



LUND
UNIVERSITY

A NEW APPROACH FOR DETECTING STRONG REFLECTIONS IN ROOMS

STINA WARGERT

Engineering
Acoustics

Master's Dissertation

DEPARTMENT OF CONSTRUCTION SCIENCES
DIVISION OF ENGINEERING ACOUSTICS

ISRN LUTVDG/TVBA--14/5045--SE (1-73) | ISSN 0281-8477

MASTER'S DISSERTATION

A NEW APPROACH FOR DETECTING STRONG REFLECTIONS IN ROOMS

STINA WARGERT

Supervisors: **JONAS BRUNSKOG**, Assoc. Prof. and **CHEOL-HO JEONG**, Assoc. Prof.,
Dept. of Electrical Engineering - Acoustic Technology, DTU, Denmark
and **GRY BÆLUM NIELSEN**, COWI A/S, Kongens Lyngby, Denmark.

Examiner: **DELPHINE BARD**, Assoc. Prof., Div. of Engineering Acoustics, LTH, Lund.

Copyright © 2014 by Division of Engineering Acoustics,
Faculty of Engineering (LTH), Lund University, Sweden.

Printed by Media-Tryck LU, Lund, Sweden, April 2014 (*PI*).

For information, address:

Div. of Engineering Acoustics, LTH, Lund University, Box 118, SE-221 00 Lund, Sweden.

Homepage: <http://www.akustik.lth.se>

Abstract

In this thesis the influence of strong reflections in rooms is investigated. A strong reflection can have a negative influence on speech or music that is transmitted from the source to the receiver. Impulse responses including different strength of a reflection are modelled using a statistical model. From the impulse response, different parameters describing the strength of the reflection can be calculated and the detection thresholds for the parameters are found from listening tests. Three different parameters were compared, of which one is an existing parameter by Dietsch and Kraak and the other two are new parameters based on the energy decay curve of the impulse response. The threshold is given as the minimum level at which the subject can hear the reflection by listening to the impulse response. The detection threshold is found from impulse responses for rooms with different volumes and reverberation times and for different delay times of the strong reflection. The performance of the parameters is then discussed and a suggested technique for identifying and evaluating a strong reflection is given.

Acknowledgements

I would like to give a great thanks to my supervisors Jonas Brunskog and Cheol-Ho Jeong for their encouragement and valuable knowledge. Thanks for taken the time for discussion and giving me feedback, I really appreciate their enthusiasm and support.

My greatfull thanks also goes to my supervisor Gry Bælum Nielsen for helping me in the progress of the project with her professional working knowledge, and also for the instruction in dancing Lindy Hop! I also want to thank all the acousticians at COWI for giving me an insight of room acoustics from an engineering point of view. It has been a great encouragement to be a part of your section.

I also want to express my appreciation to the members of CAHR for letting me use the listening booths and equipment to perform the listening tests, and of course all test subjects for their patience during the listening tests.

Last but not least, to my family and friends who have been enthusiastic in listening to my thoughts and for your great help, thanks a lot!

Contents

| | |
|--|-----------|
| Abstract | ii |
| Acknowledgements | iv |
| Contents | vi |
| 1 Introduction | 1 |
| 1.1 Motivation | 1 |
| 1.2 Background | 4 |
| 1.3 Objectives | 10 |
| 2 Theory | 11 |
| 2.1 Reflections in rooms | 11 |
| 2.2 Poisson Process | 13 |
| 2.3 Broadening and absorption | 14 |
| 2.4 Mean free path | 14 |
| 2.5 Criteria for echo detection | 15 |
| 2.5.1 Criteria by Dietsch and Kraak | 15 |
| 2.5.2 Criteria based on the decay curve | 17 |
| 3 Method | 21 |
| 3.1 Simulation | 21 |
| 3.1.1 Arrival of reflections | 22 |
| 3.1.2 Reduction of the sound pressure | 23 |
| 3.1.3 Reflectogram generation for separate frequency bands | 24 |
| 3.1.4 Properties of the rooms | 26 |
| 3.1.5 Modified impulse responses | 28 |
| 3.1.6 Objective investigation of the criteria | 29 |
| 3.2 Listening test | 30 |
| 3.2.1 Test design | 30 |
| 3.2.2 Test 1 | 32 |
| 3.2.3 Test 2 | 33 |
| 3.2.4 Analysis method | 34 |

| | | |
|----------|---|-----------|
| 4 | Results | 39 |
| 4.1 | Monte Carlo Simulation | 39 |
| 4.2 | Repeatability | 42 |
| 4.3 | Dependence of delay time | 44 |
| 4.3.1 | Cumulative distribution | 45 |
| 4.4 | Dependence on volume and reverberation time | 48 |
| 4.4.1 | Thresholds for the reflection strength | 48 |
| 4.4.2 | Thresholds for the criteria | 49 |
| 4.4.3 | Cumulative distribution | 54 |
| 5 | Discussion | 57 |
| 5.1 | Summary | 57 |
| 5.2 | Conclusions | 61 |
| 5.3 | Further considerations | 62 |
| | Physical Constants | 64 |
| | Symbols | 67 |
| A | Listening tests | 73 |
| A.1 | Instruction for the listening test | 73 |

Chapter 1

Introduction

1.1 Motivation

In room acoustics, the room introduces additional paths to transmit sound energy between the source and the receiver. Ideally, this should be done in such a way that the signal, which could be music or speech, is enhanced and that the properties of the room is advantageous for the type of message that is to be transmitted. What is advantageous depends on the purpose of the room, for example symphonic music is in need of different acoustical qualities than speech.

Problems with strong reflections

One important aspect for all acoustic situations is that the sound field should be free from strong reflections or echoes, since this might result in a disturbed sound picture. In case of speech this can lead to a decrease in speech intelligibility, which in turn can destroy the experience of a theatre performance or make an audience miss information in a lecture. In terms of music, the presence of strong reflections in general worsen the quality of the listening experience. However, the strong reflections that arrive with a short time difference compared to the first sound that reaches the receiver is not problematic. These early reflection has large amplitude, but instead of disturbing the sound these rather more enhance it [1].

Another important view of the echo-related problem is the experience of the performers themselves. For an orchestra to collaborate in a successful way the musicians need to clearly hear themselves as well as their colleagues, and therefore it is important that no extra attention is given to artefacts such as strong reflections. One example of a concert hall that is suffering from echo problems is Örebro Concert Hall, Sweden [2]. The concert hall was investigated by the author together with study colleagues during a course in architectural acoustics in spring 2013. One property that was found to be missing in the hall was the ability for the musicians to trust the response of the room and the problems were expressed from one of the bassoonists as

"The room lives its own life, and it is hard to know the response of your performance. The acoustics of the hall is not preserving the sound".

The hall was shown to have serious echo problems, which was taken as the main reason to this opinion. The problematic acoustics of this hall lead to curiosity and interest in the properties and detection of echoes, which is one important motive for this thesis.

In the end, the ability for an ensemble to collaborate is affecting the over all experience for all participants, both audience, sound engineers and musicians, during a concert or a production. An echo-situation is also a problem for the talker in a lecture hall. The talker will then hear his/hers your own voice coming back from the other side of the room, and this causes confusion and makes him/her lose concentration, which in turn also affects the audience.

Room impulse response

The room impulse response is a useful fingerprint of the room, containing much information about its properties, and therefore this is measured when investigating the acoustics of a room. This is done in the case of performing spaces by placing a omnidirectional loudspeaker on stage and a omnidirectional microphone somewhere in the audience area. The signal played is a signal with a flat spectrum, like a sine sweep or a gun-shot.

The energy parameters like Reverberation Time (RT), Early Decay Time (EDT) and Clarity (C80) are some of the parameters that can be found from the impulse response and these are often used to characterize the room. These can be found

by standardized measuring methods (DS-EN 3382-1) [3] and the parameters are therefore commonly used in room acoustic design to construct rooms with desirable room characteristics. The value of the parameters are often averaged over several measurement positions. The suggested number of positions is given by the standard and depends on the number of seats in the audience. The room impulse response should also be measured at three different positions for the loudspeaker placed symmetrically on the stage area.

The discussed parameters are similar in the sense that they describe how the sound energy is distributed over time. However, it is not only the energy over a large time interval that is characterizing the acoustics of the room, the temporal structure of the received energy will also have an influence. If the number of reflections that reaches the receiver suddenly increases at some position in time there is a risk of an echo, i.e., a reflection with sufficient strength and delay time to be perceived as a distinct echo [4]. These are properties of the impulse response that needs to be taken under consideration in order to reveal echo problems, and therefore these are the properties that have been investigated in this project.

Benefits of simulation method

A statistical model will be used to simulate impulse responses, which assumes a diffuse sound field, free from strong reflections. Using this approach, the room parameters such as absorption and volume can easily be changed and therefore the reverberation time is a control variable. In order to investigate the outcome of a strong reflection, the impulse response can be adjusted such that a strong reflection can occur any target point in time and with any target strength. This is a strong benefit with this modelling technique, since the time of the reflection and reflection strength are controlled variables. Another advantage with using this model is the ability to investigate the ensemble statistics of the parameters. A measured impulse response can be regarded as a outcome of a certain distribution, i.e., it will be different depending on the position of the source and receiver. This is due to randomness in the arrival of the reflections, but only within certain limits for a fixed volume and reverberation time. If many impulse responses can be realized, the robustness of the estimate can be evaluated. Consequently the estimate can be related to a certain quality which is desirable for evaluation. In

this thesis, Monte Carlo simulation will be used to investigate these aspects.

