# Sound Zones with a Cost Function based on Human Hearing

## MSc Thesis Summary

Niels de Koeijer

Delft University of Technology, The Netherlands

Bang and Olufsen, Denmark

August 30, 2021

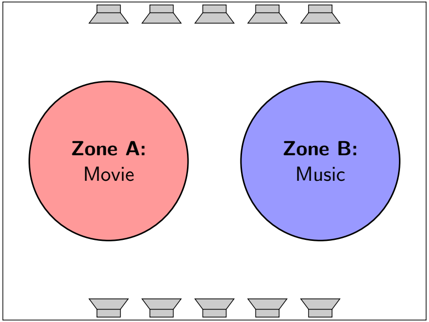
# Preface:
## About Me & this Presentation

**Niels de Koeijer**
Master Student Thesis at the Research Department at Bang & Olufsen and the Delft University of Technology

This presentation will detail the work done during my MSc thesis.

# Preface:
## The Sound Zone Problem

- **Given:**
  A room, an array of loudspeakers, and a number of zones.
- **Goal:**
  Reproduction distinct audio in the specified zones, with minimal interference.



# Preface:
## Introducing Perceptual Sound Zones

My thesis investigates improving sound zone algorithms by including a **model of the human auditory system** in the algorithms. This is a model which models how sound is perceived by humans.

The motivation for doing so is as follows:

- Sound zone algorithms typically optimize over **physical measures** such as sound pressure. Physical measures do not always correspond well with how sound is actually perceived...
- Therefore, by optimizing over a **perceptual measure** instead we are optimizing over what matters perceptually.

# Structure:
## Answering Research Questions

1. *"How can auditory perceptual models be included in sound zone algorithms?"*
    1. Determination of a suitable perceptual model for sound zone algorithms.
    2. Proposal of a perceptual sound zone framework using determined model.
    3. Proposal of perceptual sound zone algorithms through proposed framework.
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*
    1. Simulation and analysis of proposed perceptual sound zone algorithms.

# Structure:
## Answering Research Questions

1. *"How can auditory perceptual models be included in sound zone algorithms?"*
   1. **Determination of a suitable perceptual model for sound zone algorithms.**
   2. Proposal of a perceptual sound zone framework using determined model.
   3. Proposal of perceptual sound zone algorithms through proposed framework.
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*
   1. Simulation and analysis of proposed perceptual sound zone algorithms.

# Determining a Suitable Perceptual Model:
## Approach

In order to obtain a suitable perceptual model, various perceptual models from literature were considered.

- Mainly considered were algorithms that assign a **perceptually-motivated "score"** to audio. For example, such a score could quantify the perceived quality of an input audio stimuli.
- These could be used to propose sound zone algorithms that optimizes over this perceptually-motivated score.

# Determining a Suitable Perceptual Model:
## Literature Review

To this end, two categories of perceptual models were considered.

- **Objective Audio Measures:** Perceptual models that seek to predict the outcomes of listening tests, e.g. PESQ, PEAQ, Distraction, and STOI.
- **Audio Coding Models:** Perceptual models that are used to make the quantization noise introduced by audio compression as minimally disturbing as possible, e.g. the MPEG perceptual models.

# Determining a Suitable Perceptual Model:
## Introduction to the Par Distortion Detectability

From this review, the **"Par distortion detectability"**[1] is selected as the most promising model because of its ease of integration into optimization problems.

- The Par distortion detectability defines a mathematical function $D(x[n], \varepsilon[n])$ which models how easily a human listener can detect the disturbance signal $\varepsilon[n]$ in presence of the masking signal $x[n]$.
- It is used in **audio coding** to make the quantization noise introduced by compression minimally detectable.

[1] S. van de Par, A. Kohlrausch, R. Heusdens, J. Jensen, and S. H. Jensen (2005). "A perceptual model for sinusoidal audio coding based on spectral integration". In: *EURASIP Journal on Advances in Signal Processing* 2005.9, pp. 1–13

# Determining a Suitable Perceptual Model:

## Perceptual Background for the Par Distortion Detectability

The detectability $D(x[n], \varepsilon[n])$ operates using two psycho-acoustical principals: the **threshold of hearing** and the **masking properties** of the masking signal $x[n]$.

1. **Threshold of Hearing:** The sound levels as a function of frequency below which humans cannot perceive sound.
   - E.g. if a sound is below this threshold, it is not detectable.

