Lecture

Speech and Audio Signal Processing

TECHNISCHE UNIVERSITÄT DARMSTADT

Lecture 14: Sound reproduction by Wave Field Synthesis (WFS) and Higher Order Ambisonics (HOA)

Content

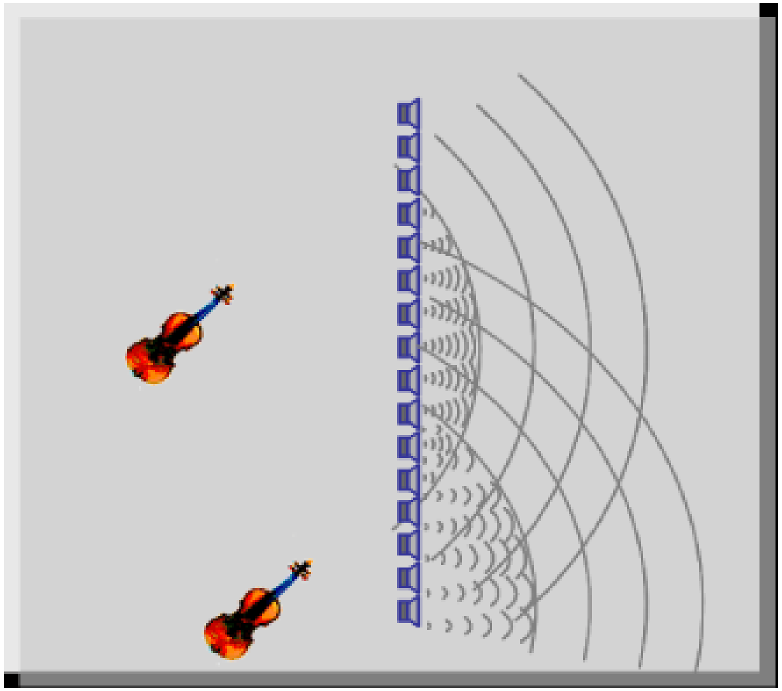
- Wave Field Synthesis (WFS)
  - Typically using a 2D loudspeaker setup, e.g., circular or rectangular array.
  - Sound generation in a plane.
  - Targets for generation a good sound field within the plane rather independent of the listening position.

- Higher Order Ambisonics (HOA)
  - Often targets for the generation of a 3D sound field.
  - Typically, a precise sound field should be generated in a sweet spot.

Wave field synthesis

WFS: Concept of Wave field synthesis

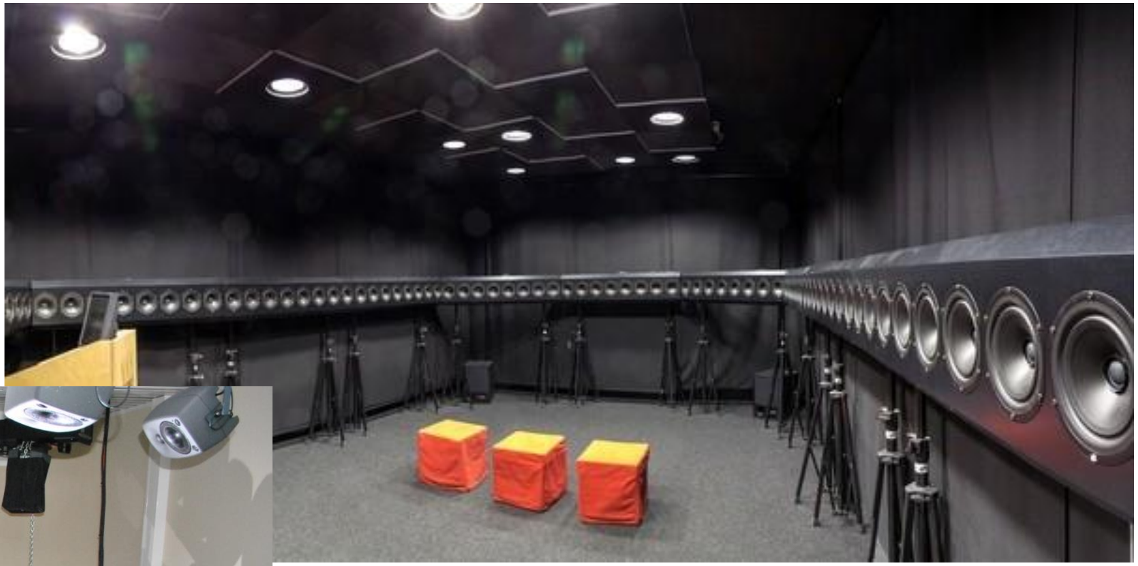
- **The basic idea:**
  Reproduction of a sound source (sound waves) by the superposition of loudspeaker signals:



- **Targets:**
  Typically, horizontal modelling by loudspeakers in a circular or quadratic, etc. array.

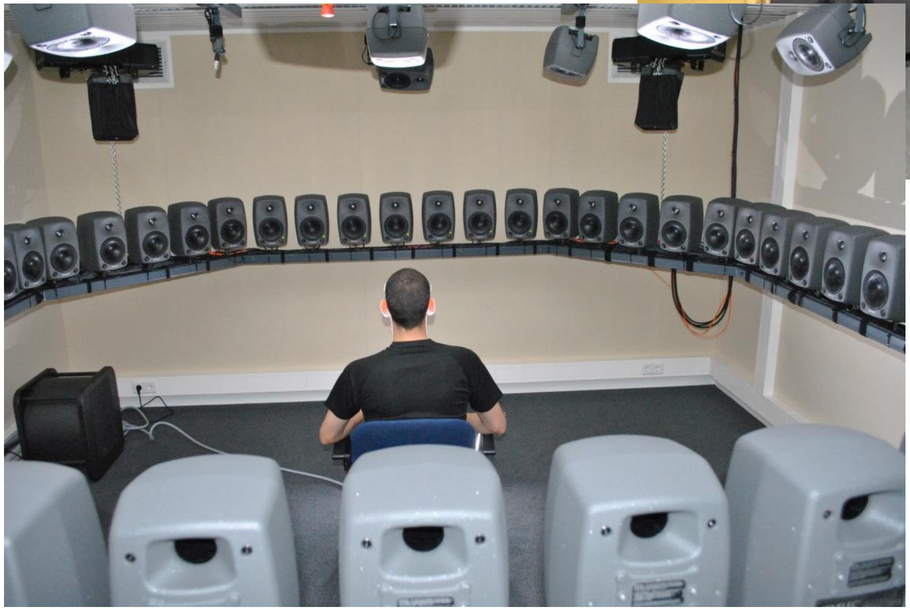
  Reproduction of sound in a rather big section of the loudspeaker array
  => no sweet spot.
  Limitations by the loudspeaker distance, i.e., spatial sampling which leads to spatial aliasing.

WFS: Concept of Wave Field Synthesis



https://www.quora.com/

Sivantos WFS System



WFS: Basics

- Based on the wave equation:

$$\triangle p(\boldsymbol{x},t) - \frac{1}{c^2}\frac{\partial}{\partial t^2}p(\boldsymbol{x},t) = 0 \qquad \boldsymbol{x} = [x,y,z]^{\mathrm{T}}$$

- Application of the Fourier Transform => Helmholtz equation :

$$\triangle P(\boldsymbol{x},\omega) + \left(\frac{\omega}{c}\right)^2 P(\boldsymbol{x},\omega) = 0 \qquad P(\boldsymbol{x},\omega) \;\bullet\!\!-\!\!\!-\!\!\circ\; p(\boldsymbol{x},t)$$

- Plane wave solutions in the time and frequency domain:

$$p(\boldsymbol{x},t) = f\left(t + \frac{1}{c} < \boldsymbol{x},\boldsymbol{n} >\right)$$

$$P(\boldsymbol{x},\omega) = F(\omega)e^{j\frac{\omega}{x}<\boldsymbol{x},\boldsymbol{n}>}$$

$< \boldsymbol{x},\boldsymbol{n} >$ scalar product between the two vectors

$\boldsymbol{n}$ Vector defining the propagation direction of the plane wave

WFS: Basics

- Contribution of a spatial sound field of a point source (radiating in all directions):

$$\delta_{3D}(\boldsymbol{x}) = \delta(x)\,\delta(y)\,\delta(z)$$

$$q(\boldsymbol{x},t) = q_0(\boldsymbol{x}_0,t)\,\delta_{3D}(\boldsymbol{x}-\boldsymbol{x}_0) \qquad Q_0(\omega,\boldsymbol{x}_0) \bullet\!\!-\!\!\!-\!\!\circ\; q_0(t,\boldsymbol{x}_0)$$

$$Q(\boldsymbol{x},\omega) = Q_0(\boldsymbol{x}_0,\omega)\,\delta_{3D}(\boldsymbol{x}-\boldsymbol{x}_0)$$

- Spatial sound field for a single point source:

$$P_0(\boldsymbol{x},\omega) = \iiint_V G_{3D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)\,Q(\boldsymbol{x}_0,\omega)d\boldsymbol{x}_0$$

- in in free-field (Green's function):

$$G^f_{3D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega) = \frac{1}{4\pi}\frac{e^{-j\frac{\omega}{c}|\boldsymbol{x}-\boldsymbol{x}_0|}}{|\boldsymbol{x}-\boldsymbol{x}_0|}$$

WFS: Basics

- *3D:* Point source, i.e., superposition of spherical waves, according to the number of point sources

$$Q_0(\boldsymbol{x},\omega) = Q_0(x,y,z,\omega)$$

- *2D*: Line source, i.e., superposition of appropriate free-field model according to the number of line sources

$$Q_1(\boldsymbol{x},\omega) = Q_1(x,y,\omega)$$

- Sound field of a line source:

Does not depend on z

$$P_1(\boldsymbol{x},\omega) = \iiint_V G_{3D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)\, Q_1(\boldsymbol{x}_0,\omega) d\boldsymbol{x}_0$$

WFS: Basics

- Sound field of a line source:

Does not depend on z

2-dim

$$P_1(\boldsymbol{x},\omega) = \iiint_V G_{3D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)\, Q_1(\boldsymbol{x}_0,\omega) d\boldsymbol{x}_0$$

$$P_1(\boldsymbol{x},\omega) = \iint_V \underbrace{\left[\int G_{3D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega) dz_0\right]}_{G_{2D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)} Q_1(\boldsymbol{x}_0,\omega) d\boldsymbol{x}_0$$