An advantage with using synthesized signals is that it is easy to change the properties of the sound field and then investigate how the new properties changes the subjective experience or a parameter. A disadvantage is that the model will show deviations from the real sound field, there is a risk that the sound field is perceived as unreal and synthetic. A lot of information about the sound field is also lost, for example the directional information.

1.2 Background

The transmission of sound between the source and receiver in a room is a good start to investigate the properties of the sound field of a room. The sound that reaches the receiver can be explained as a mixture of sound rays, which all originate from the source and each reflection is a linear combination of the original sound from the source. The geometry and size of the room and the absorption of the walls determines the strength of the reflections and from which direction they reach the receiver. It therefore modifies the original sound emitted from the source. If the properties of the room causes the reflections to become a audible repetition of the original sound, an echo has occurred. This can happen if many rays arrives at the receiver within a small time interval, but a strong reflection is not necessarily perceived as an echo.

The detection of a strong reflection depends on the delay time (t_{delay}), which is the time difference between the direct sound and the time of the reflection. If the delay time long enough, and if the strength of the reflection is above a certain threshold, the reflection is detectable. This is due to temporal masking, which is a property of our auditory system. Masking is a phenomena that can be described as *the process by which the threshold of audibility of one sound is raised by the presence of another (masking) sound* [5] and this takes place both in the frequency and time domain. In the frequency domain this occurs if two signals have similar frequency, and in time domain, if two sound events occur within a short time-interval [6]. It is the temporal resolution of the hearing system that limits our possibility to hear sounds presented at different times and this is a

important factor in echo perception. The masker can for example be the direct sound, but also neighbouring reflections to the strong reflection works as maskers. In the case of echo detection, temporal masking is an advantage since it smears out the sound. In other aspects, for example in speech perception, temporal masking affects the possibility to perceive formant transitions in speech. [7]

However, reflections arriving with a delay time shorter than 50 ms will not cause an echo [1], but rather increase the loudness of the direct sound. It also makes the source sound more extended and these early reflections are regarded as useful.

Previous work

One early study of the critical echo level was performed by Haas in 1951 [8]. In one of his experiments the subjects were presented to a speech signal with a speaking rate of 5.3 syllables per second in a room with reverberation time of 0.8 s. A second loudspeaker presented the same speech signal but with a certain delay time and level. The delay time is the time difference between the presented signals. The results (seen in Figure 1.1) shows that the percentage of subjects that was disturbed by the delayed signal increased as the delay time increases, as well as for increasing level of the echo [8].

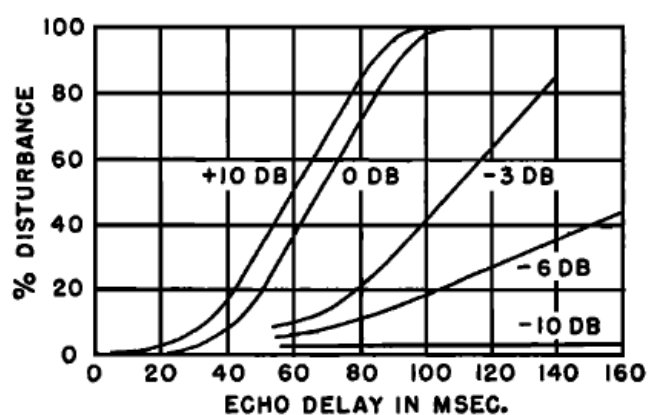


FIGURE 1.1: Percentage of annoyed test subjects as a function of delay time. For different strength of the reflection [8]

The 50% point of the curve in Figure 1.1 is referred to as the critical delay time, i.e., the delay time for the reflected signal at which 50% of the test subjects were disturbed by the reflection. By changing the reverberation time this threshold shifts, for longer reverberation times the critical delay time increases, which means that the subjects are less sensitive in reverberant conditions. Changing the speaker rate also changed the critical delay time, for slower speaker rate the subjects are less sensitive to the reflection and the critical delay time was larger. The type of signal is also of importance for the annoyance level, for example speech has a lower threshold than music [9].

Experiments were also carried out by Burgtorf, Oehlschlägel and Seraphim in 1961 [1]. Using only two reflections, the direct sound and one reflection with delay time of t_0 (the time difference between the first sound and the strong reflection), the audibility threshold ΔL for a speech signal was found to be

$$\Delta L = -0.575t_0 - 6 \text{ dB} \quad (1.1)$$

ΔL is the strength of the reflection relative the direct sound. t_0 is in milliseconds.

Using the direct sound and only adding one reflection might seem like a unfortunate approach since the impulse response consists of contribution from a large number of reflection, but it was shown by Seraphim [10] that adding more reflections doesn't change the threshold. Figure 1.2 shows his results were one, two, three and for reflections with the same strength were presented together with the direct sound. The vertical lines in Figure 1.2 are these reflections and it can be seen that the threshold of detection is unchanged as long as there are reflections occupies in time. Not until the delay time is longer than the last reflection, the detection threshold drops.

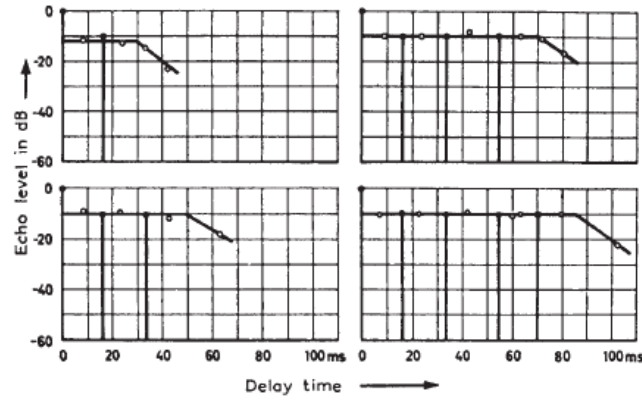
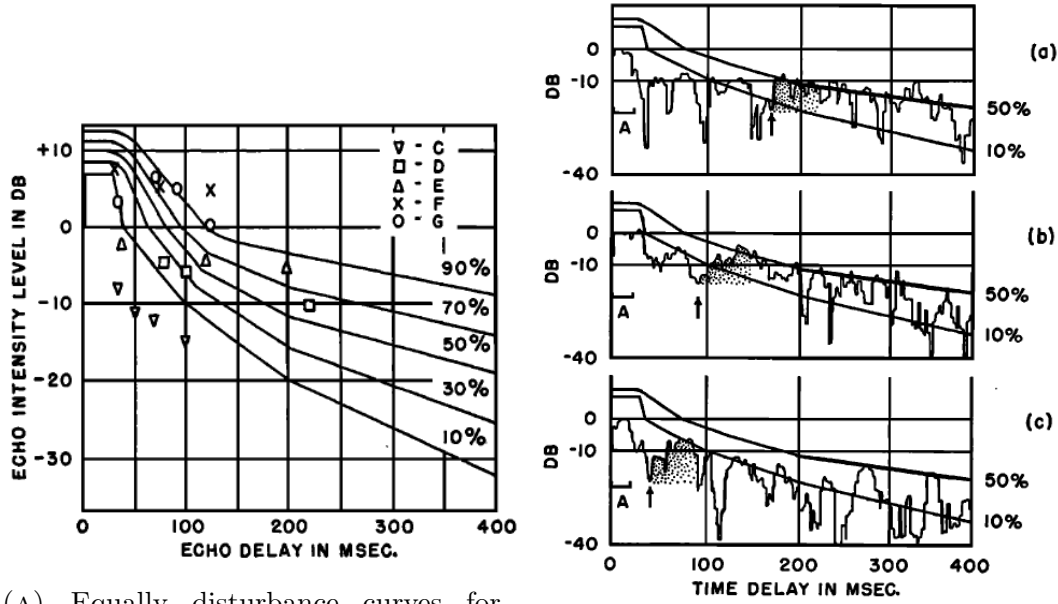


FIGURE 1.2: Level of the detection threshold for various number of reflections [1]

Bolt and Doak [11] used the results from Haas to give a suggested criteria for echo annoyance. By extrapolation, the equally disturbance curves was plotted with the echo level as a function of delay time (see Figure 1.3a). These curves were compared with the envelope of an impulse response in a theatre where a distinct echo could be heard at one position in the room. Figure 1.3b shows comparison for three different positions of the receiver in this theatre, where Figure *a* in 1.3b is the results for the problematic position. The dotted area in the signal indicates the echo. As the receiver was placed at other positions in the theatre, the time of the distinct echo decreases (Figure *b* and *c* in 1.3b). It is then no longer acknowledged as a distinct echo on a 10% annoyance level.



