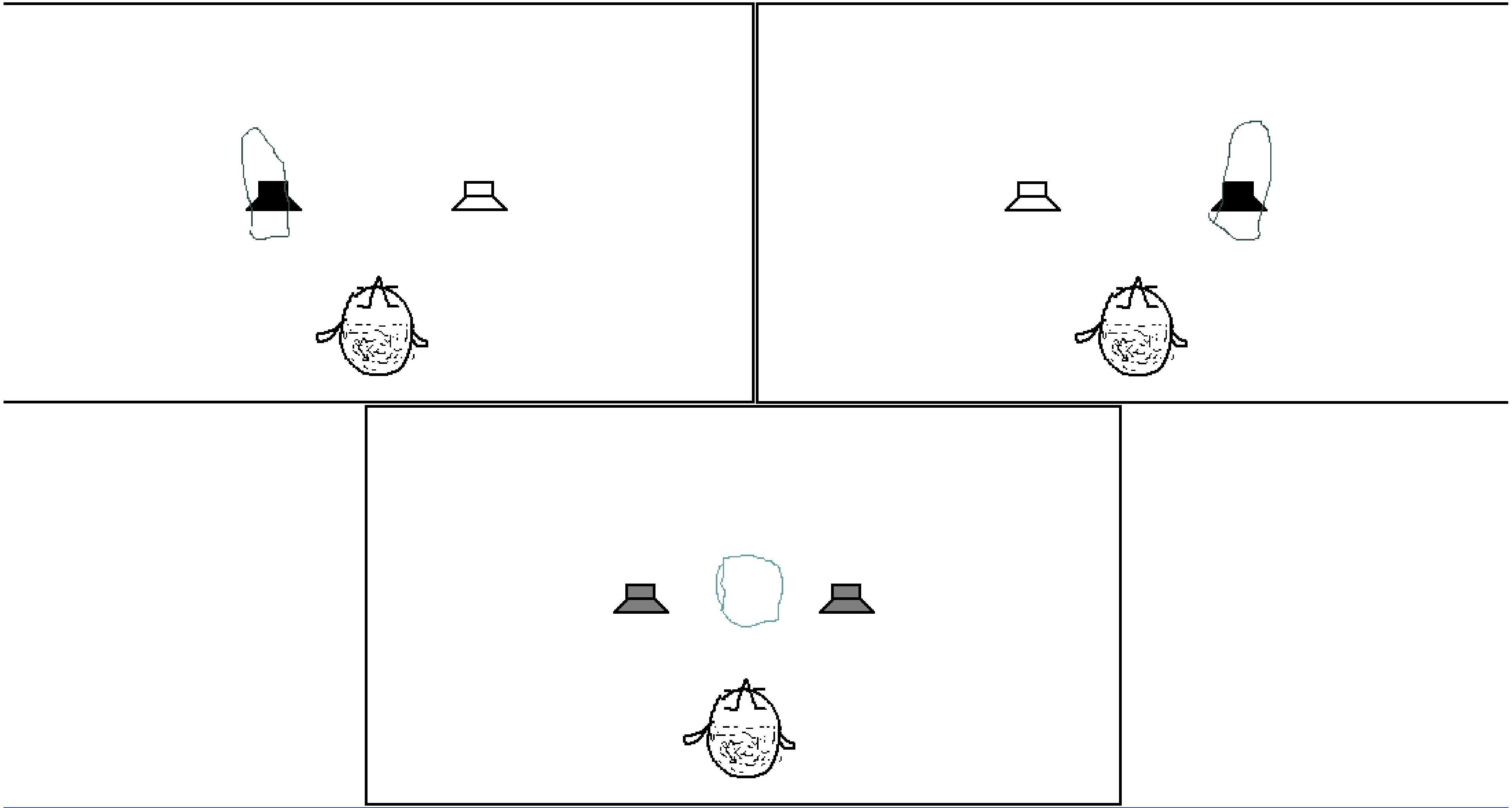
Basics of directional resoultion

Localization of coherent sound is different

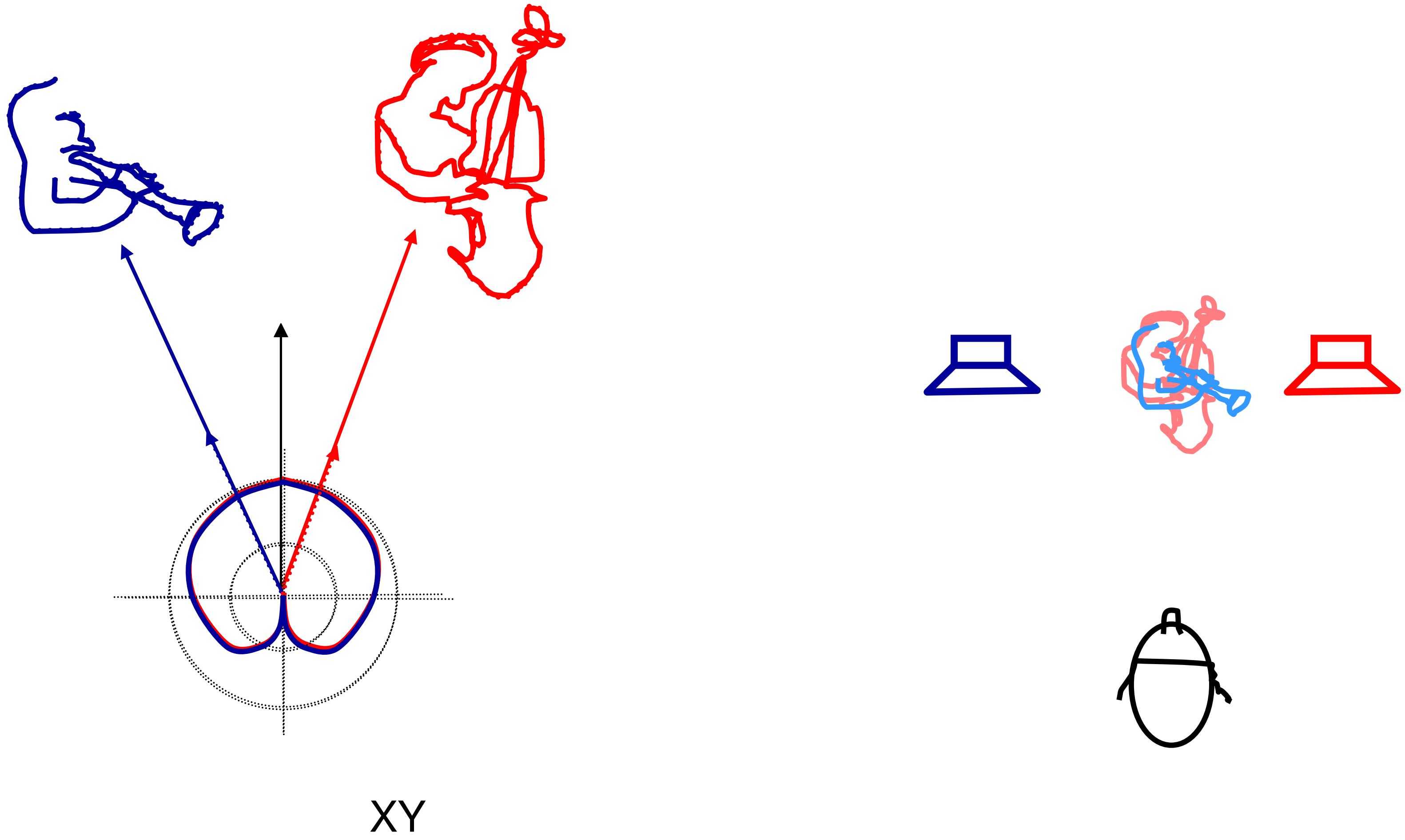


Basics of directional resoultion

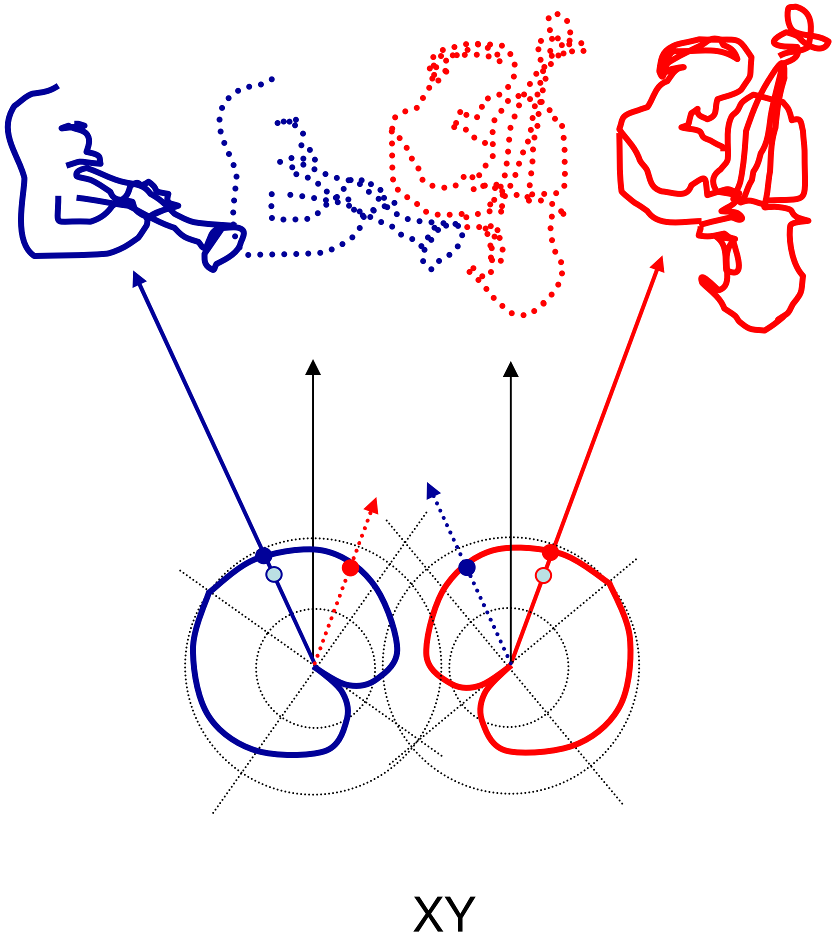
Localization of coherent sounds is different