- Resulting in the free-field sound field of a line source:

$$G_{2D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega) = \int_{-\infty}^{\infty} G_{3D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega) dz_0$$

$$G_{2D}^f(\boldsymbol{x}|\boldsymbol{x}_0,\omega) = \frac{j}{2}\, H_0^{(2)}\left(\frac{\omega}{c}\|\boldsymbol{x}-\boldsymbol{x}_0\|\right)$$

with: $H_0^{(2)}\left(\frac{\omega}{c}\rho\right) = J_0\left(\frac{\omega}{c}\rho\right) - jN_0\left(\frac{\omega}{c}\rho\right)$

WFS: Basics

- Relation of Green's functions of line and point sources for far-field assumptions (mathematical approximation) :

$$\frac{\omega}{c}\|\boldsymbol{x}-\boldsymbol{x}_0\| \gg 1$$

$$G_{3D}^f(\boldsymbol{x}|\boldsymbol{x}_0,\omega) = \frac{1}{4\pi}\,\frac{e^{-j\frac{\omega}{c}|\boldsymbol{x}-\boldsymbol{x}_0|}}{|\boldsymbol{x}-\boldsymbol{x}_0|}$$

$$G_{2D}^f(\boldsymbol{x}|\boldsymbol{x}_0,\omega) = \frac{j}{2}\,H_0^{(2)}\left(\frac{\omega}{c}\|\boldsymbol{x}-\boldsymbol{x}_0\|\right)$$

- Resulting in the approximation of the free-field sound field of a line source for the far-field:

$$G_{2D}^f(\boldsymbol{x}|\boldsymbol{x}_0,\omega) \approx \sqrt{\frac{2\,\pi\|\boldsymbol{x}-\boldsymbol{x}_0\|}{j\frac{\omega}{c}}}\,G_{3D}^f(\boldsymbol{x}|\boldsymbol{x}_0,\omega)$$

WFS: Kirchhoff-Helmholtz-Integral

- **Basic concept of WFS:**
  The sound field within the Volume V can be expressed by the known sound pressure and its gradient on the surface of the boundary volume, assuming no sound sources within the volume

$$P_0(\boldsymbol{x},\omega) = -\oiint_{\partial V} \left[G_{3D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)\frac{\partial}{\partial \boldsymbol{n}} P_0(\boldsymbol{x}_0,\omega) - P_0(\boldsymbol{x}_0,\omega)\frac{\partial}{\partial \boldsymbol{n}} G_{3D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)\right] d\boldsymbol{x}_0$$

$$\boldsymbol{x} \in V$$

- Then the sound pressure outside the volume is zero

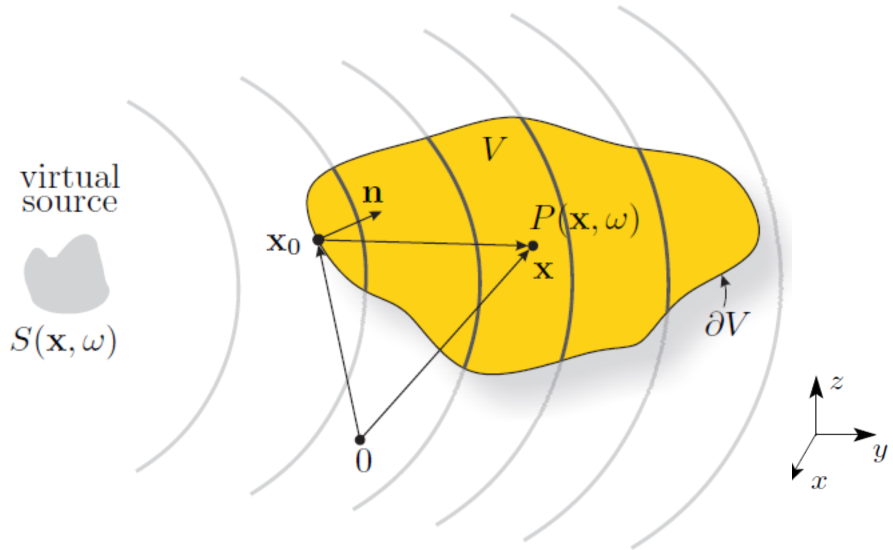
$$P_0(\boldsymbol{x},\omega) = 0 \qquad \boldsymbol{x} \notin V$$

- This is a condition which is not really necessary, since we are interested in the sound field within the volume. However, we can use a relaxation of this property for a later simplification of the Kirchhoff-Helmholtz-Integral.

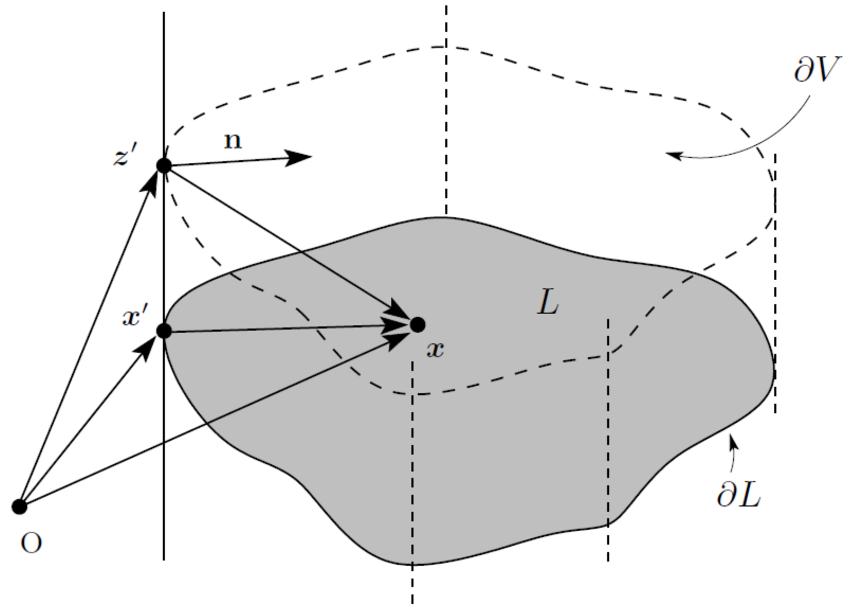
WFS: Kirchhoff-Helmholtz-Integral: 3D and 2D

❑ Basic concept of WFS:

General 3D setup



Prism 3D setup
=> leading to the 2D setup
which models the sound field in a plane L.



WFS: Kirchhoff-Helmholtz-Integral: 2D

TECHNISCHE UNIVERSITÄT DARMSTADT

- **2D WFS equation in the plane:**
  Surface integration over the prism, invariant sound field within the z-direction $P_1(\boldsymbol{x}_0,\omega) = P_0(\boldsymbol{x}_0,\omega)$

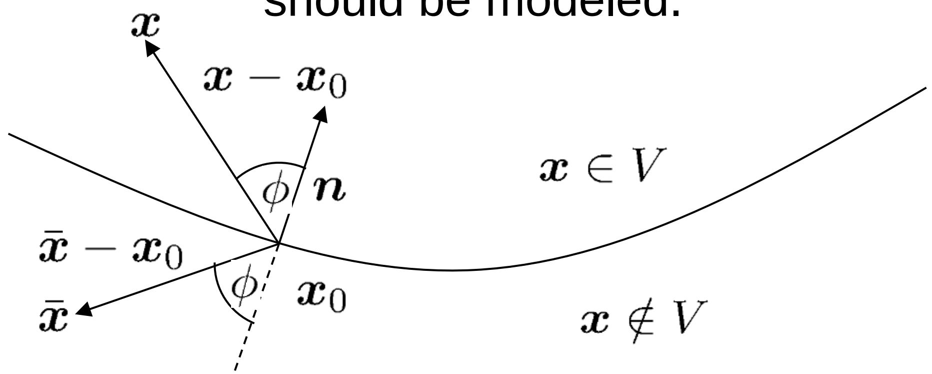
$$
\begin{aligned}
P_0(\boldsymbol{x},\omega) &= -\oiint_{\partial V} \left[G_{3D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)\frac{\partial}{\partial \boldsymbol{n}}P_0(\boldsymbol{x}_0,\omega) - P_0(\boldsymbol{x}_0,\omega)\frac{\partial}{\partial \boldsymbol{n}}G_{3D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)\right] d\boldsymbol{x}_0 \\
&= -\oint_{\partial L}\int_{z_0=-\infty}^{\infty}\left[G_{3D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)\frac{\partial}{\partial \boldsymbol{n}}P_0(\boldsymbol{x}_0,\omega)\right. \\
&\qquad \left. -P_0(\boldsymbol{x}_0,\omega)\frac{\partial}{\partial \boldsymbol{n}}G_{3D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)\right] dx_0dy_0dz_0 \\
&= -\oint_{\partial L}\left[\int_{z_0=-\infty}^{\infty} G_{3D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)dz_0\right]\frac{\partial}{\partial \boldsymbol{n}}P_0(\boldsymbol{x}_0,\omega) \\
&\qquad -P_0(\boldsymbol{x}_0,\omega)\frac{\partial}{\partial \boldsymbol{n}}\left[\int_{z_0=-\infty}^{\infty} G_{3D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)dz_0\right] dx_0dy_0
\end{aligned}
$$

$$
P_1(\boldsymbol{x},\omega) = -\oint_{\partial L}\left[G_{2D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)\frac{\partial}{\partial \boldsymbol{n}}P_1(\boldsymbol{x}_0,\omega) - P_1(\boldsymbol{x}_0,\omega)\frac{\partial}{\partial \boldsymbol{n}}G_{2D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)\right] d\boldsymbol{x}_0
$$

WFS: Choice of Green’s functions

TECHNISCHE UNIVERSITÄT DARMSTADT

- **Concept:**
  Choose Green’s functions such that it is symmetric with respect to the boundary of the surface.
  Only possible if the property $P_0(\boldsymbol{x},\omega)=0 \quad \boldsymbol{x}\notin V$
  is relaxed. => No problem: Only the sound field within the volume should be modeled.