(A) Equally disturbance curves for different level of annoyance as a function of time

(B) Equal disturbance curves compared with the envelope of impulse responses

FIGURE 1.3: Extrapolation of results found by Haas and comparison with measurements. *a* is at a position with echo disturbance. In *b* and *c* the microphone has been placed closer to the rear wall.[11]

The results from subjective tests shows a good correlation, but as pointed out by the authors this criteria is tentative and more experimental work had to be done.

A more recent criteria was proposed by Dietsch and Kraak in 1986 [12] and this was based on the center-of-gravity time of the impulse response. Thresholds for music and speech were investigated using both synthetic and measured impulse responses, but the criteria was shown by Løvstad, to give poor agreement with impulse responses generated by simulations in ODEON [13]. This lack of agreement with the subjective experience was also confirmed by the author to this thesis when using ODEON to simulate the acoustics of the concert hall in Örebro. The criteria will be investigated further in this project and therefore it will be explained more in detail in section 2.5.1.

A new parameter for strong reflections

The first milliseconds of an impulse response consists of the early reflections. These are clearly separated in time and they are important for the subjective experience of the room. But quite soon the sound field becomes diffuse, since the number of reflections collected by a receiver grows fast with time and they arrive from arbitrary directions. This can be characterized by the energy decay curve, which is the total energy remaining in the room as a function of time. For a diffuse sound field the energy decay curve (EDC) is smooth, because there are no sudden increase of the energy picked up by the receiver. But for a non-diffuse sound field, the energy decay curve becomes uneven and it has sudden changes in its slope. This roughness can be indicated by the slope ratio[14].

The slope ratio is a measure that has recently been defined by Jeong et al. [14]. This is the instantaneous slope of the energy decay curve, normalized with its average slope and has previously been used to indicate at what time a impulse response has become diffuse. But the slope ratio might also be useful in detection of strong reflection that could be perceived as an echo, and this is the main question of this project. The unevenness of the decay curve is an indicator of the diffusivity in the room, therefore this could be used to identify the time of strong reflection and in turn reveal echo problems.

1.3 Objectives

The main focus in this report is to investigate the perceptual experience of strong reflections in an impulse response which in other respects is free from echoes. With subjective experience is meant the detection threshold, which will be expressed in the strength of the reflection and other measures that can be found from the impulse response. These are explained more in Section 2.5.1 and 2.5.2. At first, a realistic model of an impulse response shall be developed, which is explained more in Chapter 3 and thereafter the behaviour of the different parameters can be investigated (Section 4.1). Finally, listening tests will be carried out, in order to find the detection threshold of strong reflections for the different parameters. The delay time of the reflection, the volume of the room and the reverberation time will be changed during the listening tests. This is described more in detail in Section 3.2.

Subsequently, the project will address the following questions:

- What properties of the energy decay curve can reveal a strong reflection?
- How can a realistic impulse response be modelled?
- How can the impulse response be modified to include strong reflections in a trustworthy way?
- What is the behaviour of the objective measures based on the impulse responses and how does this behaviour depend on volume, reverberation time and delay time?
- What is the detection threshold expressed in these parameters and how does this depend on reverberation time, volume and delay time?
- How is the threshold related to the energy of reflections in its temporal neighbourhood and to the smoothness of the impulse response?
- Which objective parameter reflects a problem with strong reflections the best?

Chapter 2

Theory

It is reasonable to suggest that a parameter for a strong reflection can be found from the room impulse response. Therefore, the impulse response is studied closer and this chapter will describe the concepts and theory behind the statistical model that has been used. Apart from that, the existing echo-criteria by Dietsch and Kraak will be described, together with a suggested criteria using a new approach.

2.1 Reflections in rooms

It is assumed that the sound emitted from a omnidirectional source is a tone burst, or ideally a dirac-delta pulse, which results in a spherical wavefront in a three dimensional coordinate system. This can also be expressed as a infinite number of sound rays equally distributed over all angles in such a space, where each ray represents a plane wave section of the spherical wave. In the reminder of this thesis, the plane wave will be described in terms of sound rays, which obeys the same rules as light rays when interacting with different media or surfaces. After a number of reflections to the walls of the room, the sound ray will finally reach the receiver. It will then have a certain *reflection order*, i.e., the ray has either been reflected 0, 1 or n times before reaching the receiver.

The first sound ray that arrives at the receiver is called the direct sound, that is reflection order = 0, and it travels to the receiver without any interaction with the

walls. After the direct sound, rays with higher reflection order will arrive and the average reflection order increases with time.

Since the sound field is described as rays and obeys the law of specular reflection, each reflection that reaches the receiver can be represented with its mirror source of the same order as the reflection order. The time of arrival of the reflections can then be found by first identifying the mirror sources and then find the delay between then mirror source and the receiver. This is a hard task for complex geometries of the room, but instead, an *average reflection density* can be found by averaging the number of reflections that are received at a position in a rectangular room. [1].

The reflection density at an arbitrary position in a rectangular room with volume V is then given by

$$\frac{dN}{dt} = 4\pi \frac{c^3 t^2}{V} \quad (2.1)$$

where N is the number of reflections and c is the speed of sound [1]. The average reflection density is inversely proportional to the volume, i.e., in a smaller room the number of reflections per second is larger for a certain position in time. Furthermore, the reflection density is proportional to t^2 which means that the first reflections are more separated in time, compared to reflections that arrive later.

At some point the reflections are no longer identified as different sound events, but contributes to the sound field as a reverberation tail. This transition time is dependent on the integration time in the auditory system of the ears. When the reflection density becomes too large, the human ear does not have the ability to separate the different reflections. The mixing time, t_{mix} , could work as an objectively measure on the transition time, although the definition of mixing time differ between the authors [15]. Reichardt [16] suggested the mixing time to be

$$t_{mix} = \sqrt{V} \quad (2.2)$$

where t_{mix} in ms and volume V in m^3 . This corresponds to a time resolution of $\frac{dN}{dt} = 507 \text{ refl/s}$, where Equation 2.1 have been used. In this project, the range of

the volumes used is $V = 450 \text{ m}^3$ to $V = 1050 \text{ m}^3$ which results in a mixing time of 21 ms and 32 ms. The time of the reflection will take place later than 100 ms later compared to the direct sounds, which means that we are in the range of the reverberation tail.

2.2 Poisson Process

The reflection density mentioned in the previous section expresses the average number of reflections as a function of time. In reality, the number of reflections per second is a discrete random number. To simulate a realistic process of reflections reaching the receiver this can be implemented with a Poisson process [14].

The Poisson distribution describes the probability that X events takes place in a time interval with a given length t_a . The events are independent on each other, and can take place at any time, which means that the duration since the last event took place does not influence the probability for a new event to occur. The probability distribution is [17]

$$p_X(k) = \frac{(\lambda t_a)^k}{k!} e^{-\lambda t_a}, \quad k = 0, 1, 2, \dots \quad (2.3)$$

if $X \in \text{Po}(\lambda t)$

where λ is the average number of events for a time interval t_a . In the case of room acoustics, an event is an arrival of a reflection at the receiver position which means that λ in this case is the reflection density, i.e.,

$$\lambda = 4\pi \frac{c^3 t_a^2}{V} \quad (2.4)$$

The probability that X reflections arrives at the receiver is then Poisson distributed. If the average number of events for the time interval is known (λ), the stochastic variable has the distribution $X \in \text{Po}(\lambda t)$.

As previously mentioned, each time a reflection arrives at the receiver can be regarded as an event, but the average reflection density is increasing according to Equation 2.1 and therefore the Poisson process has an increasing intensity λ . Figure 2.1 shows a reflectogram, where the reflections following a Poisson process

using a volume of $V = 750 \text{ m}^3$. In intervals of $\Delta t = 2.5 \text{ ms}$ the reflection density is a constant and the reflections are randomly spread in this time interval. In the beginning of this process, the probability of an event is smaller and the number of reflections for a certain time interval is low. As time goes, the intensity increases and the time between two reflections gets smaller.

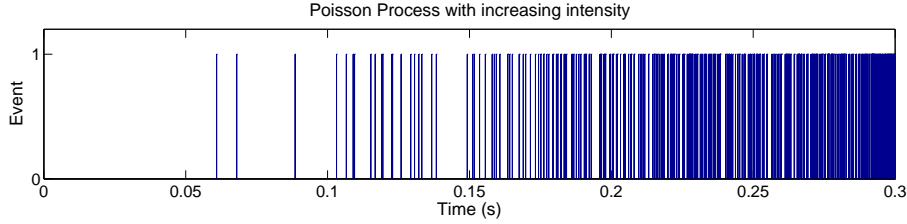


FIGURE 2.1: Poisson process with increasing intensity. The value 1 indicates that an event occurred at this position in time.

2.3 Broadening and absorption

For each wall-reflection the spectral content of the signal is changed. Often, the absorption at the wall is larger for higher frequencies and thus the reflection at the wall works as an effective low-pass filter. The reflected pulse will then be more limited in its frequency content, compared with the original signal which in the time-domain results in a broadening of the pulse. As the reflection order increases, the model expands as a filter chain of equally many low-pass filters concatenated after another. Due to the linear properties of the low-pass filter each filter in chain can be seen as broadening the pulse independently of another and thus the resulting pulse is broadened in proportion to reflection order. This leads to an increasing overlap of the reflections. For later time of arrivals, the reflections will have a larger overlap, both due to the increase in reflection density and the fact that the reflections are broader [15].

