Universität Rostock Traditio et Innovatio

Modal Bandwidth Reduction in Data-based Binaural Synthesis including Translatory Head-movements

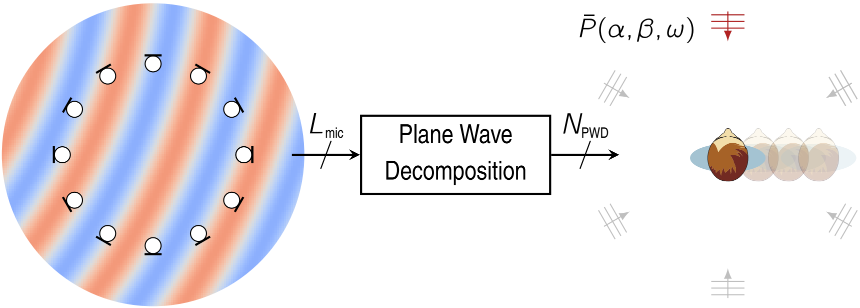
Nara Hahn and Sascha Spors

University of Rostock,Institute of Communications Engineering

41st DAGA

Nürnberg, 18. March 2015

Data-based Binaural Synthesis including Translatory Head-movements [Schultz and Spors, 2013]



$$P_{\text{L,R}}(\omega) = \int_0^{2\pi} \int_0^{\pi} \overbrace{\bar{P}(\mathbf{0}, \alpha, \beta, \omega)}^{\text{PWD coefficient}} \overbrace{H_{\text{L,R}}(\alpha, \beta, \omega)}^{\text{far-field HRTFs}} \sin\beta \, \mathrm{d}\beta \, \mathrm{d}\alpha$$

$$\bar{P}(\mathbf{x}_{\text{T}}, \alpha, \beta, \omega) = \bar{P}(\mathbf{0}, \alpha, \beta, \omega) \times \underbrace{\exp\left(-i\langle \mathbf{k}_{\text{PW}}, \mathbf{x}_{\text{T}} \rangle\right)}_{\text{PW extrapolation}}$$

Influence of Phase-shifted $\bar{P}(\alpha, \omega)$

Motivation

- localization properties studied in [Winter et al., 2014]
- physical properties not fully investigated

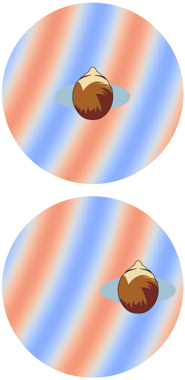
$\Rightarrow$ aim of this paper: examine the modal spectrum of a spatially translated sound field

Experimental conditions

- 2-dimensional scenario (xy-plane)
- continuous microphone array (no spatial sampling)
- parameters: modal bandwidth M and head-movement $\mathbf{x}_{\mathrm{T}} = (r_{\mathrm{T}}, \phi_{\mathrm{T}})$
- plane wave decomposition in the circular harmonics domain:

$$\bar{P}(\alpha, \omega) \underset{\mathcal{FS}_\alpha}{\overset{\mathcal{FS}_\alpha^{-1}}{\rightleftharpoons}} i^m \breve{S}_m(\omega), \quad \breve{S}_m(\omega) = \frac{\mathring{S}_m(r, \omega)}{J_m(\frac{\omega}{c} r)}$$

Spatial Translation



$$i^m \breve{S}_m(\omega) \xrightarrow{\mathcal{FS}_\alpha} \bar{P}(\alpha,\omega)$$

$$\Big\downarrow \times e^{-i\frac{\omega}{c} r_\mathrm{T} \cos(\alpha - \phi_\mathrm{T})}$$

$$i^m \breve{S}_{\mathrm{T},m}(\omega) \xleftarrow{\mathcal{FS}_\alpha^{-1}} \bar{P}_\mathrm{T}(\alpha,\omega)$$

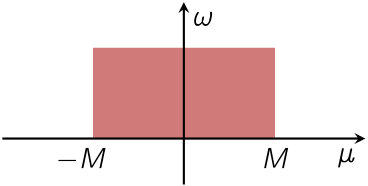
$$\breve{S}_{\mathrm{T},m}(\omega) = \sum_{\mu=-\infty}^{\infty} \breve{S}_\mu(\omega) J_{\mu-m}\left(\tfrac{\omega}{c} r_\mathrm{T}\right) e^{i(\mu-m)\phi_\mathrm{T}}$$

$\Rightarrow$ cross-correlation of $\breve{S}_m(\omega)$ and $J_m(\frac{\omega}{c} r_\mathrm{T}) e^{im\phi_\mathrm{T}}$ in the modal domain

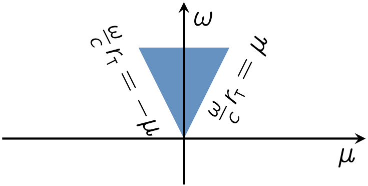
Translation of a Spatially Band-limited Sound Field

$$\breve{S}_{T,m}(\omega) \approx \sum_{\mu=-M}^{M} \breve{S}_\mu(\omega) J_{\mu-m}\left(\tfrac{\omega}{c} r_T\right) e^{i(\mu-m)\phi_T}$$