- Choice of the Green’s function at the boundary as the sum of symmetric free-field Green’s function with respect to the volume’s surface

$$G_{iD}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)=G_{iD}^{f}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)+G_{iD}^{f}(\bar{\boldsymbol{x}}|\boldsymbol{x}_0,\omega)$$

$$i\in[2,3]$$

WFS: Choice of Green's functions

TECHNISCHE UNIVERSITÄT DARMSTADT

- **Resulting in:**

Dipole: $$\frac{\partial}{\partial \boldsymbol{n}} G_{iD}(\boldsymbol{x}|\boldsymbol{x}_0,\omega) = 0$$

Monopole: $$G_{iD}(\boldsymbol{x}|\boldsymbol{x}_0,\omega) = 2\, G^f_{iD}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)$$

No dipoles, only monopoles model the Kirchhoff-Helmholtz-Integral

- **Kirchhoff-Helmholtz-Integrals only with monopoles:**

$$P_0(\boldsymbol{x},\omega) = -\oiint_{\partial V} 2\, G^f_{3D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega) \frac{\partial}{\partial \boldsymbol{n}} P_0(\boldsymbol{x}_0,\omega) d\boldsymbol{x}_0 \qquad \boldsymbol{x} \in V$$

3 dimensional

$$P_1(\boldsymbol{x},\omega) = -\oint_{\partial L} 2\, G^f_{2D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega) \frac{\partial}{\partial \boldsymbol{n}} P_1(\boldsymbol{x}_0,\omega) d\boldsymbol{x}_0 \qquad \boldsymbol{x} \in L$$

2 dimensional

WFS: Green's functions: 2D and 3D

- **Replacing line sources by point sources:**

$$G_{2D}^{f}(\boldsymbol{x}|\boldsymbol{x}_0,\omega) \approx \sqrt{\frac{2\,\pi\|\boldsymbol{x}-\boldsymbol{x}_0\|}{j\,\frac{\omega}{c}}}\,G_{3D}^{f}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)$$

- **Allows to model a 2D sound field not by line but by point sources** => desired setup for WFS in a horizontal plane generated by loudspeakers (point source):

$$P_1(\boldsymbol{x},\omega) = -\oint_{\partial L} 2\,G_{2D}^{f}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)\frac{\partial}{\partial \boldsymbol{n}}P_1(\boldsymbol{x}_0,\omega)d\boldsymbol{x}_0 \qquad \boldsymbol{x}\in L$$

2 dimensional

$$P_1(\boldsymbol{x},\omega) \approx -\oint_{\partial L} 2\sqrt{\frac{2\,\pi\|\boldsymbol{x}-\boldsymbol{x}_0\|}{j\,\frac{\omega}{c}}}\,G_{3D}^{f}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)\frac{\partial}{\partial \boldsymbol{n}}P_1(\boldsymbol{x}_0,\omega)d\boldsymbol{x}_0 \qquad \boldsymbol{x}\in L$$

WFS: Choice of Green's functions

TECHNISCHE UNIVERSITÄT DARMSTADT

- **The concept of driving signals / functions:**

$$P_1(\boldsymbol{x},\omega) \approx -\oint_{\partial L} 2\sqrt{\frac{2\,\pi\|\boldsymbol{x}-\boldsymbol{x}_0\|}{j\frac{\omega}{c}}}\, G^f_{3D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)\frac{\partial}{\partial \boldsymbol{n}}P_1(\boldsymbol{x}_0,\omega)d\boldsymbol{x}_0 \qquad \boldsymbol{x}\in L$$

$$\approx -\oint_{\partial L} G^f_{3D}(\boldsymbol{x}|\boldsymbol{x}_0,\omega)D(\boldsymbol{x}|\boldsymbol{x}_0,\omega)d\boldsymbol{x}_0$$

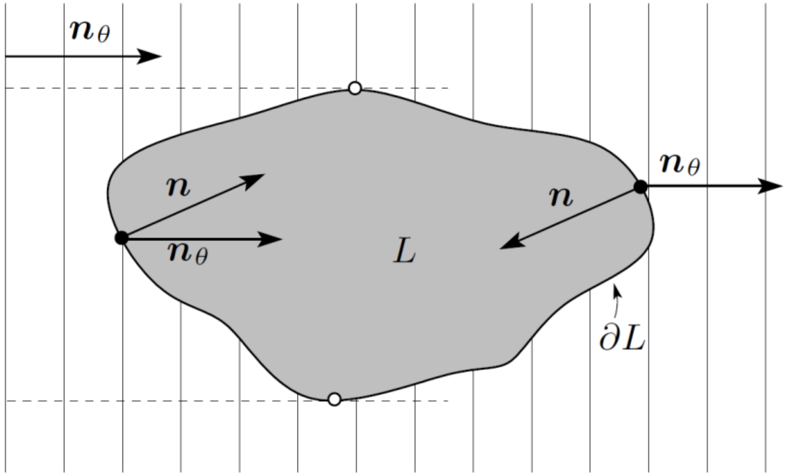
Driving signal: $D(\boldsymbol{x}|\boldsymbol{x}_0,\omega) = 2\sqrt{\frac{2\,\pi\|\boldsymbol{x}-\boldsymbol{x}_0\|}{j\frac{\omega}{c}}}\frac{\partial}{\partial \boldsymbol{n}}P_1(\boldsymbol{x}_0,\omega)$

- **Driving signals:**
  Signals of point sources to generate a sound field, i.e., the driving signals define the WFS loudspeaker signals.

- Obtaining driving signals from a recorded wave field:
  => **Data base rendering**

WFS: Boundary conditions

- **Model:** Generating a sound field by the superposition of plane waves
  Plane wave with direction $\boldsymbol{n}_\theta$



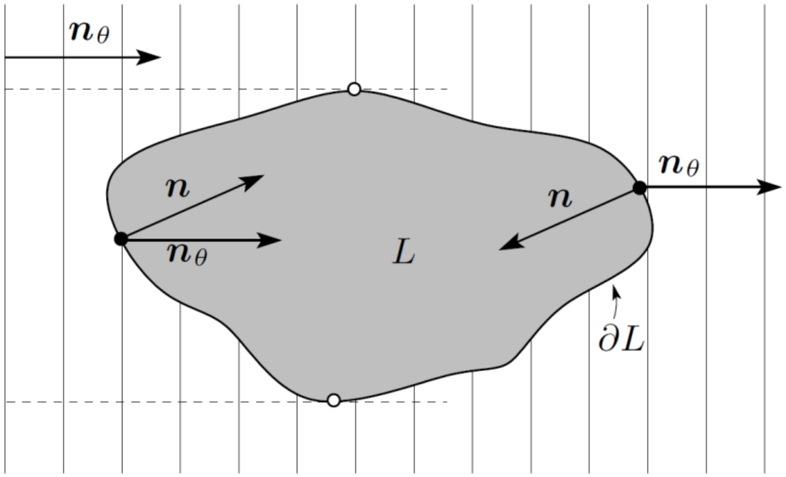
- Monopoles on the right section emanate into *L* in the opposite propagation direction than the plane wave.
- The Kirchhoff-Helmholtz equation with monopoles & dipoles avoid this problem (no sound outside *L*). With monopoles only => Countermeasures necessary!

WFS: Boundary conditions

- **Realized countermeasure:**
  Multiplication with a window function:

$$a(\boldsymbol{x}) = \begin{cases} 1 & <\boldsymbol{n}, \boldsymbol{n}_\theta> \; > 0 \\ 0 & \text{else} \end{cases}$$

=> Sources with propagation components opposite to the plane wave propagation are set to zero.



- Results in a modified driving function for monopoles in 2D:

$$D(\boldsymbol{x}|\boldsymbol{x}_0, \omega) = 2\, a(\boldsymbol{x}_0) \sqrt{\frac{2\,\pi \|\boldsymbol{x} - \boldsymbol{x}_0\|}{j\,\frac{\omega}{c}}} \, \frac{\partial}{\partial \boldsymbol{n}} P_1(\boldsymbol{x}_0, \omega)$$

WFS: Examples for driving functions

- **Plane wave:**

$$S_{pw}(\boldsymbol{x},\omega) = \hat{S}_{pw}(\omega)\, e^{-j\frac{\omega}{c}\boldsymbol{n}_{pw}^T\,\boldsymbol{x}}$$

$$\frac{\partial}{\partial \boldsymbol{n}} S_{pw}(\boldsymbol{x},\omega) = -j\frac{\omega}{c}\,\hat{S}_{pw}(\omega)\, e^{-j\frac{\omega}{c}\boldsymbol{n}_{pw}^T\,\boldsymbol{x}}\boldsymbol{n}_{pw}^T\,\boldsymbol{x}$$

- resulting in the following driving function:

$$D(\boldsymbol{x}|\boldsymbol{x}_0,\omega) = -2\,a_{pw}(\boldsymbol{x}_0)\sqrt{\frac{2\,\pi\|\boldsymbol{x}-\boldsymbol{x}_0\|}{j\frac{\omega}{c}}}\left[j\frac{\omega}{c}\,\hat{S}_{pw}(\omega)\boldsymbol{n}_{pw}^T\,\boldsymbol{x}_0\right]\, e^{-j\frac{\omega}{c}\boldsymbol{n}_{pw}^T\,\boldsymbol{x}_0}$$

WFS: Examples for driving functions

- **Spherical wave:**

$$S_{sw}(\boldsymbol{x},\omega) = \hat{S}_{sw}(\omega)\,\frac{e^{-j\frac{\omega}{c}\|\boldsymbol{x}-\boldsymbol{x}_s\|}}{\|\boldsymbol{x}-\boldsymbol{x}_s\|}$$