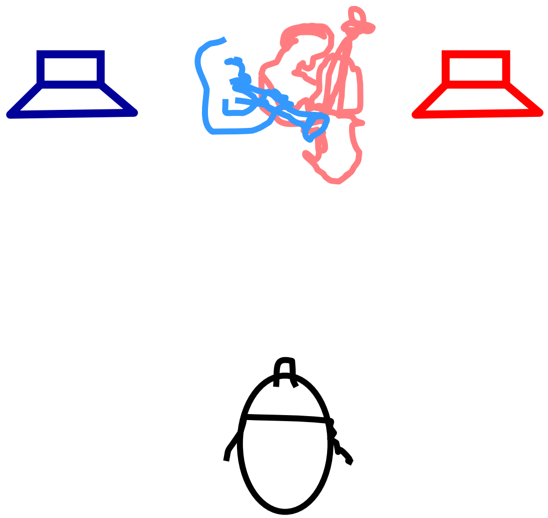


Stereophony

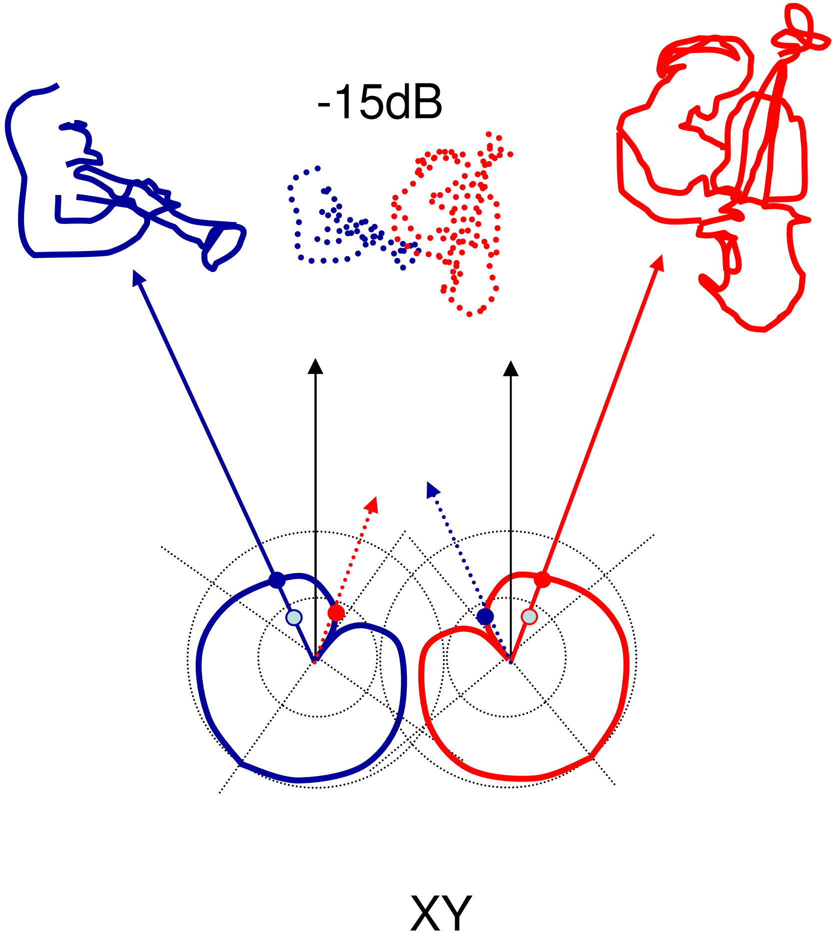


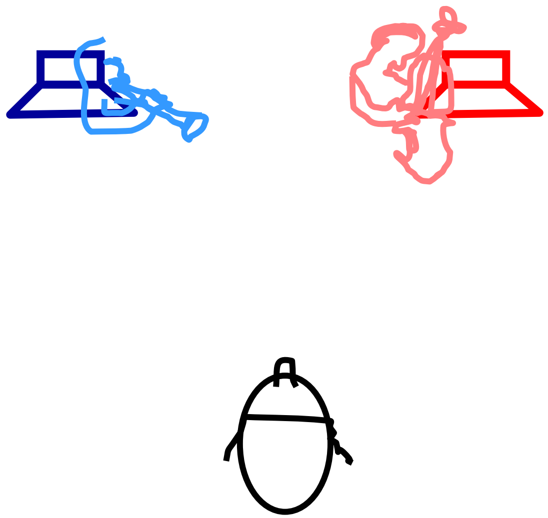
Stereophony





Stereophony

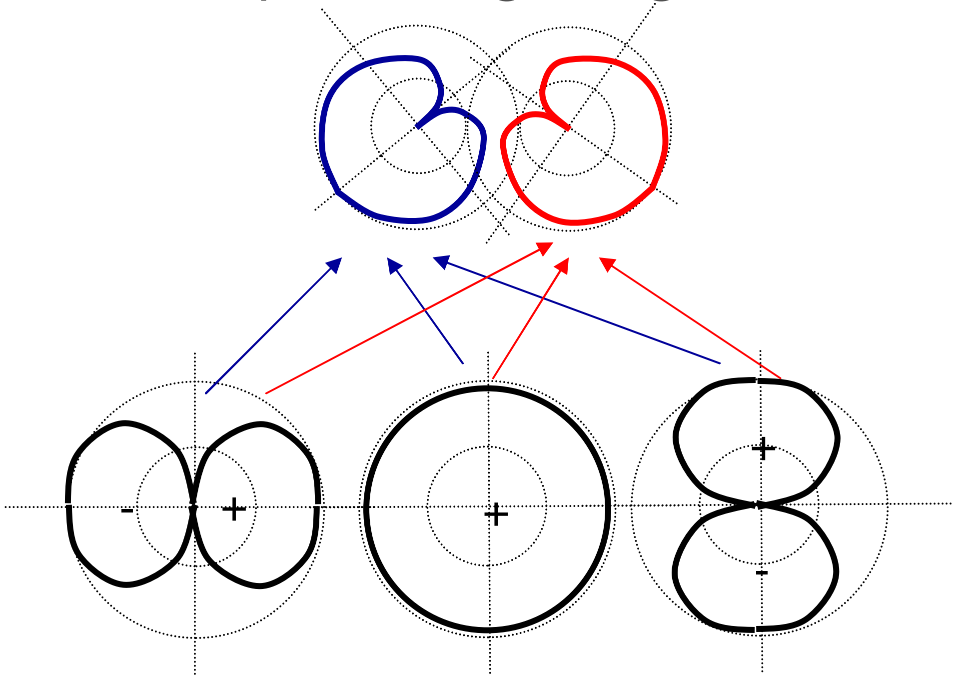


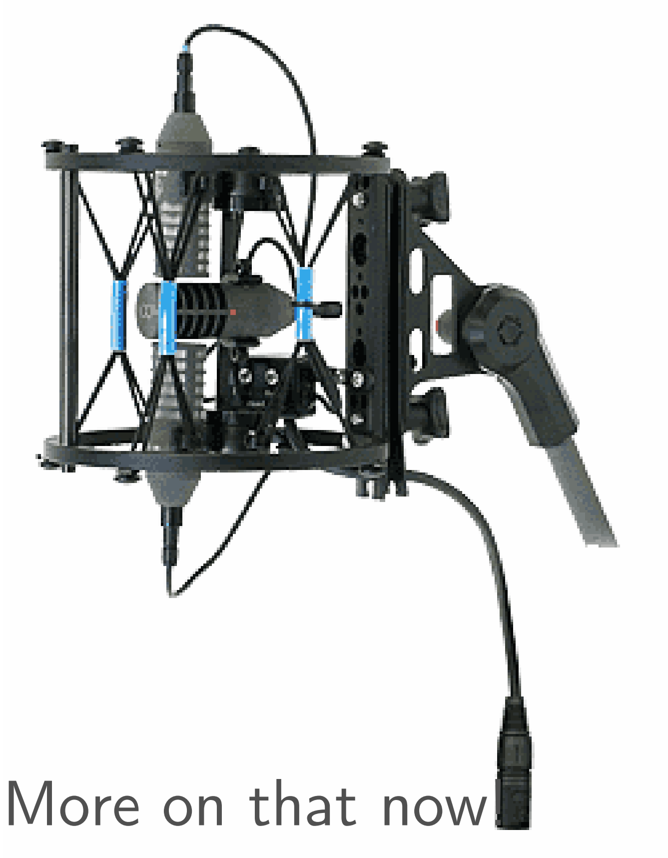


Stereophony – Schöps Doppel-MS

Adjustable stereo image by combining basic directional patterns

But also capable of getting surround image

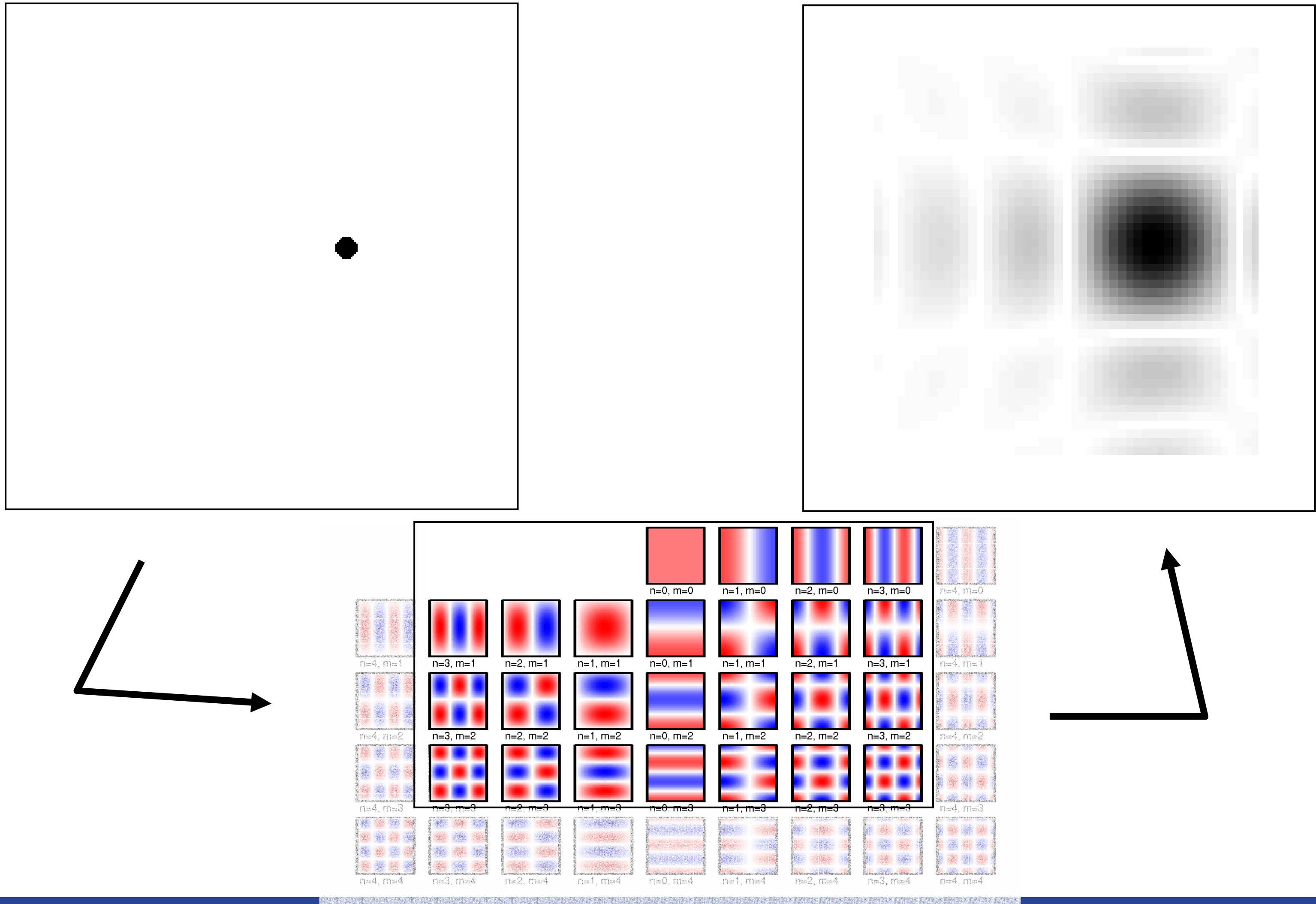




At which resolution does it image? – More on that now!

Basics of resolution on the plane

Smoothing a visual image: harmonic representation/truncation

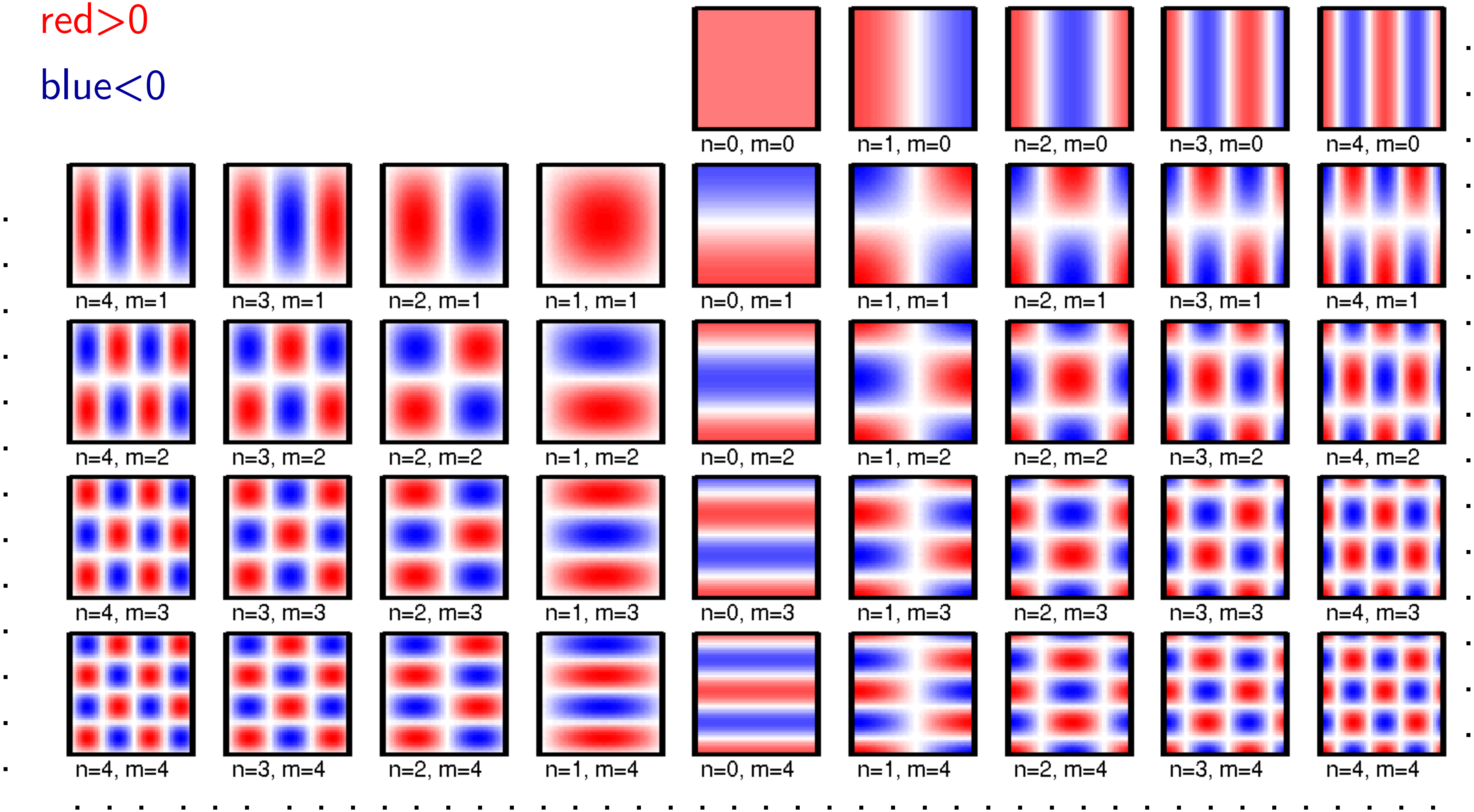


Basics of resoultion on the plane

Smoothing a visual image: harmonic representation/truncation

red>0

blue<0

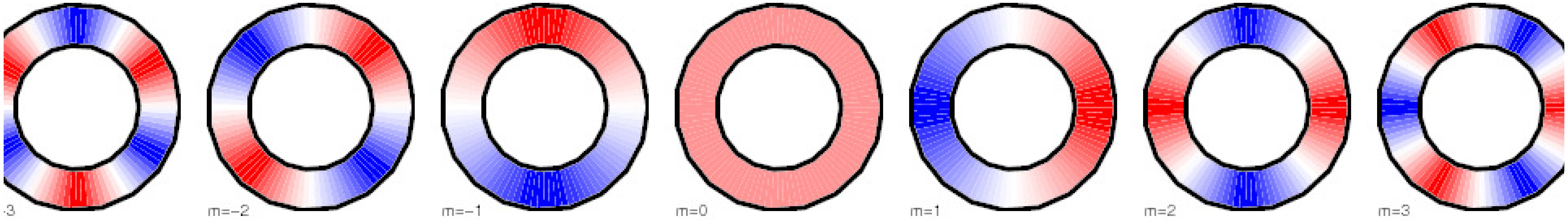


Basics of directional resoultion – horizontal surround

Smooth surround imaging: harmonic representation/truncation

red>0

blue<0

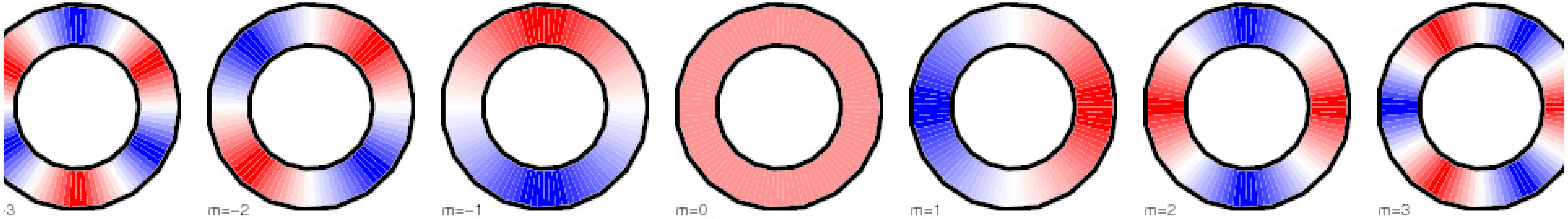


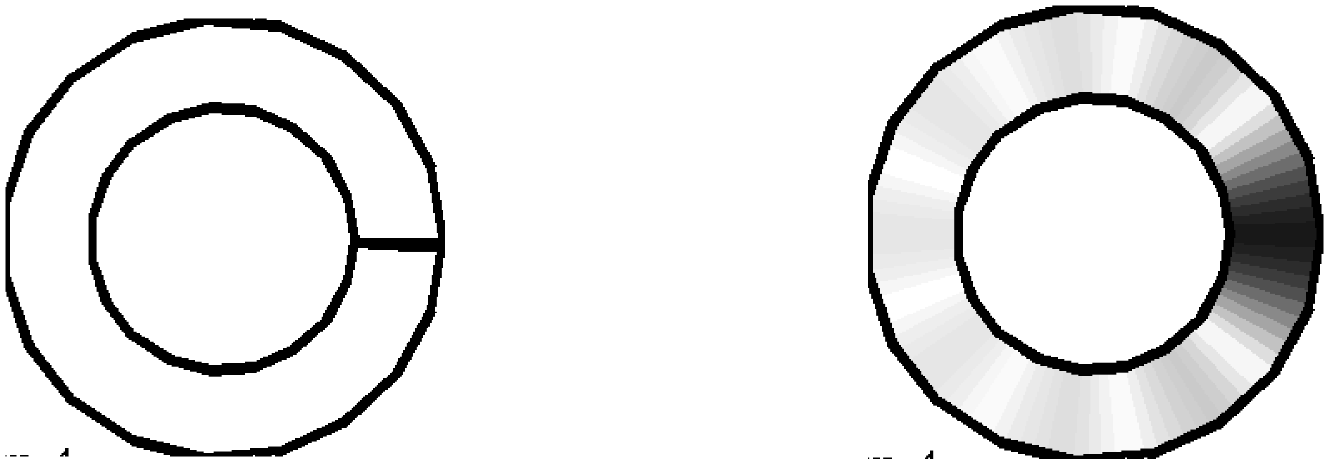
Basics of directional resoultion – horizontal surround

Smooth surround imaging: harmonic representation/truncation

red>0

blue<0



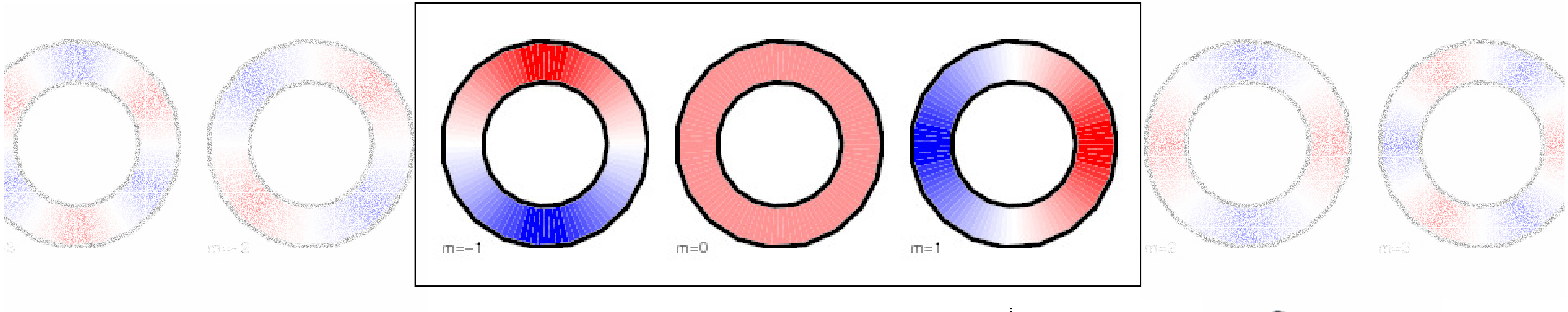


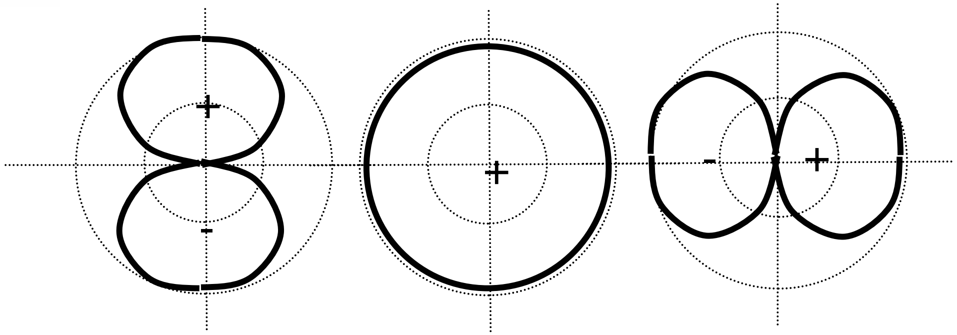
Basics of directional resoultion – horizontal surround

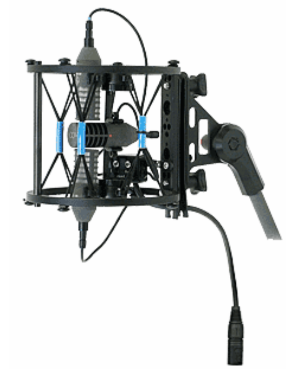
Smooth surround imaging: harmonic representation/truncation

