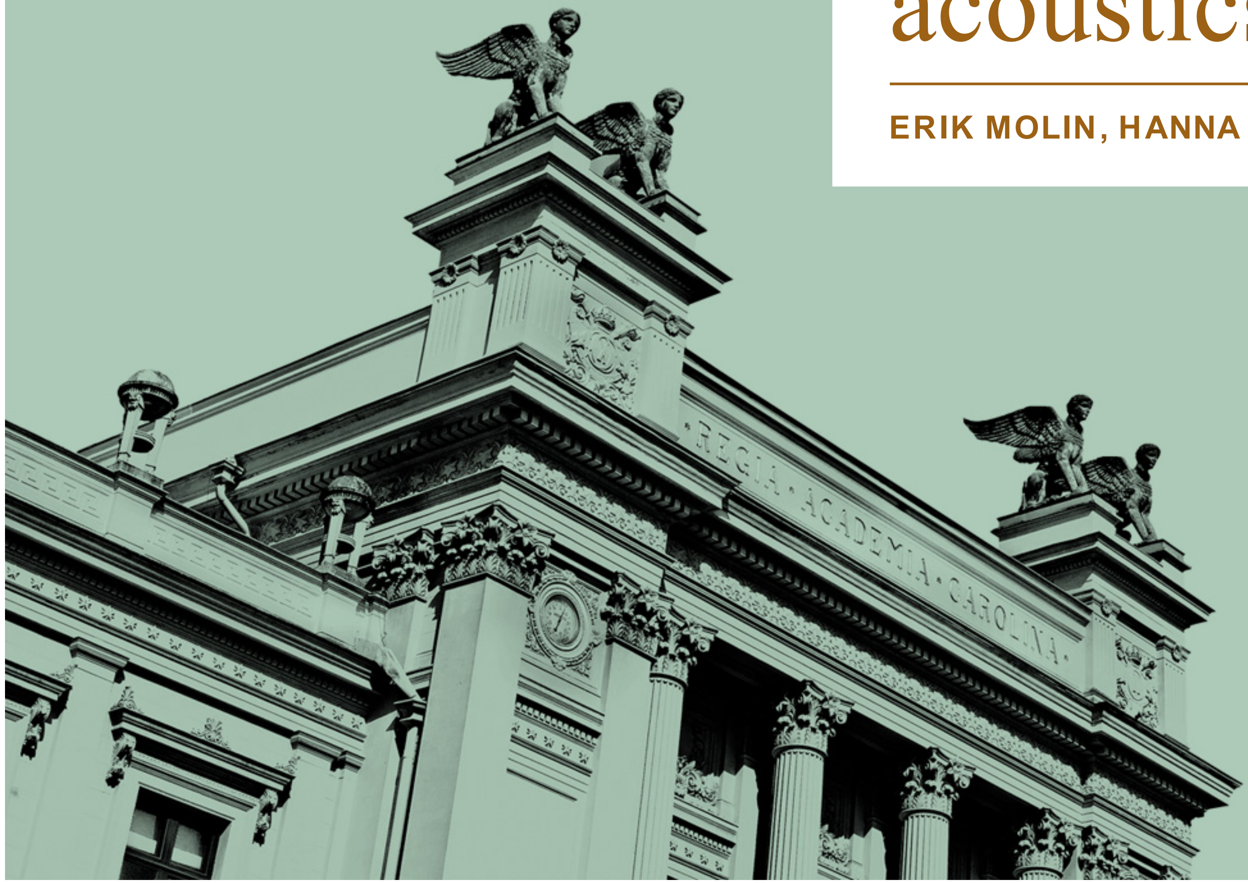
Auralization and Geometric acoustics

ERIK MOLIN, HANNA AUTIO



Auralization – what and why?

Auralization

From Wikipedia, the free encyclopedia

Auralization is a procedure designed to model and simulate the experience of acoustic phenomena rendered as a soundfield in a virtualized space. This is useful in configuring the soundscape of architectural structures, concert venues, public-spaces and in making coherent sound environments within virtual immersion systems.

- For a given acoustic situation (space, sound source(s), listener position…), what sound does the listener hear?

Auralization – what and why?

Virtual spaces –
Acoustic Design

Auralization – what and why?

Virtual spaces –
Entertainment

LUND UNIVERSITY

Overview

- Intro to Auralization
- Audio sources
- Something about systems
- Room models
  - Geometric acoustics
- Receiver modelling

Sound propagation – system model

Reality

Source | Room | Ears | Perception

LUND UNIVERSITY

Sound propagation – system model

Reality

Source	Room	Ears	Perception

Auralization

Source	RIR	HRTF	Perception

Sound propagation – system model

Reality

Source	Room	Ears	Perception

Auralization

Source	RIR	HRTF	Perception

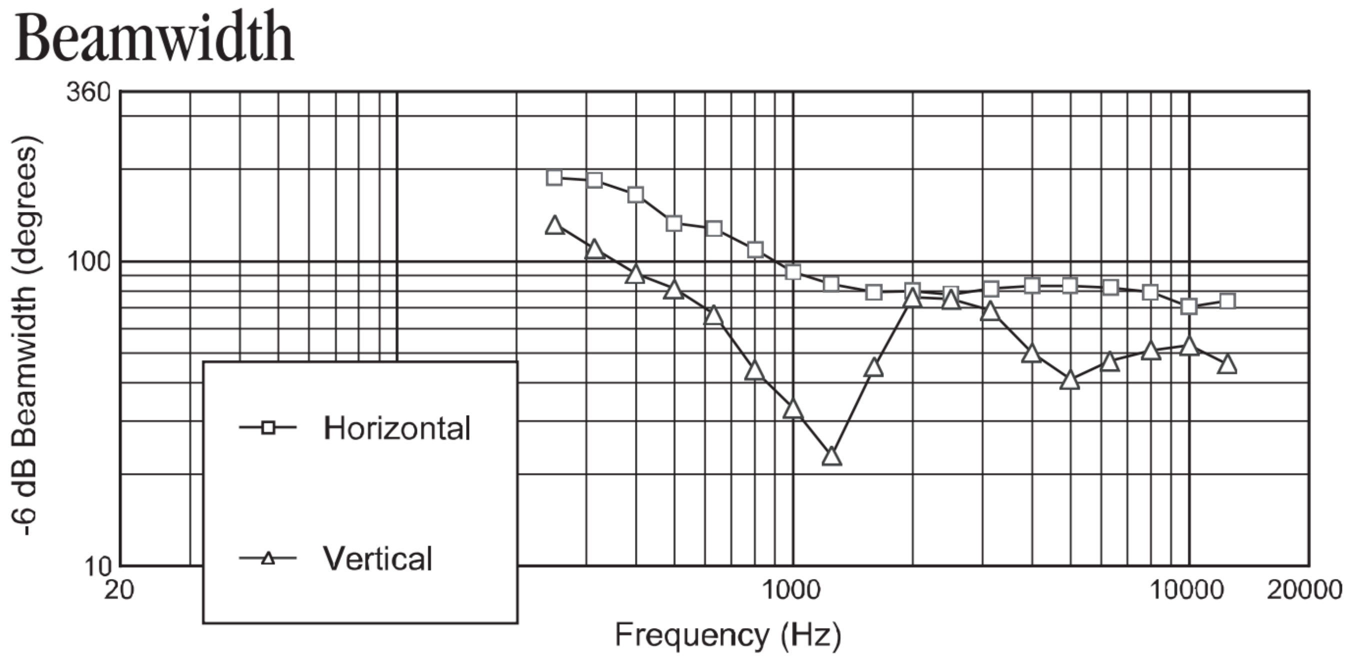
Sound sources

- What they actually sound like
  - It turns out it doesn't really matter (as we shall see later)
- Directivity d(ϕ,θ)
- Trumpet? Reasonably independent of theta
  - Generally nice with symmetry

LUND UNIVERSITY

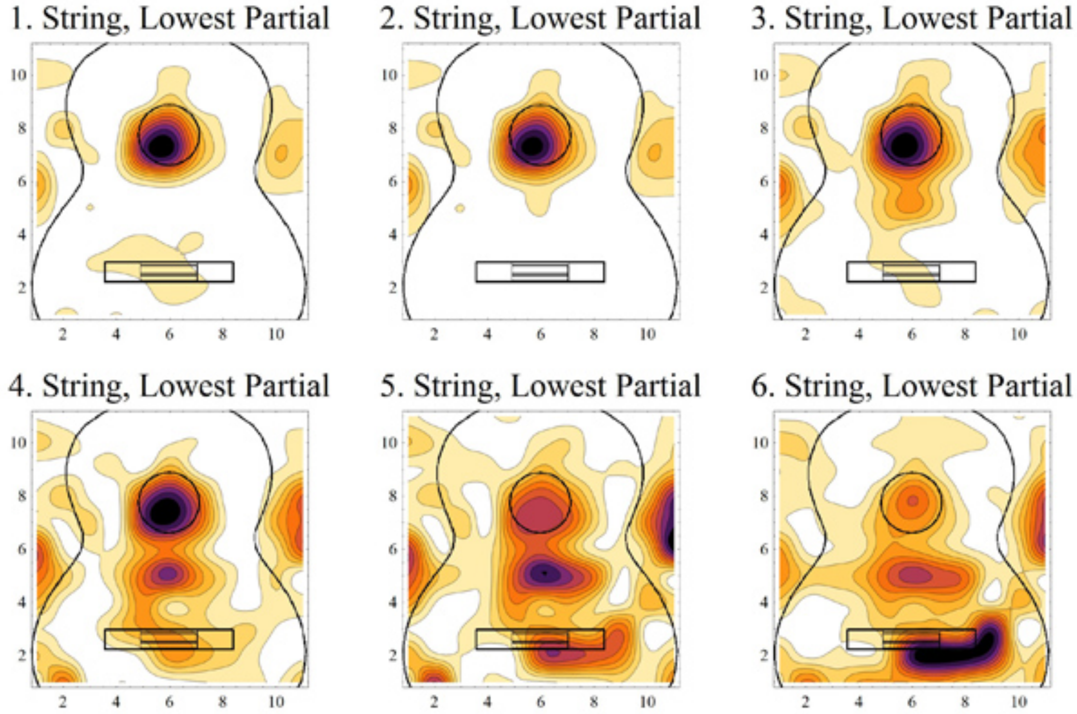
Sound sources – directivity

- Directivity is frequency dependent
- Directivity d(ϕ,θ, f)
- Loudspeaker:



