**Circuits and Systems**
Mekelweg 4,
2628 CD Delft
The Netherlands
http://ens.ewi.tudelft.nl/

**CAS-2021-??**

# M.Sc. Thesis

## Sound Zones with a Cost Function based on Human Hearing

**Niels Evert Marinus de Koeijer B.Sc.**

**Abstract**

Something about soundzones, something about perceptual models, something about perceptual sound zones.

# Sound Zones with a Cost Function based on Human Hearing

## Subtitle Compulsory?

---

Thesis

submitted in partial fulfillment of the
requirements for the degree of

Master of Science

in

Electrical Engineering

by

Niels Evert Marinus de Koeijer B.Sc.
born in Delft, The Netherlands

This work was performed in:

Circuits and Systems Group
Department of Microelectronics & Computer Engineering
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology

**Delft University of Technology**

Delft University of Technology
Department of
Microelectronics & Computer Engineering

The undersigned hereby certify that they have read and recommend to the Faculty of Electrical Engineering, Mathematics and Computer Science for acceptance a thesis entitled **"Sound Zones with a Cost Function based on Human Hearing"** by **Niels Evert Marinus de Koeijer B.Sc.** in partial fulfillment of the requirements for the degree of **Master of Science**.

Dated: September 15, 2021

Chairman: ________________________
dr.ir. R.C. Hendriks

Advisors: ________________________
dr. M. Bo Møller

________________________
dr. P. Martinez Nuevo

Committee Members: ________________________
dr. M. Mastrangeli

________________________
dr. J. Martinez Castañeda

# Abstract

Something about soundzones, something about perceptual models, something about perceptual sound zones.

# Acknowledgments

I would like to thank dr. M. Bo Møller, dr. P. Martinez Nuevo, dr.ir. R.C. Hendriks, and dr. J. Martinez Castañeda.

Niels Evert Marinus de Koeijer B.Sc.
Delft, The Netherlands
September 15, 2021

# Contents



**TODO: I am considering merging some chapters (e.g. merging the literature review and the implementation chapters).**

# List of Figures

# List of Tables

# 1 Introduction

**Skeleton of Chapter**

An introduction to the work and the problem it seeks to solve.

- An introduction to sound zones as a concept.
- Where sound zones come short, motivation to use perceptual models.
- Statement of the goal of the project:
  - Creating a perceptual multi-zone algorithm
- A description of what the rest of the document will contain.

# 2 Review of Perceptual Model Literature

**Skeleton of Chapter**

In order to build a perceptual sound zone algorithm, we review literature for perceptual models to find a suitable perceptual model.

- I will discuss the criteria which will determine which perceptual model is chosen.
  - Complexity
  - Feasibility to optimize
- I will discuss my literature review into perceptual models to find a model that best fits the criteria.
  - Dau Model
  - Detectibility Models, i.e. Par and Taal
  - Distraction Model
  - Audio quality models, PEAQ, VISQOL
  - Speech Intelligibility Based, i.e. SIIB and STOI
- I will discuss and motivate the chosen sound zone approach (pressure matching) and the chosen perceptual model (detectibility). This is done by means of summarizing the findings, and then reflecting on the criteria. From this, I will conclude that the **Par Detectibility** is best suited.
- I will discuss what perceptual models will be used for evaluation. From this I will conclude that PEAQ / VISQOL are useful for quality evaluation, and that the distraction model is useful for leakage evaluation.

## 2.1 Introduction

# 3 Implementation of Perceptual Model

**Skeleton of Chapter**

Here, I describe the implementation of the chosen perceptual models, i.e. the detectibility.

- I give a high-level description of detectibility.
- I describe the Par Detectibility.
  - The underlying perceptual ideas
  - How its implemented
  - How its calibrated
- I show that my implementations of the Par Detectibility is valid. This is done by comparing the masking curve predictions to a reference implementation of the Dau model.

## 3.1 Introduction

# 4 Review of Sound Zone Algorithms Literature

**Skeleton of Chapter**

At this point the Par Detectibility as been selected as the perceptual model of choice. In this chapter, the literature will be reviewed in order to find a suitable sound zone approach for integrating the Par Detectibility.

