Predicting localization accuracy for stereophonic downmixes in Wave Field Synthesis

Hagen Wierstorf[1], Sascha Spors[2]

[1] Assessment of IP-based Applications, TU Berlin
[2] Institut für Nachrichtentechnik, Universität Rostock

11. September 2014

TWO!EARS

berlin

Introduction

Why stereophonic downmixes in Wave Field Synthesis?

- multi-channel reproduction uses often object oriented approach:
  Wave Field Synthesis, Dolby Atmos (Robinson 2012)
- most of material available only as two channel stereo

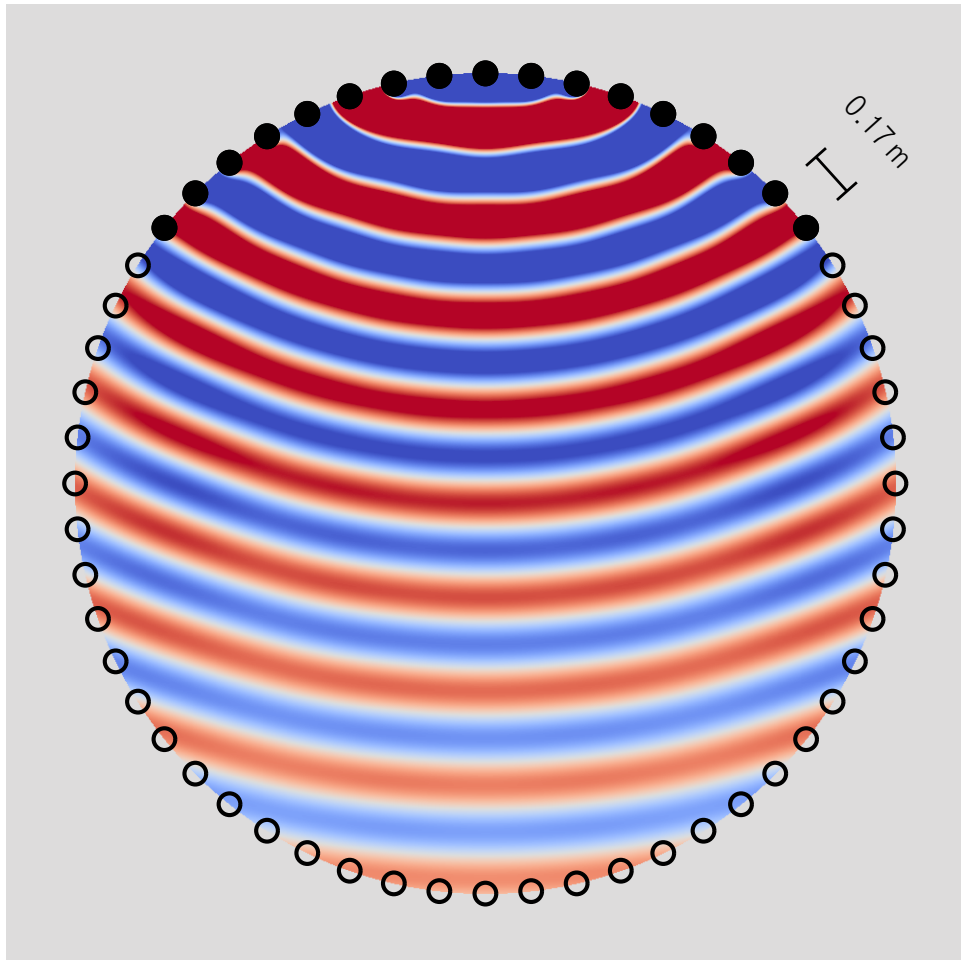
Research question

- can we increase the sweet spot of stereophony by using WFS?