2.4 Mean free path

In this project the impulse response will be modelled using a statistical model. This assumes that the sound field is diffuse, i.e., there is no direction in the room where the average intensity is larger than in other directions. In other words the sound rays are distributed in all angles with the same probability.

With the assumption of a diffuse sound field, various properties of the sound field can be used. For example the *mean free path*, which is the average distance a sound ray will travel between the interactions with the walls [1]. This is determined by the volume and the surface of the room as

$$l_m = \frac{4V}{S} \quad (2.5)$$

2.5 Criteria for echo detection

As mentioned above the impulse response contains information about the received energy distributed in time, but also temporary differences which could be detectable. Dietsch and Kraak [12] suggested one measure 1986, based on the center-of-gravity time of the impulse response. The and Kraak measure will serve as a reference through out this thesis in addition to two proposed measures that are based on the slope ratio[14].

2.5.1 Criteria by Dietsch and Kraak

Dietsch and Kraak [12] have proposed an echo detection criteria based on the center-of-gravity time of an impulse response. The criteria has been investigated for both speech an music.

The center-of-gravity time of an impulse response ($p(t)$) normalized with its total energy is

$$t_S(\tau) = \frac{\int_{t=0}^{\tau} t |p(t)|^n dt}{\int_{t=0}^{\tau} |p(t)|^n dt} \quad (2.6)$$

where $n = 2$ for the traditional center time formula. For this criteria, a suitable value of n was found to be $n = 2/3$ for speech and for $n = 1$ music. Equation 2.6 is referred to as the built-up function and this is plotted in Figure 2.2 for an impulse response with and without a strong reflection for $n = 2/3$. The time of the strong reflection is 160 ms and this results in an increase in the built up function for this impulse response around this time.

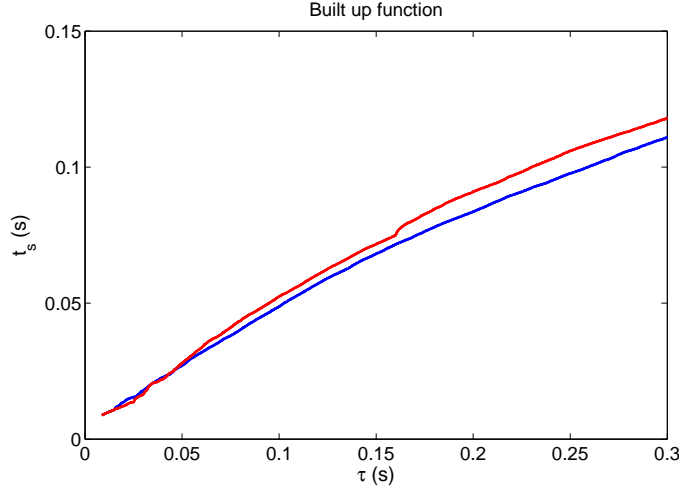


FIGURE 2.2: The built-up function 2.6 with $n = 2/3$. Blue curve is the built up function of a diffuse impulse response, while the red curve has a strong reflection at $\tau = 160$ ms. Different realizations of the Poisson process has been used in the two plots.

The criteria EK by Dietch and Kraak is finally found by taking the ratio

$$EK(\tau) = \frac{t_S(\tau) - t_S(\tau - \Delta\tau_E)}{\Delta\tau_E} \quad (2.7)$$

A suitable value of $\Delta\tau_E$ was found from listening tests and these are $\Delta\tau_E = 9$ ms for speech and 14 ms for music.

The criteria is plotted in Figure 2.3 for the same impulse responses as above and the peak around $t = 160$ ms can be seen. In this figure the criteria with $n = 2/3$ and $\Delta\tau = 9$ ms has been used, i.e., the criteria for speech. This will be the case also for further investigations of this parameter.

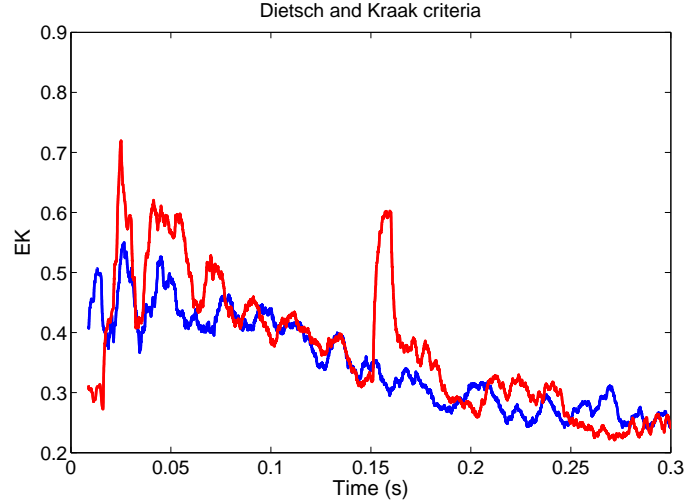


FIGURE 2.3: Echo criteria by Dietsch and Kraak. The blue line is the EK-curve without any strong reflections and the red curve is the result of a modified impulse response.

The echo annoyance threshold expressed in EK is given as the limit at which a certain percent (10% and 50%) of the test subjects were annoyed by the echo. If the ratio in Equation 2.7 is larger than $EK_{N\%}$, then $N\%$ of the listeners will be annoyed by the echo. The limits of $EK_{N\%}$ are found in Table 2.1.

| Test signal | Music | Speech |
|-------------|-------|--------|
| $EK_{10\%}$ | 1.5 | 0.9 |
| $EK_{50\%}$ | 1.8 | 1.0 |

TABLE 2.1: Threshold of the criteria EK for different annoyance levels and signals

2.5.2 Criteria based on the decay curve

It has been shown by Schroeder [18] that the energy decay curve averaged by many simulations/measurements converges to a backward integrated decay curve. The Energy Decay Curve $EDC(t)$ can then be calculated as

$$EDC(t) = \int_t^{\infty} p^2(t) dt \quad (2.8)$$

where $p(t)$ is the impulse response.

In the case of a diffuse impulse response, as discussed above, the decay curve is smooth and without larger steps, as depicted in Figure 2.4a. If the reflection density is suddenly increased in a certain position in time, due to a unfavourable geometry and/or specular reflecting surfaces, this will cause a change in the slope of the decay curve. With increasing strength of the reflection, the slope at this position becomes larger. This is illustrated in Figure 2.4b where the reflection density is increased at $t = 160$ ms.

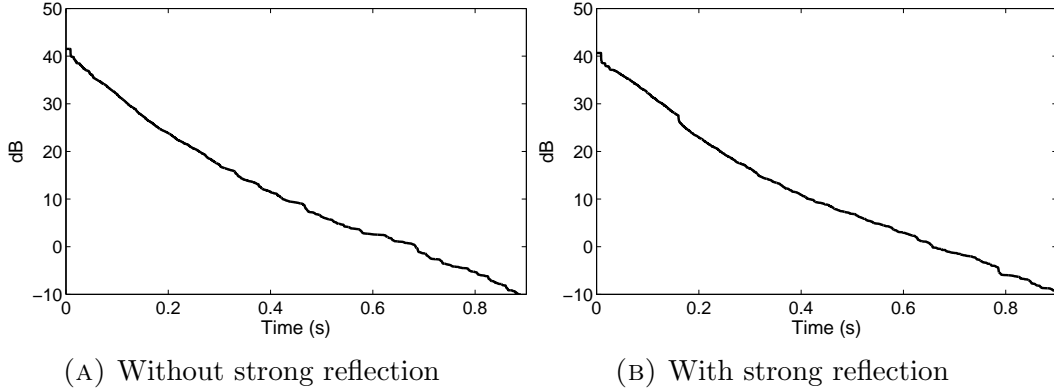


FIGURE 2.4: Energy decay curve (EDC) calculated using Equation 2.8. The right figure has a increase in reflection density at $t = 160$ ms

By taking the instantaneous slope of the decay curve L' and normalize this with the average slope of the decay curve \bar{L}' , a function that indicated distinct reflections is found. This is the slope ratio R_{slope} and can be calculated as [14]

$$L' = \frac{L(t + \Delta t) - L(t)}{\Delta t} \quad (2.9)$$

$$R_{slope} = \frac{L'}{\bar{L}'} \quad (2.10)$$

where Δt is the time difference between two samples and \bar{L}' is the mean slope of the decay curve over the range from 0 to -40 dB.

Figure 2.5a shows the slope ratio for an diffuse impulse response. In case of a strong reflection, the slope of the decay curve is suddenly increased and this results in a peak in the slope ratio, which can be seen in Figure 2.5b.