- I will define the criteria that are required for a sound zone algorithm.
  - I am yet to think of any.
- I will discuss my literature review into sound zones.
  - Pressure Matching (PM) Approaches.
  - Acoustic Contrast Control (ACC) Approaches.
  - Mode Matching Approaches.
  - The work by Tae-Woong Lee.
- I will summarize the results and reflect on the requirements. From this I will conclude that the pressure matching approach is the most suited.

## 4.1 Introduction

# 5 Implementation of Reference Sound Zone Algorithm

**Skeleton of Chapter**

In the preceding chapter it was concluded that a pressure matching approach was best suited for building a perceptual sound zone algorithm.

1. To form the basis for the perceptual algorithm.
2. To function as a reference implementation to evaluate performance.

The chapter will discuss the following

- Introduction of a data model from which the sound zone problem can be stated mathematically.
- Derivation of a multi-zone pressure matching approach from the previously introduced mathematical framework.
- Extension of the multi-zone pressure matching to perform on a frame by frame basis. This is done in order to allow for real-time sound zones.

## 5.1 Introduction

## 5.2 Data Model

**TODO: CITATION NEEDED, I have not cited anything. Probably should?** In this section the base data model will be introduced. This data model will be later used in the derivation of the sound zone algorithms.

First, in subsection 5.2.1 a spatial description of a room, the zones and loudspeakers contained within will be given. In addition to this, a signal model will then be given for the target sound pressure and loudspeaker input signals. The derived mathematical framework will then allow for posing sound zone problem more formally.

Next, subsection 5.2.2 will then define the desired target sound pressure introduced previously, and motivate this choice.

Finally, **??** will introduce the Multi-Zone Pressure Matching sound zone algorithm. This algorithm solves the sound zone problem, and forms the basis for the perceptual sound zone algorithms introduced later. It will also serve as the reference to which the perceptual models will be compared.

### 5.2.1 Room Model and Sound Zone Problem Statement

**TODO: I am not great at topology, so I would love some tips on how to write this down in a better way...** In this section, a description of the room in which sound zones are to be reproduced will be given. In general, the room can contain any number of zones, but this thesis will focus on the two zone case (the work can however be extended to a larger number of zones). The room description can then be used to pose the sound zone problem in a more formal way.

In general, the room $R$ can be modeled as a closed subset of three dimensional space, $\mathcal{R} \subset \mathbb{R}^3$. The two non-overlapping zones $\mathcal{A}$ and $\mathcal{B}$ are contained within the room $R$, i.e. $\mathcal{A} \subset \mathcal{R}$ and $\mathcal{B} \subset \mathcal{R}$ where $\mathcal{A} \cap \mathcal{B} = \emptyset$. In addition to the zones, the room $\mathcal{R}$ also contains $N_L$ loudspeakers, which can be modeled as discrete points.

The goal of the sound zone algorithm is to use the loudspeakers to realise a specified target sound pressure in the space described by zones $\mathcal{A}$ and $\mathcal{B}$. This is to be done in such a way that there is minimal interference between zones; meaning that target sound pressure intended for one zone should not be audible in the other zones. The loudspeakers can be controlled by specifying their input signals. As such, the goal of the sound zone algorithm can be reframed as finding loudspeaker input signals such that a specified target sound pressure is attained.

The rest of this section will focus on formalizing this notion mathematically. First, a way of modeling the target sound pressure will be discussed. Afterwards, a way of realizing said target sound pressure by controlling the loudspeaker inputs will be given mathematically. Finally, the data model is used to state the goal of the sound zone algorithm more formally.

