Assessment of the Perception of Synthesized Sound Fields with a Binaural Model

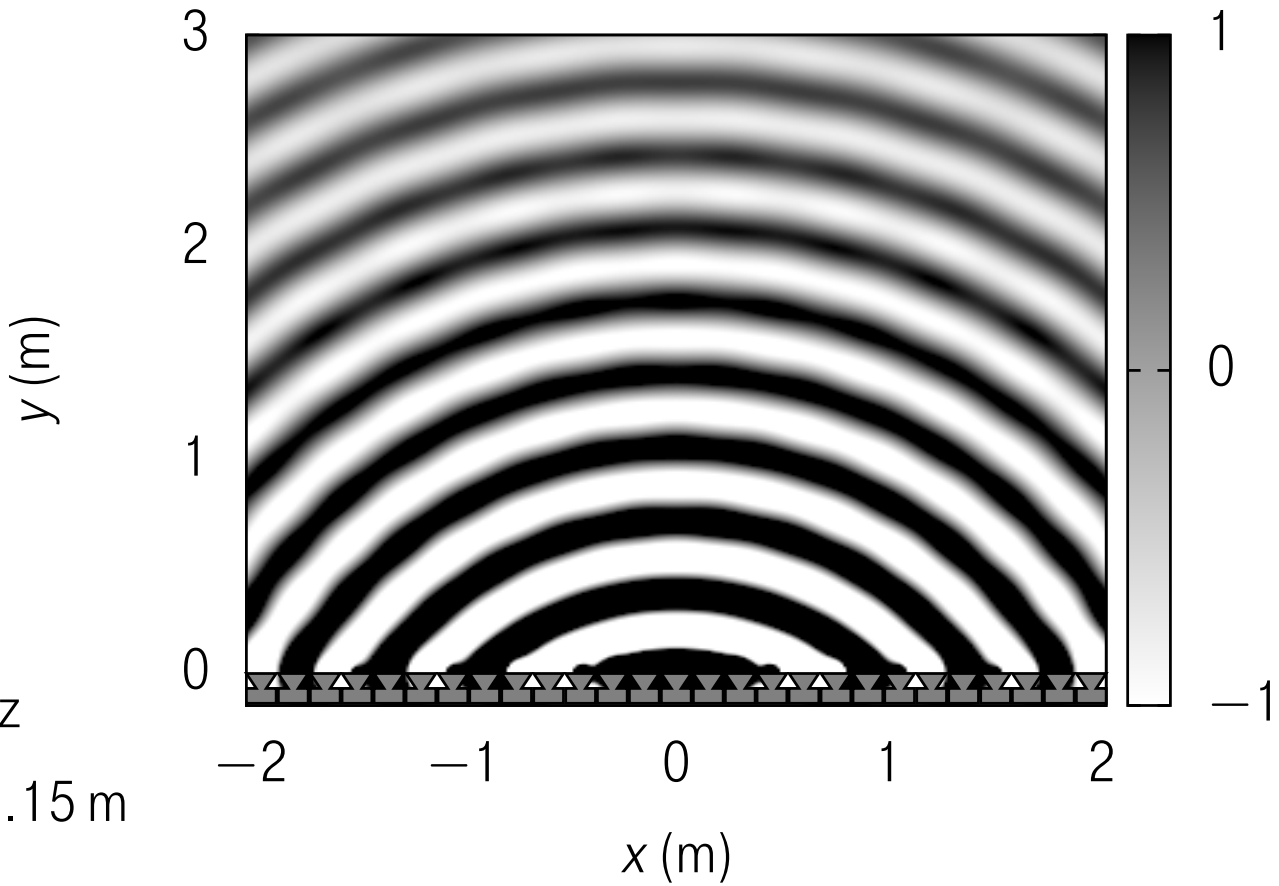
Hagen Wierstorf, Alexander Raake, Sascha Spors

Quality and Usability Lab
Deutsche Telekom Laboratories
Technische Universität Berlin

June 29, 2011

Deutsche Telekom Laboratories

Wave Field Synthesis



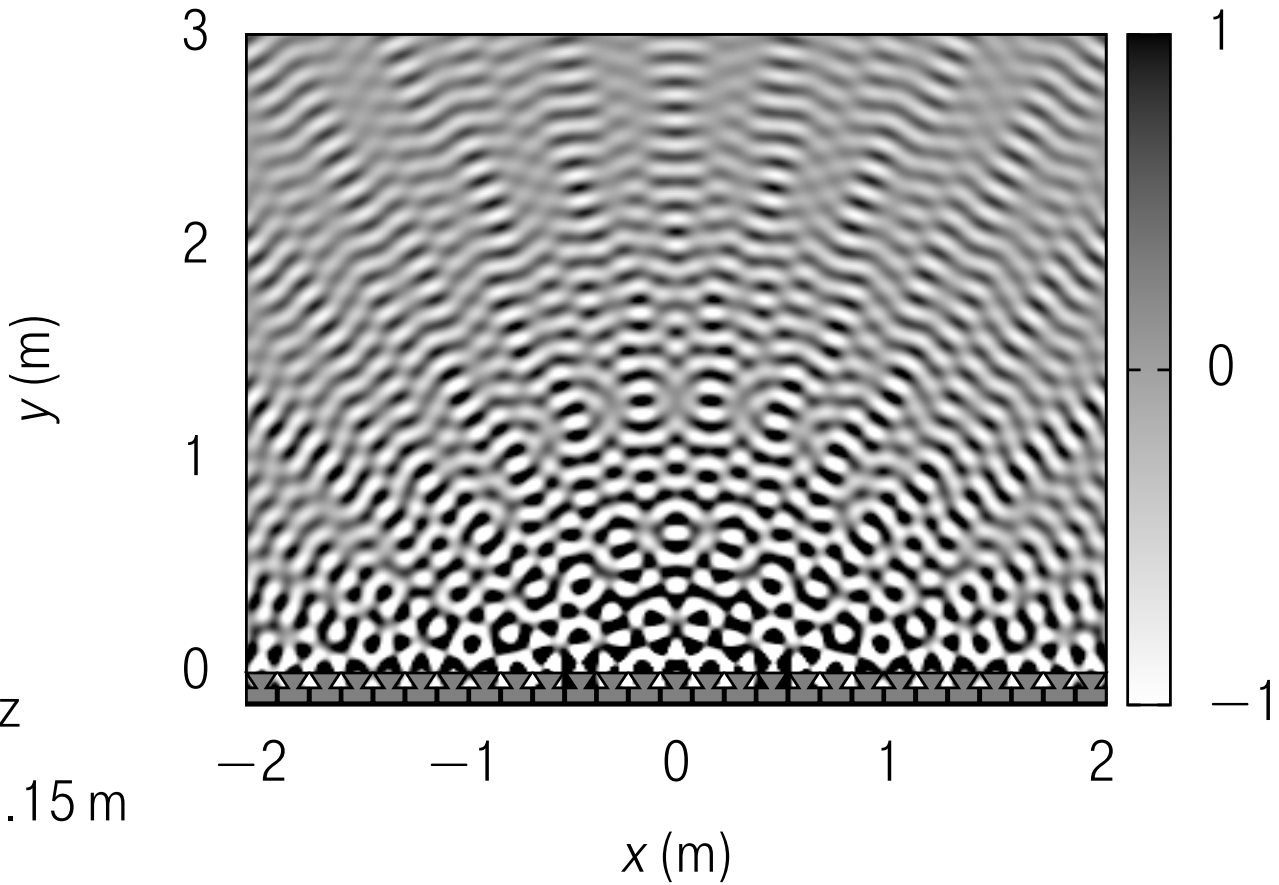
$f = 1\,\text{kHz}$

$\Delta x_0 = 0.15\,\text{m}$

- synthesis of arbitrary sound fields
- simple calculation of driving functions for source models of the sound field

Wave Field Synthesis

Spatial Sampling



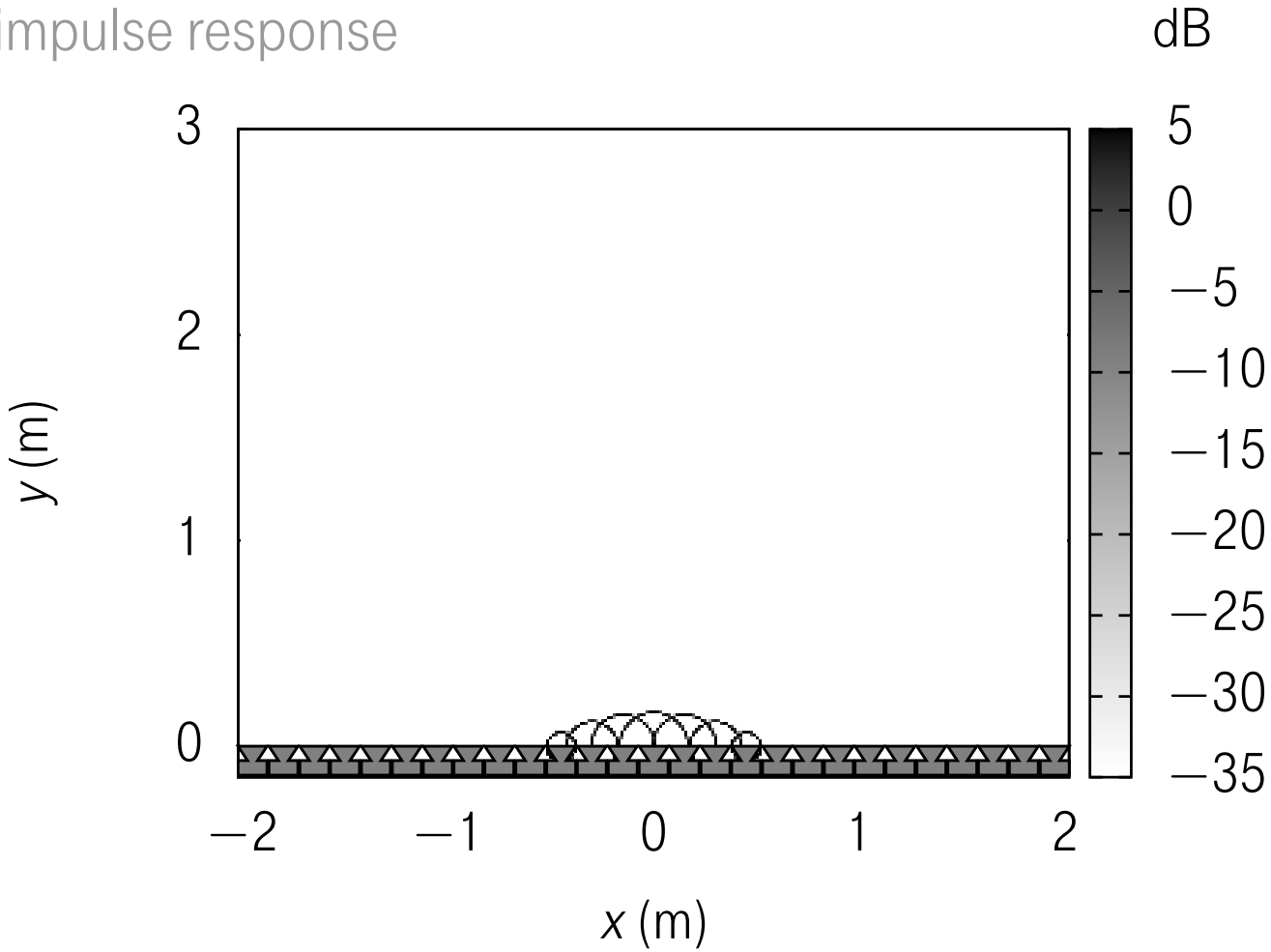
$f = 3\,\mathrm{kHz}$

$\Delta x_0 = 0.15\,\mathrm{m}$

- spatial sampling leads to aliasing above f_{al}

Wave Field Synthesis

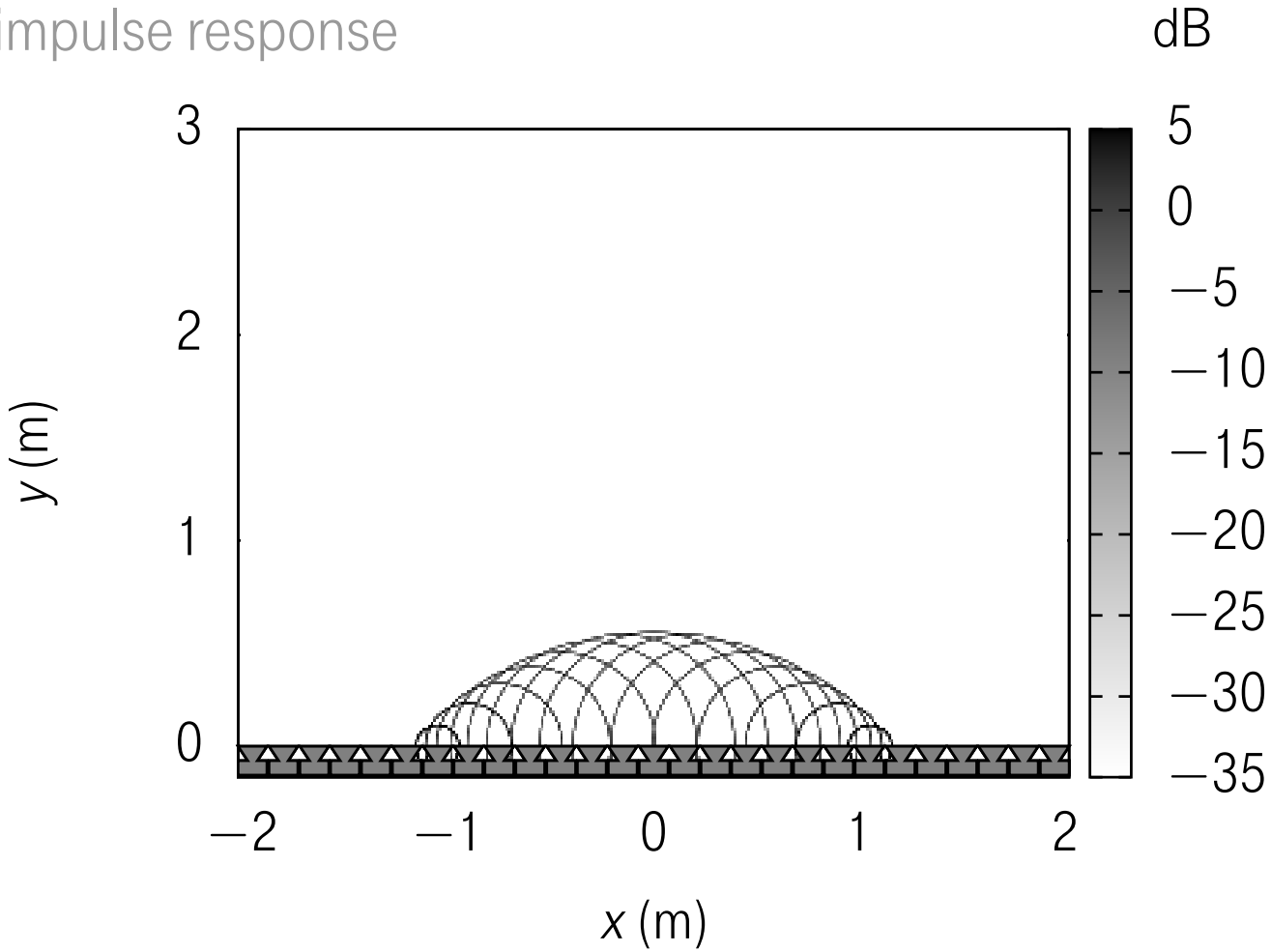
Snapshot of impulse response



- perception is dominated by the first wave front
- additional wave fronts add coloration (Wittek 2007) and an unnatural room impression