2. **Masking Properties:** The degree to which the masking signal $x[n]$ "overpowers" other sounds.
   - E.g. if $x[n]$ is loud sound, then it will other mask sounds close in frequency, and they will not be audible.

# Structure:
## Answering Research Questions

1. *"How can auditory perceptual models be included in sound zone algorithms?"*
   1. **Determination of a suitable perceptual model for sound zone algorithms.**
   2. Proposal of a perceptual sound zone framework using determined model.
   3. Proposal of perceptual sound zone algorithms through proposed framework.
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*
   1. Simulation and analysis of proposed perceptual sound zone algorithms.

# Structure:
## Answering Research Questions

1. *"How can auditory perceptual models be included in sound zone algorithms?"*
    1. Determination of a suitable perceptual model for sound zone algorithms.
    2. **Proposal of a perceptual sound zone framework using determined model.**
    3. Proposal of perceptual sound zone algorithms through proposed framework.
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*
    1. Simulation and analysis of proposed perceptual sound zone algorithms.

# Proposal of Perceptual Sound Zone Framework:
## Approach

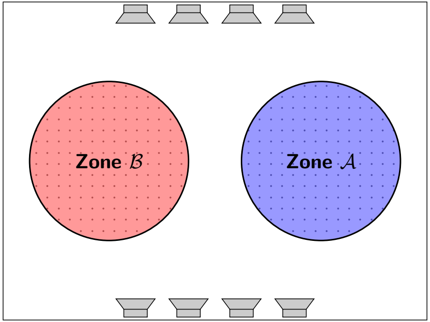
To determine how the Par distortion detectability can be used to create sound zones, sound zone literature was consulted.

- From this review it was found that a **Pressure Matching** sound zone approach can easily be adapted to use the Par distortion detectability.

# Proposal of Perceptual Sound Zone Framework:
## Review of Pressure Matching I

Pressure matching assigns **target sound pressure** that we wish to attain at **discrete points** $m$ in the zones.

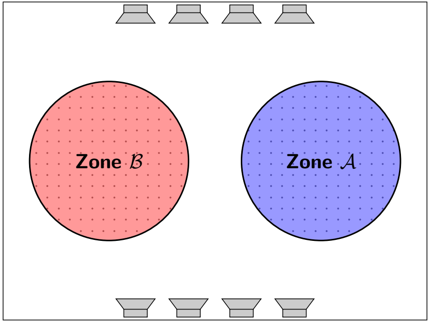
Sound zones are constructed by minimizing **sound pressure errors**.



# Proposal of Perceptual Sound Zone Framework:
## Review of Pressure Matching II

Two types of sound pressure errors can be distinguished:

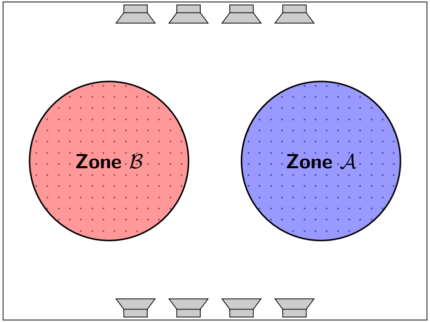
1. The **reproduction error** for control point $m$ denoted $\mathrm{RE}^{(m)}$, which quantifies the deviation from the target sound pressure.
2. The **leakage error** for control point $m$ denoted by $\mathrm{LE}^{(m)}$, which quantifies how much interference is present.



# Proposal of Perceptual Sound Zone Framework:
## Review of Pressure Matching III

The pressure matching algorithm minimizes over these sound pressure errors:

$$\arg\min \sum_{m} \left( \mathrm{RE}^{(m)} + \mathrm{LE}^{(m)} \right) \qquad (1)$$



# Proposal of Perceptual Sound Zone Framework:
## Proposal of Pressure Error Detectability Framework I

The definitions of the sound pressure errors gave inspiration for the a framework of **sound pressure error detectabilities**.