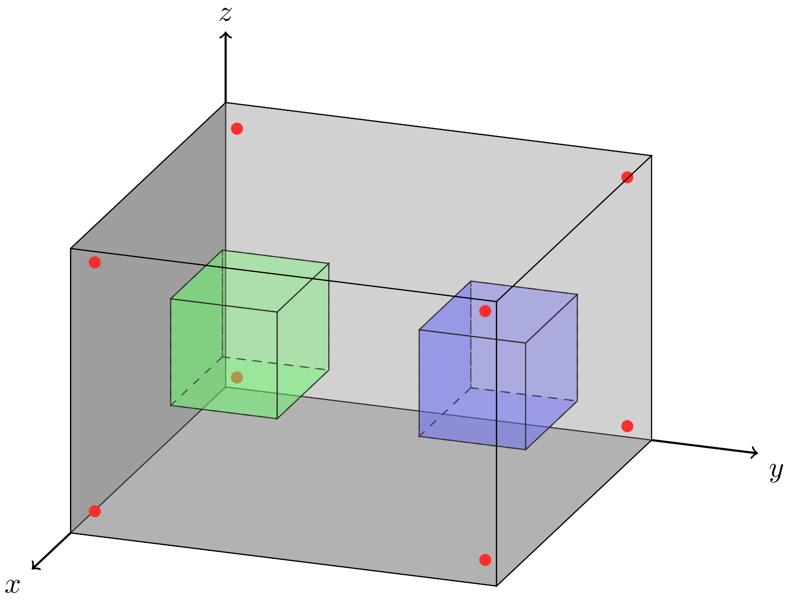


Figure 5.1: The room $\mathcal{R} \subset \mathbb{R}^3$ containing the zones $\mathcal{A} \subset \mathcal{R}$ and $\mathcal{B} \subset \mathcal{R}$ depicted in green and blue respectively. The room contains $N_L = 8$ loudspeakers, which are denoted by the red dots in the corners of the room.

#### 5.2.1.1 Defining Target Sound Pressure

As mentioned, the goal of the sound zone algorithm is to realize a specified target sound pressure in the different zones $\mathcal{A}$ and $\mathcal{B}$ in the room $R$.

The zones are given as continuous regions in space. Some sound zone approach will attempt to recreate a specified pressure in the entire region of space defined by $\mathcal{A}$ and $\mathcal{B}$. Other sound zone approaches will instead discretion the zones into so-called control points. The sound pressure is then controlled only in these control points.

In this work, the latter approach will be taken. **TODO: Why does discretizing make sense? Add more reasons...** Thus, we discretize zones $\mathcal{A}$ and $\mathcal{B}$ into a total of $N_a$ and $N_b$ control points respectively. Let $A$ and $B$ denote the sets of the resulting control points points contained within zones $\mathcal{A}$ and $\mathcal{B}$ respectively.

Now let $t^m[n]$ denote the target sound pressure at control point $m$ in either $A$ or $B$, i.e. $m \in A \cup B$. Our goal is thus to realize $t^m[n]$ in all control points $m \in A \cup B$ using the loudspeakers present in the room. The relationship between the loudspeaker input signals and the sound pressure is the topic of the next section.

#### 5.2.1.2 Realizing Sound Pressure through the Loudspeaker

The sound pressure produced by the loudspeakers can be controlled by specifying their input signals. Mathematically speaking, let $x^{(l)}[n] \in \mathbb{R}^{N_x}$ denote the loudspeaker input signal for the $l^{\text{th}}$ loudspeaker. As such, the goal of the sound zone algorithm is to find loudspeaker inputs $x^{(l)}[n]$ such that the target sound pressure $t^m[n]$ is realized for all