red>0

blue<0



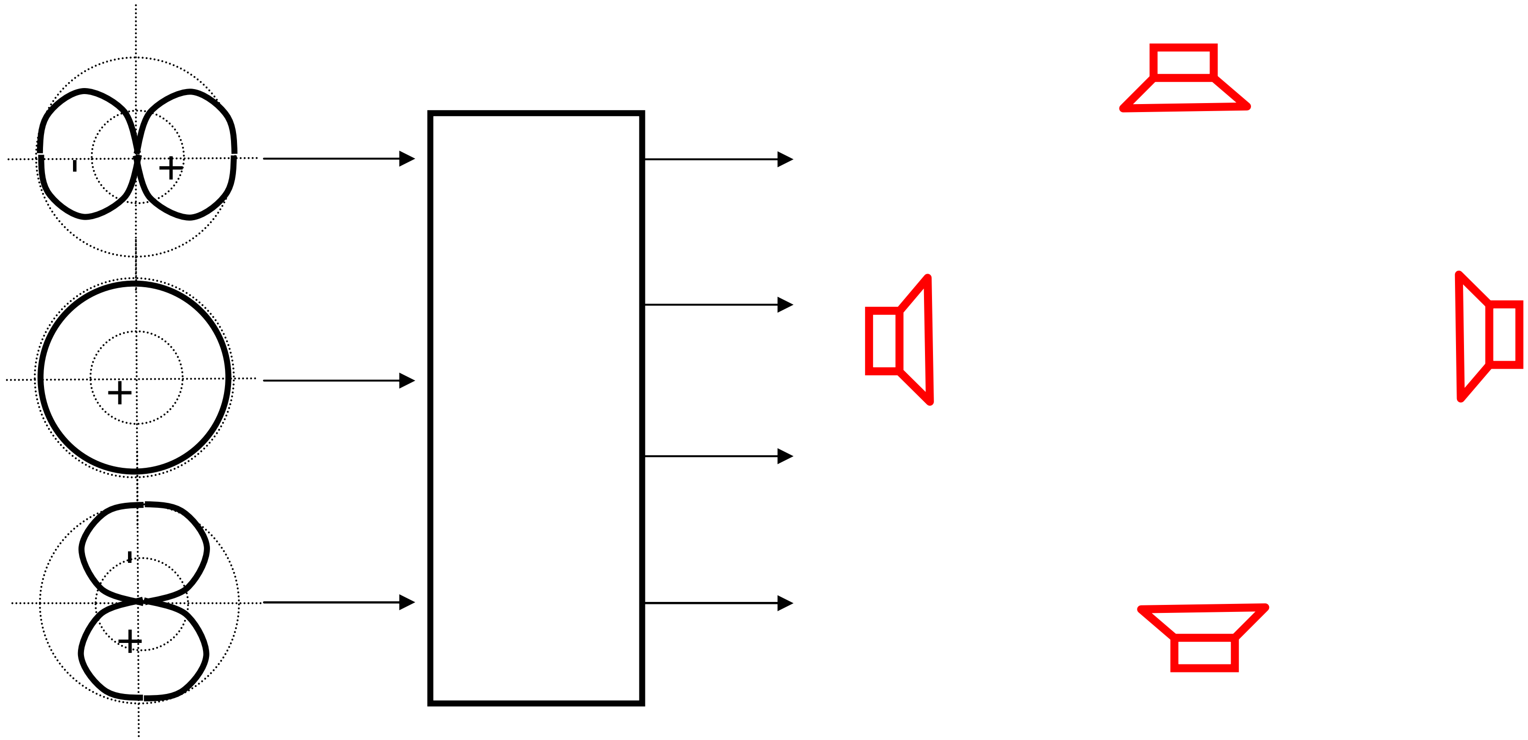




Resolution is about 180°

Ambisonics – horizontal surround

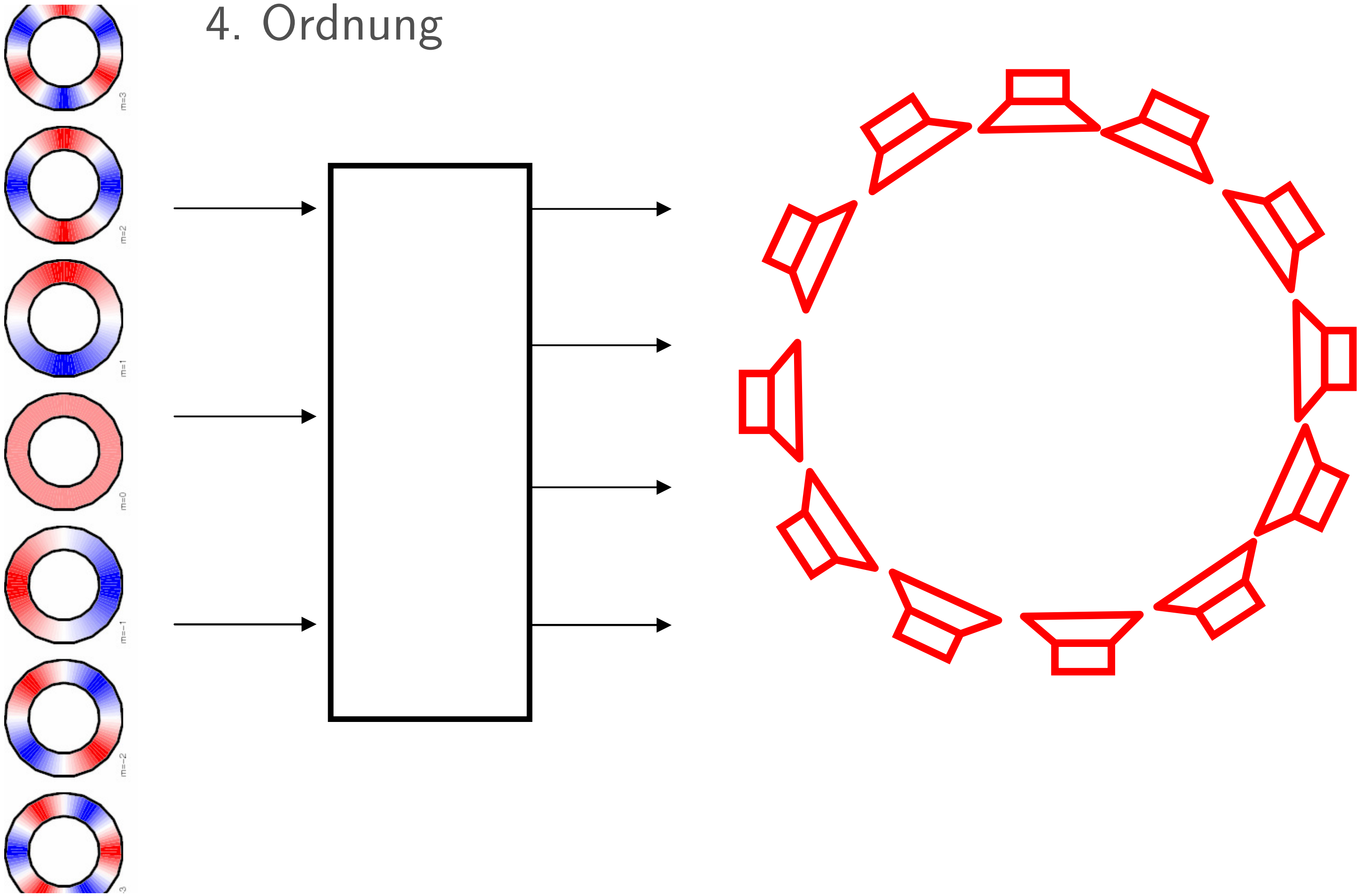
Ambisonics as a surround recording format – decoding to loudspeakers



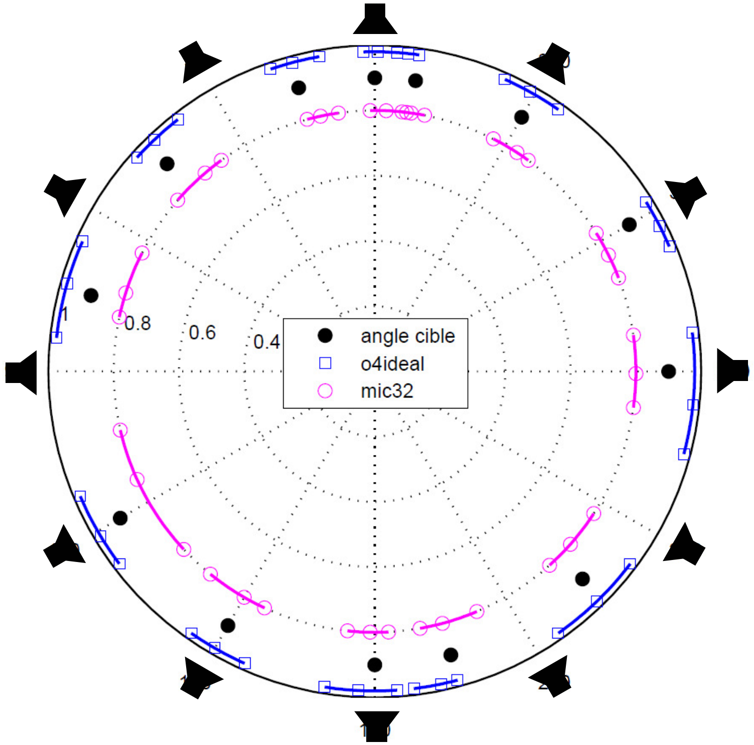
Ambisonics – horizontal surround

Artificially for surround panning

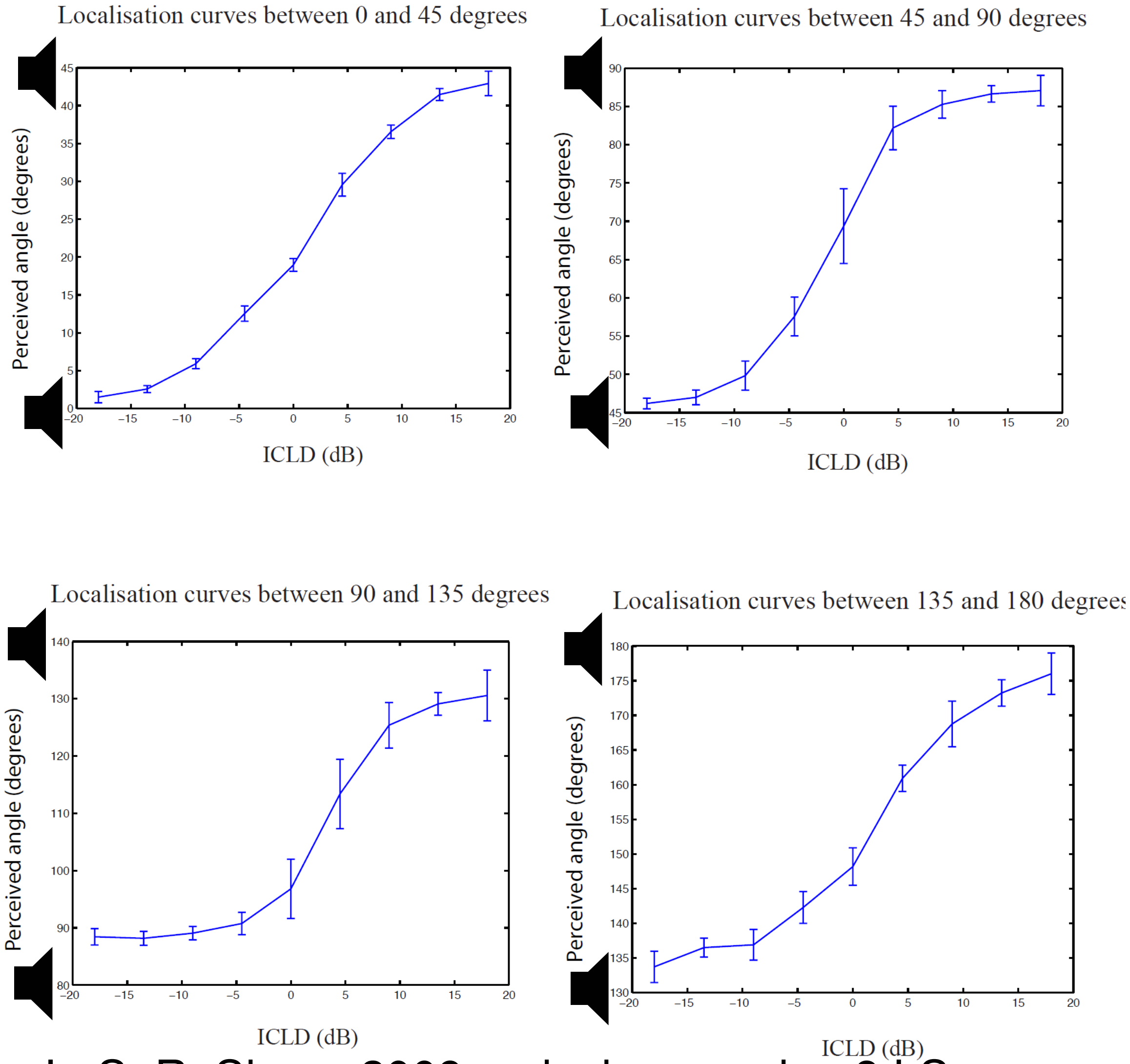
4. Ordnung



Listening tests – Stéphanie Bertet (ircam/orange-labs)

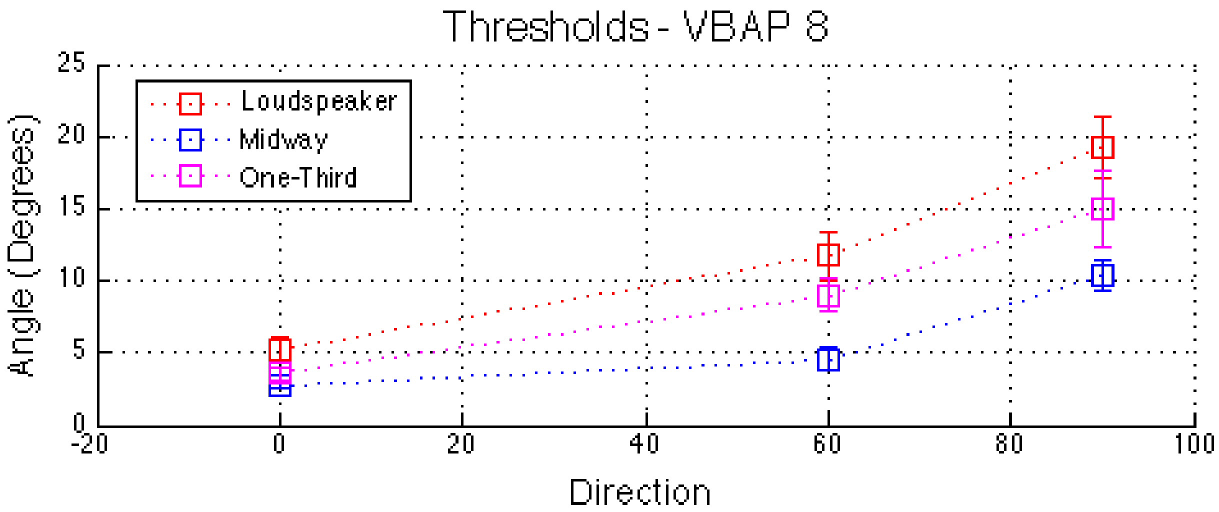


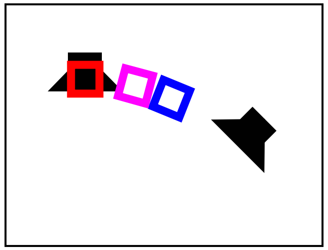
S. Bertet 2009
Ambisonics 12 LS

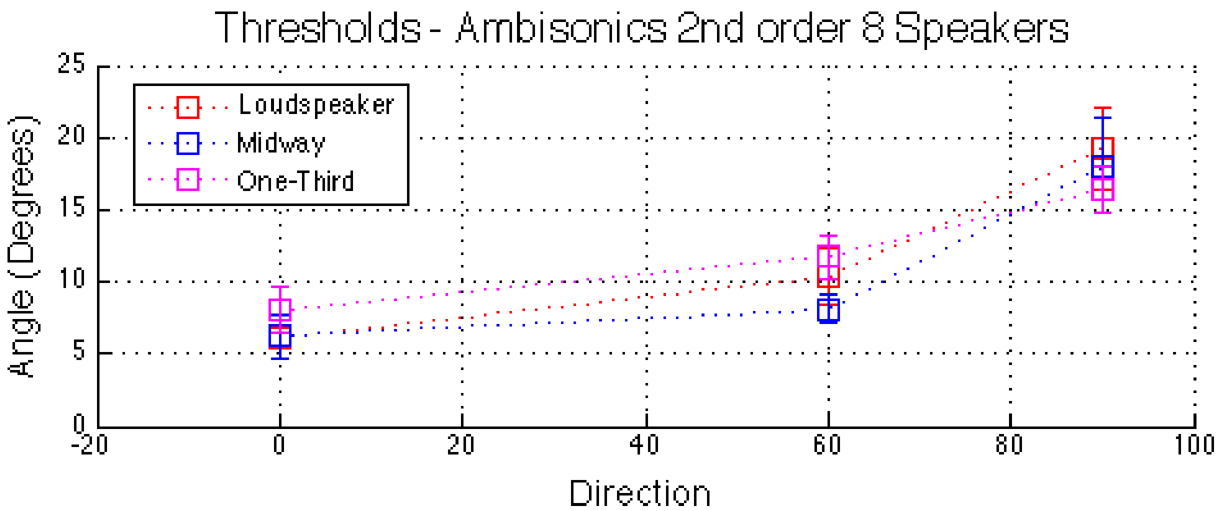


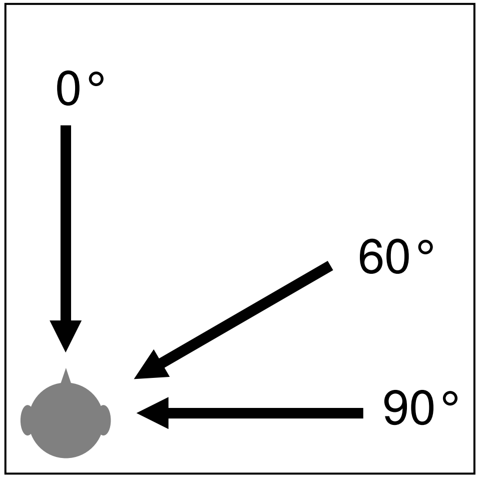
L. S. R. Simon 2009, pairwise panning 8 LS

Listening tests – Georgios Marentakis (former work at McGill)



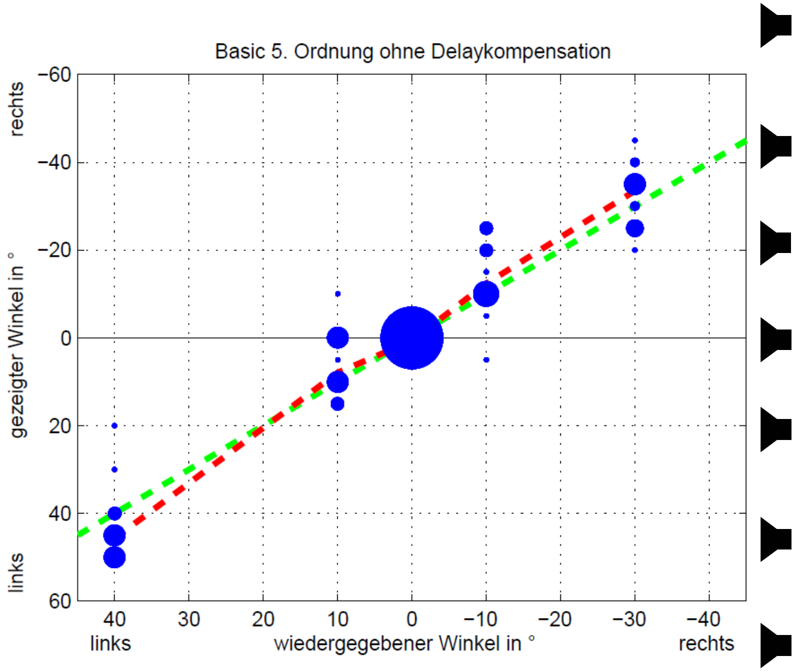




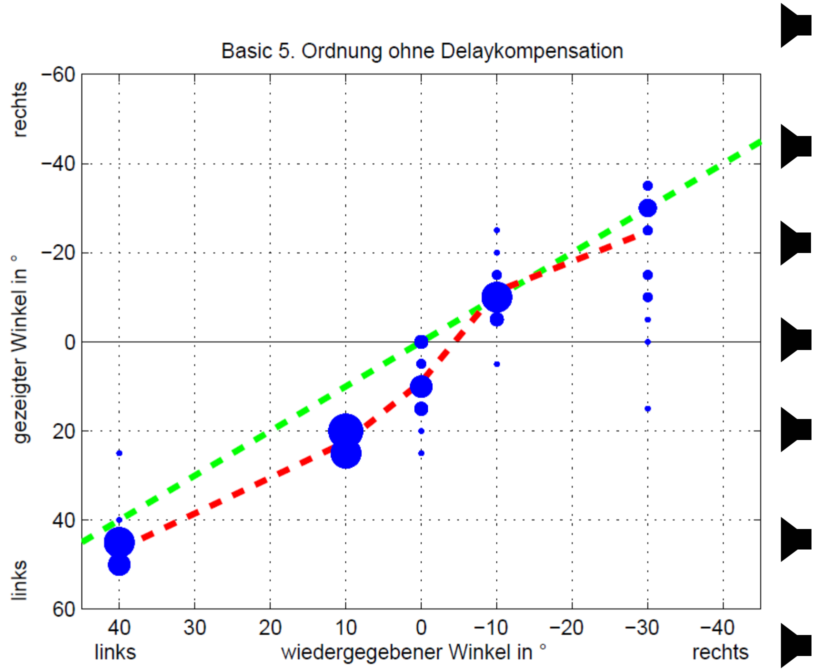


Listening tests – Matthias Frank (IEM)