Wave Field Synthesis

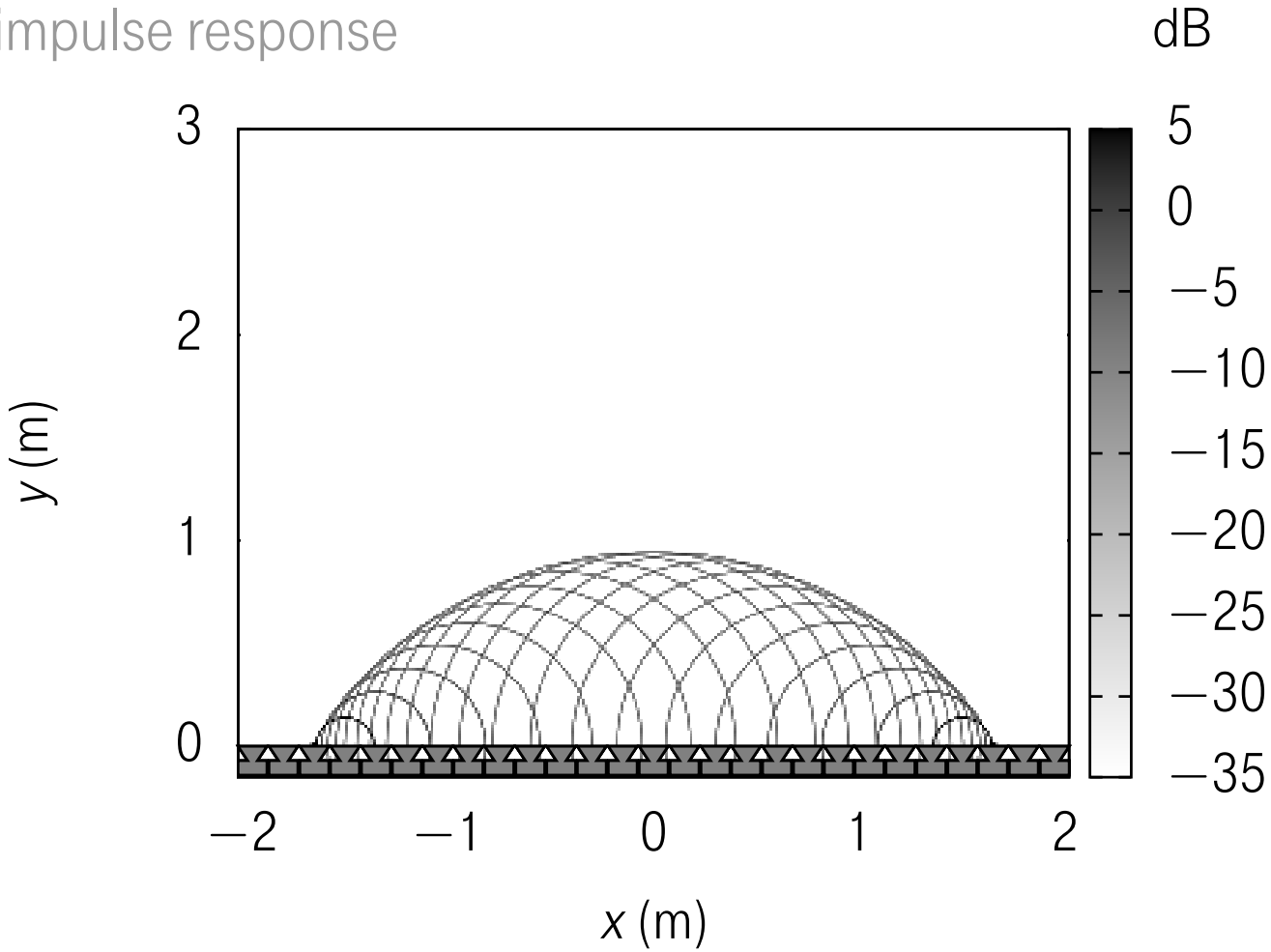
Snapshot of impulse response



- perception is dominated by the first wave front
- additional wave fronts add coloration (Wittek 2007) and an unnatural room impression

Wave Field Synthesis

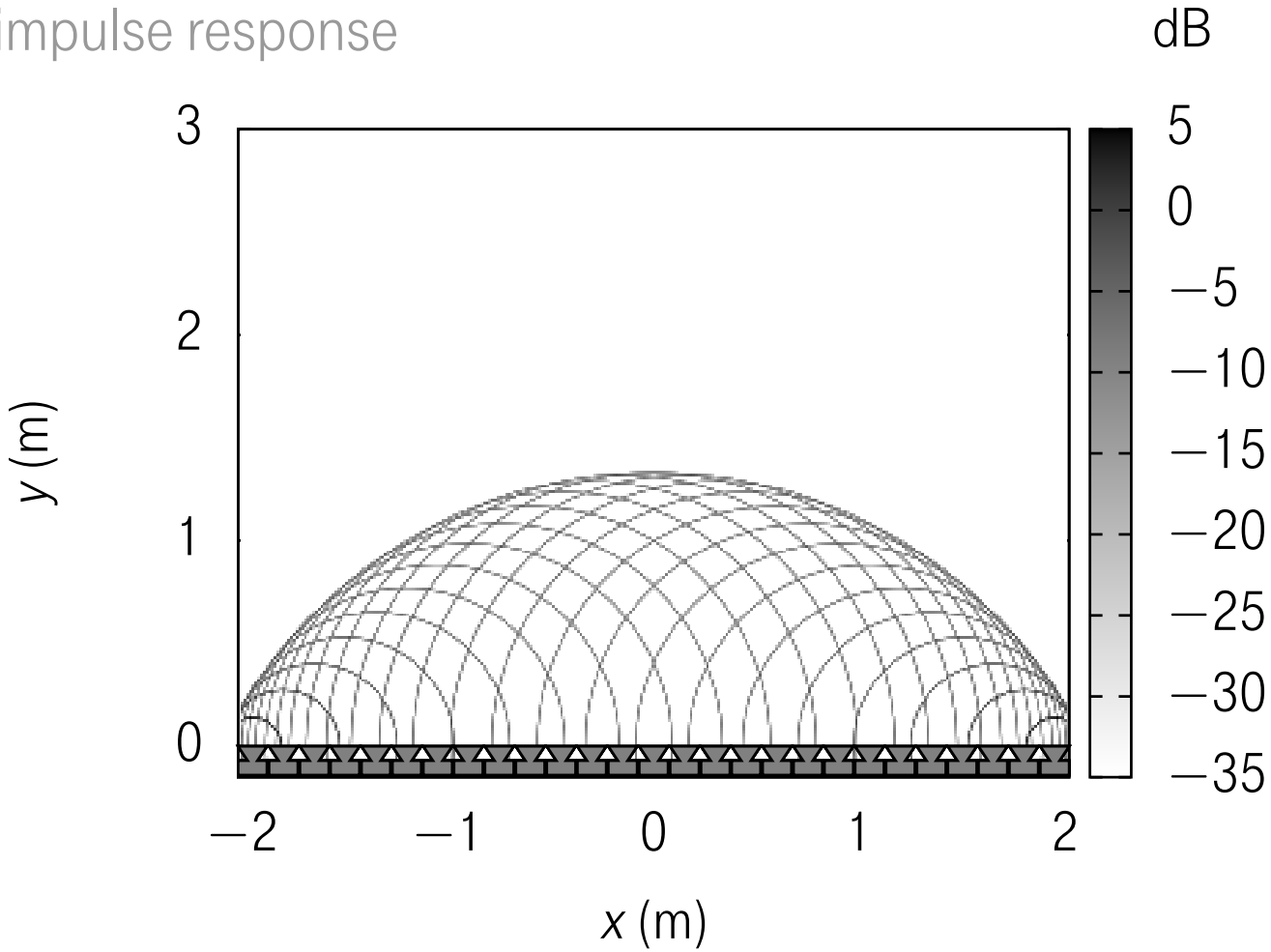
Snapshot of impulse response



- perception is dominated by the first wave front
- additional wave fronts add coloration (Wittek 2007) and an unnatural room impression

Wave Field Synthesis

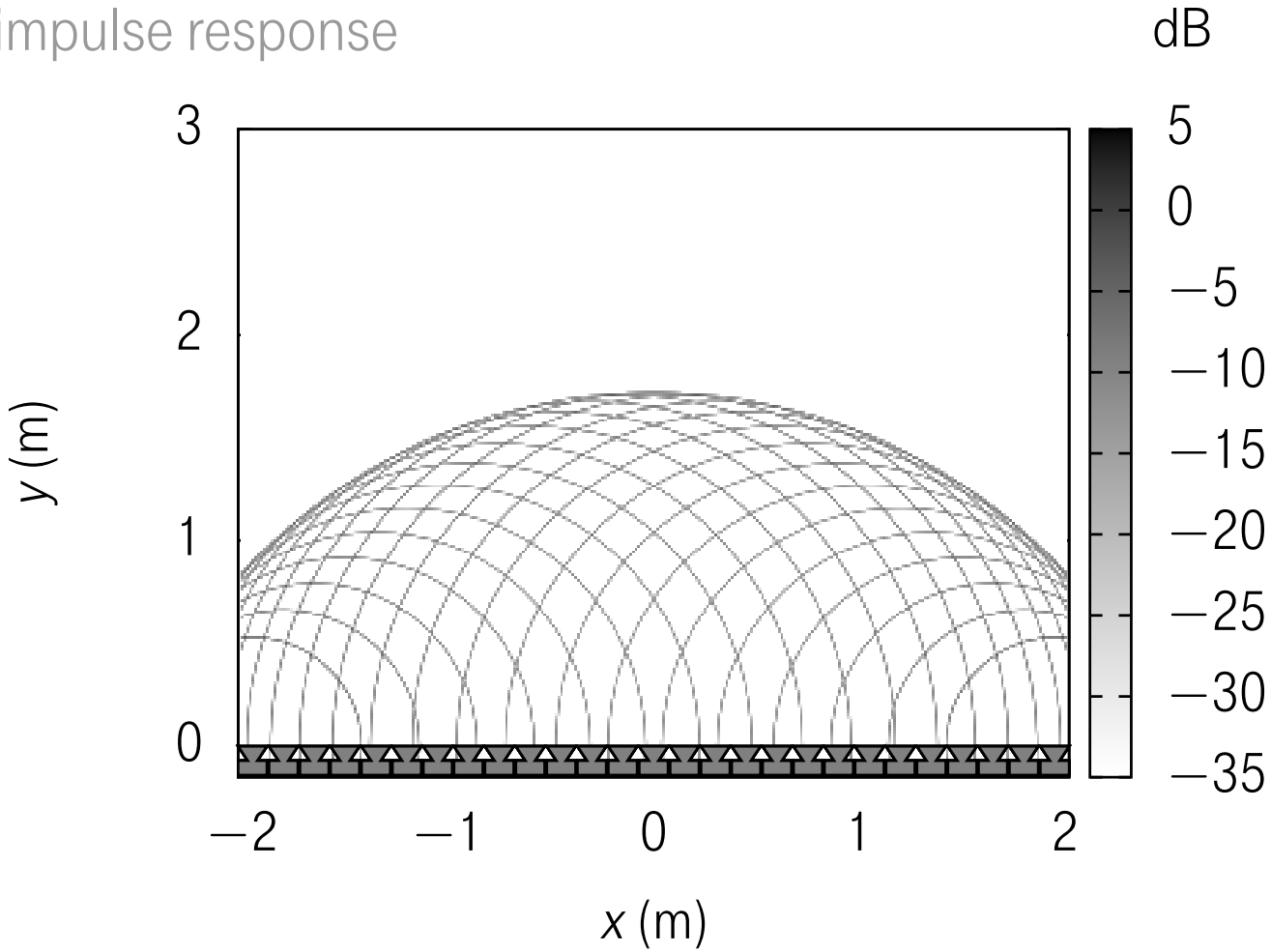
Snapshot of impulse response



- perception is dominated by the first wave front
- additional wave fronts add coloration (Wittek 2007) and an unnatural room impression