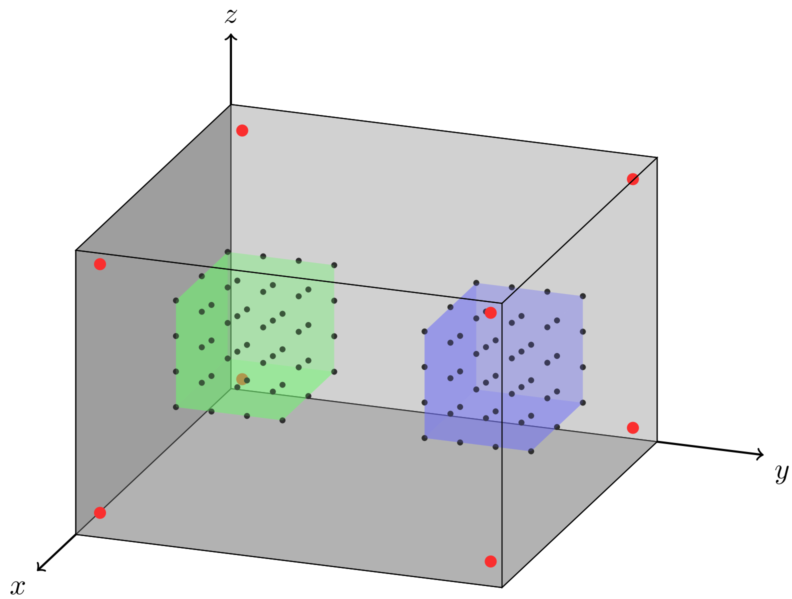


Figure 5.2: The previously introduced room $\mathcal{R}$ with zones $\mathcal{A}$ and $\mathcal{B}$ discretized.

$m \in A \cup B$.

In order to do so, a relationship must be established between the loudspeaker inputs $x^{(l)}[n]$ and the resulting sound pressure at control points $m \in A \cup B$. This relationship can be modeled by room impulse responses (RIRs) $h^{(l,m)}[n] \in \mathbb{R}^{N_h}$.

The RIRs $h^{(l,m)}[n]$ determines what sound pressure is realized at control point $m$ due to playing loudspeaker signal $x^{(l)}[n]$. Mathematically, let $p^{(l,m)}[n] \in \mathbb{R}^{N_x+N_h-1}$ represent the realized sound pressure in control point $m$ due to playing $x^{(l)}[n]$ from loudspeaker $l$:

$$p^{(l,m)}[n] = \left(h^{(l,m)} * x^{(l)}\right)[n] \tag{5.1}$$

The realized sound pressure $p^{(l,m)}[n]$ only considers the contribution of loudspeaker $l$ at reproduction point $m$. Let $p^{(l)}[n] \in \mathbb{R}^{N_x+N_h-1}$ denote the total sound pressure due to all $N_L$ loudspeakers. It can now be expressed as the sum over all contributions as follows:

$$\begin{aligned} p^{(m)}[n] &= \sum_{l=0}^{N_L} p^{(l,m)}[n] && (5.2) \\ &= \sum_{l=0}^{N_L} \left(h^{(l,m)} * x^{(l)}\right)[n] && (5.3) \end{aligned}$$

#### 5.2.1.3 Sound Zone Problem Statement

With this data model is complete and the goal of the sound zone algorithm can be restated. Namely, the goal is to find $x^{(l)}[n]$ such that the realised sound pressure $p^{(m)}[n]$ attains the target sound pressure $t^{(m)}[n]$ for all control points $m \in A \cup B$.

The rest of section 5.2 will describe how this problem can be solved in greater detail.

### 5.2.2 Choice of Target Pressure

The target sound pressure $t^{(m)}[n]$ describes the desired content for a specific control point $m$. So far, the choice of target sound pressure $t^{(m)}[n]$ has been kept general. In this section, a choice for the target pressure will be made and motivated.

Assume that the user of the sound zone system has selected loudspeaker input signals $s_{\mathcal{A}}[n] \in \mathbb{R}^{N_x}$ and $s_{\mathcal{B}}[n] \in \mathbb{R}^{N_x}$ that they wish to hear in zone $\mathcal{A}$ and $\mathcal{B}$ respectively. In order to accomodate the wishes of the user, the target sound pressure is chosen as follows:

$$t^{(m)}[n] = \sum_{l=0}^{N_L} \left(h^{(l,m)} * s_{\mathcal{A}}\right)[n] \qquad \forall\, m \in A \tag{5.4}$$

$$t^{(m)}[n] = \sum_{l=0}^{N_L} \left(h^{(l,m)} * s_{\mathcal{B}}\right)[n] \qquad \forall\, m \in B \tag{5.5}$$

This choice for target sound pressure can be understood as the sound pressure that results when separately playing the selected loudspeaker input signals $s_{\mathcal{A}}[n]$ and $s_{\mathcal{B}}[n]$ through the loudspeakers in the room.

