

# Towards usability of Higher Order Ambisonics in Digital Audio Workstations

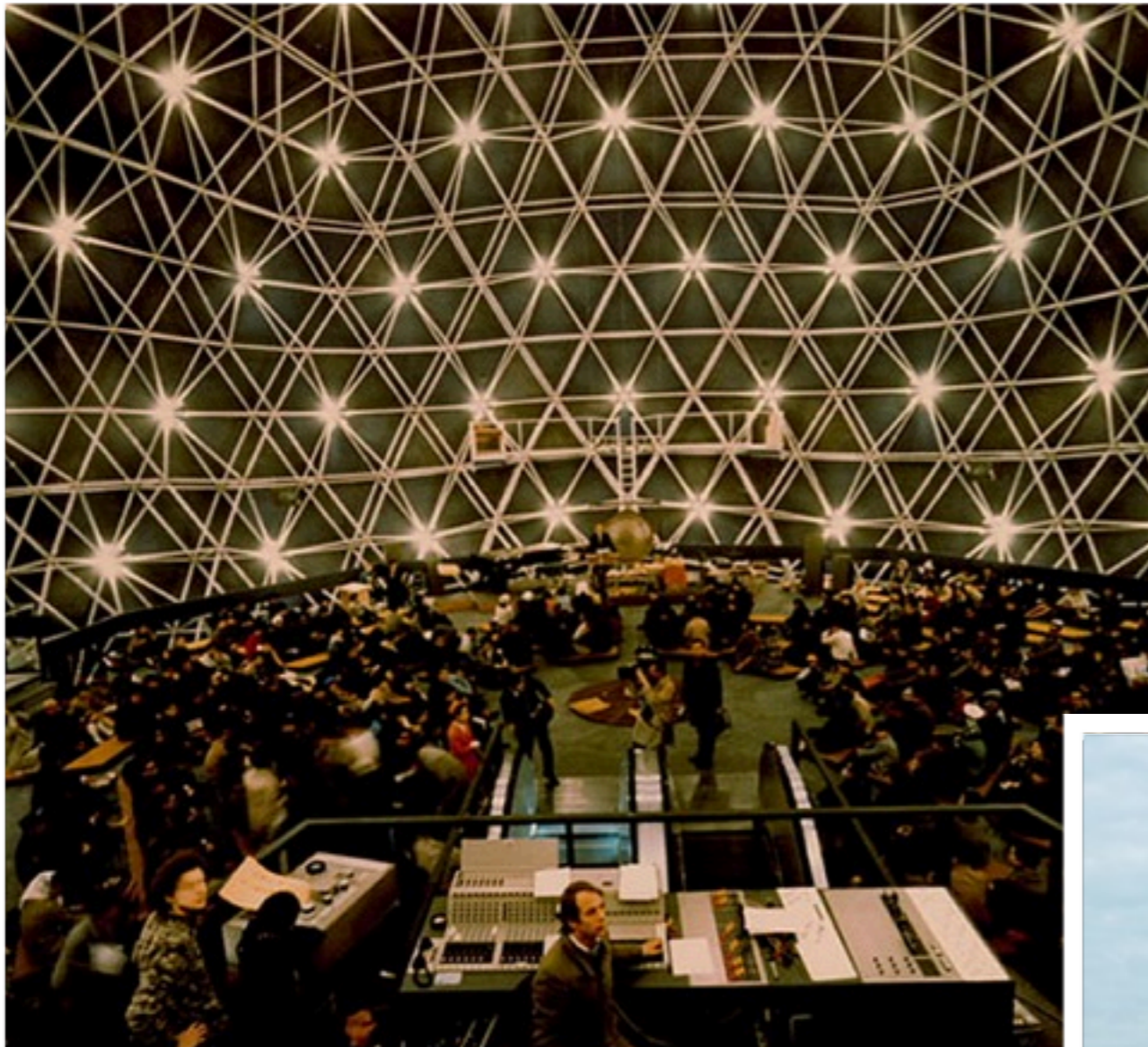
Matthias Kronlachner

Institute of Electronic Music and Acoustics  
University of Music and Performing Arts, Graz



2nd International Conference  
MUSIC AND TECHNOLOGIES  
Kaunas University of Technology,  
8-10 November 2012, Kaunas, Lithuania





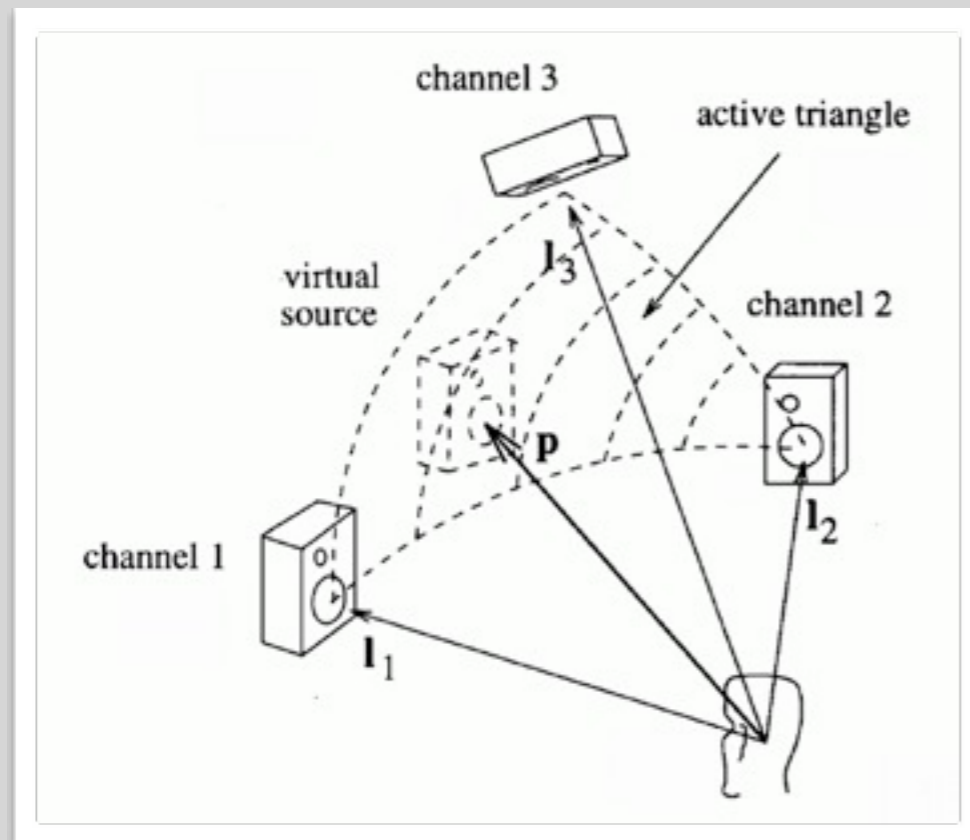
Pischelsdorf (AT), 2011-??

# What is Ambisonic(s)?

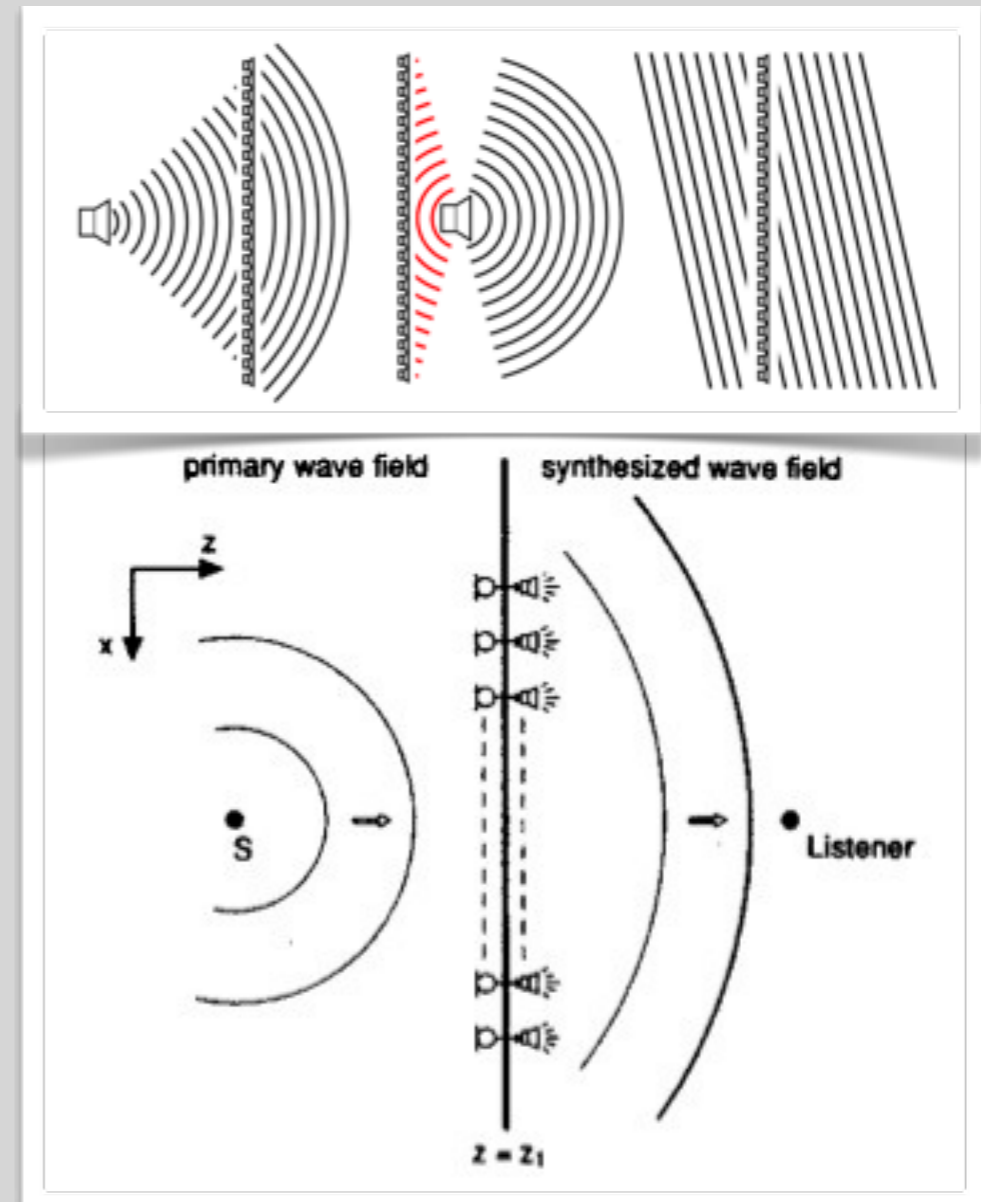
- Surround recording and playback technique, Michael Gerzon 1970s
- Little commercial success so far, patents expired
- Independent of playback loudspeaker configuration, scaleable
- 2D and 3D sound-fields can be synthesized