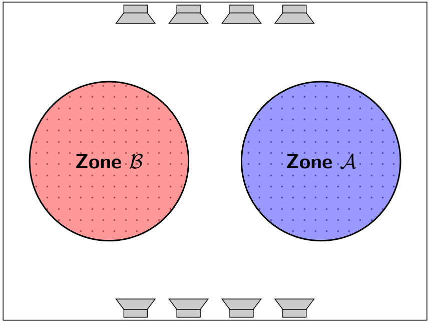
Just as the Par distortion detectability is used to minimize the detectability of quantization noise in audio coding, the proposed framework will minimize the detectabilities of the sound pressure errors from pressure matching.

This is done by **modeling the disturbance signal** $\varepsilon[n]$ of the detectability $D(x[n], \varepsilon[n])$.

# Proposal of Perceptual Sound Zone Framework:

## Proposal of Pressure Error Detectability Framework II

1. The **reproduction error detectability** for control point $m$ denoted $\mathrm{RED}^{(m)}$. Here, the disturbance signal $\varepsilon[n]$ is taken to be the deviation from the target sound pressure at point $m$.
2. The **leakage error detectability** for control point $m$ denoted by $\mathrm{LED}^{(m)}$. Here, the disturbance signal $\varepsilon[n]$ is taken to be the interference at point $m$.




In both cases, the **masking signal** $x[n]$ is taken to be the **target sound pressure**. Hence, the detectability of the errors is determined in presence of the target sound pressure.

# Structure:
## Answering Research Questions

1. *"How can auditory perceptual models be included in sound zone algorithms?"*
    1. Determination of a suitable perceptual model for sound zone algorithms.
    2. **Proposal of a perceptual sound zone framework using determined model.**
    3. Proposal of perceptual sound zone algorithms through proposed framework.
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*
    1. Simulation and analysis of proposed perceptual sound zone algorithms.

# Structure:
## Answering Research Questions

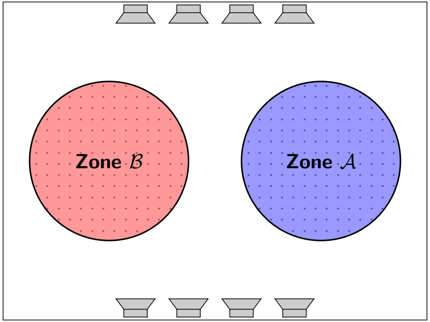
1. *"How can auditory perceptual models be included in sound zone algorithms?"*
    1. Determination of a suitable perceptual model for sound zone algorithms.
    2. Proposal of a perceptual sound zone framework using determined model.
    3. **Proposal of perceptual sound zone algorithms through proposed framework.**
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*
    1. Simulation and analysis of proposed perceptual sound zone algorithms.

# Proposal of Perceptual Sound Zone Framework:
## Algorithm 1: Unconstrained Perceptual Pressure Matching

The first perceptual sound zone algorithm is analogous to the pressure matching approach.

It simply minimizes the total sound pressure error detectability:

$$\arg\min \quad \sum_{m} \left( \mathrm{RED}^{(m)} + \mathrm{LED}^{(m)} \right) \tag{2}$$

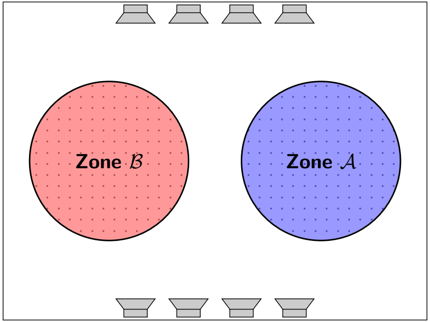


# Proposal of Perceptual Sound Zone Framework:

## Algorithm 2: Unconstrained Perceptual Pressure Matching

The second algorithm leverages the **perceptual interpretation** of the detectability. It minimizes the leakage error detectability, while constraining the reproduction error detectability.

$$\begin{aligned} \arg\min \quad & \sum_m \mathrm{LED}^{(m)} \\ \text{subject to} \quad & \mathrm{RED}^{(m)} \leq D_0 \quad \forall\, m \end{aligned} \tag{3}$$



# Structure:
## Answering Research Questions

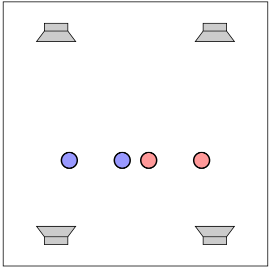
1. *"How can auditory perceptual models be included in sound zone algorithms?"*
    1. Determination of a suitable perceptual model for sound zone algorithms.
    2. Proposal of a perceptual sound zone framework using determined model.
    3. **Proposal of perceptual sound zone algorithms through proposed framework.**
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*
    1. Simulation and analysis of proposed perceptual sound zone algorithms.

