POLITECNICO DI MILANO

ISPL
Image and Sound Processing Lab

SMC Lab

Rethinking Fourier Acoustics

Augusto Sarti, Fabio Antonacci, Lucio Bianchi

DEIB – Politecnico di Milano, Italy

augusto.sarti@polimi.it, fabio.antonacci@polimi.it, lucio.bianchi@polimi.it

POLITECNICO DI MILANO

ISPL
Image and Sound Processing Lab

SMC Lab

Rethinking Fourier Acoustics

Part 0: Preliminaries

Augusto Sarti, Fabio Antonacci, Lucio Bianchi

Outline

- The acoustic field
- The D'Alambert equation and the Helmholtz equation
- General solutions of the homogeneous equation
  - Plane waves and spherical waves
- Solving the inhomogeneous equation
  - The Green's function
- Boundary conditions
- Integral representations for the acoustic field
  - The Kirchhoff-Helmholtz integral equation
  - The Single Layer Potential
- Geometrical representations for the acoustic field
  - The Eikonal equation and its solution
  - Ray geometry and examples

The acoustic field

- An **acoustic field** is a real-valued scalar function whose domain extends in
  - Time
  - Space
- The acoustic field is invariant under change of spatial coordinates
  - The value of the acoustic field in a point is independent on the coordinate system adopted to represent that point
  - Only the mathematical expression for the field as a function of the spatial coordinates varies
- The **Laplace operator** describes the curvature of the function in space

$$\nabla^2 p(\mathbf{r}, t) = \nabla \cdot (\nabla p(\mathbf{r}, t))$$

Homogeneous wave equations

- Source-free volume
- The acoustic field must satisfy the **homogeneous D'Alambert equation** (wave equation)

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

- Assuming time-harmonic behavior

$$p(\mathbf{r},t) = P(\mathbf{r},\omega)e^{-j\omega t}$$

- In the frequency domain the acoustic field must satisfy the **homogeneous Helmholtz equation**

$$\nabla^2 P(\mathbf{r},\omega) + \left(\frac{\omega}{c}\right) P(\mathbf{r},\omega) = 0$$

Plane wave solutions

- Look for solutions to the homogeneous Helmholtz equation in Cartesian coordinates
  - Complex exponential function

$$P(\mathbf{r}, \omega) = e^{j<\mathbf{k},\mathbf{r}>}$$

  - $\mathbf{k} = [k_x, k_y, k_z]^T$ is the **wavenumber vector**

- The complex exponential function is a solution only if the wavenumber vector satisfies the **dispersion relation**

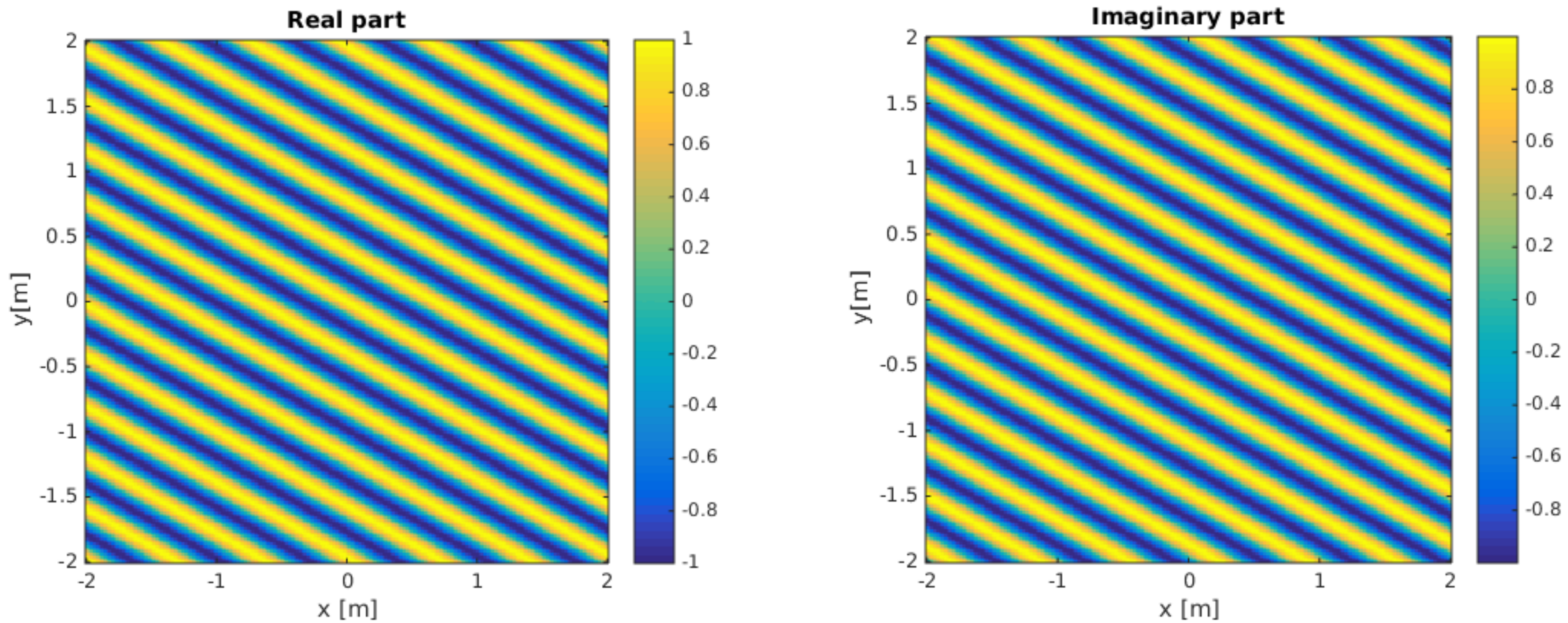
$$||\mathbf{k}||^2 = \left(\frac{\omega}{c}\right)^2$$

Propagating plane waves

- Real wavenumber vector $\mathbf{k} \in \mathbb{R}^3$
- **Wavefront** = set of points of constant phase

$$\angle P(\mathbf{r}, \omega) = \text{constant} \implies <\mathbf{k}, \mathbf{r}>= C, \quad C \in \mathbb{R}$$

  - Points of constant phase form a plane orthogonal to $\mathbf{k}$
  - The unit vector $\hat{\mathbf{k}} = \mathbf{k}/||\mathbf{k}||$ identifies the **direction of arrival** of the plane wave



Evanescent plane waves

- Assume two components of the wavenumber vector to be real

$$k_x, k_y \in \mathbb{R}$$

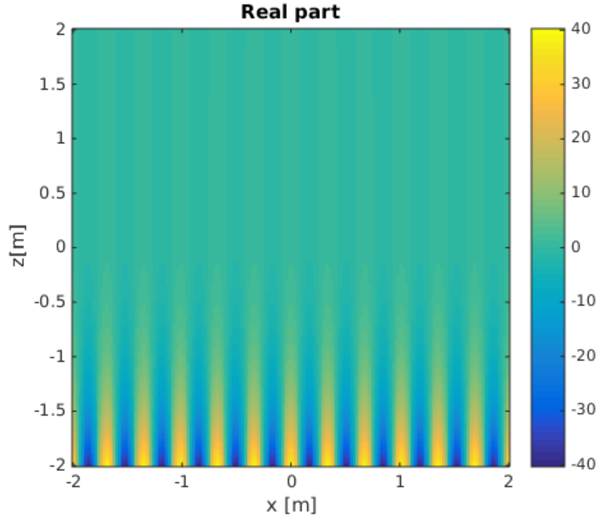
  - k_z must satisfy

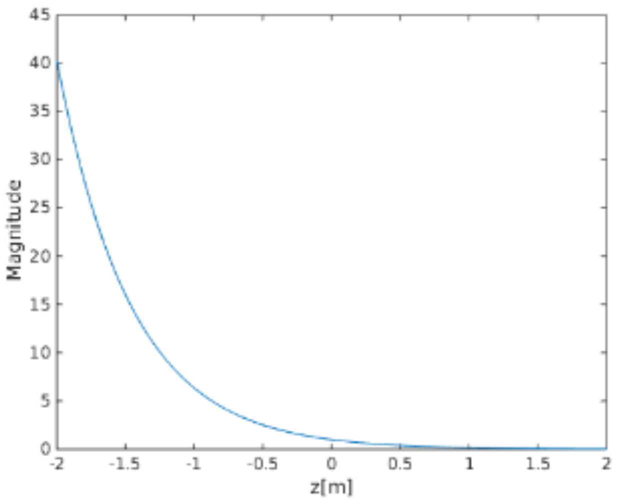
$$k_z^2 = \left(\frac{\omega}{c}\right)^2 - k_x^2 - k_y^2$$

- If $k_x^2 + k_y^2 \leq (\omega/c)^2$
  - $k_z \in \mathbb{R}$: propagating plane wave
- If $k_x^2 + k_y^2 > (\omega/c)^2$
  - $k_z = j\zeta,\ \zeta \in \mathbb{R}^+$: evanescent plane wave

$$P(\mathbf{r},\omega) = e^{j(k_x x + k_y y + j\zeta z)} = e^{j(k_x x + k_y y)} e^{-\zeta z}$$

  - The acoustic field exhibits exponential decay along the z axis





- Spherical coordinates
  - Range: r
  - Azimuth: ϕ
  - Co-elevation: θ
- Laplace operator in spherical coordinates

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

- Helmholtz equation in spherical coordinates

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial P}{\partial r}\right) + \frac{1}{r^2\sin(\theta)}\frac{\partial}{\partial\theta}\left(\sin(\theta)\frac{\partial P}{\partial\theta}\right) + \frac{1}{r^2\sin^2(\theta)}\frac{\partial^2 P}{\partial\phi^2} + \left(\frac{\omega}{c}\right)^2 P = 0$$

Spherical wave solutions

- Solutions to the homogeneous Helmholtz equation in spherical coordinates are obtained by separation of variables

$$P(\mathbf{r},\omega) = R(r)\Theta(\theta)\Phi(\phi)$$

- The angular dependency is usually summarized into the **spherical harmonic** function

$$\Theta(\theta)\Phi(\phi) = Y_l^m(\theta,\phi) = (-1)^m \sqrt{\frac{(2l+1)}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}} e^{jm\phi} \mathrm{P}_l^{|m|}(\cos(\theta))$$

  - **Frequency independent!**

- The radial dependency is given in terms of spherical Bessel or spherical Hankel functions

$$R(r) = R_1 j_l((\omega/c)r) + R_2 y_l((\omega/c)r)$$

$$R(r) = R_3 h_l^{(1)}((\omega/c)r) + R_4 h_l^{(2)}((\omega/c)r)$$

Inhomogeneous wave equations

- Volume with a source distribution
- The acoustic field must satisfy the **inhomogeneous D'Alambert equation**

$$\nabla^2 p(\mathbf{r},t) - \frac{1}{c^2}\frac{\partial^2 p(\mathbf{r},t)}{\partial t^2} = -\frac{\partial q(\mathbf{r},t)}{\partial t}$$

  - Flow per unit volume $q(\mathbf{r},t)$
- The corresponding **inhomogeneous Helmholtz equation** is

$$\nabla^2 P(\mathbf{r},\omega) + \frac{\omega^2}{c^2} P(\mathbf{r},\omega) = -j\omega Q(\mathbf{r},\omega)$$

The Green's function

- Solution to the inhomogeneous wave equations when the inhomogeneous term is a spatio-temporal impulse at $(\mathbf{r}', t')$
- Solution to the D'Alabmert equation

$$g(\mathbf{r}|\mathbf{r}', t) = \frac{1}{\|\mathbf{r}-\mathbf{r}'\|}\delta\left(t - \frac{\|\mathbf{r}-\mathbf{r}'\|}{c}\right)$$

- Solution to the Helmholtz equation

$$G(\mathbf{r}|\mathbf{r}', \omega) = \frac{e^{-j(\omega/c)\|\mathbf{r}-\mathbf{r}'\|}}{4\pi\|\mathbf{r}-\mathbf{r}'\|}$$

The boundary conditions

- Boundary conditions are imposed on solutions to the wave equations in order to consider the physical properties for the boundary of the considered domain
  - Homogeneous boundary conditions: stationary boundaries
  - Inhomogeneous boundary conditions: reacting boundaries
- Dirichlet boundary conditions
  - Conditions on the acoustic pressure field
- Neumann boundary conditions
  - Conditions on the directional derivative of the pressure field (particle velocity) in the direction normal to the boundary

$$p(\mathbf{r},\omega) = -\oint_{\partial\mathcal{V}} \left(G(\mathbf{r}|\mathbf{r}',\omega)\ \langle \nabla p(\mathbf{r},\omega), \hat{\mathbf{n}}(\mathbf{r}')\rangle|_{\mathbf{r}=\mathbf{r}'} + \right.$$

$$\left. -p(\mathbf{r}',\omega)\langle \nabla G(\mathbf{r}|\mathbf{r}',\omega), \hat{\mathbf{n}}(\mathbf{r}')\rangle\right) dA(\mathbf{r}')$$

- **Interpretation:** the sound field in the volume is uniquely determined by the sound pressure on the boundary and by its directional derivative in the direction normal to the boundary

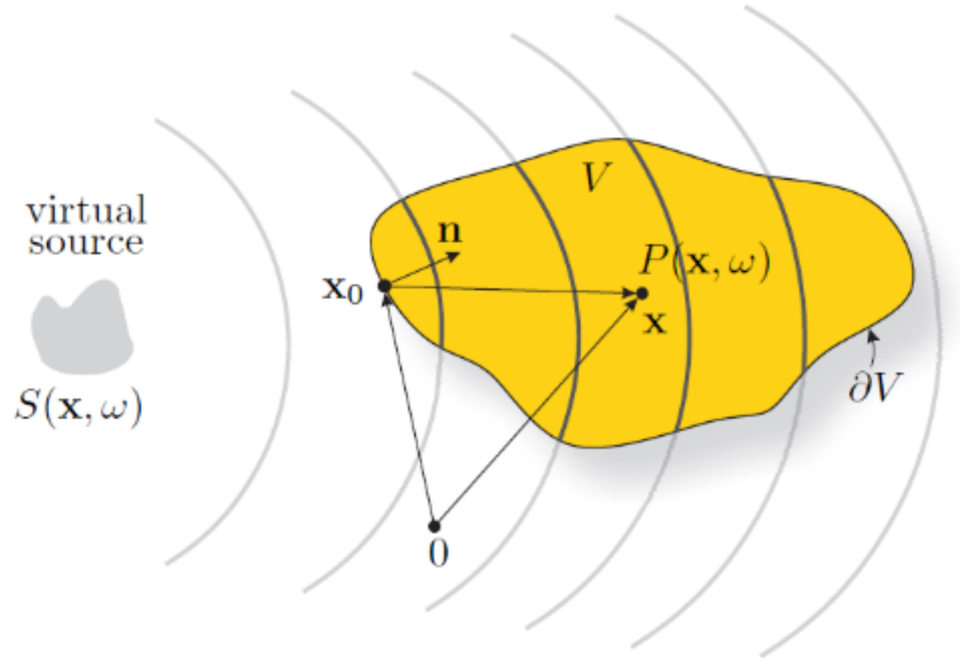


Figure from [Spors2010, Fig. 1]

The Single Layer Potential

- Simplification of the Kirchhoff-Helmholtz integral equation
  - Discard contributions propagated by the directional derivative of the Green's functions

$$p(\mathbf{r},\omega) = \int_{\mathcal{D}} G(\mathbf{r}|\mathbf{r}',\omega) D(\mathbf{r}',\omega)\, d\mathbf{r}', \quad \mathbf{r}' \in \mathcal{D}$$

  - This integral equation is known as **Single Layer Potential**

Helmholtz equation: $\nabla^2 P(\omega, \mathbf{x}) + \frac{\omega^2}{c^2} P(\omega, \mathbf{x}) = 0$

General solution $P(\omega, \mathbf{x}) = S(\omega) A(\mathbf{x}, \omega) e^{j\omega T(\mathbf{x})}$

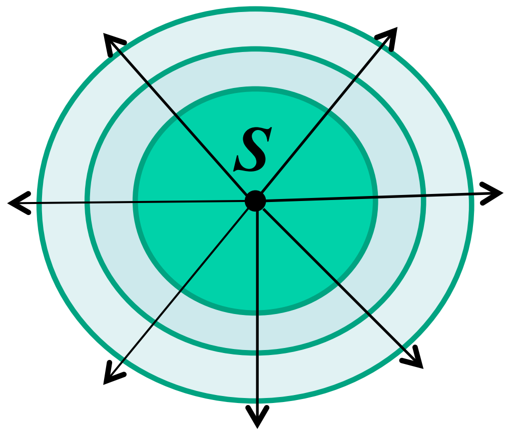
Eikonal equation:

$$[\nabla T(\mathbf{x})]^2 - \frac{1}{c^2(\mathbf{x})} = 0$$

Hp:
- High frequencies
- Non dispersive medium

Eikonal equation operates only on the direction $\nabla T(\mathbf{x})$,
i.e. the direction orthogonal to the wavefront $T(\mathbf{x})$.
This direction is referred to as **Acoustic Ray**

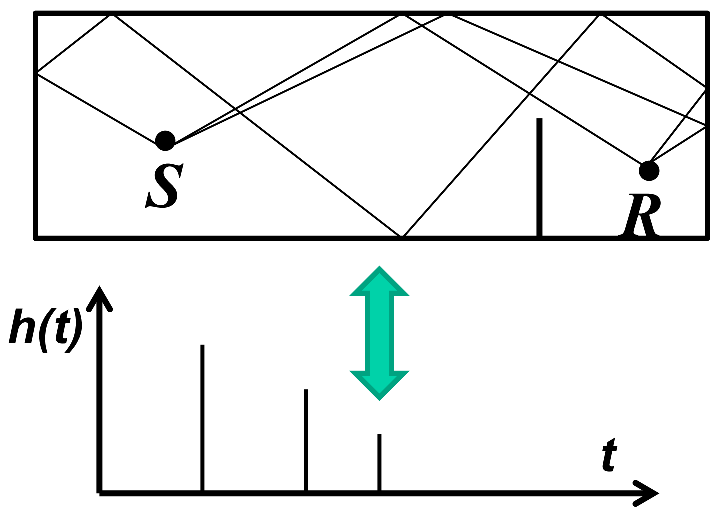
Geometrical acoustics: ray geometry and examples



The WF produced by a point-like source is described by a multitude of rays departing from the source

As they meet reflectors, rays bounce according to Snell’s law

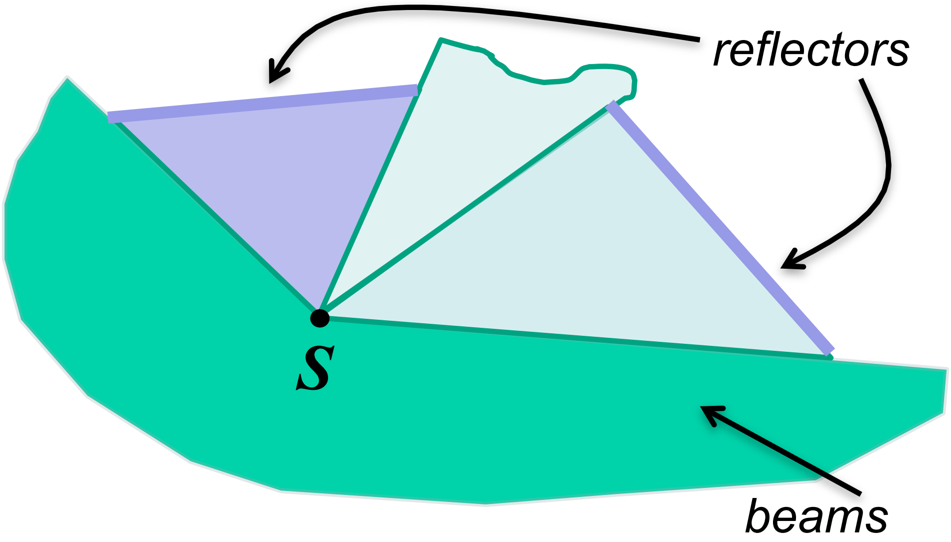
A peak in the AIR denotes an acoustic path linking source and receiver



From acoustic rays to acoustic beams

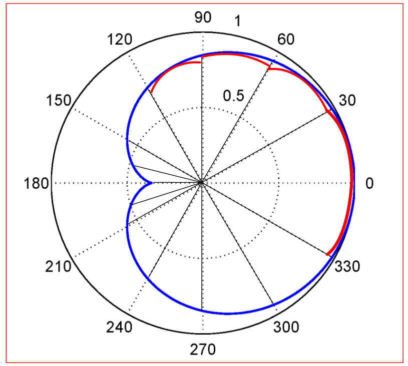
Non-planar wavefronts require a plurality of rays → we need a compact representation for bundles of rays

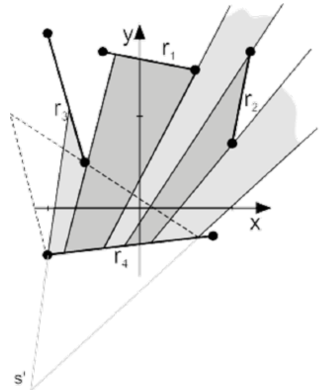
Acoustic beam: set of all rays that originate from the same pt and meet the same (planar) obstacle



Acoustic beams

- Effective for describing wave propagation in enclosures
  - *Source modeling*: any radial pattern can be described by a number of uniform beams (piece-wise constant approximation)
  - *Environment modeling*: beams split and/or branch out when they interact with reflectors
  - Suitable for modeling spherical as well as planar wavefronts

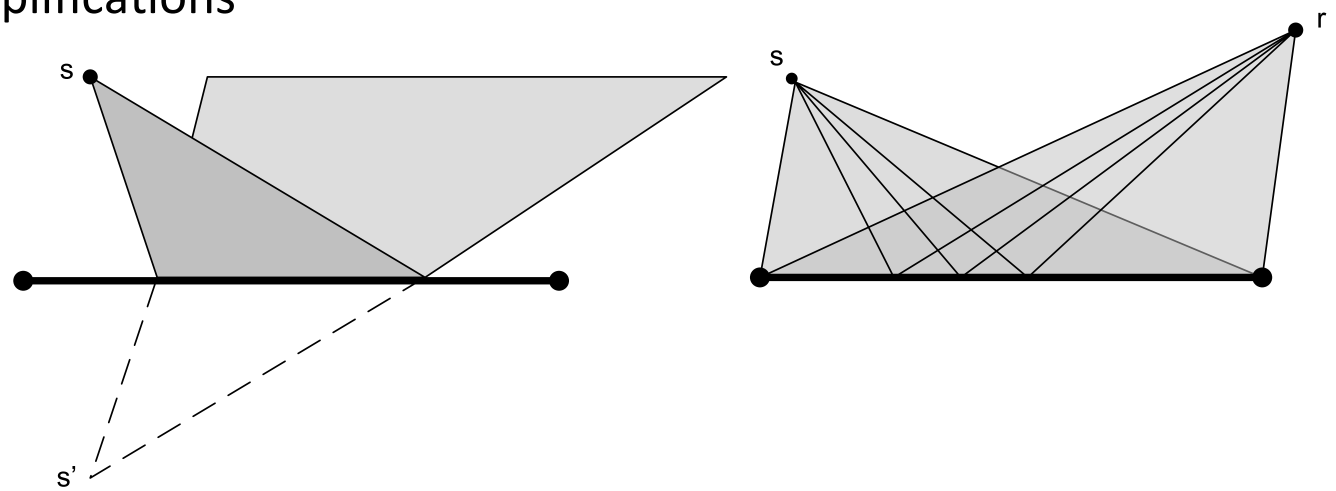




- Inherently open-loop modeling
- Natural separation btw
  - geometric aspects of propagation
  - auralization
    - filtering due to propagation, absorption, dispersion, interaction with reflectors, etc.