# What is Ambisonic(s)?

- Not using phantom sources (VBAP), but trying to recreate original sound-field (WFS)



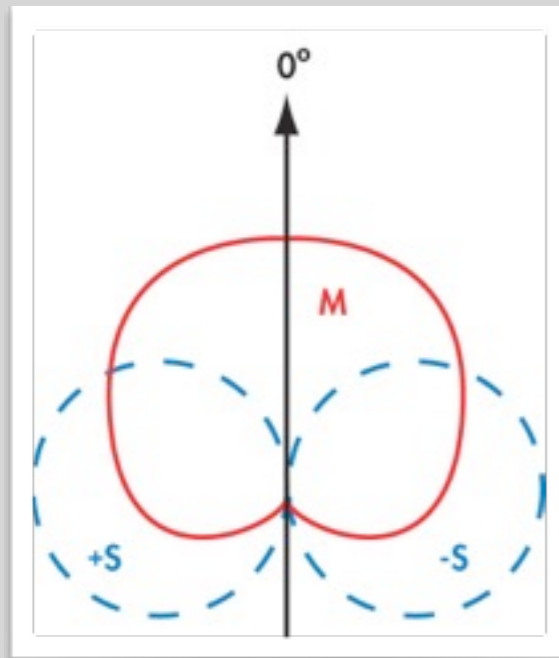
Vector Based Amplitude Panning (VBAP)



Wave Field Synthesis (WFS)

# What is Ambisonic(s)?

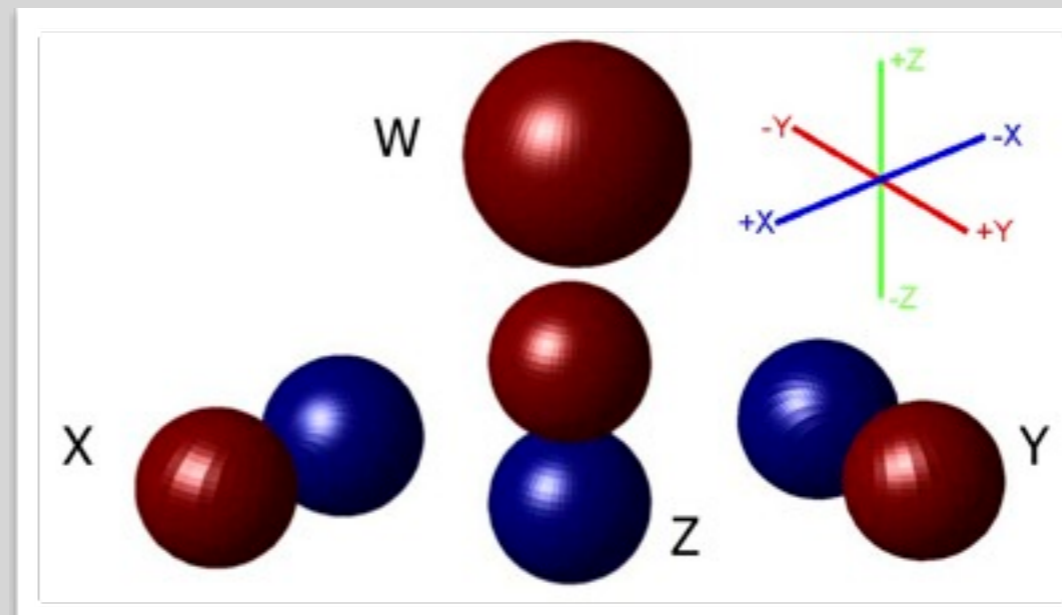
- Comparable to M/S stereo recording technique
- A sound-field is “sampled” using spherical-harmonics



## MS Stereo recording

Left =  $M+S$

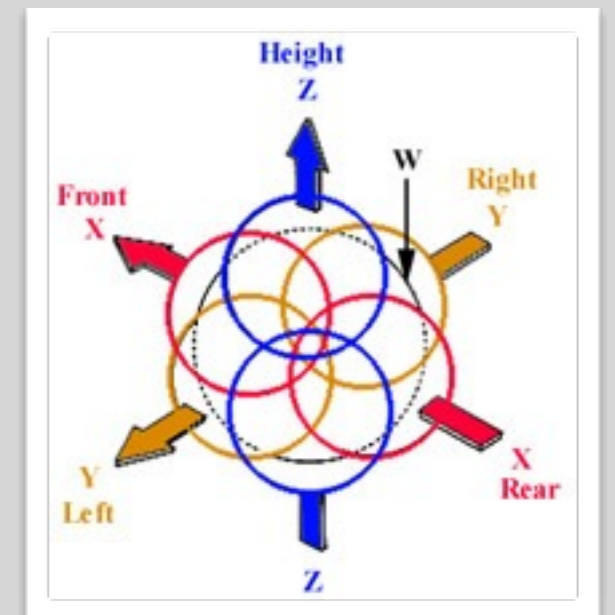
Right =  $M-S$



## First Order Ambisonics

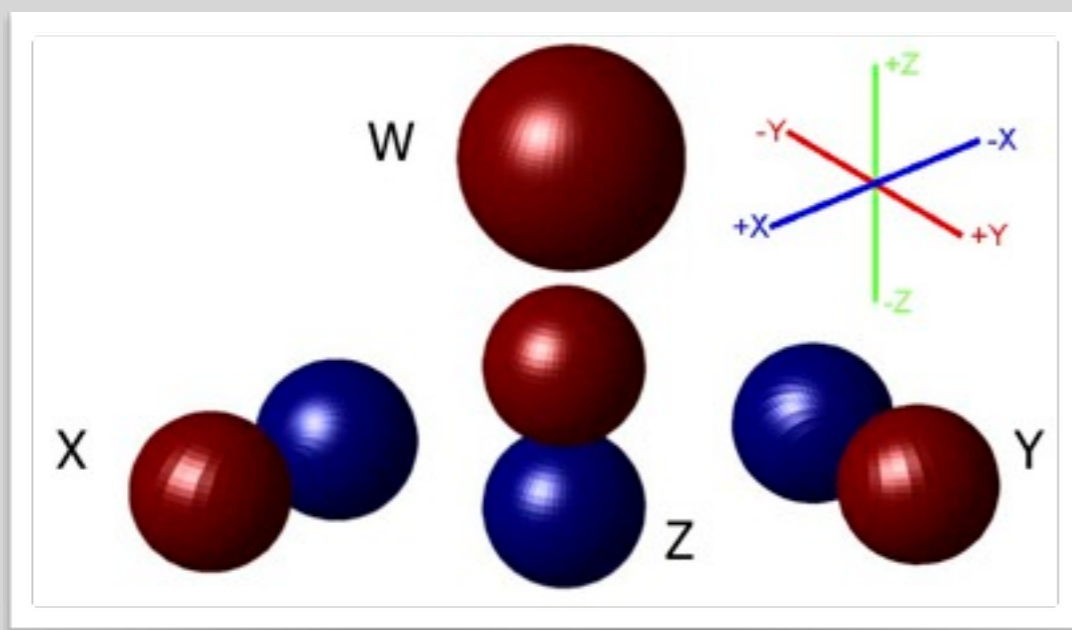
W (omni directional)

X, Y, Z (figure of eight)

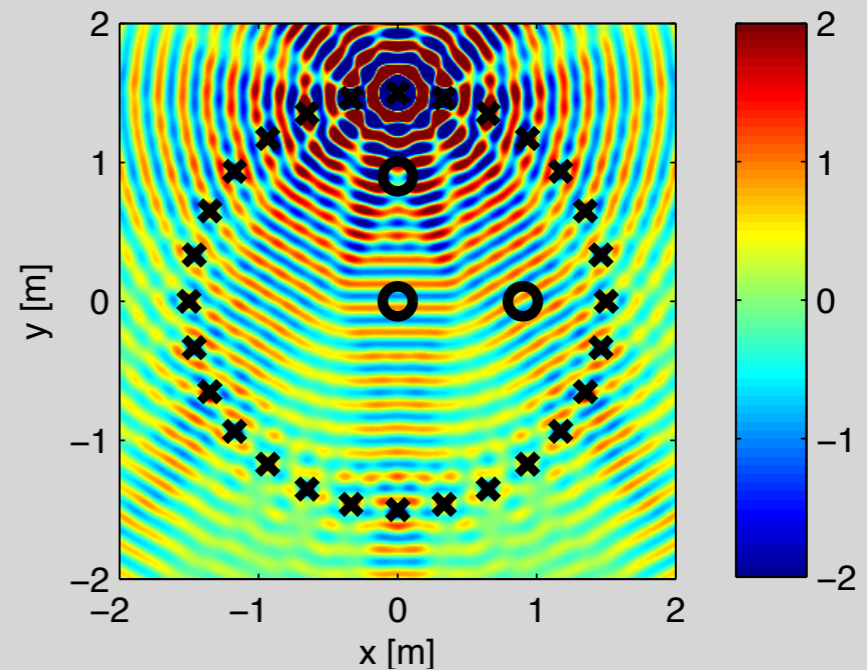


# What is Ambisonic(s)?

- 4 channel signal (W, X, Y, Z) is called B-Format
- B-Format does not represent a speaker feed neither a microphone signal
- trying to recreate original sound-field in a “sweet-spot”



B-Format (1st order Ambisonics)