# Structure:
## Answering Research Questions

1. *"How can auditory perceptual models be included in sound zone algorithms?"*
    1. Determination of a suitable perceptual model for sound zone algorithms.
    2. Proposal of a perceptual sound zone framework using determined model.
    3. Proposal of perceptual sound zone algorithms through proposed framework.
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*
    1. **Simulation and analysis of proposed perceptual sound zone algorithms.**

# Evaluation of Perceptual Sound Zone Algorithms:
## Simulation Setup

- A 5 by 5 meter square room with 4 loudspeakers is used for the evaluation.
- The zones, each consisting of two points, are assigned speech content for the simulations.
- The non-perceptual pressure algorithm previously discussed is used as a reference.



# Evaluation of Perceptual Sound Zone Algorithms:
## Evaluation Measures

In order to effectively compare the reference and the perceptual approach, perceptual measures are used.

This presentation will use the Perceptual Evaluation of Speech Quality (PESQ)[2] and Distraction[3] perceptual measures.

[2] A. W. Rix, J. G. Beerends, M. P. Hollier, and A. P. Hekstra (2001). "Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs". In: *2001 IEEE international conference on acoustics, speech, and signal processing. Proceedings (Cat. No. 01CH37221)*. Vol. 2. IEEE, pp. 749–752

[3] J. Rämö, S. Bech, and S. H. Jensen (2017). "Real-time perceptual model for distraction in interfering audio-on-audio scenarios". In: *IEEE Signal Processing Letters* 24.10, pp. 1448–1452

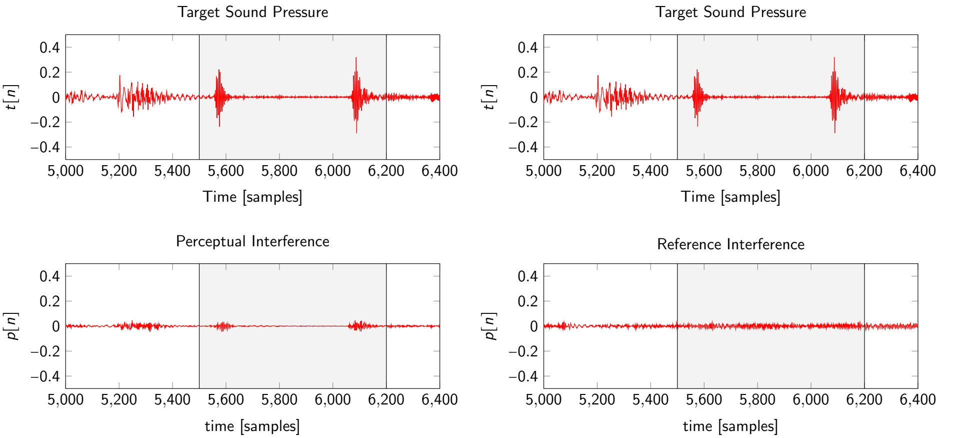
# Evaluation of Perceptual Sound Zone Algorithms:

## Evaluation of Unconstrained Perceptual Pressure Matching

| Measure | Unconstrained Perceptual PM Mean (± 95% CI) | Reference PM Mean (± 95% CI) |
|---|---|---|
| PESQ (No interference) | $3.345 \pm 0.087$ | $4.107 \pm 0.051$ |
| PESQ | $3.154 \pm 0.081$ | $2.609 \pm 0.084$ |
| Distraction | $7.828 \pm 1.868$ | $12.693 \pm 3.405$ |

The unconstrained perceptual pressure matching approach outperforms the reference when interference is taken into account.