Acoustic beams: limitations

- Only interactions with the environment that preserve the point-like nature of (real/image) sources can be directly modeled
  - In order to accommodate diffusive surfaces, we need to introduce simplifications



- Ray-based representations are not accurate at low frequency
  - Diffraction causes wavefield to work around obstacles → Geometrical Theory of Diffraction (GTD)

References

[Olver 2010] F. W. J. Olver, editor. *NIST Handbook of Mathematical Functions.* National Institute of Standards and Technology, New York, NY, USA, 2010.

[Williams1999] E. G. Williams. *Fourier Acoustics.* Academic Press, London, UK, 1999.

[Spors2008] S. Spors, R. Rabenstein, and J. Ahrens. The theory of wave field synthesis revisited. In *Proc. AES 124th Conv.,* Amsterdam, NE, May 17-20, 2008.

[Colton1992] D. Colton and R. Kress. *Inverse Acoustics and Electromagnetic Scattering Theory.* Springer-Verlag, Berlin Heidelberg, DE, 1992.

POLITECNICO DI MILANO

ISPL
Image and Sound Processing Lab

SMC Lab

Rethinking Fourier Acoustics

Part 1: The Fourier Acoustics Toolbox

Augusto Sarti, Fabio Antonacci, Lucio Bianchi

Outline

- The Fourier expansion (inverse Fourier transform) of a multidimensional signal: Plane Wave Expansion
  - Whittaker's and Weyl's representations
  - Beamforming as a Fourier transform
- The Fourier series expansion of a multidimensional signal: the Spherical Harmonics Expansion
  - Interior and exterior acoustic fields
  - Truncation of the Spherical Harmonics Expansion
- Relations between plane waves and spherical harmonics
  - Impact of truncation on the Plane Wave Expansion
- Applications
  - PW-based analysis and rendering
  - SH-based analysis and rendering
  - PW-based near field acoustic holography

The multidimensional Fourier transform

- Consider a multidimensional signal of continuous spatial variables

$$x(\mathbf{r}),\ x : \mathbb{R}^D \rightarrow \mathbb{C}$$

- The Fourier transform over space is

$$X(\mathbf{k}) = \int_{-\infty}^{\infty} x(\mathbf{r}) e^{-j\langle \mathbf{k},\mathbf{r} \rangle}\, d\mathbf{r}, \quad \mathbf{k} \in \mathbb{R}^D$$

- The Inverse Fourier transform over the spatial frequencies is

$$x(\mathbf{r}) = \left(\frac{1}{2\pi}\right)^D \int_{-\infty}^{\infty} X(\mathbf{k}) e^{j\langle \mathbf{k},\mathbf{r} \rangle}\, d\mathbf{k}, \quad \mathbf{r} \in \mathbb{R}^D$$

The Whittaker's Plane Wave Expansion

- Propagating plane waves are a solution to the homogeneous Helmholtz equation in Cartesian coordinates

$$p(\mathbf{r}, \omega) = e^{j\langle \mathbf{r}, \mathbf{k} \rangle}, \quad \mathbf{k} \in \mathbb{R}^3$$

- The **inverse multidimensional Fourier transform** of the acoustic pressure field is

$$p(\mathbf{r}, \omega) = \left(\frac{1}{2\pi}\right)^3 \iiint_{\mathcal{D}} A(\mathbf{k}) e^{j\langle \mathbf{k}, \mathbf{r} \rangle} \, d^3\mathbf{k},$$

$$\mathcal{D} = \left\{ \mathbf{k} \in \mathbb{R}^3 : \|\mathbf{k}\| = \frac{\omega}{c} \right\}$$

  - The wavenumber vector plays the role of the spatial frequency
  - This expansion is known as **Whittaker's representation**

Geometric interpretation of Whittaker's representation

- Factorize the wavenumber vector as

$$\mathbf{k} = (\omega/c)\hat{\mathbf{k}} \qquad \hat{\mathbf{k}} = [\sin(\theta)\cos(\phi), \sin(\theta)\sin(\phi), \cos(\theta)]^T$$

- Substitute the angular factorization for the unit wavenumber vector in the Whittaker's representation

$$p(\mathbf{r},\omega) = \left(\frac{1}{2\pi}\right)^3 \iint_{\mathcal{S}} A(\theta,\phi,\omega) e^{j\frac{\omega}{c}\langle\hat{\mathbf{k}},\mathbf{r}\rangle} \sin(\theta)\, d\theta\, d\phi,$$

$$\mathcal{S} = \{\theta \in [0,\pi], \phi \in [0, 2\pi)\}$$

  - The acoustic field is written as a superposition of propagating plane waves
    - Directions of propagation cover a sphere
  - The function $A(\theta,\phi,\omega)$ encodes magnitude and phase for each plane wave
    - Does not depend on the observation point
    - Is known as **Herglotz density**

The Weyl's Plane Wave Expansion

- Consider a non-zero excitation function

$$q(\mathbf{r},\omega) \neq 0, \quad \text{for } \mathbf{r} \in \mathcal{D}$$

- The resulting acoustic pressure field must satisfy the inhomogeneous Helmholtz equation

$$\nabla^2 p(\mathbf{r},\omega) + \left(\frac{\omega}{c}\right)^2 p(\mathbf{r},\omega) = -4\pi q(\mathbf{r},\omega)$$

- The solution to the inhomogeneous Helmholtz equation is constructed by means of a superposition of Green's functions

$$p(\mathbf{r},\omega) = \int_{\mathcal{D}} q(\mathbf{r}',\omega) \frac{e^{-j(\omega/c)\|\mathbf{r}-\mathbf{r}'\|}}{4\pi\|\mathbf{r}-\mathbf{r}'\|} d^3\mathbf{r}'$$

The Weyl's Plane Wave Expansion

- If we substitute the Weyl's identity in the free-field Green's function, exchange the order of integration and rearrange the terms, we obtain

$$p(\mathbf{r},\omega) = -\frac{j}{8\pi^2}\iint_{\mathbb{R}^2} e^{j\langle \mathbf{k},\mathbf{r}\rangle} Q(\mathbf{k})\, dk_x\, dk_y$$

  - Twofold integral over real variables
  - The function $Q(\mathbf{k}) = \int_{\mathcal{D}} q(\mathbf{r}',\omega) e^{j\langle \mathbf{k},\mathbf{r}'\rangle}\, d^3\mathbf{r}'$ is called **angular spectrum**

Geometric interpretation of Weyl's representation

- Let us introduce the change of variables

$$k_x = \frac{\omega}{c}\sin(\alpha)\cos(\beta), \quad k_y = \frac{\omega}{c}\sin(\alpha)\sin(\beta), \quad k_z = \frac{\omega}{c}\cos(\alpha)$$

  - If $k_x^2 + k_y^2 \leq (\omega/c)^2$ (propagating plane waves)
    - α and β are the spherical angles related to the direction of propagation of the plane wave

$$\alpha = \frac{\pi}{2} + j\alpha', \quad -\infty < \alpha' < 0 \quad \beta = \phi \in [0, 2\pi)$$

  - If $k_x^2 + k_y^2 > (\omega/c)^2$ (evanescent plane waves)
    - α must be a complex angle

$$\alpha = \frac{\pi}{2} + j\alpha', \quad -\infty < \alpha' < 0 \quad \beta = \phi \in [0, 2\pi)$$

Fourier expansion as beamforming

- Discretize the inverse Whittaker's representation considering a finite number of field points $\mathbf{r}_n,\ n = 1, \ldots, N$

$$A(\theta, \phi, \omega) \propto \sum_{n=1}^{N} p(\mathbf{r}_n, \omega) e^{-j\langle \hat{\mathbf{k}}(\theta,\phi), \mathbf{r}_n \rangle}$$

  - This operation is widely known in the array signal processing literature as **beamforming**
  - Observations at individual sensors are modulated in order to align in phase the directional contribution from direction (θ, ϕ)

Spherical Harmonics Expansion

- Basis solution to Helmholtz equation in spherical coordinates

$$p(\mathbf{r},\omega) = R(r)\Theta(\theta)\Phi(\phi) = R(r)Y_l^m(\theta,\phi)$$

  - **Spherical harmonic waves**
- A general field can be written as the summation of infinite spherical harmonic waves
  - Radial dependence expressed with spherical Hankel functions

$$p(\mathbf{r},\omega) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left(A_{lm}(\omega)h_l^{(1)}((\omega/c)r) + B_{lm}(\omega)h_l^{(2)}((\omega/c)r)\right)Y_l^m(\theta,\phi)$$

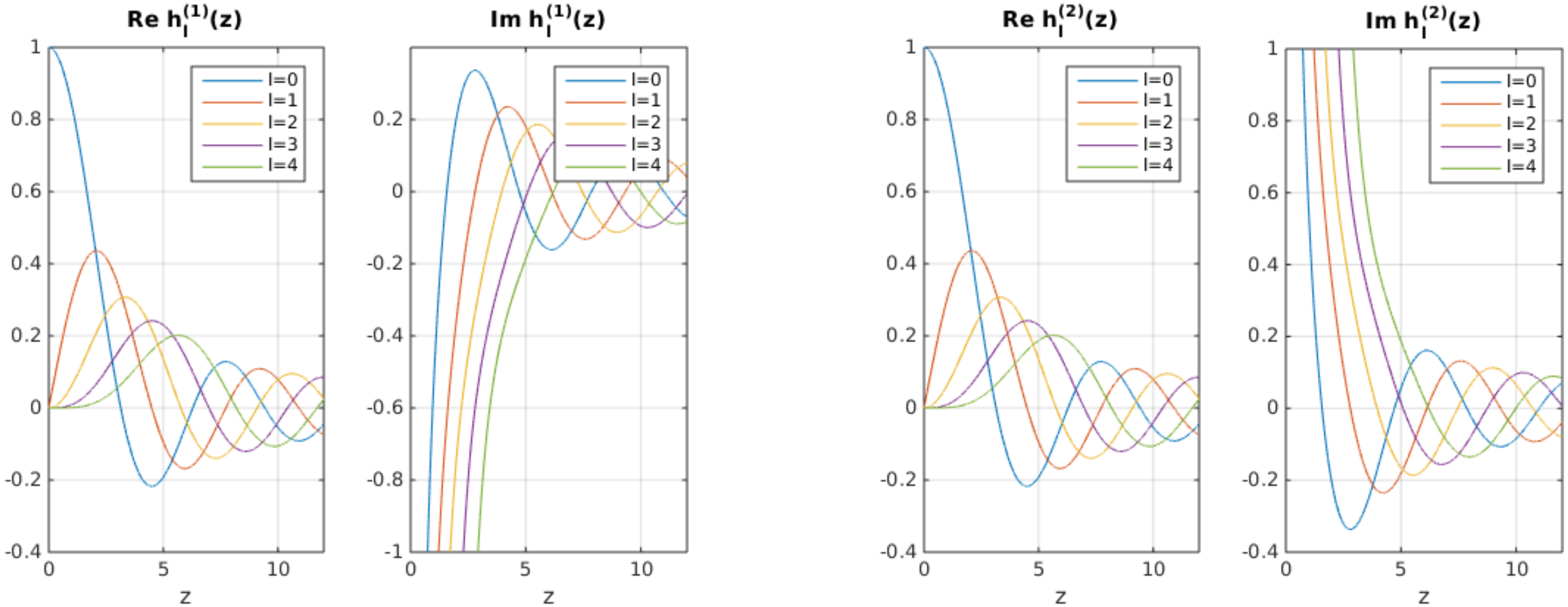
  - Radial dependency expressed with spherical Bessel functions

$$p(\mathbf{r},\omega) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left(C_{lm}(\omega)j_l((\omega/c)r) + D_{lm}(\omega)y_l((\omega/c)r)\right)Y_l^m(\theta,\phi)$$

- In both cases the acoustic field is characterized by sets of coefficients

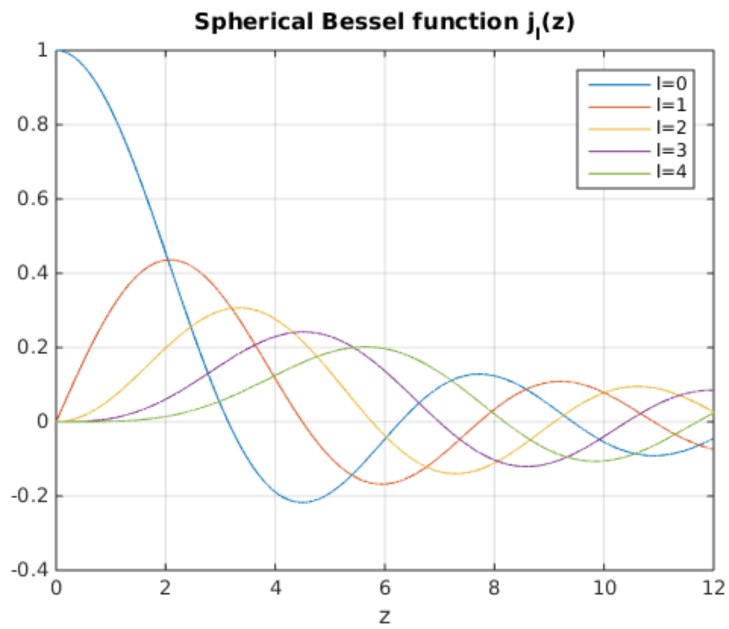
Considerations on the radial functions

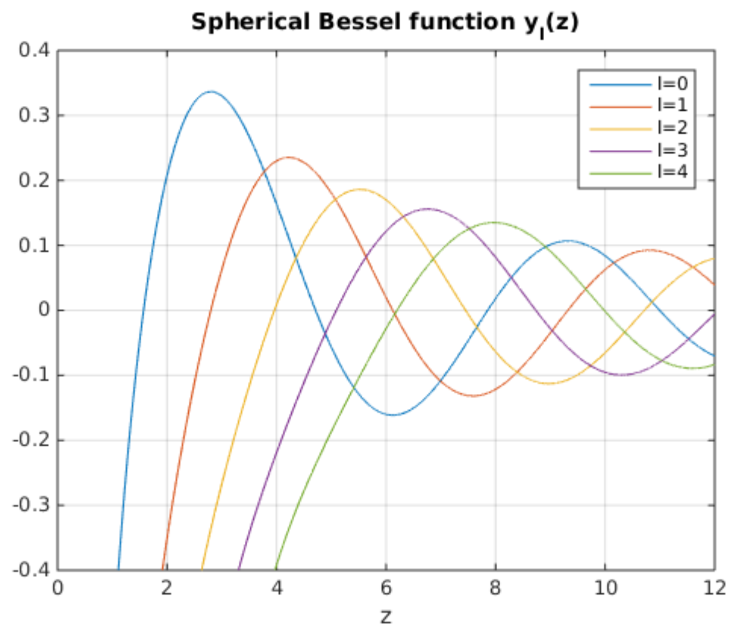
- For $z \to 0$: $y_l(z) \to \infty$, $h_l^{(1)}(z) \to \infty$ and $h_l^{(2)}(z) \to \infty$
  - These function are suitable to represent acoustic field due to **sources near the origin**
  - The function $j_l(z)$ is suitable to represent acoustic fields in a source free region around the origin



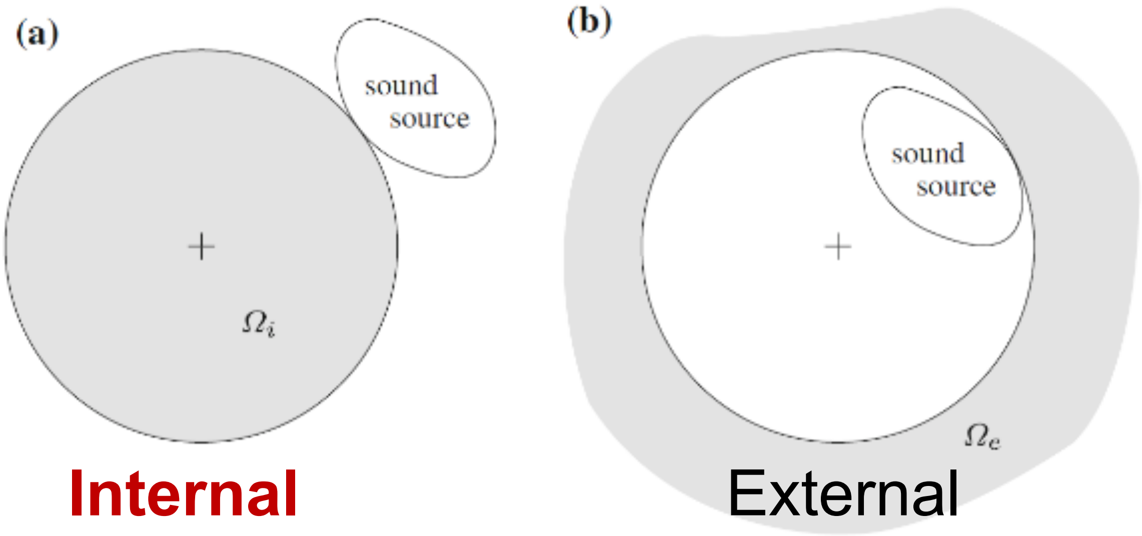
Considerations on the radial functions

- For $z \to 0$: $y_l(z) \to \infty$, $h_l^{(1)}(z) \to \infty$ and $h_l^{(2)}(z) \to \infty$
  - These function are suitable to represent acoustic field due to sources near the origin
  - The function $j_l(z)$ is suitable to represent acoustic fields in a **source free region around the origin**





Internal acoustic field



Internal

External

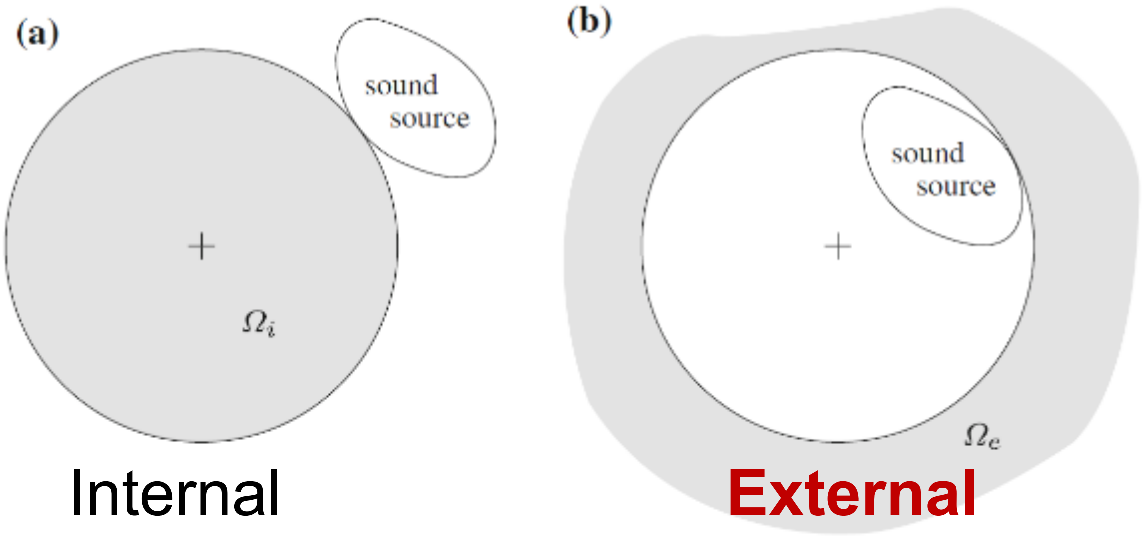
- Inverse Spherical harmonics expansion (synthesis)

$$p(\mathbf{r},\omega) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} C_{lm}(\omega) j_l((\omega/c)r) Y_l^m(\theta,\phi)$$

- Spherical harmonics expansion (analysis)

$$C_{lm}(\omega) = \frac{1}{j_l((\omega/c)r)} \int_0^{2\pi} \int_0^{\pi} p(\mathbf{r},\omega) Y_l^{-m}(\theta,\phi) \sin(\theta)\, d\theta\, d\phi$$

External acoustic field



Internal External

- Inverse Spherical harmonics expansion (synthesis)

$$p(\mathbf{r}, \omega) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} B_{lm}(\omega) h_l^{(2)}((\omega/c)r) Y_l^m(\theta, \phi)$$

- Spherical harmonics expansion (analysis)

$$B_{lm}(\omega) = \frac{1}{h_l^{(2)}((\omega/c)r)} \int_0^{2\pi} \int_0^{\pi} p(\mathbf{r}, \omega) Y_l^{-m}(\theta, \phi) \sin(\theta)\, d\theta\, d\phi$$

Bandlimited spherical harmonics expansion

- Limit the spherical harmonics expansion to order L-1

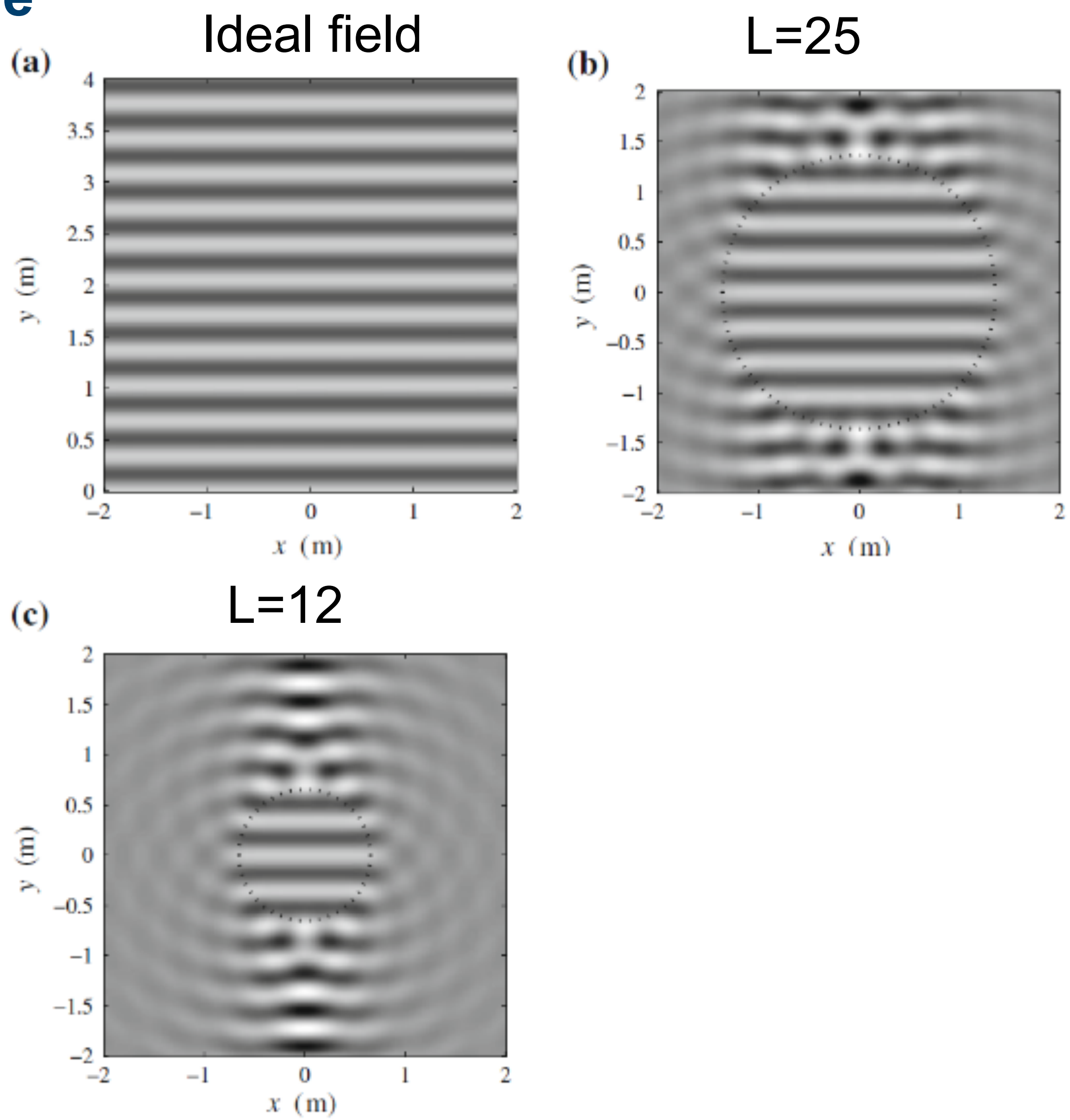
$$p(\mathbf{r}, \omega) \approx \sum_{l=0}^{L-1} \sum_{m=-l}^{l} B_{lm}(\omega) h_l^{(2)}((\omega/c)r) Y_l^m(\theta, \phi)$$