Wave Field Synthesis

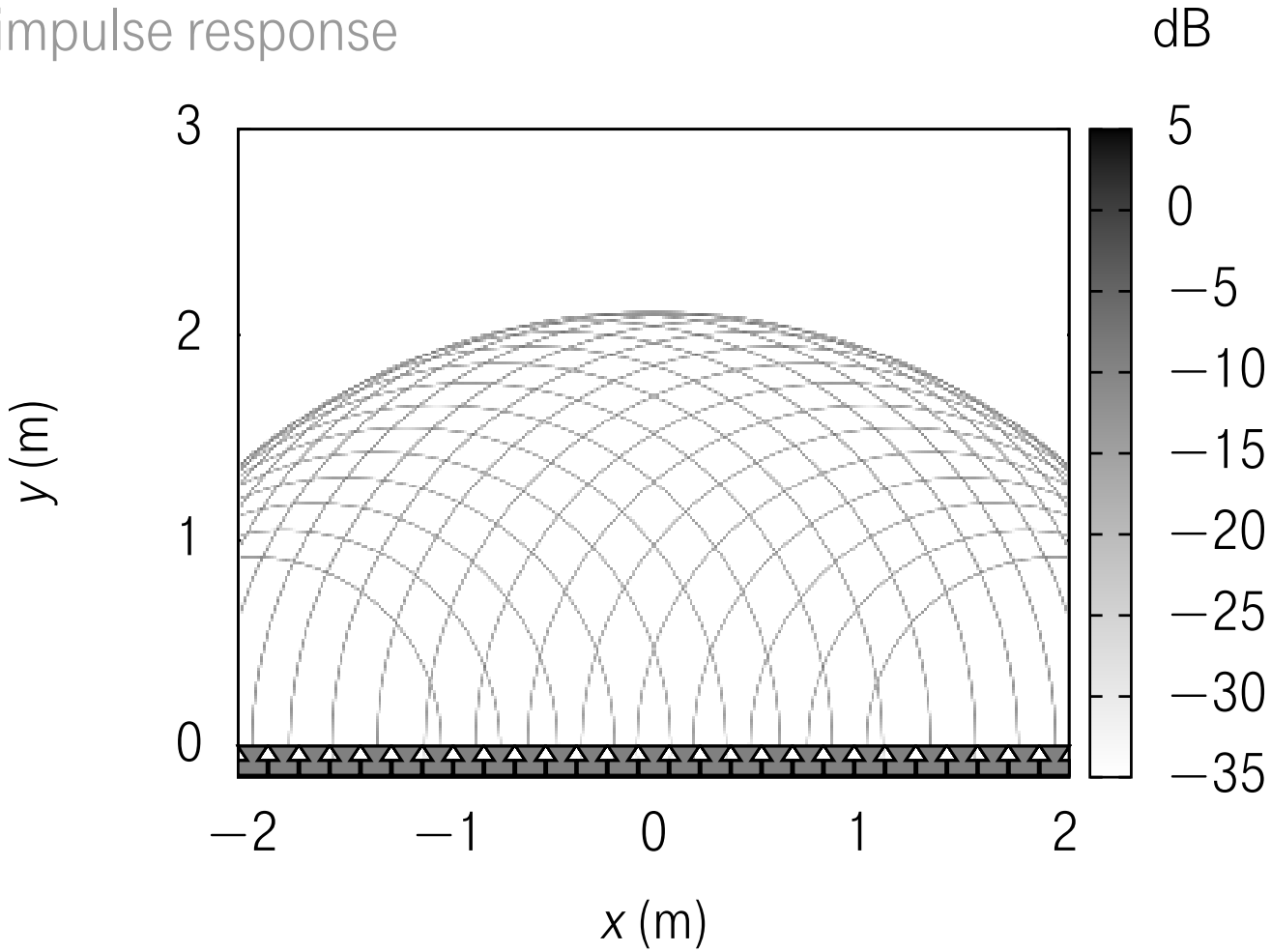
Snapshot of impulse response



- perception is dominated by the first wave front
- additional wave fronts add coloration (Wittek 2007) and an unnatural room impression

Wave Field Synthesis

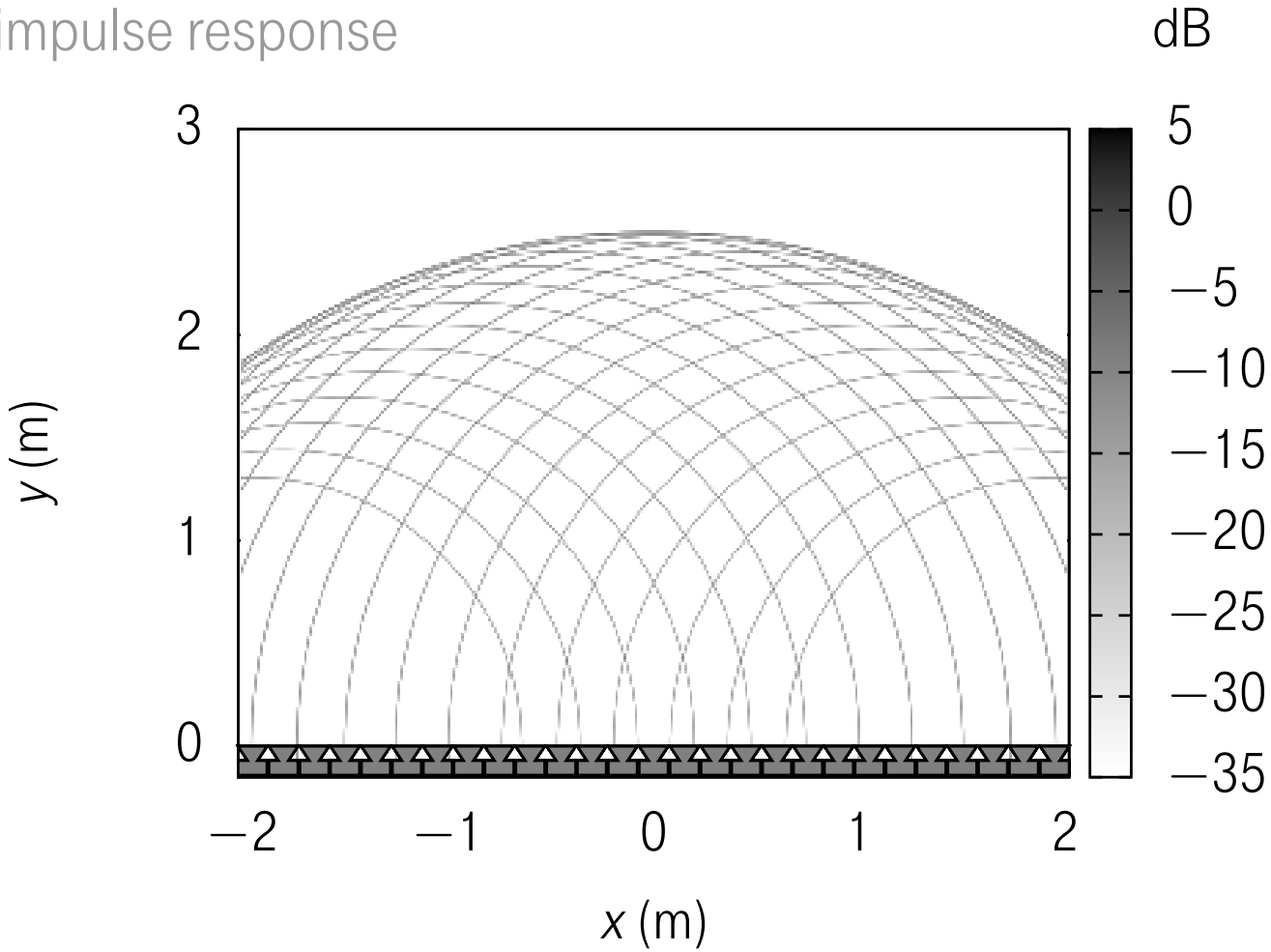
Snapshot of impulse response



- perception is dominated by the first wave front
- additional wave fronts add coloration (Wittek 2007) and an unnatural room impression

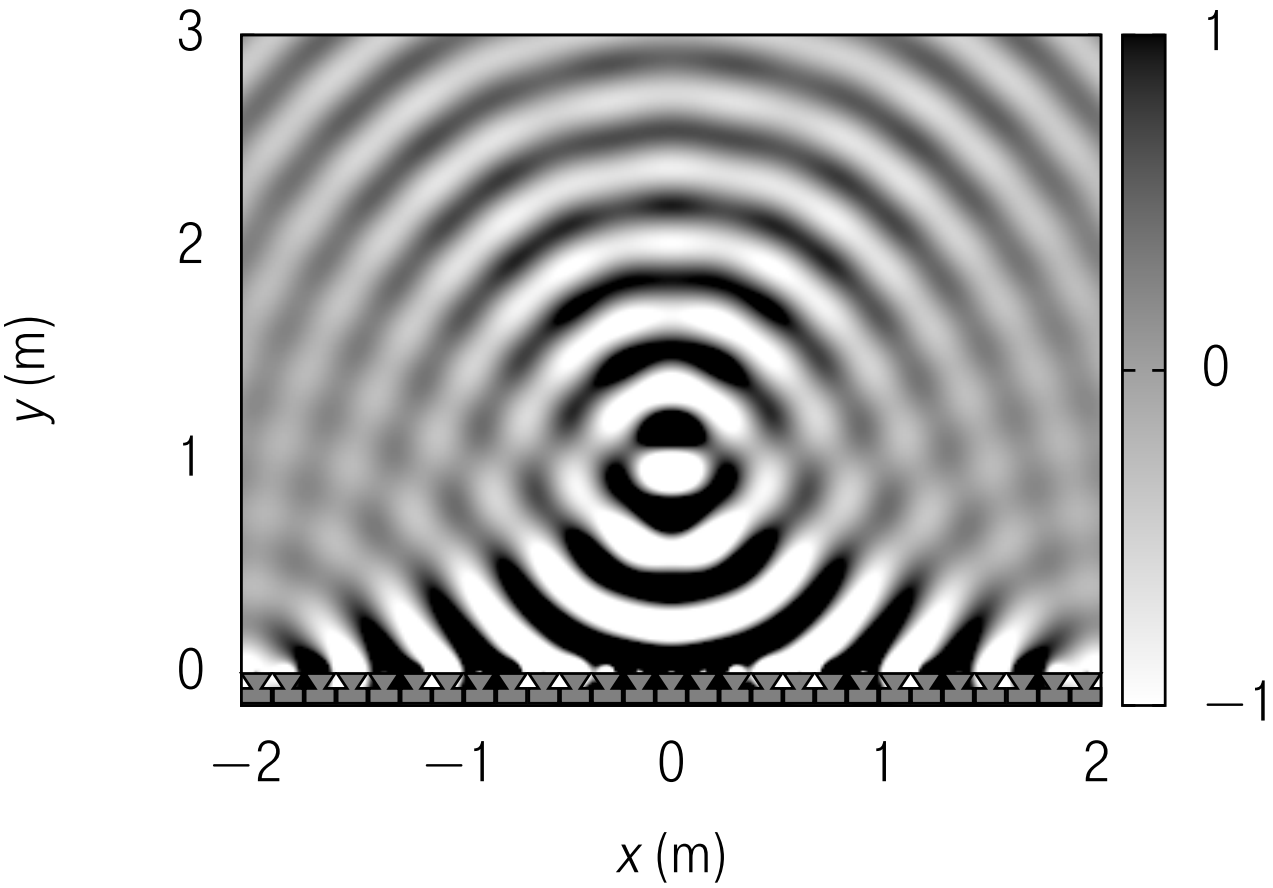
Wave Field Synthesis

Snapshot of impulse response



- perception is dominated by the first wave front
- additional wave fronts add coloration (Wittek 2007) and an unnatural room impression

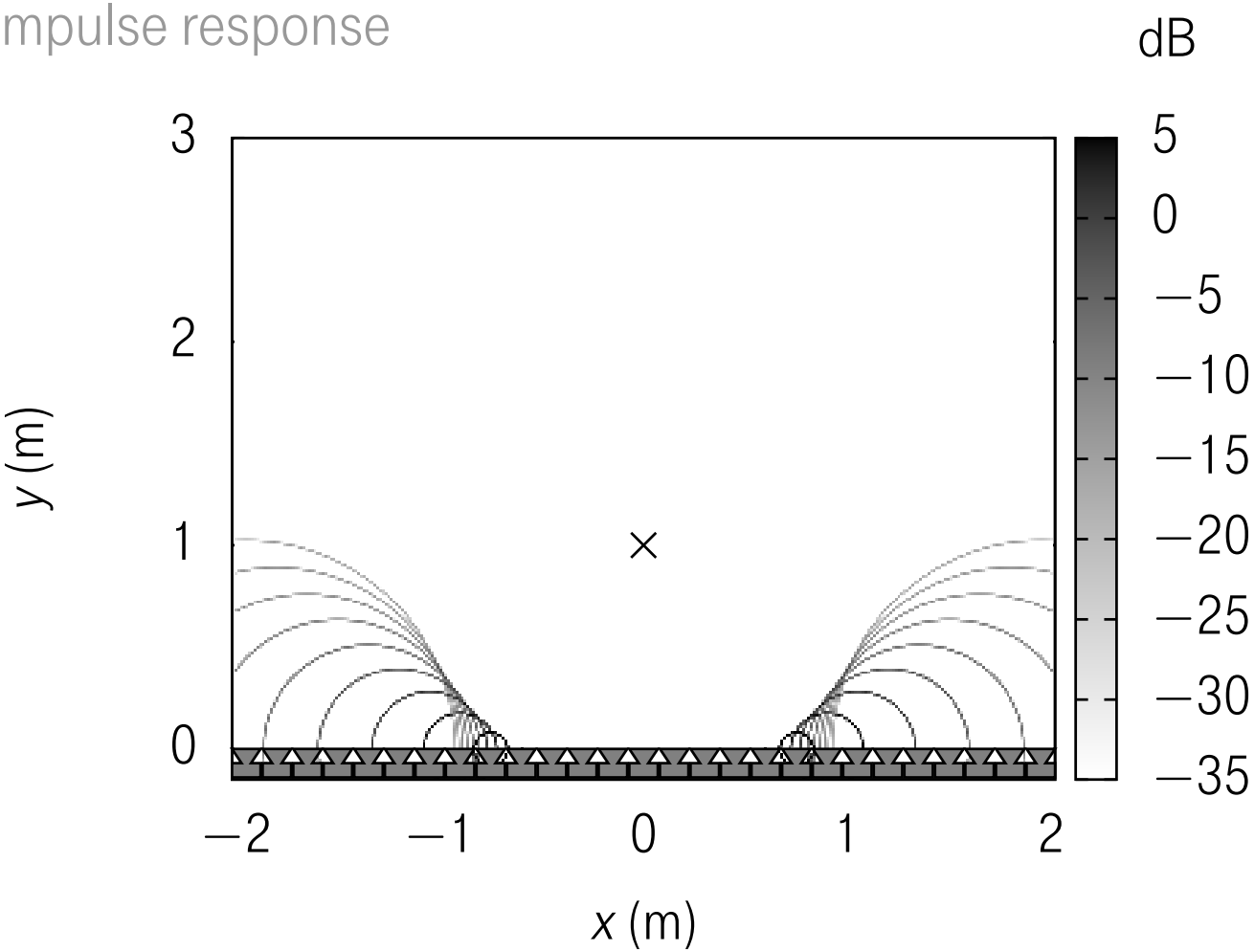
Focused Sources



- focused source located between listener and loudspeaker array
- converges to the focal point and diverges afterwards