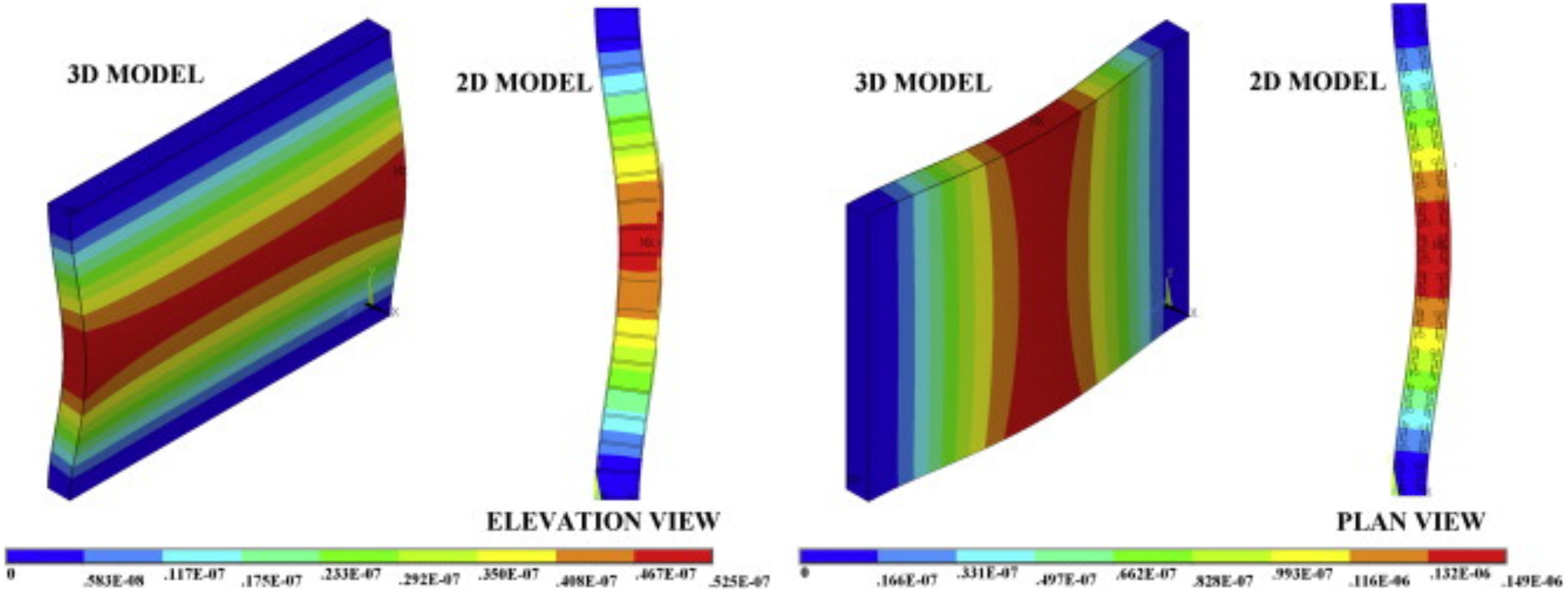
Sound sources – spatial extension

- The trumpet can be modeled as a point – but what about a guitar?



Sound sources – spatial extension

- What about a wall?



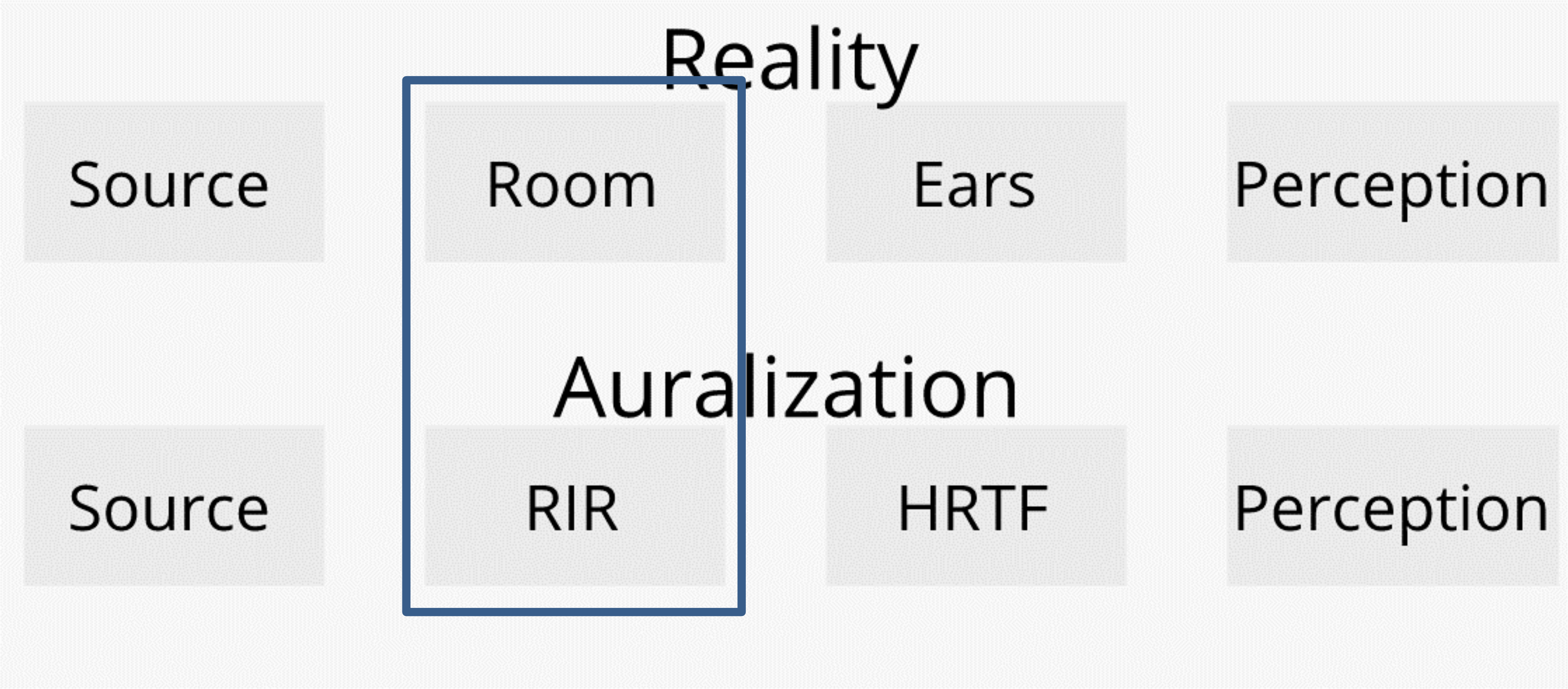
LUND UNIVERSITY

Sound sources – conclusion

- Hard problem!
  - Current solutions mainly directed point sources
  - No current software accurately models transmission

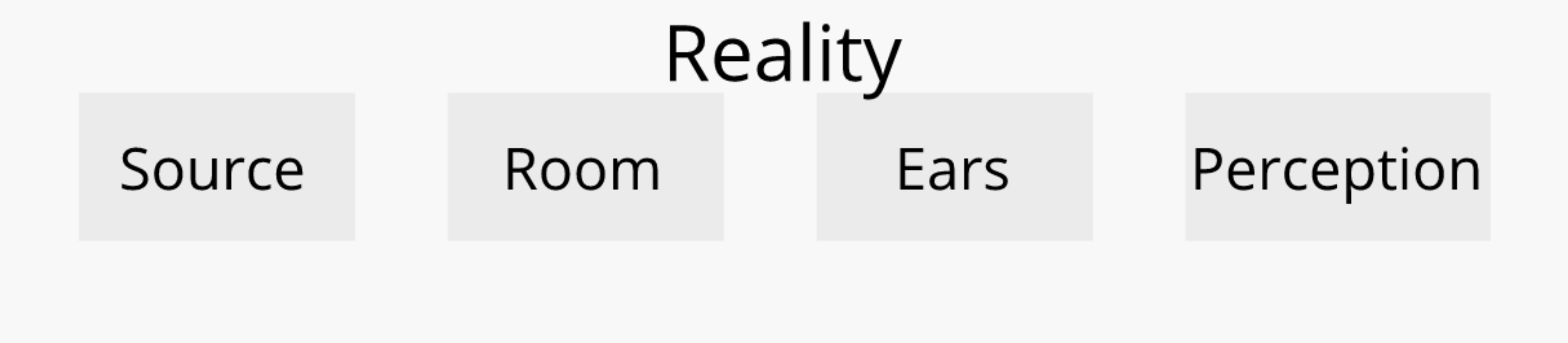
LUND UNIVERSITY

Sound propagation – system model



Background - Filters

- Input → filter → output

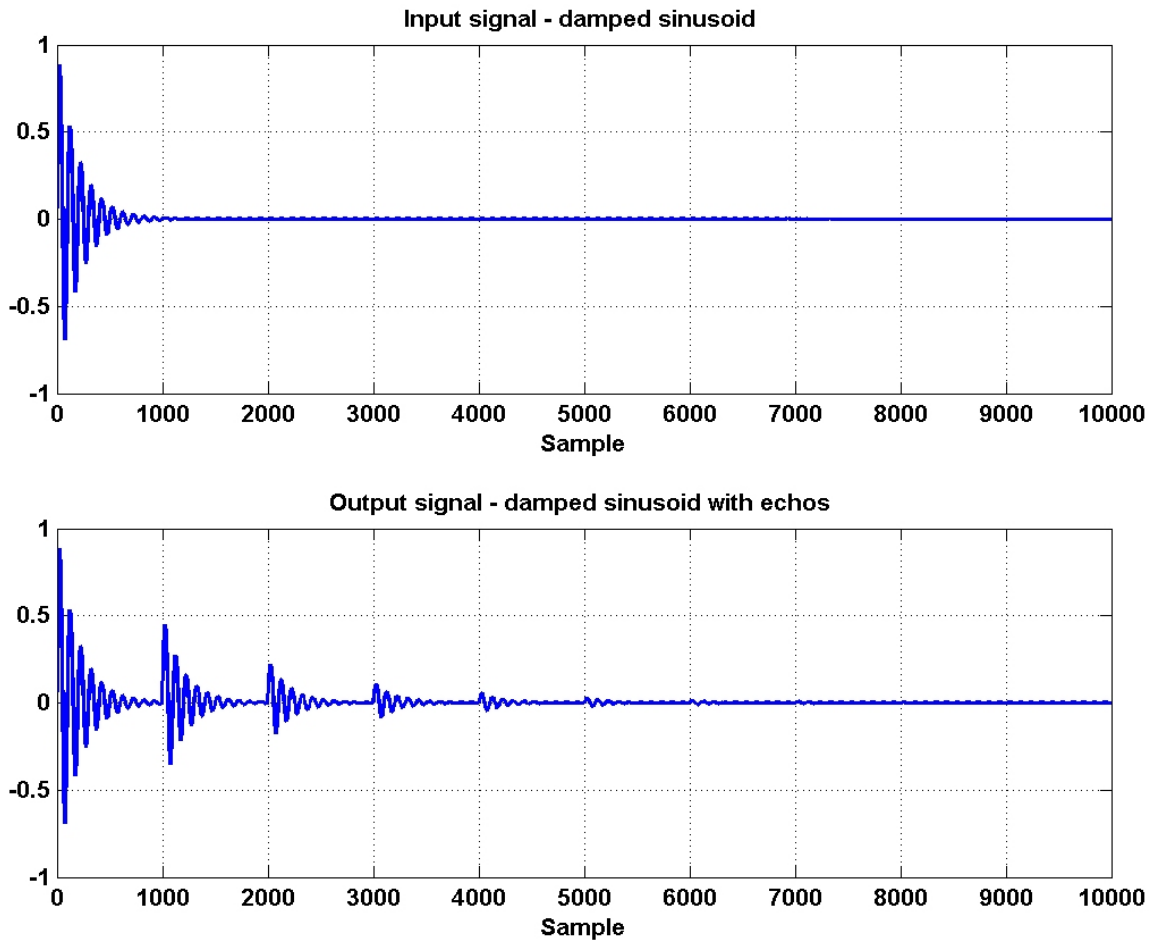


Background - Filters

- Equaliser:
  – Perceptively - Equaliser changes frequency content.