$\breve{S}_\mu(\omega)$



$J_\mu\left(\tfrac{\omega}{c} r_T\right) e^{i\mu\phi_T}$



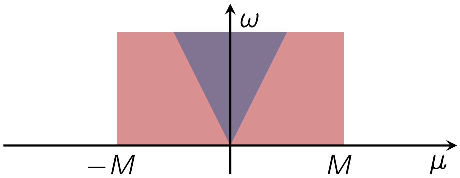
Translation of a Spatially Band-limited Sound Field

$(m = 0)$

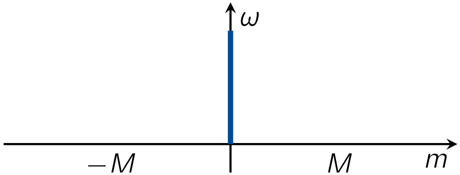
$$\breve{S}_\mu(\omega)$$

$$\times$$

$$J_{\mu-m}\left(\tfrac{\omega}{c} r_\mathrm{T}\right) e^{i(\mu-m)\phi_\mathrm{T}}$$



$$\breve{S}_{\mathrm{T},m}(\omega)$$



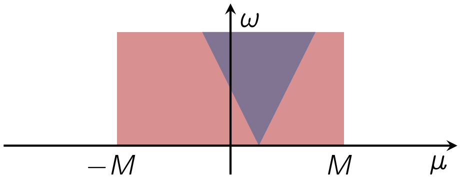
Translation of a Spatially Band-limited Sound Field

$(0 < m < M - \frac{\omega}{c} r_\mathsf{T})$

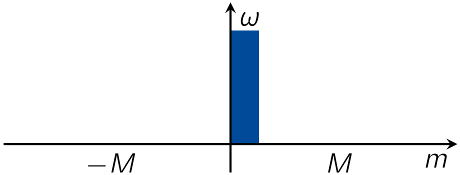
$\breve{S}_\mu(\omega)$

$\times$

$J_{\mu-m}(\frac{\omega}{c} r_\mathsf{T}) e^{i(\mu-m)\phi_\mathsf{T}}$



$\breve{S}_{\mathsf{T},m}(\omega)$



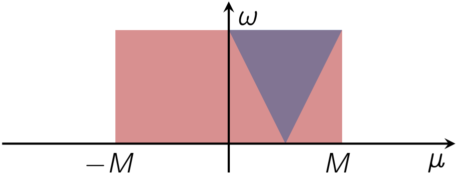
Translation of a Spatially Band-limited Sound Field

$(m = M - \frac{\omega}{c} r_\mathrm{T})$

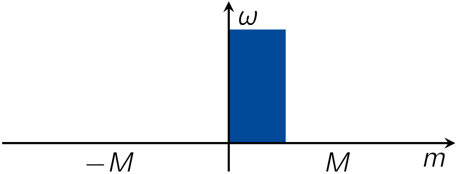
$$\breve{S}_\mu(\omega)$$

$$\times$$

$$J_{\mu-m}(\tfrac{\omega}{c} r_\mathrm{T}) e^{i(\mu-m)\phi_\mathrm{T}}$$



$$\breve{S}_{\mathrm{T},m}(\omega)$$



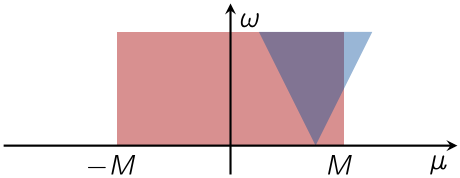
Translation of a Spatially Band-limited Sound Field

$(M - \frac{\omega}{c} r_\mathrm{T} < m < M)$

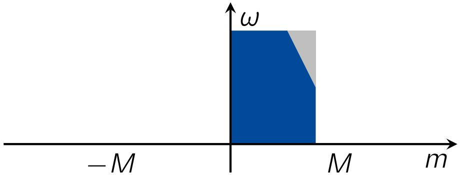
$$\breve{S}_\mu(\omega)$$

$$\times$$

$$J_{\mu-m}\left(\frac{\omega}{c} r_\mathrm{T}\right) e^{i(\mu-m)\phi_\mathrm{T}}$$



$$\breve{S}_{\mathrm{T},m}(\omega)$$



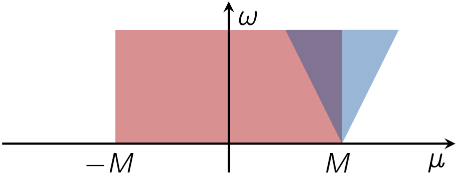
Translation of a Spatially Band-limited Sound Field

$(m = M)$

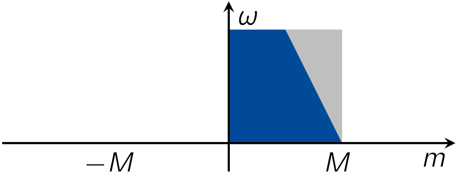
$$\breve{S}_\mu(\omega)$$

$$\times$$

$$J_{\mu-m}\left(\frac{\omega}{c} r_{\mathrm{T}}\right) e^{i(\mu-m)\phi_{\mathrm{T}}}$$



$$\breve{S}_{\mathrm{T},m}(\omega)$$



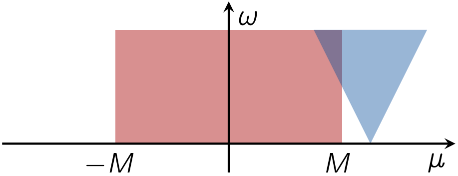
Translation of a Spatially Band-limited Sound Field