(h) HOA 28 2000Hz

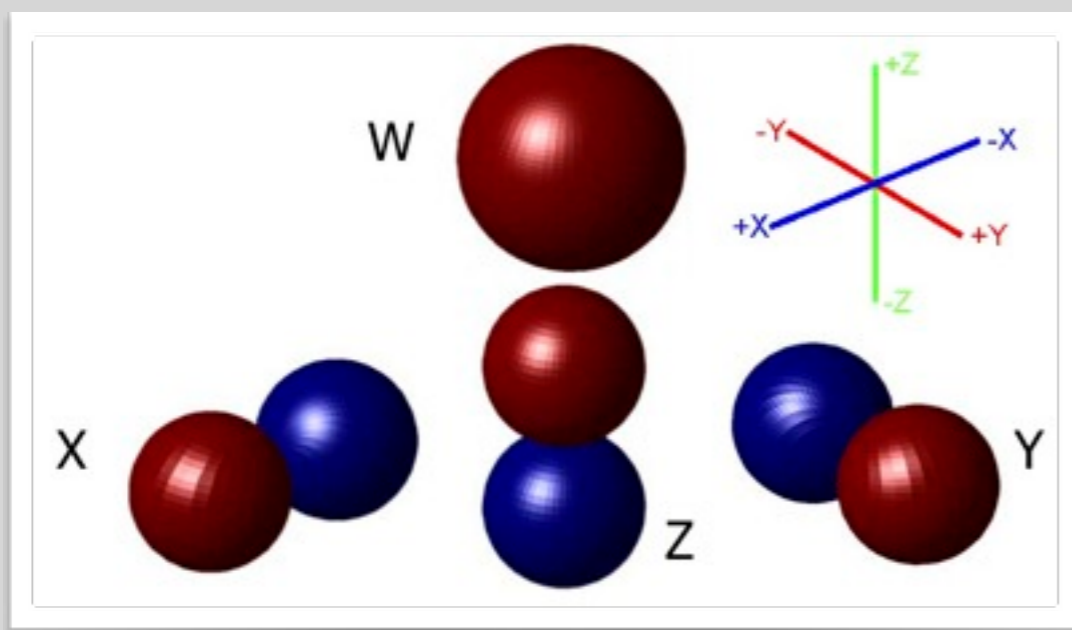
plane wave from 0° (M. Frank)

# What is Ambisonic(s)?

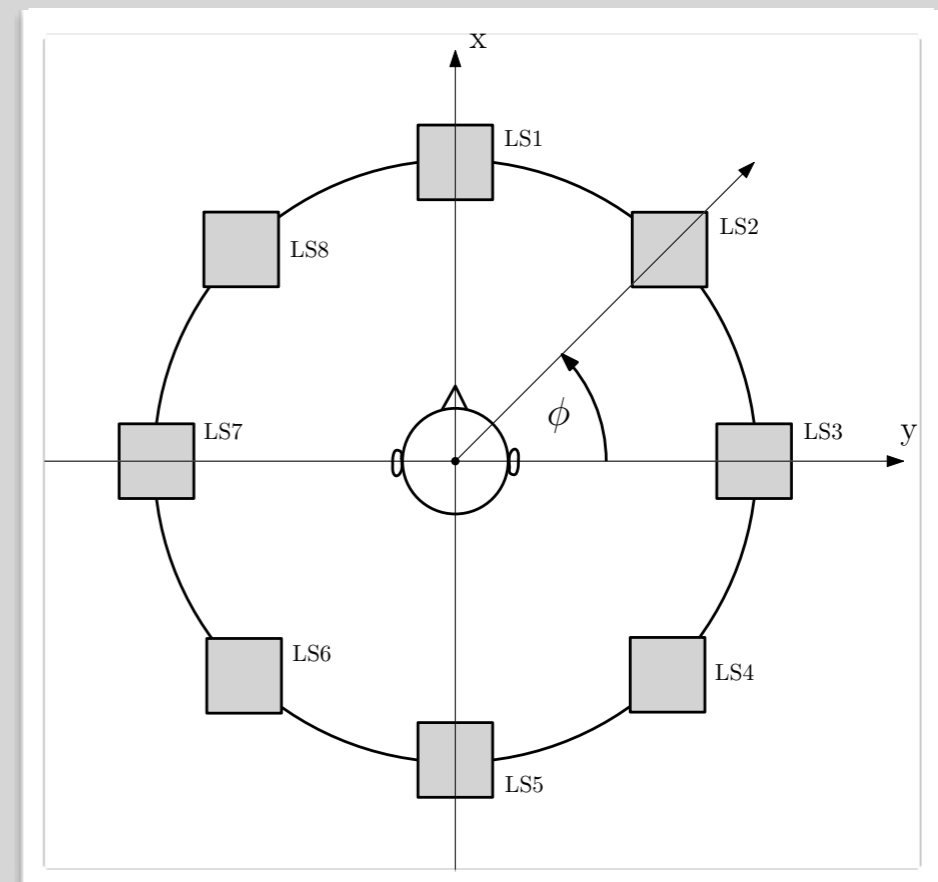
- Loudspeaker feeds are a linear combination of the B-Format signals (Decoder Matrix) e.g.:

$$p_j = \frac{1}{L} \left[ W \cdot \left( \frac{1}{\sqrt{2}} \right) + X \cdot (\cos \phi_j \cos \theta_j) + Y \cdot (\sin \phi_j \cos \theta_j) + Z \cdot (\sin \theta_j) \right]$$

- ALL speakers work together to synthesize the sound-field



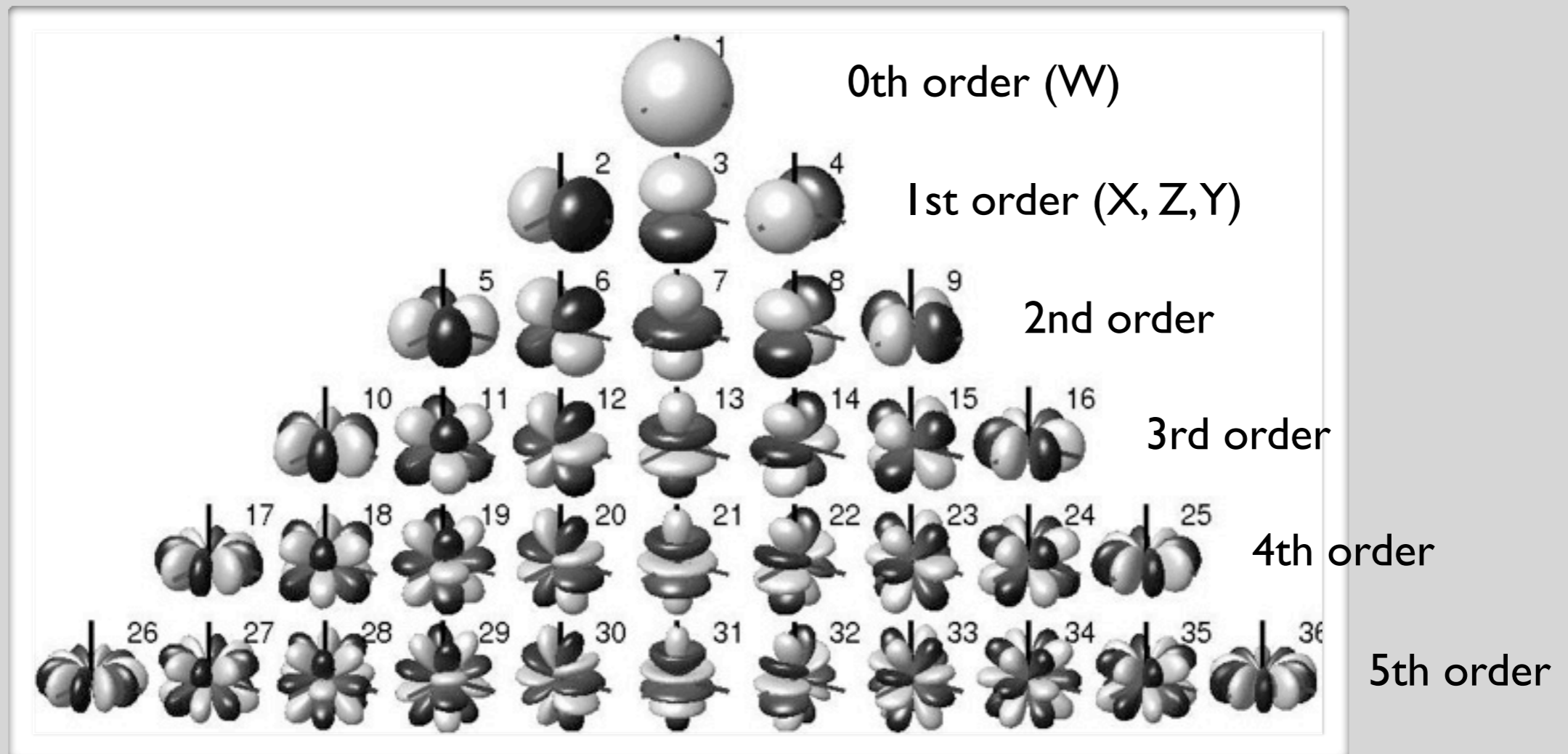
B-Format (1st order Ambisonics)



Regular 2D loudspeaker placement

# What is Ambisonic(s)?

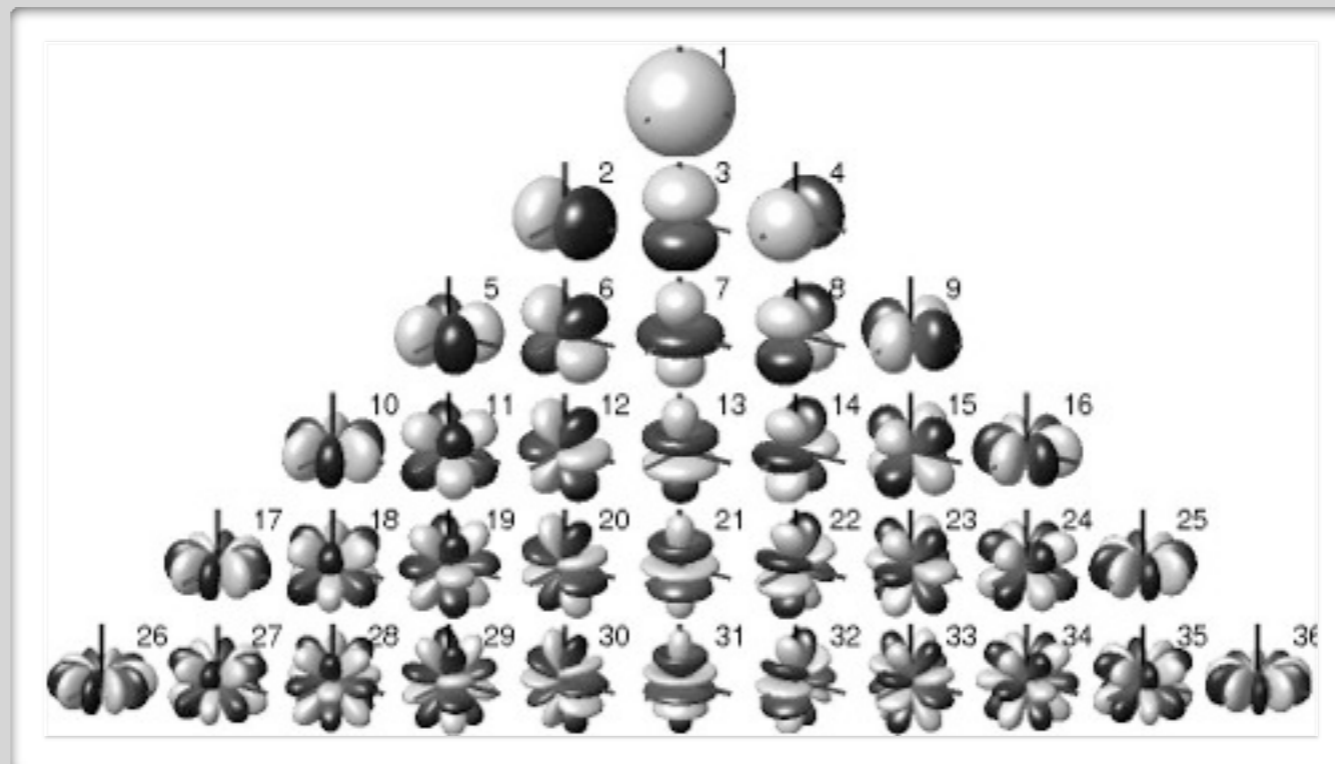
Higher Order Ambisonics (HOA) extends this approach to a bigger set of spherical harmonics