  - The acoustic field is described by L^2 coefficients
- For an internal acoustic field
  - Rule of thumb: the bandlimited expansion provides a reasonable approximation is

$$\frac{\omega}{c} r_{L-1} < (L-1)$$

    - r_{L-1} is the radius of the internal region
    - Fixed the maximum order of the expansion, the radius of the region of validity is inversely proportional to frequency

Bandlimited spherical harmonics expansion - example



From [Ahrens2012, Fig. 2.7]

Relation between plane waves and spherical waves

- Spherical harmonics and plane waves are related through the Gegenbauer expansion

$$e^{j\langle \mathbf{k},\mathbf{r}\rangle} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{l} j^l Y_l^{-m}(\theta,\phi) j_l(kr) Y_l^m(\theta_r,\phi_r)$$

$$j_l(kr) Y_l^m(\theta_r,\phi_r) = \frac{1}{4\pi} j^{-l} \int_0^{2\pi}\int_0^{\pi} e^{j\langle \mathbf{k},\mathbf{r}\rangle} Y_l^m(\theta,\phi)\, d\theta\, d\phi$$

- Expansion coefficients in spherical harmonics and plane wave expansion are related by

$$A(\theta,\phi,\omega) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} j^{-l} A_{lm}(\omega) Y_l^m(\theta,\phi)$$

$$C_l^m(\theta,\phi) = j^l \int_0^{2\pi}\int_0^{\pi} A(\theta,\phi) Y_l^{-m}(\theta,\phi)\, d\theta\, d\phi$$

Impact of truncation on the plane wave expansion

- In practice, the infinite summation in the Gegenbauer expansion is replaced with a finite summation up to mode L

$$e^{j\langle \mathbf{k},\mathbf{r}\rangle} = 4\pi \sum_{l=0}^{L} \sum_{m=-l}^{l} j^l Y_l^{-m}(\theta,\phi) j_l(kr) Y_l^m(\theta_r,\phi_r)$$

- The mode-limited plane wave coefficients are

$$A(\theta,\phi,\omega) = \sum_{l=0}^{L} \sum_{m=-l}^{l} j^{-l} A_{lm}(\omega) Y_l^m(\theta,\phi)$$

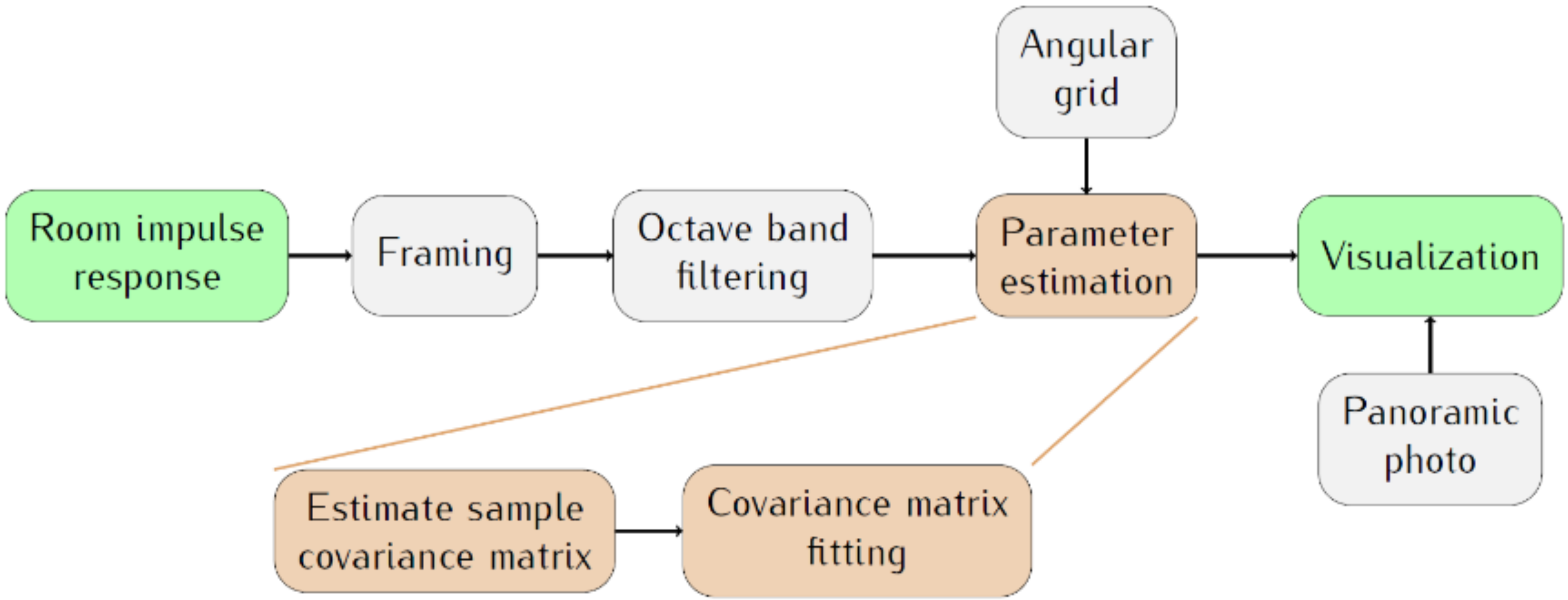
- The two representations are bandlimited to $\mathcal{O}(L^2)$
  - Gibbs phenomena arise in the plane wave spectrum

Applications

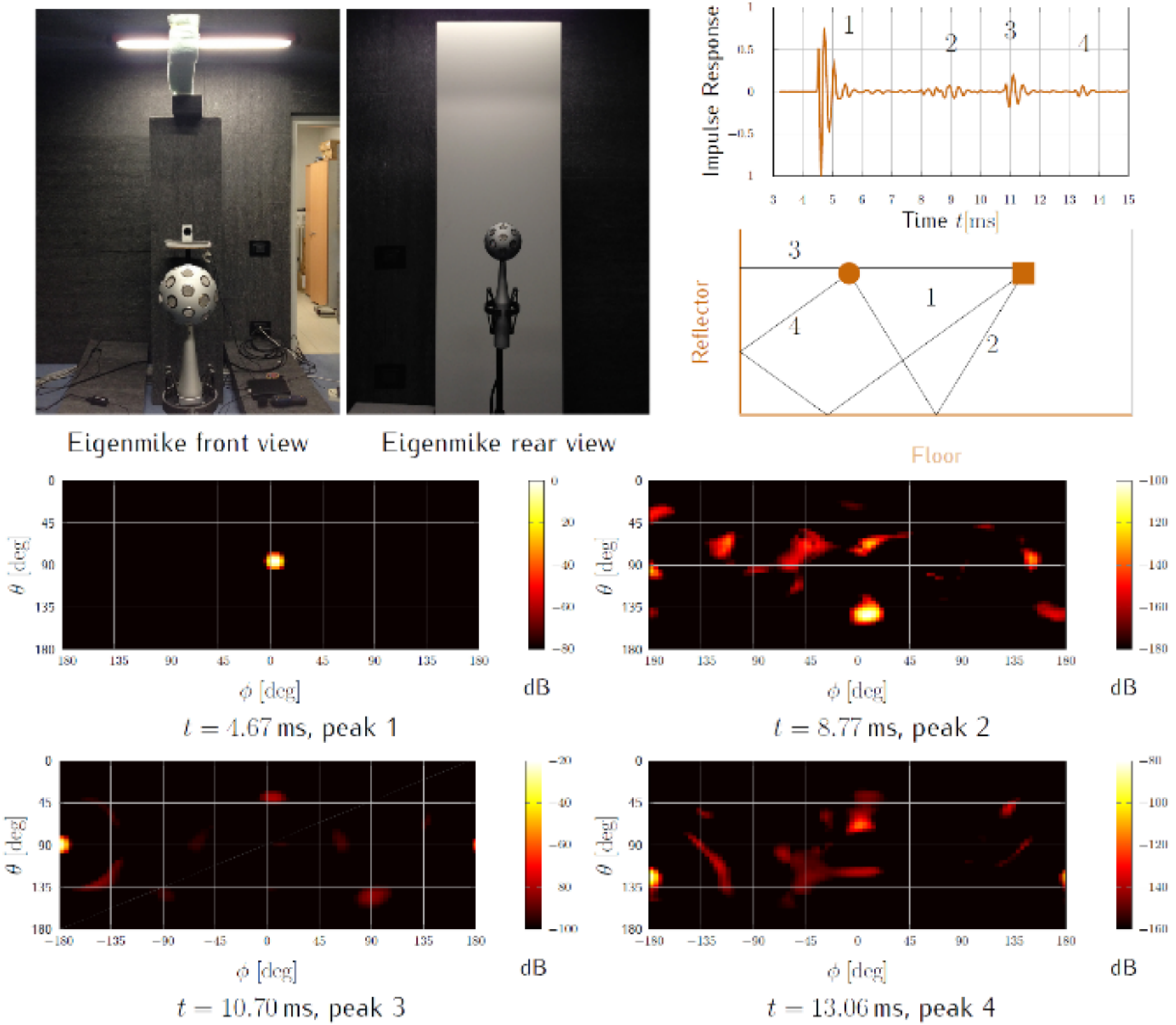
- PW-based analysis and rendering
- SH-based analysis and rendering
- PW-based near field acoustic holography

- Adopt Whittaker’s plane wave representation in the short-time scale
  - The acoustic field within a short-time frame is represented as an integral of propagating plane waves
- Discretize the integral representation
- Estimate strength, direction of arrival and time of arrival of plane wave components
  - Plane waves are associated to direct sound and early reflections

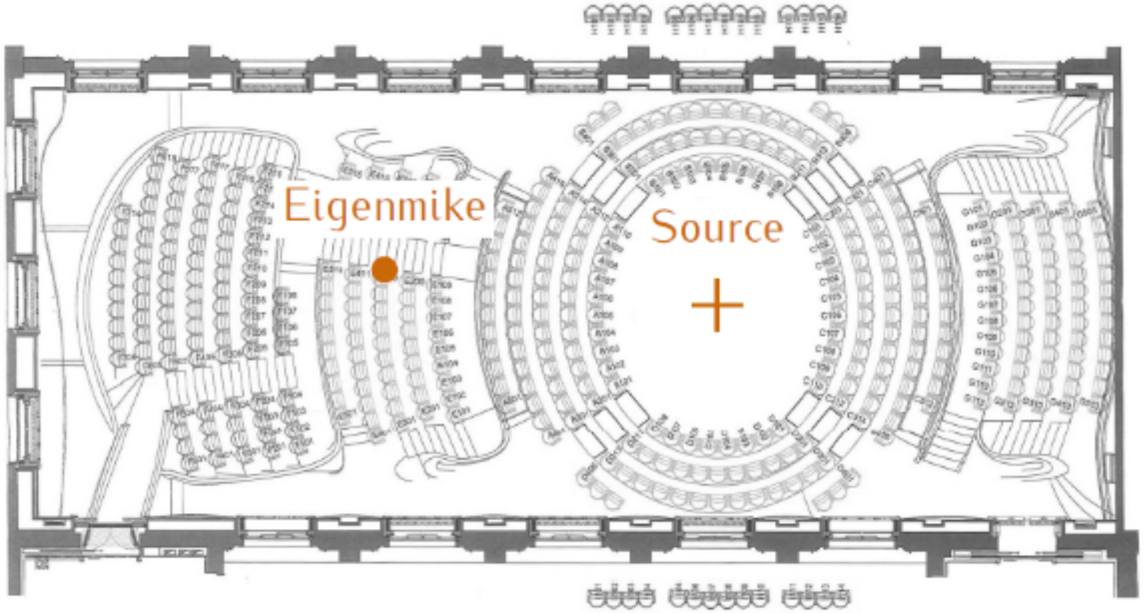
High Resolution Acoustic Camera

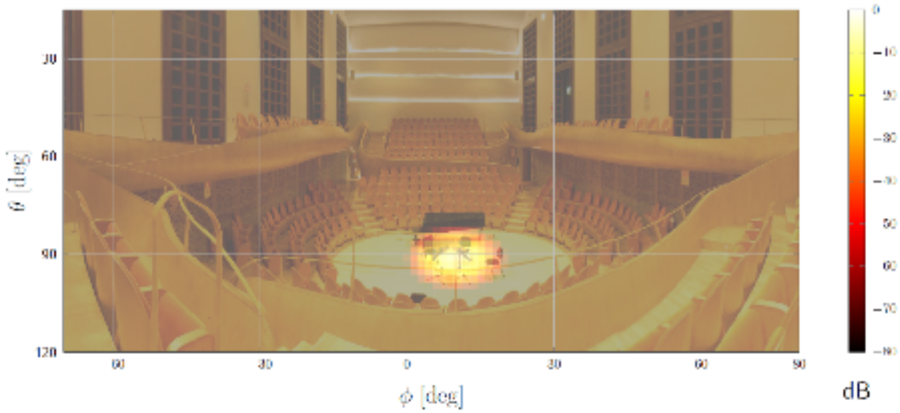


Imaging in a controlled environment

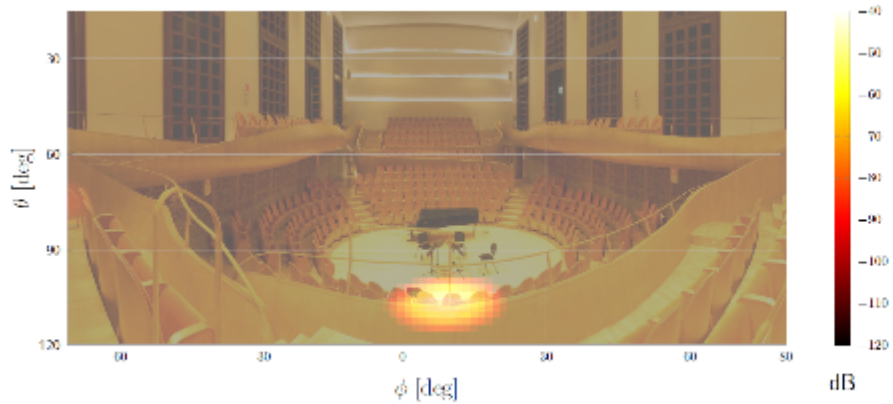


From [Bianchi2015], fig. 3





$t = 30.32$ ms, direct path



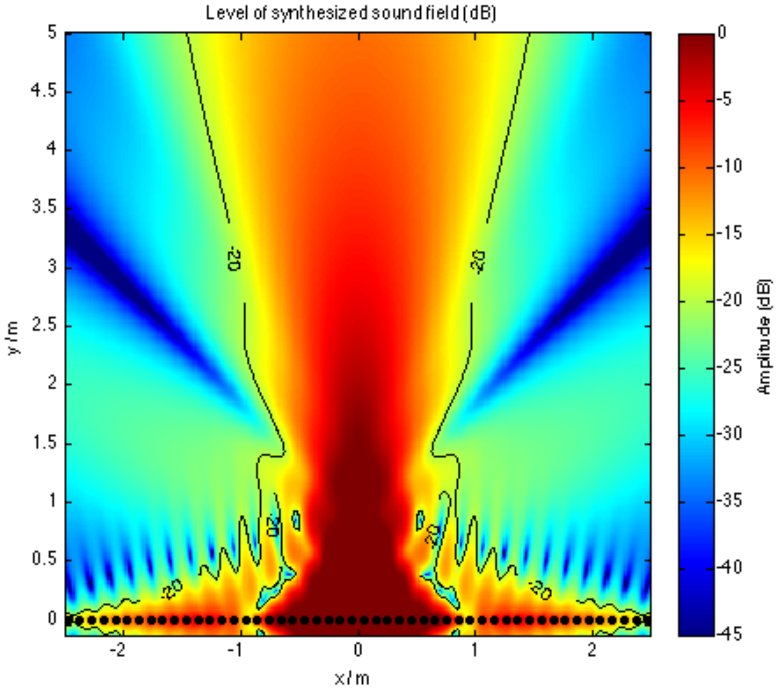
$t = 33.59$ ms, railing diffraction

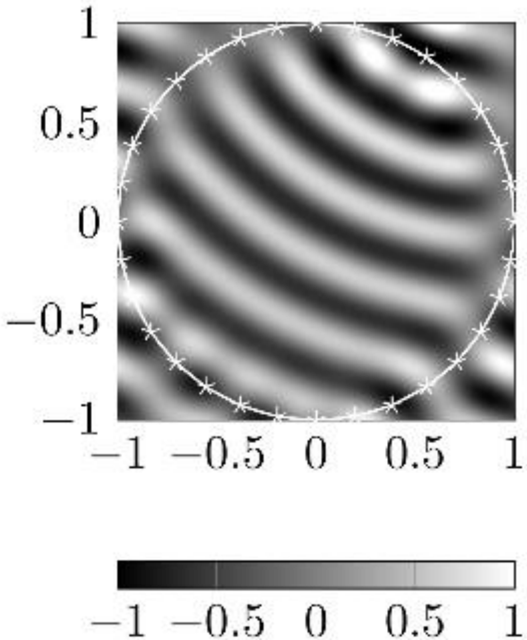
From [Bianchi2015], fig. 5

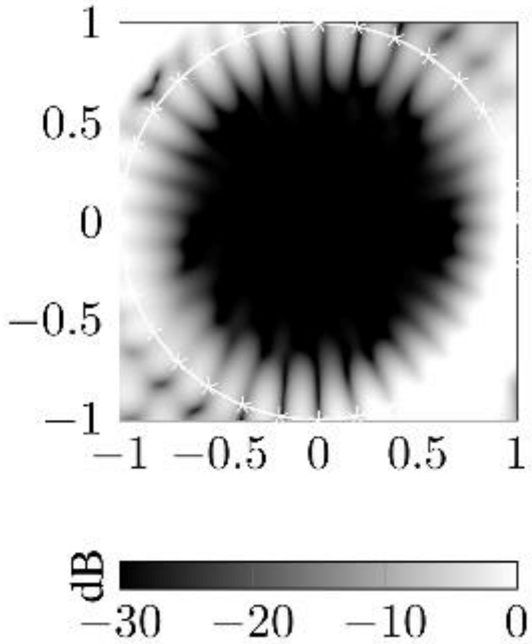
PW-based rendering – Acoustic displays

- Dual to acoustic cameras
- Implemented as arrays of loudspeakers that, through beamforming, generate plane-wave components
  - Can be used to focus acoustic energy to specific directions


In collaboration with B&C Speakers







From [Bianchi2015a], fig. 8

SH-based analysis and rendering

- Sound field at point r is expressed in terms of spherical harmonics (SH) of order n and mode m

$$p(\mathbf{r},\omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} C_{nm}(\omega) j_n\left(\frac{\omega}{c} r\right) Y_n^m(\theta,\phi)$$

  - Coefficients are independent of location
  - If one could record them, then the acoustic field could be rendered through this expression
- Ideally, the spherical harmonic coefficients are computed by

$$C_{nm}(\omega) = \frac{1}{j_n\left(\frac{\omega}{c} r\right)} \int_0^{2\pi} \int_0^{\pi} p(\mathbf{r},\omega) Y_n^{-m}(\theta,\phi)\, d\theta\, d\phi$$

  - Valid if r is on the surface of a sphere of radius r

- Approximate the integration by a finite summation
- Use Q omnidirectional microphones on a rigid sphere of radius R to obtain pressure measurements $p(R, \theta_q, \phi_q, \omega),\ q = 1, \ldots, Q$

$$\hat{C}_{nm}(\omega) = \frac{1}{j_n\left(\frac{\omega}{c}R\right)} \sum_{q=1}^{Q} p(R, \theta_q, \phi_q) Y_n^{-m}(\theta_q, \phi_q) w_q$$

  - w_q are suitable weights that depend on the quadrature rule adopted to sample the sphere with microphones

- An acoustic field bandlimited to N has $(N+1)^2$ harmonic components
- In order to accurately reconstruct a field up to order N one needs Q microphones, where

$$Q \geq (N+1)^2$$

- Consider equiangular spacing for microphones
  - More dense packing near poles
  - $(N+1)^2$ microphones are not sufficient, one needs at least

$$Q \geq (2N-1)^2$$

    - Weights are $w_q = 2\pi/Q$

- Adopt Weyl’s representation for the acoustic field measured on a plane at height z

$$p(x,y,z,\omega) = -\frac{j}{8\pi^2}\iint_{\mathbb{R}^2} e^{j(k_x x + k_y y)} P(k_x, k_y, z)\, dk_x\, dk_y$$

- In the wavenumber (spatial frequency) domain

$$P(k_x, k_y, z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x,y,z,\omega) e^{-j(k_x x + k_y y)}\, dx\, dy$$

Holographic prediction

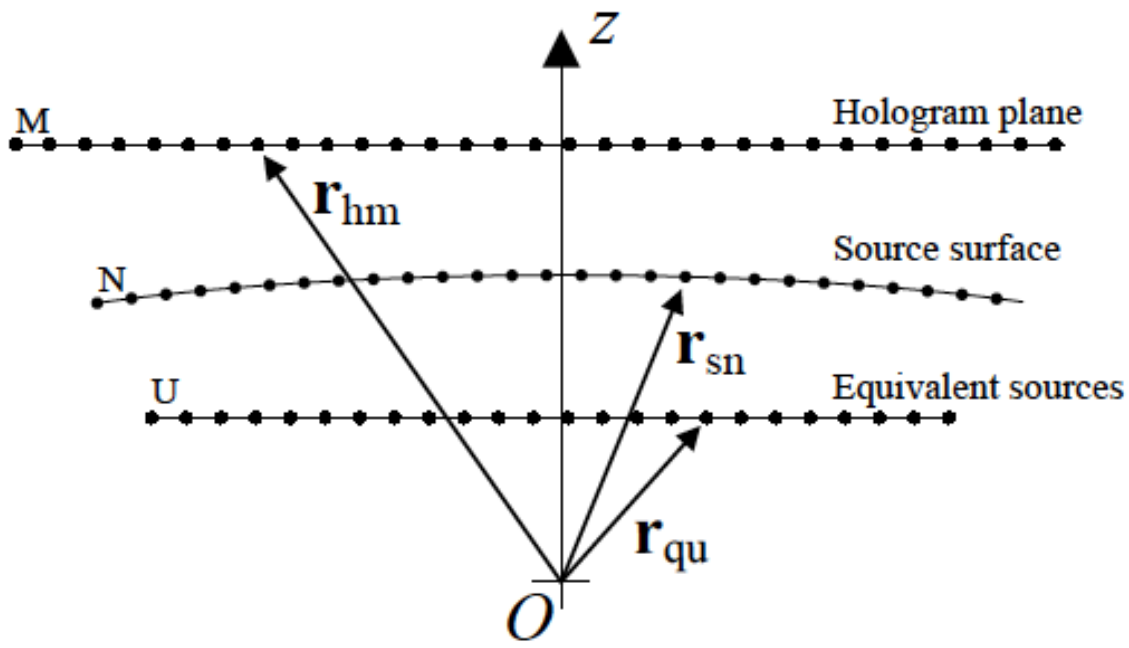
- The angular spectrum on the prediction plane z_p can be expressed as a function of the angular spectrum on the holographic plane z_h

$$P(k_x, k_y, z_p) = P(k_x, k_y, z_h)G(k_x, k_y, z_p - z_h)$$

- The plane wave propagator is

$$G(k_x, k_y, z_p - z_h) = \begin{cases} e^{j\sqrt{k^2 - k_x^2 - k_y^2}}(z_p - z_h),\ k_x^2 + k_y^2 \leq k^2 \\ e^{-j\sqrt{k_x^2 + k_y^2 - k^2}}(z_p - z_h),\ k_x^2 + k_y^2 > k^2 \end{cases}$$