$(M < m < M + \frac{\omega}{c} r_\mathrm{T})$

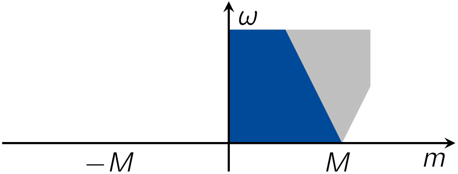
$$\breve{S}_\mu(\omega)$$

$$\times$$

$$J_{\mu-m}(\tfrac{\omega}{c} r_\mathrm{T}) e^{i(\mu-m)\phi_\mathrm{T}}$$



$$\breve{S}_{\mathrm{T},m}(\omega)$$



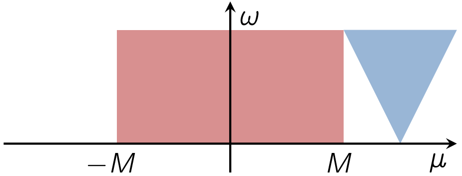
Translation of a Spatially Band-limited Sound Field

$(m = M + \frac{\omega}{c} r_\mathrm{T})$

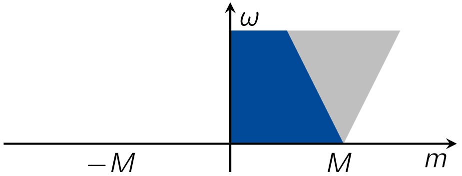
$$\breve{S}_\mu(\omega)$$

$$\times$$

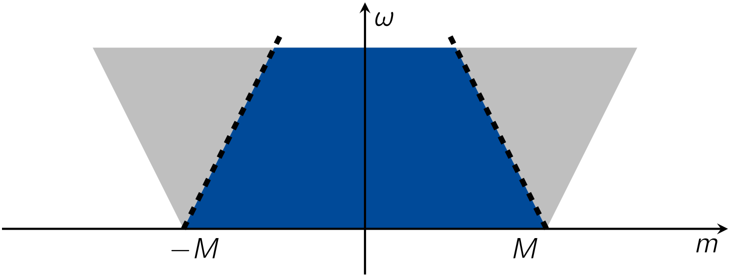
$$J_{\mu-m}(\tfrac{\omega}{c} r_\mathrm{T}) e^{i(\mu-m)\phi_\mathrm{T}}$$



$$\breve{S}_{\mathrm{T},m}(\omega)$$



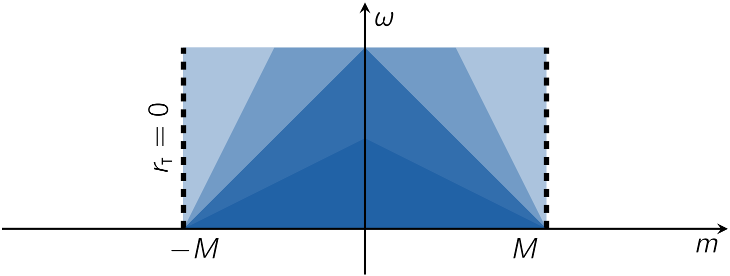
Modal Spectrum



- trapezoidal / triangular error-free region
- slope: $\pm \frac{\omega}{c} r_{\mathrm{T}}$

$\Rightarrow$ For a given ω_0, the modal spectrum at $\mathbf{x}_{\mathrm{T}}$ is accurate only up to $M_{\mathrm{T}} = M - \lceil \frac{\omega_0}{c} r_{\mathrm{T}} \rceil$.

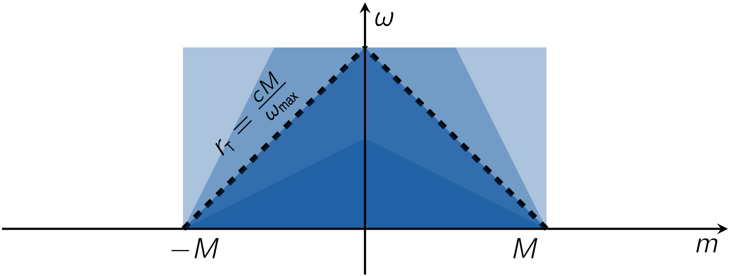
Modal Spectrum



- trapezoidal / triangular error-free region
- slope: $\pm \frac{\omega}{c} r_T$

$\Rightarrow$ For a given ω_0, the modal spectrum at $\mathbf{x}_T$ is accurate only up to $M_T = M - \lceil \frac{\omega_0}{c} r_T \rceil$.