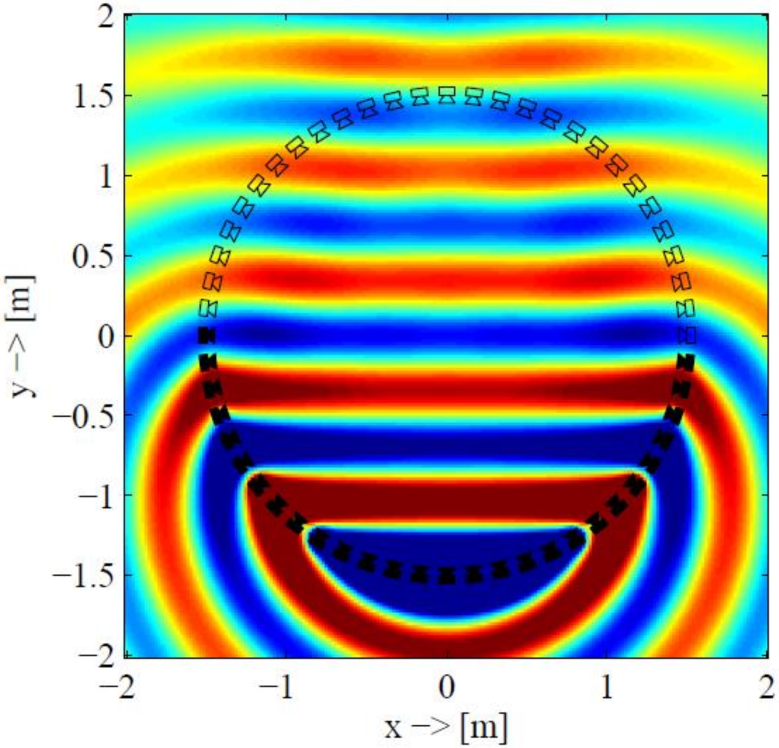
$$\frac{\partial}{\partial \boldsymbol{n}} S_{sw}(\boldsymbol{x},\omega) = -\frac{[\boldsymbol{x}-\boldsymbol{x}_s]^T\,\boldsymbol{n}(\boldsymbol{x})}{\|\boldsymbol{x}-\boldsymbol{x}_s\|^2}\,\left[\frac{1}{\|\boldsymbol{x}-\boldsymbol{x}_s\|}+\frac{j\omega}{c}\right]\hat{S}_{sw}(\omega)\,e^{-j\frac{\omega}{c}\|\boldsymbol{x}-\boldsymbol{x}_s\|}$$

- resulting in the following driving function:

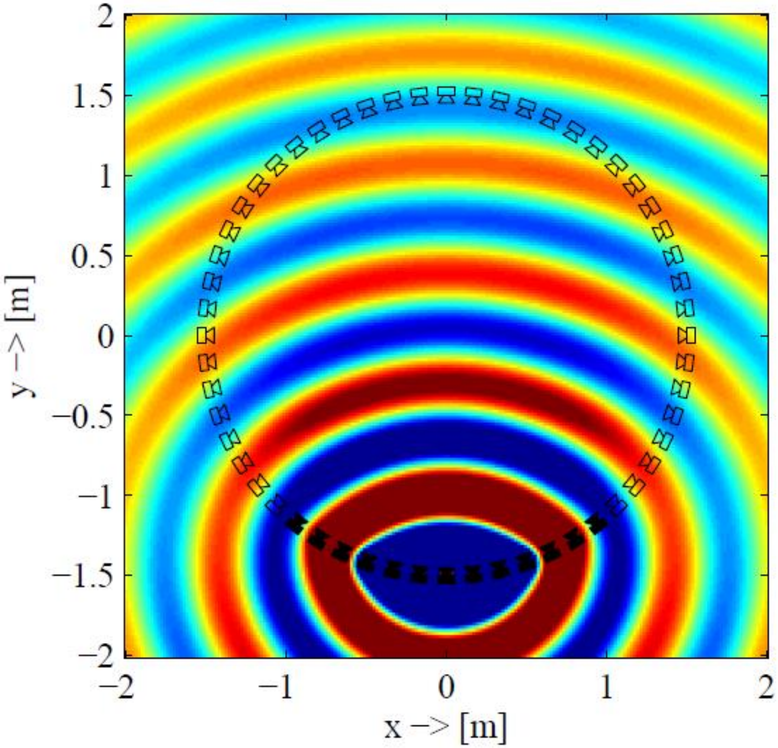
$$D(\boldsymbol{x}|\boldsymbol{x}_0,\omega) = -2\,a_{sw}(\boldsymbol{x}_0)\,\sqrt{\frac{2\,\pi\|\boldsymbol{x}-\boldsymbol{x}_0\|}{j\frac{\omega}{c}}}\,\frac{[\boldsymbol{x}_0-\boldsymbol{x}_s]^T\,\boldsymbol{n}(\boldsymbol{x}_0)}{\|\boldsymbol{x}_0-\boldsymbol{x}_s\|^2}\,\left[\frac{1}{\|\boldsymbol{x}_0-\boldsymbol{x}_s\|}+\frac{j\omega}{c}\right]\hat{S}_{sw}(\omega)\,e^{-j\frac{\omega}{c}\|\boldsymbol{x}_0-\boldsymbol{x}_s\|}$$

TECHNISCHE UNIVERSITÄT DARMSTADT

- **Plane & Spherical waves generated with a circular loudspeaker array:**



(a) plane wave ($f_{\mathrm{pw}} = 500$ Hz, $\mathbf{n}_{\mathrm{pw}} = [0\,1]^T$)



(b) spherical wave ($f_{\mathrm{sw}} = 500$ Hz, $\mathbf{x}_S = [0\ \ -2]^T$ m)

WFS: Dependency of the driving function

- Driving function (signals of the WFS loudspeakers) depend on the loudspeaker position $\boldsymbol{x}_0$ and the **listener position** $\boldsymbol{x}$ within the plane *L*.

$$D(\boldsymbol{x}|\boldsymbol{x}_0,\omega)$$

- However, the loudspeaker signals should be independent of the listeners' position (otherwise a location tracking of listeners would be necessary!).

- Fortunately, the dependence on the listeners position is rather mild.

- Typically, the driving function is fixed for a fixed listener's position:

$$D_L(\boldsymbol{x}_0,\omega) = D(\boldsymbol{x}_L|\boldsymbol{x}_0,\omega)$$

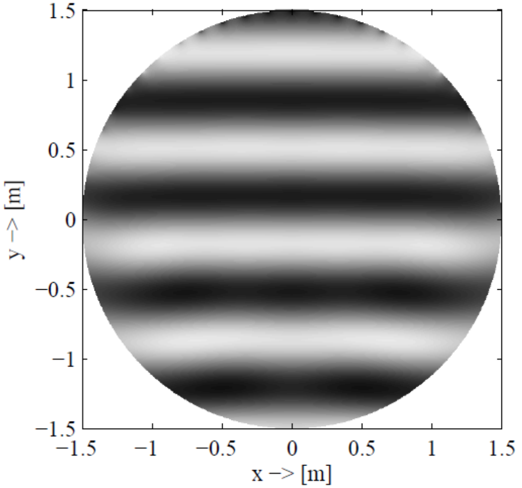
- In case the listener is close to the loudspeakers => rather big error.

WFS: Artifacts / aliasing

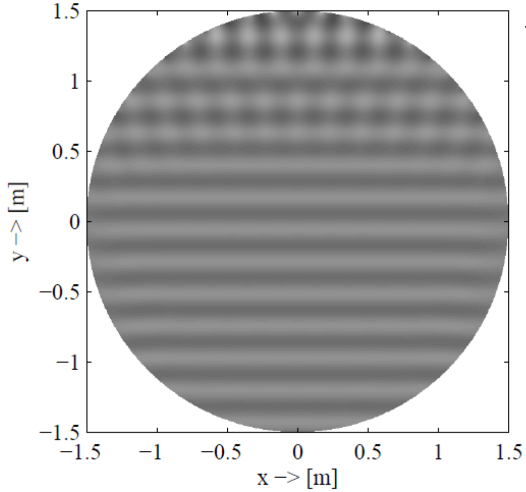
TECHNISCHE UNIVERSITÄT DARMSTADT

- WFS is generated with loudspeakers at discrete positions.
- => Spatial aliasing occurs. However, the analysis of the aliasing is rather complicated and dependent on the specific geometry / aperture of the WFS system.
- In general:
  - Spatial aliasing increases with increasing frequency, i.e., bandwidth
  - Ex.: For typical loudspeaker distances of 10-30 cm => Aliasing > 1kHz
  - However, listeners are not too sensitive to spatial aliasing, typically a spatial coloration occurs.
  - Pre-filtering with the term $\sqrt{\frac{j\omega}{c}}$ in the driving functions should only be performed below the spatial aliasing frequency (~1 kHz).
  - WFS systems with monopoles generate sound field outside L or V => in case of a WFS in a room => reflections from the wall generate artifacts.

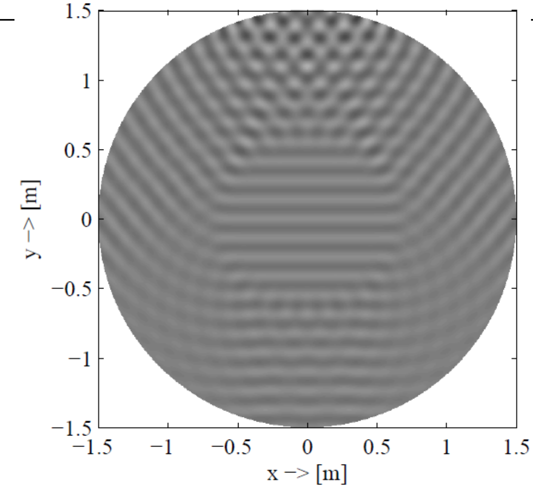
WFS: Spatial aliasing: Plane waves generated by circ. array



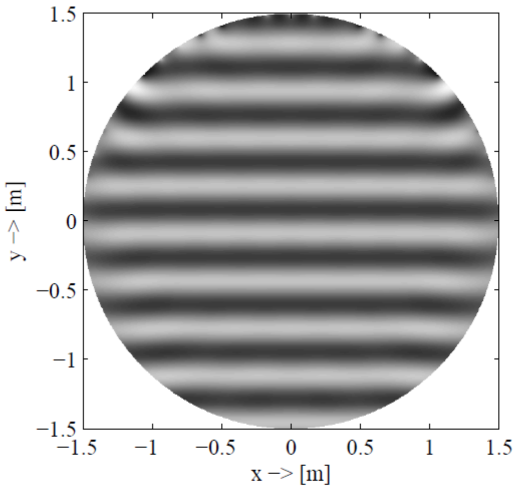
(a) WFS ($f_{\mathrm{pw}} = 500$ Hz)