Background - Filters

- Echo:
  - Perceptively – echo changes time content.

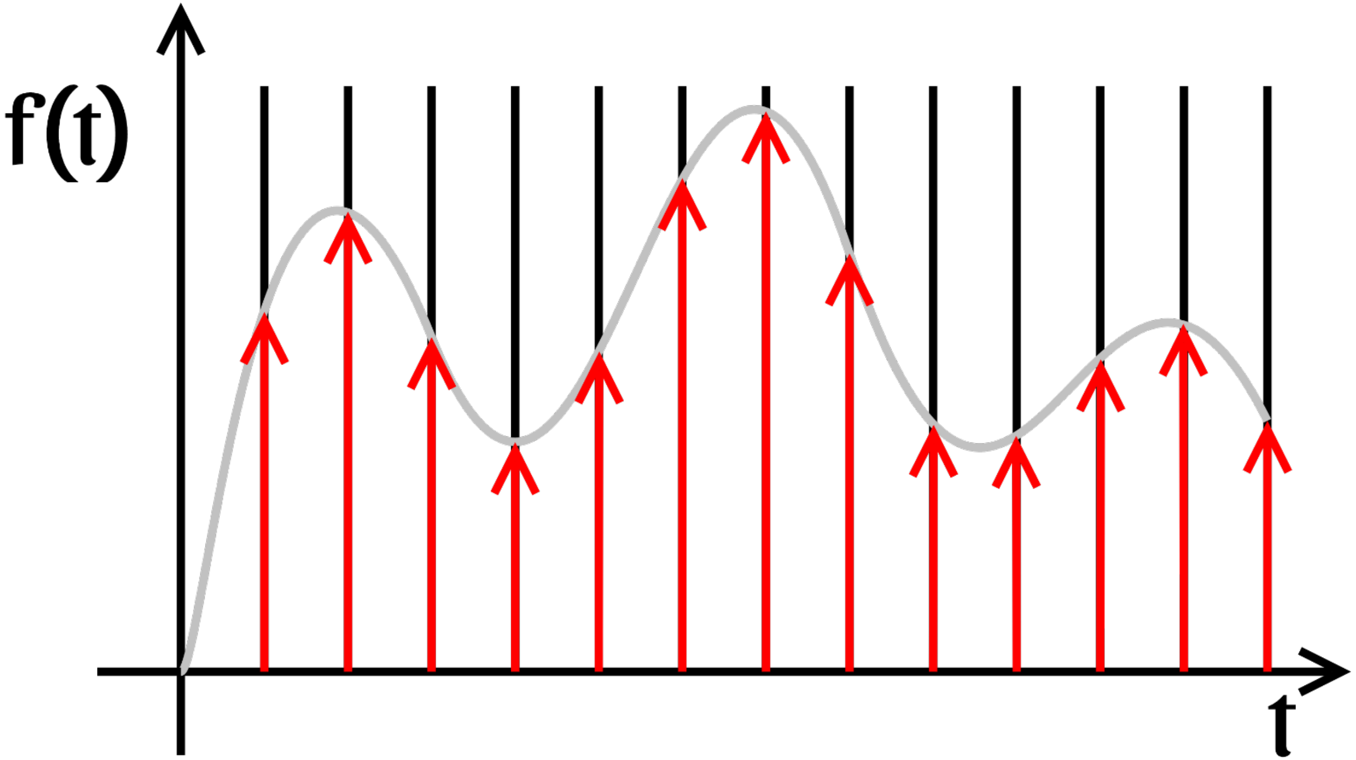


Background - Filters

- Fourier analysis → changes reflected in both time and frequency.
- Fourier transform FT
- Signal(t) → (FT) → signal(f) → FT^-1 → Signal (t)
  - Time and frequency domains are interchangable

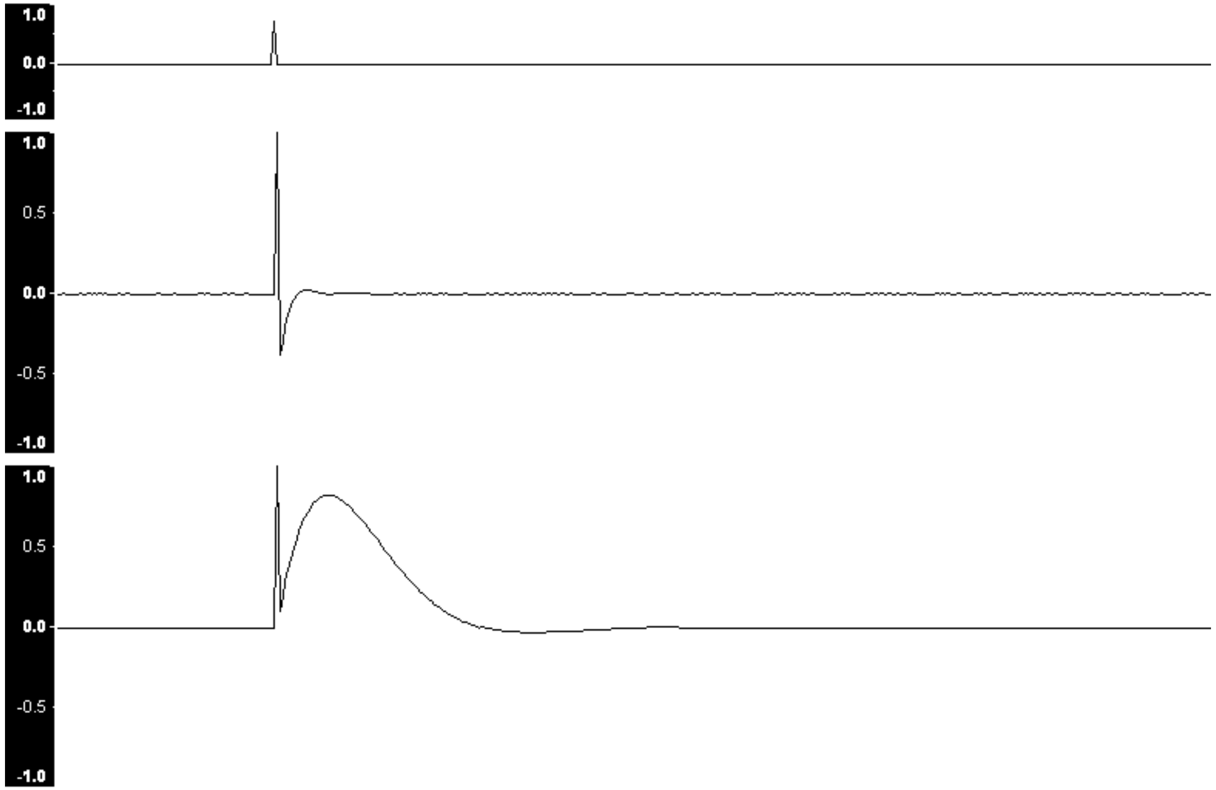
Signals

- Analog signal: continuous, physical: voltage, pressure, …
- Digital signal: sequence of numbers x_n



Impulse responses

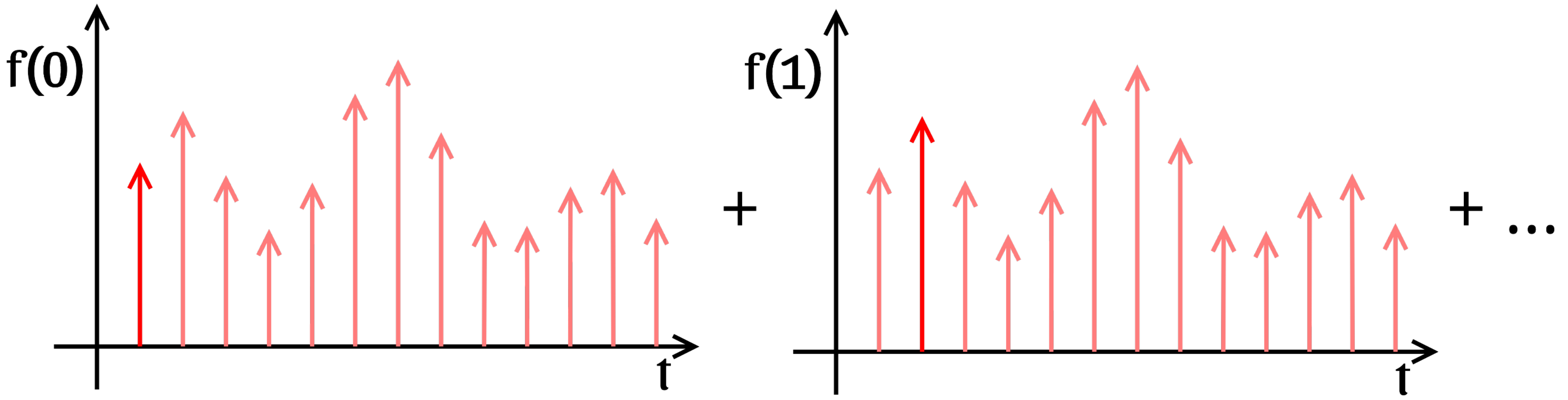
- δ(t) = 1 if t = 0, 0 otherwise.
- Impulse response I
  of filter/system S: I = S(δ)



LUND UNIVERSITY

Impulse responses

- Why is this important?
  - All signals can be expressed as sum of $n_a * \delta(x - a)$!