inside sweet spots



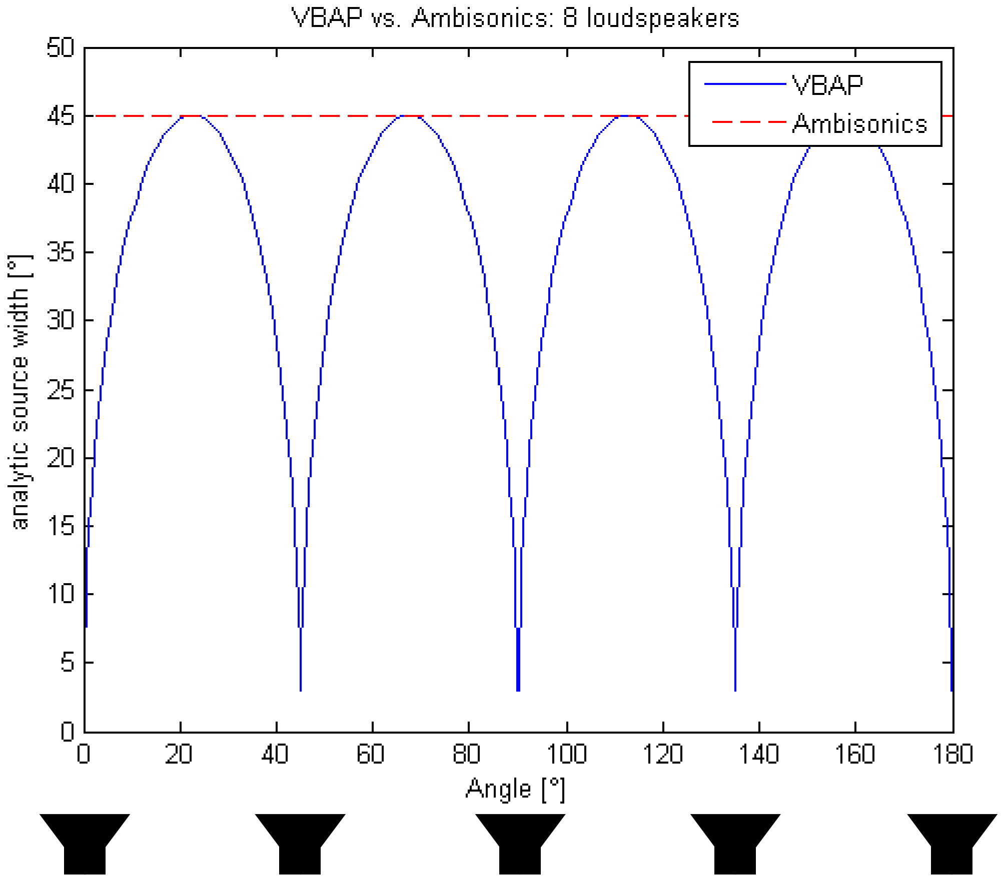
outside sweet spot



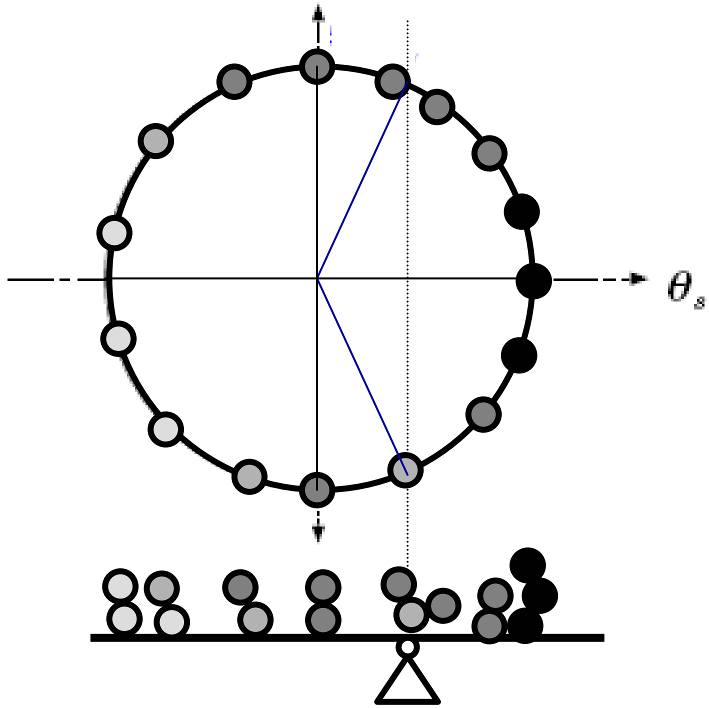
Ambisonics 5th order

Listening tests – Matthias Frank (IEM)

M. Frank 2011 ☺

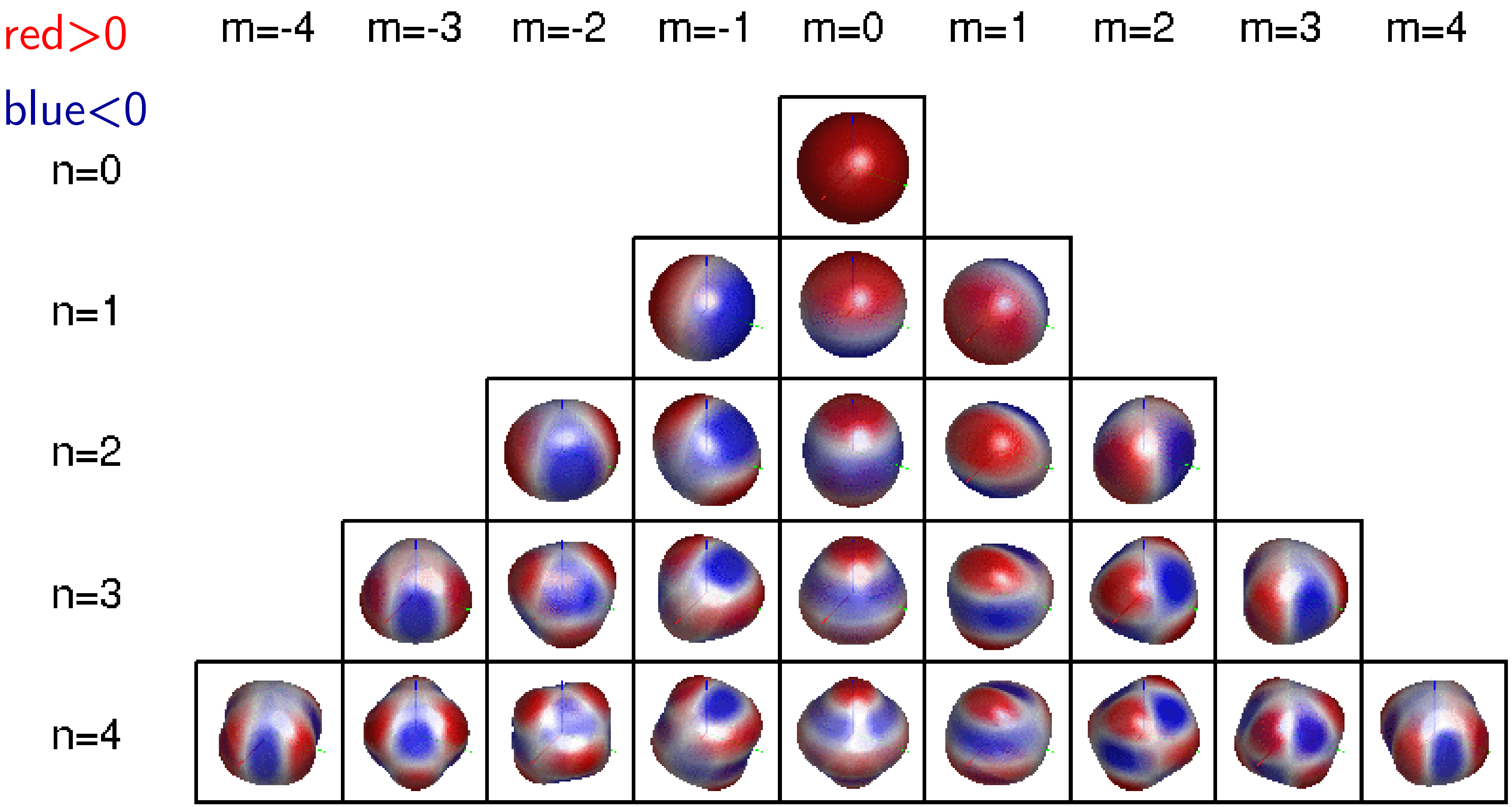


source width estimation



Basics of directional resoultion – surround (with height)

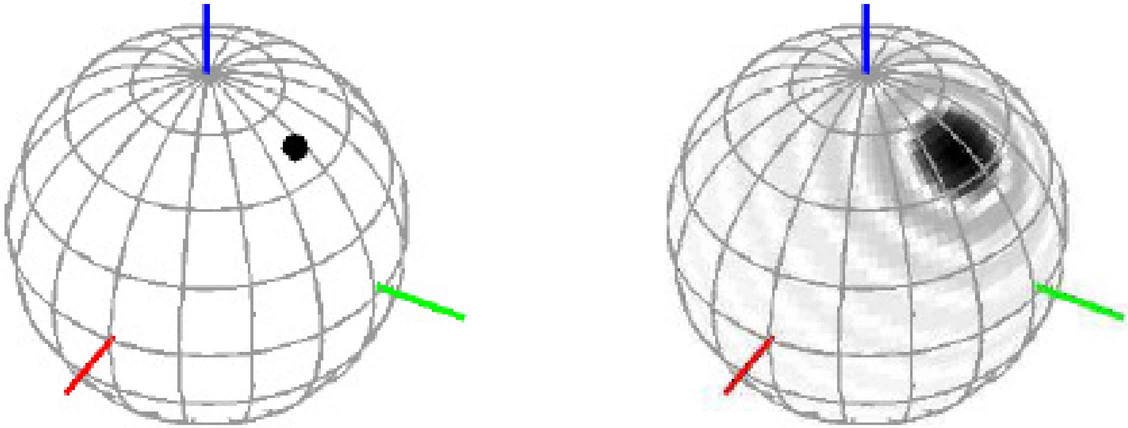
Smooth surround-with-height imaging:
harmonic representation/truncation



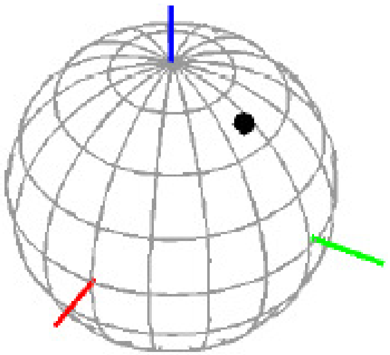
. . . .

Basics of directional resoultion – surround (with height)

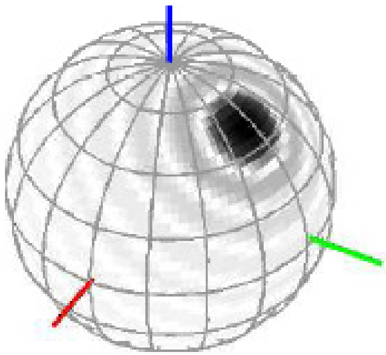
Smooth surround-with-height imaging:
harmonic representation/truncation



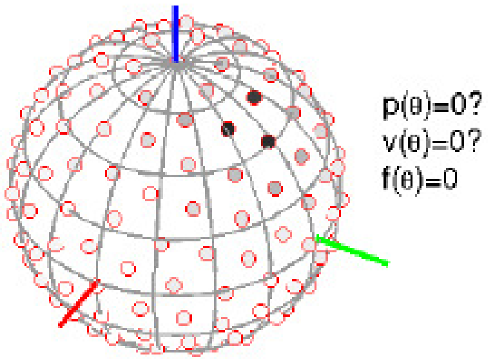
Ambisonics



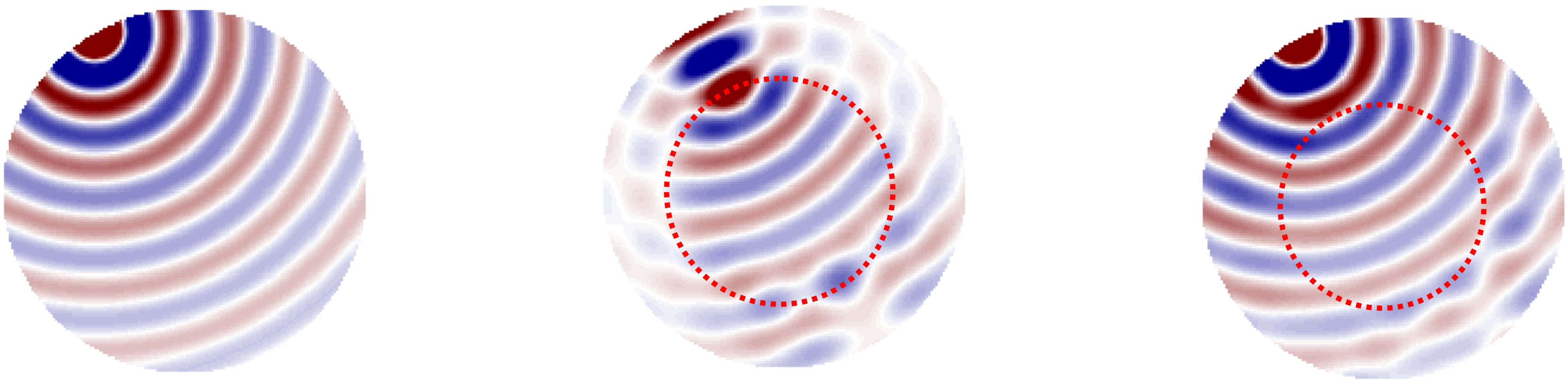
(a) Continuous distribution



(b) Angularly band-limited distr.



(c) Discretized distribution

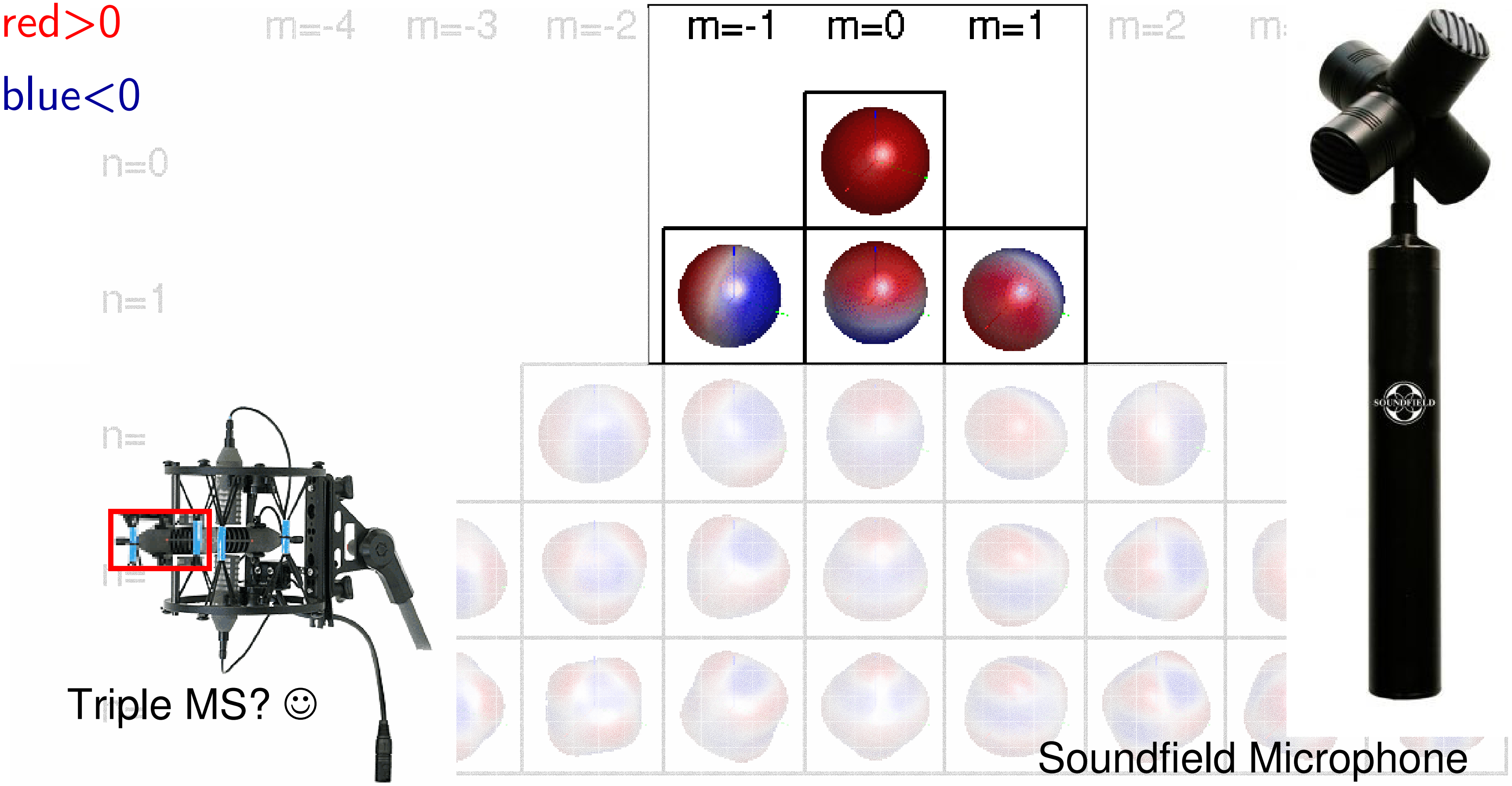


Basics of directional resoultion – surround (with height)

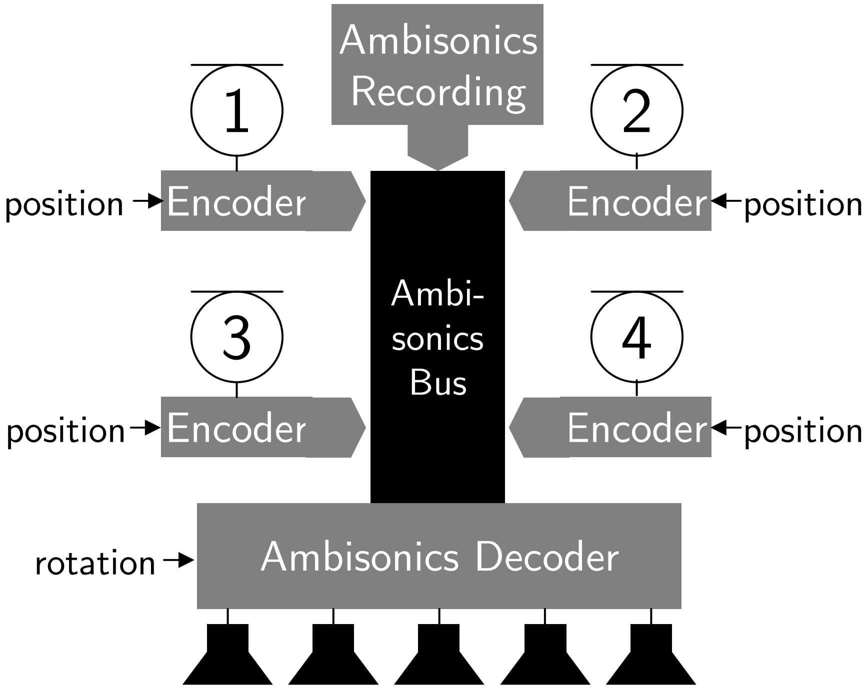
Smooth surround-with-height imaging:
harmonic representation/truncation