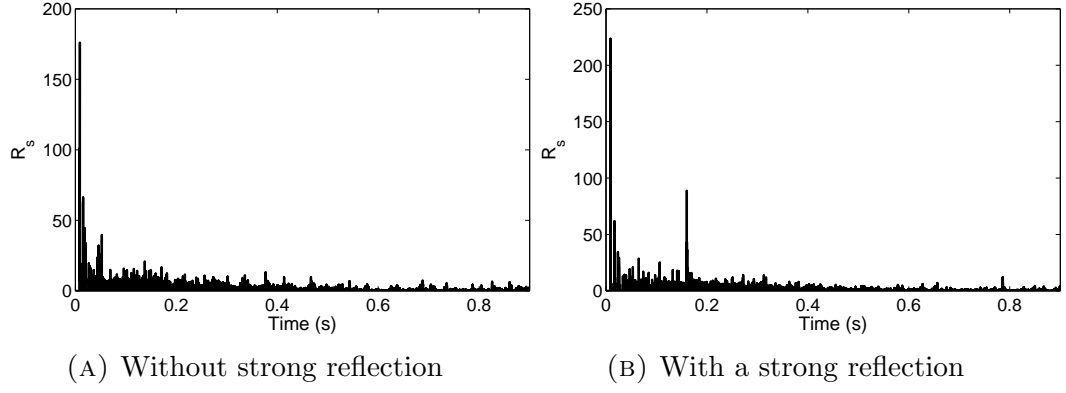
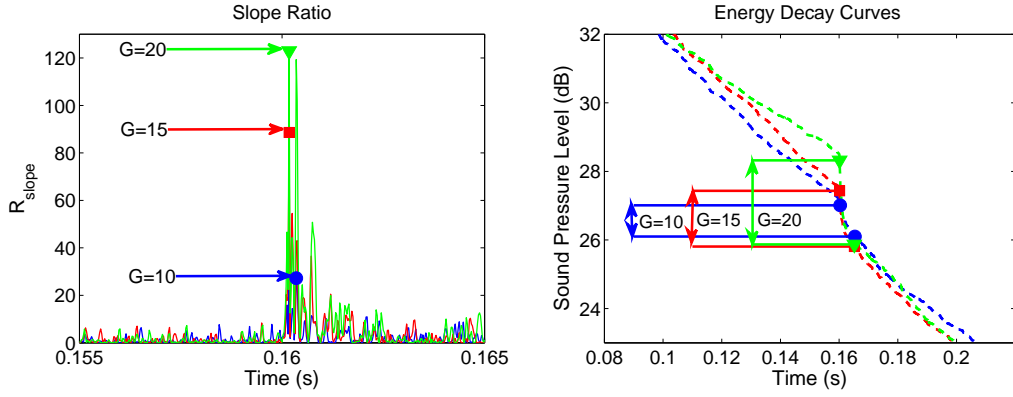


FIGURE 2.5: Slope Ratio R_s calculated using Equation 2.9 and 2.10

It can be seen in Figure 2.5 that the slope ratio is large also for shorter delay times. This is due to the low reflection density in the beginning of the impulse response. The time difference between the reflections are then relatively long and each reflection has a large influence on the received energy. Since the detection threshold in delay time is often said to be 50 ms the peaks in slope ratio before this time is not taken into consideration when finding the maximum value of R_{slope} , which is notated R_{slope}^{max} .

Another property that differs between the impulse response free from a strong reflection and the one containing a strong reflection, is the total drop in energy around the time of the strong reflection. This drop is notated ΔEDC . The comparison between the two measures can be seen in figure 2.6 where the peak of the slope ratio and the drop in the energy decay curve has been zoomed into. The average reflection density is increased with a factor $G = 10, 15$ and 20 at a time of $t = 160$ ms. Figure 2.6b shows the corresponding energy decay curve, and it can be seen that the drop in energy increases with the strength of the reflection.



(A) Slope ratio at the time of the strong reflection (B) Energy Decay Curve at the time of the strong reflection

FIGURE 2.6: Closer look at the energy decay curve and slope ratio in case of a reflection with various strength G . The increase in reflection density is for the blue curve 10, the red curve 15 and the green curve 20.

For a time interval of 5 ms around the peak, the slope ratio has significantly larger values compared to other positions in time (2.6a). The suggested parameter ΔEDC is therefore found by

- 1) Identifying the time t_{max} for the maximum slope as the peak value of R_{slope}
- 2) Taking the difference in the EDC over the interval $t_{max} + 5$ ms

Chapter 3

Method

Under the assumption that the sound field is diffuse, the statistical properties of the sound field described in the previous section can be used. It can therefore be assumed that the average reflection density increases according to Equation 3.1, the event history of the reflections follows a Poisson process and the mean free path l_m can be used. In this chapter it will be explained how these assumptions can be implemented to simulate a diffuse impulse response. Furthermore it will be explained how to modify the impulse responses so that they contain a strong reflection at a certain position in time. It will also be described how the echo criteria can be studied objectively and finally how the subjective thresholds can be found from a listening test.

3.1 Simulation

The impulse response is characterized by the time of arrival and the strength of the reflection. As time goes, the strength of the reflections reaching the receiver has in general a lower intensity, since it has a larger reflection order, which means that more energy has been absorbed, and it has travelled a longer distance. However, the reflection density is increasing as discussed in the previous section, resulting in a contribution of sound energy at the receiving point.

3.1.1 Arrival of reflections

The number of reflections ΔN that will arrive at an arbitrary position in the room in a certain time interval Δt is given by

$$\Delta N = 4\pi \frac{c^3}{V} t^2 \Delta t \quad (3.1)$$

where V is the volume and $c = 343 \text{ m/s}$ is the speed of sound. This equation was introduced in section 2.1. The sampling frequency is set to $f_s = 40 \text{ kHz}$ which gives a time interval of $\Delta t = \frac{1}{f_s} = 25 \text{ ms}$.

The time of arrival of the reflections at the receiver can be simulated with a Poisson process with an intensity that is changed corresponding to the reflection density, see section 2.2. Figure 3.1 shows the number of reflections that reaches the receiver for each sampling interval Δt . This can be compared with Figure 2.1 in Section 2.2, with the difference that the y-axis in this figure is the number of reflections within a sampling interval, compared to figure 2.1 where all reflections, or events, that arrives to the receiver within a sampling interval were displayed.

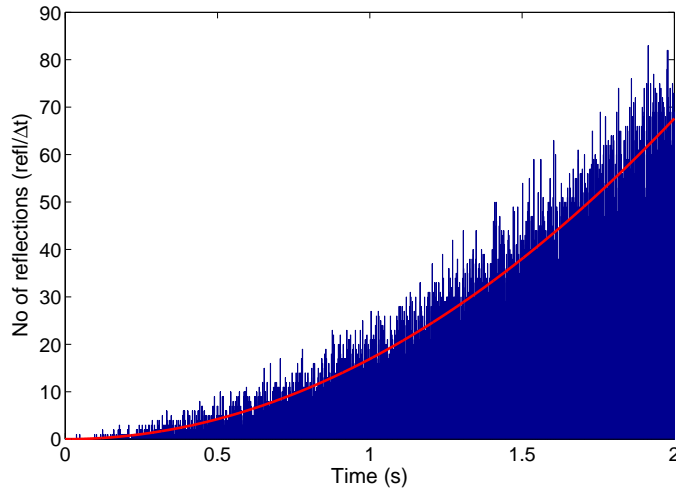


FIGURE 3.1: Number of reflections per time interval $\Delta t = 25 \text{ ms}$ for $V = 750 \text{ m}^3$ (blue bars). The red curve is the average reflection density.

Apart from the Poisson process, a direct sound is added to the diffuse sound by adding a pulse at a time of 9 ms, which corresponds to a distance of 3 m between the source and receiver. Consequently, the Poisson process starts immediately after 9 ms.

3.1.2 Reduction of the sound pressure

A reflection that arrives after a certain time t , has been reduced due to *air absorption*, *geometrical spread* and *absorption at the surfaces*.

Absorption at the surfaces

The distance travelled is of large importance and is simply $r = c \cdot t$. Together with the mean free path l_m (see Equation 2.5) the average number of reflections that the sound ray has gone through at this time, i.e., the reflection order N_r , can be found by dividing the distance travelled with the mean free path

$$N_r = \frac{r}{l_m} \quad (3.2)$$

Each time a sound ray interacts with a surface of the room the sound pressure is reduced by a factor $\sqrt{1 - \alpha}$ and after N_r reflections this factor becomes $\sqrt{(1 - \alpha)^{N_r}}$. For simplicity, α is the average absorption coefficient of the surfaces in the room.

Geometrical spread

Assuming that the source is omnidirectional, the sound pressure at a distance r from the source is given as

$$p(r) = \frac{j\omega\rho Q e^{j(\omega t - kr)}}{4\pi r} \quad (3.3)$$

where $j\omega Q$ is the volume acceleration of the source, Q is the volume velocity, ρ is the density of air, ω is the angular frequency and k is the wavenumber [19]. The reduction factor due to geometrical spread is then proportional to $1/r$. The proportional factor is the same for all frequencies, since the volume acceleration is supposed to be a constant. Furthermore, the strength of the response is adjusted in the end of the simulation and the proportional factor is therefore not of interest.