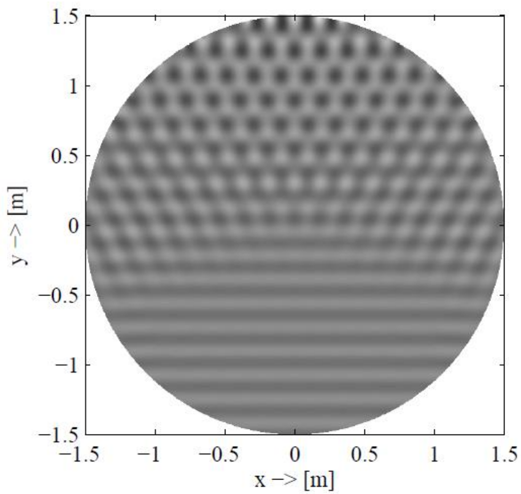
(e) WFS ($f_{\mathrm{pw}} = 1500$ Hz)



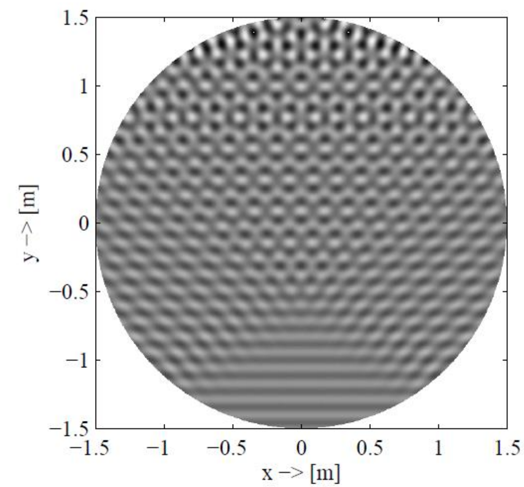
(d) HOA ($f_{\mathrm{pw}} = 2500$ Hz)



(c) WFS ($f_{\mathrm{pw}} = 1000$ Hz)



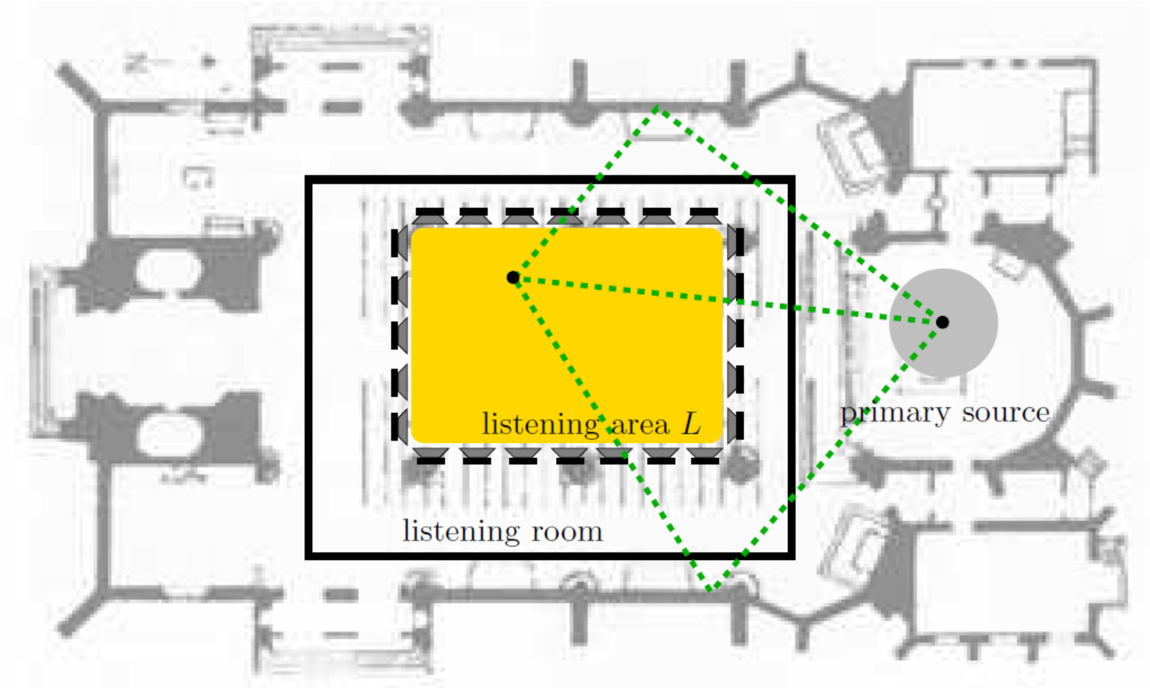
(a) WFS ($f_{\mathrm{pw}} = 2000$ Hz)



(e) WFS ($f_{\mathrm{pw}} = 3000$ Hz)

WFS: Example of virtual acoustics

- Reproduction of a sound field of a church:
  - Virtual scene model is based on the decomposition of the sound signal in the church into plane waves.
  - Image model of reverberation: Sound in the listening area is a multitude of point sources. => Superposition of plane waves.

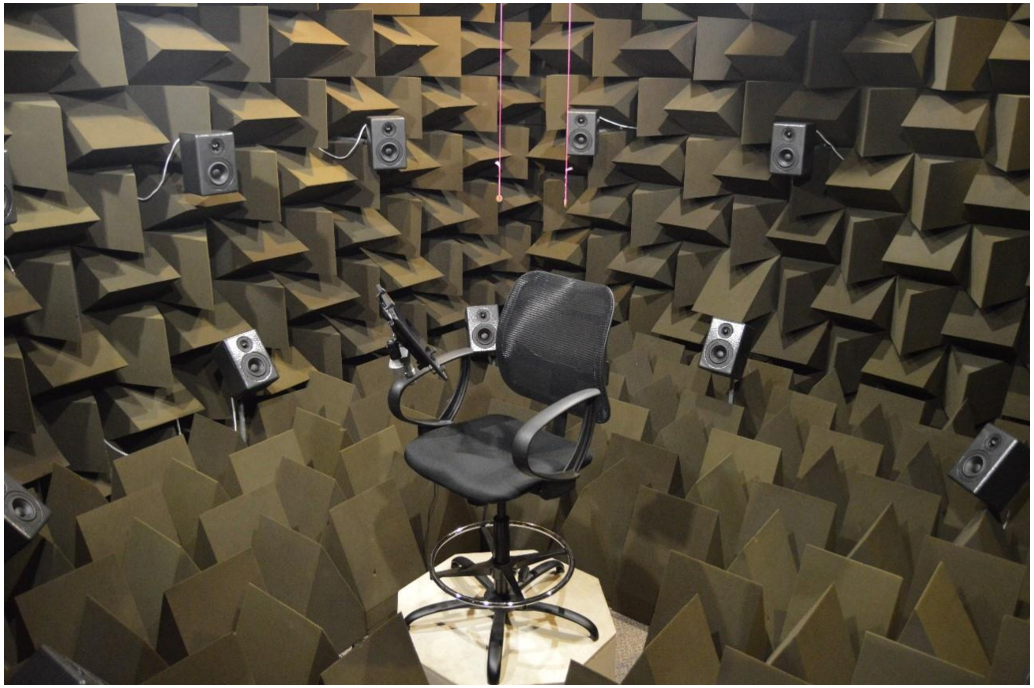


Higher Order Ambisonics

HOA: Examples

Source: https://www.southampton.ac.uk

Source: http://sites.psu.edu/spral/current-projects/auras/



HOA: Setup

- **The basic idea:**
  - Decomposition of a wave field into spherical modes.
  - Performing a transform into these modes.
  - Mode description independent of the recording (microphone) or the generation (loudspeaker) setup.

  - Order determines the precision of the decomposition / generation.
  - For an order M there are $(M+1)^2$ microphones / loudspeakers necessary for its recording / generation.

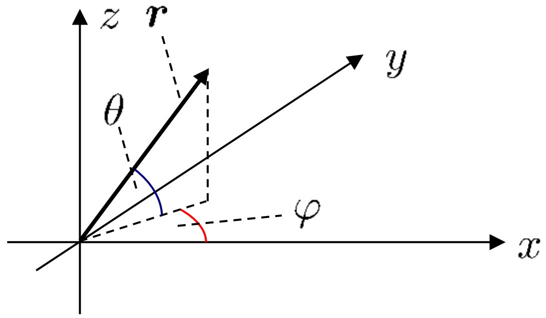
  - The setup is generally used to reproduce a 3D sound field precisely at a sweet spot within a loudspeaker setup.

  - The higher the order is, the bigger is the sweet spot or the max. frequency given the radius of the target sphere.

HOA: Theoretical sound field decomposition

- Definition of a subspace Ω_1 where the sources are located and a subspace Ω_2 where no sources are located (listing area).

- General setup: All sources outside a radius, R_2 i.e., the listening area is inside.



- Comparable to sinus components for the Fourier transform, the sound field can be decomposed into “spherical harmonics”

HOA: Theoretical sound field decomposition

- The decomposition is expressed as follows:

HOA components

Spherical functions

$$p(\boldsymbol{r},\omega) = \sum_{m=0}^{\infty} i^m\, j_m(kr) \sum_{n=0}^{m} \sum_{\sigma=\pm 1} B_{mn}^{\sigma}(\omega)\, Y_{mn}^{\sigma}(\varphi,\theta) \qquad r = |\boldsymbol{r}| \qquad k = \frac{\omega}{c}$$

- with the following spherical functions:

$j_m(kr)$ Spherical Bessel function of first kind

$$Y_{mn}^{\sigma}(\varphi,\theta) = \sqrt{(2m+1)\,\epsilon_n\,\frac{(m-n)!}{(m+n)!}}\; P_{mn}(\sin\theta)\,\mathrm{sc}(n\,\varphi)$$

$$\mathrm{sc}(n\,\varphi) = \begin{cases} \cos(n\,\varphi) & : \quad \sigma = +1 \\ \sin(n\,\varphi) & : \quad \sigma = -1 \end{cases} \qquad \epsilon_n = \begin{cases} 1 & : \quad n = 0 \\ 2 & : \quad n > 0 \end{cases}$$

$$P_{mn}(\sin\theta) = \frac{d^n\, P_m(\sin\theta)}{d(\sin\theta)^n}$$ Legendre function n. order, m. degree