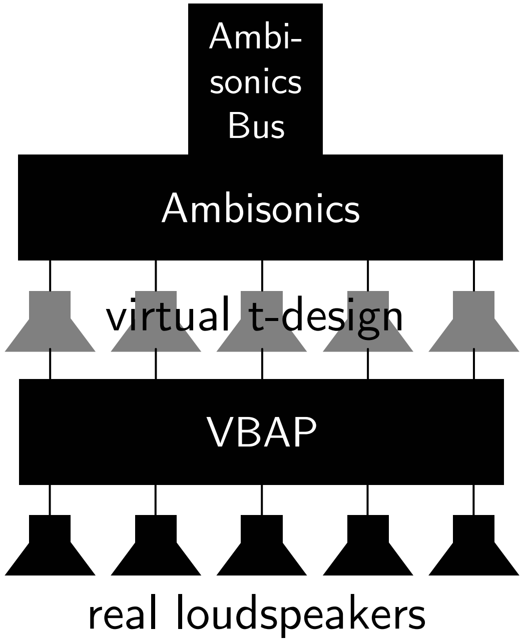
red>0

blue<0

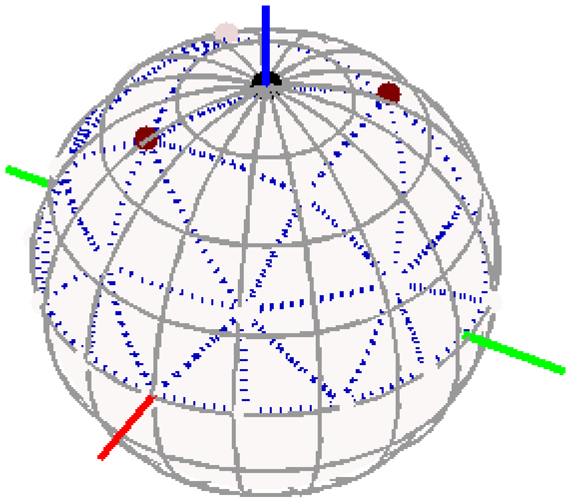


Ambisonics – recording and production format





Comparison: VBAP/pairwise panning



+ constant loudness
+ arbitrary layout
-- varying spread

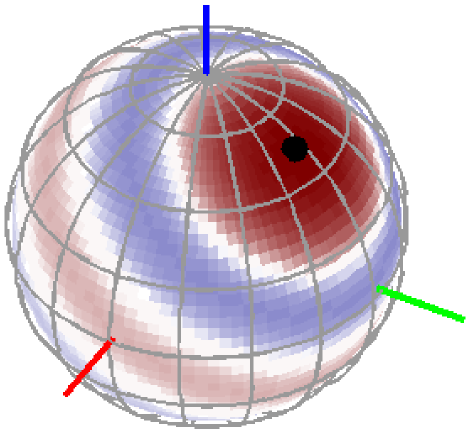
Ambisonics decoding – mode-matching approach

Similar desires as for pairwise/VBAP panning

- Additionally: constant source width

Ambisonics decoding – mode-matching approach

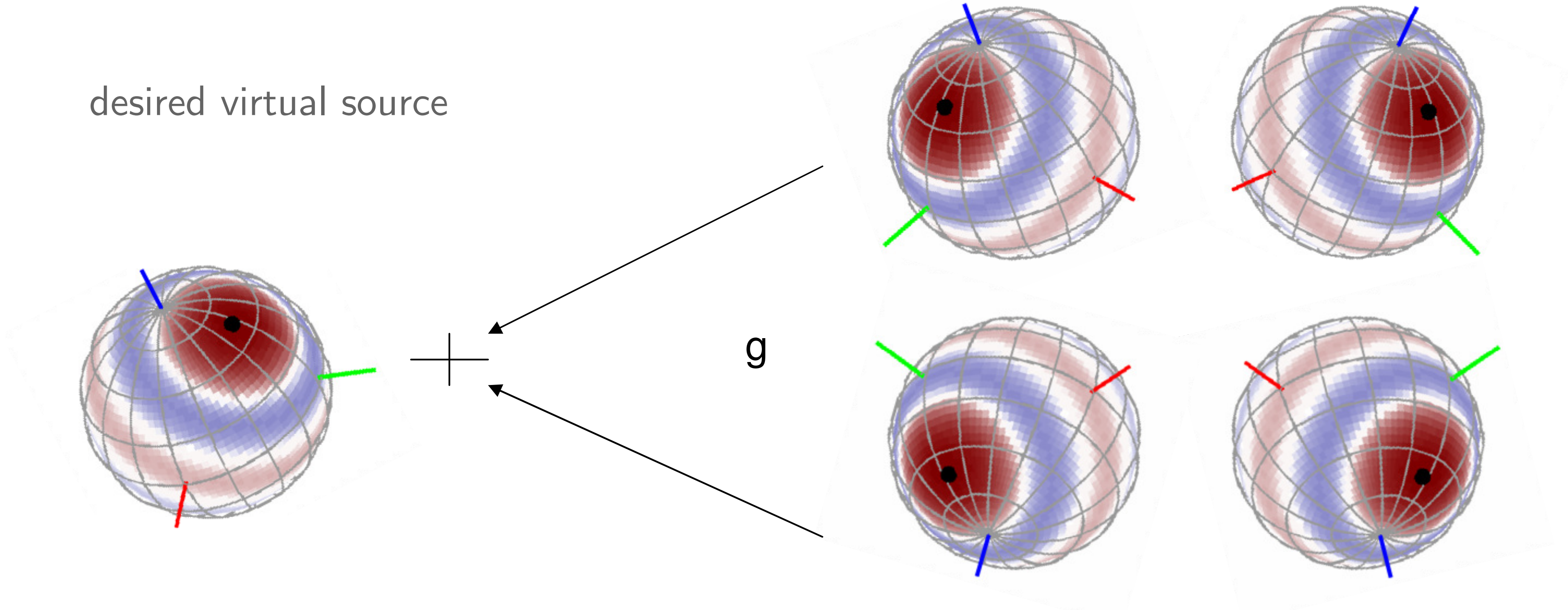
- Hypotheses:
  - localization spread can be controlled
  - assuming ideal limited representation of loudspeakers



$$\mathcal{B}_{\mathrm{N}}\{\delta(\boldsymbol{\theta}-\boldsymbol{\theta}_s)\} = \sum_{n=0}^{\mathrm{N}} \sum_{m=-n}^{n} Y_n^m(\boldsymbol{\theta}) Y_n^m(\boldsymbol{\theta}_s)$$

Ambisonics decoding – mode-matching approach

desired virtual source

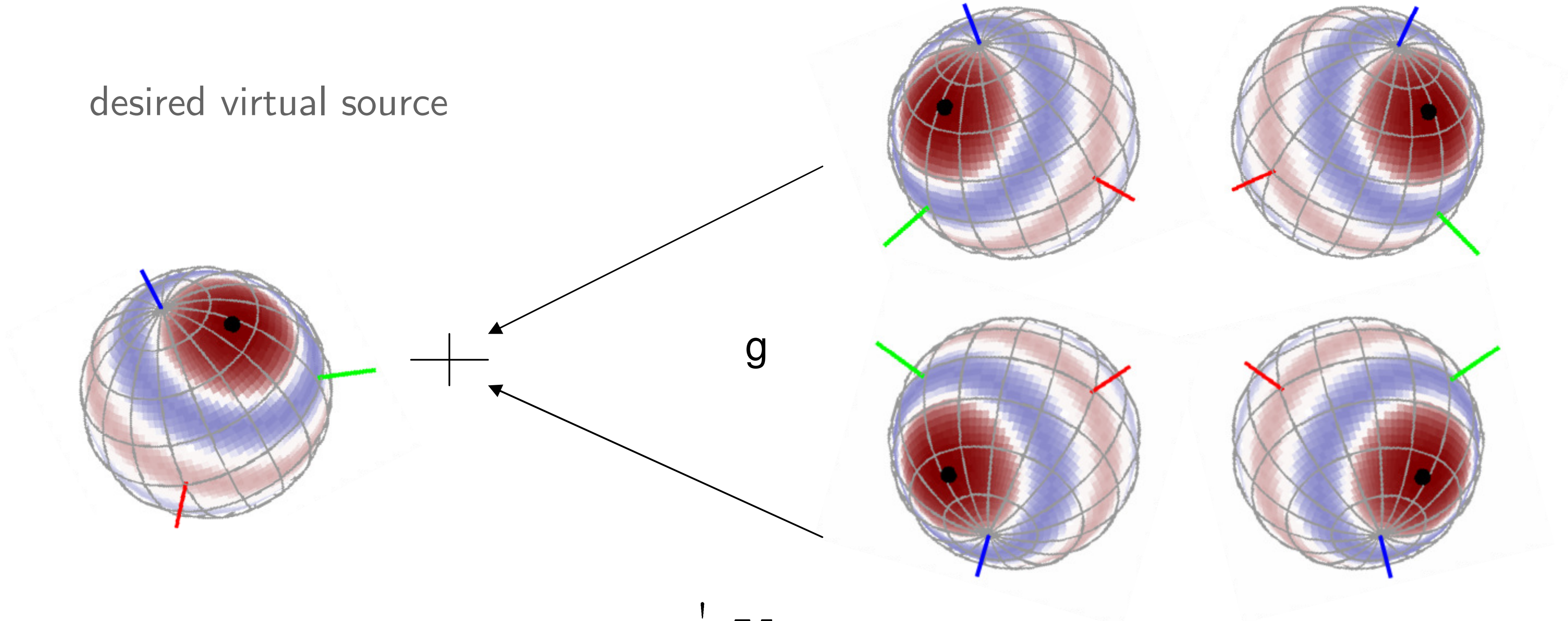


given loudspeakers

$$Y_n^m(\boldsymbol{\theta}_s) \overset{!}{=} \sum_{l=0}^{\mathrm{L}} Y_n^m(\boldsymbol{\theta}_l)\; g_l$$

$$\forall 0 \geq n \geq \mathrm{N}, -n \leq m \leq n$$

Ambisonics decoding – mode-matching approach



given loudspeakers

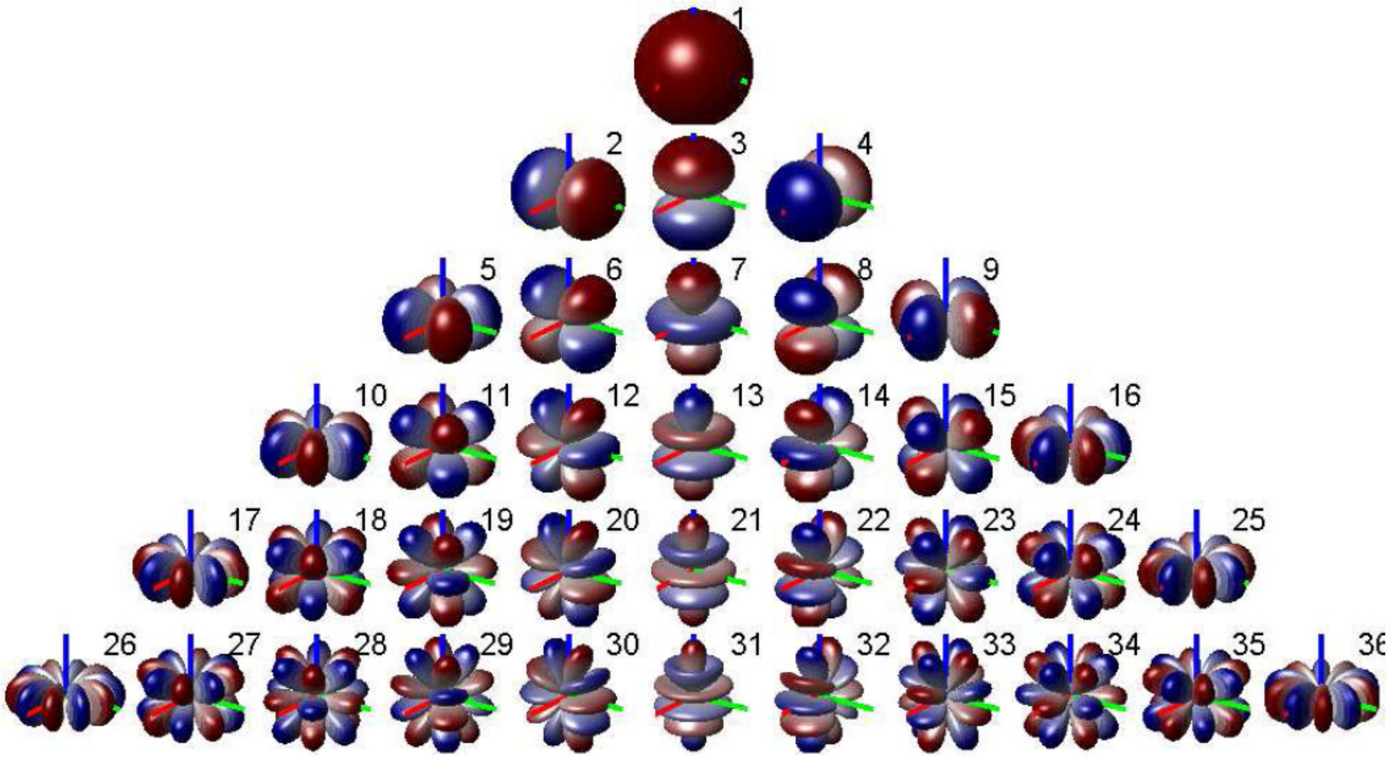
$$\boldsymbol{y}_{\mathrm{N}} \stackrel{!}{=} \boldsymbol{Y}_{\mathrm{N}}\,\boldsymbol{g}$$

$$\Rightarrow \boldsymbol{g} = \boldsymbol{Y}_{\mathrm{N}}^{-1}\,\boldsymbol{y}_{\mathrm{N}}(\boldsymbol{\theta})$$

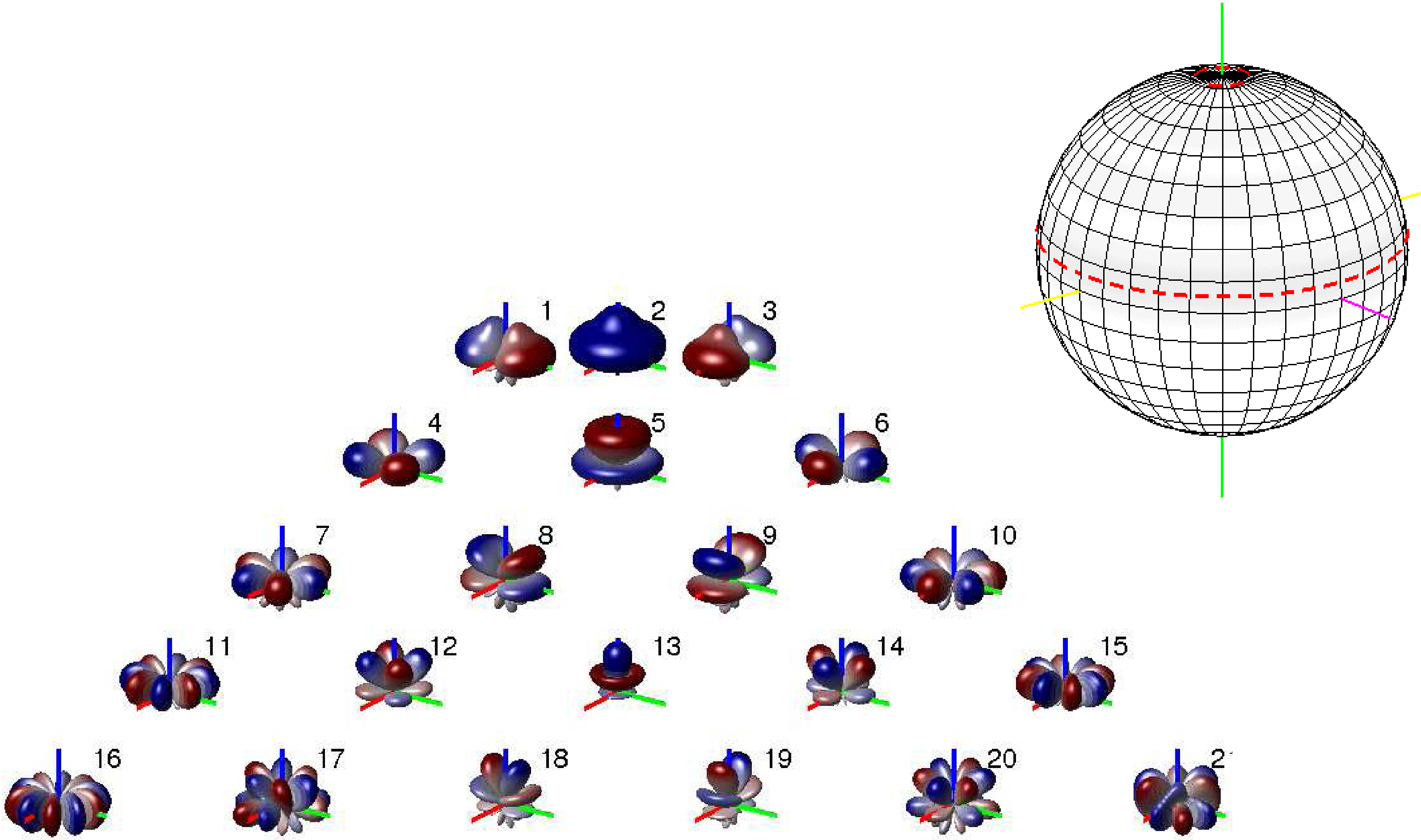
~- constant loudness
-- arbitrary layout
+ invariant spread

1) This inverse is hard to compute for arbitrary loudspeaker geometries
2) And still loudness may vary

Ambisonics – hemisphere decoding using orthogonalization



Ambisonics – hemisphere decoding using orthogonalization



Ambisonics with t-design loudspeaker layout

- spherical t-designs (particular arrangements)

$$\boldsymbol{Y}_{\mathrm{N}}^{-1} = \frac{4\pi}{L} \boldsymbol{Y}_{\mathrm{N}}^{\mathrm{T}}$$

$$\Rightarrow \boldsymbol{g} = \frac{4\pi}{L} \boldsymbol{Y}_{\mathrm{N}}^{\mathrm{T}} \boldsymbol{y}_{\mathrm{N}}(\boldsymbol{\theta}) \qquad \|\boldsymbol{g}\|^2 = \mathrm{const}$$