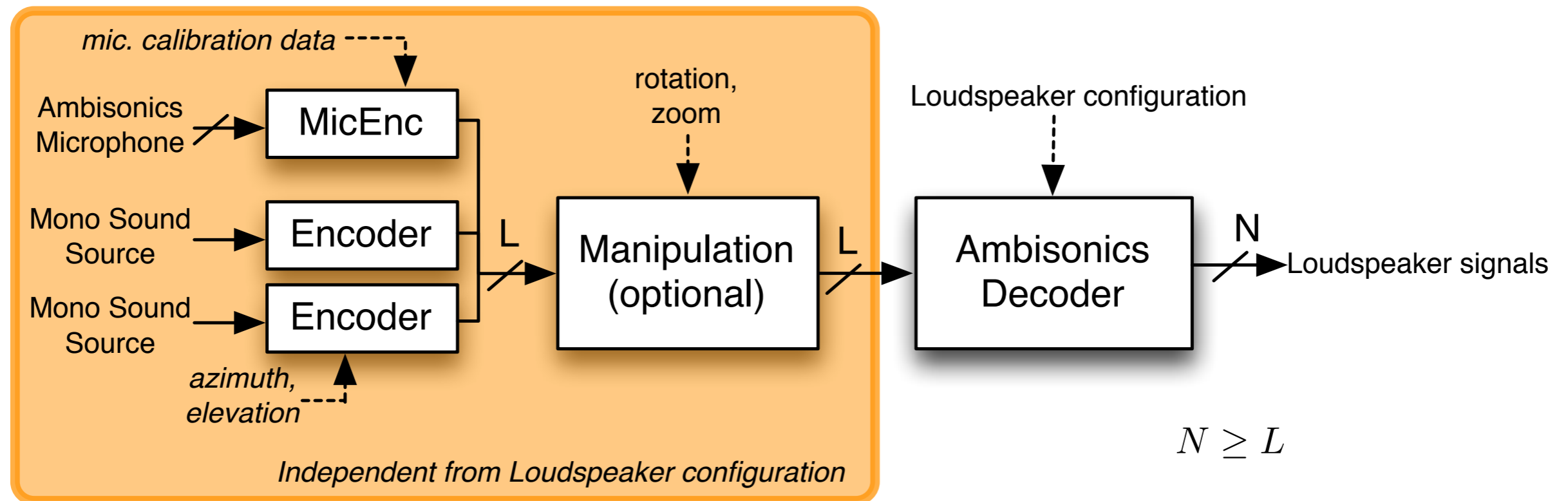
Spherical harmonics up to 5th order 3D

# Higher Order Ambisonics

- more accurate spatial resolution in a larger area (“sweet-spot”)
- i need more transmission channels, loudspeakers



# Ambisonics in practice



$$L_{3D} = (M + 1)^2$$

$$L_{2D} = 2M + 1$$

M... Ambisonics Order

L... Number of Ambisonics Channels

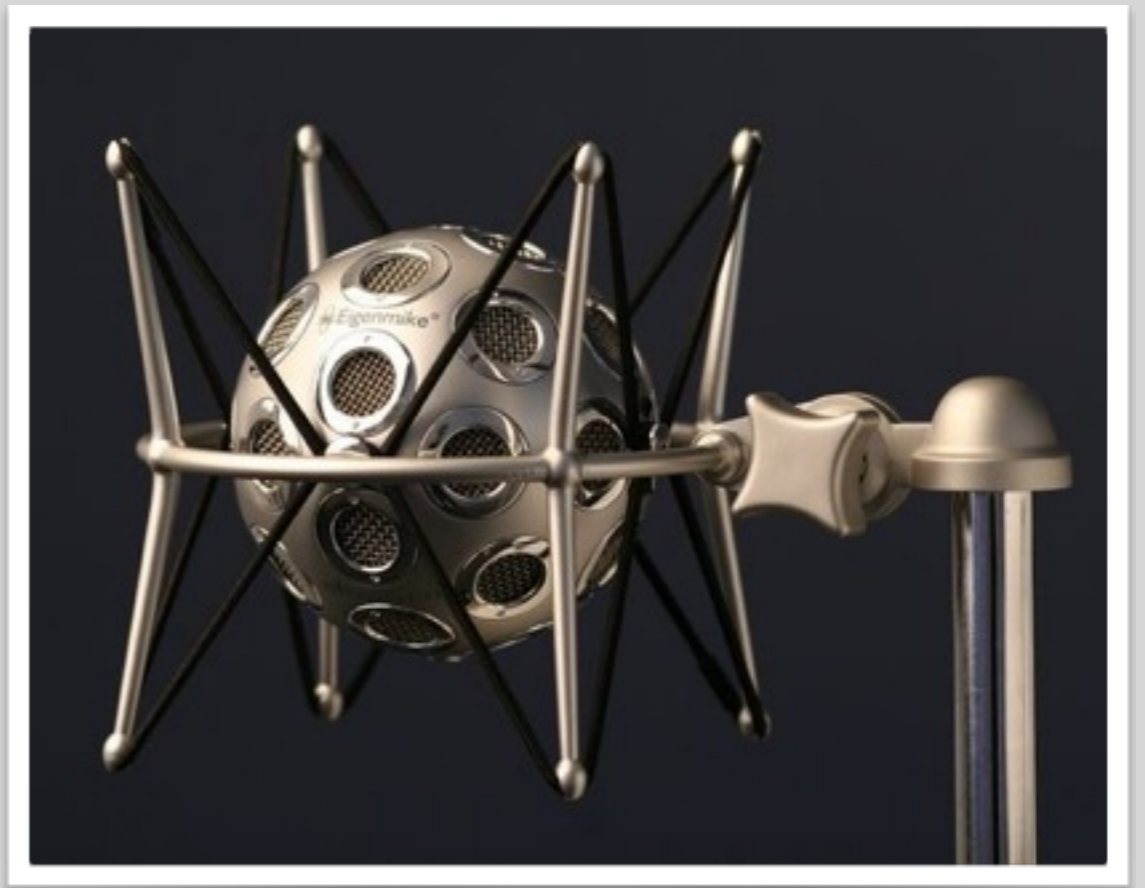
N... Number of Loudspeakers

Overview of an Ambisonics Playback System

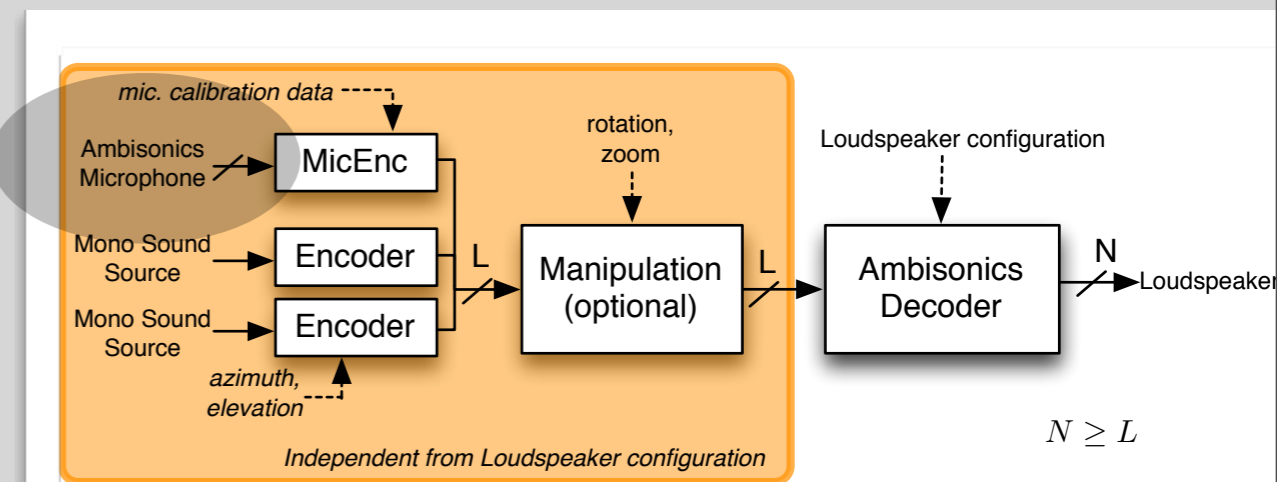
# Microphone arrays



Core Sound Tetramic, Soundfield and  
DIY First Order Microphone (4 cardioids)