Equivalent source method:



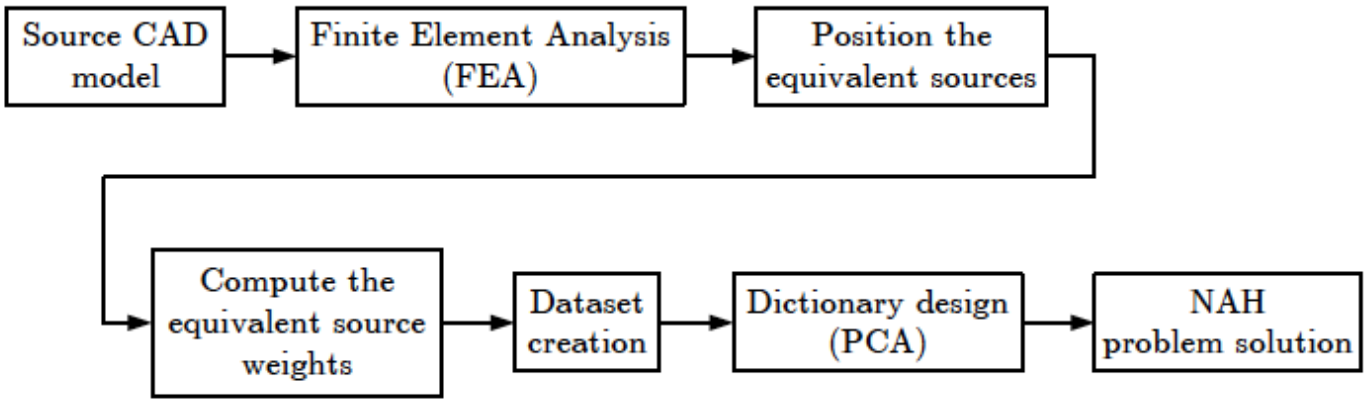
Equivalent source method involves two steps:

1. Given the pressure measurement on the hologram plane, estimate the equivalent sources;
2. equivalent sources are propagated to compute the acoustic pressure field on the source plane

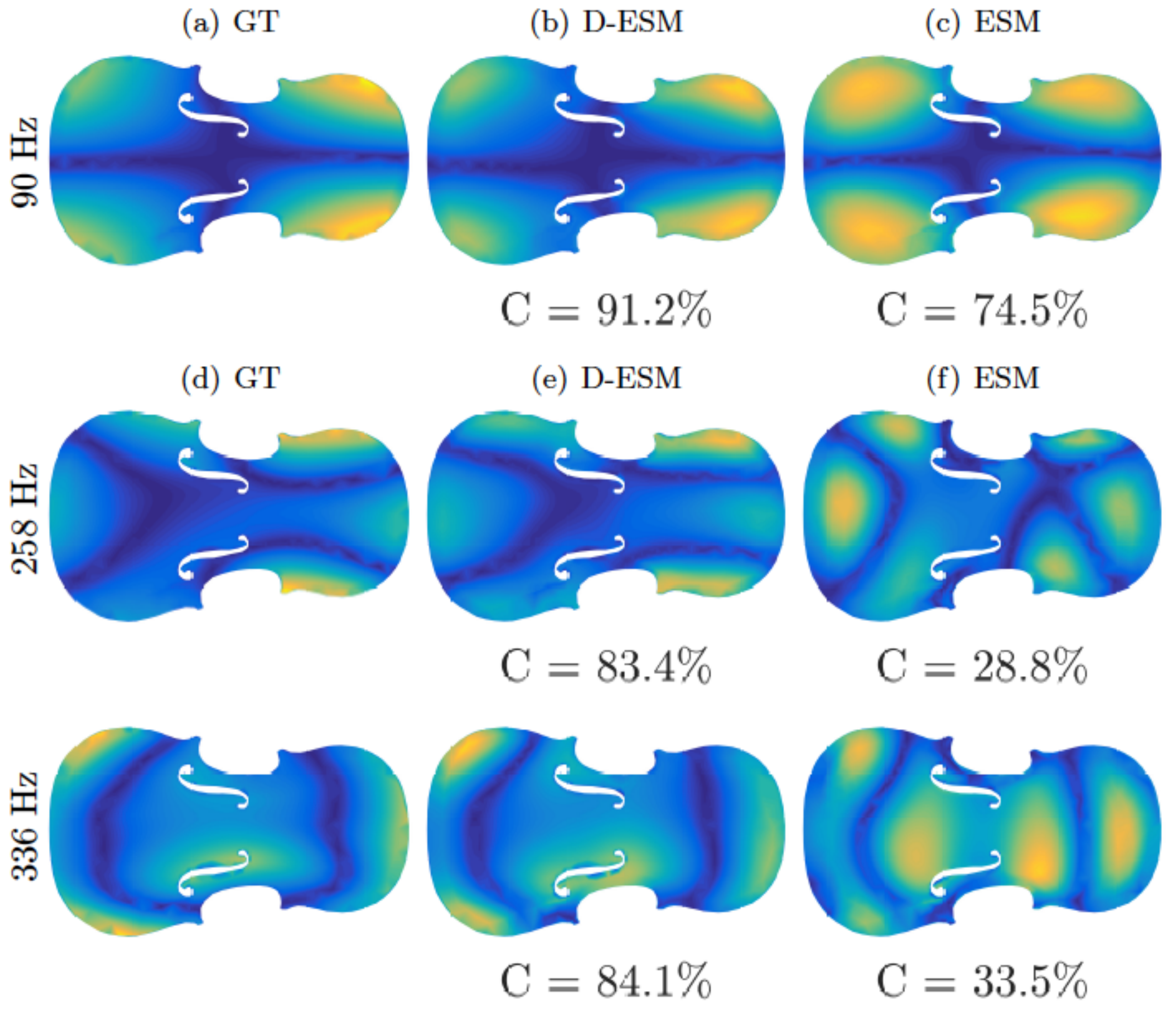
ESM cannot guarantee a high level of accuracy if only a few measurement points are available and noise is present.

Solution:

1. build a dictionary of vibrating modes under variation of the material properties (Poisson ratio, velocity, stiffness, etc.)
2. Find the combination of dictionary modes that best explains the measurements.



Dictionary-based ESM



GT: ground truth

D-ESM: Dictionary-based ESM

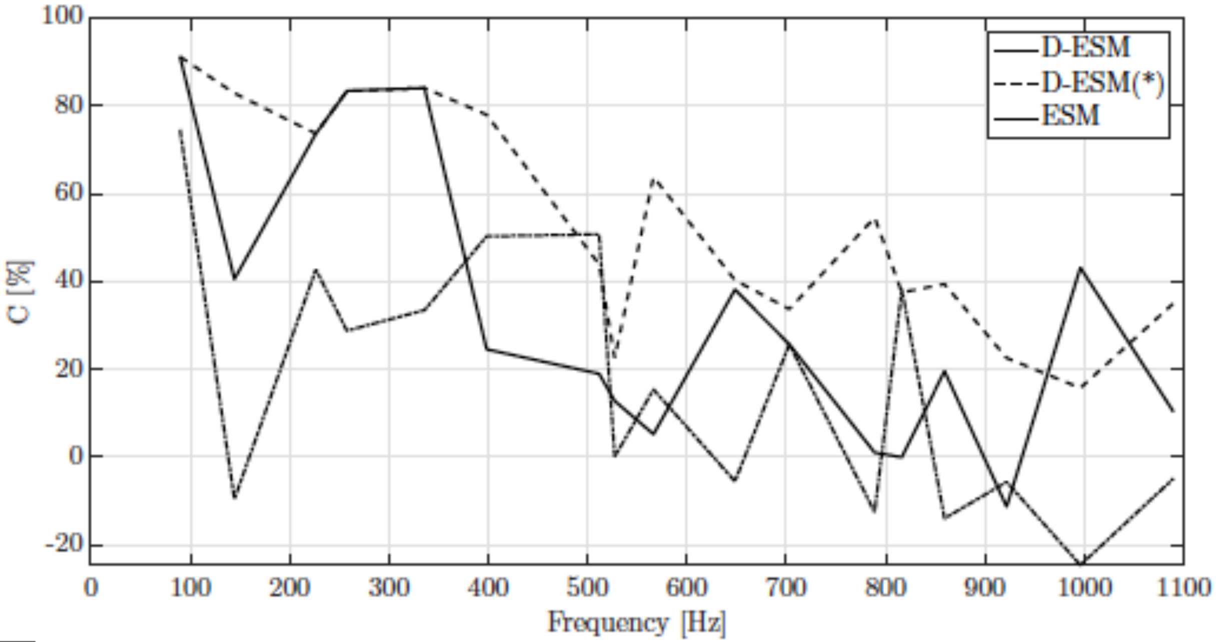
ESM: equivalent source method

C: correlation index with GT

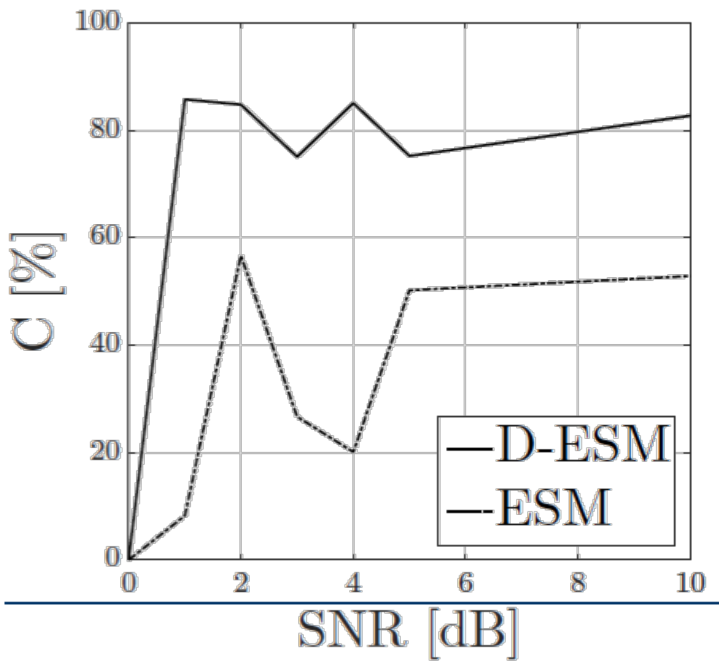
Comment: D-ESM turns out to overcome ESM

Dictionary-based ESM

Correlation index of different NAH techniques as a function of frequency



D-ESM(*): modification of D-ESM in which a single mode is selected from the dictionary



Correlation index of different NAH techniques as a function of the Signal to Noise Ratio at the microphone array

References

[Colton 1992] D. Colton and R. Kress. *Inverse Acoustics and Electromagnetic Scattering Theory.* Springer-Verlag, Berlin Heidelberg, DE, 1992.

[Williams1999] E. G. Williams. *Fourier Acoustics.* Academic Press, London, UK, 1999.

[Stoica2004] P. Stoica and R. Moses, *Spectral Analysis of Signals.* Prentice Hall, Upper Saddle River, NJ, USA, 2004.

[Olver2010] F. W. J. Olver, editor. *NIST Handbook of Mathematical Functions.* National Institute of Standards and Technology, New York, NY, USA, 2010.

[Kennedy2007] R. A. Kennedy, P. Sadeghi, T. D. Abhayapala, and H. M. Jones. Intrinsic limits of dimensionality and richness in random multipath fields. *IEEE Trans. Signal Process.,* 55(6): 2542-2556, June 2007.

[Ahrens2012] J. Ahrens. *Analytic Methods of Sound Field Synthesis.* Springer-Verlag, Berlin, DE, 2012.

[Zotkin2010] D. N. Zotkin, R. Duraiswami, and N. A. Gumerov. Plane-wave decomposition of acoustical scenes via spherical and cylindrical microphone arrays. *IEEE Trans. Audio, Speech, Lang. Process.,* 18(1): 2-16, Jan. 2010.

References

[Bianchi2015] L. Bianchi, M. Verdi, F. Antonacci, A. Sarti, S. Tubaro, "High Resolution Imaging of acoustic reflections with Spherical Microphone Array", in proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), 2015, NY, Oct. 2015, pages 1-5

[Bianchi 2015a] L. Bianchi, F. Antonacci, A. Sarti, S. Tubaro, "Model-Based Acoustic Rendering based on Plane Wave Decomposition", Applied Acoustics 104 (2016) Elsevier , pp. 127-134,

POLITECNICO DI MILANO

Image and Sound Processing Lab

Rethinking Fourier Acoustics

Part 2: Sharpening the Tools

Augusto Sarti, Fabio Antonacci, Lucio Bianchi

Outline

- Sharpening the Spherical Harmonic tools
  - Translation operator
  - Application of the translation operator for arrays of higher order microphones
- Sharpening the Plane Wave tools
  - Fusing information coming from local PWD’s
  - The plenacoustic function
  - Representation of the plenacoustic function using geometrical acoustics: the ray space
  - Geometric primitives in the ray space
  - Fusing multiple sound field images in the projective ray space

The validity of SHD is only local

- Spherical Harmonic expansion of an acoustic field (referred to a given global reference frame centered in $\mathcal{O}$)

$$P(R,\vartheta,\varphi,k)=\sum_{n=0}^{\infty}\sum_{m=-n}^{n}C_{nm}(k)j_n(kR)Y_{nm}(\vartheta,\varphi)$$

- Consider now a new reference frame whose orientation and coordinates of the origin are $(R_q,\vartheta_q,\varphi_q)$. With respect to this new frame, the SH expansion can be written as

$$P(r,\theta,\phi)=\sum_{\nu=0}^{\infty}\sum_{\mu=-\nu}^{\nu}B_{\nu\mu}(k)j_\nu(kr)Y_{\nu\mu}(\theta,\phi)$$

- **Problem:** how do these two expressions relate to each other?

Translator operator for SHD

- Relation between the coefficients of the reference frames centered in $\mathcal{O}$ and $\mathcal{O}_q$ [Chen2015]

$$B_{\nu\mu} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} C_{nm} \hat{S}_{n\nu}^{m\mu}(\mathbf{R}_q)$$

$$\hat{S}_{n\nu}^{m\mu}(\mathbf{R}_q) = 4\pi i^{\nu-n} \sum_{\ell=|\mu-m|}^{n+\nu+1} i^{\ell}(-1)^{2m-\mu} j_\ell(kR_q) Y_{\ell(\mu-m)}^{*}(\vartheta_q, \varphi_q) W$$

$$W = \sqrt{\frac{(2n+1)(2\nu+1)(2\ell+1)}{4\pi}} W_1 W_2$$

$$W_1 = \begin{pmatrix} n & \nu & \ell \\ 0 & 0 & 0 \end{pmatrix}, W_2 = \begin{pmatrix} n & \nu & \ell \\ m & -\mu & \mu - m \end{pmatrix}$$

The translation operator gives us the coefficients of the local reference frame as a function of those of the global reference frame

Application of the translation operator for SHD: higher order microphone arrays

Scenario: spatial distribution of high-order microphones, each using a local reference frame

Goal: reconstruct the global sound field coefficients C_{nm} from the knowledge of the local coefficients $B_{\nu\mu}(\varphi_q)$

$$\frac{1}{Q}\sum_{q=1}^{Q} B_{\nu\mu}(\varphi_q)E_{\mu-m}(\varphi_q) \approx \sum_{n=|m|}^{\infty} C_{nm}H_{n\nu}^{m\mu}(R_s,\vartheta_s) \qquad \alpha_{\nu\mu}^{m} = \frac{1}{Q}\sum_{q=1}^{Q} B_{\nu\mu}(\varphi_q)E_{(m-\mu)}(\varphi_q)$$

$$\boldsymbol{\alpha}_m = \mathbf{H}_m\mathbf{C}_m \quad \boldsymbol{\alpha}_m = \begin{bmatrix}\alpha_{00}^{m} & \alpha_{1-1}^{m} & \alpha_{10}^{m}\dots\alpha_{\nu\mu}^{m}\end{bmatrix}^T$$

$$\mathbf{C}_m = \begin{bmatrix}C_{|m|m} & C_{(|m|+1)m}\dots C_{Nm}\end{bmatrix}^T$$

$$\mathbf{H}_m = \begin{bmatrix} H_{|m|0}^{m0} & H_{(|m|+1)0}^{m0} & \cdots & H_{N0}^{m0} \\ H_{|m|1}^{m(-1)} & H_{(|m|+1)1}^{m(-1)} & \cdots & H_{N1}^{m(-1)} \\ \vdots & \vdots & \ddots & \vdots \\ H_{|m|\nu}^{m\mu} & H_{(|m|+1)\nu}^{m\mu} & \cdots & H_{N\nu}^{m\mu} \end{bmatrix}$$

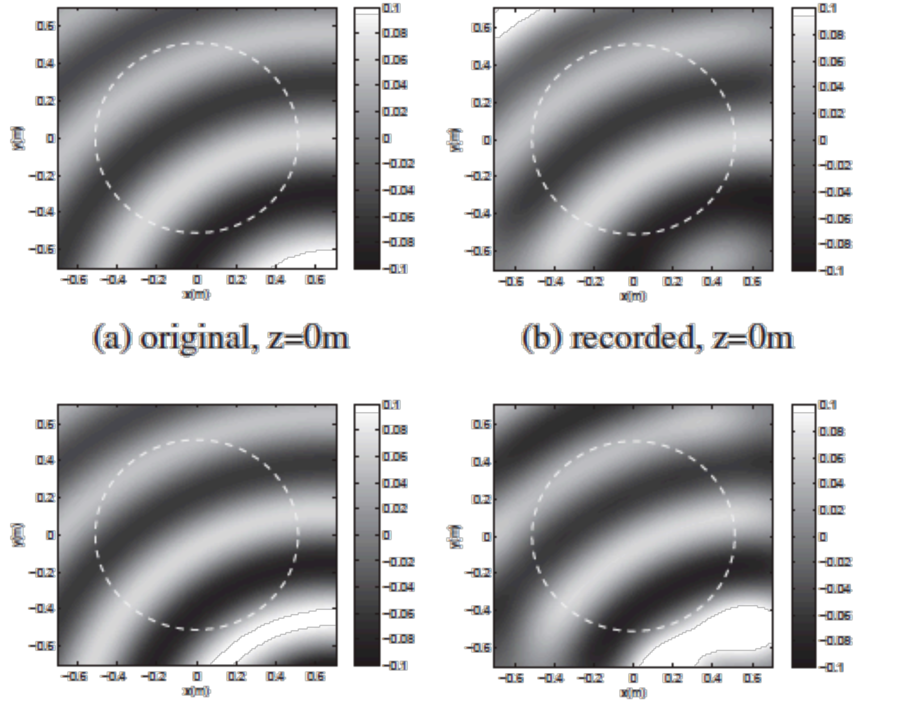
Application of the translation operator for SHD: higher order microphone arrays

- The direct application of the pseudoinverse yields unstable results.
- By applying *K* circular higher order microphone arrays the robustness and precision of the microphone system is increased:

$$\mathbf{C}_m = (\hat{\mathbf{H}}_m^* \hat{\mathbf{H}}_m + \lambda \mathbf{I})^{-1} \hat{\mathbf{H}}_m^* \hat{\alpha}_m$$

$$\hat{\mathbf{H}}_m = [\mathbf{H}_{m;1}{}^T \mathbf{H}_{m;2}{}^T \ldots \mathbf{H}_{m;K}{}^T]^T$$

$$\hat{\alpha}^m = [\alpha_{m;1}{}^T \alpha_{m;2}{}^T \ldots \alpha_{m;K}{}^T]^T$$



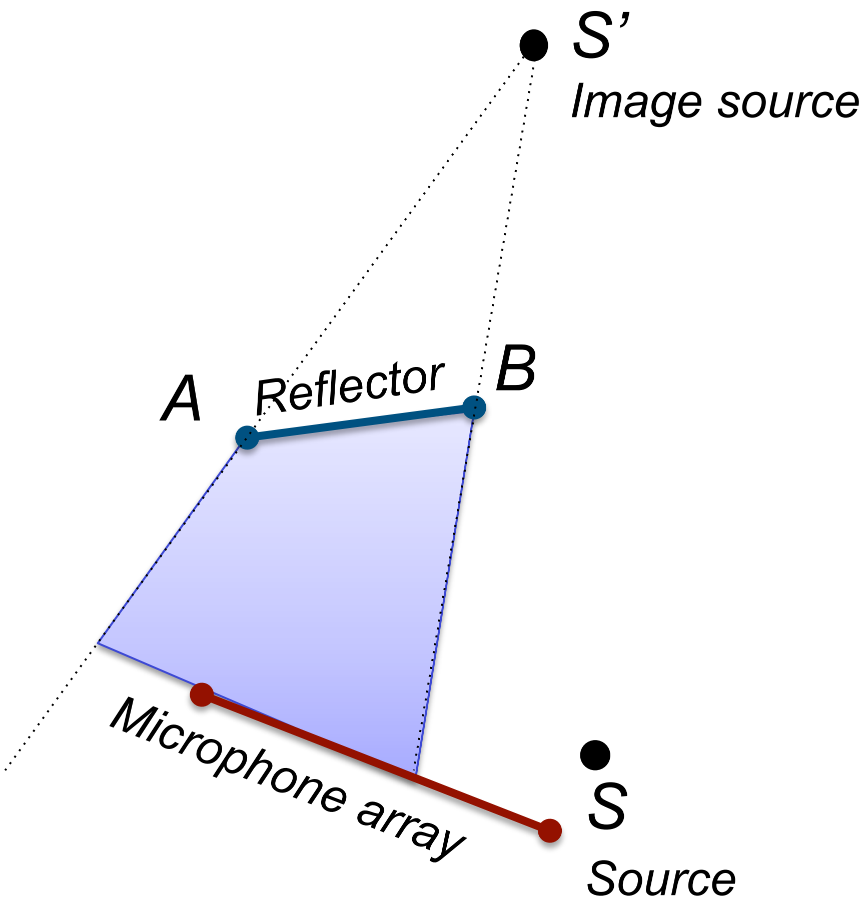
(a) original, z=0m (b) recorded, z=0m

(c) original, z=0.2m (d) recorded, z=0.2m

1st order mics arranged into four circular arrays, placed in (R_s,ϑ_s) = (0.4, 90°), (0.34, 72°), (0.28, 108°) and (0.22, 72°), the number of first order mics on each array is 17, 15, 13 and 11, respectively [Chen2015]

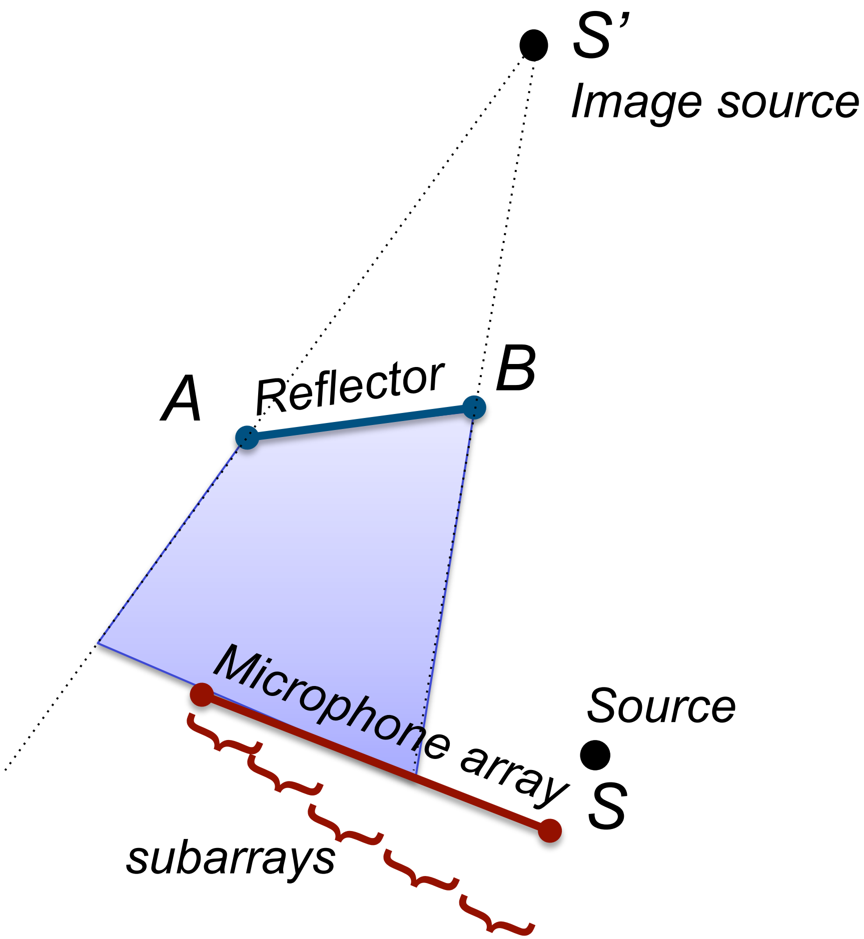
Limits of the PWD

Consider the following acoustic scene:



- The image source *S'* is only visible to some of the mics of the array, therefore a PW analysis using the whole array would fail
- More generally, PW analysis fails when the components are not space-invariant [Lalor1968].