Focused Sources

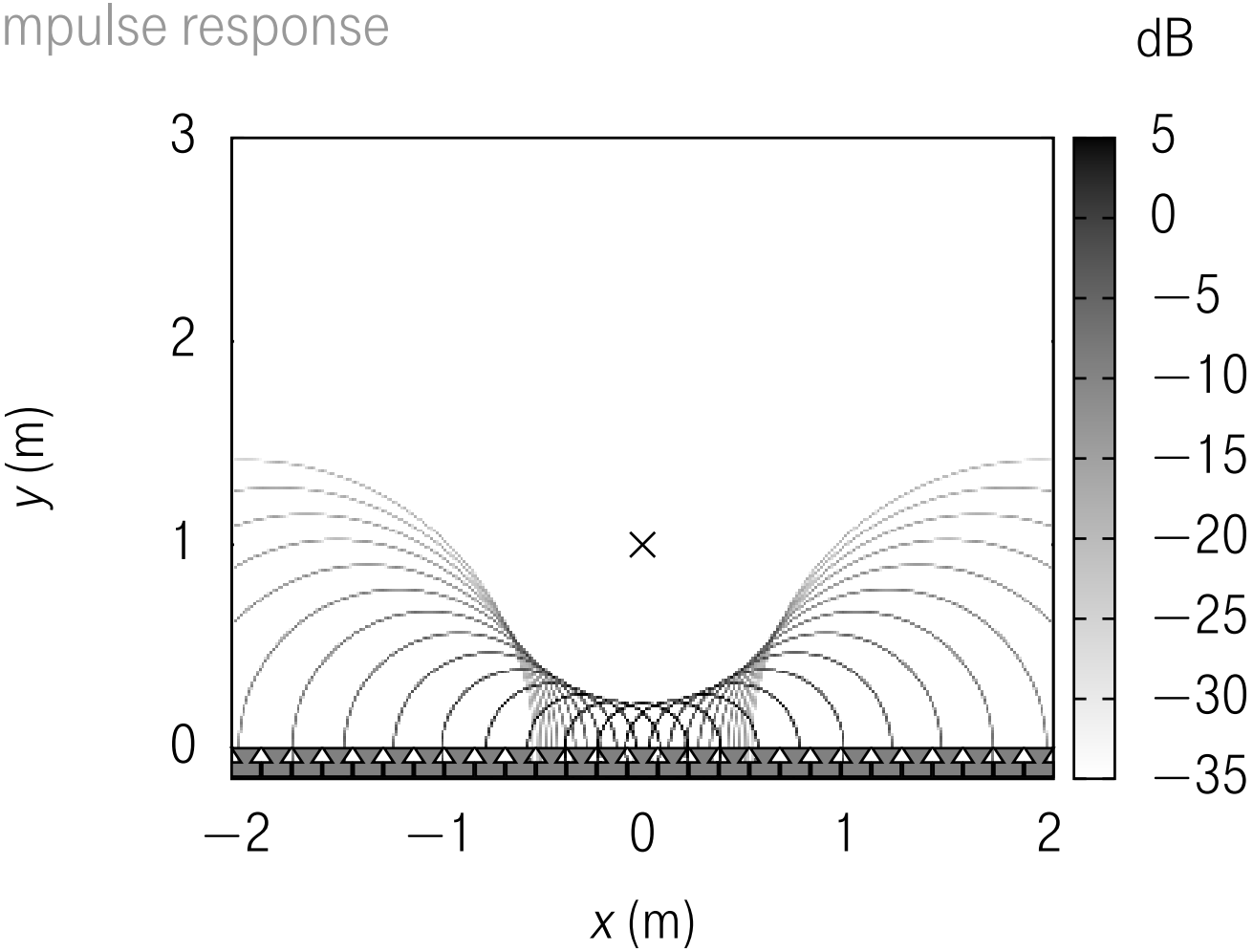
Snapshot of impulse response



- additional wave fronts arrive at the listener before the desired wave front
- different perception ⇒ echo artifacts

Focused Sources

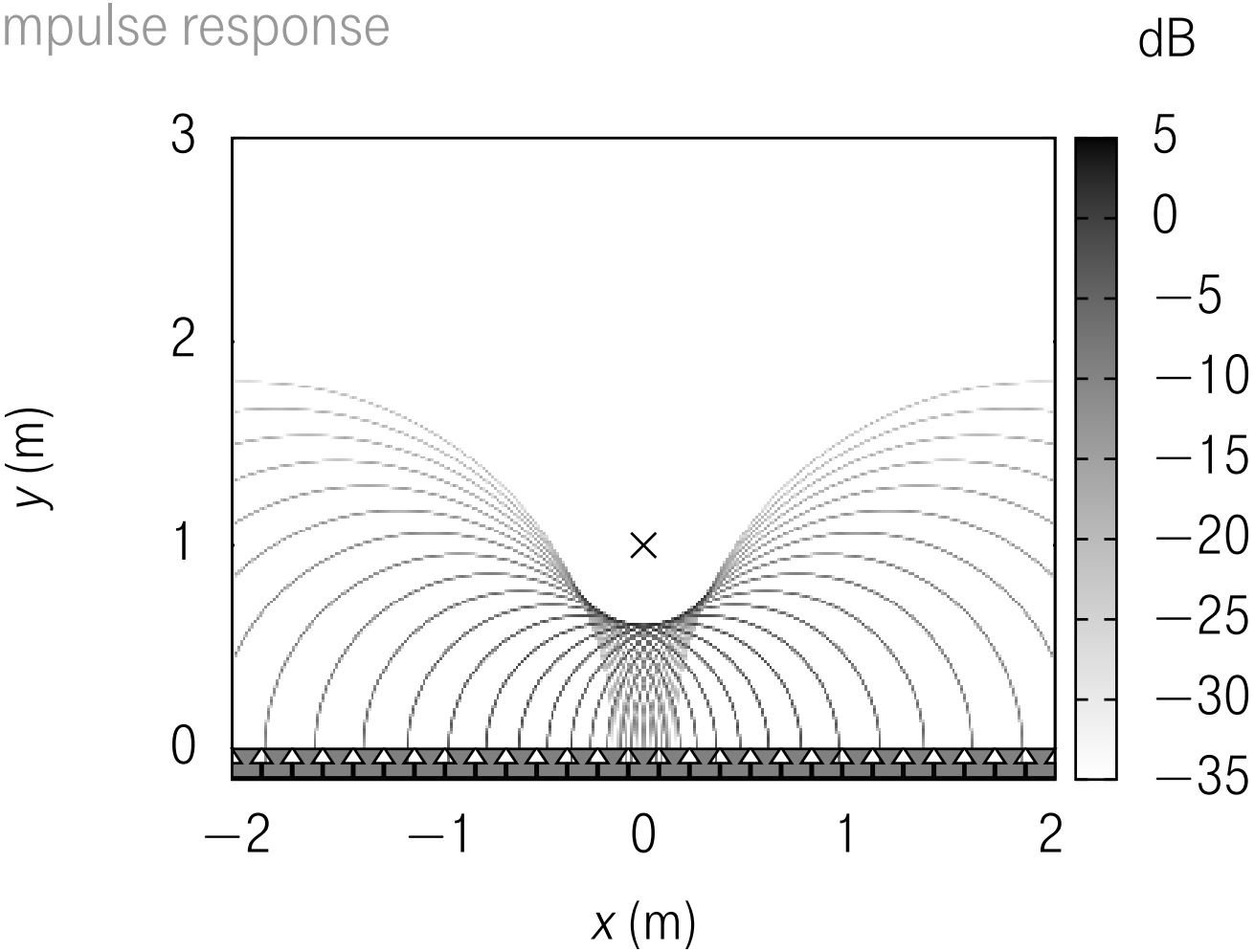
Snapshot of impulse response



- additional wave fronts arrive at the listener before the desired wave front
- different perception ⇒ echo artifacts

Focused Sources

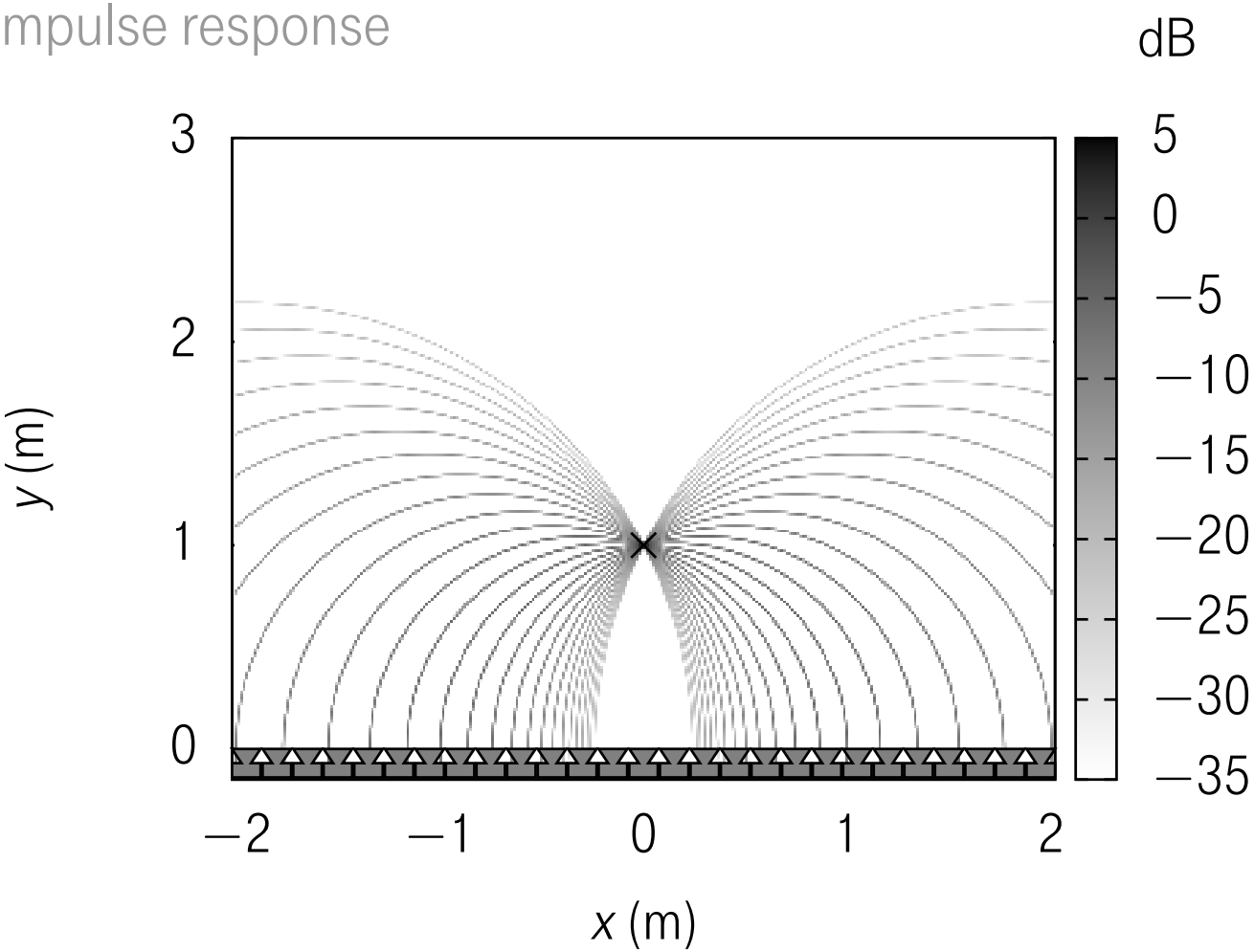
Snapshot of impulse response



- additional wave fronts arrive at the listener before the desired wave front
- different perception ⇒ echo artifacts

Focused Sources

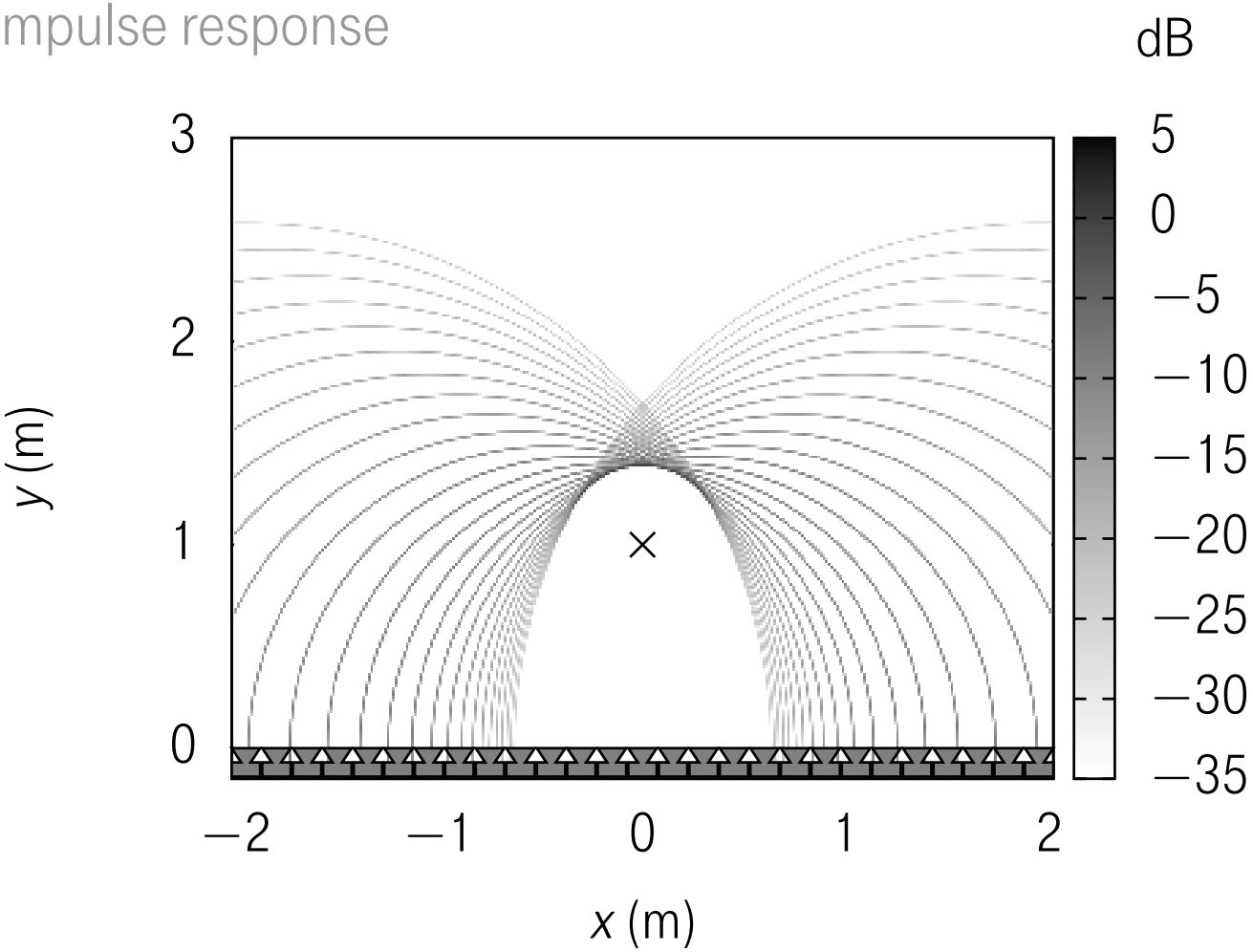
Snapshot of impulse response



- additional wave fronts arrive at the listener before the desired wave front
- different perception ⇒ echo artifacts