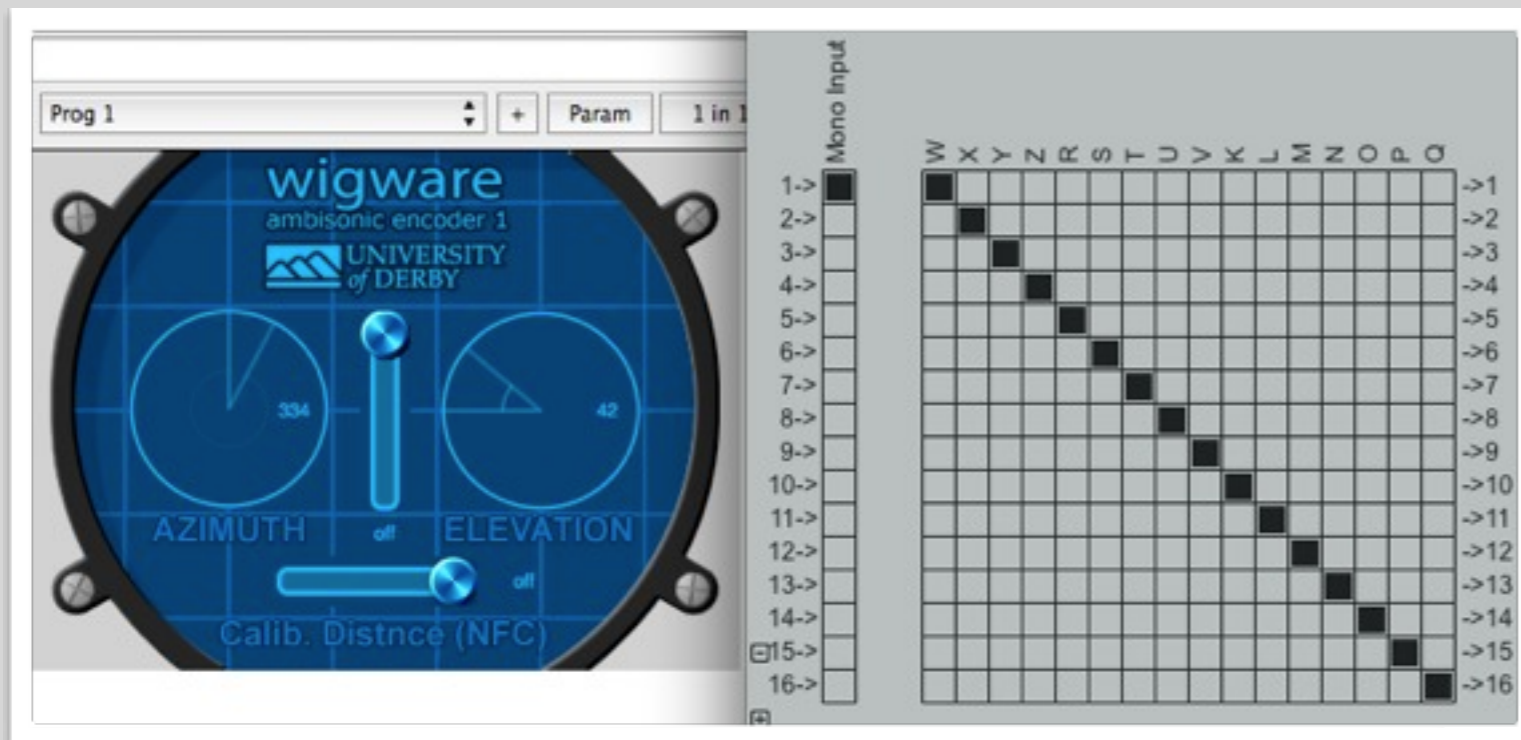
Eigenmike (32 microphone capsules) for  
Higher Order Recordings



$$N \geq L$$

# Encoder (Panner)

Mono Signal sources can be positioned defining their position with azimuth ( $\varphi$ ) and elevation ( $\theta$ ).



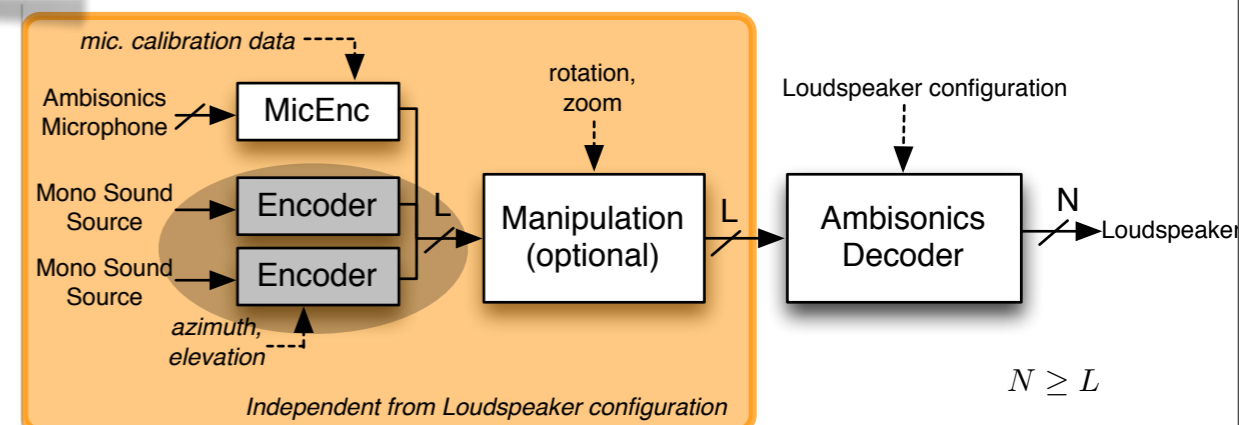
Wigware VST Plug-Ins (Win, Mac OS)  
up to 3rd order 3D

Channel Routing in Reaper

- Mono Source Input
- 3rd order 3D output:

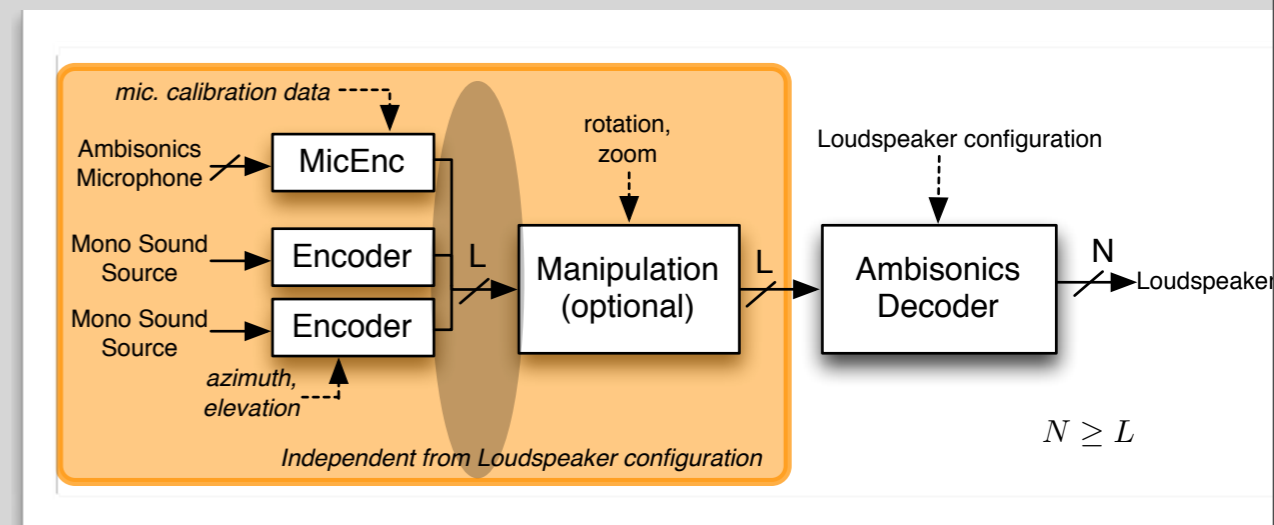
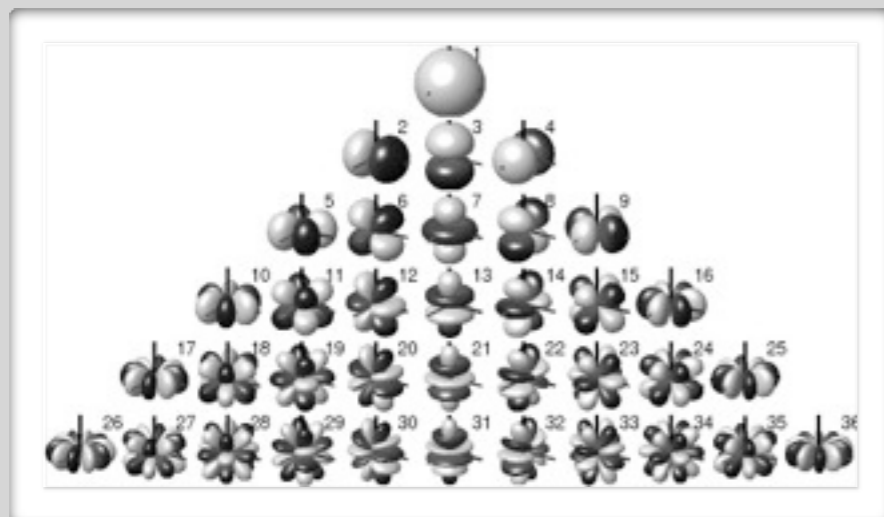
$$L_{3D} = (M + 1)^2 = 16$$

$$L_{2D} = 2M + 1 = 7$$



# Ambisonics Domain

the recording or streaming of Ambisonics signals is independent from the final loudspeaker placement...