Work around the PWD limits



- Performing PW analysis on subarrays alleviates the problem: space invariance will now concern only one of the subarrays
- **New issues**:
  - how do we merge the information acquired by each sub-array?
  - Which representation of the sound field should we use?

- We need a representation of the sound field that describes the plane wave components as a function of the spatial location: the **plenacoustic function** $p(x, y, \theta, \omega, t)$ [Ajdler2003, Ajdler2005, Ajdler2006]
- The most immediate parameterization of the plenacoustic function is in terms of acoustic rays: we can think of the plane wave component passing through (x, y) and with direction θ as an acoustic ray

Sound field map: representation of the plenacoustic function using the tools of geometrical acoustics

- Do we need to use three variables to describe the sound field map, or may we reduce the dimensionality?
  - The **Radiance Invariance Law** (RIL) states that the acoustic radiance is invariant along a ray, i.e. the plenacoustic function is constant along the line of direction θ , passing through (x, y)
  - The dimensionality of the space is therefore the same as that of the parameters that describe the ray, i.e. 2

We need to find a suitable parameterization for the acoustic rays

Parameterizations of acoustic rays

Acoustic ray

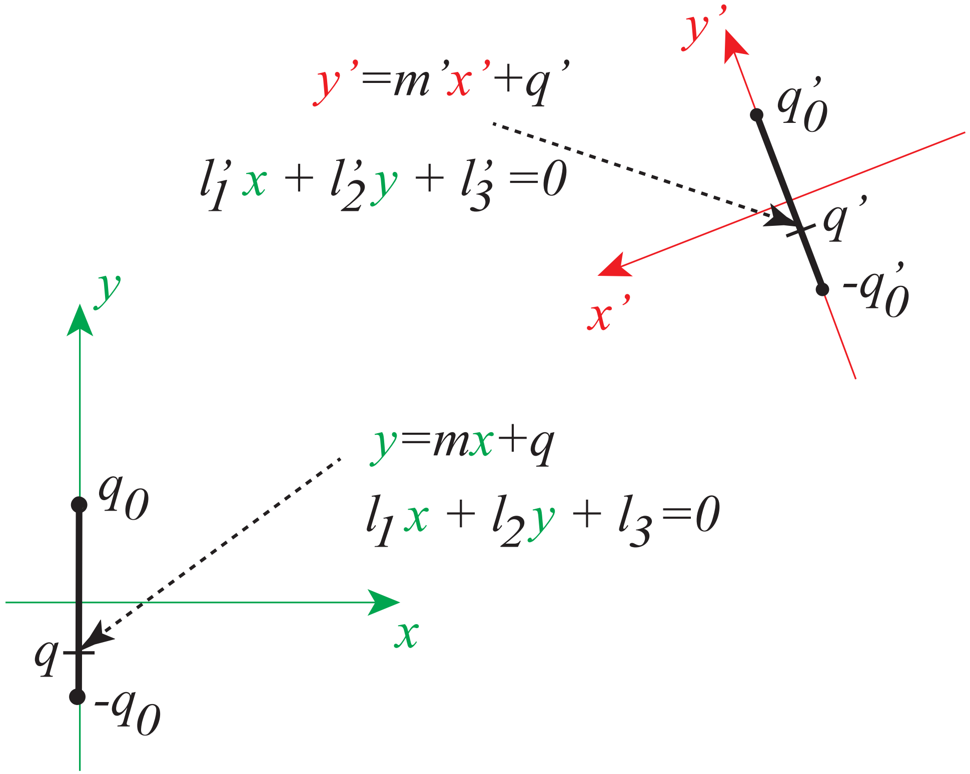
- Different ways to parameterize rays in 2D:

Line parameters *(global)*

$$l_1 = k\sin(\theta)$$
$$l_2 = -k\cos(\theta)$$
$$l_3 = k[y\cos(\theta) - x\sin(\theta)],\ k > 0$$

Slope and intercept *(local)*

$$m = -\frac{l_1}{l_2},\ q = -\frac{l_3}{l_2}$$

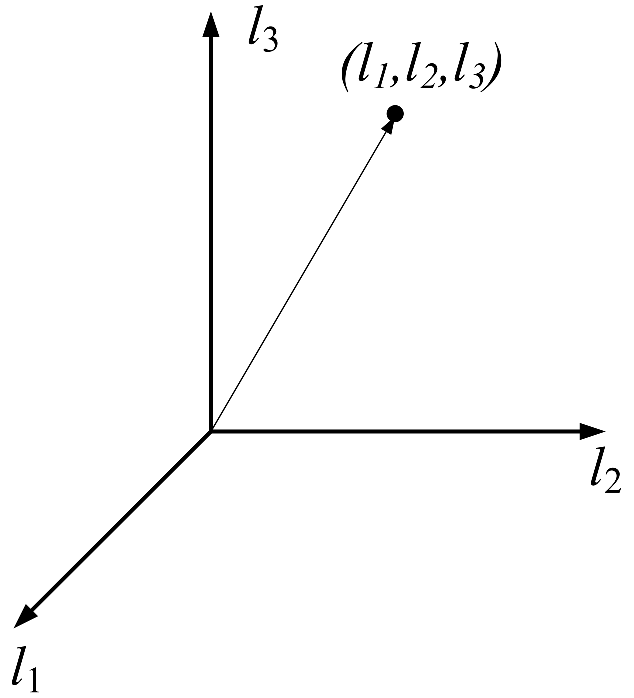


The Ray Space

A **ray** is a point in the projective space P^2

Its coordinates $[l_1,l_2,l_3]^T$ form a class of equivalence, as $[kl_1, kl_2, kl_3]^T, k\neq 0$ are all the same ray

The Euclidean space spanned by such homogeneous coordinates of lines is called **ray space**



Geometric primitives in the Ray Space: point (sources and receivers)

Let $[x_1,x_2]^T$ be the Euclidean coordinates of a **point** in the geometric space. Its homogeneous coordinates in $\mathbb{P}^2$ are $\mathbf{x}=[x_1,x_2,1]^T$

This point lies on the line $\mathbf{l}=[l_1,l_2,l_3]^T$ iff $\mathbf{x}^T\mathbf{l}=0$

If **x** lies in the intersection between $\mathbf{l}_1$ and $\mathbf{l}_2$, it will also lie in the intersection between $k\mathbf{l}_1$ and $k\mathbf{l}_2$ $(k\neq 0)$

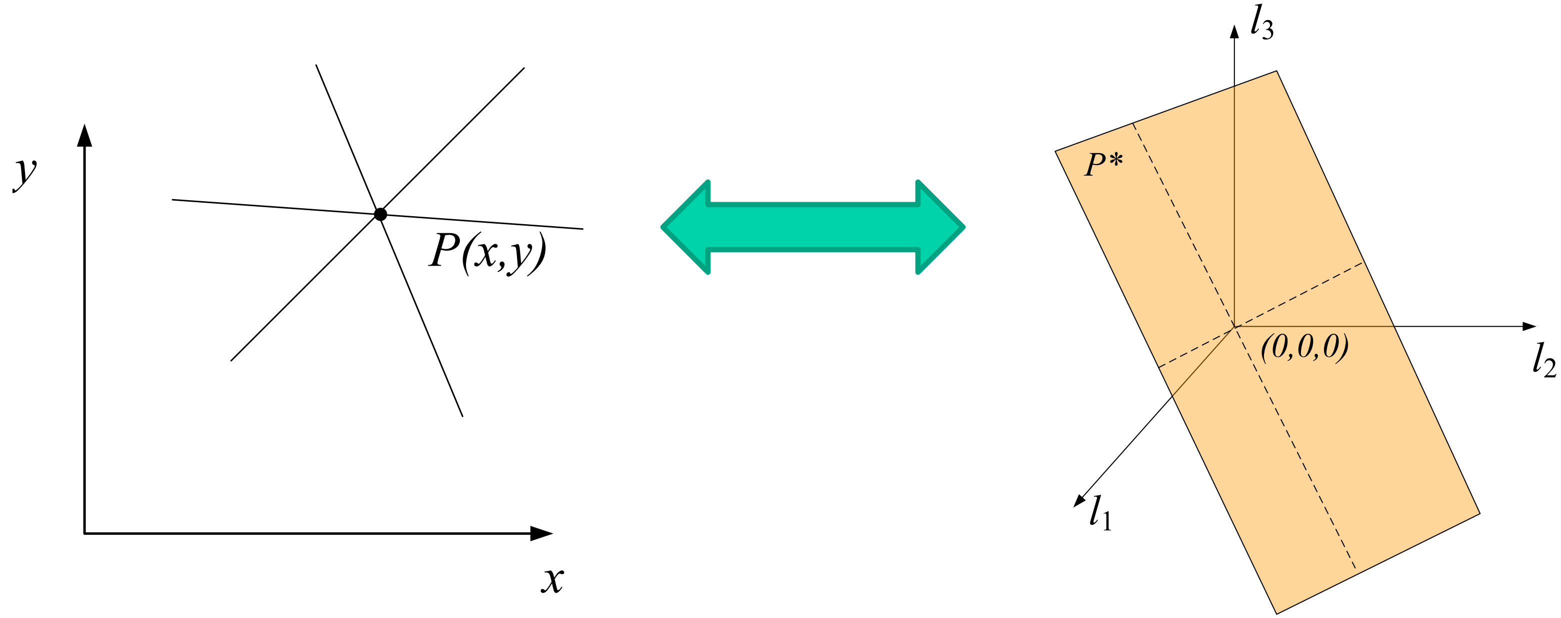
$$\begin{array}{c} \mathbf{x}^T\left(k\mathbf{l}_1\right)=0 \\ \mathbf{x}^T\left(k\mathbf{l}_2\right)=0 \end{array} \Rightarrow \begin{array}{c} \left(k\mathbf{x}^T\right)\mathbf{l}_1=0 \\ \left(k\mathbf{x}^T\right)\mathbf{l}_2=0 \end{array}$$

therefore $k\mathbf{x}$ will be the same point as $\mathbf{x}$

This means that $\mathbf{x}$ is homogeneous as well

Geometric primitives in the Ray Space: Point (sources and receivers)

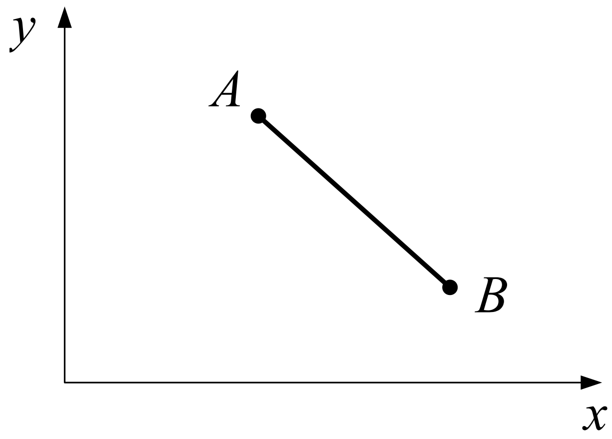
A **point** is identified by the set of all rays that pass through it



$$P^* = \{ (l_1, l_2, l_3) \mid l_1 x_1 + l_2 x_2 + l_3 = 0 \}$$

In the ray space this set of lines corresponds to a plane passing through the origin, whose normal is $[x_1, x_2, 1]^T$

Geometric primitives in the Ray Space: Acoustic Reflector

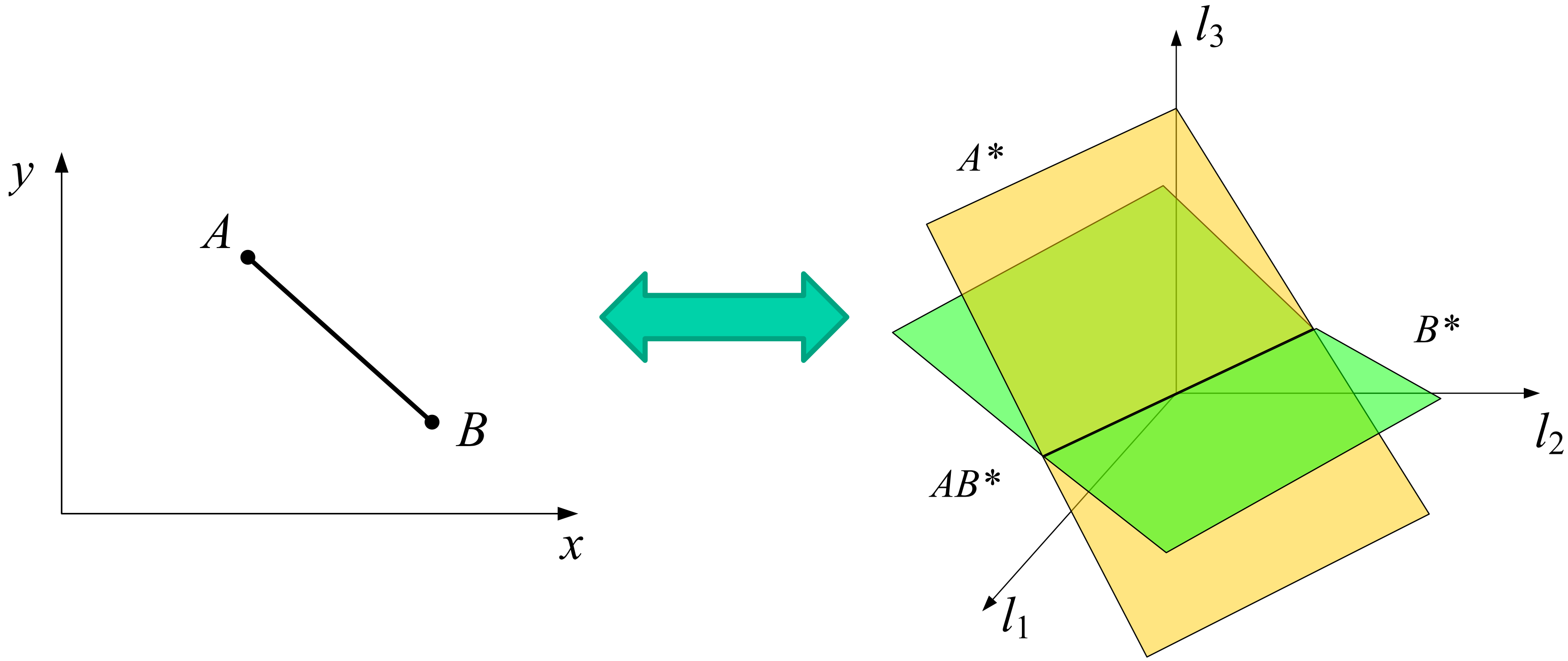


A (bounded) **reflector** is identified by its endpoints *A* and *B*

In the ray space this reflector is represented by the set of rays that pass through the intermediate points between A and B

Geometric primitives: Acoustic reflector

Geometric vs. ray-space representation of the **reflector**

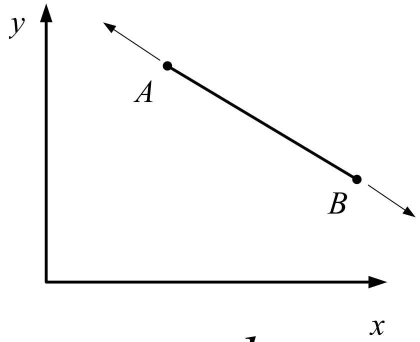


Geometric space

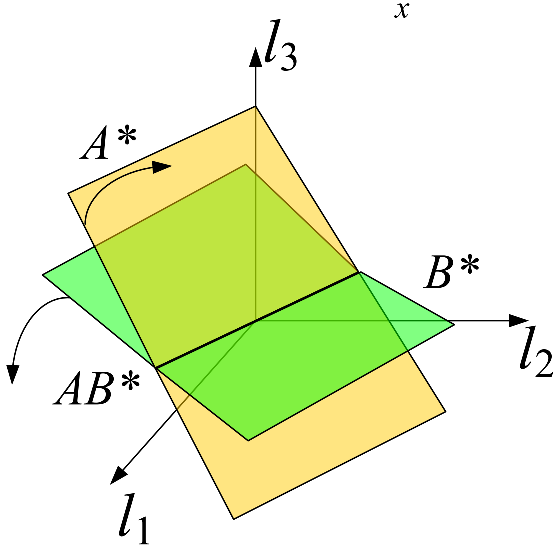
Ray space

Geometric primitives: Unbounded reflector

An **unbounded reflector** can be thought of as obtained by indefinitely stretching a bounded reflector



Stretching the reflector corresponds to widening the corresponding wedge (between *A** and *B**)



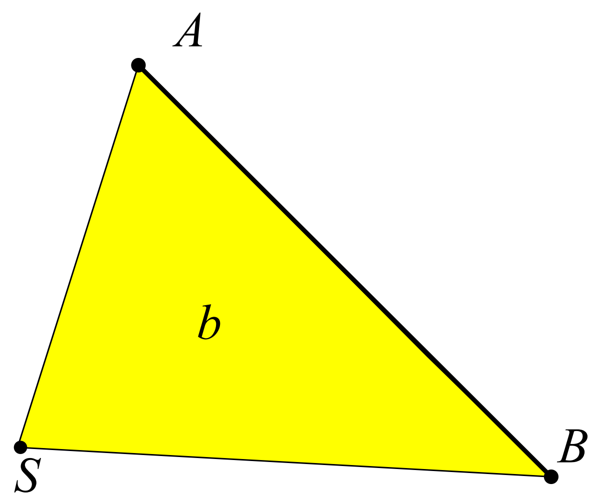
At the end the wedge becomes the whole space

- Any ray meets an infinite plane (possibly at infinity)

Geometric primitives: Acoustic beam

A **beam** is a *connected* bundle of rays that originate from the same point (source) and fall onto the same reflector (or a portion of it)

- Note: the reflector can also be at infinity



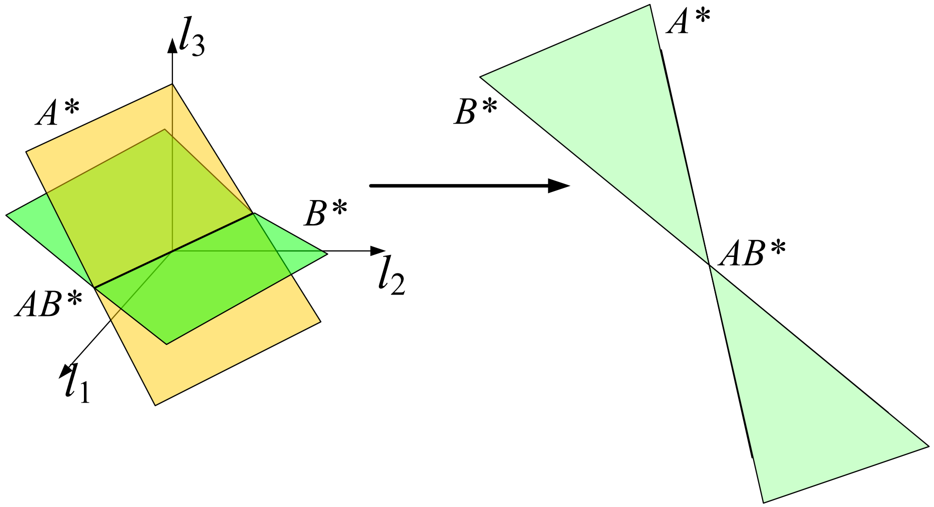
As a primitive, the **beam** is the intersection between other primitives

- Set of rays originating from the source
- Connected region of the reflector illuminated by the source

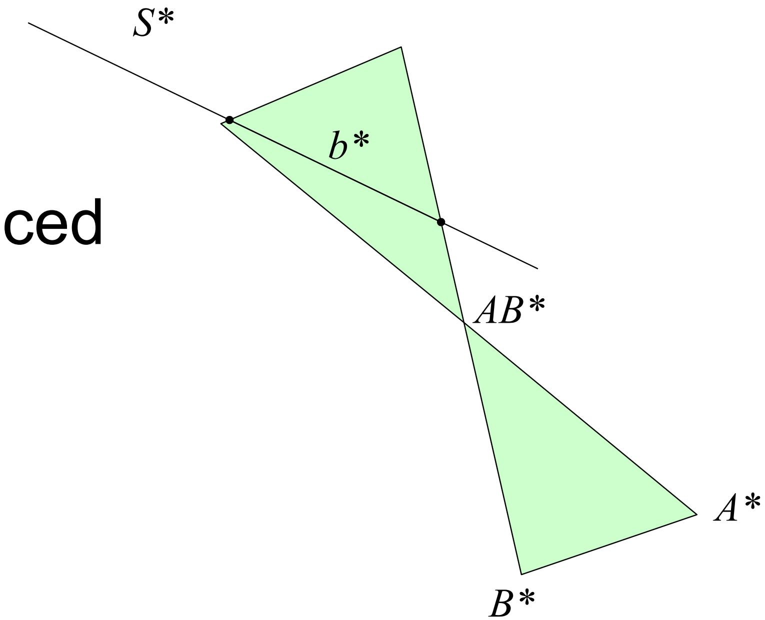
Intersection in geometric (primal) space corresponds to an intersection in ray (dual) space

Geometric primitives: The reduced ray space

A reflector is more easily represented in a reduced (2D) space, obtained through an arbitrary cross-section