Impulse responses

- Why is this important?
  - All signals can be expressed as sum of $\delta(x - a)$

  If system linear, all outputs can be expressed as sum of I

LUND UNIVERSITY

Impulse responses

- Why is this important?
  - All signals can be expressed as sum of δ(x - a)

  If system linear, all outputs can be expressed as sum of I

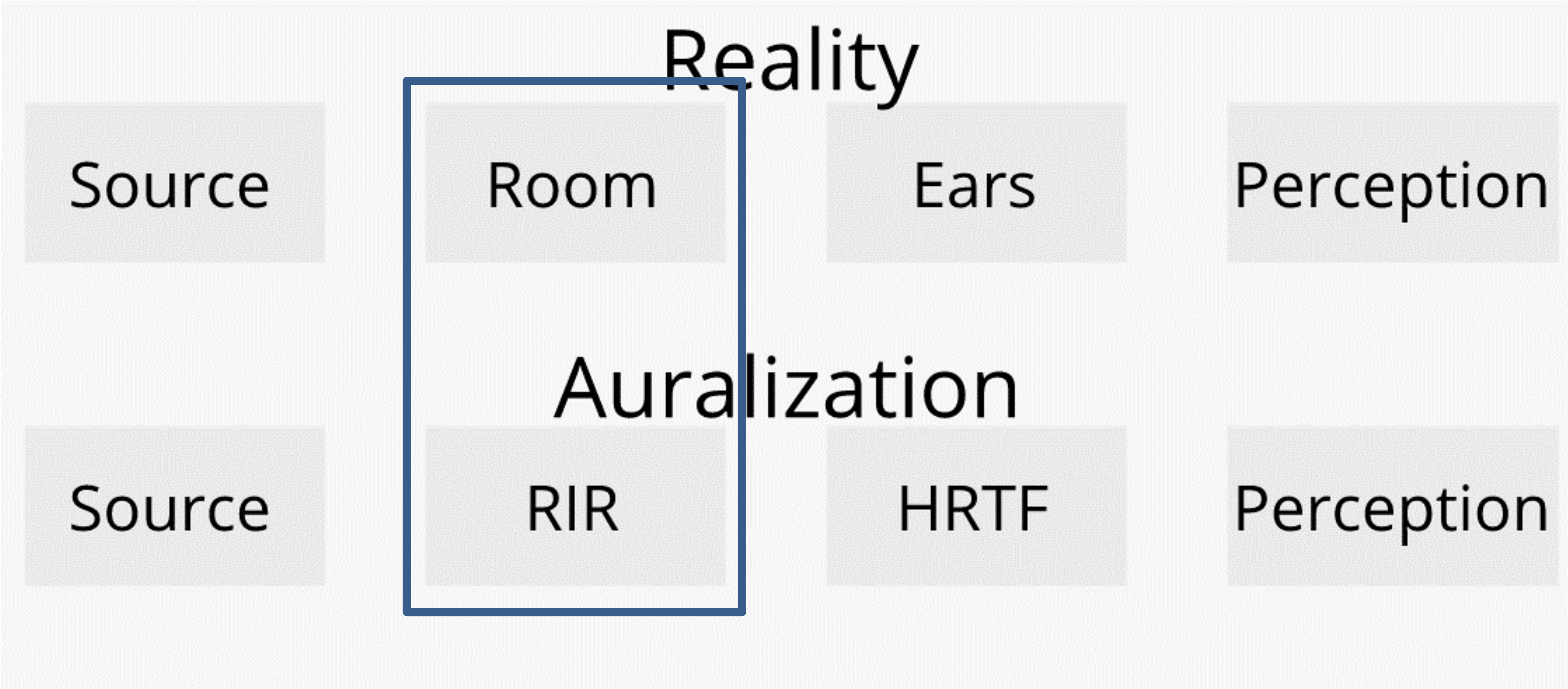
  This lets us run simulation with δ input, and use this result to calculate result for any input

Linear systems overview

		Input signal	LTI system	Output signal
A			Response function	
B	**Time domain:**	$u(t)$	$h(t)$	$y(t)$
C	Linearity:	$u_1(t)$		$y_1(t)$
		$u_2(t)$		$y_2(t)$
		$au_1(t)+bu_2(t)$		$ay_1(t)+by_2(t)$
D	Time invariance:	$u(t-T)$		$y(t-T)$
E	Input–output relation:		$\int u(\tau)h(t-\tau)\cdot d\tau = y(t)$	
F	**Frequency domain:**	$U(\omega)$	$H(\omega)$	$Y(\omega)$
G	Input–output relation:		$U(\omega)\times H(\omega)=Y(\omega)$	

LUND UNIVERSITY

Sound propagation – system model



LUND UNIVERSITY

Room Impulse Response (RIR)

- Can be measured (recall reverberation time measurements)

Room Impulse Response (RIR)

- Can be measured (recall reverberation time measurements)
- Can also be simulated:
  - RIR(x_r, x_s, G)
    - » X_r – receiver position
    - » X_s – source position (+directivity)
    - » G – geometry of the room – including acoustic data

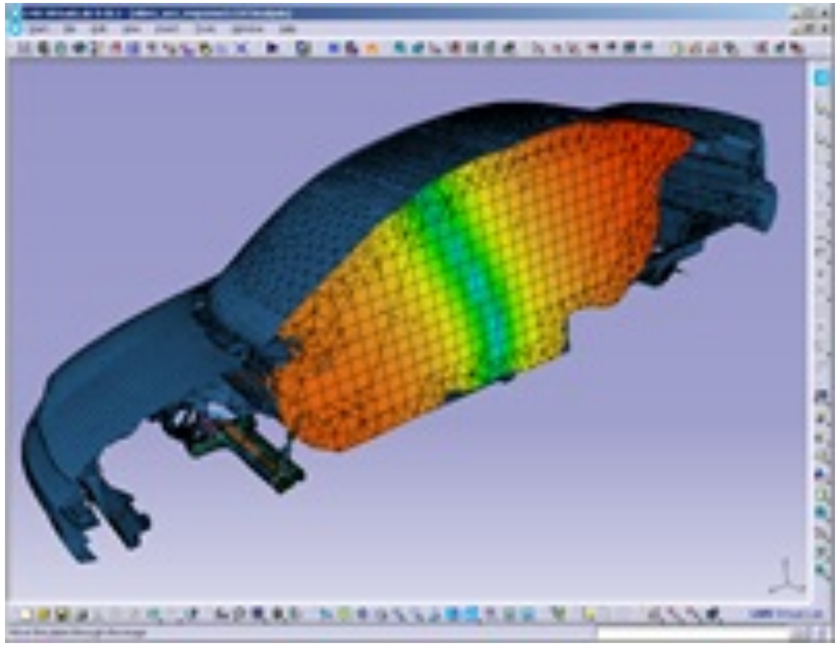
RIR – acoustic parameters

- Some measure of absorption
- Some measure of reflection
- More?

LUND UNIVERSITY

How to create the RIR?

- Insane method: Full pressure field simulation from 20Hz to 20kHz – lots of time and/or money.
  - Sampling theorem: mesh must match frequency.



How to create the RIR?

- Insane method: Full pressure field simulation from 20Hz to 20kHz – lots of time and/or money.
- Less insane: use what we know!
  - Wave phenomena only very important below schroeder frequency!
  - Wave models below that, simpler models above
  - Combine somehow…

Geometric acoustics

- Has been used since ~1990. (when computers became powerful enough to do it)
- Models sound as rays – no wave phenomena

Image source method

- Make a mirror image in each surface

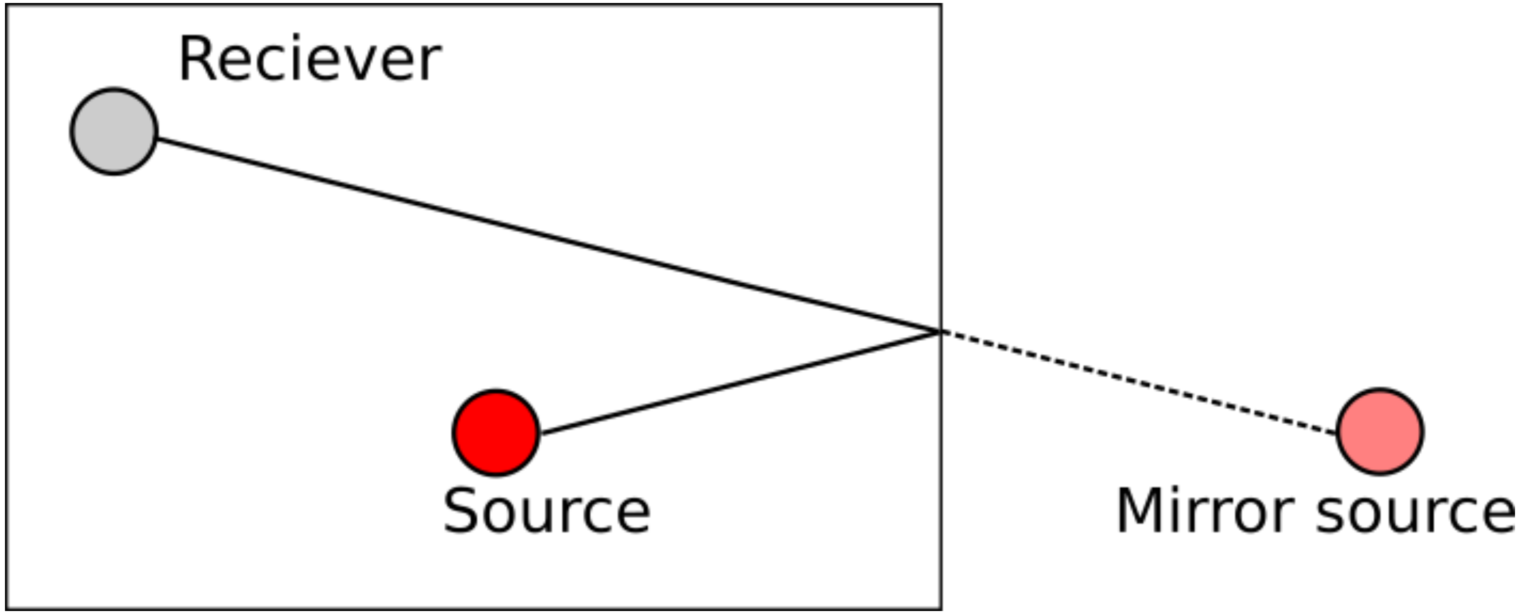
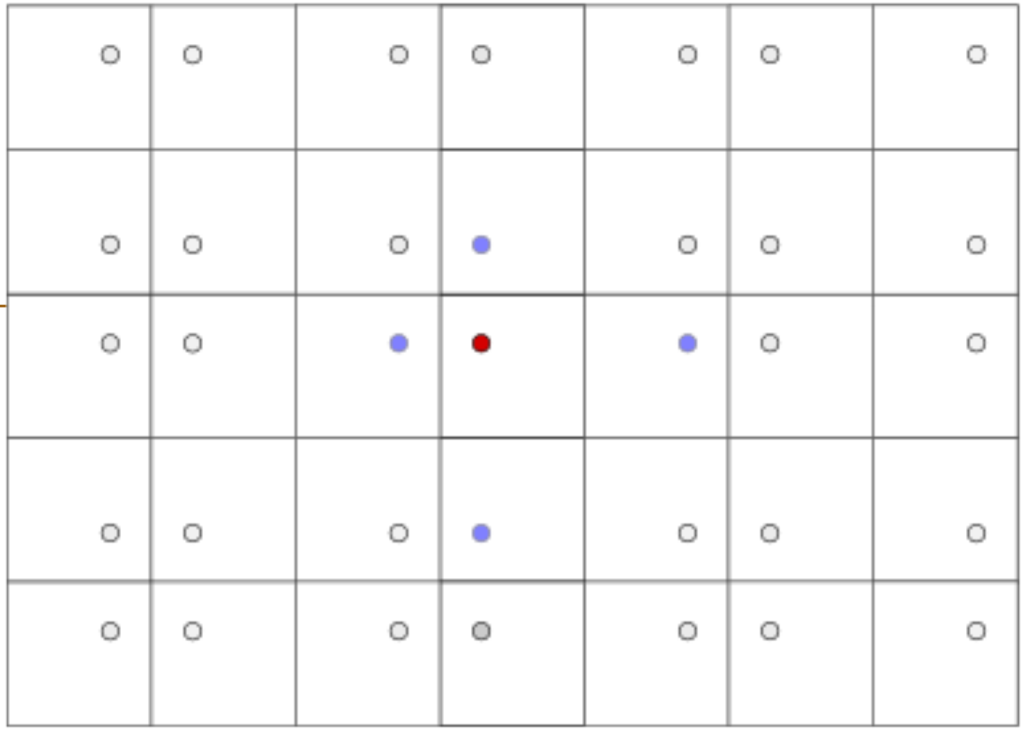