Robinson et al. (2012), Scalable Format and Tools to Extend the Possibilities of Cinema Audio, SMPTE Motion Image Journal

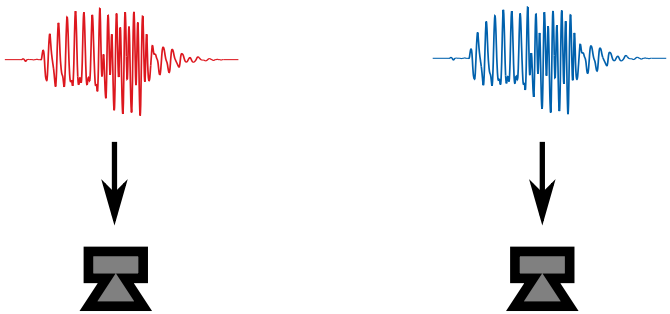
Wave Field Synthesis

$f = 1000\,\text{Hz}$

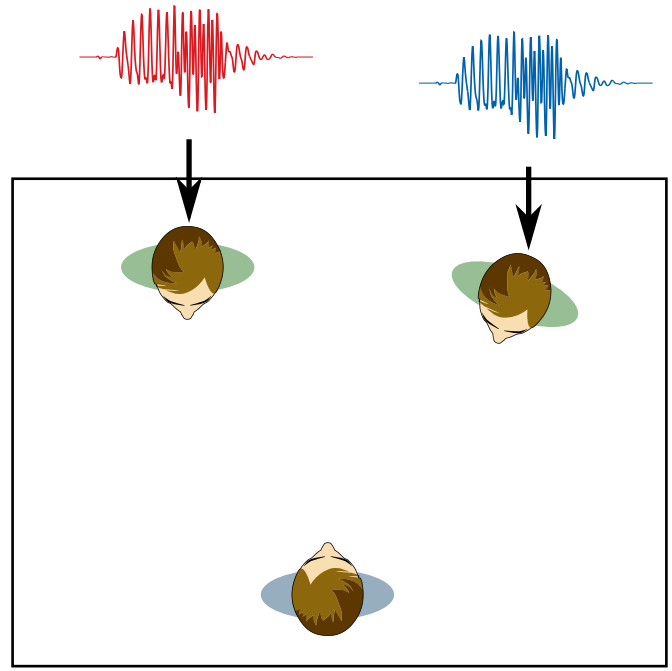


Channel Based vs. Object Based

Representing an Audio Scene



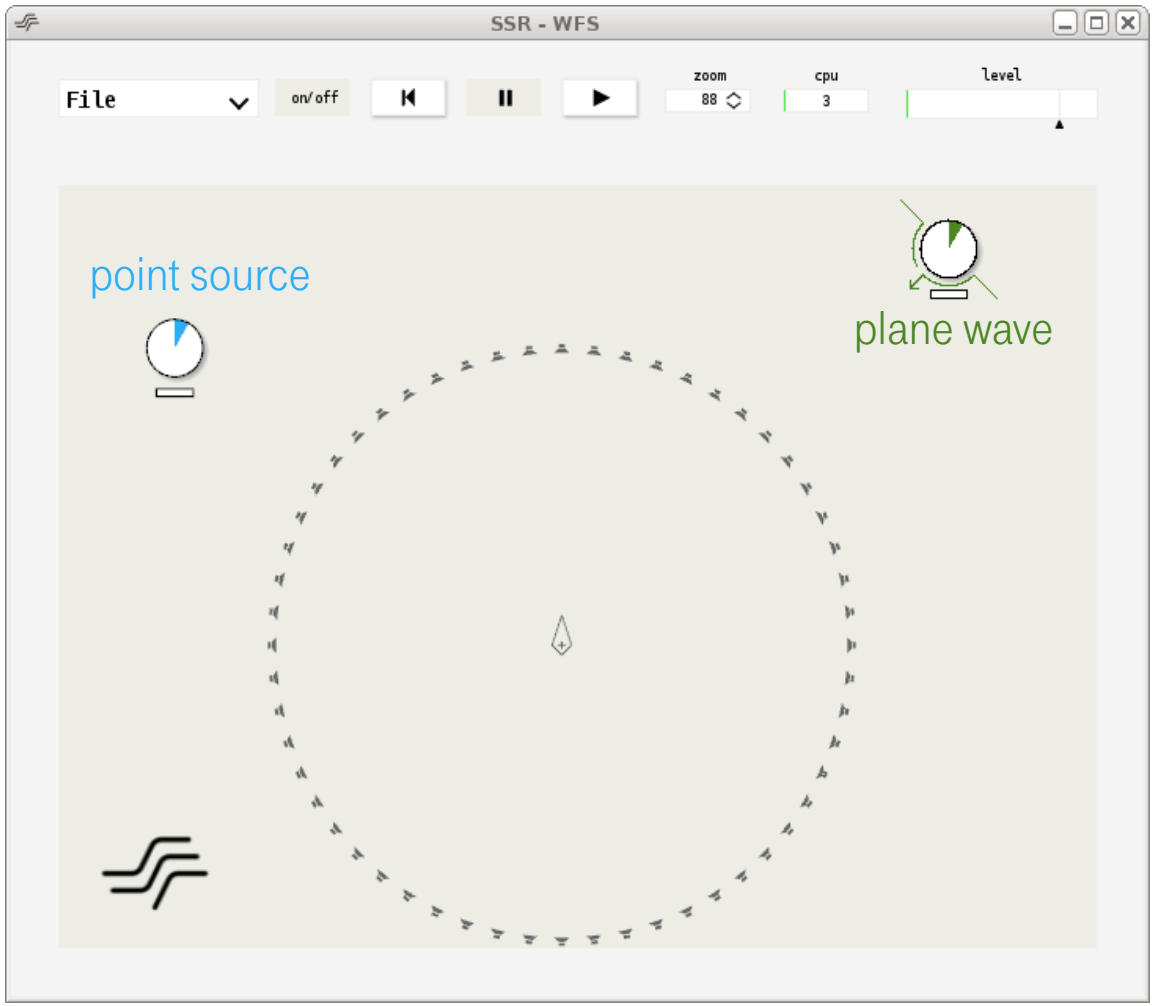