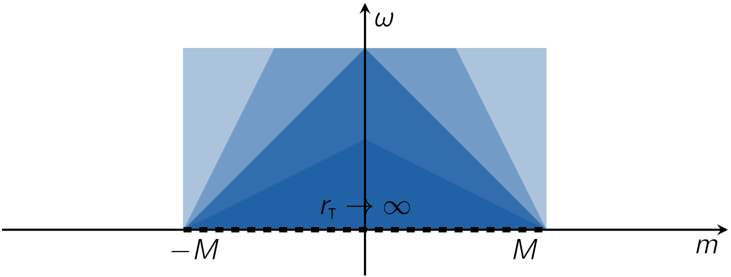
Modal Spectrum



- trapezoidal / triangular error-free region
- slope: $\pm \frac{\omega}{c} r_T$

$\Rightarrow$ For a given ω_0, the modal spectrum at $\mathbf{x}_T$ is accurate only up to $M_T = M - \lceil \frac{\omega_0}{c} r_T \rceil$.

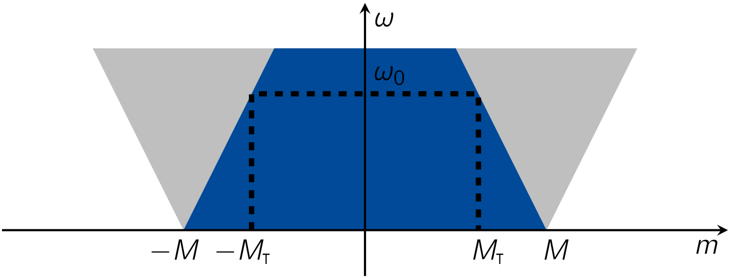
Modal Spectrum



- trapezoidal / triangular error-free region
- slope: $\pm \frac{\omega}{c} r_\mathrm{T}$

$\Rightarrow$ For a given ω_0, the modal spectrum at $\mathbf{x}_\mathrm{T}$ is accurate only up to $M_\mathrm{T} = M - \lceil \frac{\omega_0}{c} r_\mathrm{T} \rceil$.

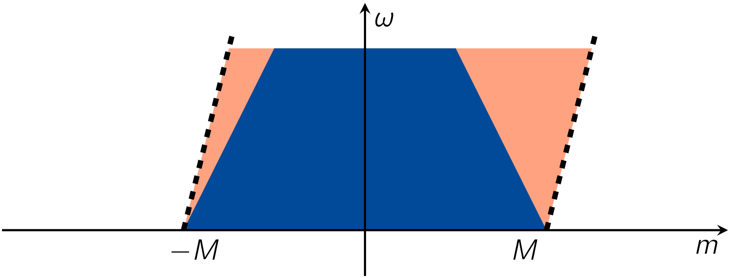
Modal Spectrum



- trapezoidal / triangular error-free region
- slope: $\pm \frac{\omega}{c} r_{\mathrm{T}}$

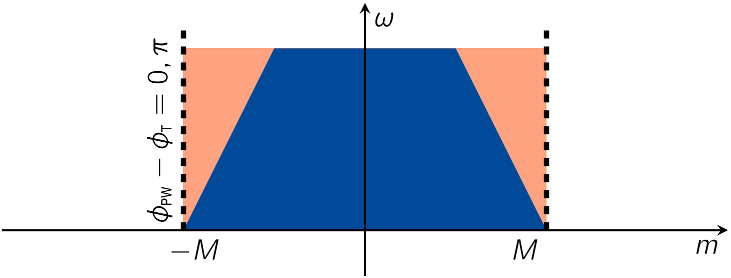
$\Rightarrow$ For a given ω_0, the modal spectrum at $\mathbf{x}_{\mathrm{T}}$ is accurate only up to $M_{\mathrm{T}} = M - \lceil \frac{\omega_0}{c} r_{\mathrm{T}} \rceil$.

Modal Spectrum



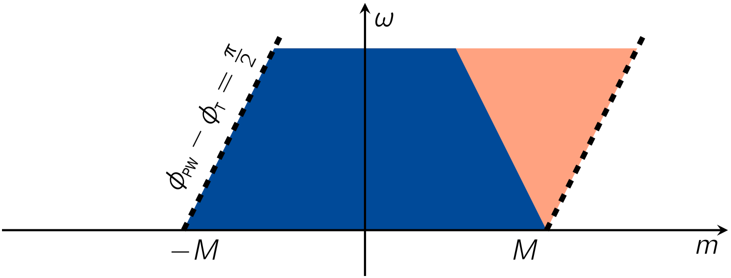
- parallelogram-shaped region with moderate error (≈ -20 dB)
- slope: $\frac{\omega}{c} r_{\mathrm{T}} \sin(\phi_{\mathrm{PW}} - \phi_{\mathrm{T}})$
- energy along the direction of plane wave propagation ($\rightarrow$)