In other words, the sound pressure in a zone is chosen such that it matches the sound pressure arises in said zone when playing only the desired input signal. For example, when in zone $A$, the target sound pressure is set equal to the sound pressure corresponding to playing only $s_{\mathcal{A}}[n]$ from the loudspeakers.

The motivation for choosing this target is that it physically attainable with the given loudspeakers and room. **TODO: Expand this motivation with a couple more arguments.**

## 5.3 Multi-Zone Pressure Matching Solution Approach

The "Pressure Matching" (PM) is widely used in literature to solve the sound zone problem. In this section, a "Multi-Zone Pressure Matching" (MZ-PM) algorithm will be derived. The motivation for discussing it is that it will be used as the foundation on which the perceptual sound zone algorithm will be built, as it was found that perceptual information was easily intergratable into the pressure matching framework. **TODO: Expand this motivation with a couple more arguments...?**

In the typical PM approach, the resulting loudspeaker input signals $x^{(l)}[n]$ are determined for just a single zone. If the solution for multiple zones is desired, than multiple PM problems must be solved and their resulting loudspeaker input signals combined. In the MZ-PM approach, the loudspeaker input signals are instead determined for jointly for all zones.

In a two zone approach, the loudspeaker input signals are decomposed into two parts as follows:

$$x^{(l)}[n] = x_{\mathcal{A}}^{(l)}[n] + x_{\mathcal{B}}^{(l)}[n] \tag{5.6}$$

Here, $x_{\mathcal{A}}^{(l)}[n]$ and $x_{\mathcal{B}}^{(l)}[n]$ are the parts of the loudspeaker input signal responsible for reproducing the target sound pressure in zone $\mathcal{A}$ and $\mathcal{B}$ respectively. Now, it is possible to consider the sound pressure that arises due to the separate loudspeaker input signals:

$$p_{\mathcal{A}}^{(m)}[n] = \sum_{l=0}^{N_L} \left( h^{(l,m)} * x_{\mathcal{A}}^{(l)} \right)[n] \tag{5.7}$$

$$p_{\mathcal{B}}^{(m)}[n] = \sum_{l=0}^{N_L} \left( h^{(l,m)} * x_{\mathcal{B}}^{(l)} \right)[n] \tag{5.8}$$

Here, $p_{\mathcal{A}}^{(m)}[n]$ and $p_{\mathcal{B}}^{(m)}[n]$ can be understood to be the pressure that arises due to playing loudspeaker input signals $x_{\mathcal{A}}^{(l)}[n]$ and $x_{\mathcal{B}}^{(l)}[n]$ respectively.

The idea in this approach is to chose $x_{\mathcal{A}}^{(l)}[n]$ and such that the resulting pressure $p_{\mathcal{A}}^{(m)}[n]$ attains the target sound pressure $t^{(m)}[n]$ in all $m \in A$. At the same time however, $p_{\mathcal{A}}^{(m)}[n]$ should not attain any sound pressure in all $m \in B$. Any sound pressure resulting from $x_{\mathcal{A}}^{(l)}[n]$ in zone $\mathcal{B}$ is essentially leakage or cross-talk between zones. Similar arguments can be given for $x_{\mathcal{B}}^{(l)}[n]$: it should reproduce the target sound pressure for $m \in B$ but no sound pressure for $m \in A$.

In the MZ-PM approach, the loudspeaker weights $x_{\mathcal{A}}^{(l)}[n]$ and $x_{\mathcal{B}}^{(l)}[n]$ that achieve this goal are found by minimizing the difference between the intended pressure and the realized pressure as follows:

$$\underset{x_{\mathcal{A}}^{(l)}[n],\, x_{\mathcal{B}}^{(l)}[n]\, \forall\, l}{\arg\min} \quad \sum_{m \in A} \left|\left| p_{\mathcal{A}}^{(m)}[n] - t^{(m)}[n] \right|\right|_2^2 + \sum_{m \in A} \left|\left| p_{\mathcal{B}}^{(m)}[n] \right|\right|_2^2 + \tag{5.9}$$

$$\sum_{m \in B} \left|\left| p_{\mathcal{B}}^{(m)}[n] - t^{(m)}[n] \right|\right|_2^2 + \sum_{m \in B} \left|\left| p_{\mathcal{A}}^{(m)}[n] \right|\right|_2^2 \tag{5.10}$$

Here, the first two terms can be understood as the reproduction error and the leakage for zone $\mathcal{A}$. Similarly, the last two terms are the reproduction error and leakage for zone $\mathcal{B}$. Typically, this approach results in trade-off between minimizing the reproduction errors and leakages. Some pressure matching approaches attempt to control this trade-off by introducing weights for the different error terms, or constraints.