- information is stored in loudspeaker signals
- reproduction system is (implicitly) part of the stored information



- information is stored in source signals and audio scene description
- reproduction system is not part of the stored information

Object Based

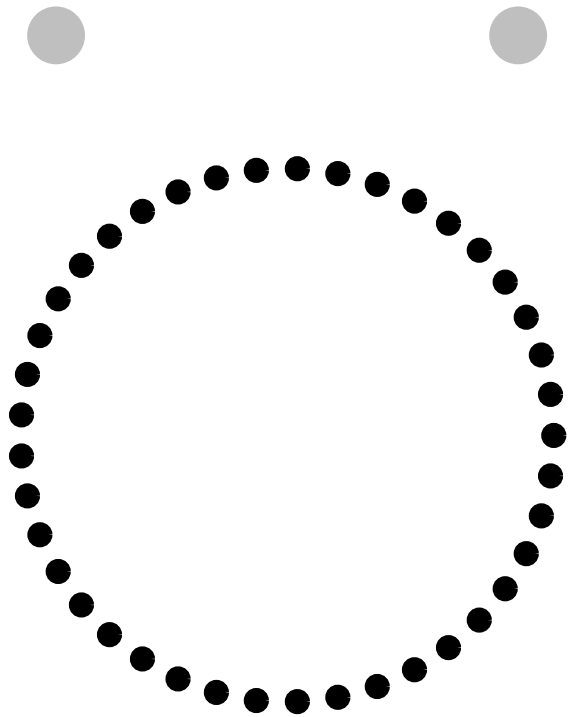
Interactivity + Independence of Reproduction System