Image source method

- Store distance, amplitude of each mirror source
- Energy rather than pressure



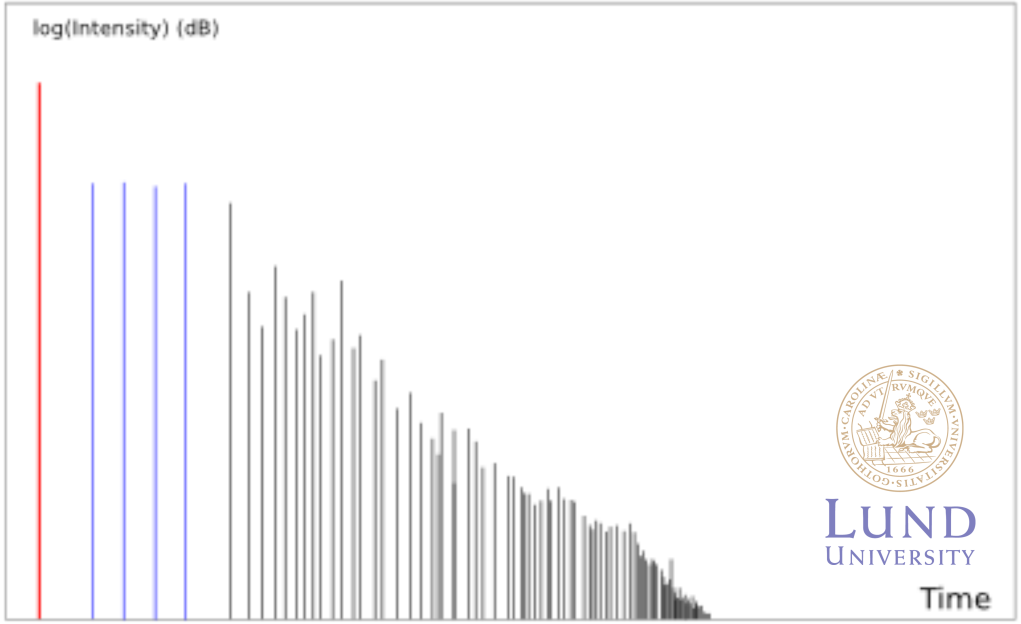


Image source method

- Non-convex rooms need visibility lookup

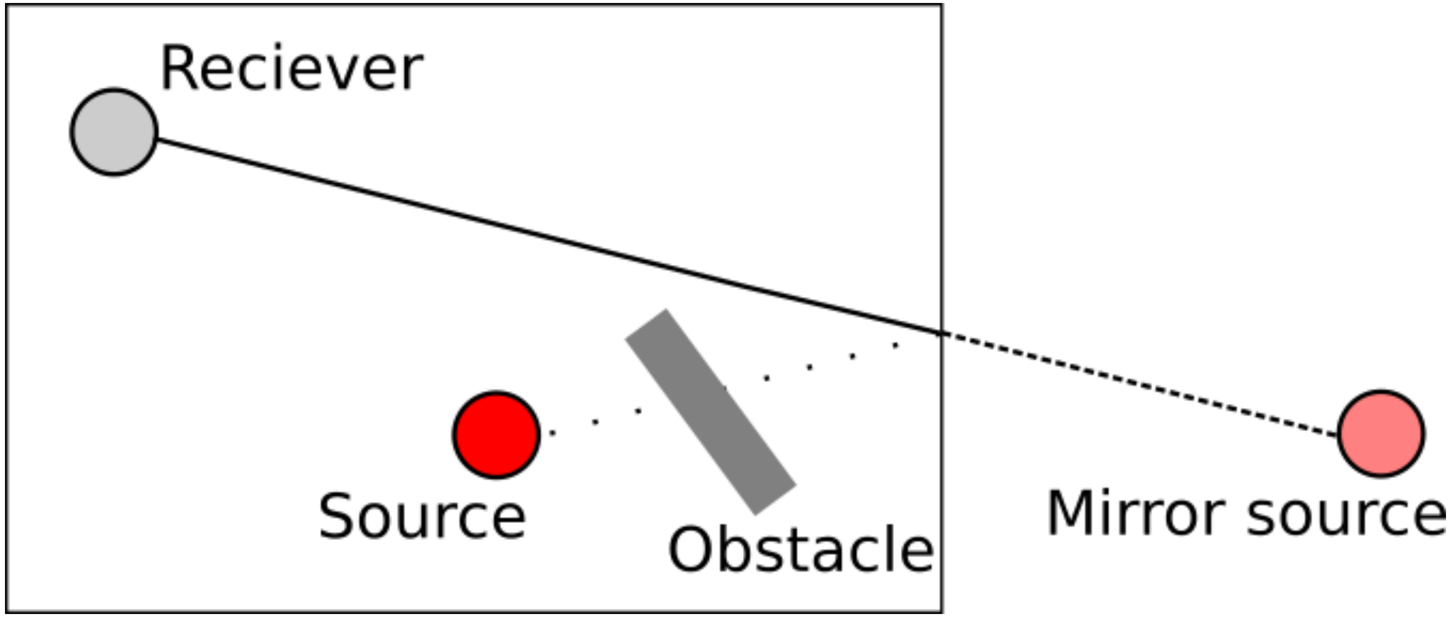


Image source method

- Exact solution for flat hard surfaces, convex rooms

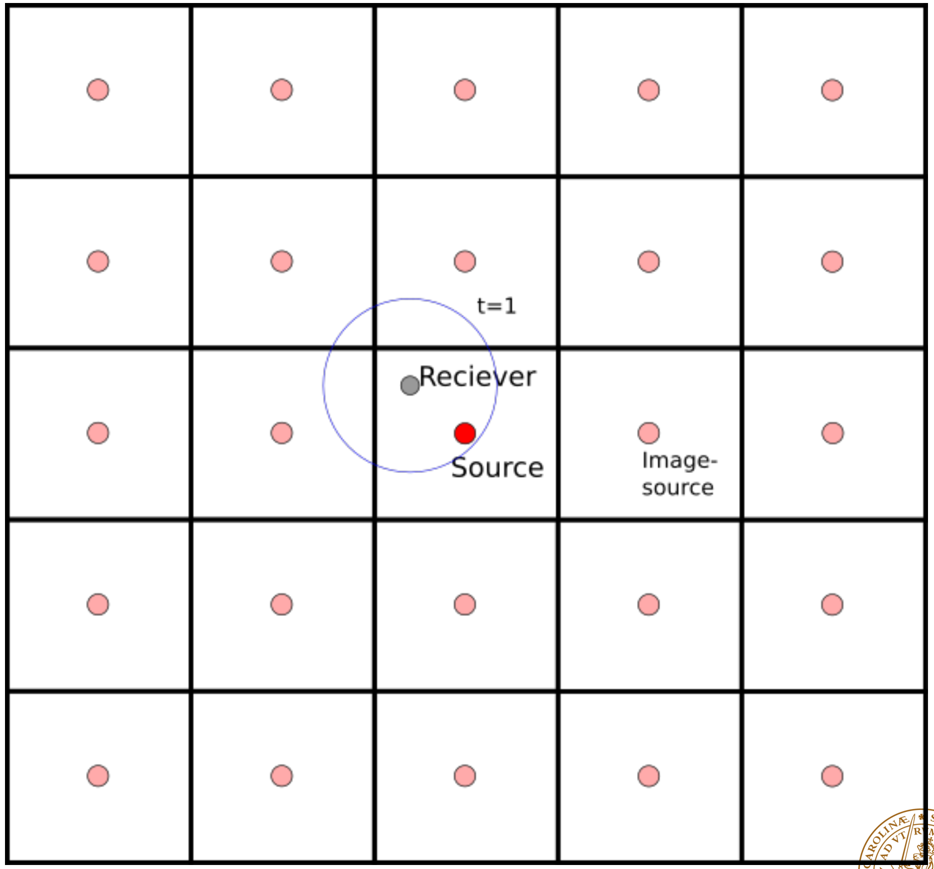
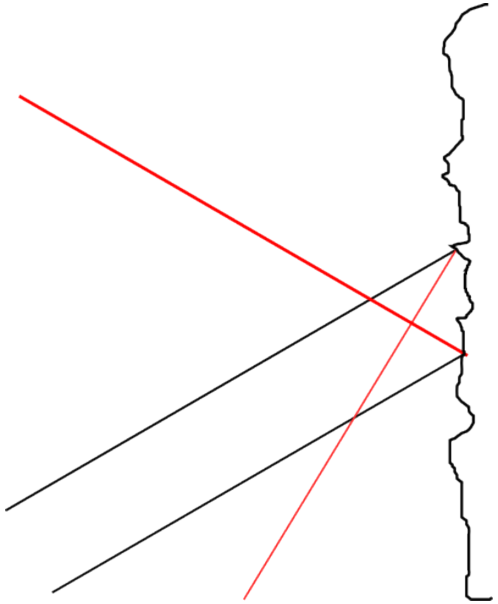