Focused Sources

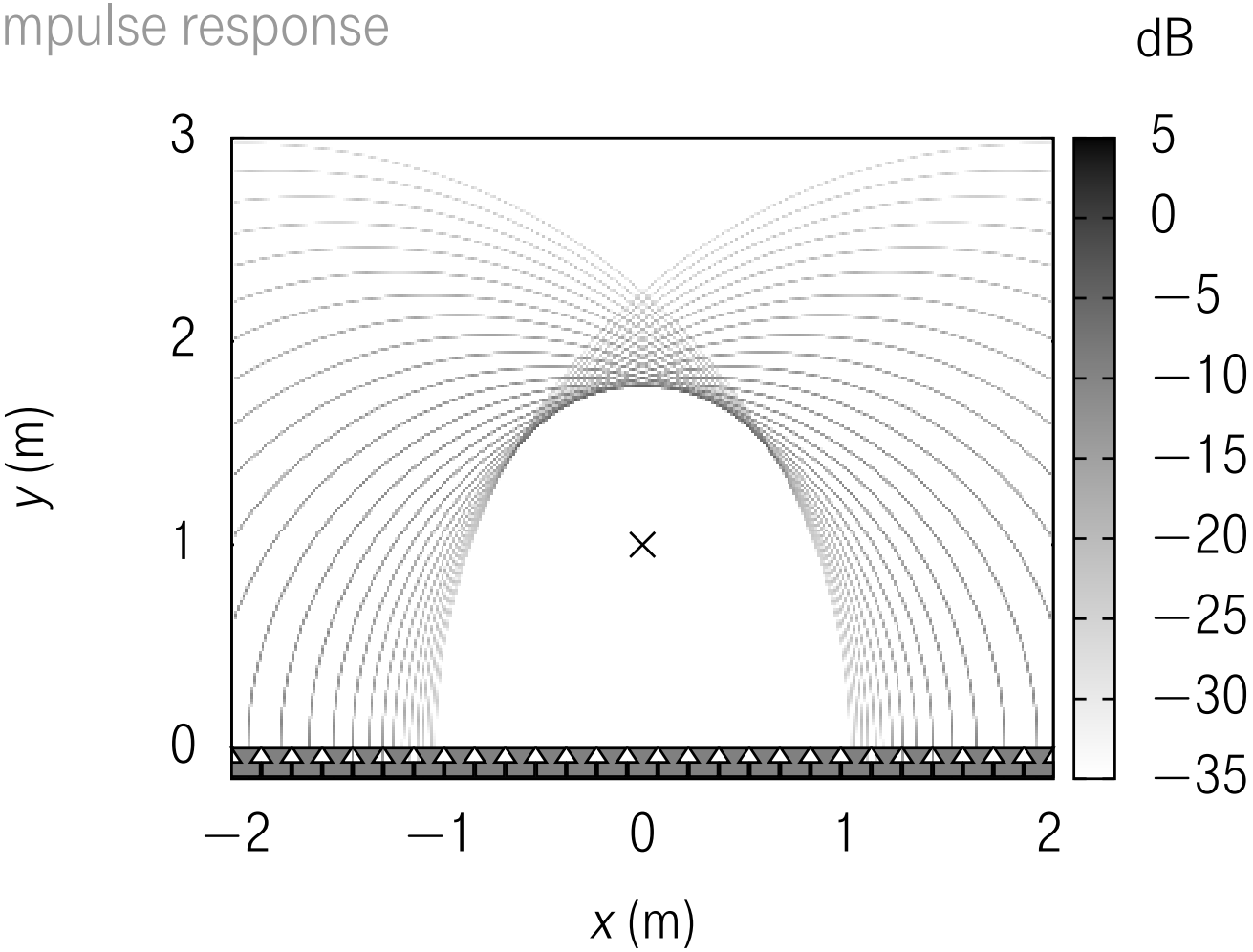
Snapshot of impulse response



- additional wave fronts arrive at the listener before the desired wave front
- different perception ⇒ echo artifacts

Focused Sources

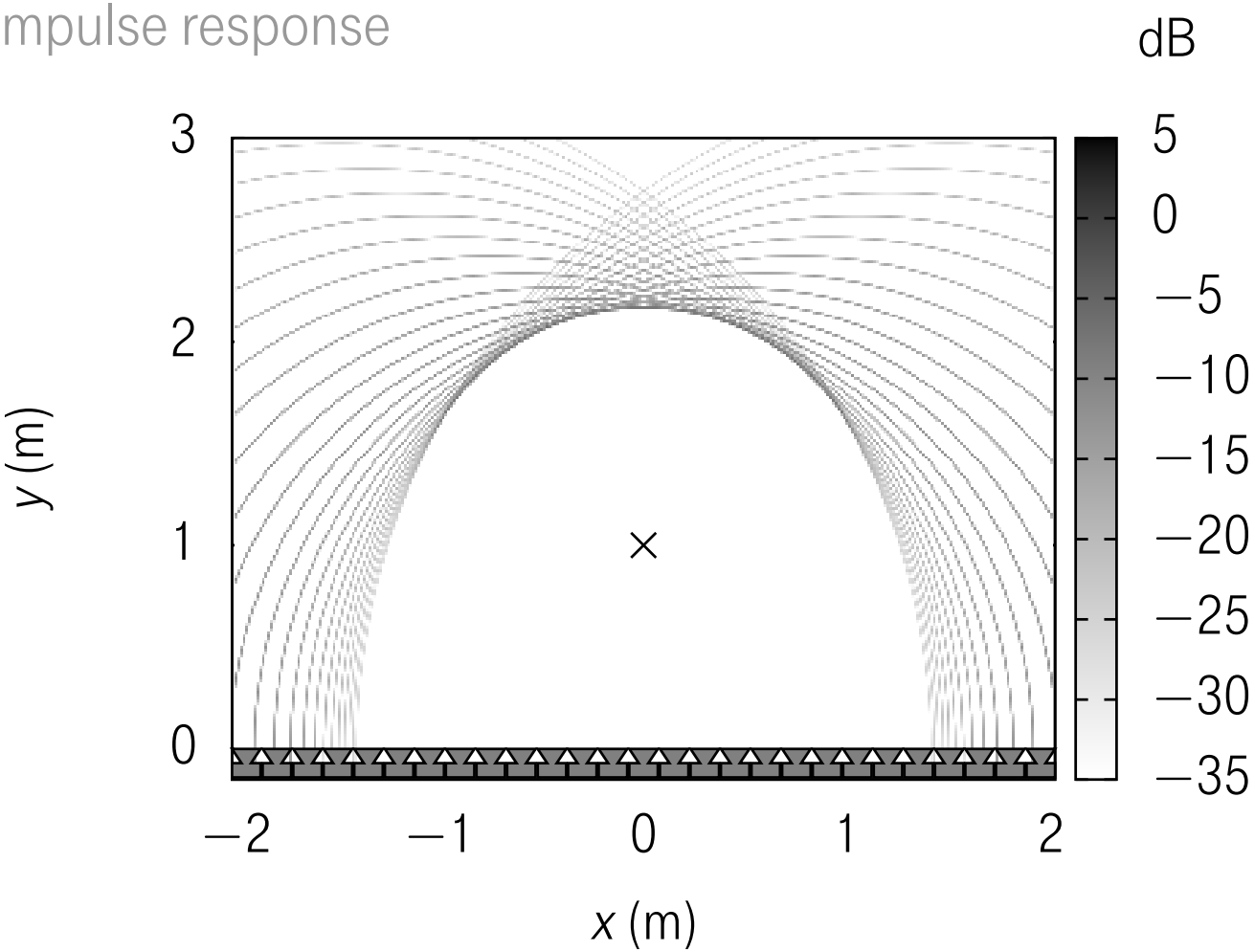
Snapshot of impulse response



- additional wave fronts arrive at the listener before the desired wave front
- different perception ⇒ echo artifacts

Focused Sources

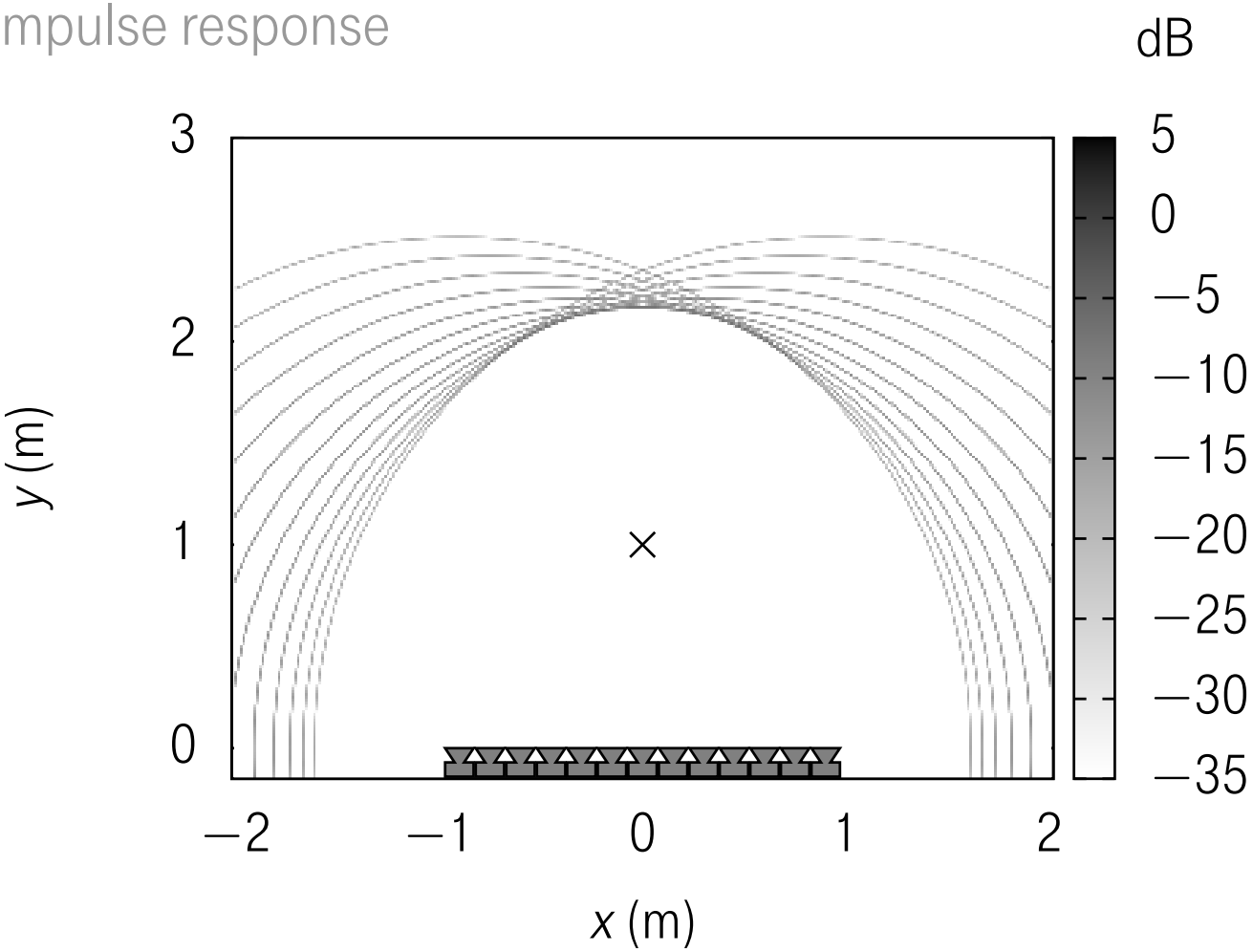
Snapshot of impulse response



- additional wave fronts arrive at the listener before the desired wave front
- different perception ⇒ echo artifacts

Focused Sources

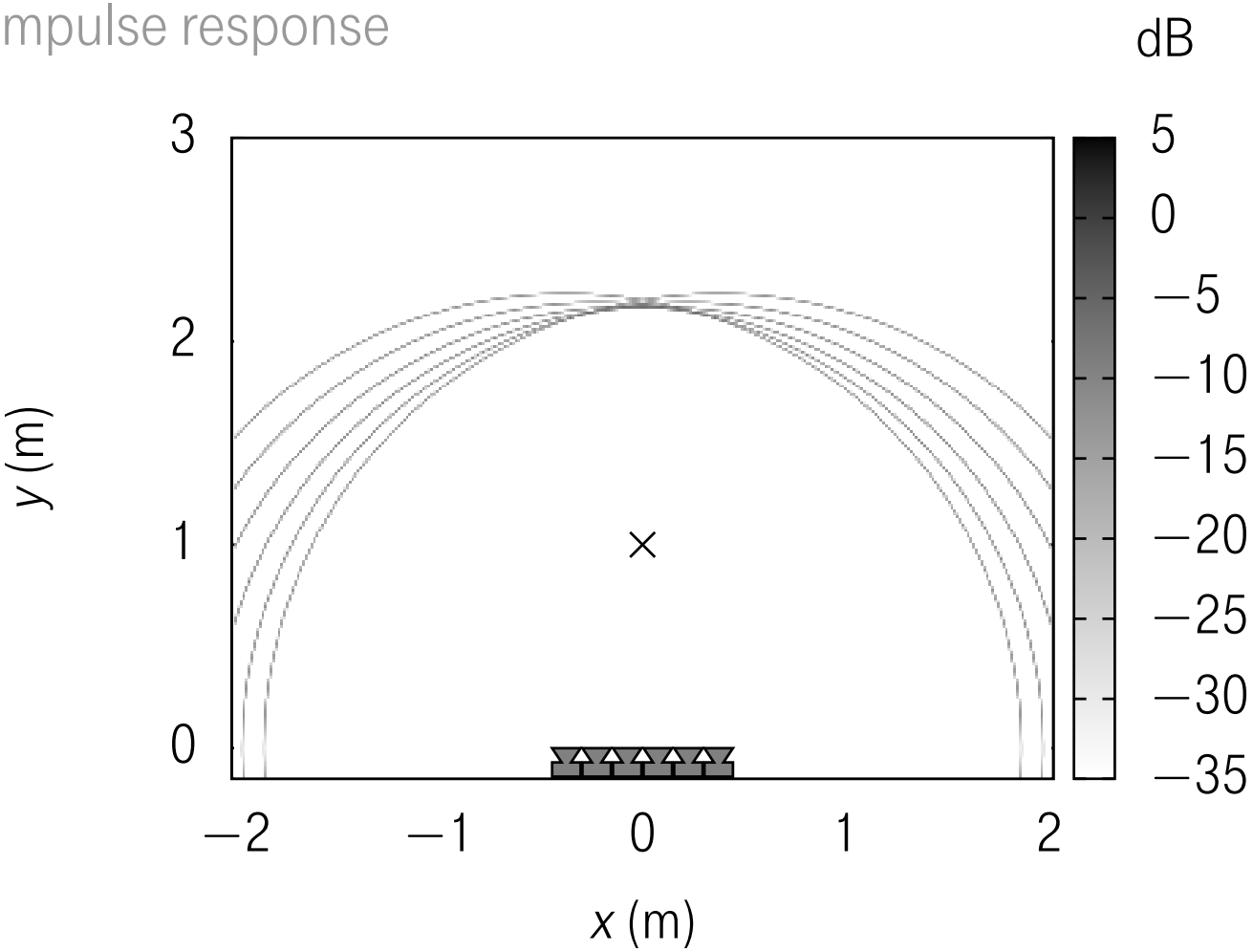
Snapshot of impulse response



- one additional wave front per loudspeaker
- ⇒ less audible artifacts for smaller arrays (Wierstorf 2010)