Wave Field Synthesis

Downmixing Stereo

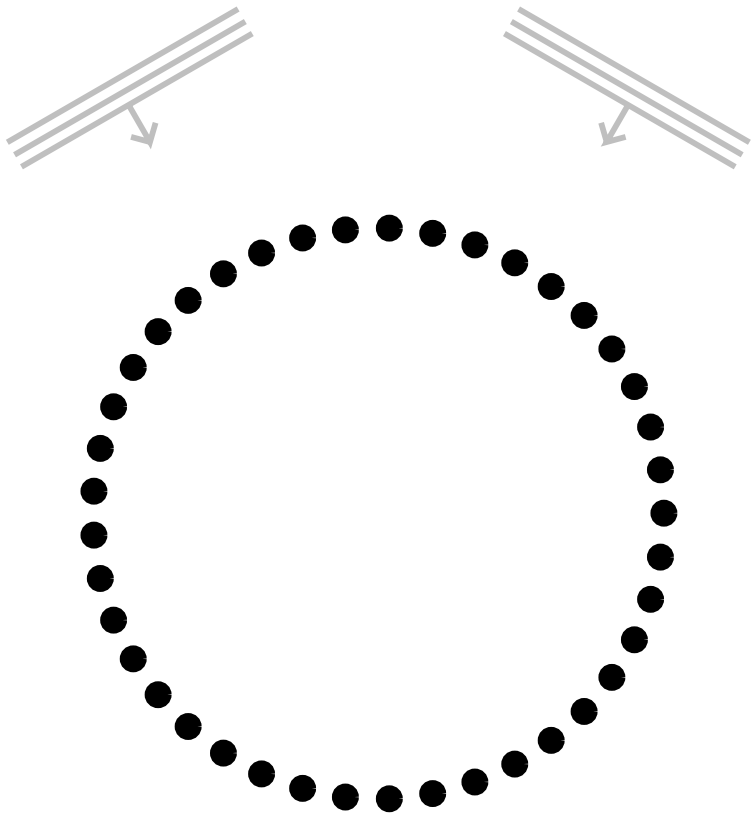
Virtual panning spots (Theile 2003)



Theile et al. (2003) Potential wave-field synthesis applications in the multichannel stereophonic world, 24th AES Conf.