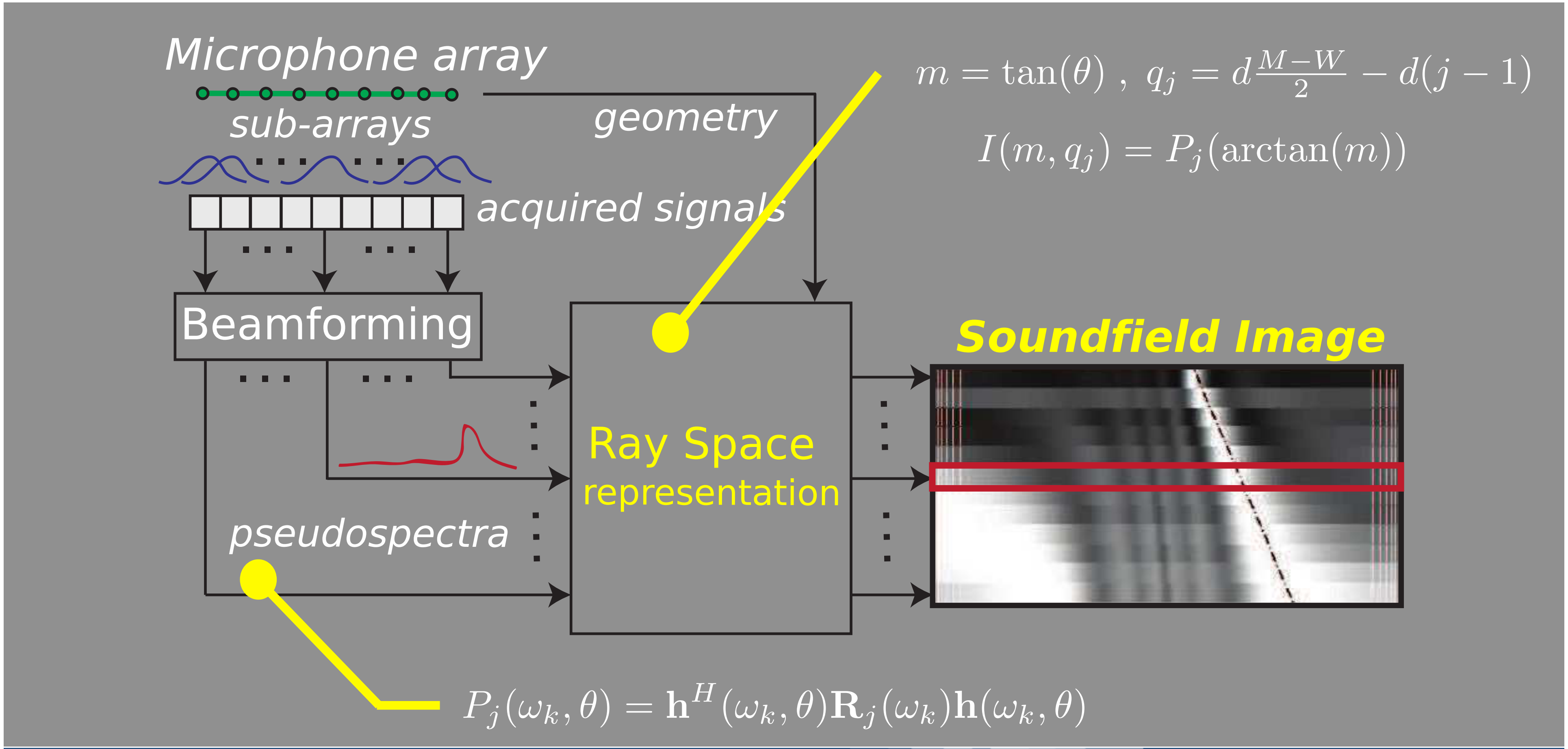


Intersection in the reduced space

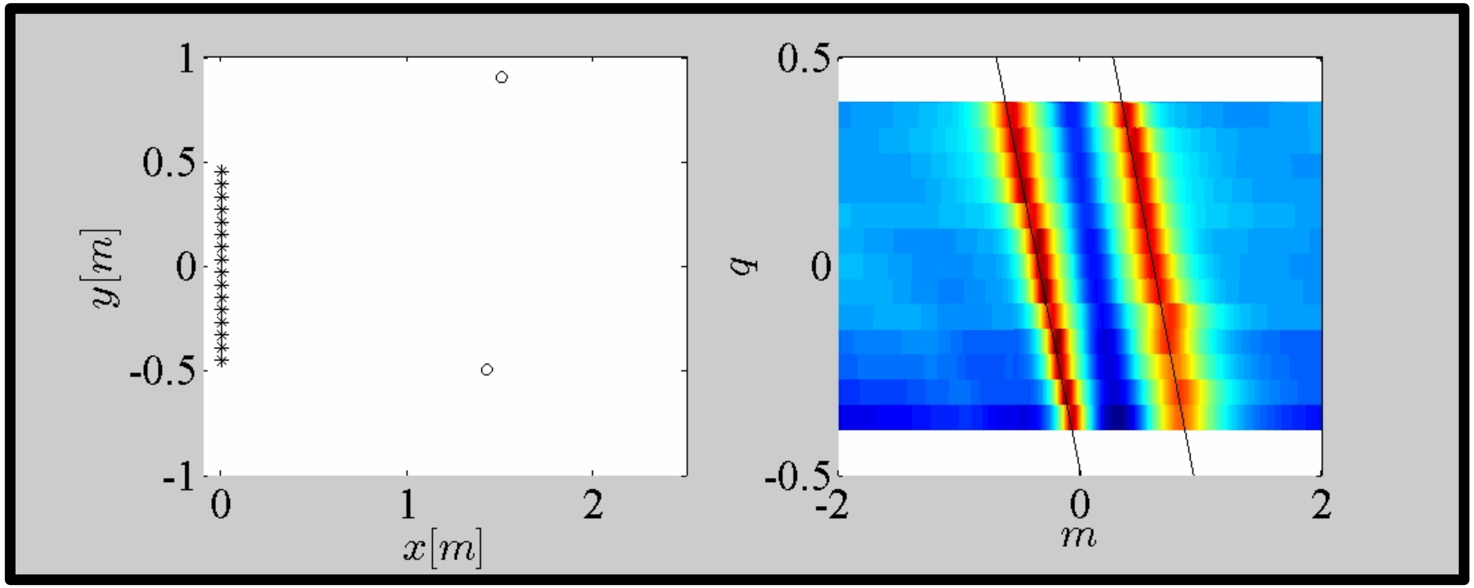


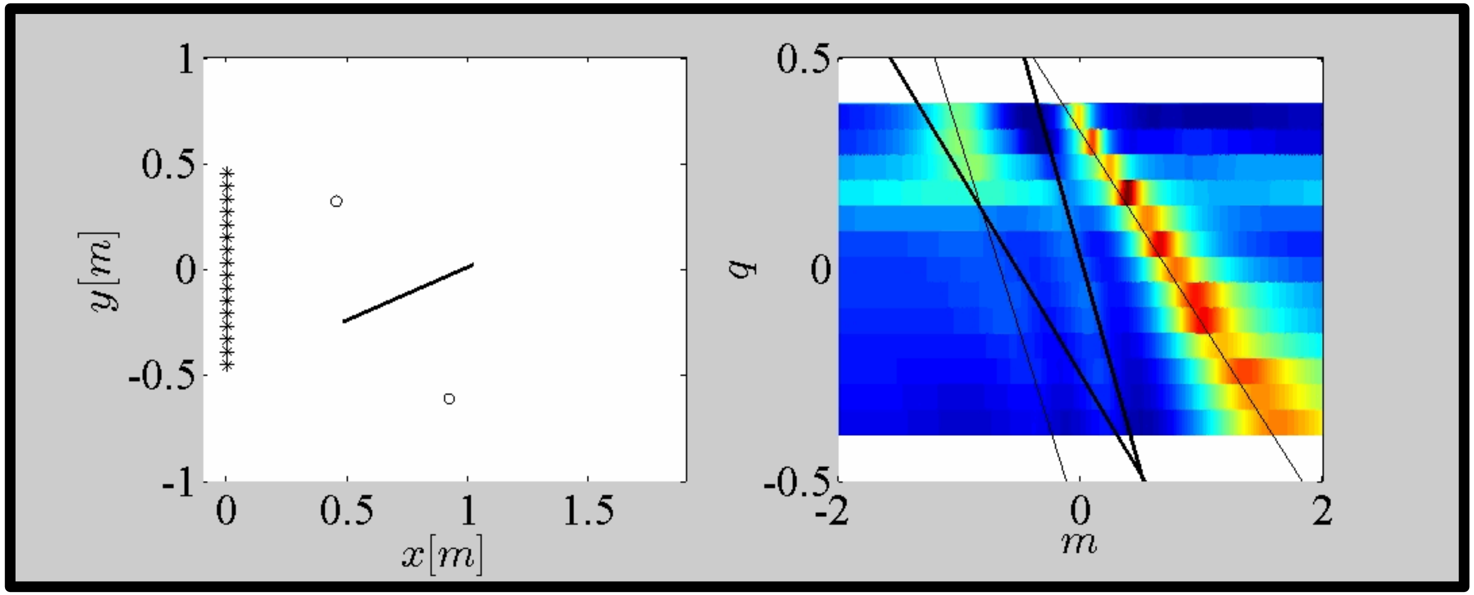
Measuring the sound field map: the Sound Field Camera

- In order to measure the sound field map we need to devise a measuring methodology
- Recall: analogy between the beamforming operation and the plane wave decomposition

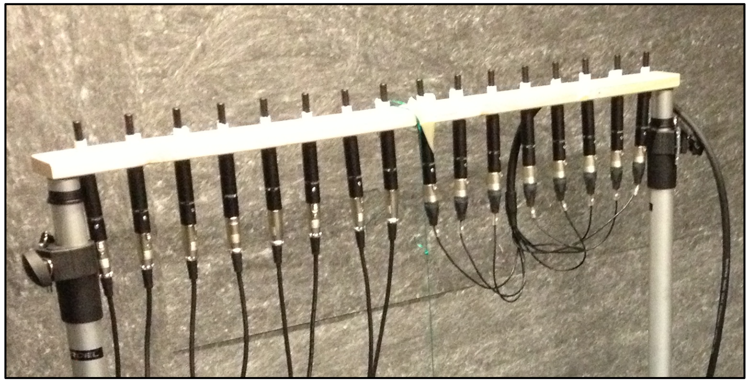


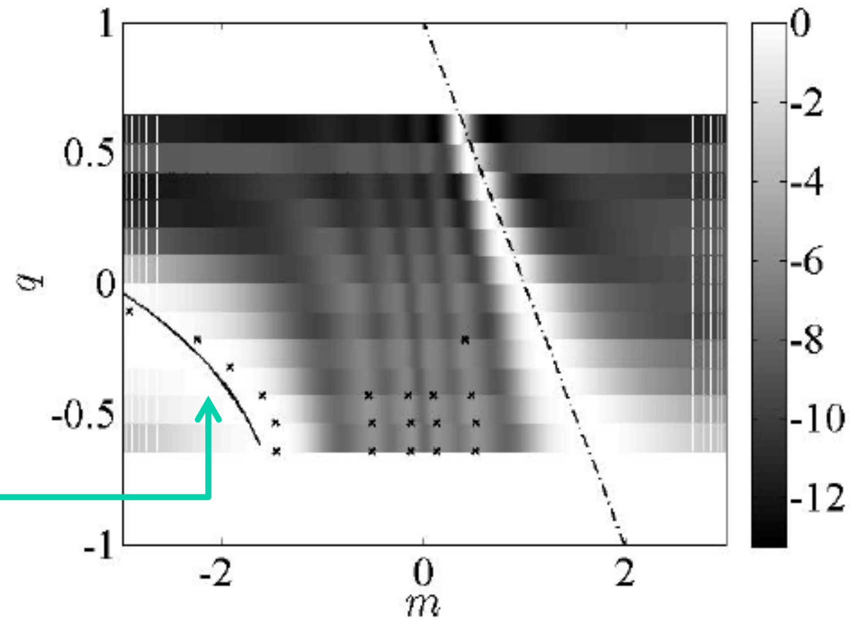
Examples of sound field images





Issues of sound field imaging: Aliasing





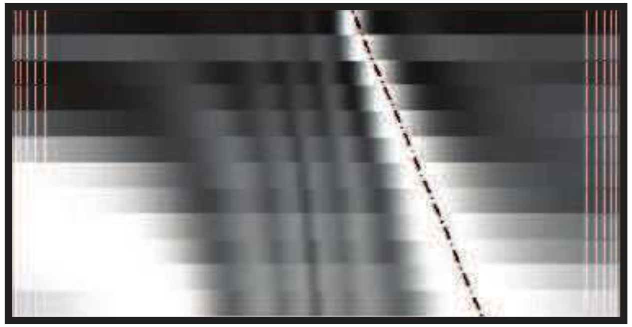
From [Marković2013]

Aliasing patterns can be easily identified:

- Nonlinear
- Frequency-dependent

$$\theta = \arcsin\left(\frac{l\lambda}{d} + \sin\theta_i\right)\ , -\pi/2 \leq \theta < pi/2\ , l \in \mathbb{Z}$$

Issues of sound field imaging: resolution



$$H_j(\omega_k,\theta) = ||\mathbf{a}(\omega_k,\theta)^H \mathbf{a}(\omega_k,\theta_{j0}d_{jo})/W||^2$$

$$H_j(\theta) = \frac{1}{W^{2K}} \left\| \prod_{k=1}^{K} \mathbf{a}(\omega_k,\theta)^H \mathbf{a}(\omega_k,\theta_{j0}d_{jo}) \right\|^2$$

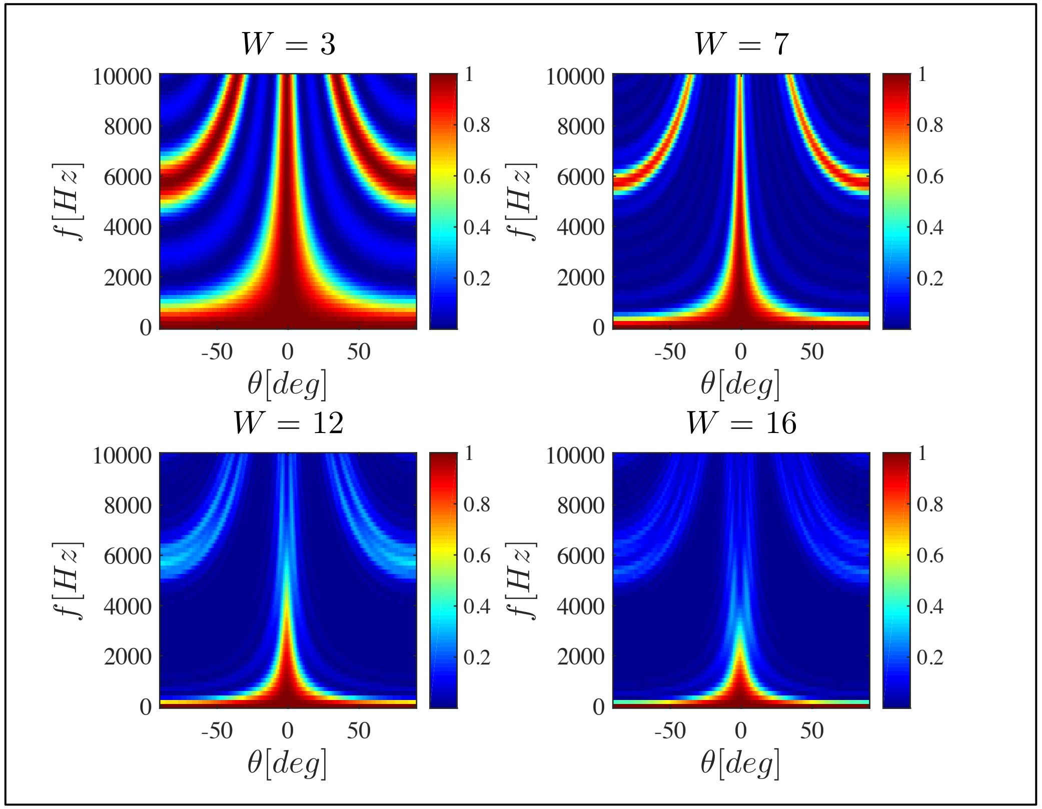
$$H_j(\omega_k,\theta) = \frac{1}{W^2} \frac{\sin\left[\frac{\omega_k dW}{2c}(\sin\theta - \sin\theta_{j,0})\right]^2}{\sin\left[\frac{\omega_k d}{2c}(\sin\theta - \sin\theta_{j,0})\right]^2}$$

$$H_j(\omega_k,\theta) = \left\| \frac{1}{W} \sum_{i=1}^{W} e^{j\omega_k \frac{\left[d(i-\frac{W+1}{2})\sin(\theta) - \Delta d_i(\theta_{j,0},d_{j,0})\right]}{c}} \right\|^2$$

Resolution improves with the number W of mics
of the sub-arrays. But is it always true?

Issues of sound field imaging: resolution

Beampattern (beamformer tf function) of a source at frequency *f* [Hz] coming from θ_0=0°, for different subarray sizes



From [Marković2015]

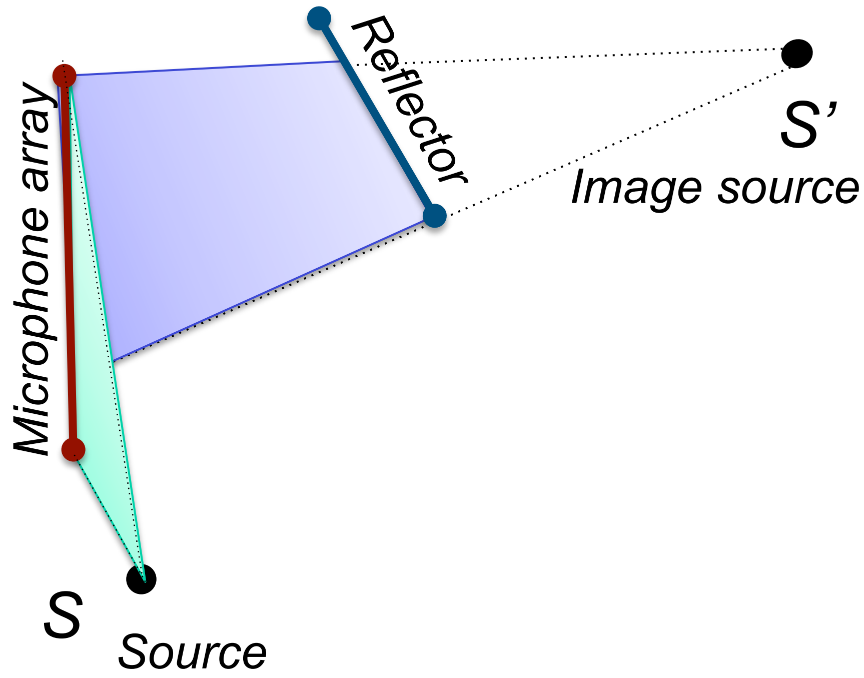
For large subarrays and high frequencies the beampattern exhibits attenuation: the assumption of the sub-array in the far-field fails!

Issues of sound field imaging: resolution

Property	Impacts	Related to
Blurring ($W\nearrow$)	Ability to discern different objects	Lobe of the point-spread function $2[\theta_{j,0} - \arcsin(\frac{2\pi c}{\omega_k dW} + \sin(\theta_{j,0}))]$
Sampling ($W\searrow$)	The field of view	The total number of the sub-arrays $M - W + 1$
Focus ($W\searrow$)	Ability to observe the object at certain distance	Attenuation $H_j(\omega_k, \theta_{j,0})$

From [Marković2015]

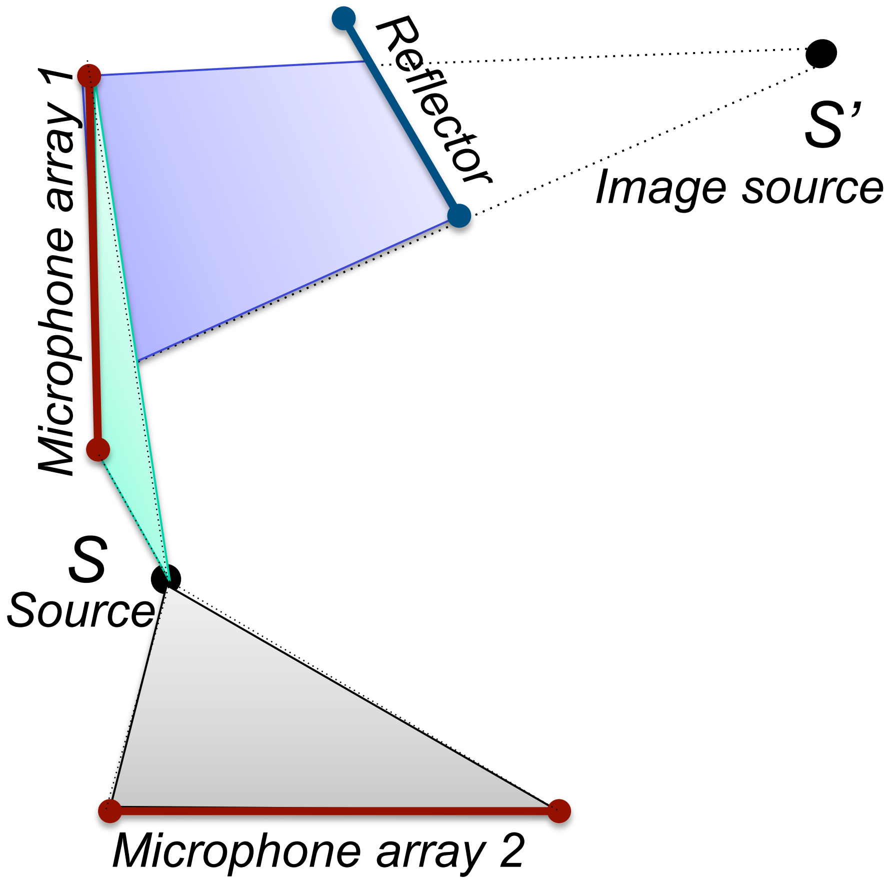
Fusing multiple ray spaces: the projective ray space



The array "sees" the source *S* from a "difficult" angle (*m* tends to infinity), which causes loss of resolution in the sound field image

Using multiple soundfield cameras we can be sure that sources are always effectively "viewed"

Fusing multiple ray spaces: the projective ray space



- We need a ray space where we can fuse the information coming from both cameras
- The reduced ray space *(m,q)* is not suitable, as not all rays from both cameras can be represented (it has blind regions)
- Use of the projective parameterization [3]:

$$\mathbf{l} = [l_1, l_2, l_3]^T$$

$$\mathcal{L} : \{\mathbf{x} = [x, y]^T : \mathbf{x}^T\mathbf{l} = 0\}$$

Fusing multiple ray spaces: the projective ray space

How can we merge the information coming from multiple cameras?