Focused Sources

Snapshot of impulse response



- one additional wave front per loudspeaker
- ⇒ less audible artifacts for smaller arrays (Wierstorf 2010)

Focused Sources

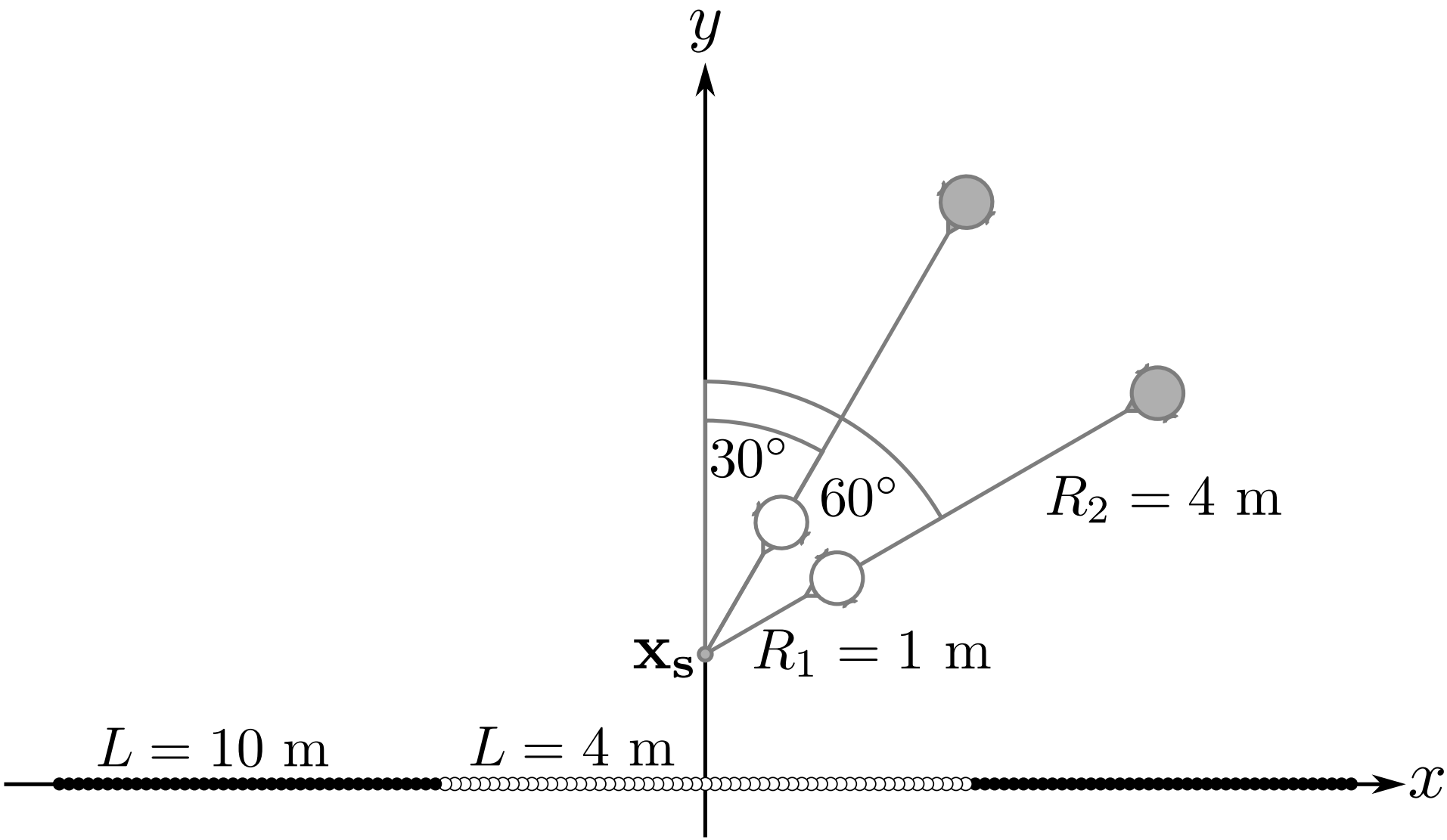
Localization

The localization of focused sources is affected by spatial sampling.
The additional wave fronts may trigger the ...

- precedence effect
  $\Rightarrow$ the perceived location is dominated by the ends of the loudspeaker array
- summing localization
  $\Rightarrow$ the perceived location is somewhere between the loudspeaker array and the focused source

Experiment

Geometry



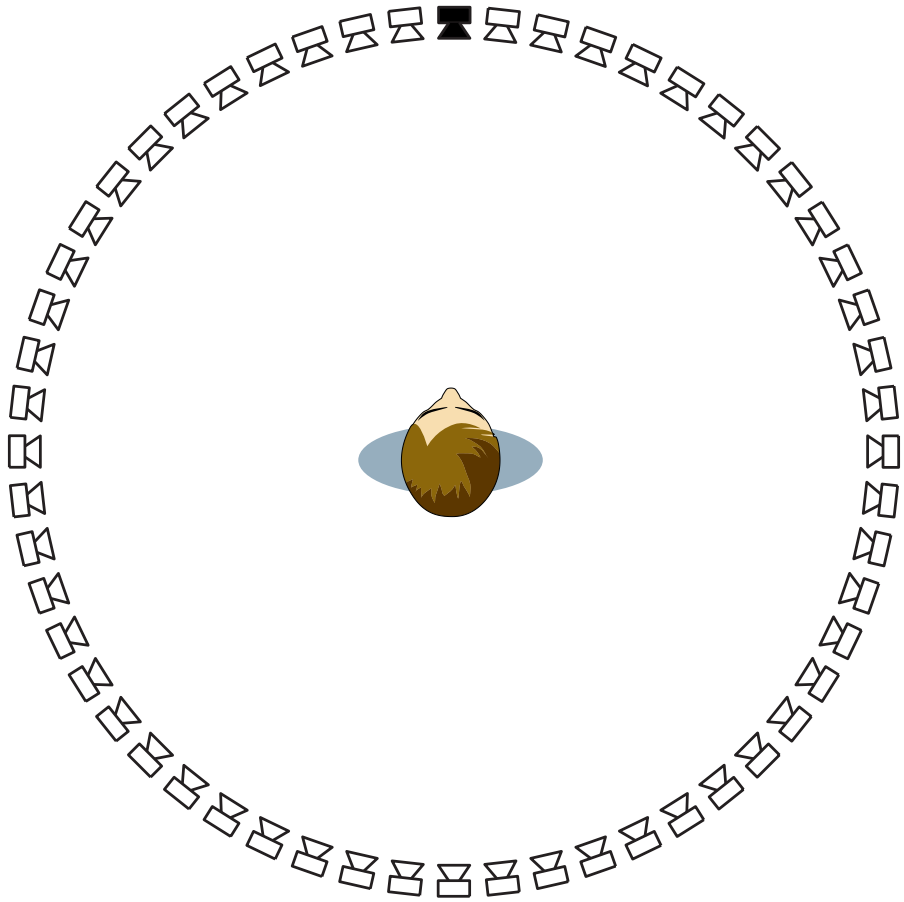
Experiment

Method

- length of loudspeaker arrays: 10 m, 4 m, 1.8 m, 0.75 m, 0.3 m with a distance of 0.15 m between speakers, single loudspeaker as reference (ref)
- dynamic (head tracking) binaural resynthesis of virtual loudspeakers
- 6 subjects
- rating of the attribute pair left vs. right
- castanets and speech as signal
- SoundScape Renderer and example stimuli can be found on our BLOG: `http://audio.qu.tu-berlin.de`

Experiment

Dynamic Binaural Resynthesis



- measured set of head related transfer functions (HRTF)