Air absorption

The air absorption α_{air} is dependent on the room temperature and the relative humidity. Here it is assumed that the temperature is $T = 20^\circ\text{C}$ and a relative humidity is 50%. The air absorption using these assumptions is found in Table 3.1 [21] and is taken into account through multiplying with a factor $10^{-\alpha_{air}r/20}$. α_{air}

has the unit [dB/m].

| Center frequency (Hz) | 63 | 125 | 250 | 500 | 1000 | 2000 | 4000 | 8000 |
|---|-------|-------|------|------|------|------|------|------|
| α_{air} (dB/m · 10 ⁻³) | 0.123 | 0.445 | 1.32 | 2.63 | 4.65 | 9.86 | 29.4 | 104 |

TABLE 3.1: Air absorption at 20° C and humidity 50%

All together, the amplitude of the sound pressure becomes proportional to

$$A(t_n) \propto \Delta N(t_n) \sqrt{(1 - \alpha)^{N_r}} \cdot 10^{-\alpha_{air} r / 20} / r \quad (3.4)$$

where $\Delta N(t_n)$ is the number of reflections in the time interval at t_n .

Effect of the head

Since the receiver in the listening test is a human, the head related transfer function for a diffuse field is taken into account. Table 3.2 [22] shows the effect of the human torso and head as the difference between the sound pressure level measured with or without the presence of the torso and head. This factor was only added to the impulse responses used in the listening tests. The receiver for the measured impulse response is assumed to be omnidirectional.

| Center frequency (Hz) | 63 | 125 | 250 | 500 | 1000 | 2000 | 4000 | 8000 |
|------------------------|----|-----|-----|-----|------|------|------|------|
| ΔL_{HRTF} , dB | 0 | 0 | 0 | 2 | 4 | 11 | 13 | 13 |

TABLE 3.2: The difference in sound pressure due to the effect of the head and torso in a diffuse sound field [22]

3.1.3 Reflectogram generation for separate frequency bands

The relative pressure amplitude that arrives at the receiver within a time interval Δt can be plotted as a function of time and this is referred to as the reflectogram. Since the absorption at the walls and in the air is frequency dependent the reflectogram has to be calculated for each octave band (n) separately. Figure 3.2 shows the result for $f = 1000$ Hz, that is Equation 3.4 where $\alpha_{1000 \text{ Hz}}$ and $\alpha_{air 1000 \text{ Hz}}$ have been used.

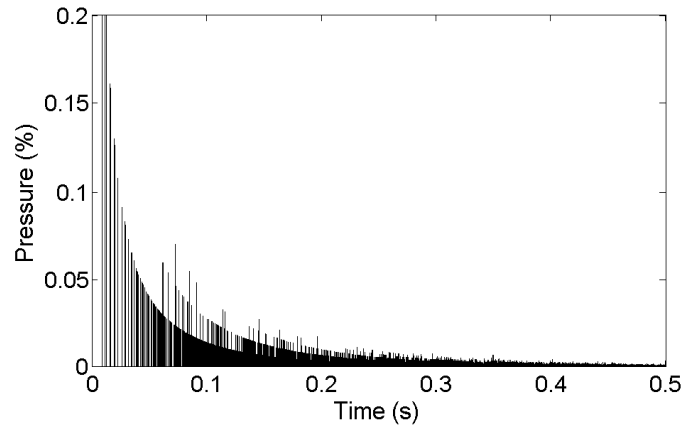


FIGURE 3.2: Reflectogram for $f = 1000$ Hz. The amplitude of the sound pressure relative the direct sound, $A(t_n)$

Each reflectogram was convolved with a signal (h), band limited to the corresponding octave band. Octave bands 63-8000 Hz was used in the simulation and the bandlimited signal h was represented by 6th order bandpass Butterworth filters, which are plotted in figure 3.3. The bandwidth of the filters were determined such that the upper cut-off frequency for one filter coincides with the lower cut-off frequency with the next filter.

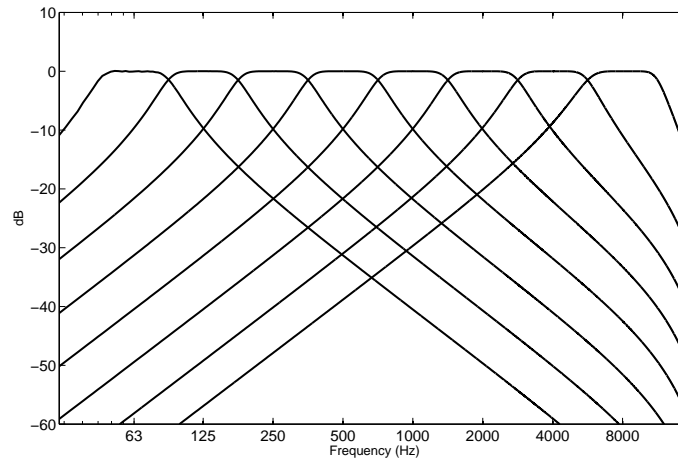


FIGURE 3.3: Filterbank of 6th order butterworth filters with center frequencies 63 – 8000Hz

The convolution with h was implemented in MATLAB by filtering the reflectogram with the corresponding butterworth filter. Figure 3.4 shows the responses for each band separately, i.e., the reflectogram has been filtered with the corresponding filter. These responses could then be added to get the total impulse response.

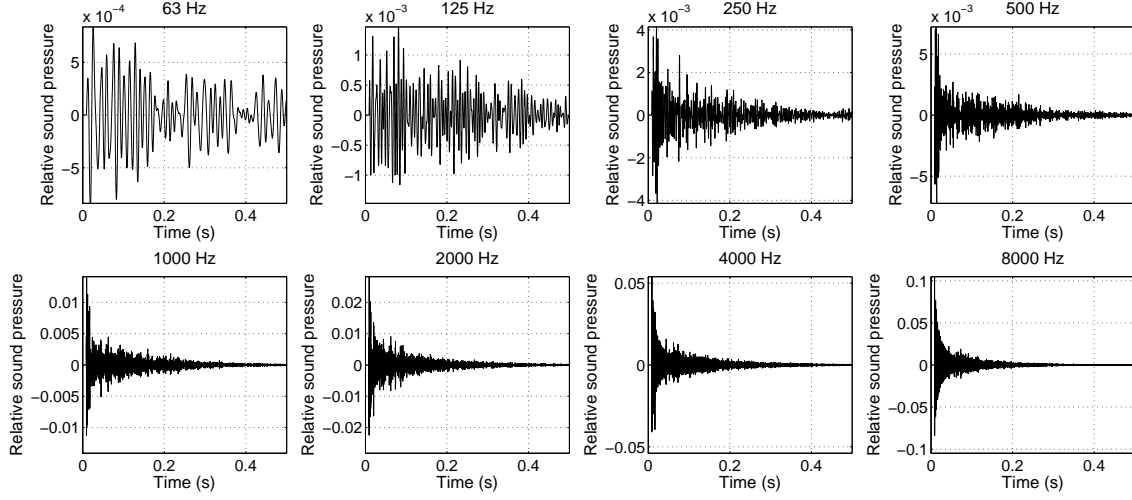


FIGURE 3.4: Impulse responses separate bands

Phase shift for different frequencies

Lower frequencies has, on average a larger phase difference to the direct sound, compared to higher frequencies. The average time delay for for a angular frequency ω is given by

$$t_{shift} = \frac{\bar{\varphi}}{\omega} \quad (3.5)$$

where $\bar{\varphi} = \angle(-i\omega t_n)$, t_n is the time of sample n and $\omega = 2\pi f$ [23]. Subsequently, before adding the responses, these were shifted a time t_{shift} , i.e., the time corresponding to the phase difference.

The total impulse response can now be written

$$p(t) \propto \sum_{band=1}^8 \sum_{n=1}^N A_{band}(t_n) \delta(t - t_n - t_{shift_{band}}) * h_{band}(t) \quad (3.6)$$

where $A_{band}(t_n)$ is the amplitude of the reflection at time t_n for a certain frequency band, δ is the dirac delta function and N is the number of samples.

3.1.4 Properties of the rooms

In the first of the two listening tests, the room has the dimensions $5 \times 10 \times 15$ which gives the volume $V = 750 \text{ m}^3$ and a total surface $S = 550 \text{ m}^2$. The material of the walls and ceiling is plaster and the floor area had an absorption

coefficient corresponding to upholstered seats. Absorption coefficients for these materials were found from ODEON where material no 4003 [Bobran, 1973] and 11006 [Beranek, 196] were used. In Table 3.3 the absorption coefficients are listed together with the average absorption of the surfaces of the room.

| Frequency Hz) | 63 | 125 | 250 | 500 | 1000 | 2000 | 4000 | 8000 |
|--------------------|------|------|------|------|------|------|------|------|
| $\alpha_{plaster}$ | 0.02 | 0.02 | 0.03 | 0.04 | 0.05 | 0.07 | 0.08 | 0.08 |
| α_{seats} | 0.44 | 0.44 | 0.60 | 0.77 | 0.89 | 0.82 | 0.70 | 0.70 |
| $\bar{\alpha}$ | 0.13 | 0.13 | 0.19 | 0.24 | 0.28 | 0.27 | 0.25 | 0.25 |

TABLE 3.3: Absorption coefficients of the surfaces used in the first listening test

Using this absorption and volume results in a reverberation time of $RT = 0.93\text{ s}$ using Eyring's formula [1]

$$T = \frac{0.16V}{S\alpha^* + 4mV} \quad (3.7)$$

where $\alpha^* = \ln\left(\frac{1}{1-\bar{\alpha}}\right)$ and $\bar{\alpha} = \frac{1}{S} \sum_i \alpha_i S_i$. S is the total surface of the room and V is the volume of the room. m is the air absorption coefficient, which is found from Table 3.1. But this is the coefficient with e , the base of the natural logarithm, as a base. The air absorption coefficient m was therefore found as

$$10^{-\alpha_{air}ct/20} = e^{-mct/2} \rightarrow m = \frac{1}{10} \ln(10) \cdot \alpha_{air} \quad (3.8)$$

where α_{air} is the value for air absorption found in Table 3.1.