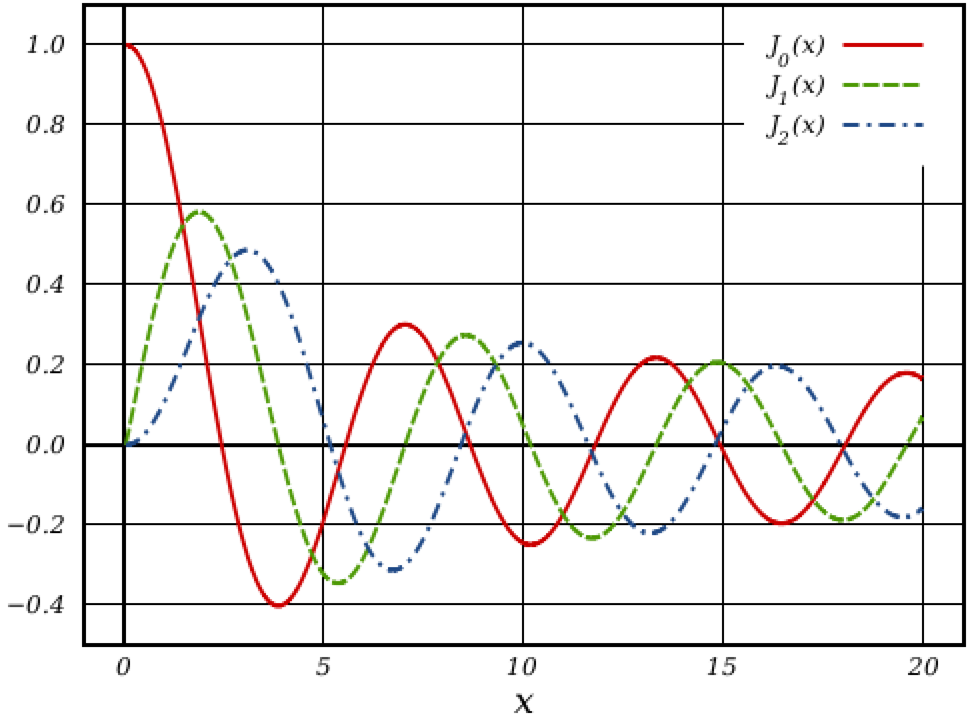
$$P_n(z) = \frac{1}{2^n\, n!}\,\frac{d^n}{dz^n}(z^2-1)^n$$ Legendre function n. order, 1st degree

- and the HOA components $B_{mn}^{\sigma}(\omega)$ i.e., a universal format to encode a sound field.

HOA: Spherical Bessel functions

$j_m(kr)$ Spherical Bessel function of first kind

$$j_m(x) = \frac{1}{2\pi} \int_{-pi}^{\pi} e^{i(m\tau - x \sin \tau)} \, d\tau$$

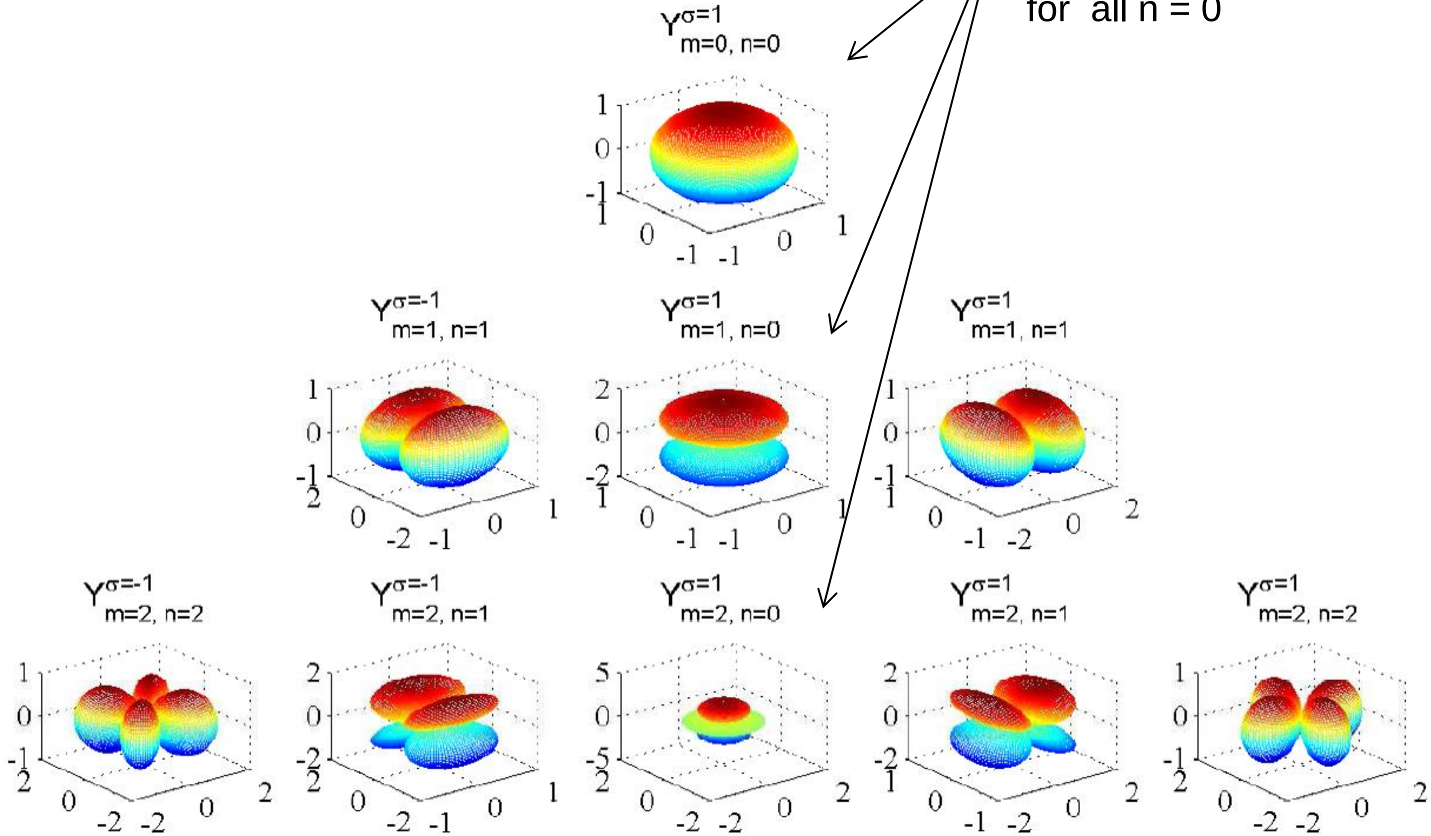


Source: Wikipedia

HOA: Theoretical sound field decomposition

- Plot of the spherical functions:

Only one component for all n = 0

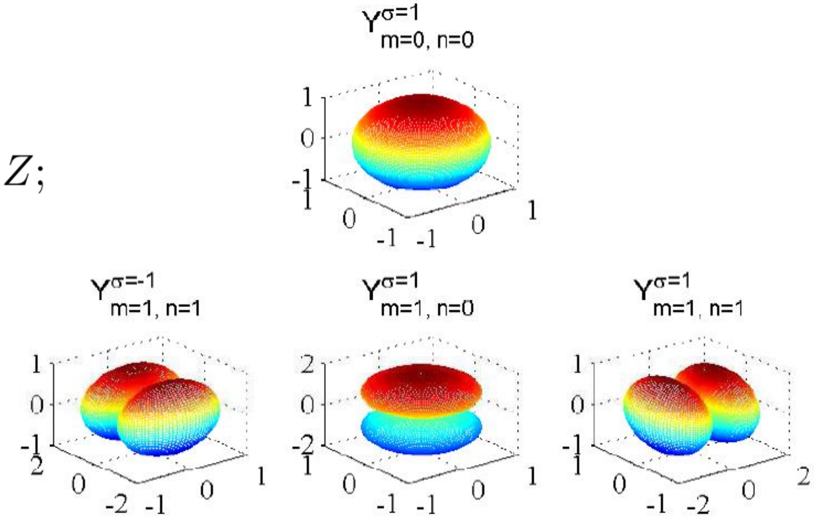


HOA: Theoretical sound field decomposition

- Truncation of the infinite sum to M = 1, results in the well-known B-Format according to Gerzon, i.e., one onmi and three dipole components:

$$B_{00}^{1} = W;$$

$$B_{11}^{1} = X; \qquad B_{11}^{-1} = Y; \quad B_{10}^{1} = Z;$$



- When limiting the sum, a truncation / reconstruction error occurs. The error can assumed to be negligible for $kr \leq M$

=> Sweet spot:

$$r_{\max} = \frac{M\,c}{\omega_0}$$

Given the radius of sphere of targeted reconstruction, the reconstruction is only valid up to:

$$\omega_{\max} = \frac{M\,c}{r_0}$$

HOA: Encoding

- The HOA components can in general be calculated as follows:

$$B_{mn}^{\sigma}(\omega) = EQ(k\, r_M)\, \frac{1}{4\pi\, r_M^2} \int_{\varphi=0}^{2\pi} \int_{\theta=-\frac{\pi}{2}}^{\frac{\pi}{2}} p(r_M, \varphi, \theta, \omega)\, Y_{mn}^{\sigma}(\varphi, \theta) \cos\theta \, d\theta \, d\varphi$$

$$EQ(k\, r_M) = \frac{1}{i^m\, j_m(k\, r_M)}$$: normalization

$p(r_M, \varphi, \theta, \omega)$: recorded sound field over a shere with radius r_M

- Important to note: The HOA components are totally independent of the recording setup, including the radius of the recorded sphere.

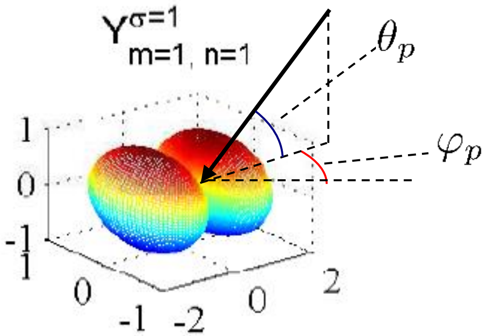
HOA: Encoding for specific cases

- For specific cases, i.e., plane and spherical waves the HOA components can be directly calculated based on the spherical functions: $Y^{\sigma}_{mn}(\varphi, \theta)$

- 1) Plane waves arriving from a direction φ_p, θ_p with amplitude O_p

$$\boxed{B^{\sigma}_{mn}(\omega) = \frac{O_p}{4\pi} Y^{\sigma}_{mn}(\varphi_p, \theta_p)}$$

$B^{\sigma}_{mn}(\omega)$ can be directly calculated by evaluating $Y^{\sigma}_{mn}(\varphi_p, \theta_p)$ at the direction of arrival of the plane wave.