Modal Spectrum



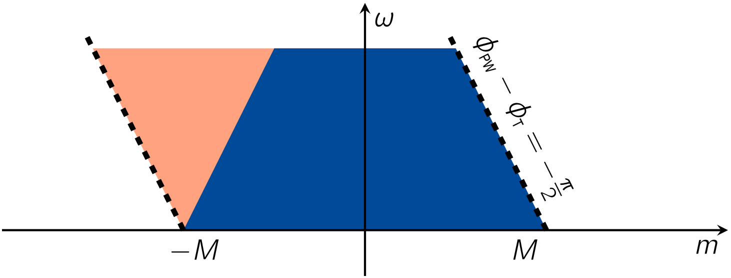
- parallelogram-shaped region with moderate error (≈ -20 dB)
- slope: $\frac{\omega}{c} r_T \sin(\phi_{PW} - \phi_T)$
- energy along the direction of plane wave propagation ($\rightarrow$)

Modal Spectrum

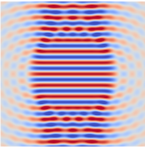


- parallelogram-shaped region with moderate error (≈ -20 dB)
- slope: $\frac{\omega}{c} r_T \sin(\phi_{PW} - \phi_T)$
- energy along the direction of plane wave propagation ($\longrightarrow$)

Modal Spectrum

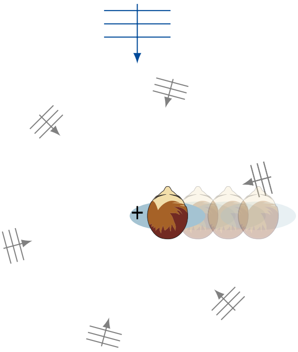


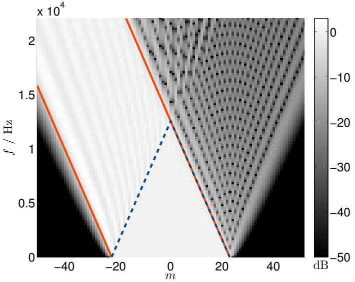
- parallelogram-shaped region with moderate error (≈ -20 dB)
- slope: $\frac{\omega}{c} r_{T} \sin(\phi_{PW} - \phi_{T})$
- energy along the direction of plane wave propagation ($\rightarrow$)

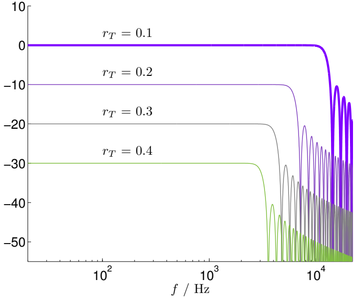


($f = 1$ kHz, $M = 23$)

Evaluation



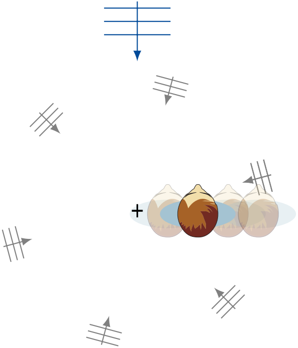


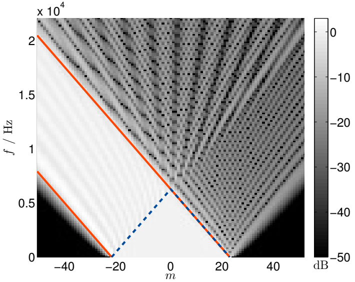


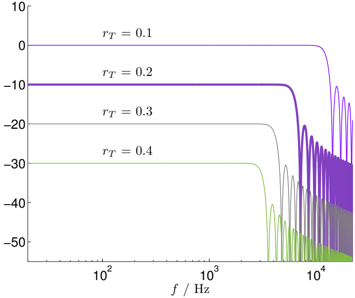
- sound field: $\phi_{\mathrm{PW}} = -\frac{\pi}{2}$, $M = 23$
- translation: $r_{\mathrm{T}} = \{0.1, 0.2, 0.3, 0.4\}$, $\phi_{\mathrm{T}} = 0$

→ bandwidth: $M_{\mathrm{T}}^{(1\ \mathrm{kHz})} = \{21, 19, 17, 15\}$

Evaluation



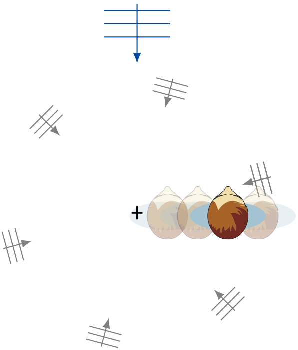


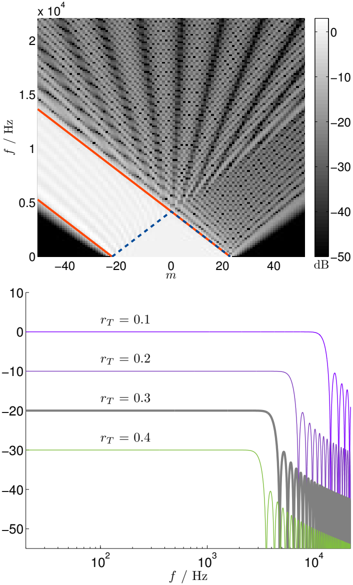


- sound field: $\phi_{\mathrm{PW}} = -\frac{\pi}{2}$, $M = 23$
- translation: $r_{\mathrm{T}} = \{0.1, 0.2, 0.3, 0.4\}$, $\phi_{\mathrm{T}} = 0$