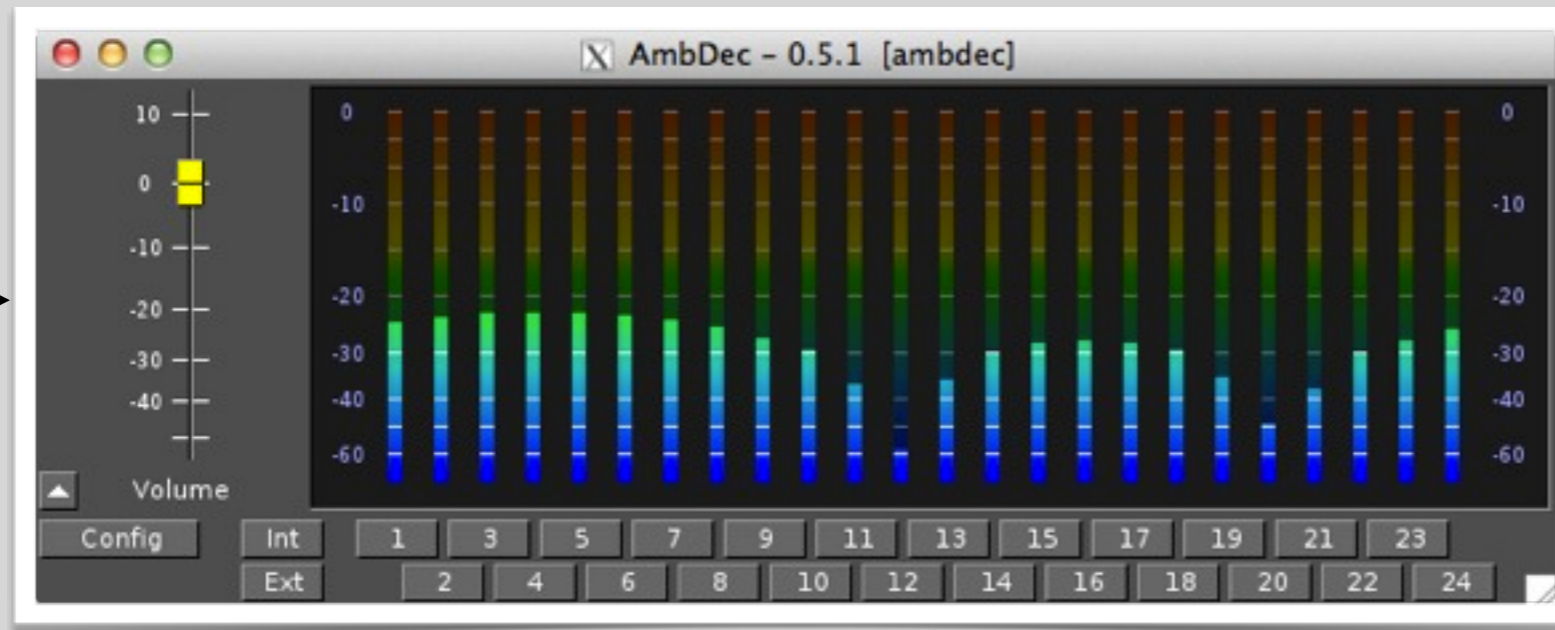


@ Mumuth Graz

# Decoder

- Computing loudspeaker signals from B-Format

$L$   
 $\swarrow$  →  
 B-Format

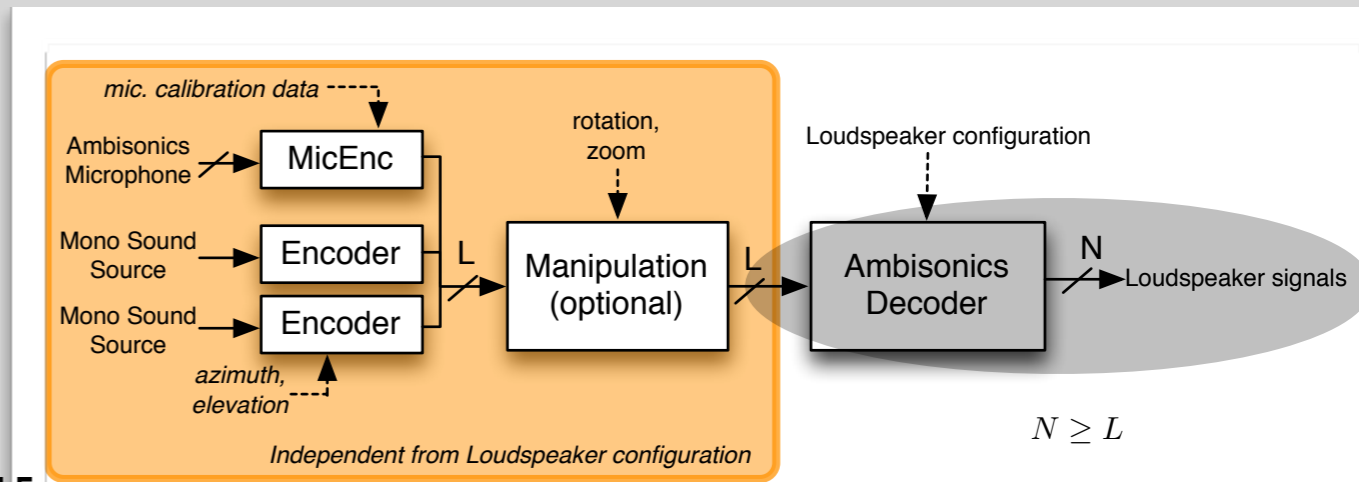


$N$  loudspeaker signals  
 $\swarrow$  →

Ambdec (Linux, Mac OS)  
 Standalone Jack Application



VST Plug-In Decoder (Windows, Mac OS)

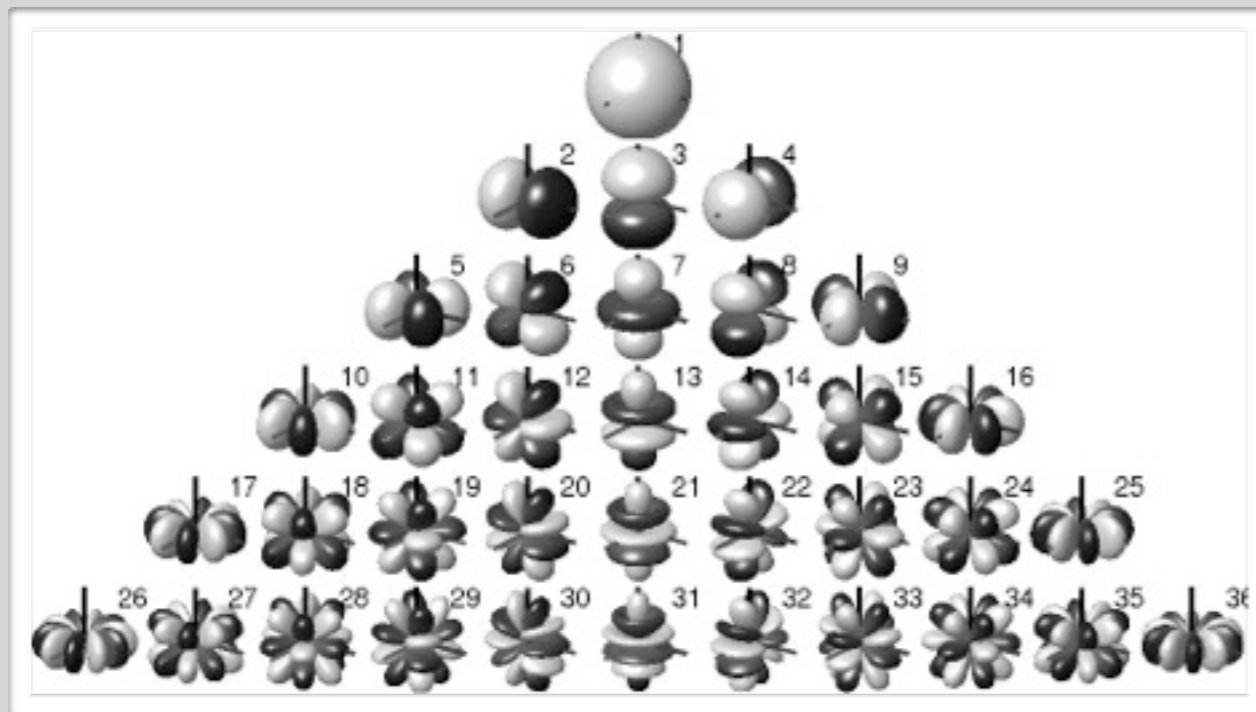


# Problems with existing solutions

- Operating system compatibility (Win, Mac OS, Linux)  
Plug-In Standards (VST, AU, AAX, RTAS, LV2)
- Listening/producing at home?
- jumps in angular representation between  $-180^\circ$  and  $180^\circ$   
(especially when encoding - panning)
- Different standards about channel sequence and channel scaling

# Ambisonics standards

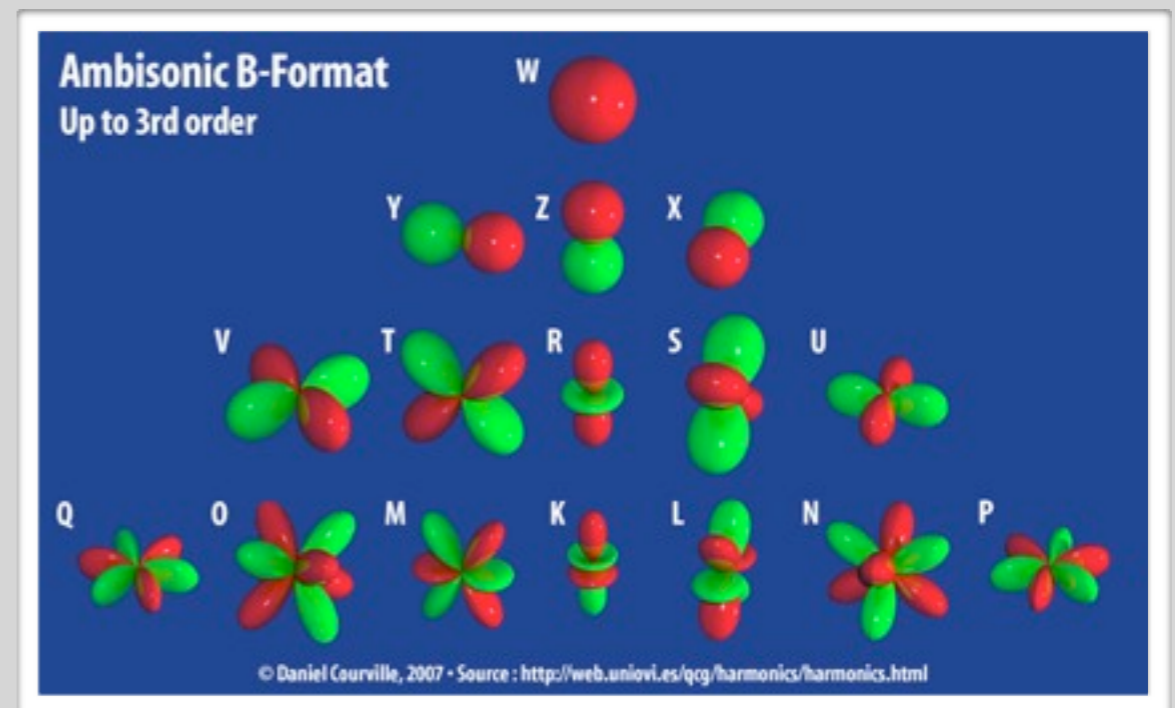
Problem: different standardization of channel order and normalization (scaling) of channels