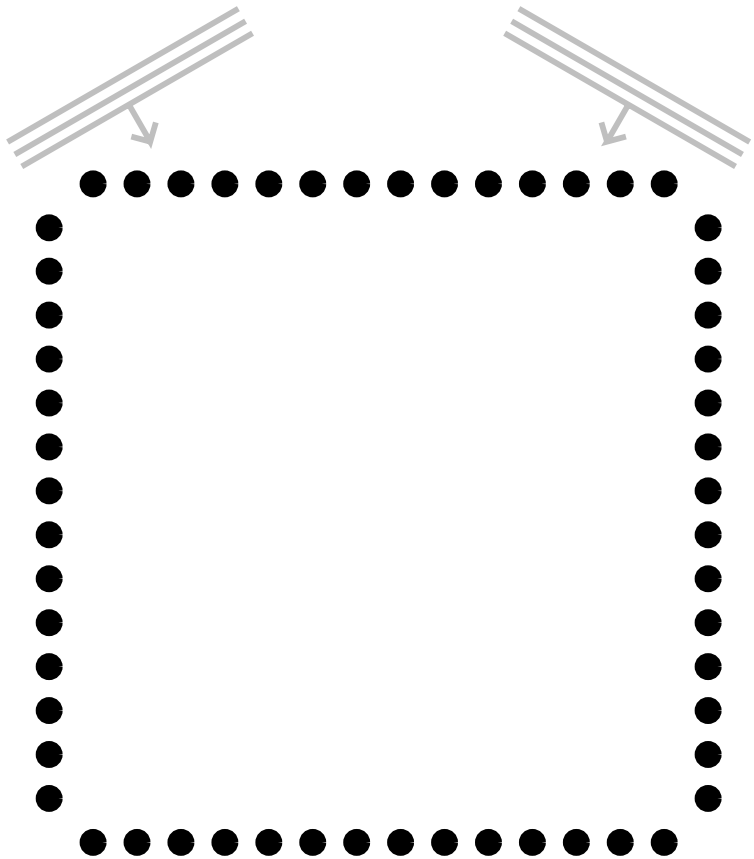
Wave Field Synthesis

Downmixing Stereo



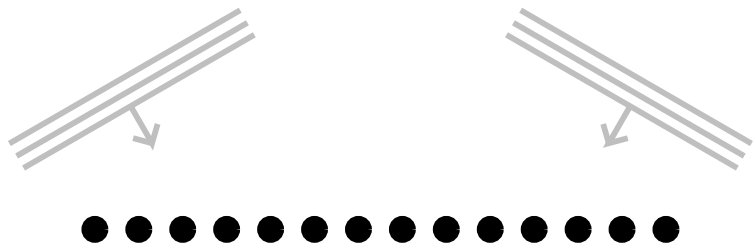
Wave Field Synthesis

Downmixing Stereo



Wave Field Synthesis

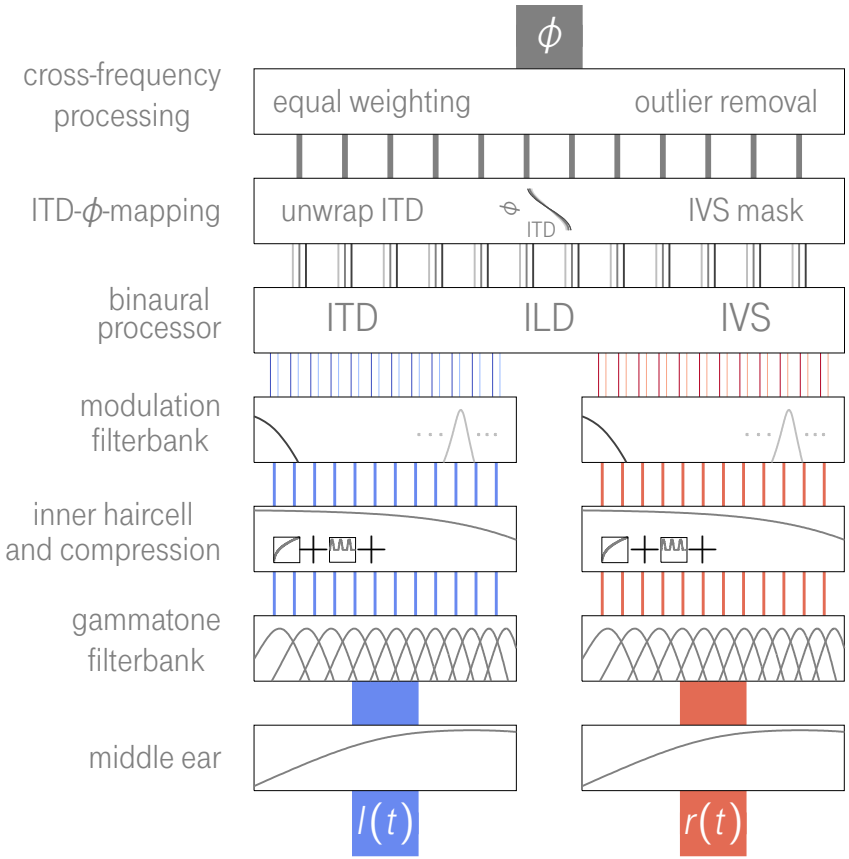
Downmixing Stereo



Predicting Localization Accuracy

Binaural Model

- binaural model (Dietz 2011)
- prediction of localization in WFS
- binaural simulation of ear signals
- accuracy of model around 1.5° (Wierstorf 2013)