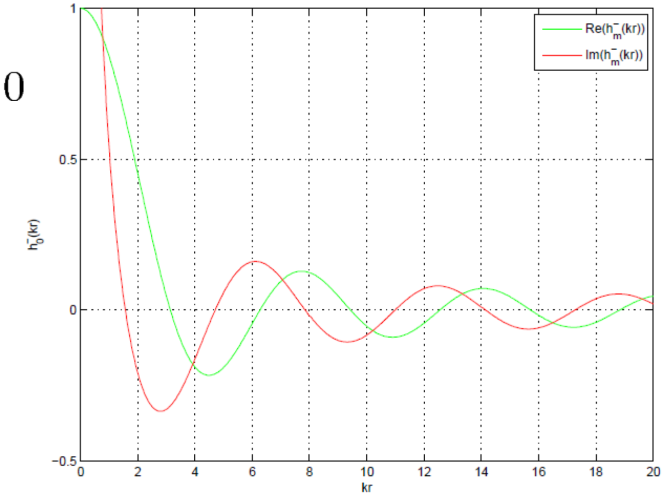
- 2) Spherical waves where the point of origin $\boldsymbol{r}_s = (r_s, \varphi_s, \theta_s)$ is placed outside of the listening area with radius R_2, i.e. $r_s > R_2$ with the amplitude O_s

$$\boxed{B^{\sigma}_{mn}(\omega) = \frac{O_s}{4\pi} \, i^{-(m+1)} \, \frac{h^{-}_{m}(k\, r_s)}{k} \, Y^{\sigma}_{mn}(\varphi_s, \theta_s)}$$

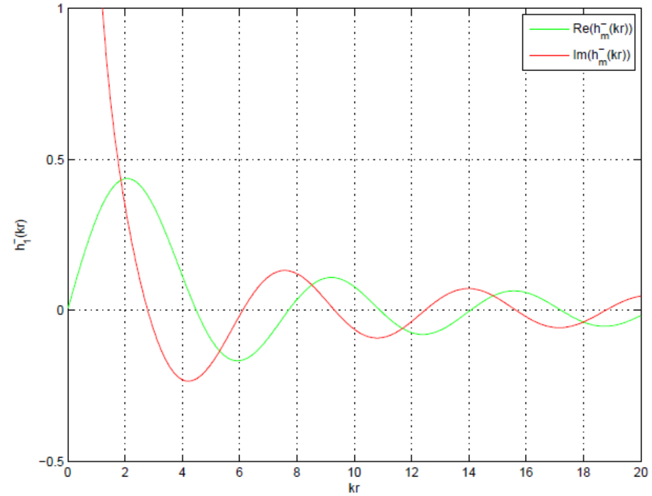
$h^{-}_{m}(k\, r_s)$: spherical Hankel functions of second kind: $h^{-}_{m}(k\, r_s) = H^{(2)}_{m}(k\, r_s)$

HOA: Spherical Hankel functions of second kind

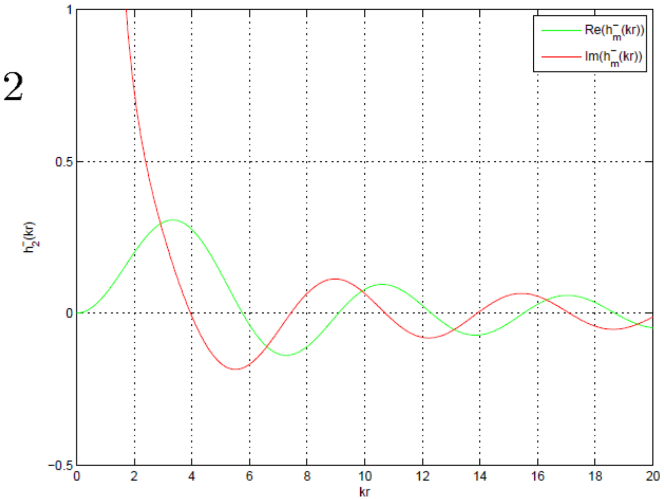
$m = 0$



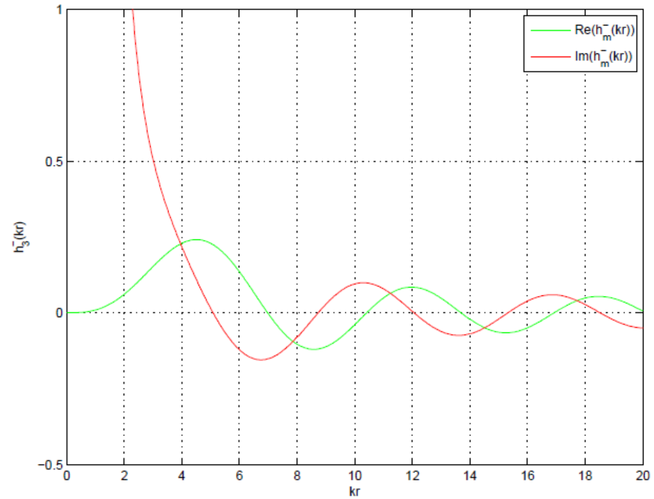
$m = 1$



$m = 2$



$m = 3$



HOA: Encoding by spatial sampling

- **Task:** Record a sound field and calulate the HOA components
- Concept: Using N_M microphones positioned on a sphere with radius r_M, where q is the position of the q^{th} microphone.

$$p_q(\omega) = p(r_M, \varphi_q, \theta_q, \omega)$$
$$p_q(\omega) = \sum_{m=0}^{\infty} i^m \, j_m(k\, r_M) \sum_{n=0}^{m} \sum_{\sigma=\pm 1} B_{mn}^{\sigma}(\omega)\, Y_{mn}^{\sigma}(\varphi_q, \theta_q)$$

HOA: Encoding by spatial sampling

- For N_M microphones one can obtain N_M equations; approximating the sound field up to order *M*.

$$p_q(\omega) = \sum_{m=0}^{M} i^m \, j_m(k\, r_M) \sum_{n=0}^{m} \sum_{\sigma=\pm 1} B_{mn}^{\sigma}(\omega)\, Y_{mn}^{\sigma}(\varphi_q, \theta_q) \qquad N_M \geq (M+1)^2$$

- In Matrix vector notation these equations can be written as follows:

$$\boldsymbol{p} = \boldsymbol{Y}_M \boldsymbol{W}_M \boldsymbol{b}$$

with:

$$\boldsymbol{Y}_M = \begin{bmatrix} Y_{00}^1(\varphi_1, \theta_1) & Y_{10}^1(\varphi_1, \theta_1) & \cdots & Y_{MM}^{-1}(\varphi_1, \theta_1) \\ Y_{00}^1(\varphi_2, \theta_2) & Y_{10}^1(\varphi_2, \theta_2) & \cdots & Y_{MM}^{-1}(\varphi_2, \theta_2) \\ \vdots & \vdots & \ddots & \vdots \\ Y_{00}^1(\varphi_{N_M}, \theta_{N_M}) & Y_{10}^1(\varphi_{N_M}, \theta_{N_M}) & \cdots & Y_{MM}^{-1}(\varphi_{N_M}, \theta_{N_M}) \end{bmatrix}$$

$$\boldsymbol{b} = \begin{bmatrix} B_{00}^1(\omega) & B_{10}^1(\omega) & B_{11}^1(\omega) & B_{11}^{-1}(\omega) & \cdots & B_{MM}^{-1}(\omega) \end{bmatrix}^{\mathrm{T}}$$

$$\boldsymbol{W}_M = \mathrm{diag}\left\{ \begin{bmatrix} j_o(k\, r_M) & i^1 j_1(k\, r_M) & i^1 j_1(k\, r_M) & i^1 j_1(k\, r_M) \cdots & i^M j_M(k\, r_M) \end{bmatrix} \right\}$$

HOA: Encoding by spatial sampling

- Size analysis of vectors and matices:

$$\boldsymbol{p} = \boldsymbol{Y}_M \, \boldsymbol{W}_M \, \boldsymbol{b}$$

$\boldsymbol{b}$: Vector of length $(M+1)^2$

$\boldsymbol{W}_M$: Matrix of size $(M+1)^2$ x $(M+1)^2$

$\boldsymbol{Y}_M$: Matrix of size N_M x $(M+1)^2$

- Least squares solutions:

$$\boldsymbol{W}_M \, \hat{\boldsymbol{b}} = (\boldsymbol{Y}_M^{\mathrm{T}} \boldsymbol{Y}_M)^{-1} \boldsymbol{Y}_M^{\mathrm{T}} \boldsymbol{p}$$

$$\Rightarrow \boxed{\hat{\boldsymbol{b}} = \boldsymbol{E}_M \, (\boldsymbol{Y}_M^{\mathrm{T}} \boldsymbol{Y}_M)^{-1} \boldsymbol{Y}_M^{\mathrm{T}} \boldsymbol{p}}$$

With the diagonal normalization matrix:

$$\boldsymbol{E}_M = \mathrm{diag}\left\{ \left[\frac{1}{j_o(k\,r_M)} \quad \frac{1}{i^1 j_1(k\,r_M)} \quad \frac{1}{i^1 j_1(k\,r_M)} \quad \frac{1}{i^1 j_1(k\,r_M)} \cdots \quad \frac{1}{i^M j_M(k\,r_M)} \right] \right\}$$

HOA: Encoding by spatial sampling

- If the spatial distribution of the microphones fulfills the orthonormality property:

$$(\boldsymbol{Y}_M^{\mathrm{T}}\,\boldsymbol{Y}_M) = \boldsymbol{1}$$

$$\hat{\boldsymbol{b}} = \boldsymbol{E}_M\,(\boldsymbol{Y}_M^{\mathrm{T}}\,\boldsymbol{Y}_M)^{-1}\,\boldsymbol{Y}_M^{\mathrm{T}}\,\boldsymbol{p} \quad \Rightarrow \quad \hat{\boldsymbol{b}} = \boldsymbol{E}_M\,\boldsymbol{Y}_M^{\mathrm{T}}\,\boldsymbol{p}$$