## ACN Order

(Ambisonics Channel Numbering)

> should be new standard



## Furse-Malham Ordering

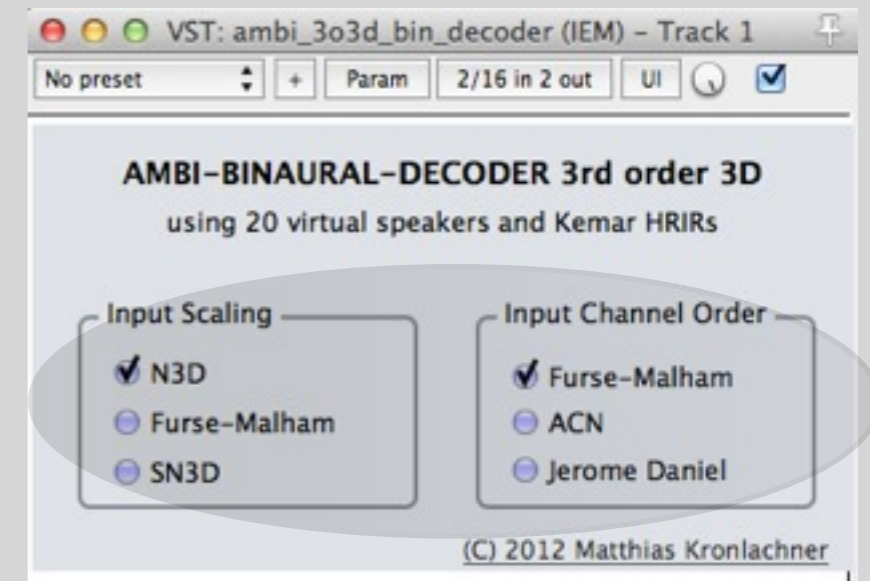
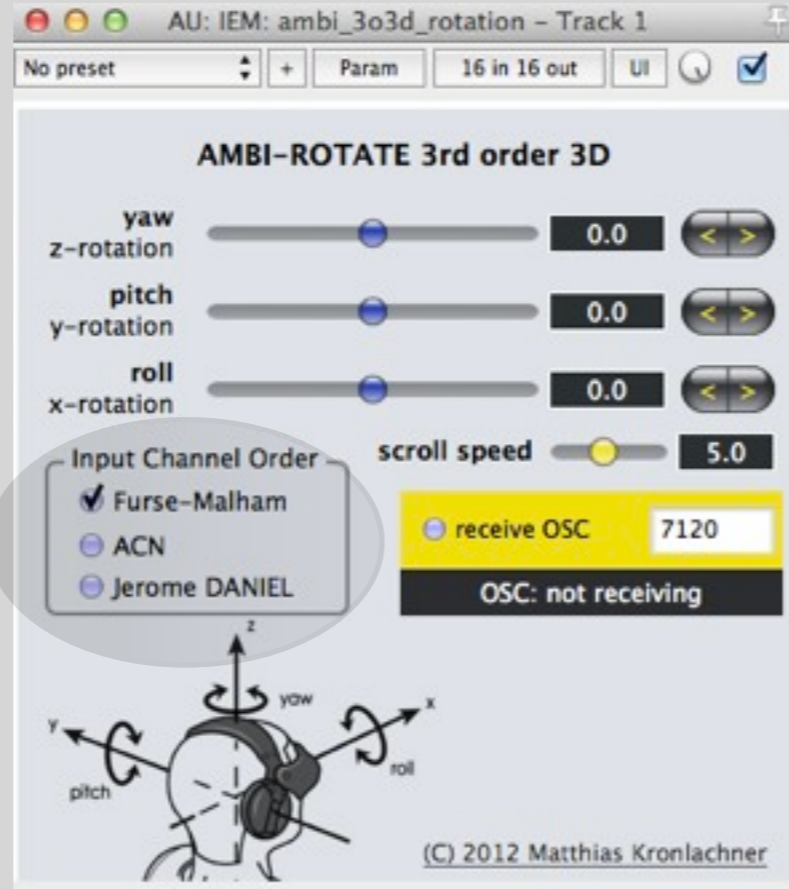
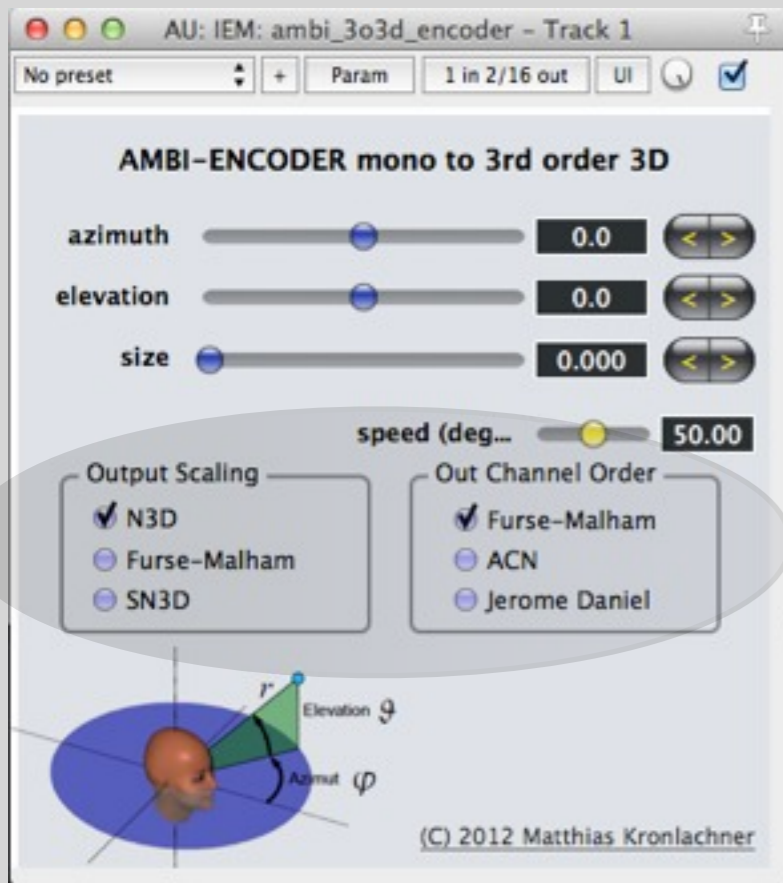
1st order: W, X, Y, Z