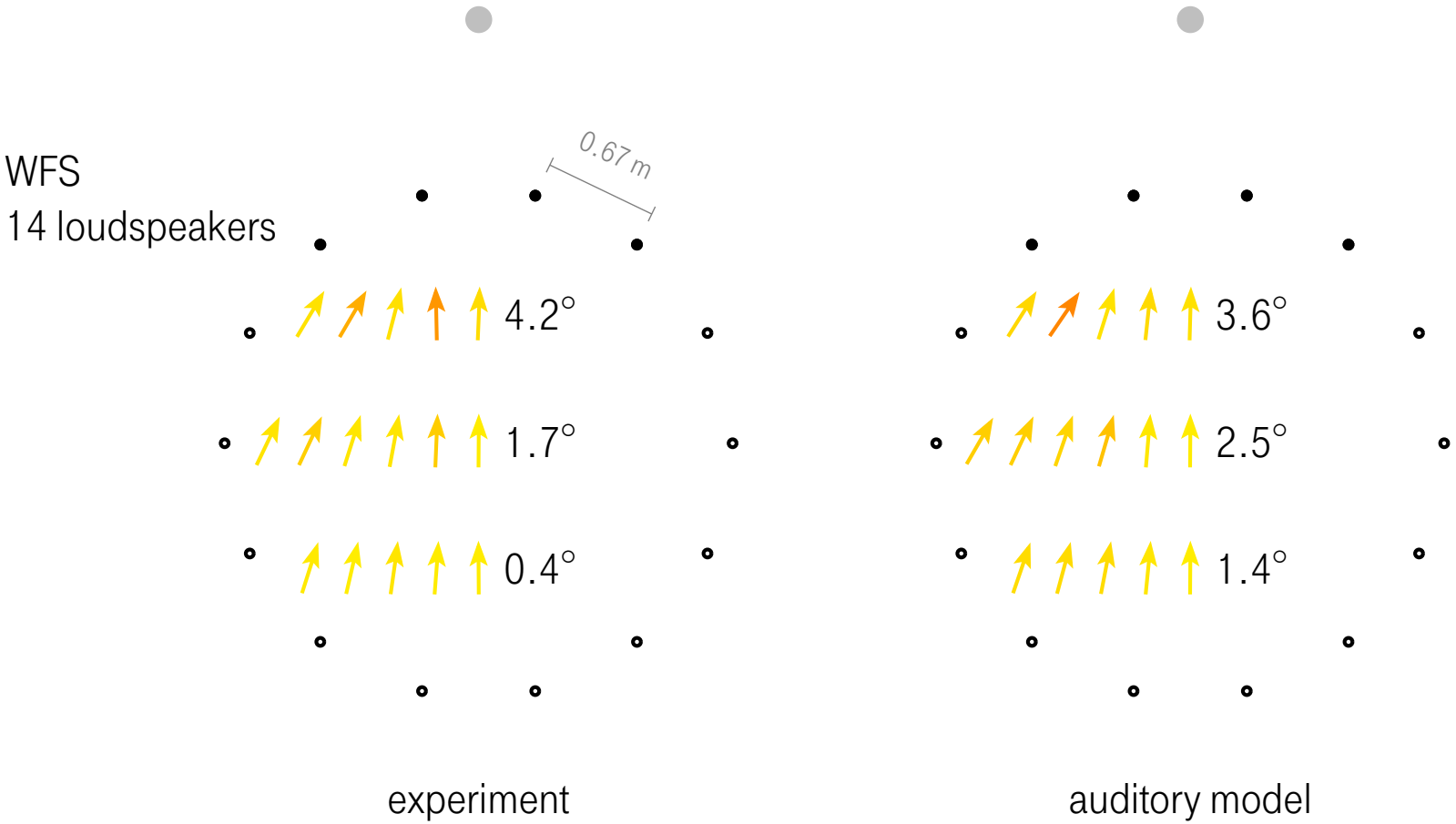


Dietz et al. (2011), Auditory model based direction estimation of concurrent speakers from binaural signals, Speech Communication

Wierstorf et al. (2013), Binaural Assessment of Multichannel Reproduction, in *The Technology of Binaural Listening*