In a second listening test different volumes and reverberation times were used. The reverberation time was changed by a scaling of the absorption coefficient and both the absorption coefficient for the floor and the walls were subsequently changed with the same factor. The volume remained unchanged. When changing the volume only, the absorption coefficient also had to be changed in order to keep the reverberation time constant. The dimensions of the room was kept the same. The average absorption coefficient for these room set-up is found in Table 3.4

| Volume (m ³) | RT (s) | Frequency (Hz) | | | | | | | |
|--------------------------|--------|----------------|------|------|------|------|------|------|------|
| | | 63 | 125 | 250 | 500 | 1000 | 2000 | 4000 | 8000 |
| 450 | 0.93 | 0.11 | 0.11 | 0.16 | 0.21 | 0.24 | 0.24 | 0.21 | 0.21 |
| 750 | 0.56 | 0.21 | 0.21 | 0.29 | 0.37 | 0.42 | 0.42 | 0.39 | 0.43 |
| | 0.93 | 0.13 | 0.13 | 0.19 | 0.24 | 0.28 | 0.27 | 0.25 | 0.25 |
| | 1.30 | 0.10 | 0.10 | 0.14 | 0.18 | 0.21 | 0.20 | 0.18 | 0.15 |
| 1050 | 0.93 | 0.15 | 0.15 | 0.21 | 0.26 | 0.31 | 0.30 | 0.27 | 0.27 |

TABLE 3.4: Average absorption coefficients of the surfaces for the different room set-ups

3.1.5 Modified impulse responses

For certain delay times the reflection density was modified in order to simulate a strong reflection. Instead of using the number of reflections from the Poisson process in a time interval, Equation 3.1 was used and this average number of reflection for one time interval was multiplied with a factor G . The value of α used for this time interval was α_{walls} from Table 3.3, since it is more probable that a strong reflection is due to a reflection at a hard surface. Figure 3.6 shows three modified reflectograms with different strength of the reflection and figure 3.5 shows corresponding impulse responses. The reflection density has been increased with a factor $G = 10, 15$ and 20 .

As discussed in section 1.2, it is the strength of the reflection that is important for the perception, and not the number of strong reflections [10]. It was therefore decided to only change the reflection density in on sample interval.

The strength of the reflection at a certain point in time compared to the direct sound will be different for different volumes for the same value of G , since the average reflection density is different for different volumes. In the remaining of this thesis, the strength of the reflection will therefore be referred to as the reflection strength relative the direct sound in decibel and it is denoted as P_{rel} .

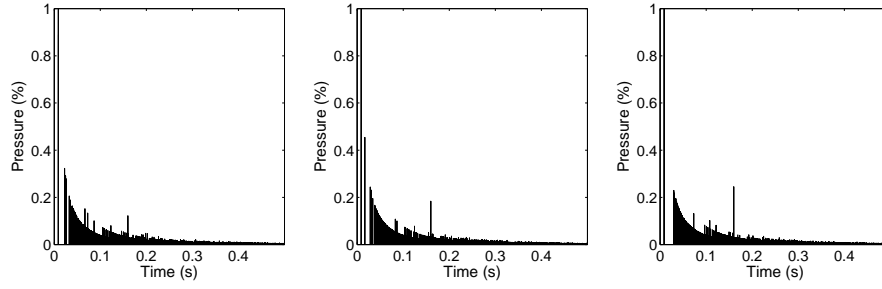


FIGURE 3.5: Reflectogram for the modified impulse responses. The reflection density was changed at $t_{delay} = 150$ ms. The reflectograms are based on different realizations of the Poisson process.

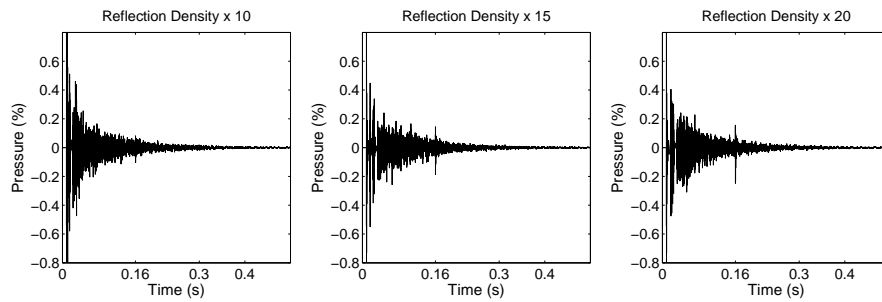


FIGURE 3.6: Impulse responses with modified reflection density at $t_{delay} = 150$ ms. The impulse responses are based on different realizations of the Poisson process.

By convolving the simulated impulse responses with anechoic signal of speech the authenticity of this modelling approach could be investigated. A speech signal was obtained from ODEON with a man's voice was used for this purpose. The convolved signal sounded realistic when played over headphones and the impulse responses are accepted in its ability to simulate a realistic sound.

3.1.6 Objective investigation of the criteria

Since the impulse response is a result of a random process, the value of R_{slope}^{max} will differ between each realization. Therefore, a Monte Carlo simulation was performed in order to investigate the relationship between the parameter R_{slope}^{max} and the relative strength of the reflection P_{rel} . The simulation was repeated until the confidence interval of the mean value was sufficiently small and the simulation was performed for rooms with various volumes, reverberation times and delay times. The same simulation was performed for the ΔEDC and the criteria by Dietsch and Kraak, EK to determine which parameter is the most robust one.

3.2 Listening test

In order to find a subjective threshold for detecting the reflection in terms of various parameters, two listening tests were performed. In the first listening test the repeatability of the test and the dependence on delay time was investigated. Therefore, the test sequence was repeated 4 times for each test subject and for three different delay times, where all other parameters remained fixed. In the second test, the dependence of volume and reverberation time was investigated. The two listening tests are described in more detail later. The test design was the same for both listening tests. Finally, the results could be analysed with ANOVA, which is described in Section 3.2.4.

3.2.1 Test design

Finding the detection threshold

In the listening tests, the subject listened to three impulse responses in each presentation. Two of the signals were free from strong a reflection (the reference) and one was the impulse response that included a strong reflection (the test signal). The task was then to identify the test signal. Even if the subject could not distinguish which of the three signals that contained the strong reflection, the subject had to make a decision, i.e a *3 alternative forced choice method (3AFC)* was used.

The listening tests were performed with an adaptive method, which means that each presentation is dependent on the previous answer. After two correct answers in a row, with the same test signal, the strength of the reflection was decreased. In case of one incorrect answer, the strength of the reflection was increased. With this test design the final threshold will correspond to the 70% point on the psychometric function¹ [25].

The test starts with a large value of the reflection strength for the test signal and the strength of the reflection was then 40 times the average strength. In case of a correct answer (signal was detected) the strength was changed with a step size of

¹The psychometric function is the probability that the subject answers "Yes" to the question if he/she could identify the test signal as a function of the strength of the test signal

4 for the next presentation. After one reversal (change in the direction because of an incorrect answer) the step size was decreased to 2, and after another reversal the step size was 1. This was also the point where the measuring phase of the test started. After four reversals the test was stopped and the final threshold is the median of the levels presented during the measuring phase.

Preparation of the signal

The impulse responses were calculated on beforehand and saved. For each combination of volume, reverberation time and delay time 40 impulse responses were created with reflection strength, $G = 1$ to 40. This was done both for impulse responses where the head related transfer function had been taken into account (used for playback) and for impulse responses without the HRTF (used in order to calculate the parameter). The event history of the reflections were identical for the impulse response used for playback and for the responses used to calculate the parameter, i.e., the same realization of the Poisson process was used. But for different values of G , delay time, reverberation time and volume a different realization of the Poisson process was used.

In order to calculate the parameters, the event history of the Poisson process plays a role and these could only be found from those impulse responses already created and saved. The threshold from the test therefore has to be rounded to the nearest integer (1 to 40) to find the corresponding parameter.

The listening test was designed using PSYLAB [24] which is a free software containing scripts written in MATLAB. With help of these scripts, the test could be designed in terms of step size, maximum of reversals, start values and presenting order as described above.