The problem can be solved in the time and the frequency domain. In frequency domain approaches, the convolutions become inner products, which typically results in a lower computational complexity.

The algorithm above will form the basis of the perceptual algorithms to be introduced in the following sections.

## 5.4 Frame-Based Processing Framework

In the preceding section it is assumed that the desired playback signals $s_{\mathcal{A}}[n]$ and $s_{\mathcal{B}}[n]$ were known in their entirety. In practice however, this is not a valid assumption as a user can change the desired playback content in real-time. In this section, the Multi-Zone Pressure Matching approach introduced in **??** will be adjusted in order to accommodate the real-time processing constraint. **TODO: This really needs better motivation...?**

As mentioned, it cannot be assumed that the desired playback signals $s_{\mathcal{A}}[n]$ and $s_{\mathcal{B}}[n]$ are known for all $n$ if the system is to accommodate real-time playback. Instead, at time $m$ it is assumed that $s_{\mathcal{A}}[n]$ and $s_{\mathcal{B}}[n]$ are only known from $\infty \leq n \leq m$, i.e. the most recent and all previous input samples.

The loudspeaker input signals $x^{(l)}[n]$ must thus be found in real-time. In order to do so, a block-processing based approach is taken where $x^{(l)}[n]$ is computed in blocks of hop size $H$. When $H$ new samples of $s_{\mathcal{A}}[n]$ and $s_{\mathcal{B}}[n]$ are revealed to the system, $H$ new samples of $x^{(l)}[n]$ can be computed. As such, this approach will introduce a delay of at least $H$, as $H$ new samples of the desired playback signals are required before computation can begin. **TODO: This really needs better motivation...?**

The processing occurs in steps of $H$ samples. Let $\mu = \lfloor n/H \rfloor$ index is the current step at a time $n$. Thus at a time $n$, up to and including the $\mu^{\text{th}}$ block of desired playback signals is known.

Adapting the previously derived data model to the chosen block-based approach will be the topic of this section.

### 5.4.1 Implications for Computing Target Pressure

As the playback signals are revealed in blocks of size $H$, it makes sense to update the target sound pressure $t^{(m)}[n]$ in blocks of $H$. Consider the following rewrite of the target signal for the desired sound pressure of zone $\mathcal{A}$:

$$t^{(m)}[n] = \sum_{l=0}^{N_L-1} \left(h^{(l,m)} * s_{\mathcal{A}}\right)[n] \tag{5.11}$$

$$= \sum_{l=0}^{N_L-1} t^{(l,m)}[n] \tag{5.12}$$

Here, $t^{(m,l)}[n]$ is the contribution of the $l^{\text{th}}$ loudspeaker to the target sound pressure at reproduction point $m$. Now, consider the following rewrite:

$$
\begin{aligned}
t^{(m,l)}[n] &= \sum_{b=n-N_h+1}^{n} h^{(l,m)}[n-b] s_{\mathcal{A}}[b] && (5.13)\\
&= \sum_{b=n-N_h+1}^{n} h^{(l,m)}[n-b] s_{\mathcal{A}}[b] \sum_{k=-\infty}^{\infty} w[b-kH] && (5.14)\\
&= \sum_{b=n-N_h+1}^{n} h^{(l,m)}[n-b] \sum_{k=-\infty}^{\infty} s_{\mathcal{A}}[b] w[b-kH] && (5.15)\\
&= \sum_{b=n-N_h+1}^{n} h^{(l,m)}[n-b] \sum_{k=-\infty}^{\infty} s_{\mathcal{A},k}[b] && (5.16)
\end{aligned}
$$

Here, $w[n] \in \mathbb{R}^{N_w}$ is a window that satisfies the COLA condition for a hop size $H$. The window is defined to be non-zero for $-N_w + 1 \leq n \leq 0$, as such it is non-causal. Furthermore, the windows are overlapping, thus $N_w > H$.