Experiment

Dynamic Binaural Resynthesis

- interpolate and extrapolate HRTF set for desired loudspeaker positions

Experiment

Dynamic Binaural Resynthesis

- repeat this for every possible head orientation

Experiment

Dynamic Binaural Resynthesis

- repeat this for every possible head orientation

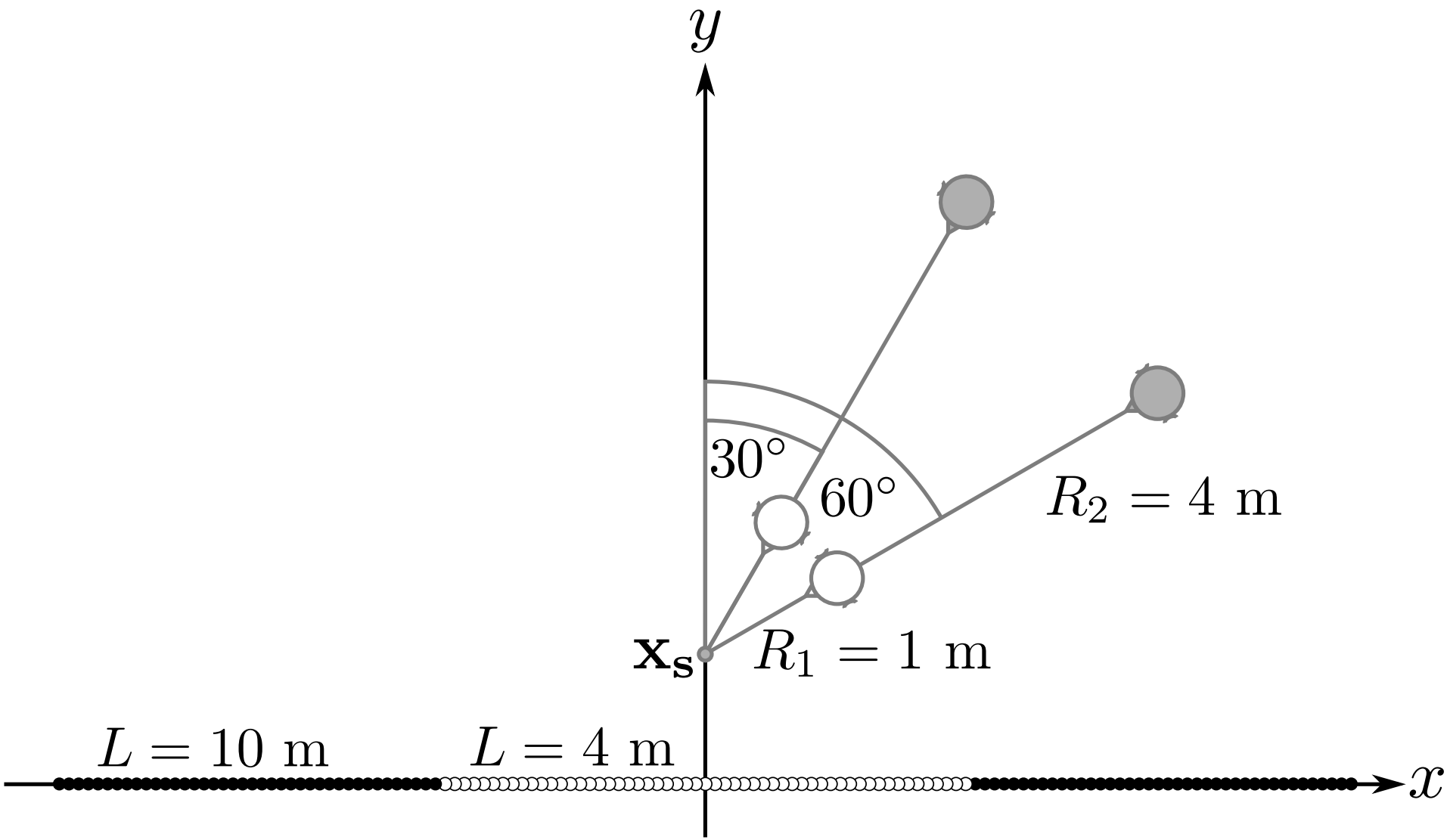
Experiment

Dynamic Binaural Resynthesis

- repeat this for every possible head orientation

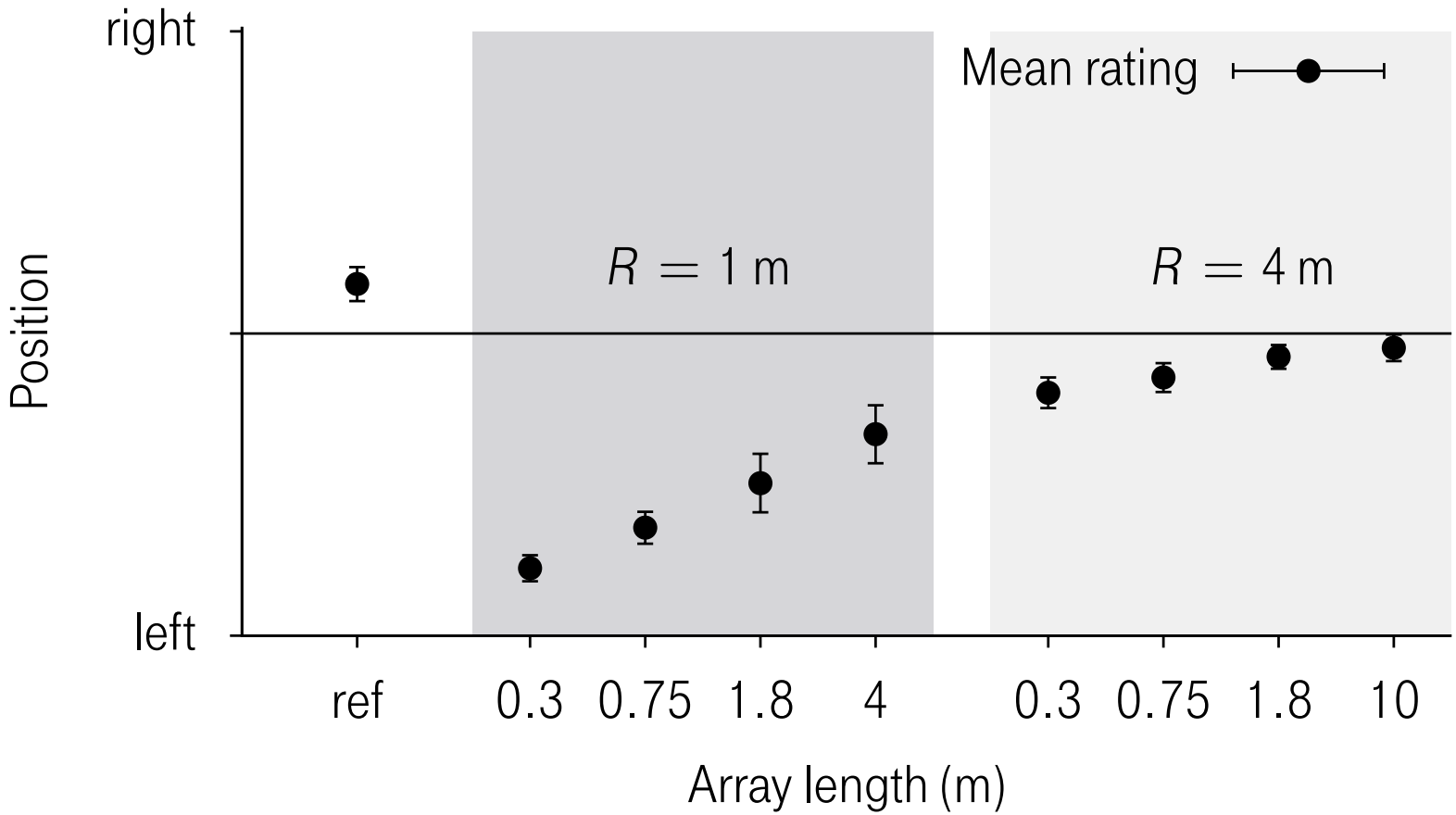
Experiment

Geometry



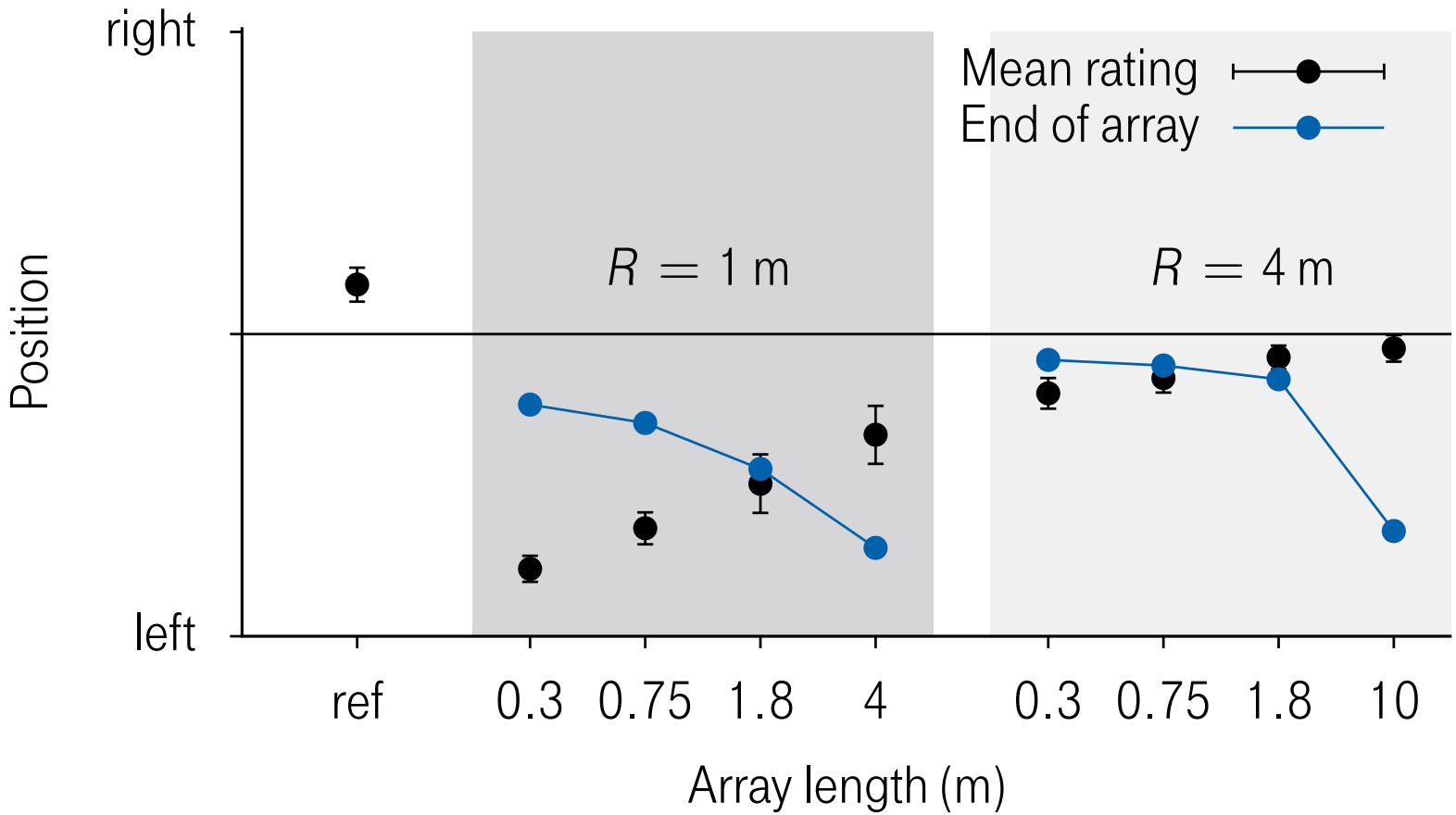
Experiment

Results



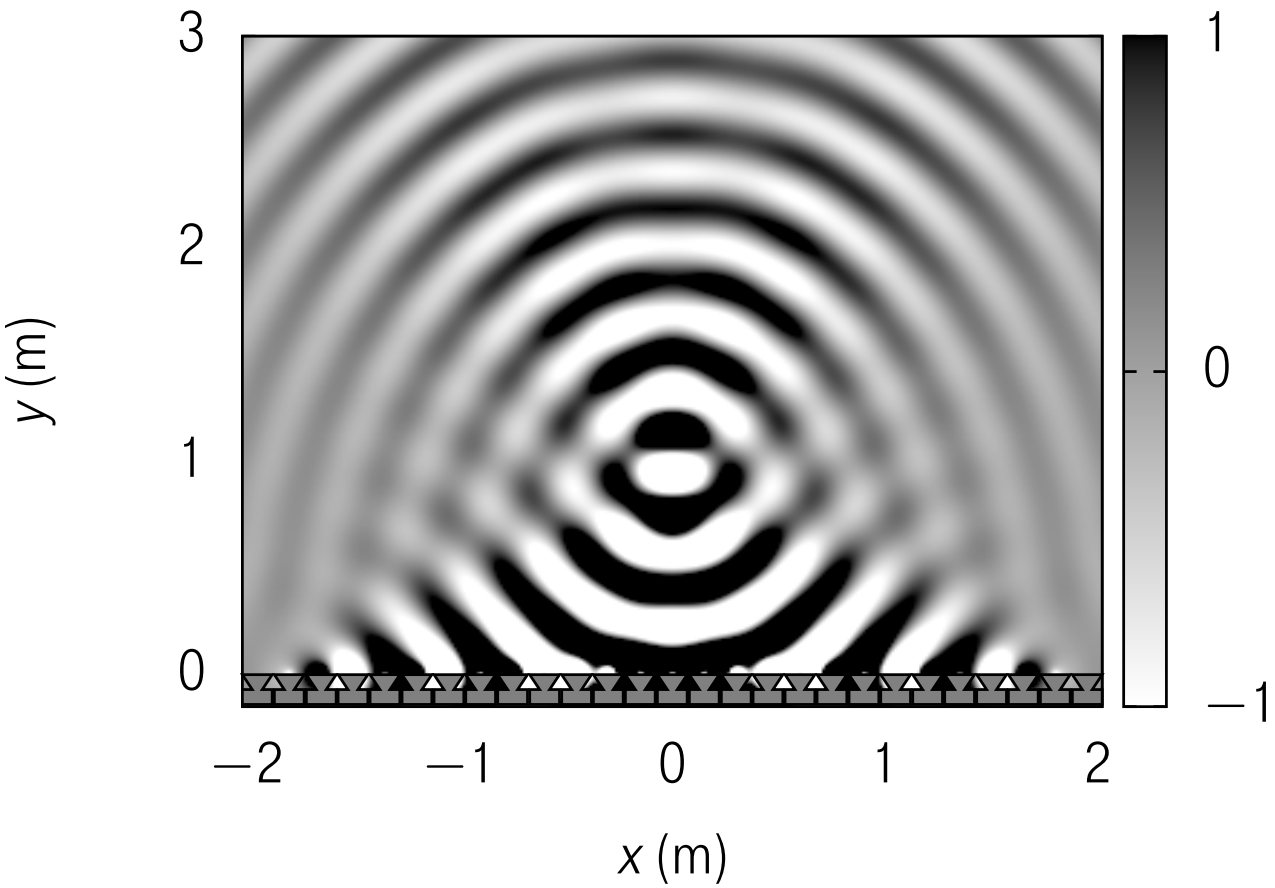
Experiment

Results



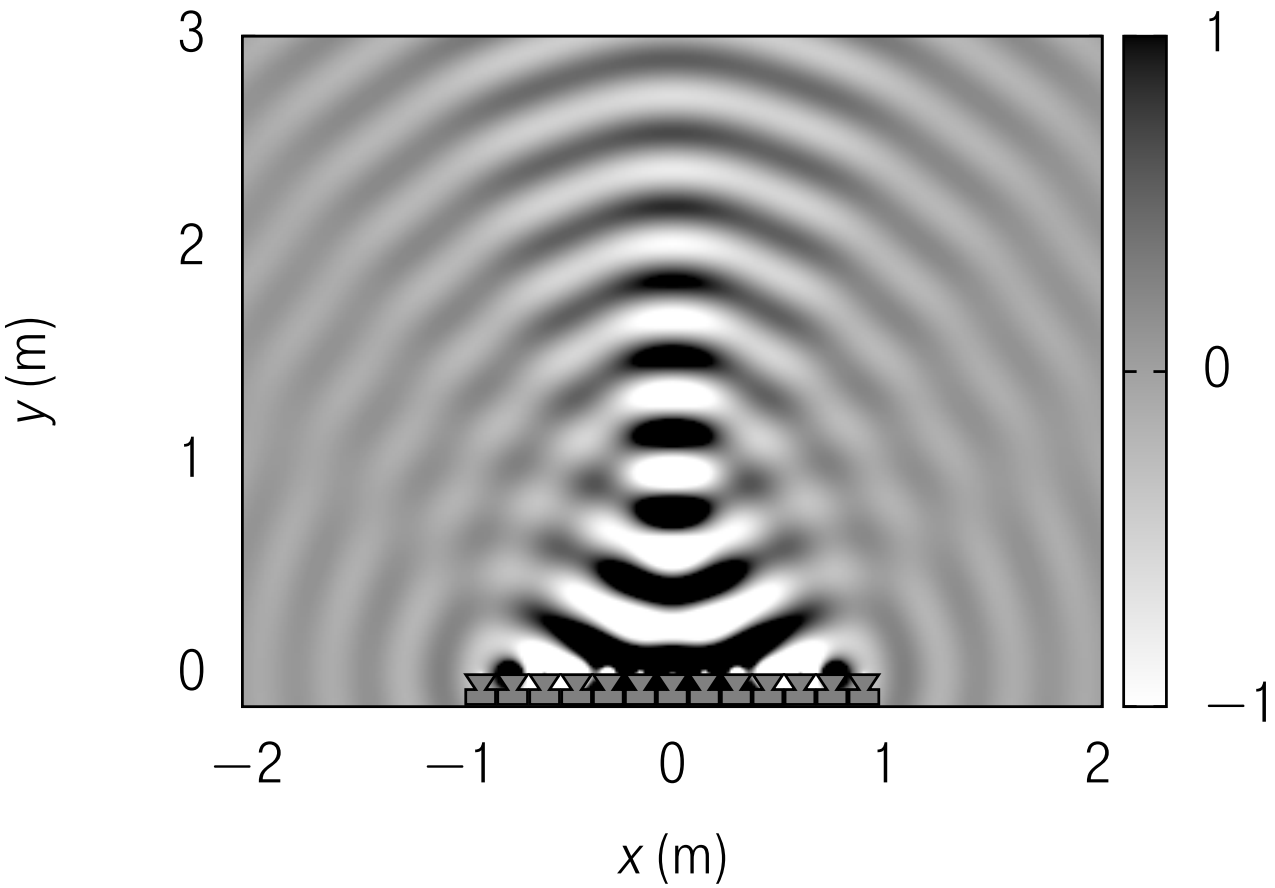
- results can not be explained by precedence effect or summing localization