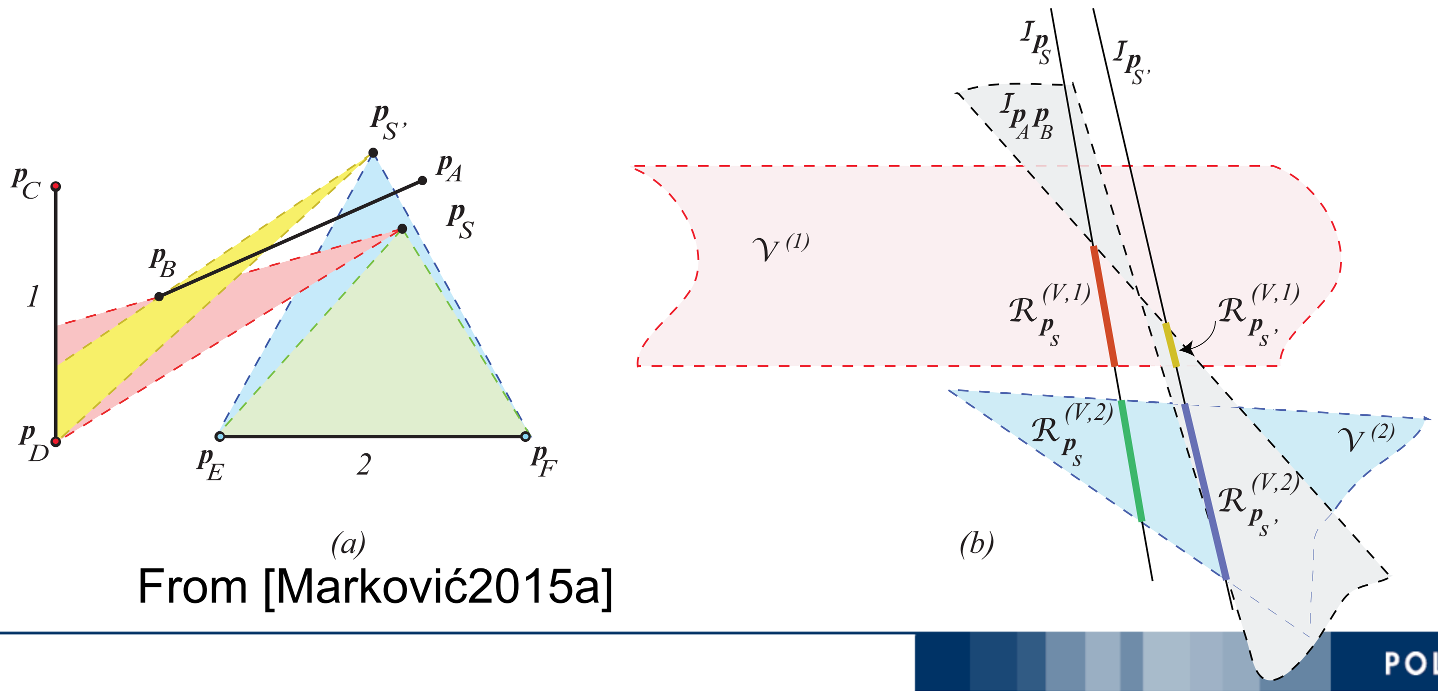
- Change of reference frame for projective coordinates:

$$\mathbf{p}_A = \mathbf{H}^{(i)}\mathbf{p}_A^{(i)} \quad \mathbf{H}^{(i)} = \begin{bmatrix} \mathbf{R}^{(i)} & \mathbf{t}^{(i)} \\ \mathbf{0} & 1 \end{bmatrix}$$

- Change of reference frame for the ray space:

$$\mathbf{l} = (\mathbf{H}^{(i)})^{-T}\mathbf{l}^{(i)} \Rightarrow m = -\frac{l_1}{l_2},\ q = -\frac{l_3}{l_2}$$

Geometric space *Reduced ray space*



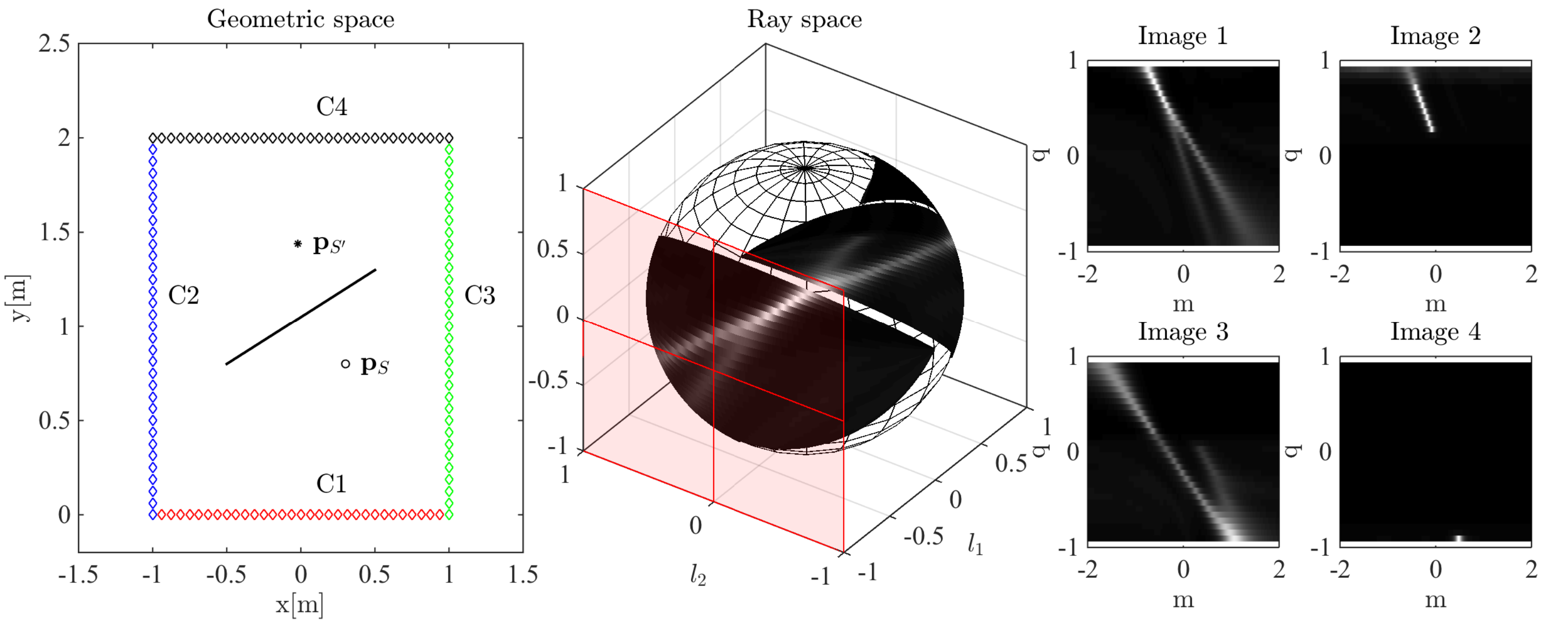
(a) (b)

From [Marković2015a]

Fusing multiple ray spaces: the projective ray space

$$\mathbf{p}_A = \mathbf{H}^{(i)}\mathbf{p}_A^{(i)}$$

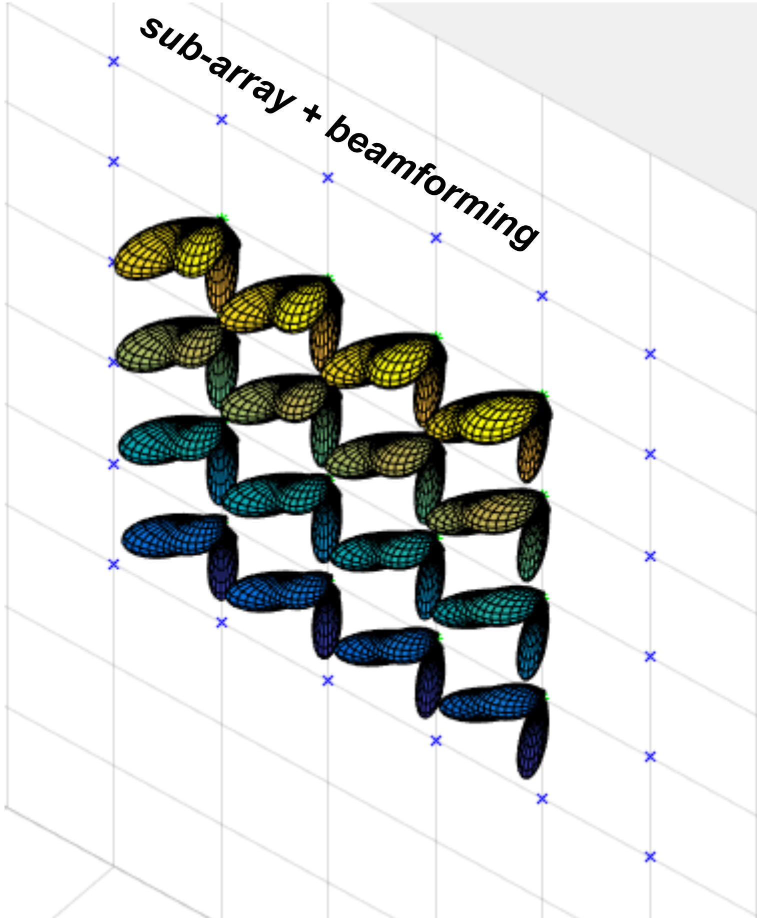
$$\mathbf{H}^{(i)} = \begin{bmatrix} \mathbf{R}^{(i)} & \mathbf{t}^{(i)} \\ \mathbf{0} & 1 \end{bmatrix} \Rightarrow \mathbf{l} = (\mathbf{H}^{(i)})^{-T}\mathbf{l}^{(i)} \Rightarrow m = -\frac{l_1}{l_2},\ q = -\frac{l_3}{l_2}$$

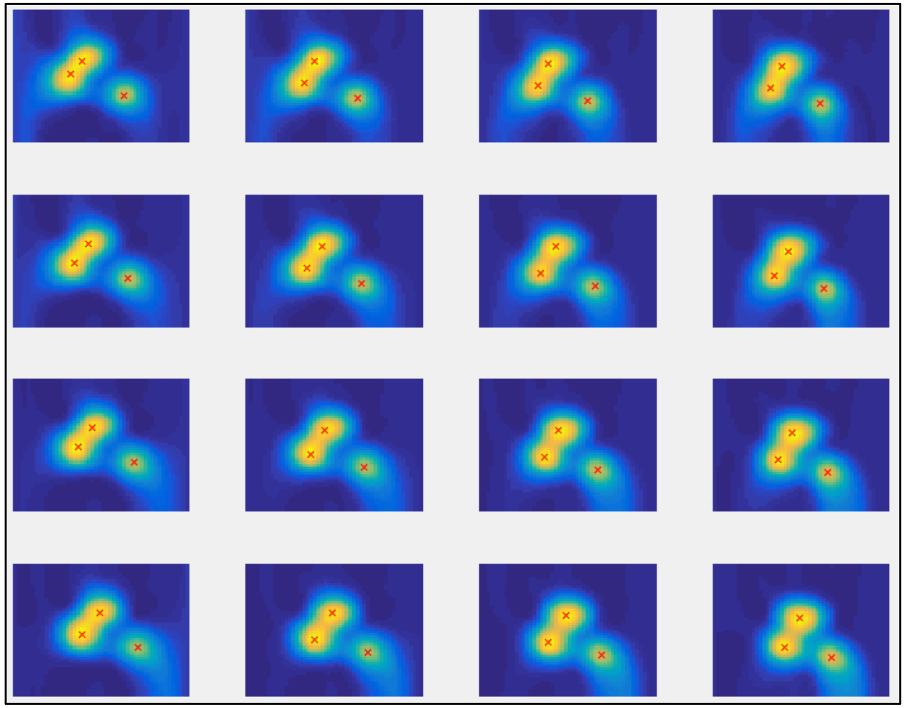


From [Marković2015a]

From 2D to 3D

How can we extend from 2D to 3D?





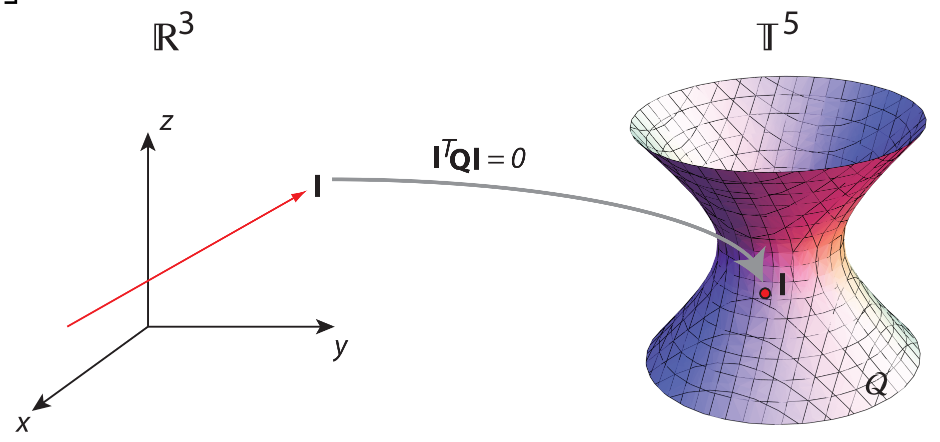
(feature detection = local maxima = morphological dilation)

How can we represent rays that propagate in a 3D world?

- Rays in 3D are identified by at least four parameters (see lightfield representation).
- We adopt the (redundant) Plücker parameterization of the acoustic rays:

$$\mathbf{l} = k \begin{bmatrix} \mathbf{p}_B - \mathbf{p}_A \\ \mathbf{p}_A \times \mathbf{p}_B \end{bmatrix} \qquad \mathbf{p}_A = \begin{bmatrix} x_i + \sin(\phi)\cos(\theta) \\ y_i + \sin(\phi)\sin(\theta) \\ z_i + \cos(\phi) \end{bmatrix} \qquad \mathbf{p}_B = \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix}$$

$$\Rightarrow \quad \mathbf{l} = \begin{bmatrix} \mathbf{l}^d \\ \mathbf{l}^m \end{bmatrix}, \; s.t. \; (\mathbf{l}^d)^T \mathbf{l}^m = 0 \quad Q = \{\mathbf{l} \in \mathbb{T}^5 \mid \mathbf{l}^T \mathbf{Q} \mathbf{l} = 0\}, \quad \mathbf{Q} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_3 \\ \mathbf{I}_3 & \mathbf{0} \end{bmatrix}$$



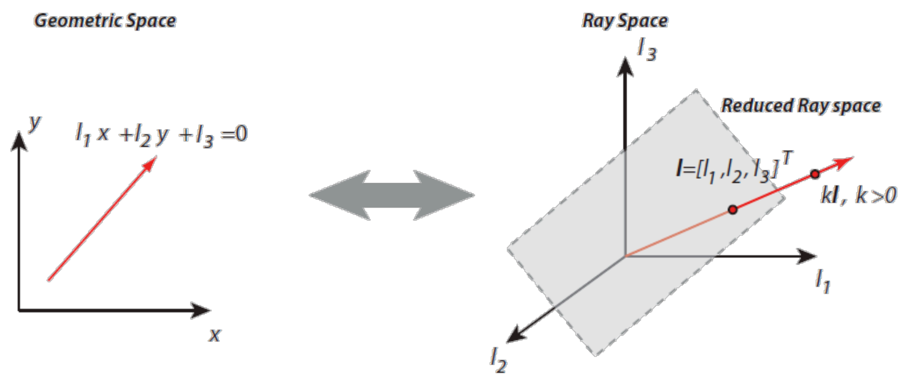
(a) (b)

The Ray Space

2D

Homogeneous coefficients of lines

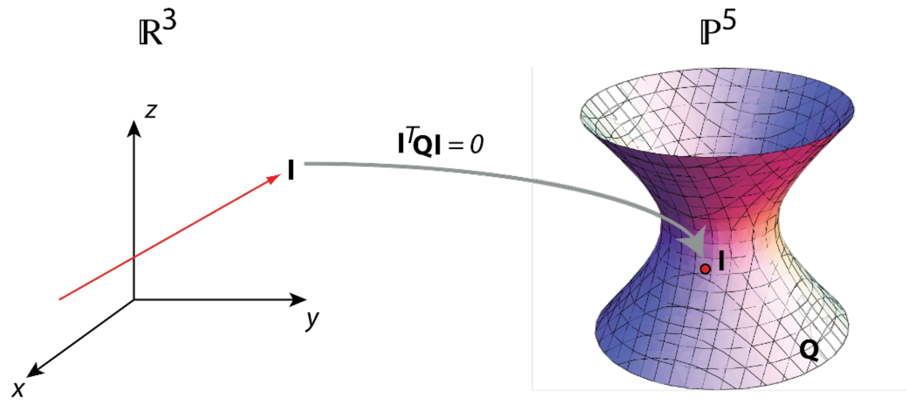
Oriented projective space $\mathbf{P}^2$



3D

Plücker coordinates of lines

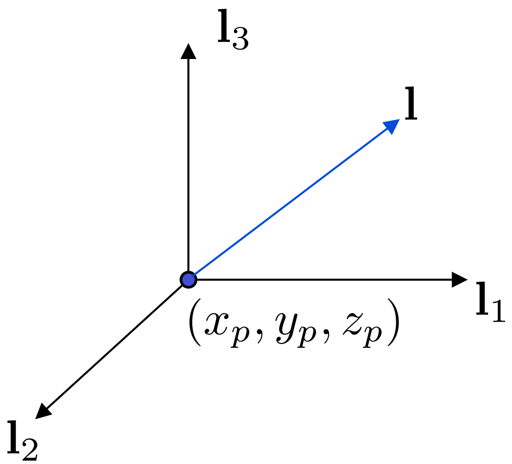
Oriented projective space $\mathbf{P}^5$



(reduced dimensionality representation)

Given a point $\mathbf{p} = [x_p, y_p, z_p]^T$ we define the matrix $\mathbf{L} = \left[\widetilde{l}_1, \widetilde{l}_2, \widetilde{l}_3\right]^T$

$$\mathbf{L} = \begin{bmatrix} 0 & z_P & -y_P & 1 & 0 & 0 \\ -z_P & 0 & x_P & 0 & 1 & 0 \\ y_P & -x_P & 0 & 0 & 0 & 1 \end{bmatrix}$$



$$\mathbf{U}_P = \begin{bmatrix} 0 & 1/z_P & x_P/z_P \\ 1/x_P & y_P/(x_P z_P) & y_P/z_P \\ 0 & 0 & 1 \\ -z_P/x_P & -y_P/x_P & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\mathbf{l} = \mathbf{u}_1 a_1 + \mathbf{u}_2 a_2 + \mathbf{u}_3 a_3 = \mathbf{U}_P \mathbf{a}$$

$$\mathcal{I}_{\mathbf{p}_P} = \{\mathbf{l} \in \mathbb{T}^5 | \mathbf{l} = \mathbf{U}_P \mathbf{a}, \mathbf{a} \in \mathbb{T}^2\}$$

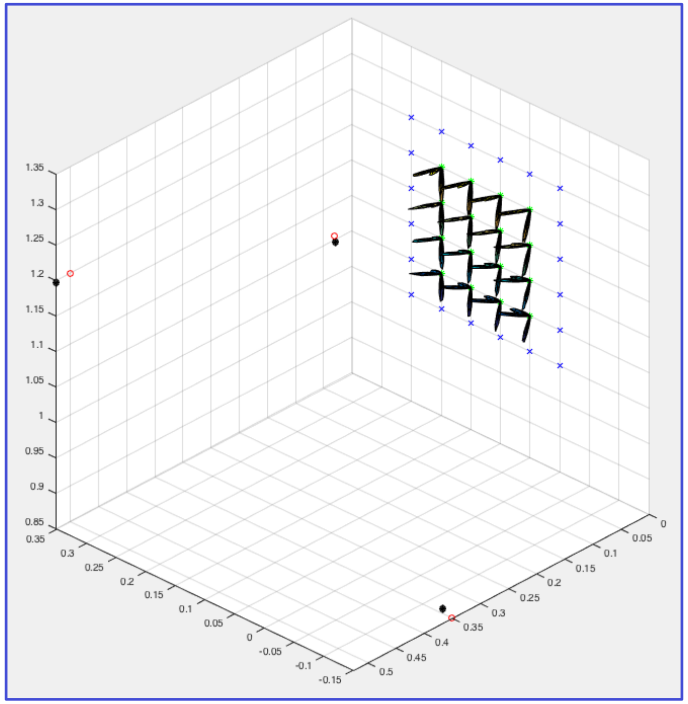
SUBSPACE FITTING

"Given the clusters how do we find the source position?"

$$\mathbf{X} = [\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_N}]$$

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^T$$

$$\mathbf{U}_P = [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3]$$



$$\mathbf{L} = \begin{bmatrix} 0 & z_P & -y_P & 1 & 0 & 0 \\ -z_P & 0 & x_P & 0 & 1 & 0 \\ y_P & -x_P & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{U}_P = \begin{bmatrix} 0 & 1/z_P & x_P/z_P \\ 1/x_P & y_P/(x_P z_P) & y_P/z_P \\ 0 & 0 & 1 \\ -z_P/x_P & -y_P/x_P & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

- Consider the sound field map of a point source
  - The length of each subarray is assumed small compared to the distance of the acoustic source
  - Assume far field propagation at subarray level, so that the field observed by the i^{th} subarray is

$$f_i(z,\omega) = \exp\left(jkz\sin(\theta_i')\right)$$

    - θ_i' is the angle under which the source is observed by the ith subarray

- Sound field map analysis

$$[\mathbf{F}]_{i,w}(\omega) = d\sum_{l=0}^{L-1} \exp(jkz\Theta_{i,w})\mathrm{rect}\left(\frac{z-q_i}{\nu}\right)$$

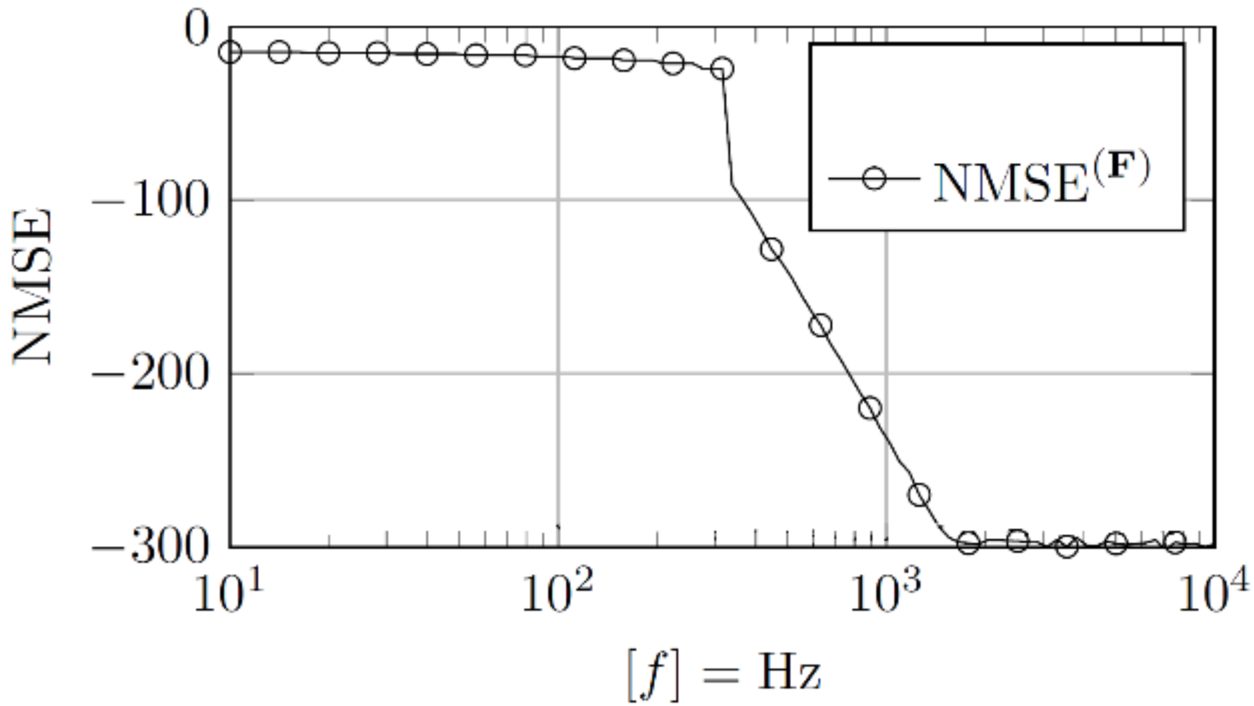
$$\Theta_{i,w} = \sin(\theta_i') - \sin(\theta_w)$$

- Sound field map synthesis

$$p^{(\mathbf{F})}(z,\omega) = \sum_{i=0}^{I-1}\sum_{w=0}^{W-1}[\mathbf{F}]_{i,w}(\omega)e^{jk\frac{m_w}{\sqrt{1+m_w^2}}}\tilde{\mathrm{rect}}\left(\frac{z-q_i}{\nu}\right)$$

  - The function $\tilde{\mathrm{rect}}(\cdot)$ has infinite length, thus it must be truncated

- NMSE due to sound field map synthesis as a function of frequency



From [Bianchi2016]

  - Large error at mid-low temporal frequencies
    - Model mismatch: the sound source is not in the far field of the subarrays
    - Errors due to the truncation of $\widetilde{\mathrm{rect}}(\cdot)$

References

[Chen2015] H. Chen, T. D. Abhayapala and W. Zhang, 3D sound field analysis using circular higher-order microphone array, 23rd European Signal Processing Conference (EUSIPCO), 2015, Nice, 2015, pp. 1153-1157.