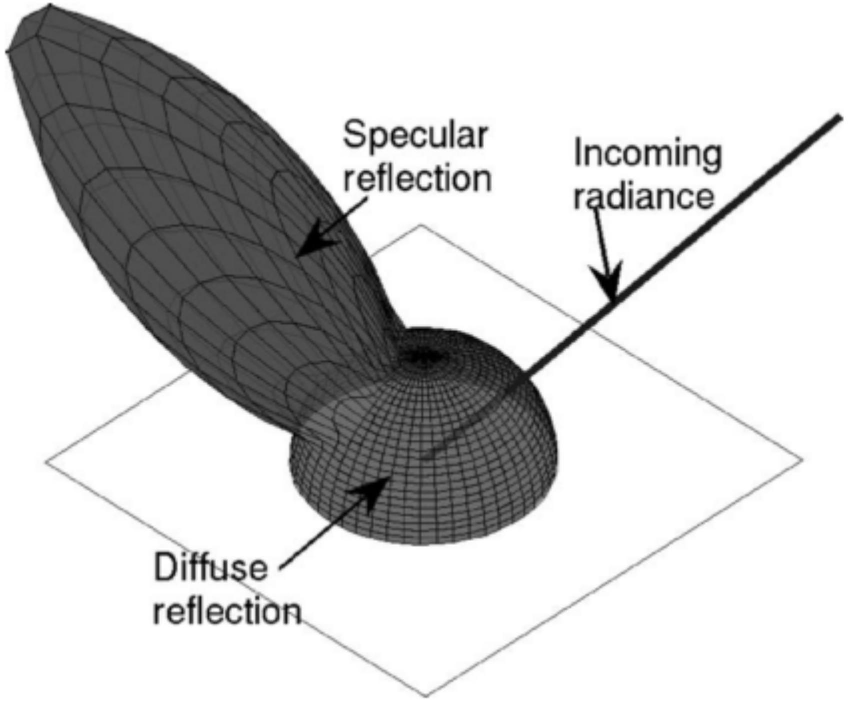
Image source method

- ...but real rooms have soft & scattering surfaces



Reflection modelling: BRDF

- Bidirectional Reflectance Distribution Function

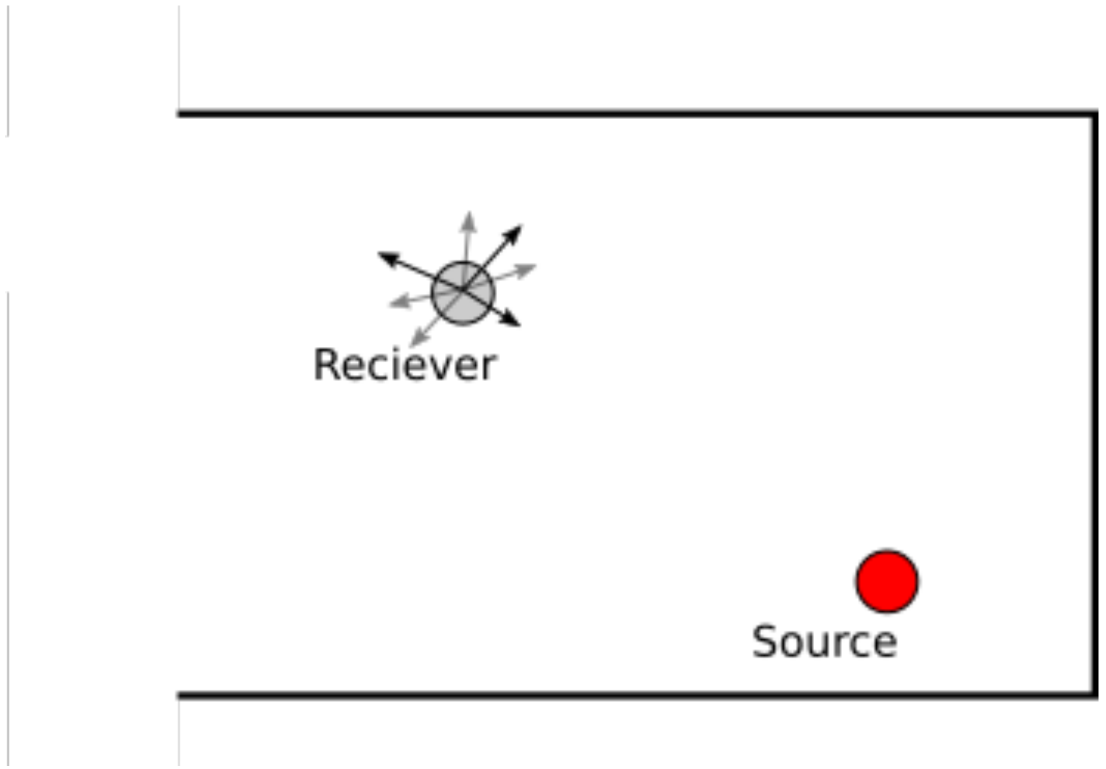


Stochastic Ray tracing

- Computationally heavier than IS – but scales better
- Can handle scattering
- Not exact.

Stochastic Ray tracing

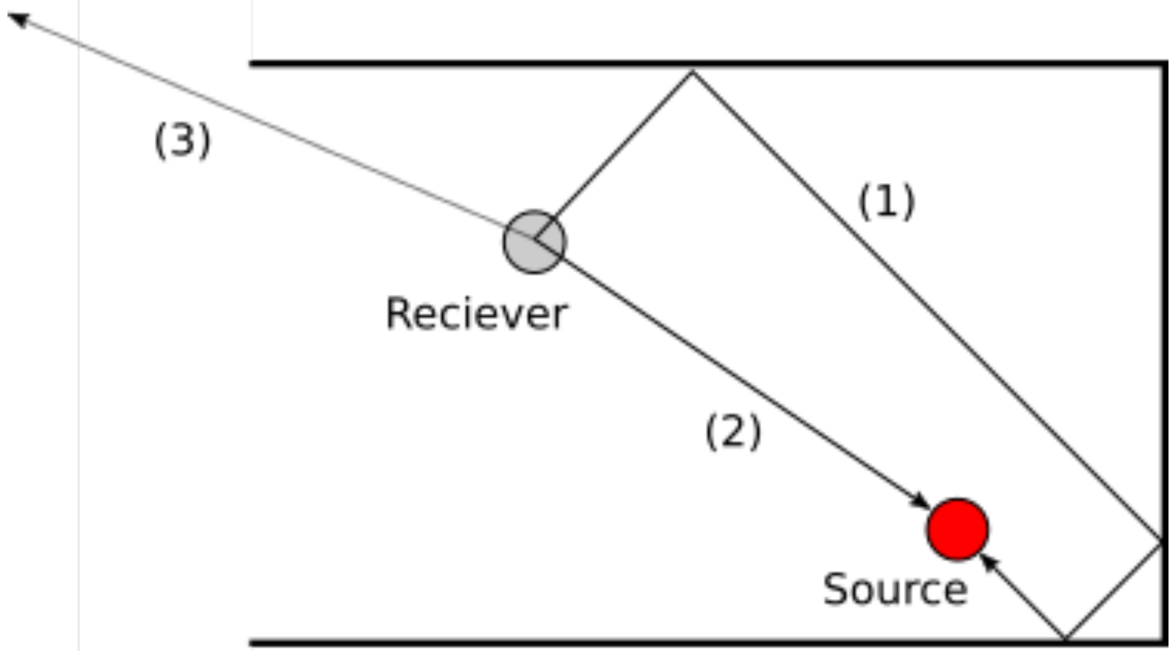
- Fire rays in a spherical distribution, and see what they hit.



Stochastic Ray tracing

- Fire rays in a spherical distribution, and see what they hit.

- (1) reflected sound
- (2) direct sound
- (3) miss – wasted computation

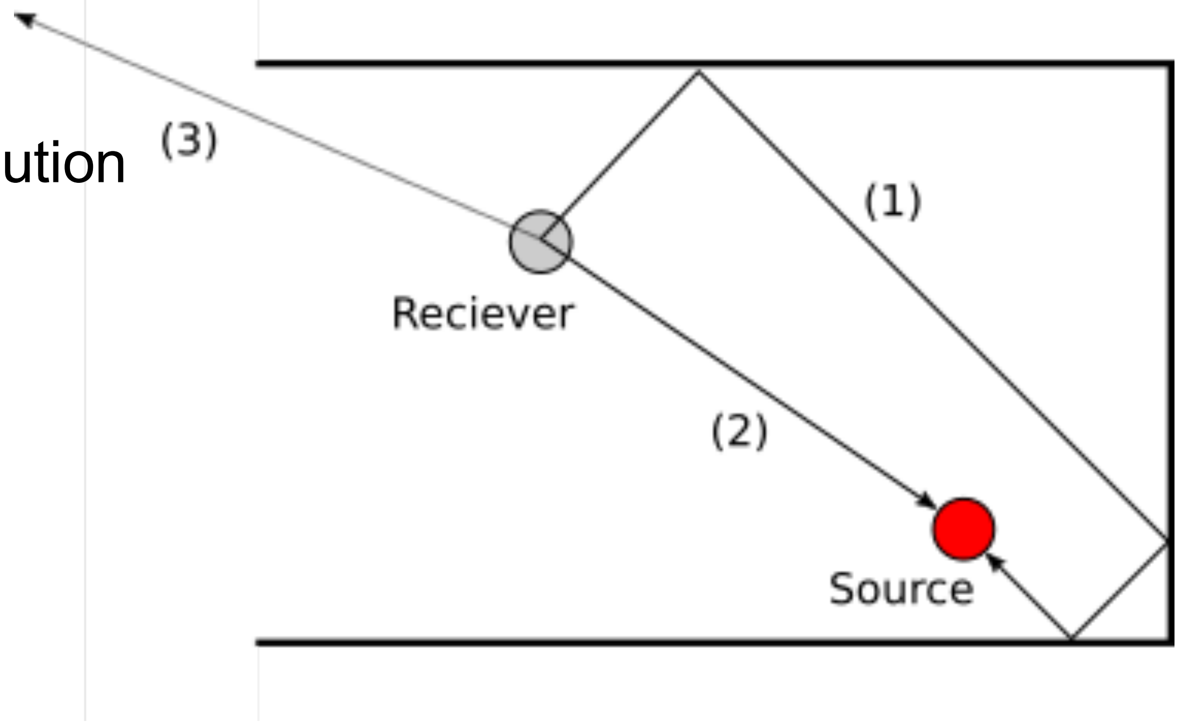


LUND UNIVERSITY

Stochastic Ray tracing

- Fire rays in a spherical distribution, and see what they hit.

- (1) reflected sound
- Uses BRDF as probability distribution for reflected ray



Acoustic radiance transfer

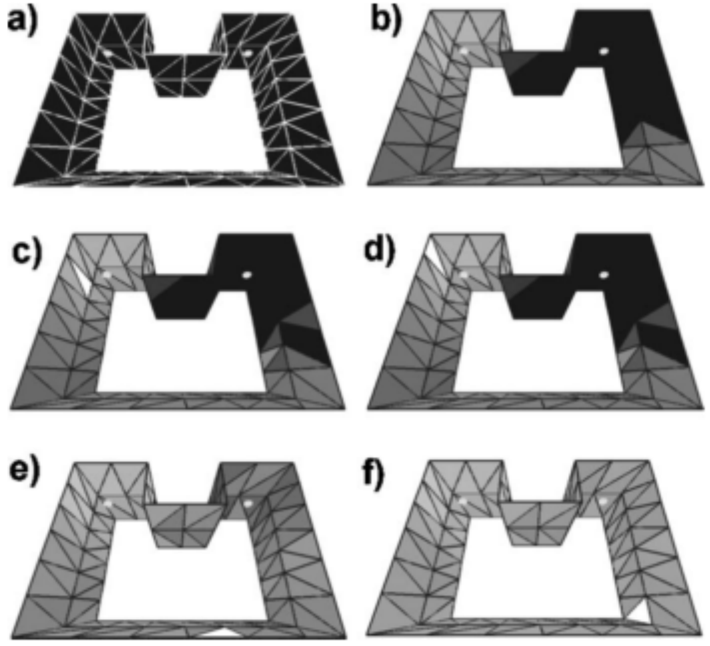
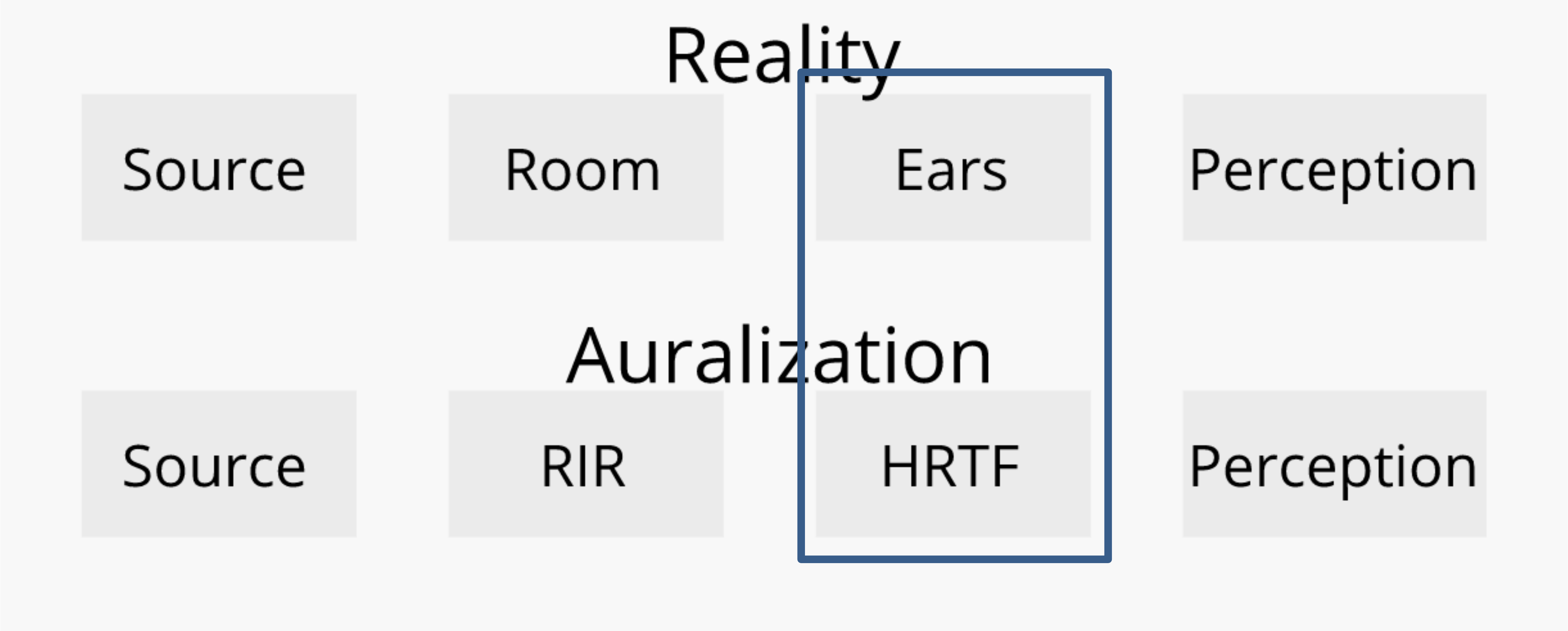