Diffraction



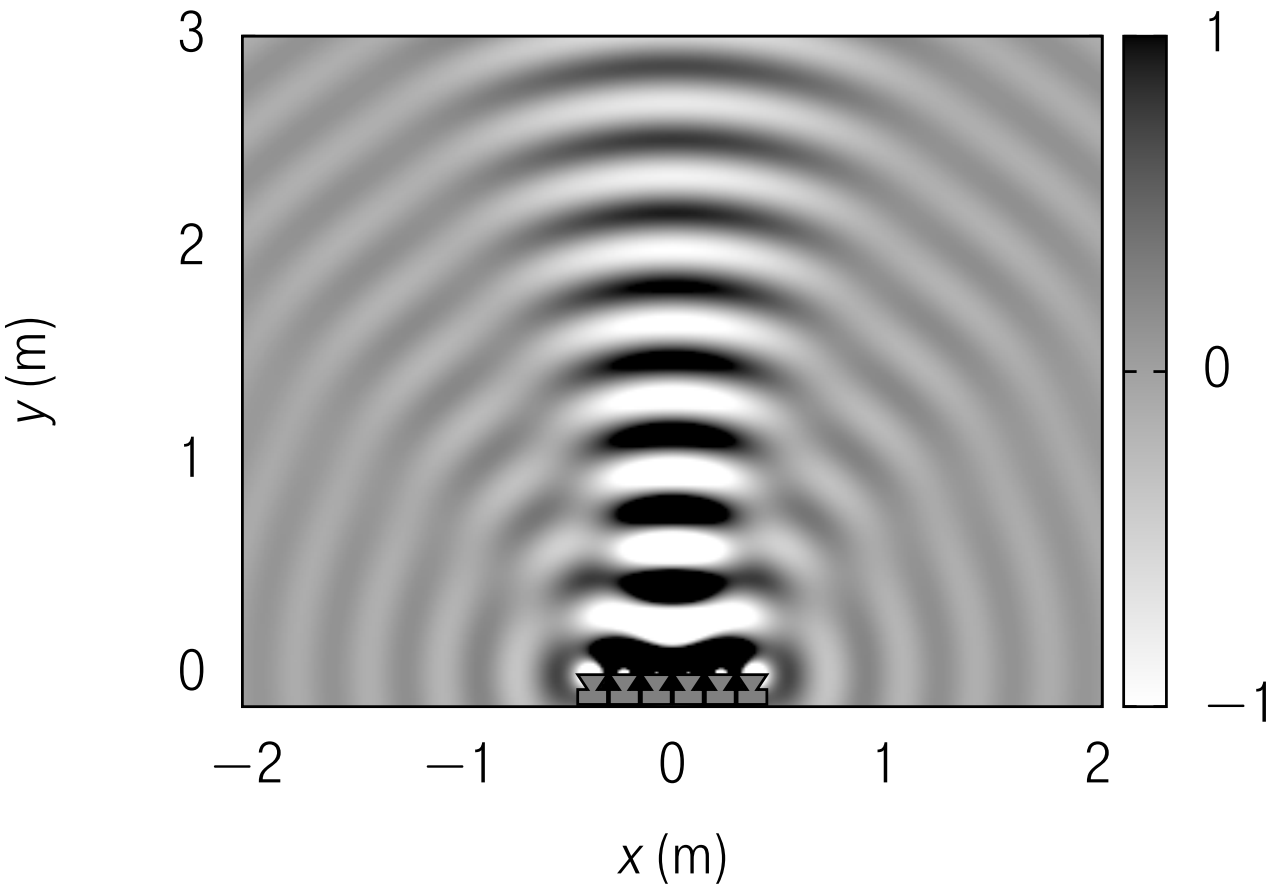
- diffraction depends on the frequency
- diffraction changes ILD

Diffraction



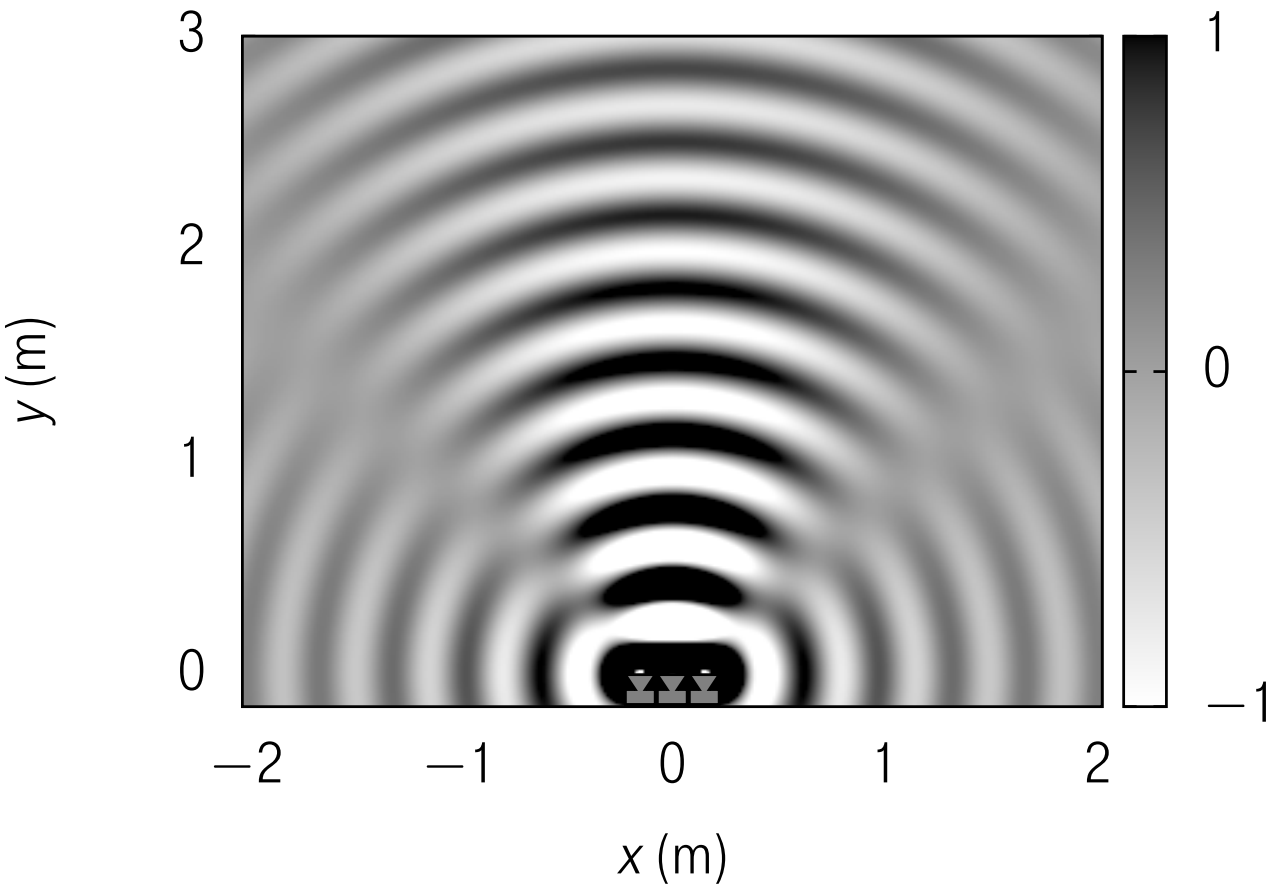
- diffraction depends on the frequency
- diffraction changes ILD

Diffraction



- diffraction depends on the frequency
- diffraction changes ILD

Diffraction



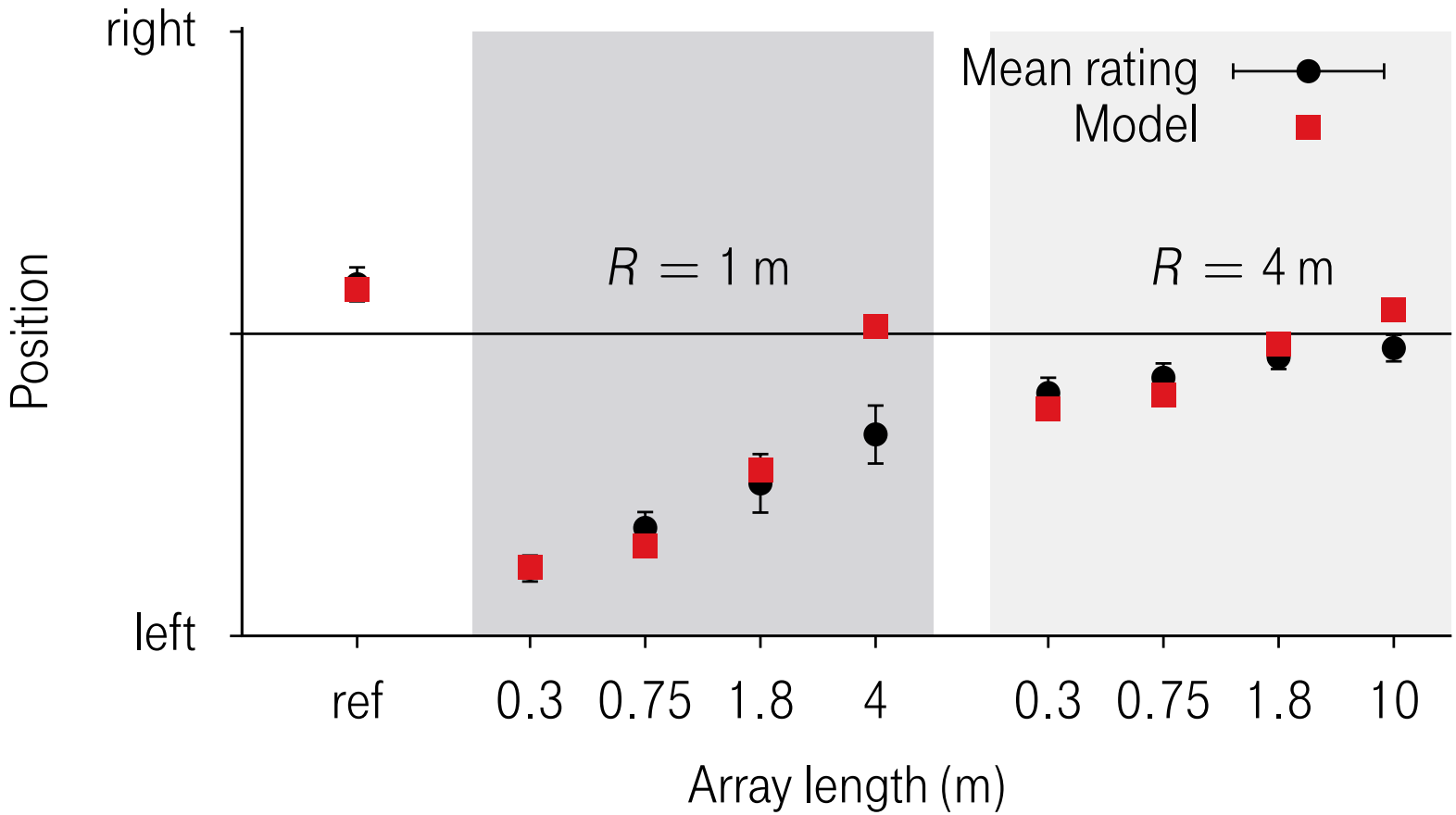
- diffraction depends on the frequency
- diffraction changes ILD

Binaural Model

- investigate the causes of observed undesired localization
- binaural model after Lindemann (1986a)
- same parameters used as in the original paper
- model is part of the Auditory Modelling Toolbox (Søndergaard 2011): `http://amtoolbox.sourceforge.net/`

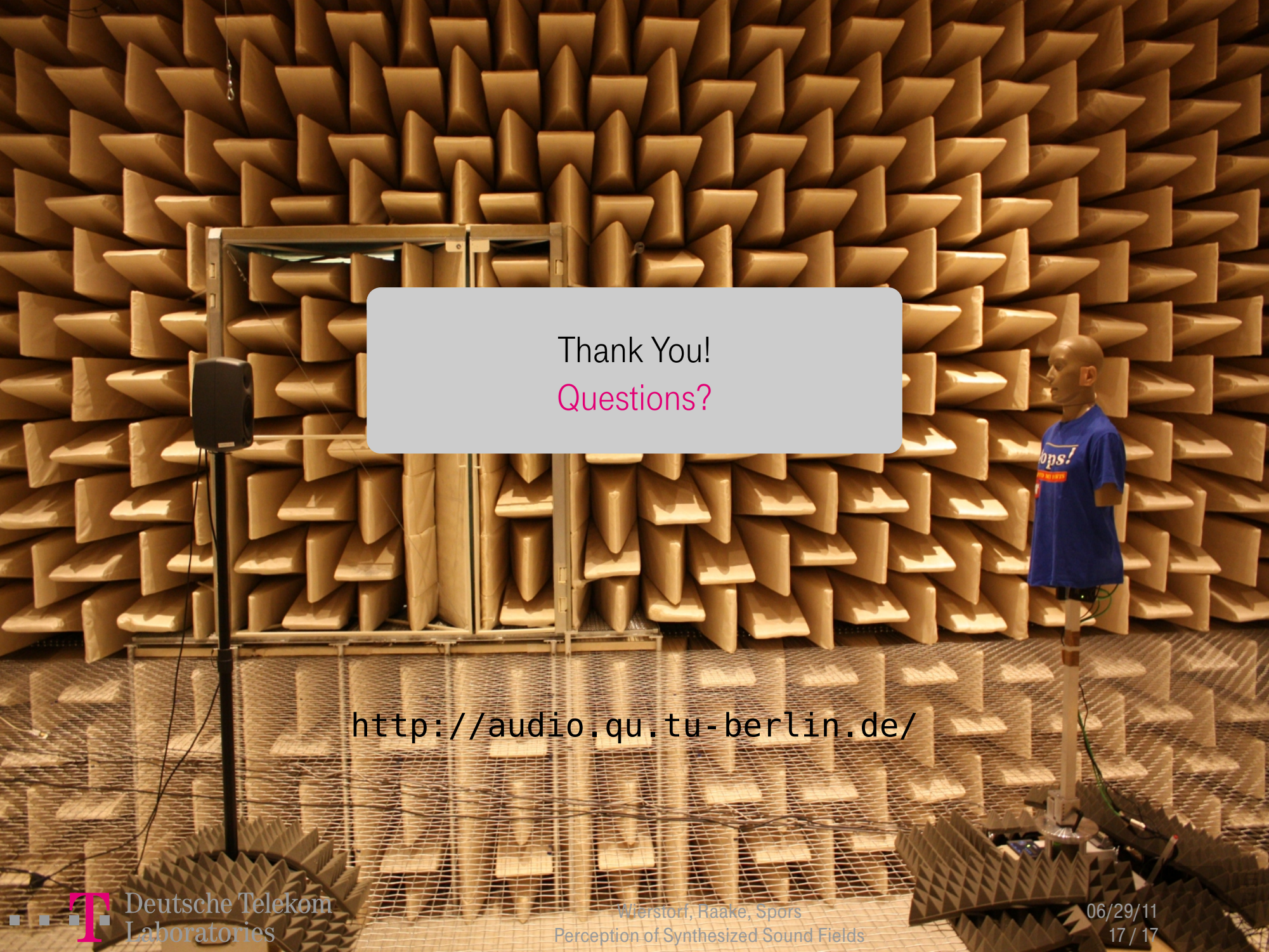
Binaural Model

Results



Summary

- spatial sampling leads to additional wave fronts in WFS
- the repetitions are more critical for focused sources
- diffraction due to the length of the array causes wrong localization of synthesised sources
- binaural model is able to predict wrong localization due to diffraction
- binaural model gives a hint for which conditions the precedence effect has to be considered
- open question: coloration due to the additional wave fronts


Thank You!
Questions?
http://audio.qu.tu-berlin.de/