- which results in the discrete version of the general coding equation:

$$\boxed{B_{mn}^{\sigma}(\omega) = \frac{1}{j_m(k\,r_M)}\,\frac{1}{N_M}\sum_{q=0}^{N_M} p_q(\omega)\,Y_{mn}^{\sigma}(\varphi_q,\theta_q)}$$

- Reminder: general coding equation:

$$B_{mn}^{\sigma}(\omega) = EQ(k\,r_M)\,\frac{1}{4\pi\,r_M^2}\int_{\varphi=0}^{2\pi}\int_{\theta=-\frac{\pi}{2}}^{\frac{\pi}{2}} p(r_M,\varphi,\theta,\omega)\,Y_{mn}^{\sigma}(\varphi,\theta)\cos\theta\,d\theta\,d\varphi$$

$$EQ(k\,r_M) = \frac{1}{i^m\,j_m(k\,r_M)}$$: normalization

HOA: Encoding by spatial sampling
Thoughts about the recording setup

- The matrix ϵ quantifies the "non-orthogonality" of the microphone distribution (i.e., recording setup, e.g., the „Eigenmike"):

$$\boldsymbol{\epsilon} = \mathbf{1} - \frac{1}{N_M}(\boldsymbol{Y}_M^{\mathrm{T}}\,\boldsymbol{Y}_M)$$

$$p_q(\omega) = \sum_{m=0}^{\infty} i^m\, j_m(k\, r_M) \sum_{n=0}^{m} \sum_{\sigma=\pm 1} B_{mn}^{\sigma}(\omega)\, Y_{mn}^{\sigma}(\varphi_q, \theta_q)$$

- Lower spatial aliasing with **reduced radius** r_M :

  The Bessel functions act as a low-frequency filter which reduces spatial sampling like an anti-aliasing filter.
  The cutoff frequency increases with r_M. Therefore, decreasing the radius will minimize the spatial aliasing.

- Better conditioned inverse matrix with **increased radius** r_M:

$$(\boldsymbol{Y}_M^{\mathrm{T}}\,\boldsymbol{Y}_M)$$

HOA: Decoding

- **HOA decoding task:** Calculation of the loudspeaker signals of an HOA loudspeaker array with N_L loudspeakers.
- Loudspeaker positions:

$$\boldsymbol{r}_{L,l} = (r_{L,l}, \varphi_{L,l}, \theta_{L,l})$$

- Sound pressure of the l-th loudspeaker at position $\boldsymbol{r}$ where $s_l(\omega)$ is the driving signal: $s_l(\omega)\, p_l(\boldsymbol{r}, \omega)$

- HOA decomposition of the signal radiated by the l-th loudspeaker signal. I.e., $L^{\sigma}_{l,mn}(\omega)$ is known by the measurement of the loudspeaker array.

$$p_l(\boldsymbol{r}, \omega) = \sum_{m=0}^{\infty} i^m\, j_m(k\, r) \sum_{n=0}^{m} \sum_{\sigma=\pm 1} L^{\sigma}_{l,mn}(\omega)\, Y^{\sigma}_{mn}(\varphi_q, \theta_q)$$

1 loudspeaker

$$p_{\mathrm{LS}}(\boldsymbol{r}, \omega) = \sum_{l=1}^{N_L} s_l(\omega)\, p_l(\boldsymbol{r}, \omega)$$

All loudspeakers

$$= \sum_{l=1}^{N_L} s_l(\omega) \sum_{m=0}^{\infty} i^m\, j_m(k\, r) \sum_{n=0}^{m} \sum_{\sigma=\pm 1} L^{\sigma}_{l,mn}(\omega)\, Y^{\sigma}_{mn}(\varphi_q, \theta_q)$$

TECHNISCHE UNIVERSITÄT DARMSTADT

- **Mode matching**

  Target: calculate the excitation signal $s_l(\omega)$ of the loudspeakers in order to approximate the sound field given by an HOA description $B_{mn}^{\sigma}(\omega)$

$$p(\boldsymbol{r},\omega) = \sum_{m=0}^{\infty} i^m j_m(kr) \sum_{n=0}^{m} \sum_{\sigma=\pm 1} B_{mn}^{\sigma}(\omega) Y_{mn}^{\sigma}(\varphi,\theta)$$

General description

$$p_{\mathrm{LS}}(\boldsymbol{r},\omega) = \sum_{m=0}^{\infty} i^m j_m(k\,r) \sum_{n=0}^{m} \sum_{\sigma=\pm 1} \sum_{l=1}^{N_L} s_l(\omega) L_{l,mn}^{\sigma}(\omega)\, Y_{mn}^{\sigma}(\varphi_q,\theta_q)$$

LS description

$$\boxed{=> B_{mn}^{\sigma}(\omega) \stackrel{!}{=} \sum_{l=1}^{N_L} s_l(\omega)\, L_{l,mn}^{\sigma}(\omega) \qquad \text{equal to: } \boldsymbol{b} = \boldsymbol{L}\,\boldsymbol{s}}$$

$$\boldsymbol{L} = \begin{bmatrix} L_{1,00}(\varphi_1,\theta_1) & L_{2,00}(\varphi_2,\theta_2) & \cdots & L_{N_L,00}(\varphi_{N_L},\theta_{N_L}) \\ L_{1,10}(\varphi_1,\theta_1) & L_{2,10}(\varphi_2,\theta_2) & \cdots & L_{N_L,10}(\varphi_{N_L},\theta_{N_L}) \\ \vdots & \vdots & \ddots & \vdots \\ L_{1,MM}(\varphi_1,\theta_1) & L_{2,MM}(\varphi_2,\theta_2) & \cdots & L_{N_L,MM}(\varphi_{N_L},\theta_{N_L}) \end{bmatrix}$$

$$\boldsymbol{s} = \left[s_1(\omega)\; s_2(\omega)\; \cdots\; s_{N_L}(\omega)\right]^{\mathrm{T}}$$: excitation signal vector

$$\boldsymbol{b} = \left[B_{00}^{1}(\omega) \quad B_{10}^{1}(\omega) \quad B_{11}^{1}(\omega) \quad B_{11}^{-1}(\omega) \quad \cdots \quad B_{MM}^{-1}(\omega)\right]^{\mathrm{T}}$$

HOA: Decoding

TECHNISCHE UNIVERSITÄT DARMSTADT

- Size analysis of vectors and matices:

$$\boldsymbol{b} = \boldsymbol{L}\,\boldsymbol{s}$$

Vector of length N_L

Matrix of size $(M+1)^2$ x N_L

- Least squares solutions:

$$\hat{\boldsymbol{s}} = \boldsymbol{L}^{\mathrm{T}}\,(\boldsymbol{L}\,\boldsymbol{L}^{\mathrm{T}})^{-1}\,\boldsymbol{b}$$

- Decoding matrix:

$$\boldsymbol{D} = \boldsymbol{L}^{\mathrm{T}}\,(\boldsymbol{L}\,\boldsymbol{L}^{\mathrm{T}})^{-1}$$

- Regular loudspeaker distribution on a sphere of radius r_L
- Regular distribution in case the following equation is fulfilled: $(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}) = \mathbf{1}$
- Then the decoding matrix reduces to $\boldsymbol{D} = \boldsymbol{L}^{\mathrm{T}}$

HOA: Decoding

- Special cases for loudspeakers radiating specific types of waves

  - Spherical waves:

$$\boldsymbol{L} = \boldsymbol{W}_L^{\mathrm{T}} \boldsymbol{Y}_L$$

$$\hat{\boldsymbol{s}} = \left(\boldsymbol{W}_L^{\mathrm{T}} \boldsymbol{Y}_L \right)^{\mathrm{T}} \boldsymbol{b}$$

$$\boldsymbol{Y}_L = \begin{bmatrix} Y_{00}^1(\varphi_1, \theta_1) & Y_{00}^1(\varphi_2, \theta_2) & \cdots & Y_{00}^1(\varphi_{N_L} \theta_{N_L}) \\ Y_{10}^1(\varphi_1, \theta_1) & Y_{10}^1(\varphi_2, \theta_2) & \cdots & Y_{10}^1(\varphi_{N_L}, \theta_{N_L}) \\ \vdots & \vdots & \ddots & \vdots \\ Y_{MM}^{-1}(\varphi_1, \theta_1) & Y_{MM}^{-1}(\varphi_2, \theta_2) & \cdots & Y_{MM}^{-1}(\varphi_{N_L}, \theta_{N_L}) \end{bmatrix}$$

$$\boldsymbol{W}_L = \mathrm{diag}\left\{ [-i \quad - \frac{h_1^-(kr_L}{k} i^{-M-1} \quad - \frac{h_M^-(kr_L}{k}] \right\}$$

$$w_m = i^{-(m+1)} \frac{h_m^-(k\, r_s)}{k}$$

  - Plane waves:

$$\boldsymbol{L} = \boldsymbol{Y}_L$$

$$\hat{\boldsymbol{s}} = \boldsymbol{Y}_L^{\mathrm{T}} \boldsymbol{b}$$

Summary

- Introduction of WFS and HOA for sound field generation of loudspeaker arrays.

- Typically specific fields of applications:

- HOA: 3D sound in a sweet spot
  => more precise reproduction

- WFS: 2D in larger plane => listening area
  => more natural sound experience

References

[1] R. Nicol , “Sound spatialization by higher order ambisonics: Encoding and decoding a sound scene in practice from a theoretical point of view,” in *Proceedings of the 2nd International Symposium on Ambisonics and Spherical Acoustics*, Paris, France (May 6–7, 2010)

[2] S. Spors, J. Ahrens. "A comparison of wave field synthesis and higher-order ambisonics with respect to physical properties and spatial sampling." Audio Engineering Society Convention 125. Audio Engineering Society, 2008.

[3] S. Spors, R. Rabenstein, and Jens Ahrens. "The theory of wave field synthesis revisited." 124th AES Convention. 2008.

[4] R. Rabenstein, S. Spors, and P. Steffen. "Wave field synthesis techniques for spatial sound reproduction." Topics in Acoustic Echo and Noise Control 5 (2006): 517-545.

[5] A. Sontacchi. Dreidimensionale Schallfeldreproduktion für Lautsprecher-und Kopfhöreranwendungen. Dissertation, Uni Graz, 2003.

Final Hints

- **In case of questions**:
  - Please contact me by eMail.

- **Lecture in Summer Term**: „Adaptive Filters“: Theory and applications