FIG. 4. The progressive radiance transfer solution. (a) The model without floor and ceiling before the initial shooting, where the source is the light dot on the left and the receiver is the dot on the right. (b) The model after the initial shooting, where the lightness of the patch is proportional to the amount of energy it has received. (c) The patch with the highest energy is emphasized, and the energy reflected from it has been added to the patches visible to that patch. This patch now no longer has the highest energy. (d) The next patch with the highest unshot energy is shown, and the energy is propagated similarly to (d). (e) The situation several steps later, and (f) the energy distribution in the model when the solution has converged. The energy from patches visible to the receiver is collected into it in the final gathering phase.

Conclusion – modelling techniques

- Full wave simulations are still computationally expensive
- Several approximate methods exist – each with strengths and drawbacks
- Current research: ray tracing, wave simulation, radiosity.
  - Goal: interactive auralizations for VR

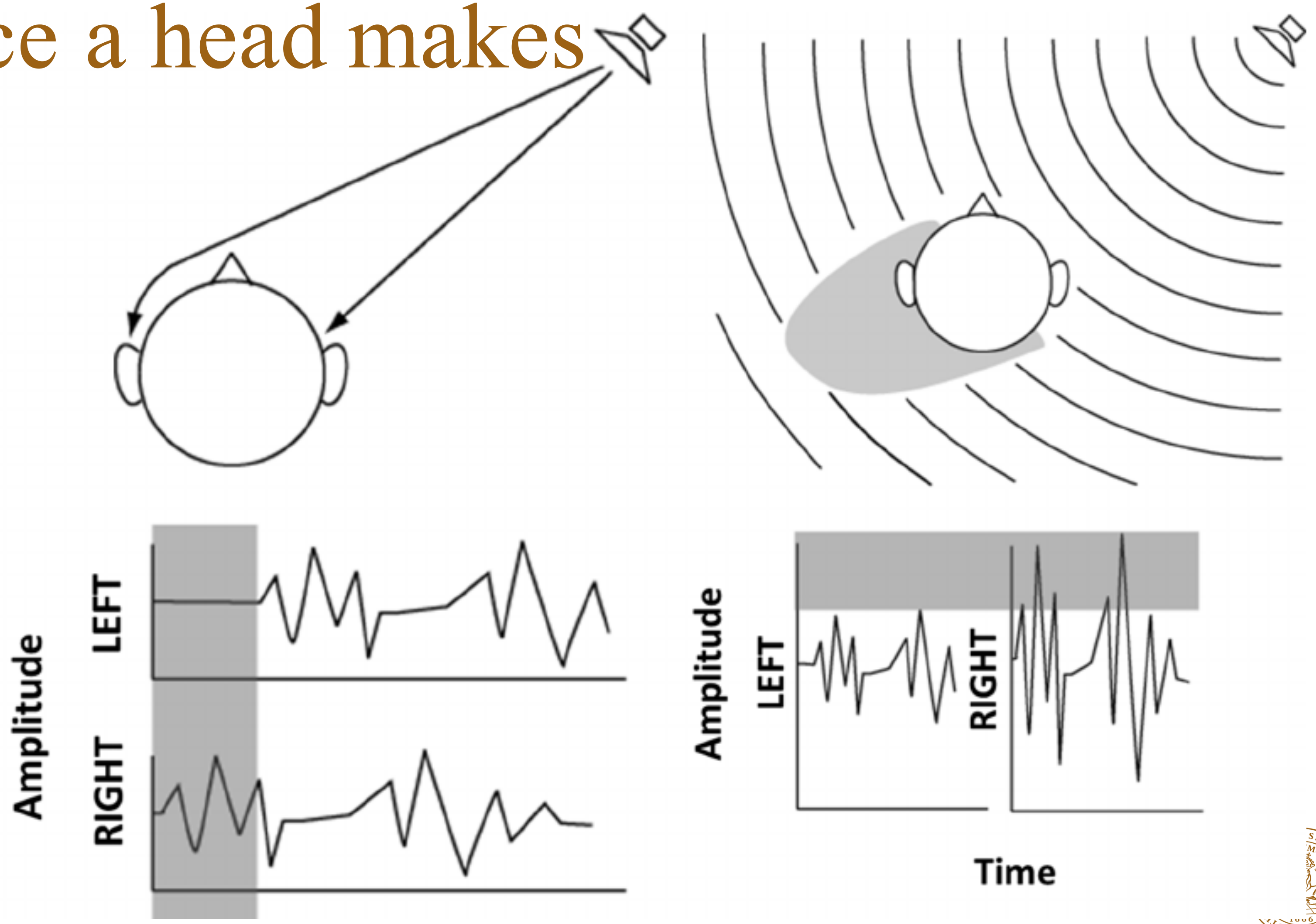
Sound propagation – system model



What a difference a head makes

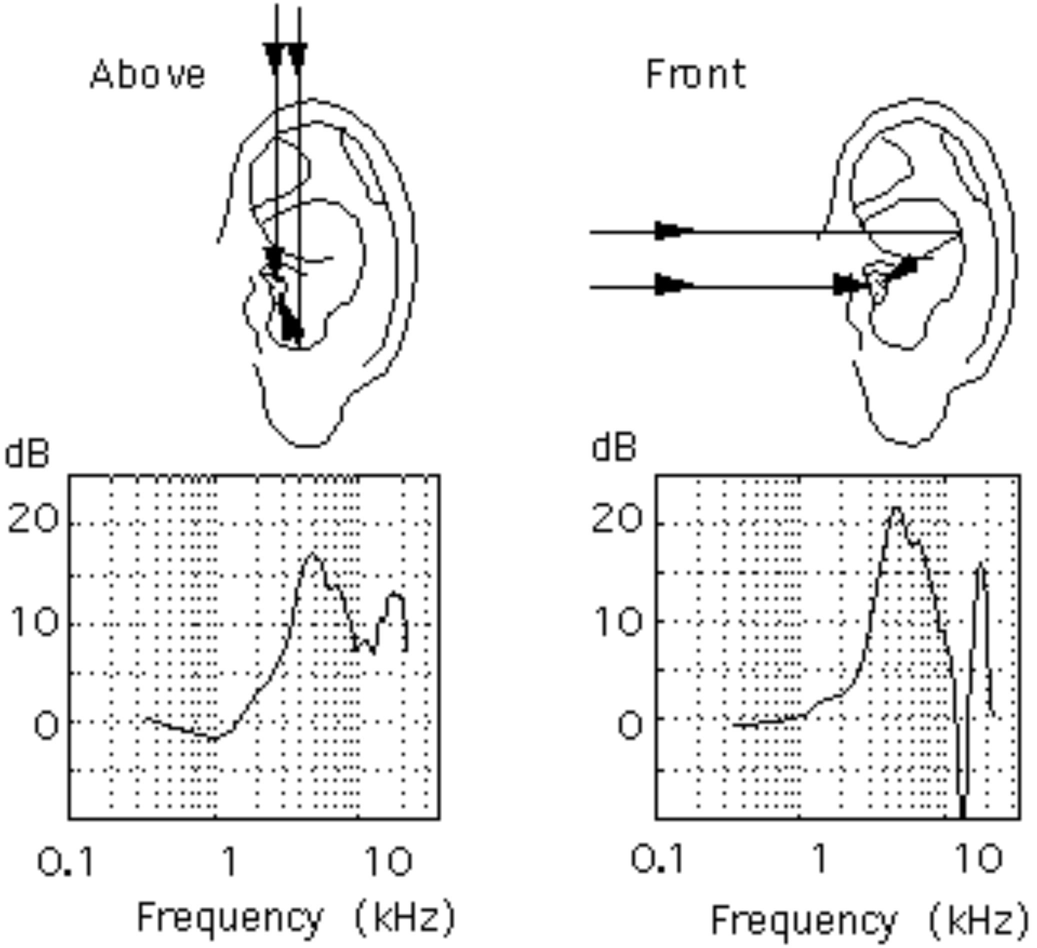
Interaural Time Difference,

Interaural Level Difference



What a difference an ear makes

Elevation dependent filter



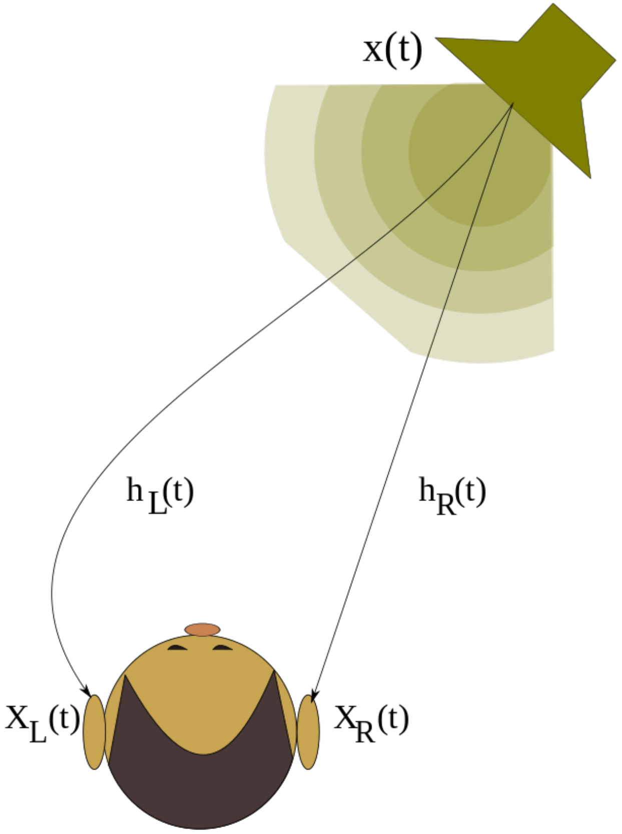
LUND UNIVERSITY

Directional audio (DTU Copenhagen)