Localization Accuracy

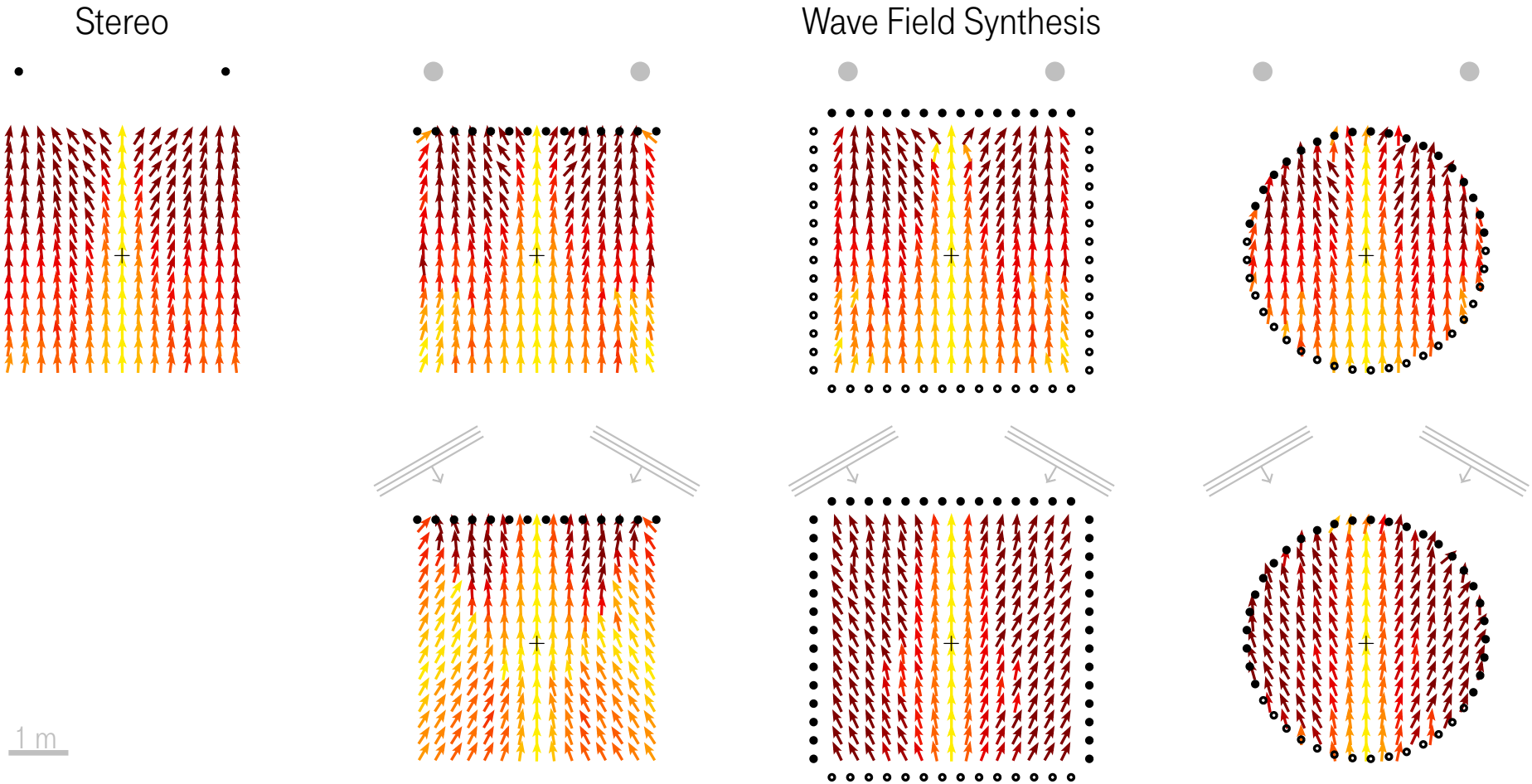
Wave Field Synthesis



Wierstorf et al. (2014), Wahrnehmung künstlich erzeugter Schallfelder, DAGA

Localization Accuracy

Stereo Downmix



Sweet-Spot

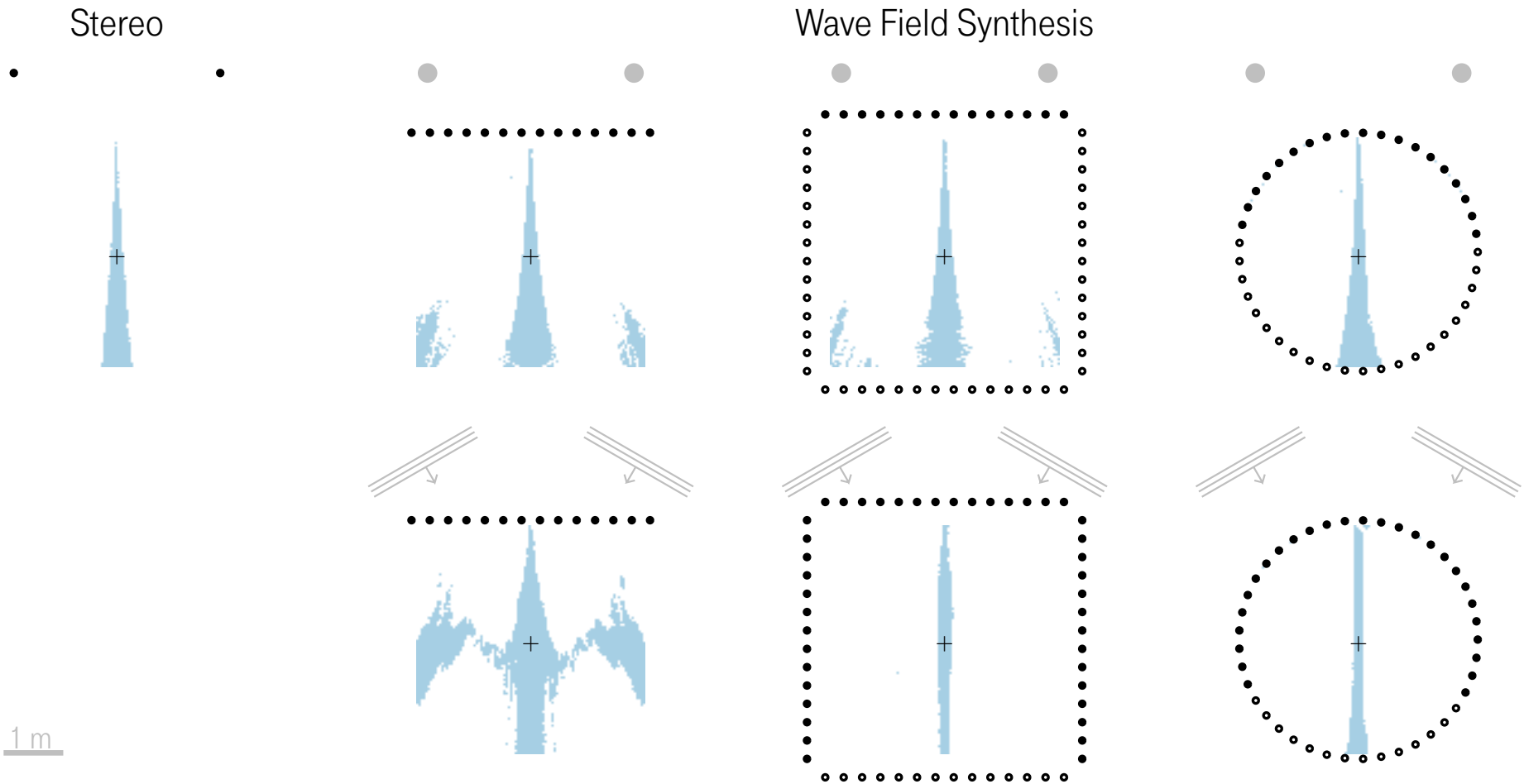
Definition

- normally not defined in literature
- area in which localization is *as intended*
- coloration not included
- not a point (e.g. Merchel 2010)
- used definition: localization error $< 5°$

Merchel and Groth (2010), Adaptively Adjusting the Stereophonic Sweet Spot to the Listener's Position, JAES

Sweet-Spot

Results



Conclusion

- virtual loudspeakers allow the object based representation of stereo
- point-sources: sweet-spot identical to real setup
- plane-wave: sweet-spot size depends on array setup
- inheritance of WFS problems like coloration and amplitude decay

Questions?

http://twoears.eu
http://spatialaudio.net
http://gnuplotting.org