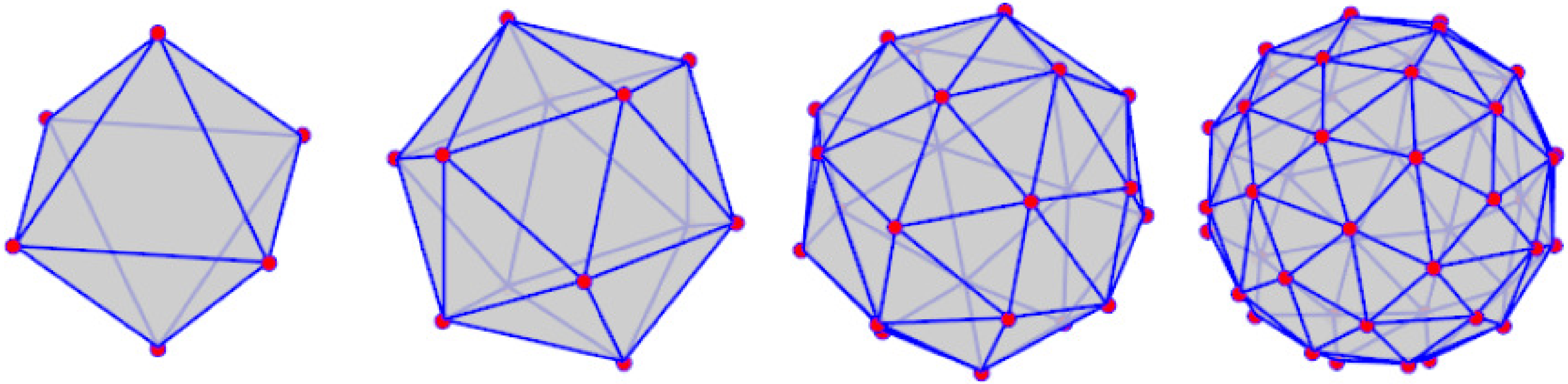
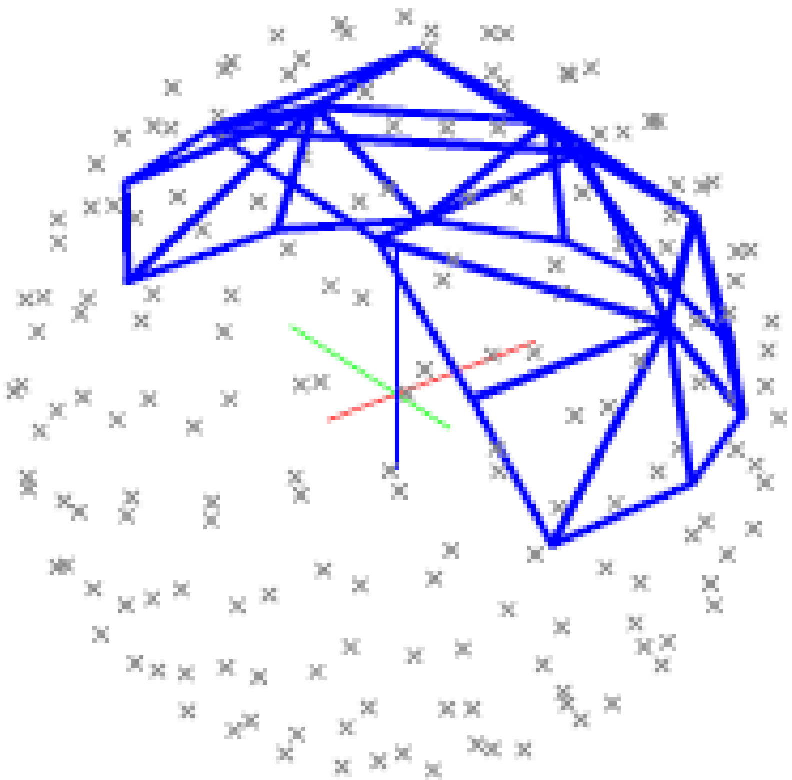


Figure 4: t-designs: $t = 3, 5, 7, 9$

\+ constant loudness
-- very restricted layout
\+ invariant spread

Virtual t-design Ambisonics-Rig Using VBAP

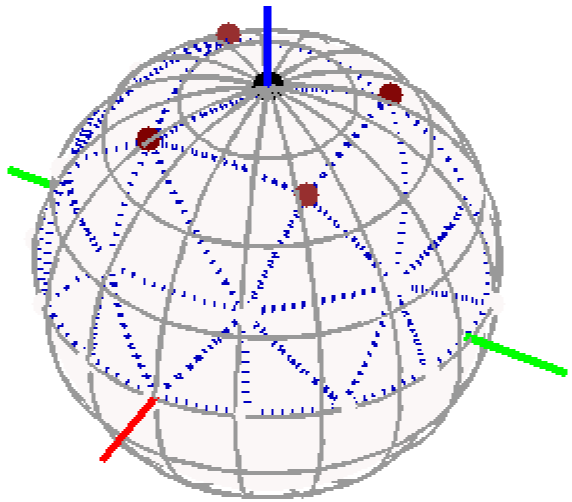
cp. also Batke and Keiler 2010



(a) Virtual 18-design.

HOPE?
+ constant loudness
+ arbitrarz layouts
+ invariant spread

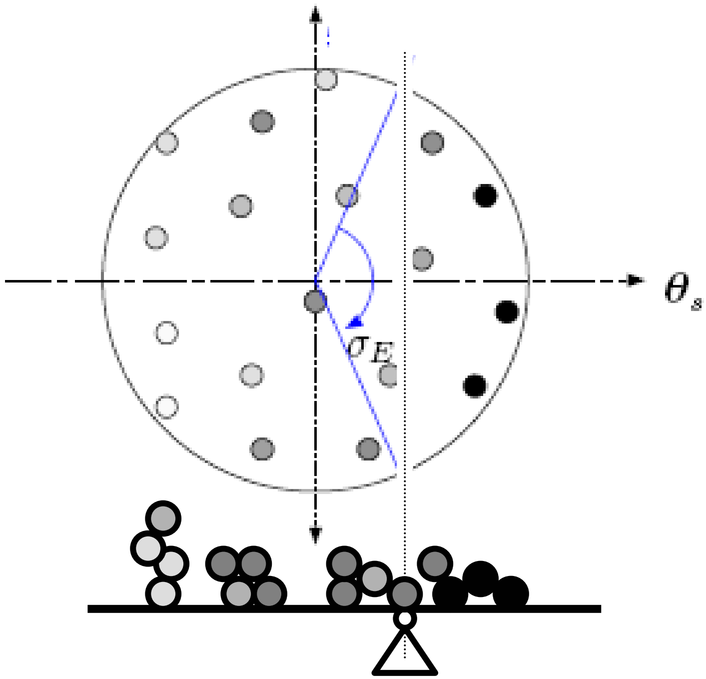
T-design Ambisonics using VBAP



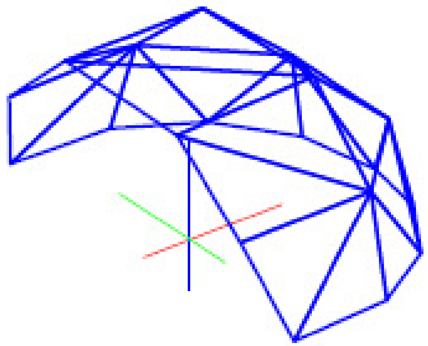
~+ constant loudness
+ arbitrary layout
~+ invariant spread

Measures

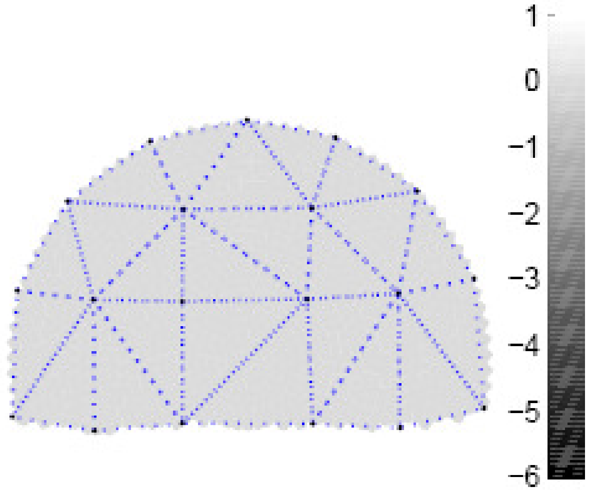
- Loudness Measure: Energy
  - average in space: statistically independent superposition of ls signals at mid and high frequencies
- Spread: based on *energy-vector* (center of mass along source direction)



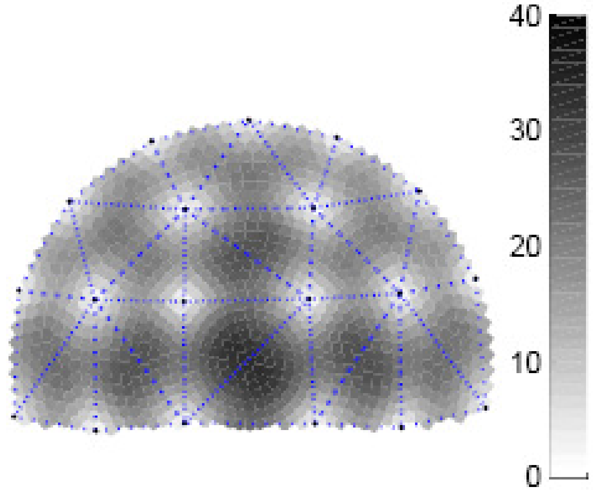
VBAP vs Virtual T-Dsgn Ambi Using VBAP



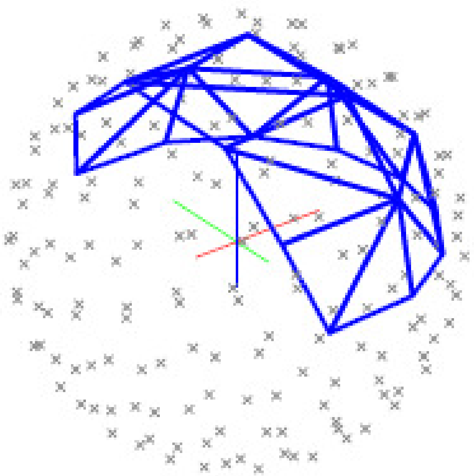
(a) Triangulation.



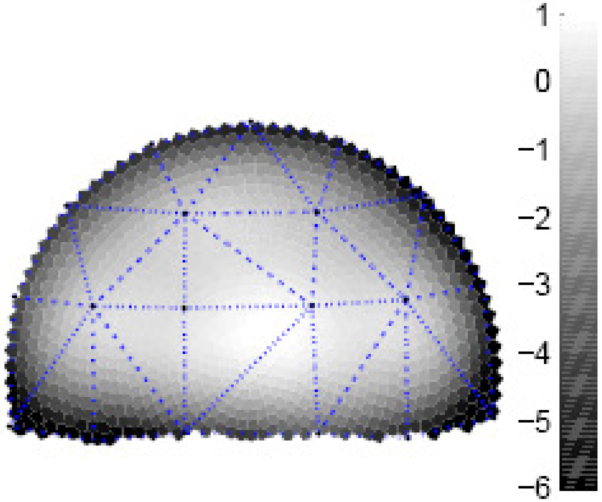
(b) $E = 0[\mathrm{dB}]$.



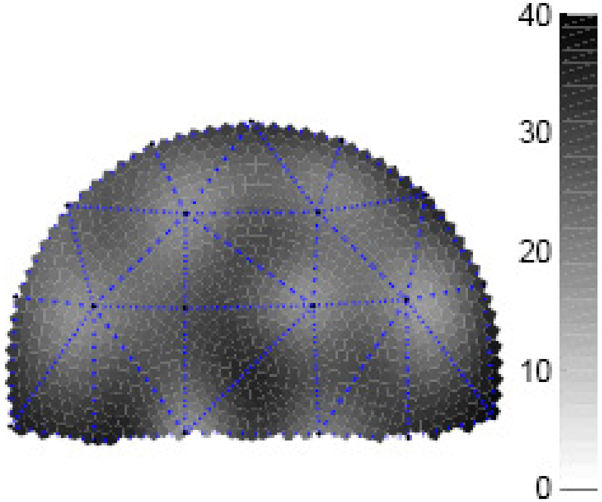
(c) $\sigma_{\mathrm{E}} = 0 \ldots 33[^\circ]$.



(a) Virtual 18-design.



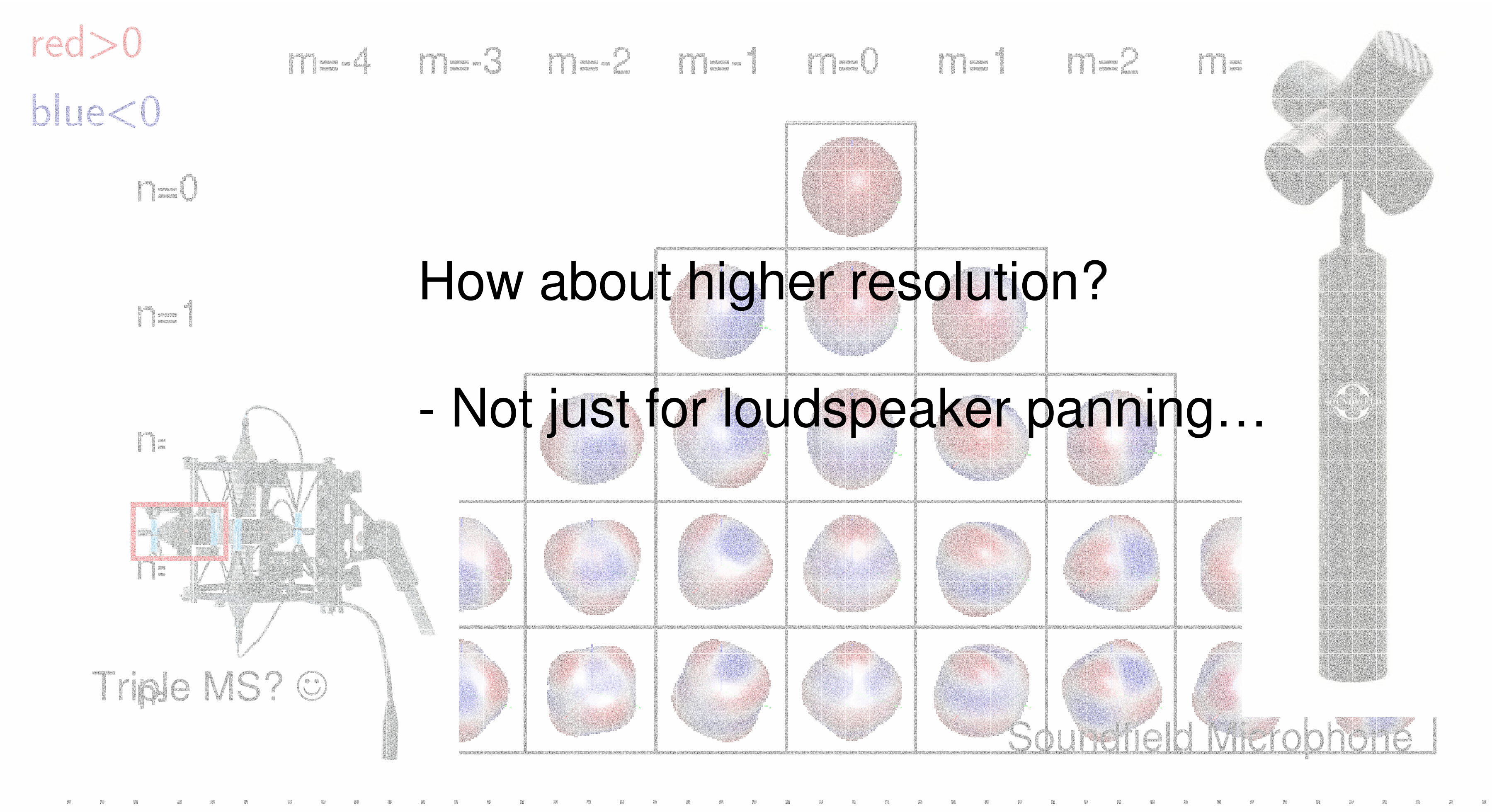
(b) $E = -6 \ldots 1[\mathrm{dB}]$.



(c) $\sigma_{\mathrm{E}} = 25 \pm 10[^\circ]$.

Basics of directional resoultion – surround (with height)

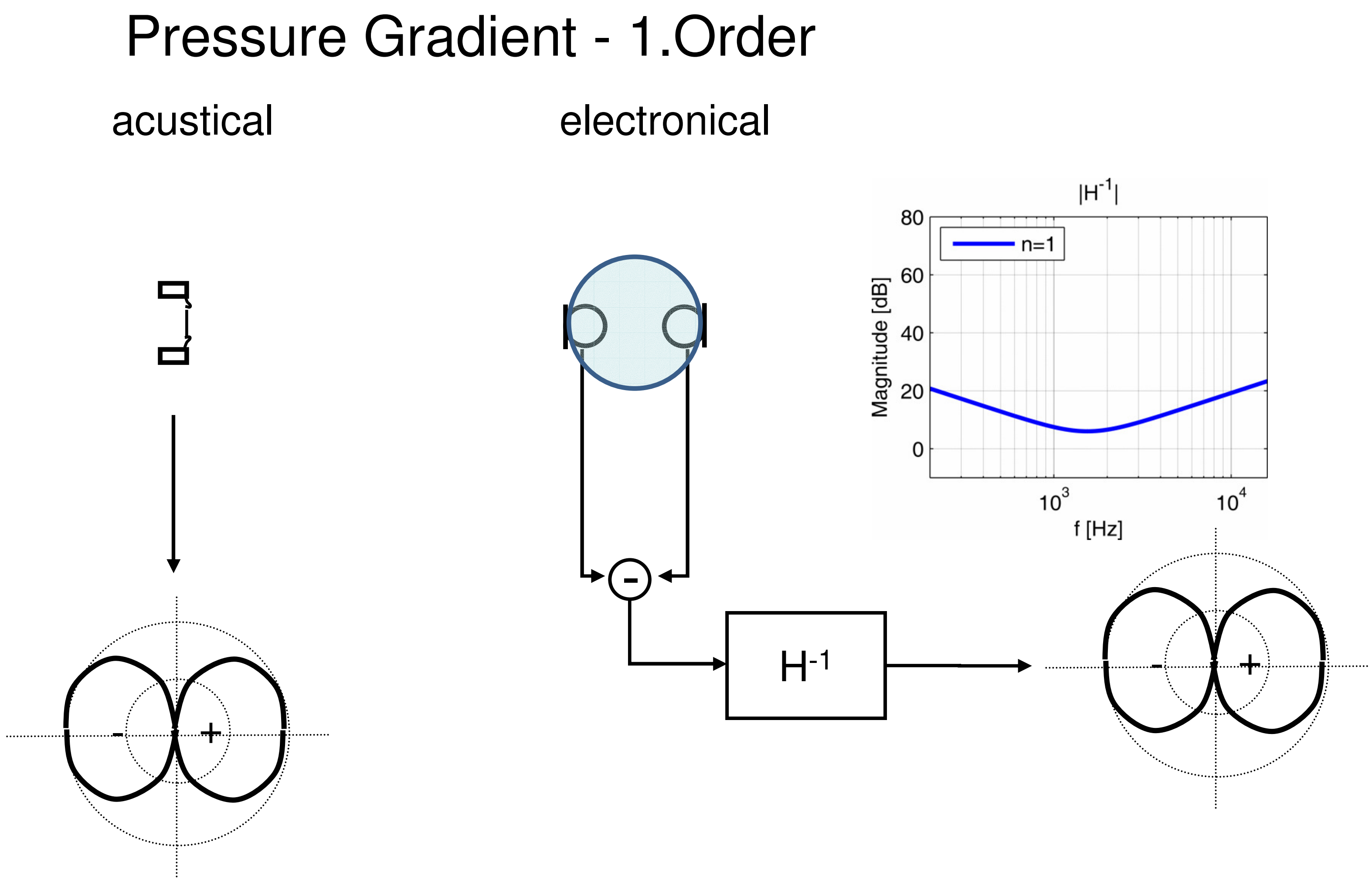
Smooth surround-with-height imaging:
harmonic representation/truncation



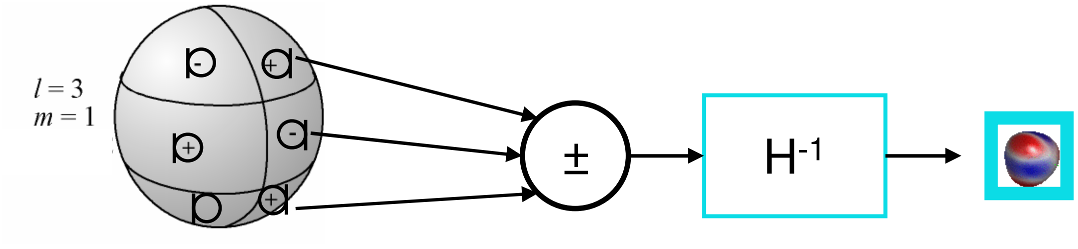
How about higher resolution?