LUND UNIVERSITY

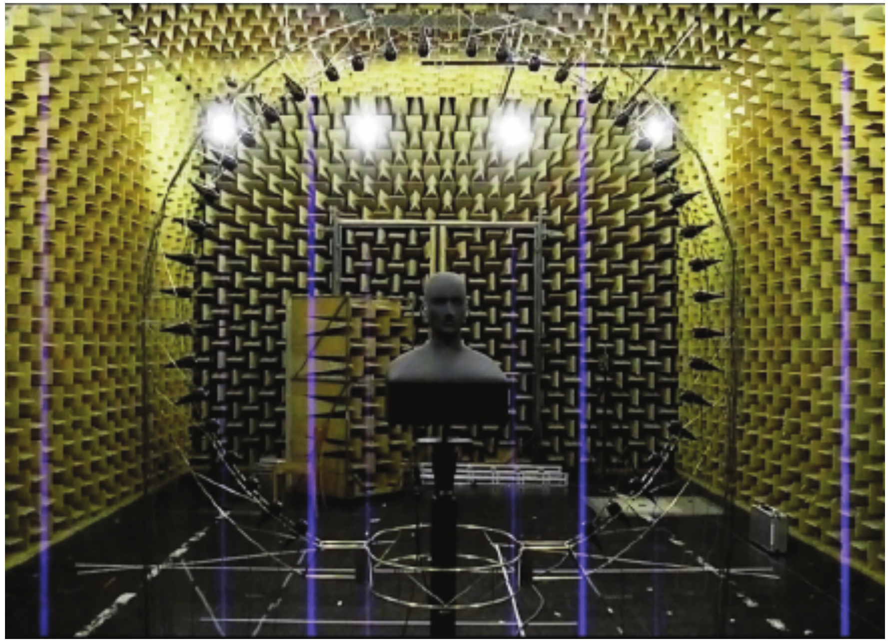
Head related impulse response - HRIR
Head related transfer function - HRTF

- HRTF = FFT(HRIR) – terminology difference
- Captures the effect of the head and torso on sound



How do we obtain the HRTF?

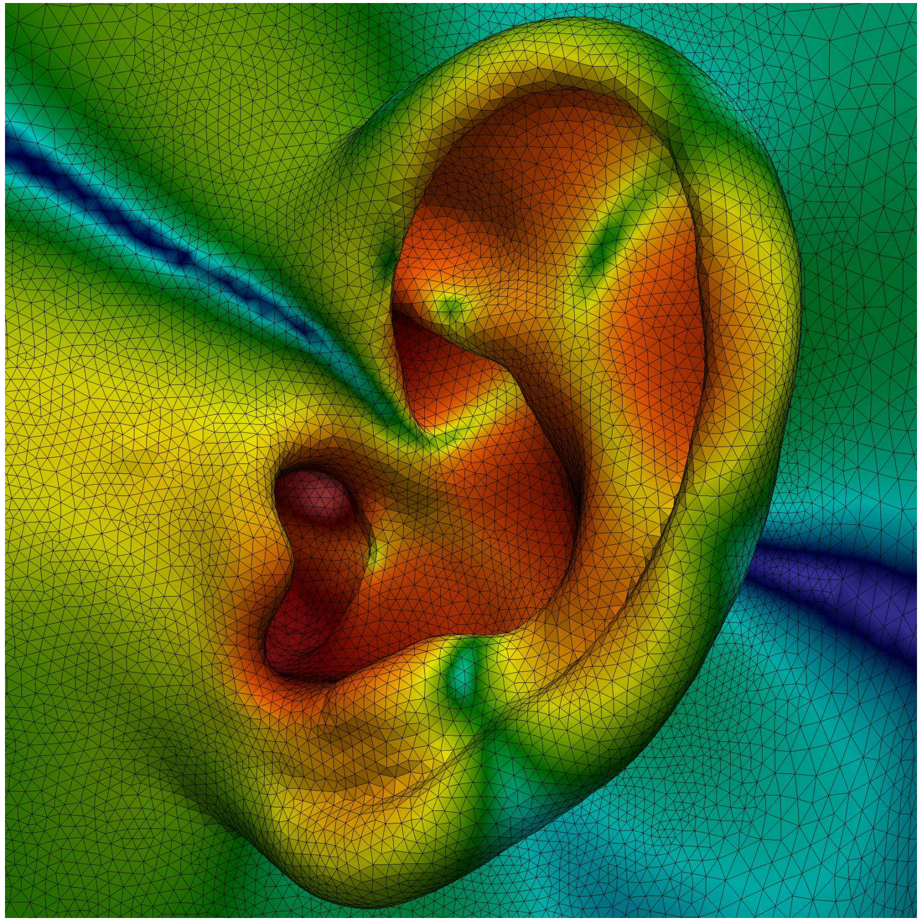
- Measurement in anechoic chamber
- Cumbersome and expensive
  - HRTFs are individual!



LUND UNIVERSITY

How do we obtain the HRTF?

- Simulation – 3d scan of torso + solve wave equation
- Less cumbersome, but still quite



LUND UNIVERSITY

How do we obtain the HRTF?

- HRTF individual -
  but is governed by body dimensions

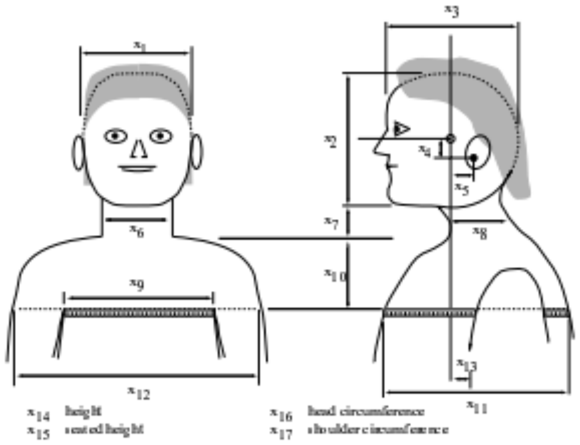


Figure 2: *Head and torso measurements*

Var	Measurement	μ	σ	%
x_1	head width	14.49	0.95	13
x_2	head height	21.46	1.24	12
x_3	head depth	19.96	1.29	13
x_4	pinna offset down	3.03	0.66	43
x_5	pinna offset back	0.46	0.59	254
x_6	neck width	11.68	1.11	19
x_7	neck height	6.26	1.69	54
x_8	neck depth	10.52	1.22	23
x_9	torso top width	31.50	3.19	20
x_{10}	torso top height	13.42	1.85	28
x_{11}	torso top depth	23.84	2.95	25
x_{12}	shoulder width	45.90	3.78	16
x_{13}	head offset forward	3.03	2.29	151
x_{14}	height	172.43	11.61	13
x_{15}	seated height	88.83	5.53	12
x_{16}	head circumference	57.33	2.47	9
x_{17}	shoulder circumference	109.43	10.30	19
d_1	cavum concha height	1.91	0.18	19
d_2	cymba concha height	0.68	0.12	35
d_3	cavum concha width	1.58	0.28	35
d_4	fossa height	1.51	0.33	44
d_5	pinna height	6.41	0.51	16
d_6	pinna width	2.92	0.27	18
d_7	intertragal incisure width	0.53	0.14	51
d_8	cavum concha depth	1.02	0.16	32
θ_1	pinna rotation angle	24.01	6.59	55
θ_2	pinna flare angle	28.53	6.70	47

Table 1. Anthropometric statistics, $\% = 100(2\,\sigma/\mu)$

LUND UNIVERSITY

How do we assign the HRTF?

- HRTF individual -
  but is governed by body dimensions

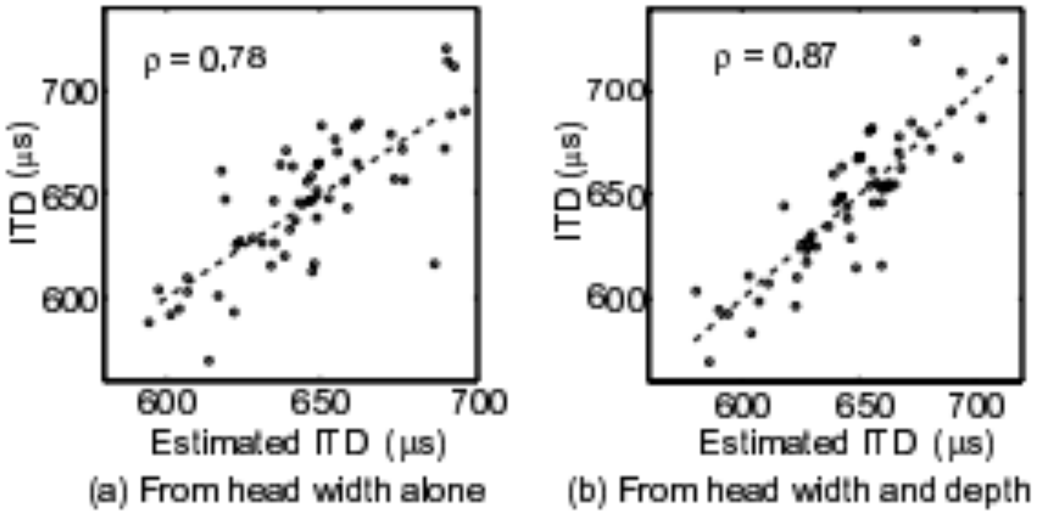


Figure 6: *Scatterplots for estimation of the ITD*

LUND UNIVERSITY

How do we assign the HRTF?

- Microsoft hololens – Augmented Reality headset
- Measures some head parameters, maps to interpolation of HRTF measurements.

LUND UNIVERSITY

Conclusion

- Many open problems
- Renewed interest with VR development – especially for HRTF.

LUND
UNIVERSITY

Bonus Material:

- Microsoft TRITON. Full wave simulation of game levels
  - Insane indeed

Triton = baked wave simulation

- Wave simulation: Accurate & reliable results on complex scenes
- Runtime = lookup + interpolation. Light on CPU.
- Need *dynamic* source & listener: large RAM!
- Baking is restricted to static geometry
  - Feasible first step
  - Dynamic scenes (doors/destruction): could layer heuristics on top, like lighting

Auto-layout **adaptively**-sampled *player probes*

After slide by Erik Molin