2nd order: R, S, T, U, V

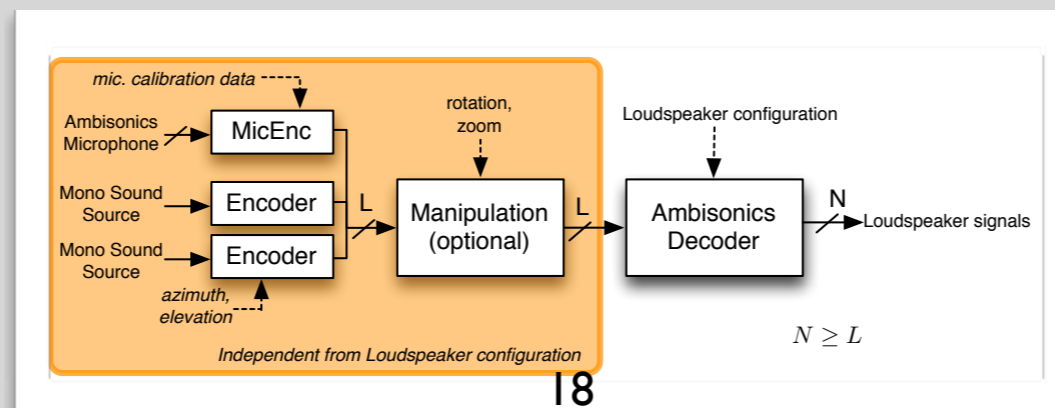
3rd order: K, L, M, N, O, P, Q

*de facto* standard for order  $\leq 3$

# Different standards

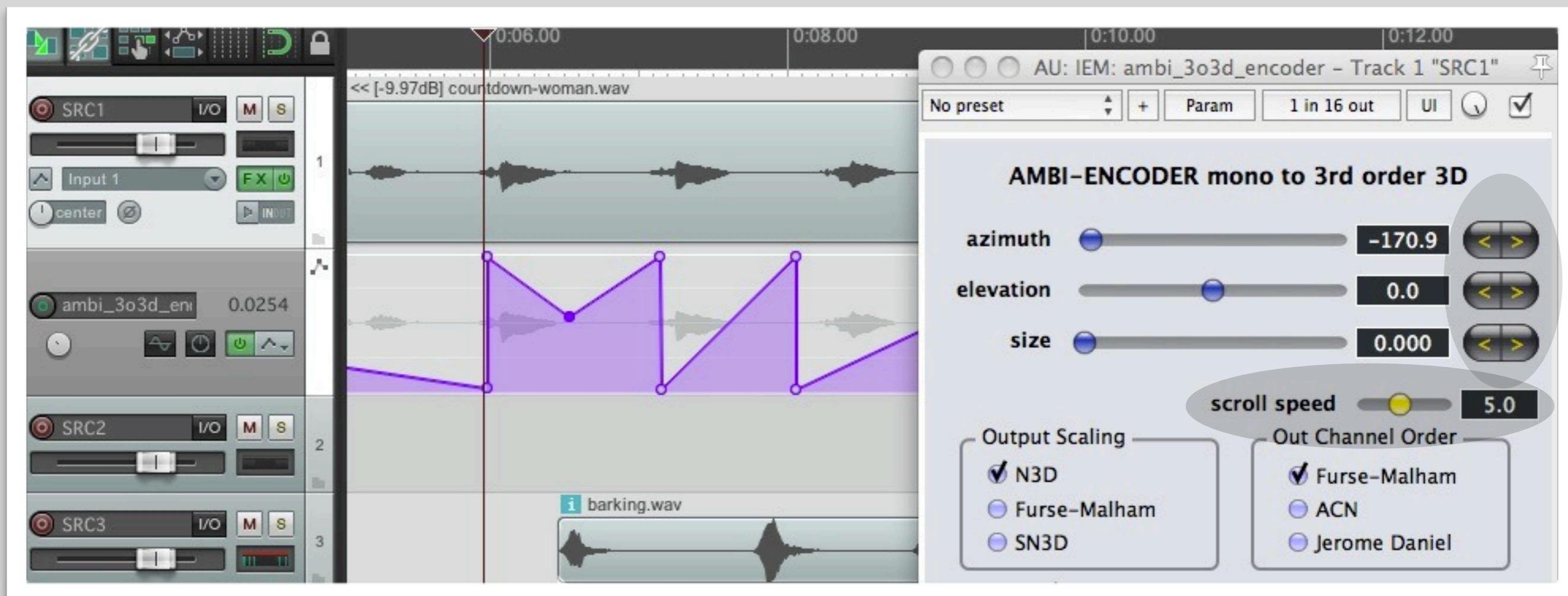


adjust scaling and channel order in every stage



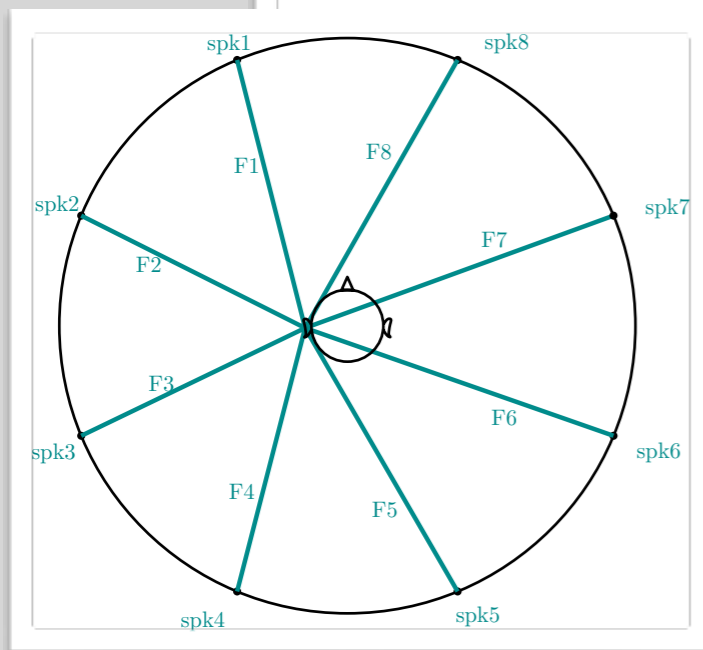
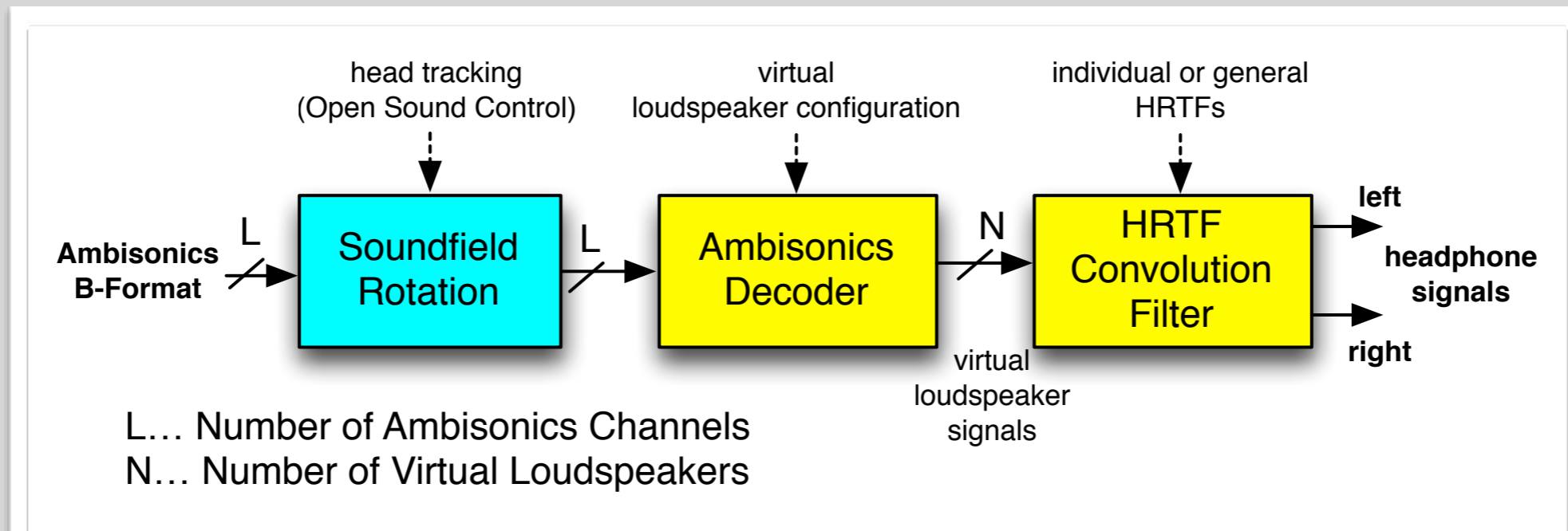
# Jumping angular representation

- Visual representation of automation curve not intuitive



# Listening at home

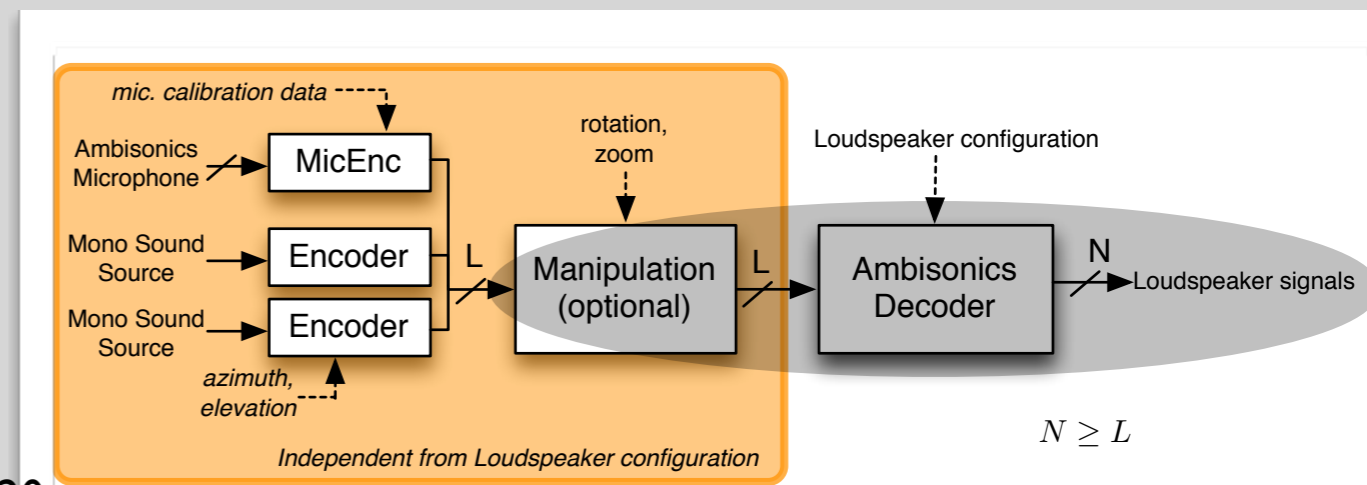
Listening to recordings without a large number of loudspeakers possible using headphones and headtracking



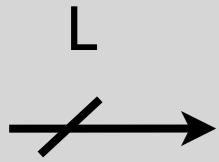
loudspeakers - left ear transfer path

Related Transfer Function

## Ambisonics Binaural Decoder



# Binaural Decoder

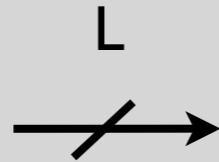


AMBI-ROTATE 3rd order 3D

yaw: -37.1  
pitch: 19.9  
roll: 11.4

Input Channel Order:  
 Furse-Malham  
 ACN  
 Jerome DANIEL

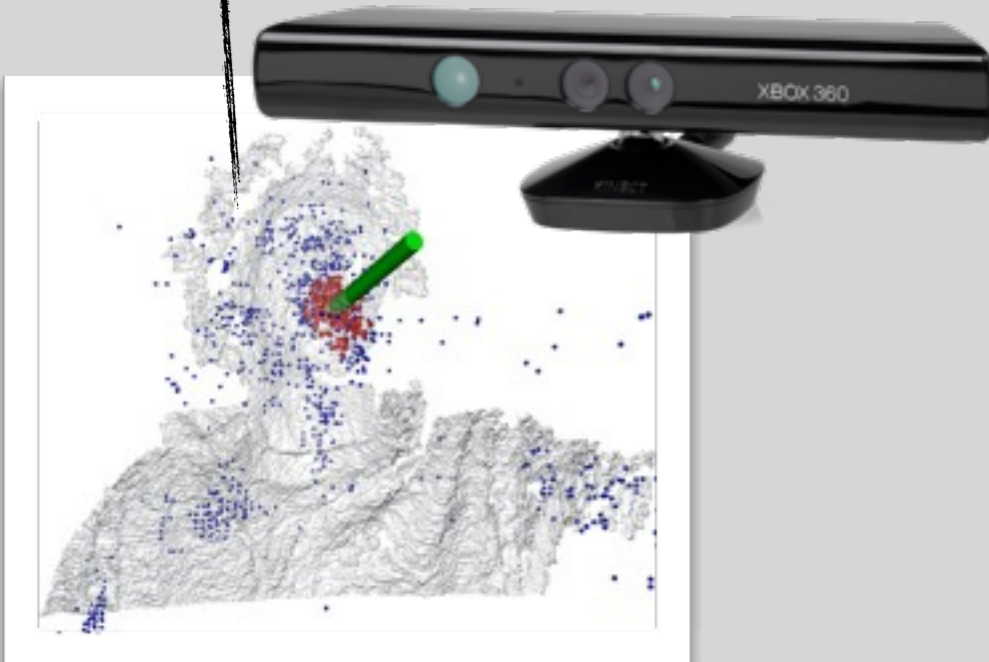
OSC: receiving on port 7120



AMBI-BINAURAL-DECODER 3rd order 3D  
using 20 virtual speakers and Kemar HRIRs

Input Scaling:  
 N3D  
 Furse-Malham  
 SN3D

Input Channel Order:  
 Furse-Malham  
 ACN  
 Jerome Daniel



# Things to do...

## **... for me**

- more sophisticated panning options and visualization
- order  $> 3$
- CPU efficiency

## **... for the community**

- standardize channel sequence and scaling for order  $> 3$
- Affordable DIY Higher Order Microphones

**Ačīū!**