→ bandwidth: $M_{\mathrm{T}}^{(1\,\mathrm{kHz})} = \{21, 19, 17, 15\}$

Evaluation

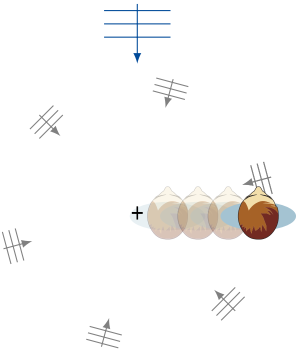


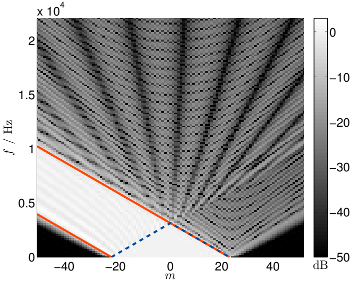


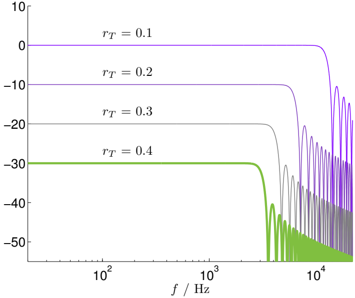
- sound field: $\phi_{\mathrm{PW}} = -\frac{\pi}{2}$, $M = 23$
- translation: $r_{\mathrm{T}} = \{0.1, 0.2, 0.3, 0.4\}$, $\phi_{\mathrm{T}} = 0$

→ bandwidth: $M_{\mathrm{T}}^{(1\,\mathrm{kHz})} = \{21, 19, 17, 15\}$

Evaluation



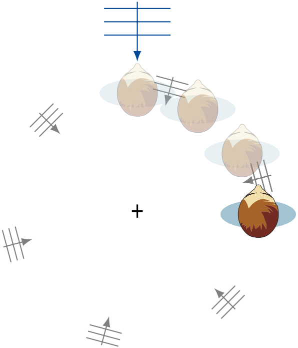


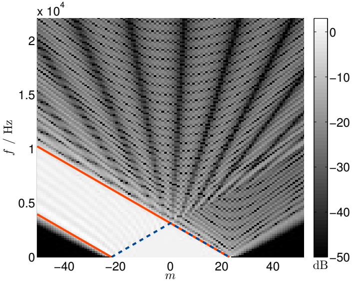


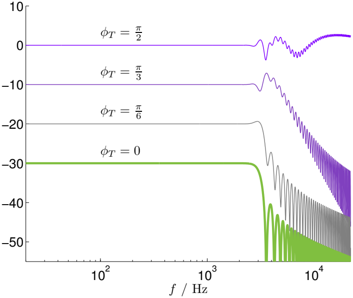
- sound field: $\phi_{\mathrm{PW}} = -\frac{\pi}{2}$, $M = 23$
- translation: $r_{\mathrm{T}} = \{0.1, 0.2, 0.3, 0.4\}$, $\phi_{\mathrm{T}} = 0$

→ bandwidth: $M_{\mathrm{T}}^{(1\text{ kHz})} = \{21, 19, 17, 15\}$

Evaluation



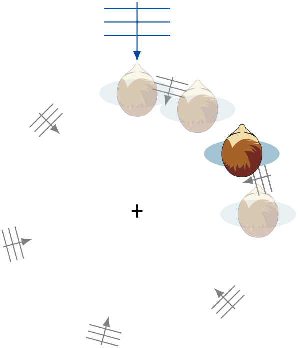


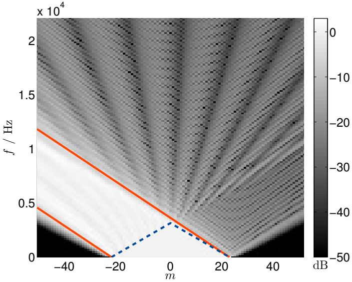


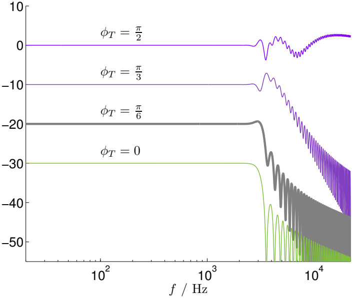
- sound field: $\phi_{\mathrm{PW}} = -\frac{\pi}{2}$, $M = 23$
- translation: $r_{\mathrm{T}} = 0.4$, $\phi_{\mathrm{T}} = \{0, \frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}\}$