In the rewrite above, the desired playback signal $s_{\mathcal{A}}[n]$ is projected onto a basis of overlapping frames of size $N_w$. The projection results in a sum of frames $s_{\mathcal{A},k}[n]$. The support of $s_{\mathcal{A},k}[n]$ is defined by the shifted window that is used to synthesize it, i.e. it is non-zero for $-N_w+1+kH \leq n \leq kH$. As the COLA condition is met for the chosen window, the sum over all frames reconstructs $s_{\mathcal{A}}[n]$ perfectly.

At a time $n = \mu H$, the frames up to $k = \mu$ can be computed. Let $t_{\mu}^{(m,l)}[n]$ represent the target using frames up to $k = \mu$:

$$
\begin{aligned}
t_{\mu}^{(m,l)}[n] &= \sum_{b=n-N_h+1}^{n} h^{(l,m)}[n-b] \sum_{k=-\infty}^{\mu} s_{\mathcal{A},k}[b] && (5.17)\\
&= \sum_{b=n-N_h+1}^{n} h^{(l,m)}[n-b] s_{\mathcal{A},\mu}[b] + \sum_{b=n-N_h+1}^{n} h^{(l,m)}[n-b] \sum_{k=-\infty}^{\mu-1} s_{\mathcal{A},k}[b] && (5.18)\\
&= \sum_{b=n-N_h+1}^{n} h^{(l,m)}[n-b] s_{\mathcal{A},\mu}[b] + t_{\mu-1}^{(m,l)}[n] && (5.19)
\end{aligned}
$$

As can be seen, $t_{\mu}^{(m,l)}[n]$ can expressed as the contribution of the current frames and the contribution of all previous frames. In addition, note how the computation can be performed recursively. To compute $t_{\mu}^{(m,l)}[n]$, we compute the convolution of the current frame $s_{\mathcal{A},\mu}[n]$ with the RIRs, and then add the history of previous frames.

Thus, $t_{\mu}^{(m,l)}[n]$ can be considered an estimation of the target given frames up to $\mu$. As new frames are revealed, the target estimation can be updated. Note that this definition converges to the "true" target estimation: $t_{\infty}^{(m,l)}[n] = t^{(m,l)}[n]$.

Essentially, this approach allows for the real-time computation for the target signal.

### 5.4.2 Implications for Computing Loudspeaker Inputs

Just as the target is comuted as new frames are revealed, the loudspeaker input should also be computed this way. **TODO: I need to motivate this?** Consider the following rewrite of the realized sound pressure $p^{(l)}[n]$:

$$p^{(m)}[n] = \sum_{l=0}^{N_L-1} \left(h^{(l,m)} * x^{(l)}\right)[n] \tag{5.20}$$

$$= \sum_{l=0}^{N_L-1} p^{(l,m)}[n] \tag{5.21}$$

$p^{(m,l)}[n]$ is the contribution of the $l^{\text{th}}$ loudspeaker to the realized sound pressure at reproduction point $m$. Consider the following:

$$p^{(m,l)}[n] = \sum_{b=n-N_h+1}^{n} h^{(l,m)}[n-b]x^{(l)}[b] \tag{5.22}$$

$$= \sum_{b=n-N_h+1}^{n} h^{(l,m)}[n-b]x^{(l)}[b] \sum_{k=-\infty}^{\infty} w[b-kH] \tag{5.23}$$

$$= \sum_{b=n-N_h+1}^{n} h^{(l,m)}[n-b] \sum_{k=-\infty}^{\infty} x^{(l)}[b]w[b-kH] \tag{5.24}$$

$$= \sum_{b=n-N_h+1}^{n} h^{(l,m)}[n-b] \sum_{k=-\infty}^{\infty} x_k^{(l)}[b] \tag{5.25}$$