[Lalor 1968] Éamon Lalor, "Conditions for the Validity of the Angular Spectrum of Plane Waves*," J. Opt. Soc. Am. 58, 1235-1237 (1968)

[Bianchi2016] L. Bianchi, F. Antonacci, A. Sarti, and S. Tubaro. The Ray Space Transform: a new Framework for Wavefield Processing. *IEEE Trans. Signal Process.,* doi: 10.1109/TSP.2016.2591500.

[Ajdler2003] T.Ajdler and M.Vetterli, "The Plenacoustic function and its sampling", in proc. of Workshop on Applications of Signal Processing to Audio and Acoustics, WASPAA, 2003

[Ajdler2006] T.Ajdler, L.Sbaiz and M.Vetterli, "The Plenacoustic Function and Its Sampling," IEEE Transactions on Signal Processing, vol.54, no.10, pp.3790-3804, Oct. 2006

[Ajdler2005] T. Ajdler , L. Sbaiz , M. Vetterli "The plenacoustic function on the circle with application to HRTF interpolation", In proceedings of IEEE ICASSP (2005)

References

[Marković2013] D. Marković, F. Antonacci, A. Sarti, S. Tubaro, "Soundfield imaging in the ray space" IEEE/ACM Transactions on Audio, Speech and Language Processing, volume 21, issue 12

[Marković2015] D. Marković, F. Antonacci, A. Sarti, S. Tubaro, "Resolution issues in Soundfield Imaging: a multiresolution approach to sound source localization, in proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), 2015, NY, Oct. 2015, pages 1-5

[Marković2015a] D. Markovic, F. Antonacci, A. Sarti, S. Tubaro, "Multiview Soundfield Imaging in the Projective Ray Space," IEEE/ACM Transactions on Audio, Speech, and Language Processing , vol. 23, no.6, pp.1054,1067, June 2015 doi: 10.1109/TASLP.2015.2419076

[Bianchi2016] L. Bianchi; F. Antonacci; A. Sarti; S. Tubaro, "The Ray Space Transform: a New Framework for Wave Field Processing," in IEEE Transactions on Signal Processing , doi: 10.1109/TSP.2016.2591500

POLITECNICO DI MILANO

ISPL
Image and Sound Processing Lab

SMC Lab

Rethinking Fourier Acoustics

Part 3: Expanding the toolbox

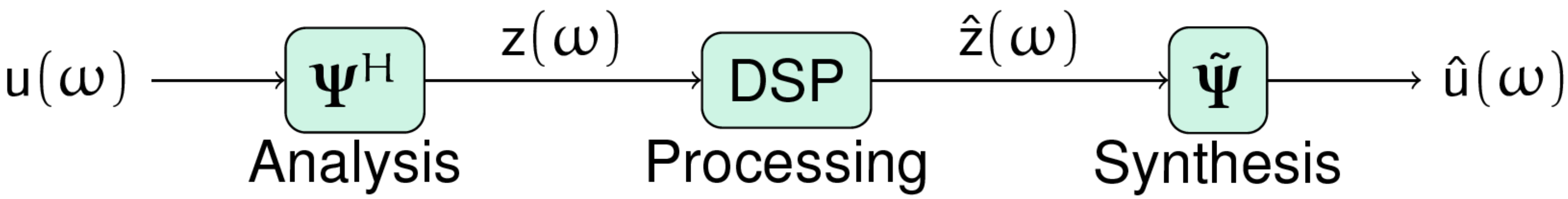
Augusto Sarti, Fabio Antonacci, Lucio Bianchi

Outline

- From Fourier decomposition to Gabor frames
- Frame-based analysis of acoustic fields
  - The Ray Space Transform (RST) and its inverse
- The RST as a wavefield decomposition into tapered beams
- Applications
  - RST-based nearfield plenacoustic cameras
  - IRST-based nearfield plenacoustic projectors
- Beyond acoustics: applications to EM signals
- Retrospective
- Perspectives

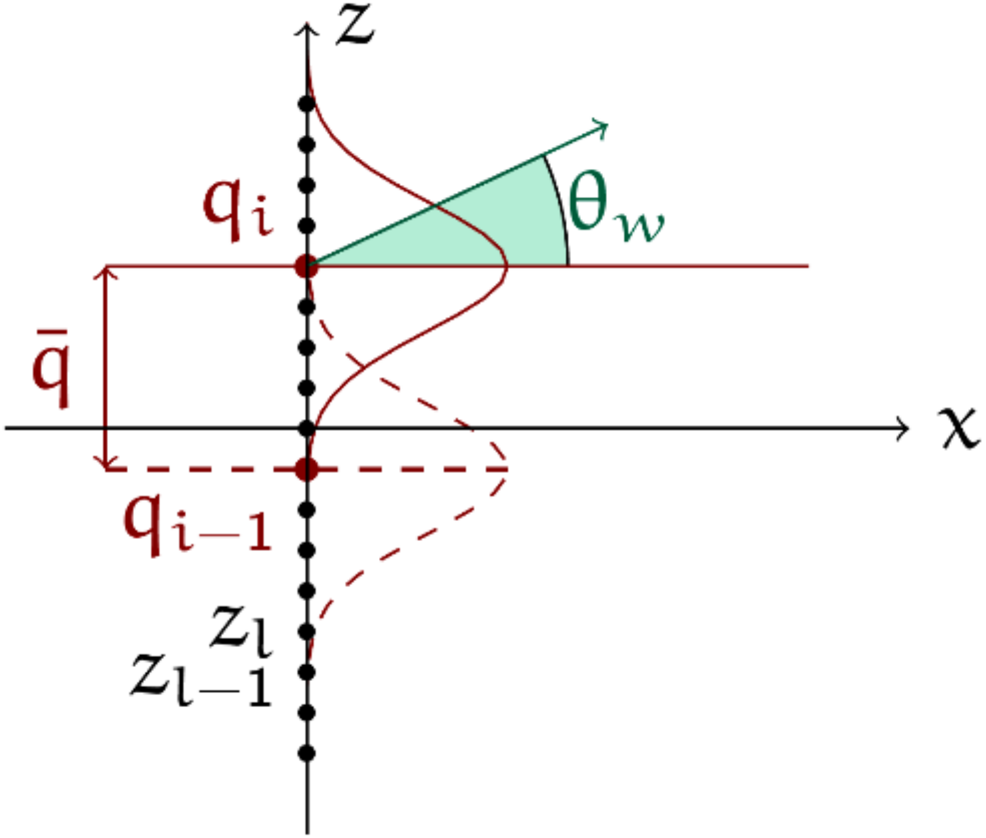
From Fourier decomposition to Gabor frames

- The inverse spatial Fourier transform describes a global decomposition of the acoustic field into plane waves
  - Spatial events are no longer discernible
  - In the case of Fourier representations of signals, we can overcome this problem by using a STFT
  - Similarly, we can window the acoustic field in the spatial domain before applying the Fourier transform, to obtain a more compact representation
- We need to define and organize translations of the spatial windows which allow us to retain the ability to discern events in space
- If we do things right, we should end up with a decomposition of the acoustic field in terms of **local directional wave objects**



Frame-based analysis of acoustic fields

- Sample the acoustic field with a linear array of microphones
  - Apply a series of translated spatial window to array data
  - Modulate each translated window to estimate directional contributions



The Ray Space Transform: preliminaries

- Adopt the ray space as the domain of the transformation
  - Parametrize directions θ by $m = \tan(\theta)$
- Phase shift at position z due to a directional contribution from θ

$$z \sin(\theta) = \frac{zm}{\sqrt{1+m^2}}$$

- Adopt uniform grid for sampling the (m, q) plane

$$q_i = (i - (I-1)/2)\,\bar{q}, \quad i = 0, \ldots, I-1$$

$$m_w = (w - (W-1)/2)\,\bar{m}, \quad w = 0, \ldots, W-1$$

The Ray Space transform of a continuous aperture

- Gabor transform of aperture data
  - Evaluate the similarity between the captured acoustic field and shifted and modulated copies of a prototype function

- Analysis equation

$$[\mathbf{Z}]_{i,w}(\omega) = \int_{-q_0}^{q_0} p(z,\omega) e^{-\frac{jkzm_w}{\sqrt{1+m_w^2}}} \psi^*_{i,w}(z)\, dz$$

$$\psi(z) = e^{-\pi\frac{z^2}{\sigma^2}}, \quad \sigma \in \mathbb{R}$$

- Synthesis equation

$$p^{(\mathbf{Z})}(z,\omega) = \sum_{i=0}^{I-1} \sum_{w=0}^{W-1} [\mathbf{Z}]_{i,w}(\omega) e^{\frac{jkzm_w}{\sqrt{1+m_w^2}}} \tilde{\psi}_{i,w}(z)$$

- Analysis equation

$$[\mathbf{Z}]_{i,w}(\omega) = d\sum_{l=0}^{L-1} p(z_l,\omega)e^{-\frac{jkz_l m_w}{\sqrt{1+m_w^2}}}e^{-\frac{\pi(z_l-\bar{q}i)^2}{\sigma^2}}$$

- The discrete Ray Space Transform (RST) can be conveniently written in matrix form upon introducing the discrete Gabor frame operators

$$[\mathbf{\Psi}]_{l,i+wI+1} = e^{j\frac{kz_l m_w}{\sqrt{1+m_w^2}}}e^{-\frac{\pi(z_l-q_i)^2}{\sigma^2}} \qquad \tilde{\mathbf{\Psi}} = \left(\mathbf{\Psi}\mathbf{\Psi}^H\right)^{-1}\mathbf{\Psi}$$

- Analysis equation in matrix form

$$\mathbf{z} = \mathbf{\Psi}^H\mathbf{p} \qquad [\mathbf{z}]_{i+wI+1} = [\mathbf{Z}]_{i,w}$$

- Synthesis equation in matrix form

$$\mathbf{p}^{(\mathbf{Z})} = \tilde{\mathbf{\Psi}}^H\mathbf{z}$$

RST Interpretation

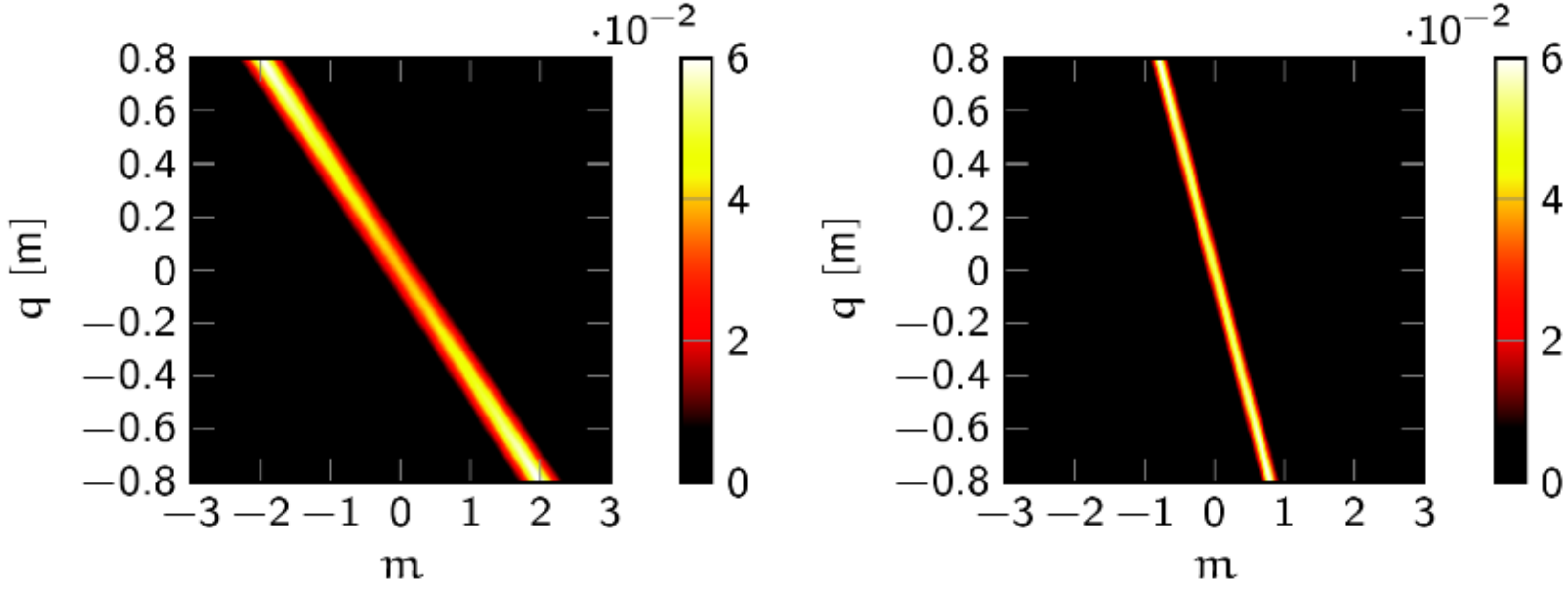
- Consider the *i*-th spatial window
- The RST can be interpreted as the beamforming operation applied to array data
  - Before beamforming, array data are weighed by a Gaussian spatial window function centered at q_i
- The *i*-th row of the matrix $\mathbf{Z}$ collects the outputs of multiple beamforming operations

Example: RST of a spherical wave

- Acoustic field generated by a point source at $\mathbf{r}' = [x', z']^T$, observed on the z axis

$$p(z,\omega) = \frac{\exp(-jk\sqrt{x'^2 + (z'-z)^2})}{4\pi\sqrt{x'^2 + (z'-z)^2}}$$

- Magnitude of ray space coefficients



(a) $x' = 1\,\mathrm{m}$, $z' = 0\,\mathrm{m}$. (b) $x' = 2.5\,\mathrm{m}$, $z' = 0\,\mathrm{m}$.

RST as a wavefield decomposition into tapered beams

- Acoustic pressure field generated by a continuous distribution of point sources on the z axis
  - Solution to

$$\left(\nabla^2 + k^2\right) p(\mathbf{r};k) = -\delta(x)\delta(y)u(z;k)$$

    - The function $u(z;k)$ is the source strength

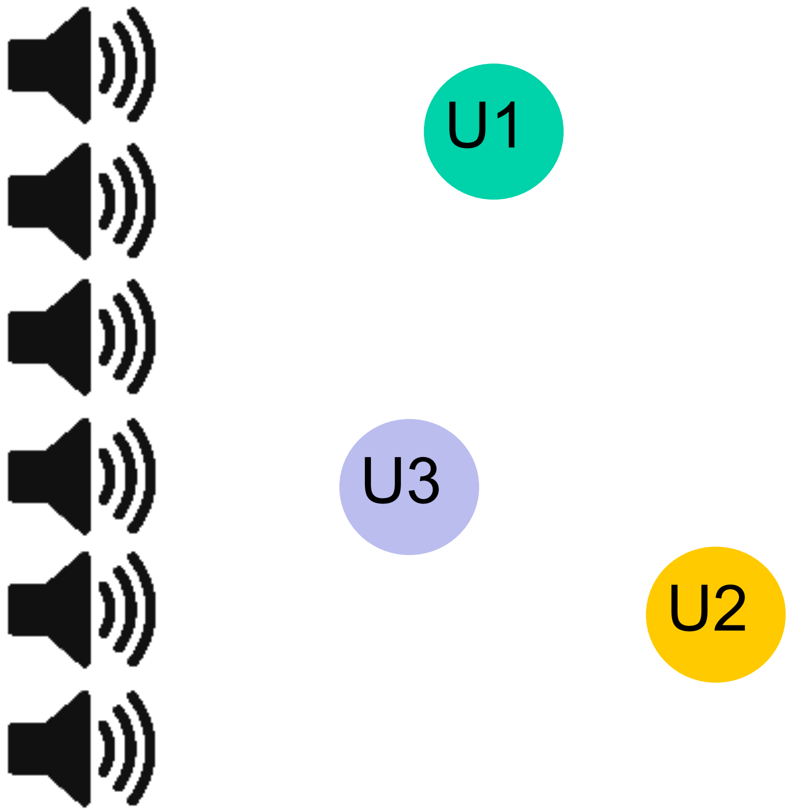
  - The solution can be expressed as the Raileigh first integral equation

$$p(\mathbf{r};k) = -j\rho_0 ck \int_{-\infty}^{\infty} u(z';k)g(\mathbf{r}|z';k)\,dz'$$

    - The function $g(\mathbf{r}|z';k)$ is the propagation function from a point source in z' to the field point r

Application example: multiuser sound field rendering

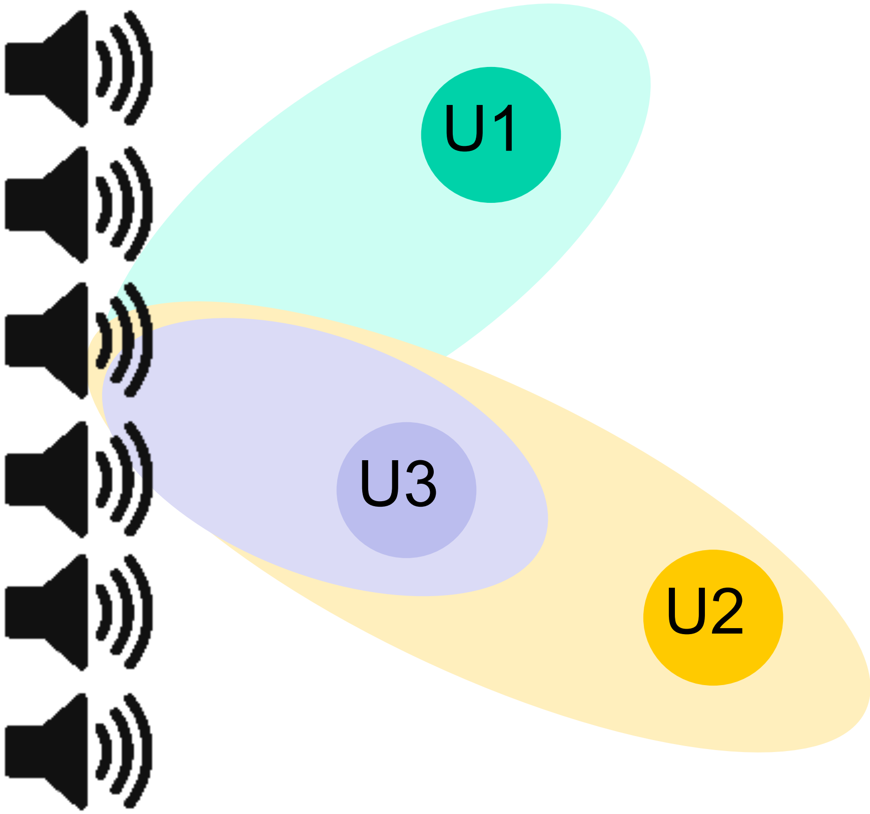
Provide multiple users with different audio contents using a single loudspeaker array



Application example: multiuser sound field rendering

Provide multiple users with different audio contents using a single loudspeaker array

- Conventional solutions: array considered as a whole

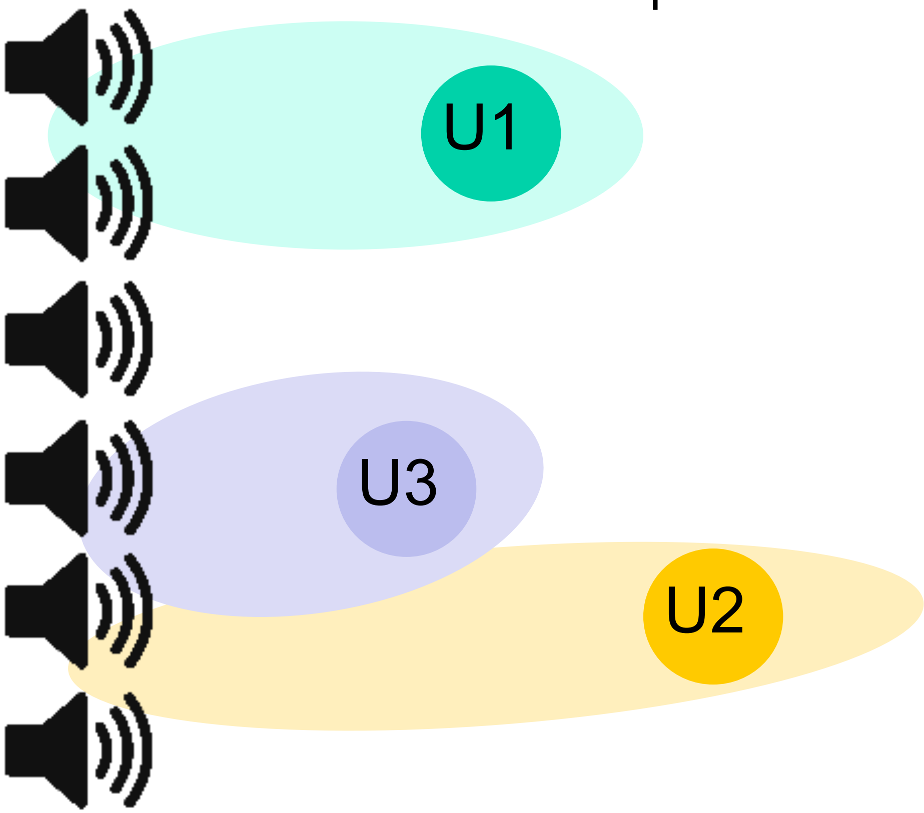


- User 3 is overwhelmed by the content intended to user 2
- Unable to manage users occlusions

Application example: multiuser sound field rendering

Provide multiple users with different audio contents using a single loudspeaker array

- Beam-based solution: freedom to chose beam amplitudes, directions and initiation points



- Every user is able to listen to its intended sound content
- Beam parameters are chosen to minimize the beam overlap at the locations of the users

Conclusions

- Ray-based representations are visually very powerful and effective for solving a wide range of problems in a very general fashion
- Rays, however, are generally defined and used under the conditions of Fourier Acoustics, which can be rather restrictive
- Can we enhance such representations and retain their visual power, while relaxing the operative conditions under which they function?
- In order to do so we need to replace the traditional Fourier decomposition with a different one of local validity, which can be thought of as a Short Space-Time Fourier Transform (SSTFT). This can be done using the theory of Gabor Frames
- We showed how to define a local signal decomposition for acoustic sound fields, we replaced the Fourier transform with a new mapping called Ray Space Transform, which
  - Preserves the visual representation power of Sound Field Mapping (SFM)
  - Invertible (can be used for analysis as well as synthesis)
  - Relaxes the operative conditions
    - Inherently nearfield operation
    - Extended frequency bandwidth
    - Ability to discern spatial events
    - …

What's next

- Ray space transform in the 3D domain (radiance complex)
- New Gabor frame bases (and related transforms)
- New analysis methodologies operating in the ray space
  - Egomotion estimation, self-calibration
  - Multi-camera fusion
  - Nearfield holography
  - …
- Framework development, complete with
  - Transform blocks
  - Pattern analysis blocks
  - Calibration/egomotion
  - Transcoders from and to other representations
    - SH, WFS, ambisonics, binaural, ...
- New applications that take full advantage of the ray space parameterization
  - Object-based acoustics
  - Augmented and mixed realities

References

[Bianchi2016] L. Bianchi; F. Antonacci; A. Sarti; S. Tubaro, "The Ray Space Transform: a New Framework for Wave Field Processing," in IEEE Transactions on Signal Processing , doi: 10.1109/TSP. 2016.2591500

Contributors

- Prof. Augusto Sarti
- Prof. Fabio Antonacci
- Dr. Lucio Bianchi
- Dr. Dejan Marković
- Dr. Antonio Canclini

Sound and Music Computing Lab

Image and Sound Processing Group

DEIB - *Dipartimento di Elettronica, Informazione e Bioingegneria*

Politecnico di Milano, Italy