→ bandwidth: $M_{\mathrm{T}}^{(1\,\mathrm{kHz})} = 15$

Evaluation



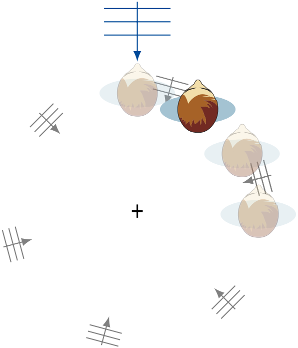


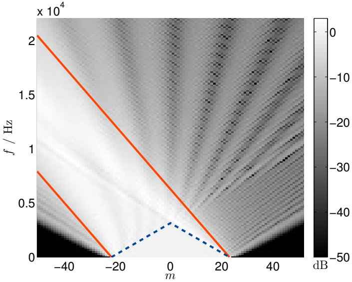


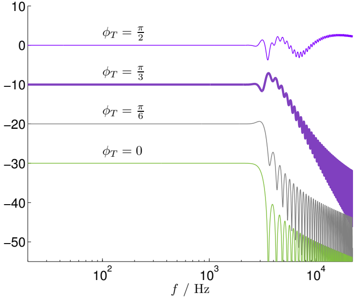
- sound field: $\phi_{\mathrm{PW}} = -\frac{\pi}{2}$, $M = 23$
- translation: $r_{\mathrm{T}} = 0.4$, $\phi_{\mathrm{T}} = \{0, \frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}\}$

→ bandwidth: $M_{\mathrm{T}}^{(1\ \mathrm{kHz})} = 15$

Evaluation



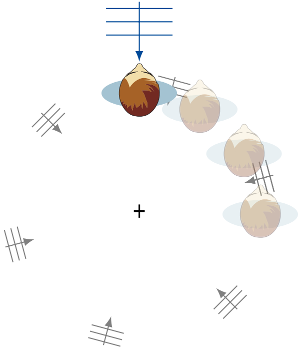


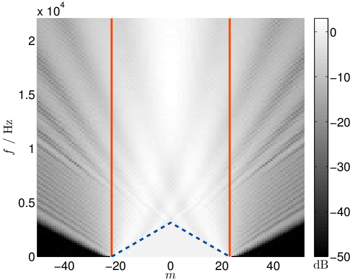


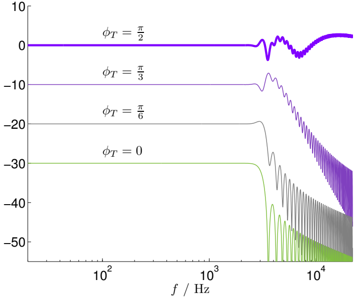
- sound field: $\phi_{PW} = -\frac{\pi}{2}$, $M = 23$
- translation: $r_T = 0.4$, $\phi_T = \{0, \frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}\}$

→ bandwidth: $M_T^{(1\text{ kHz})} = 15$

Evaluation



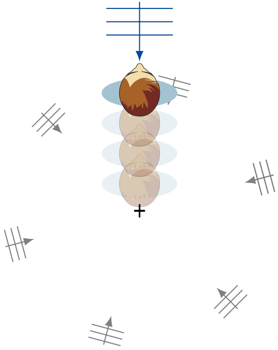


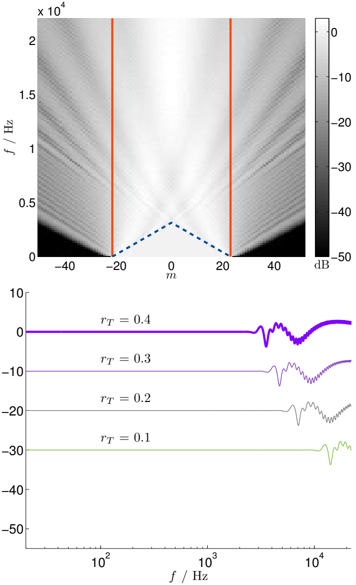


- sound field: $\phi_{\mathrm{PW}} = -\frac{\pi}{2}$, $M = 23$
- translation: $r_{\mathrm{T}} = 0.4$, $\phi_{\mathrm{T}} = \{0, \frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}\}$

→ bandwidth: $M_{\mathrm{T}}^{(1\ \mathrm{kHz})} = 15$

Evaluation

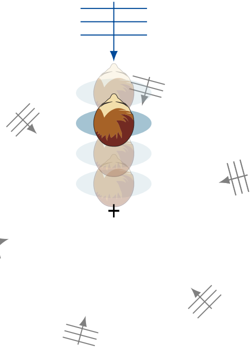


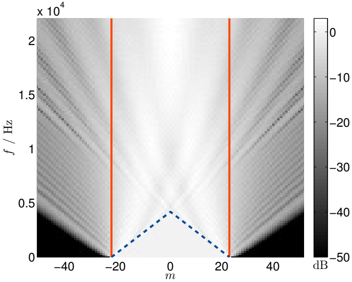


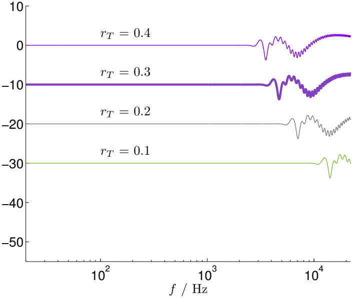
- sound field: $\phi_{\mathrm{PW}} = -\frac{\pi}{2}$, $M = 23$
- translation: $r_{\mathrm{T}} = \{0.1, 0.2, 0.3, 0.4\}$, $\phi_{\mathrm{T}} = \frac{\pi}{2}$