Analogous to the derivation for the target sound pressure $t^{(m,l)}[n]$, the loudspeaker input signal is project onto a basis consisting of windows $w[n]$. This results in frames $x_k^{(l)}[n]$. Just as with the target sound pressure, let $p_\mu^{(m,l)}[n]$ represent the realized sound pressure using frames up to $k = \mu$:

$$p_\mu^{(m,l)}[n] = \sum_{b=n-N_h+1}^{n} h^{(l,m)}[n-b] \sum_{k=-\infty}^{\mu} x_k^{(l)}[b] \tag{5.26}$$

$$= \sum_{b=n-N_h+1}^{n} h^{(l,m)}[n-b]x_\mu^{(l)}[b] + \sum_{b=n-N_h+1}^{n} h^{(l,m)}[n-b] \sum_{k=-\infty}^{\mu-1} x_k^{(l)}[b] \tag{5.27}$$

$$= \sum_{b=n-N_h+1}^{n} h^{(l,m)}[n-b]x_\mu^{(l)}[b] + p_{\mu-1}^{(m,l)}[n] \tag{5.28}$$

As such, the realized pressure can be computed recursively just like the target sound pressure. Using this data model also allows for the recursive computation of loudspeaker frames $x_\mu^{(l)}[n]$. The loudspeaker input signals must be chosen such that sound pressure realized by them must approximate the target sound pressure.

This reveals one possible solution approach: computing the loudspeaker frames $x_\mu^{(l)}[n]$ such that the target sound pressure $t_\mu^{(l)}[n]$ is attained. **TODO: Expand on this. Why does this approach make sense? There are probably other ways of doing it aswell...**

### 5.4.3 Implications for the Multi-Zone Pressure Matching Algorithm

**TODO: Expand...** In the previous sections, frame-based approaches for the target sound pressure and the realized sound pressure were derived. This allows us to rewrite the problem as follows:

$$\underset{x_{\mathcal{A},\mu}^{(l)}[n],\, x_{\mathcal{B},\mu}^{(l)}[n]\,\forall\, l}{\arg\min} \quad \sum_{m\in A} \left|\left|p_{\mathcal{A},\mu}^{(m)}[n] - t_\mu^{(m)}[n]\right|\right|_2^2 + \sum_{m\in A} \left|\left|p_{\mathcal{B},\mu}^{(m)}[n]\right|\right|_2^2 + \tag{5.29}$$

$$\sum_{m\in B} \left|\left|p_{\mathcal{B},\mu}^{(m)}[n] - t_\mu^{(m)}[n]\right|\right|_2^2 + \sum_{m\in B} \left|\left|p_{\mathcal{A},\mu}^{(m)}[n]\right|\right|_2^2 \tag{5.30}$$

The problem above is solved recursively for each frame $\mu$. The loudspeaker input signals $x_\mu^{(l)}$ can then be recombined as follows:

$$x^{(l)} = \sum_{k=-\infty}^{\infty} x_k^{(l)} \tag{5.31}$$

# 6 Perceptual Sound Zone Algorithm

**Skeleton of Chapter**

Describes the design of the perceptual sound zone algorithm.

- Introduction of detectibility to form perceptual cost function building blocks:
  - Detectibility of reproduction error
  - Detectibility of interference
- Show how building blocks can be used to form different perceptual algorithms:
  - Algorithm 1: Minimization of detectibility of reproduction error and detectibility of interference.
  - Algorithm 2: Minimization of detectibility of reproduction error subject to constraint on detectibility of interference.
  - Algorithm 3: Minimization of detectibility of interference subject to constraint on detectibility of reproduction error.
- Evaluate different algorithms:
  - Compare Reference and Algorithm 1 in terms of Distraction and PEAQ.
  - Compare Reference and Algorithm 2 when varying the constraint on the detectibility of the interference in terms of Distraction and PEAQ.
  - Compare Reference and Algorithm 3 when varying the constraint on the detectibility of the reproduction error in terms of Distraction and PEAQ.

## 6.1 Introduction

# 7 Conclusion

**Skeleton of Chapter**

A conclusion about the work.

# 8 Bibliography