# Evaluation of Perceptual Sound Zone Algorithms:
## Evaluation of Unconstrained Perceptual Pressure Matching



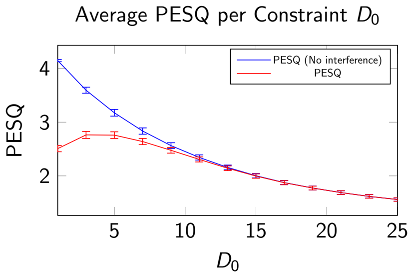
# Evaluation of Perceptual Sound Zone Algorithms:

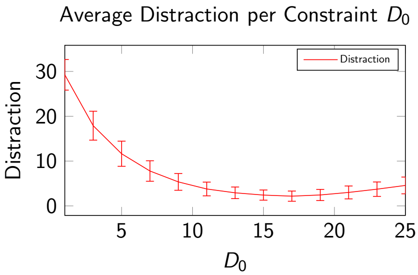
## Evaluation of Unconstrained Perceptual Pressure Matching

To conclude,

- The unconstrained perceptual pressure matching algorithm outperforms the reference for both perceptual measures when taking interference into account.
- One possible reason for this is the perceptual approach makes a **better perceptual trade-off**.

# Evaluation of Perceptual Sound Zone Algorithms:
## Evaluation of Constrained Perceptual Pressure Matching

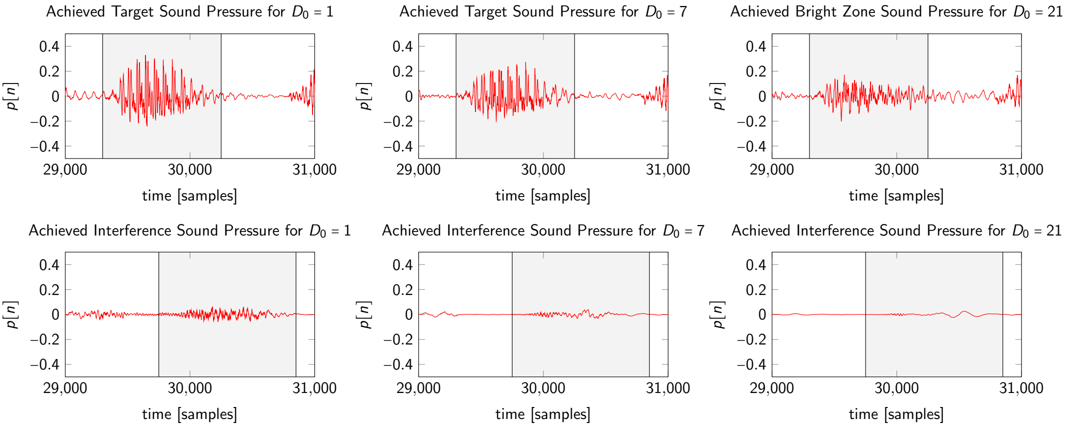



The plots show the PESQ and Distraction perceptual measures as a function of constraint $D_0$. The error bars denote 95% confidence intervals. As can be seen, the constraint allows for a degree of control over the reproduced quality.

# Evaluation of Perceptual Sound Zone Algorithms:

## Evaluation of Constrained Perceptual Pressure Matching



# Evaluation of Perceptual Sound Zone Algorithms:
## Evaluation of Constrained Perceptual Pressure Matching

To conclude,

- The constraint is shown to correlate well with the other perceptual measures, which allows for a degree of control over the quality of the reproduced audio.

# **Structure:**
## Answering Research Questions

1. *"How can auditory perceptual models be included in sound zone algorithms?"*
    1. Determination of a suitable perceptual model for sound zone algorithms.
    2. Proposal of a perceptual sound zone framework using determined model.
    3. Proposal of perceptual sound zone algorithms through proposed framework.
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*
    1. **Simulation and analysis of proposed perceptual sound zone algorithms.**

# Future Work:

- Further proposal of new perceptual sound zone algorithms through the proposed framework.
- Current algorithms are very slow, as such it is of interest to speed them up!

# Sources:


Par, S. van de, A. Kohlrausch, R. Heusdens, J. Jensen, and S. H. Jensen (2005). "A perceptual model for sinusoidal audio coding based on spectral integration". In: *EURASIP Journal on Advances in Signal Processing* 2005.9, pp. 1–13.

Rämö, J., S. Bech, and S. H. Jensen (2017). "Real-time perceptual model for distraction in interfering audio-on-audio scenarios". In: *IEEE Signal Processing Letters* 24.10, pp. 1448–1452.

Rix, A. W., J. G. Beerends, M. P. Hollier, and A. P. Hekstra (2001). "Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs". In: *2001 IEEE international conference on acoustics, speech, and signal processing. Proceedings (Cat. No. 01CH37221)*. Vol. 2. IEEE, pp. 749–752.