→ bandwidth: $M_{\mathrm{T}}^{(1\,\mathrm{kHz})} = \{21, 19, 17, 15\}$

Evaluation



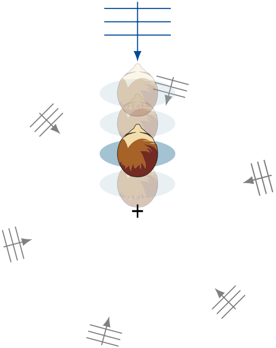


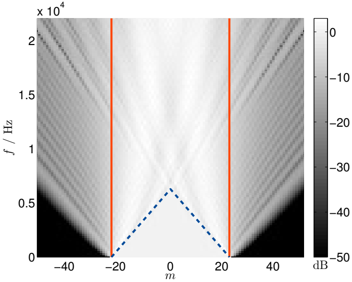


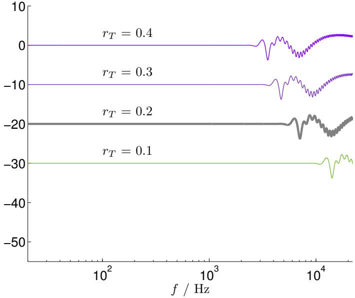
- sound field: $\phi_{\mathrm{PW}} = -\frac{\pi}{2}$, $M = 23$
- translation: $r_{\mathrm{T}} = \{0.1, 0.2, 0.3, 0.4\}$, $\phi_{\mathrm{T}} = \frac{\pi}{2}$

→ bandwidth: $M_{\mathrm{T}}^{(1\ \mathrm{kHz})} = \{21, 19, 17, 15\}$

Evaluation



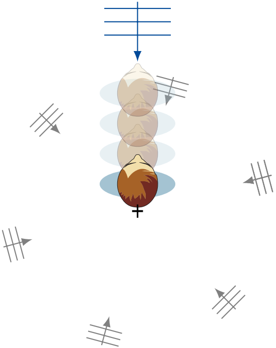


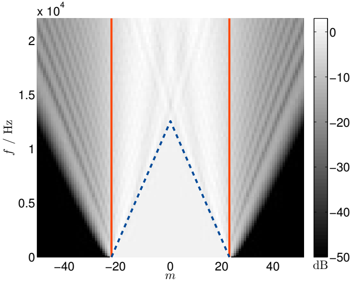


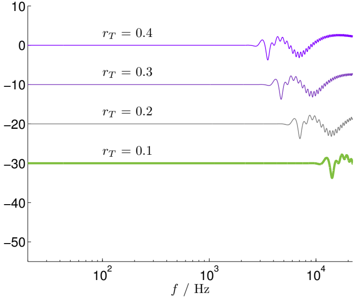
- sound field: $\phi_{\mathrm{PW}} = -\frac{\pi}{2}$, $M = 23$
- translation: $r_{\mathrm{T}} = \{0.1, 0.2, 0.3, 0.4\}$, $\phi_{\mathrm{T}} = \frac{\pi}{2}$

→ bandwidth: $M_{\mathrm{T}}^{(1\ \mathrm{kHz})} = \{21, 19, 17, 15\}$

Evaluation



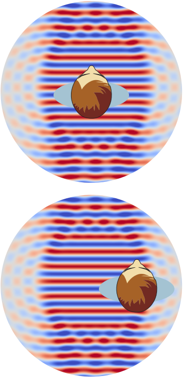


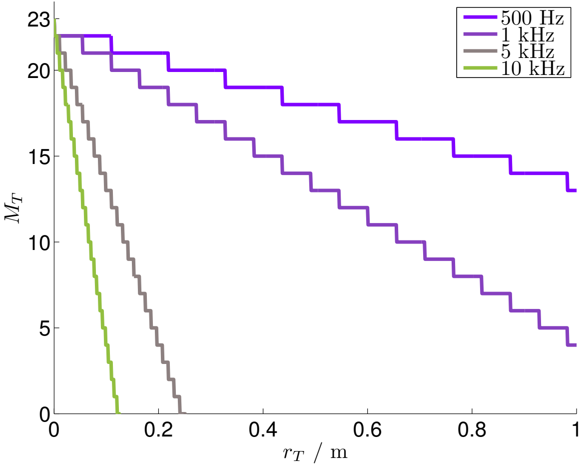


- sound field: $\phi_{\mathrm{PW}} = -\frac{\pi}{2}$, $M = 23$
- translation: $r_{\mathrm{T}} = \{0.1, 0.2, 0.3, 0.4\}$, $\phi_{\mathrm{T}} = \frac{\pi}{2}$

→ bandwidth: $M_{\mathrm{T}}^{(1\,\mathrm{kHz})} = \{21, 19, 17, 15\}$

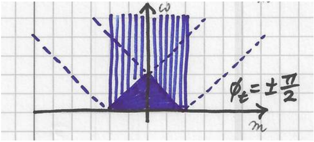
Conclusion



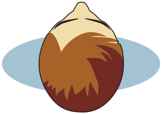


Outlook

- temporal and spectral properties
- perceptual properties: coloration and ~~localization~~
- influence of discrete plane wave decomposition



THANK YOU!



http://spatialaudio.net