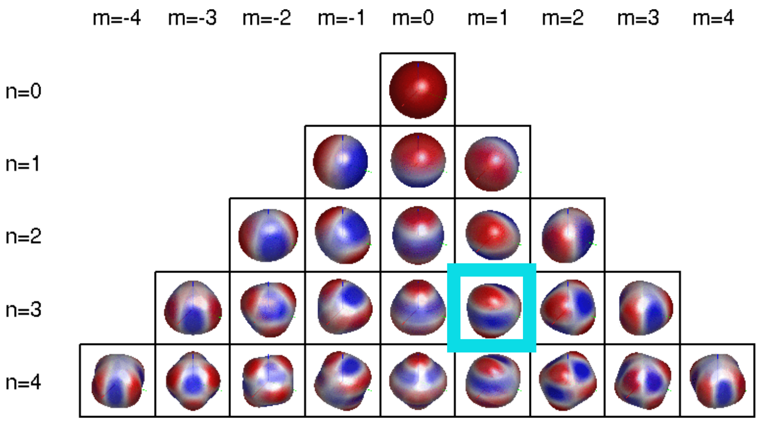
- Not just for loudspeaker panning…

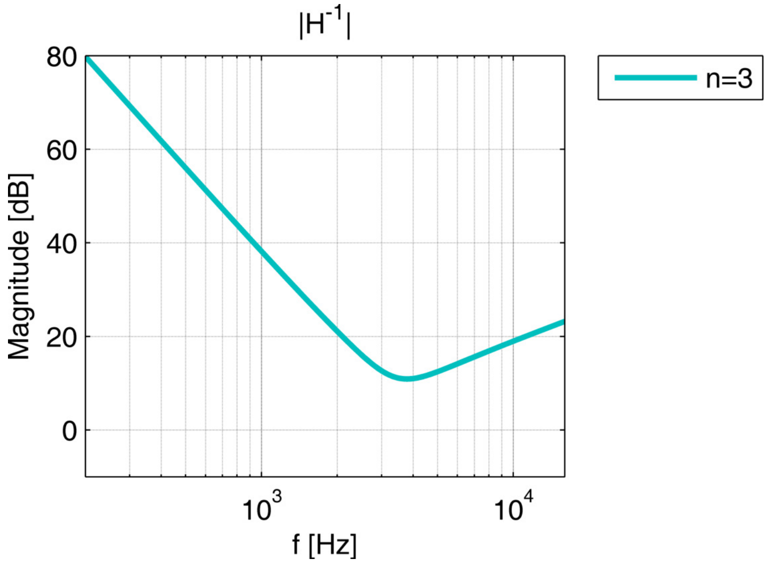
Higher-order Microphones



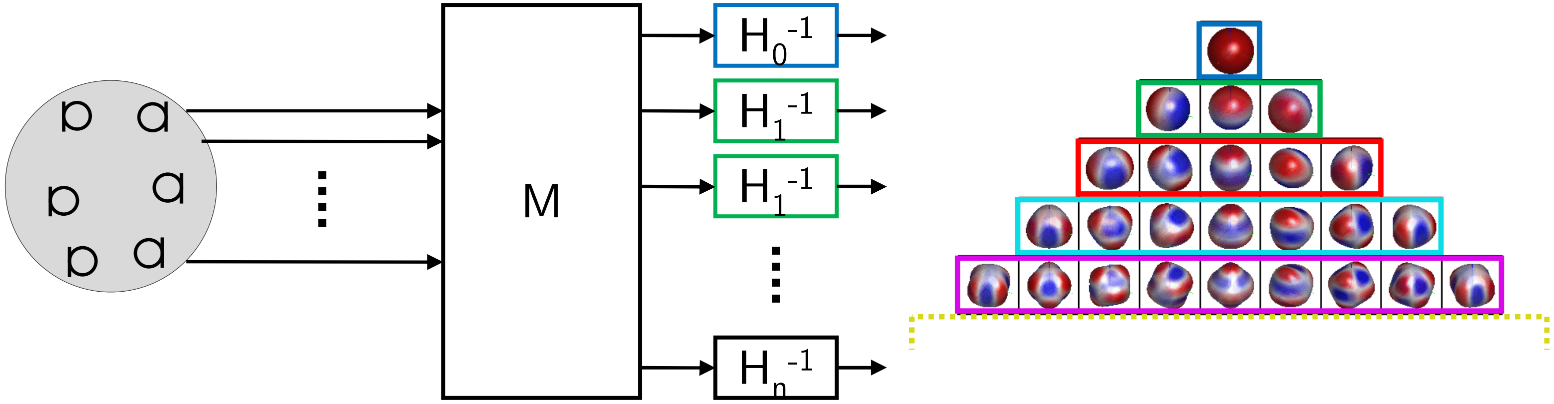
Higher-order Microphones



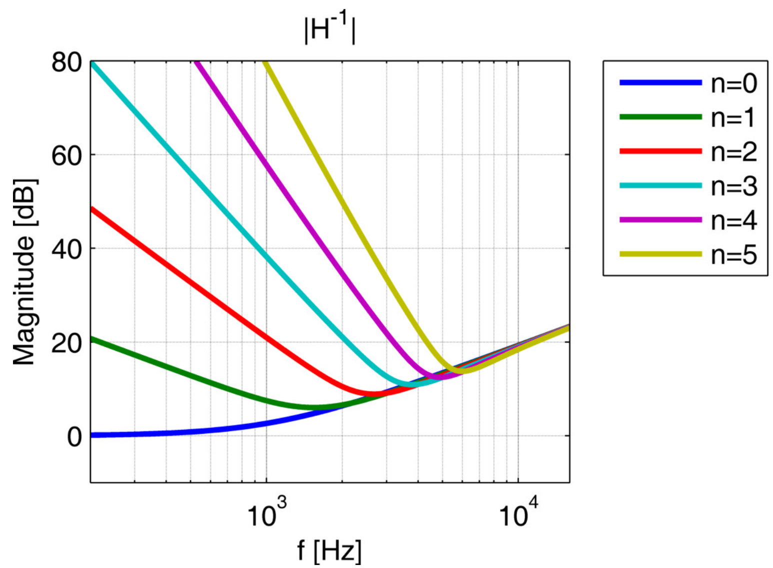




Higher-order Microphones

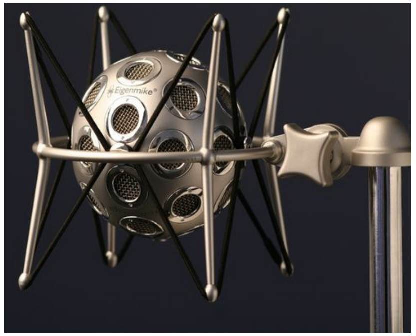


Holographic filters:

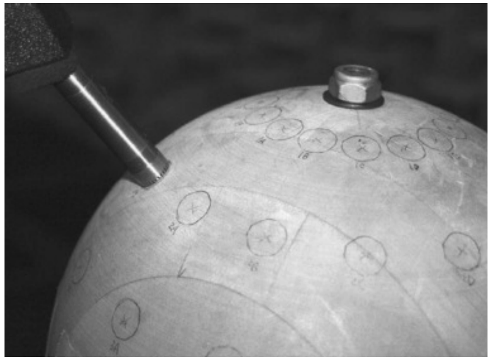


Higher order microphone arrays

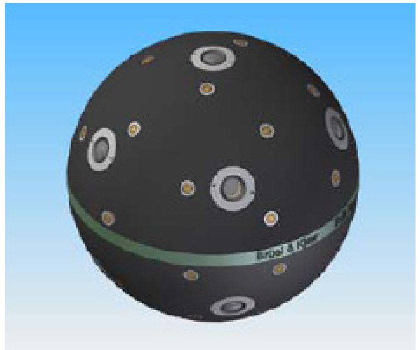
Meyer, Elko
2000-2010



Rafaely&Park,
2004/2010

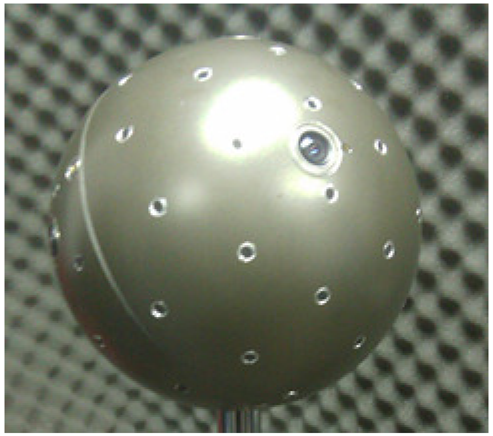


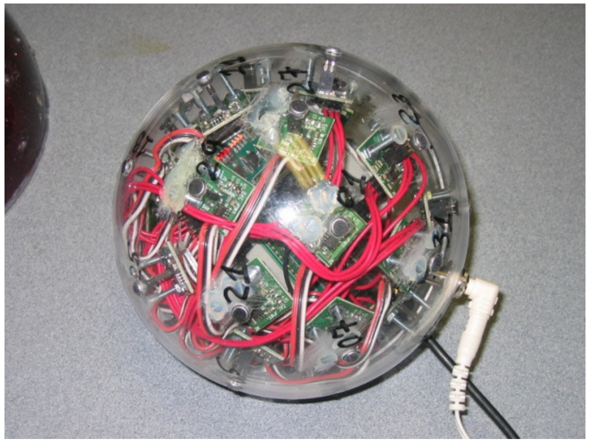
Petersen, Moller-Juhl,
B&K, 2003-2010



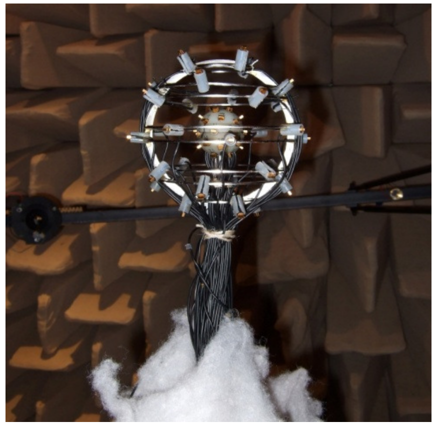
Takashima, Nakagawa,
Williams,2008

Li, Duraiswami, O‘Donavan, Grassi,
2004-2010

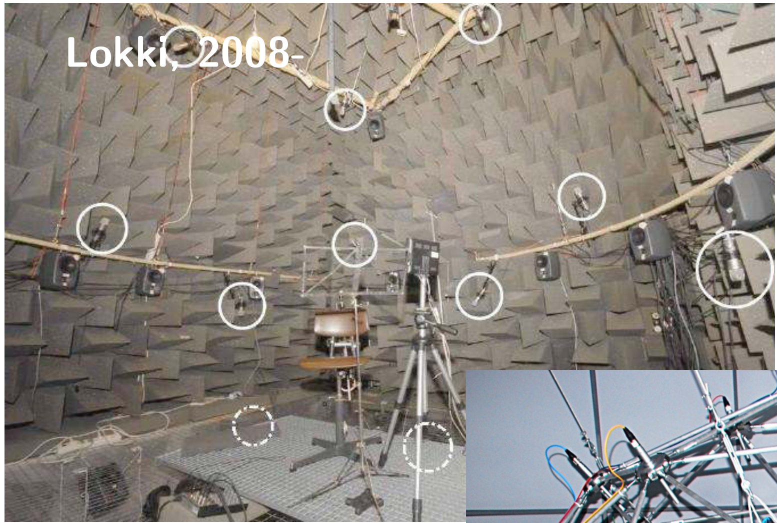


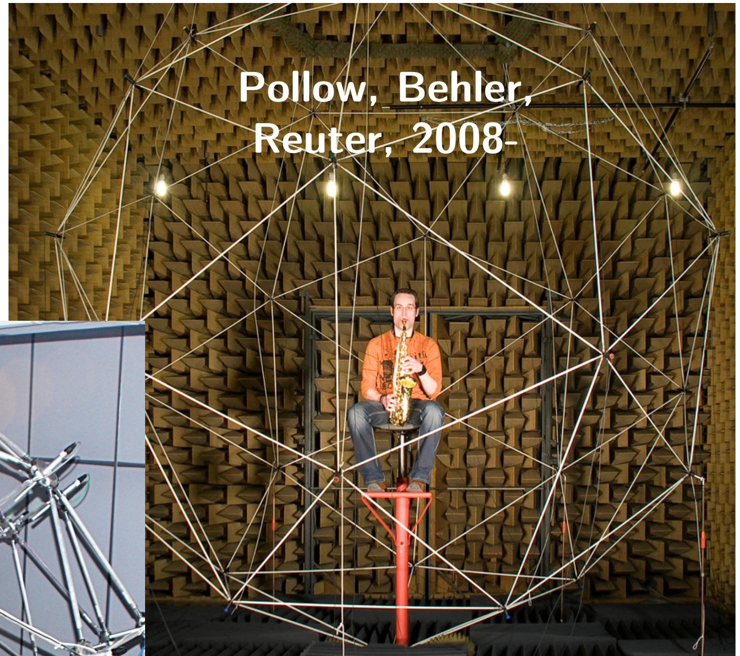


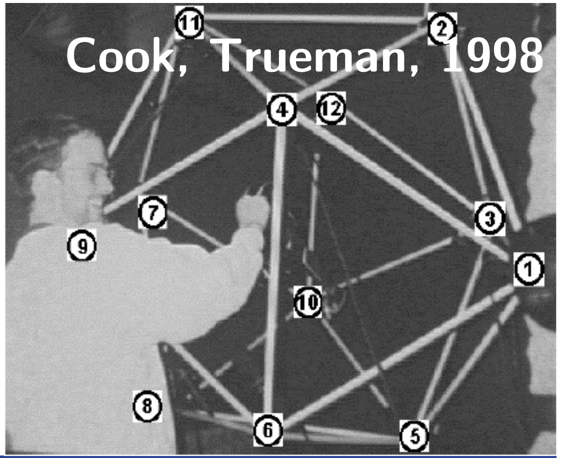
Jin, v.Schaik,
2006-2010



Sound radiation: surrounding spherical microphone arrays



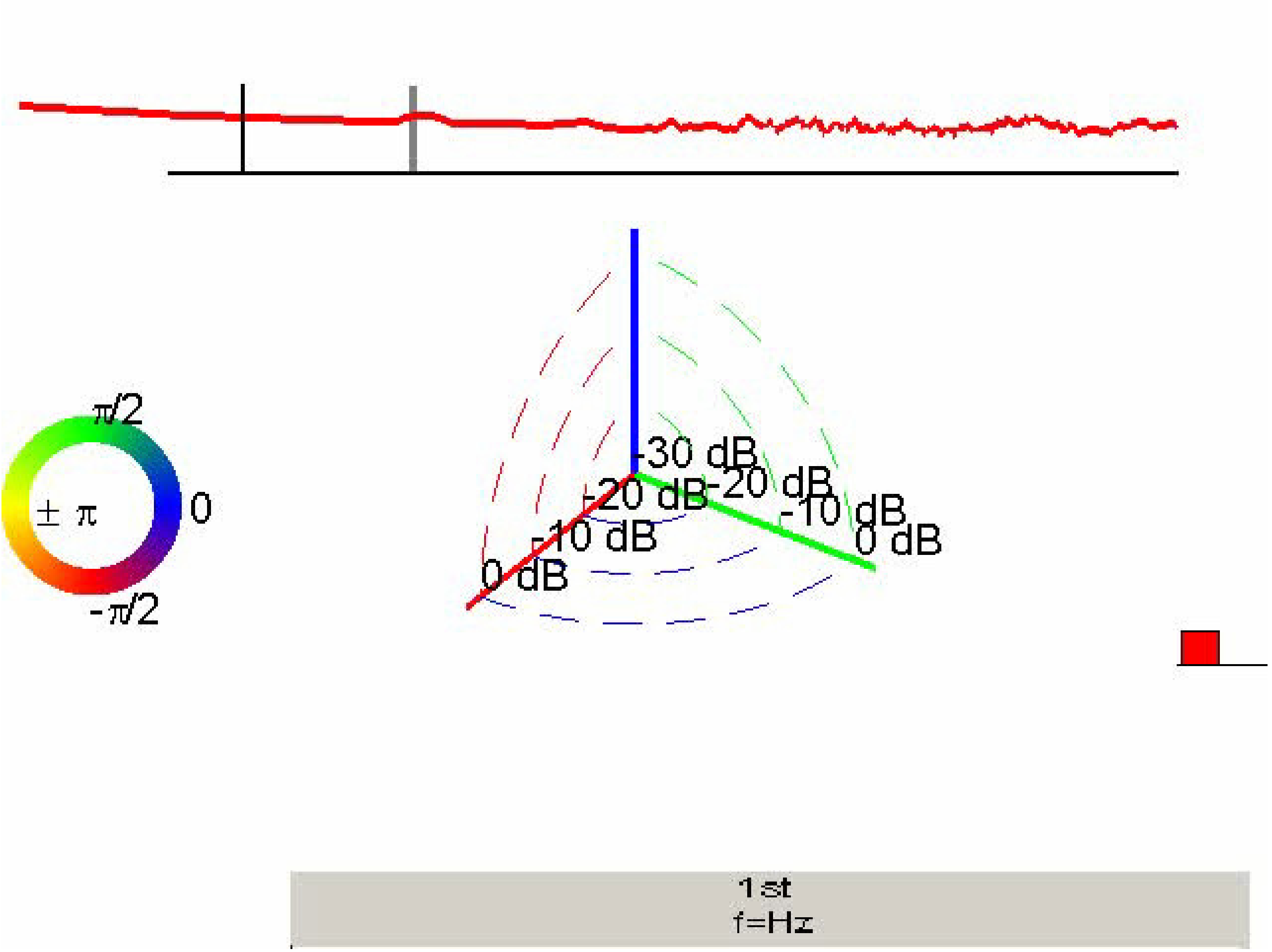




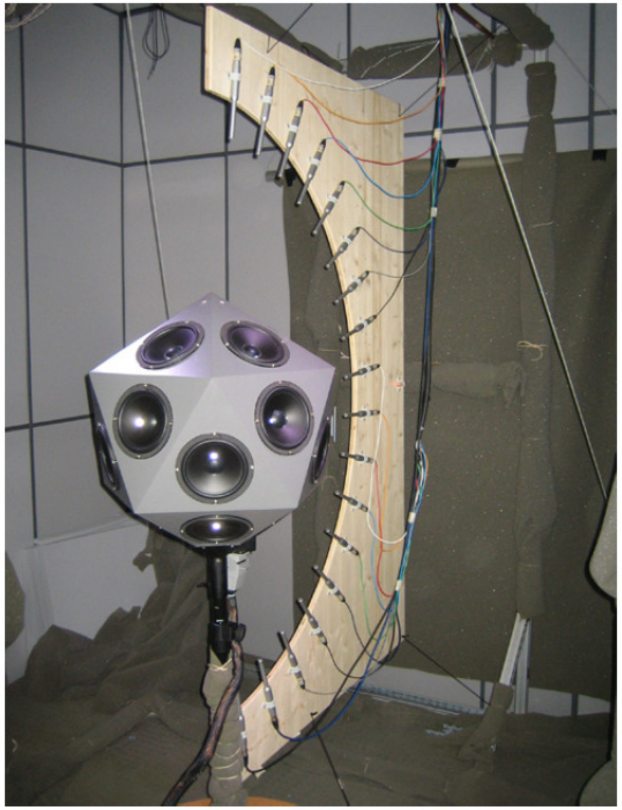
Analysis of Radiation Patterns

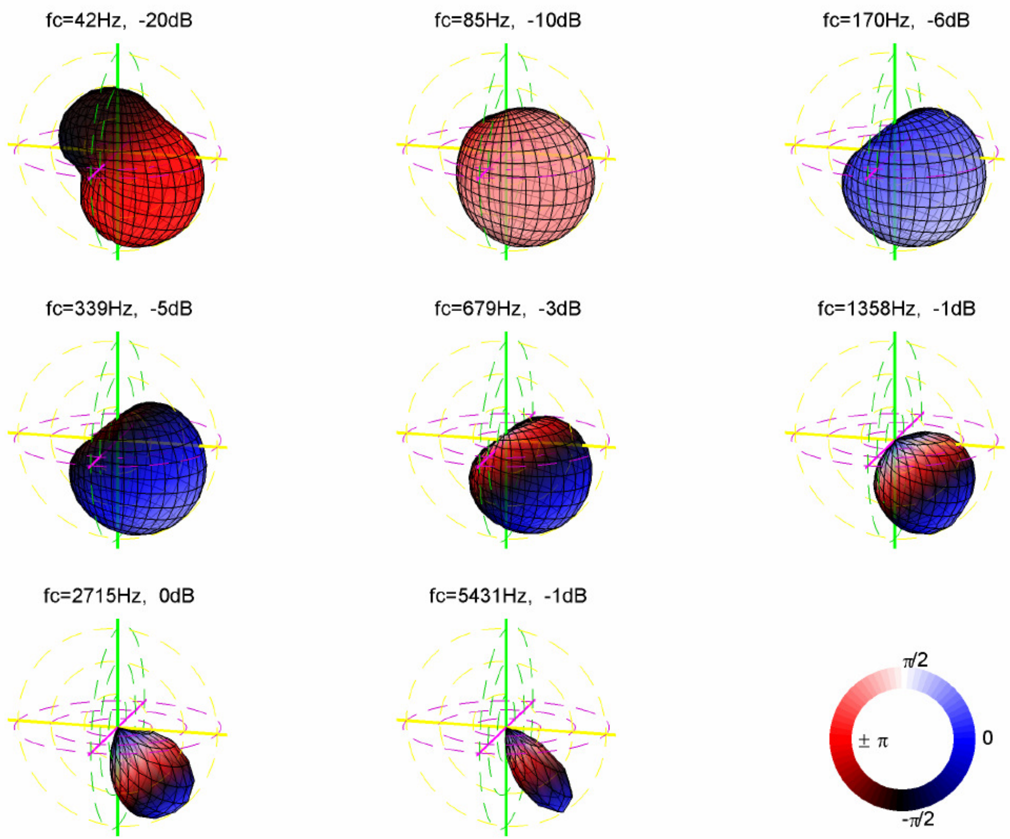
- **example:** Saxophone

First partial

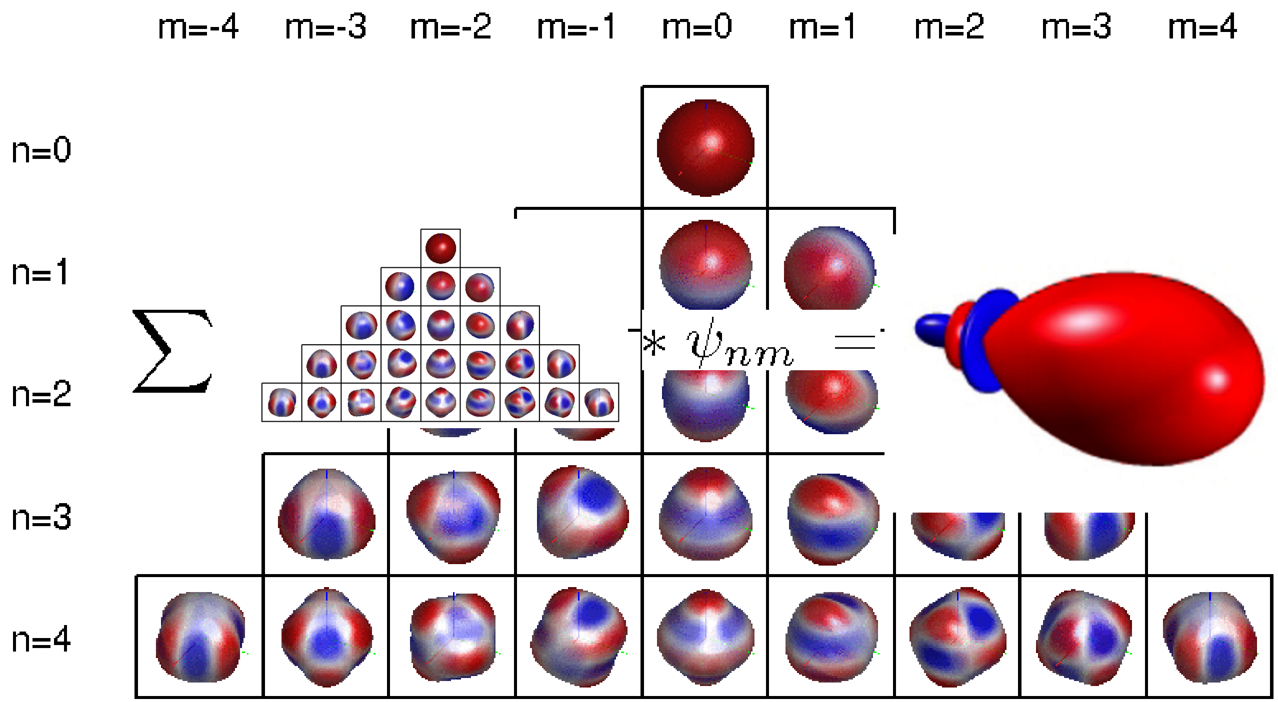


Compact Spherical Loudspeaker Array



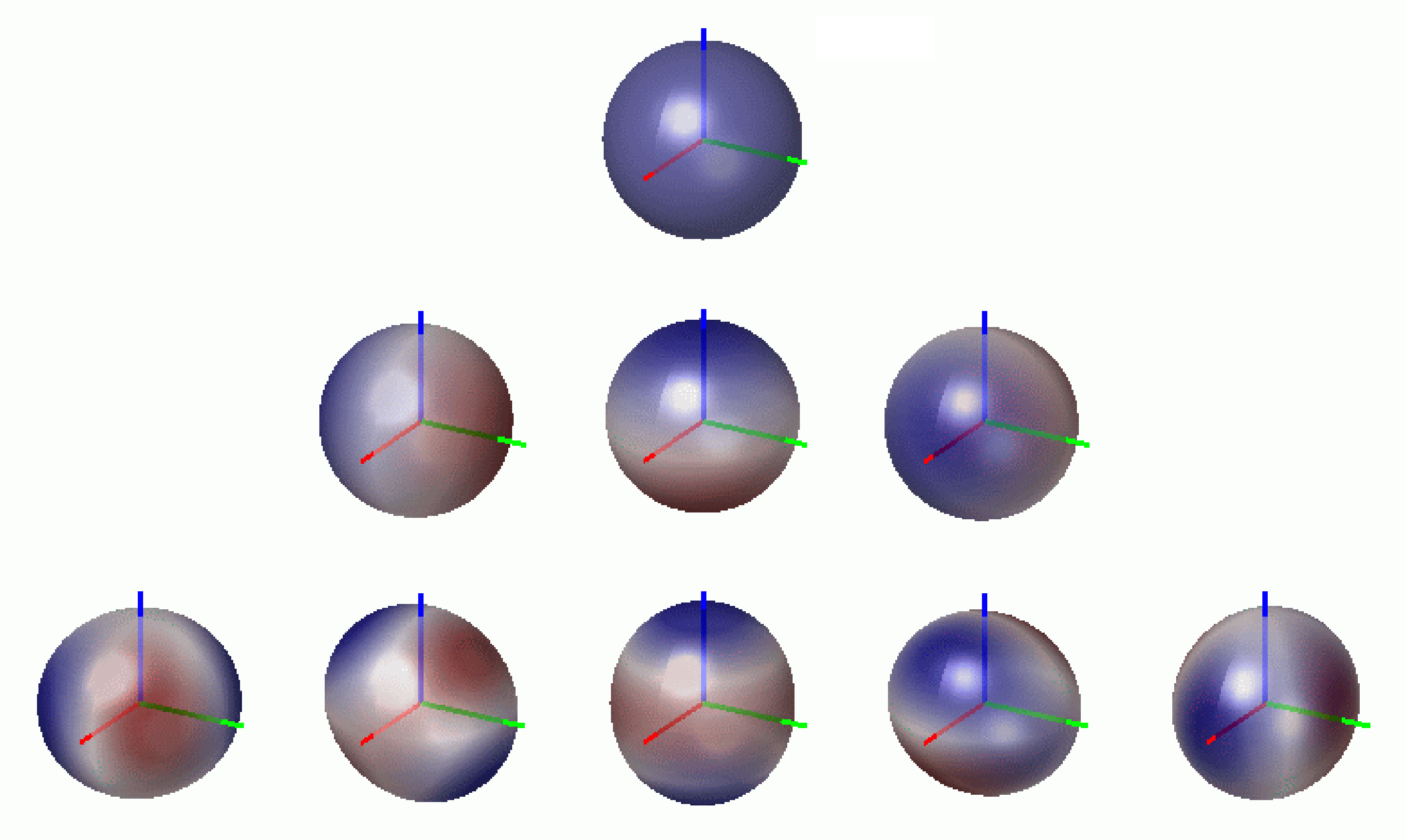


Spherical Harmonics

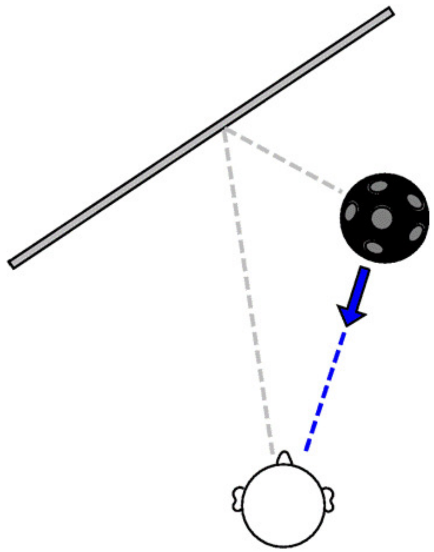


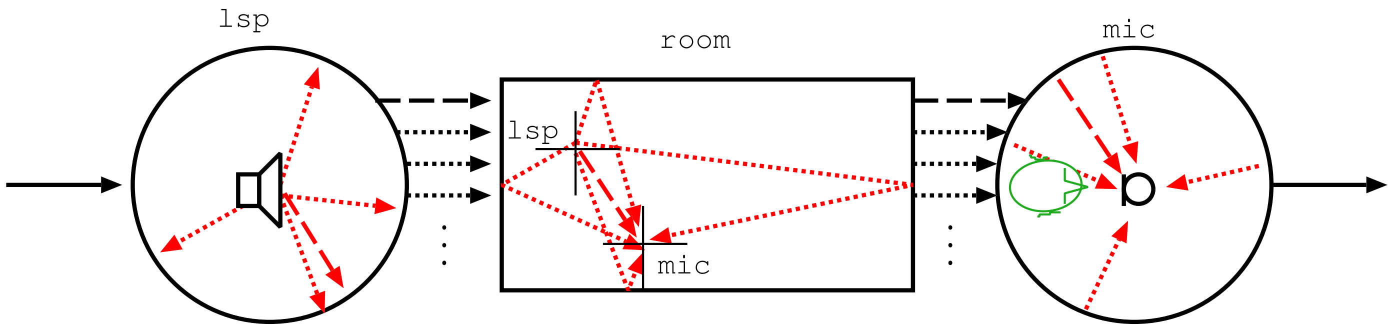
$$Y_n^m(\varphi,\vartheta) = N_n^m P_n^m(\cos(\vartheta)) \begin{cases} \sin(m\varphi), & \text{for } m < 0 \\ \cos(m\varphi), & \text{for } m \geq 0 \end{cases}$$

Spherical Harmonics



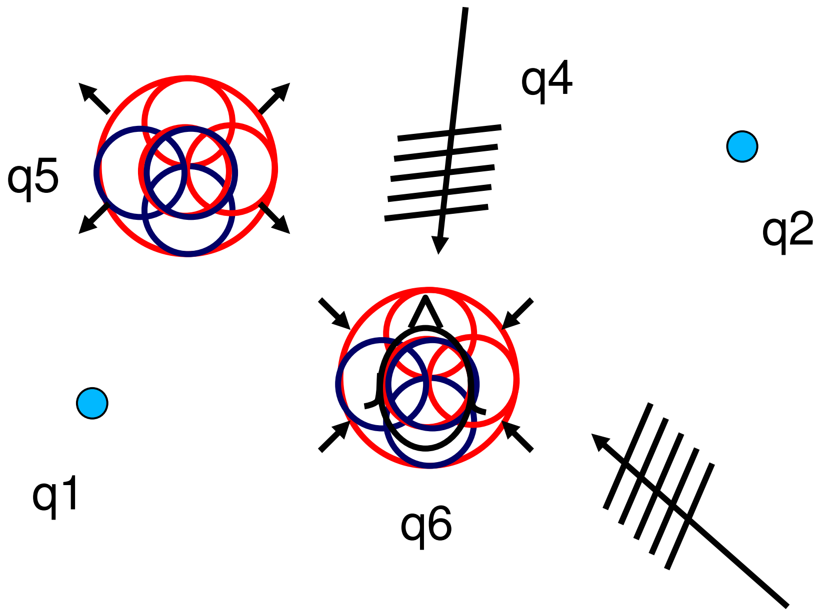
Applications

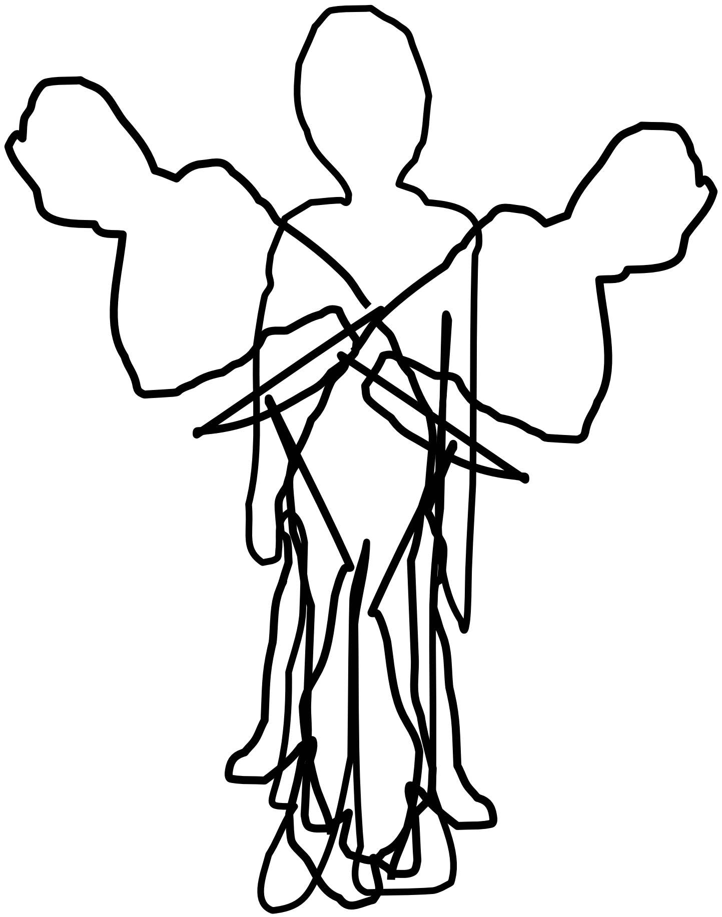




Ambisonics as a source/surround representation format

Object based mixing (for whatever renderer)





Vielen Dank!

Mitwirkende:

Hannes Pomberger
Matthias Frank
Georgios Marentakis