Preparation for the subjects

For both listening tests 18 test subjects, 8 female and 10 male, in the age 23-30 were participating. All the subjects had previous experience in this type of listening tests. Before the test started, the subject got familiar with the test signal through listening to a sequence of impulse responses with and without a strong reflection. The test-subject also did a test-run to get some training. This was done in order to get rid of the training effect of the first runs of the experiment, as the first run would otherwise most likely have a larger threshold than the following runs. After

6 runs, the test subject was asked to take a break in order to avoid a tiring effects. The full instruction, given to each test subject can be found in Appendix A.1.

Signal strength

Since the direct sound had the same strength in all different cases, this determined the level during playback. By using an ear-simulator and measuring the level with a small integration time, this level could be measured and set to 70 dB.

3.2.2 Test 1

The configuration of the impulse responses used for listening test 1 can be found in Table 3.5. As mentioned earlier, the purpose of this test is to investigate the dependence on delay time and the repeatability of the test. 16 test subjects participated and they all listened to same 4 realizations of the impulse responses which resulted in 64 measuring points for each delay time. All impulse responses used in this test were based on different realizations of the Poisson process, but the same impulse responses were presented to the different subjects. With this experimental set-up, it could be investigated whether the subjects could hear the difference between impulse responses based on different realizations of the Poisson process, and if the realization of the Poisson process has an influence on the detection threshold. Furthermore, each presentation during the test (the test signals and the two reference signals) were based on a different realization of the Poisson process. The resulting threshold is therefore not dependent on one specific Poisson process, which would be a bad idea if the test subject could hear the difference between different realizations. If the same realization is used between the different presentations, the value of the parameters will also be biased since the surrounding reflections always have the same strength, but by using different realizations, the randomness of the parameters is taken into account.

Test 1

| Volume (m ³) | Reverberation Time (s) | Delay time (s) | Realizations |
|--------------------------|------------------------|----------------|--------------|
| 750 | 0.93 | 100 | 4 |
| | | 150 | 4 |
| | | 200 | 4 |

TABLE 3.5: Properties of the impulse responses used in test 1

The 12 different runs (4 repetitions three different delay times) were presented in a random order for each subject.

From this test, the dependence on delay time and the repeatability of the test was analysed using ANOVA, which is described further in Section 3.2.4.

3.2.3 Test 2

In the second listening test, the dependence of reverberation time and volume were investigated. The volume was changed while using the same reverberation time as in test 1, and the reverberation time was changed keeping the volume from test 1. In Table 3.6 the configuration of the impulse responses can be found. In comparison to test 1, each subject were now presented to impulse responses based on different realizations of the reflectogram.

| Test 2 | | |
|---------------------------|--------|-----------------|
| Volume (m ³) | RT (s) | Delay time (ms) |
| 450 | 0.93 | 100 |
| | | 150 |
| | | 200 |
| 750 | 0.56 | 100 |
| | | 150 |
| | | 200 |
| | 1.30 | 110 |
| | | 150 |
| | | 200 |
| 1050 | 0.93 | 100 |
| | | 150 |
| | | 200 |

TABLE 3.6: Properties of the impulse responses used in test 1

In test 2, the presenting order of the 12 different runs were decided with help by a Latin Square Matrix. This was to ensure that two runs will not appear in the same order after each other for any of the subjects [25]. With this pattern of presenting the different runs, the dependence on the previous presentation has less influence

and his decreases the risk of bias of the response from the test subjects. For the first subject the presentation order can be found as

1, 2, n , 3, $n-1$, 4, $n-2$...

where n is the number of runs.

The pattern is repeated for the next subject but the start value is changed. The second subject will be presented with the sequence

2, 3, n , 4, $n-1$, 5, $n-2$...

This pattern continued for 8 different sequences, since 8 subjects participated in the test.

3.2.4 Analysis method

Analysis of Variance

For test 2, an one-way ANOVA was performed in order to investigate how significant the difference is between detection thresholds for different volumes, reverberation times and delay times. With a F-test, the significance of the difference between two groups of test-values can be found by the corresponding p-value. F is the ratio between the mean squared sum for the 'within a group-variance' MS_W , and the 'between groups variance' MS_B , which is further the sum of squares SS divided by the degree of freedom df [26], i.e.,

$$F = \frac{MS_B}{MS_W} = \frac{SS_B/df_B}{SS_W/df_W} \quad (3.9)$$

If the ratio is large, the variance within the groups are more dominant than the variance within the groups. What value of F that is considered as large is decided by the p-value. Under the null-hypothesis, " H_0 -no difference between the two groups", the test statistics follows a F -distribution. By finding the corresponding p-value for the expected F -distribution, the significance level α is found. Here, significance is a p -value $\alpha < 0.05$ will be used, i.e., on a 5% confidence level.

When comparing two or more groups, for example three different reverberation times, the F -test gives a significant results if at least one group is different from the others. The test was therefore comparing the groups two by two in most of the cases.

Repeatability

From the results in test 1 the repeatability will be investigated. The test has more than one factor that can contribute to the variance in the data, that is the variation between the different subjects and for different realizations of the test signal. Either the subjects have different thresholds, or the realizations of the tests is the main reason to the variance and if the variance is too large for different realizations, the repeatability is not sufficient. In order to analyse from where the variance originates a two-way ANOVA can be used, since this test indicates for which factor the variance is largest.

The two factors used here is then, *Realization* and *Subject*. If there is sufficient repeatability in the test, the variance within different realizations is smaller than the variance within the subjects [27]. The basic idea is to use the same principle as a one way ANOVA, but for the two different factors separately. The two hypotheses for this test is:

" H_{01} - The data categorized as different Realizations all belong to the same distribution

" H_{02} - The data categorized as different Subjects all belong to the same distribution

If a hypothesis is rejected, there is a significant difference for this factor.

Cumulative distribution

When analysing the dependence of different factors (volume, reverberation time, delay time), the cumulative distribution of the thresholds will also be used. This is found by calculating the percentage of the subject that detected the reflection up to a certain threshold. This curve starts at a value 0%, for low values of the parameter, none of the subjects detected the reflection, and then it increases up 100% where all the subject has detected the reflection. The level at which 50% of the test subjects detected the reflection will be used as the threshold that represents all the subjects. The 50% point was found from a fitted cumulative distribution.

A logistic function has similar shape as the cumulative function, and it is very similar to the cumulative distribution of the normal probability density [25]. This function was fitted to the cumulative distribution and it has the form

$$c(x) = \frac{1}{1 + e^{a+bx}} \quad (3.10)$$

where a and b are the coefficients to be found. Coefficient b determines the steepness of the curve, while a only determines its position on the x-axis. Figure 3.7 show this function for different values of b and $a = 10$.

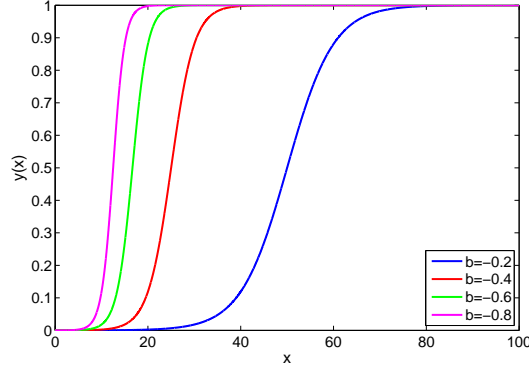


FIGURE 3.7: Equation 3.10 for different values of b

Regression analysis

When the 50% point on the cumulative distributions was found, for various volumes, reverberation times and delay times, the dependence of the threshold on these factors could be investigated. For this, a linear regression curve was fitted to the results, using the least square method. This method finds the coefficients β_0 and β_1 in the regression line $\mu = \beta_0 + \beta_1 x$ that minimizes the residual

$$Q(\beta_0, \beta_1) = \sum_{i=1}^n (y_i - \mu_i)^2 \quad (3.11)$$

where $\mu_i = \beta_0 + \beta_1 x_i$ [28].

The significance of the coefficient β_1 is further analysed with ANOVA. The null hypothesis is in this case " H_0 - the slope of the fitted line β_1 is equal to zero. The p-value is the probability that the slope takes a more extreme value than the estimated value of β_1 , under the assumption that H_0 is true. For a small p-value, the null hypothesis can be rejected. [28]

Variance comparison

When comparing the variance of different parameters expressed in different units the relative variation, %RSD is used. This is calculated by taking the standard deviation divided by the mean

$$\%RSD = \frac{\sigma^2}{\mu} \quad (3.12)$$

Another useful tool when comparing two parameters X and Y of different units, is the correlation coefficient [28].

$$\rho(X, Y) = \frac{E(XY) - E(X)E(Y)}{\sqrt{V(X)}\sqrt{V(Y)}} \quad (3.13)$$

Chapter 4

Results

In this chapter the results from the Monte Carlo simulation and the listening tests are displayed together with the results from the ANOVA and regression analysis. The outcome of the listening test are first presented in terms of the strength of the reflection relative the direct sound, P_{rel} , which is the parameter that was found from the listening test.

This is followed by the results in detection thresholds expressed in terms of the different parameters *peak value of the slope ratio*, R_{slope}^{max} , *drop in the decay curve*, ΔEDC and the *criteria by Dietsch and Kraak*, EK . As described in Section 3.2 the impulse responses were saved in forehand and the parameters were calculated from that impulse response with a reflection strength corresponding to the threshold found from the listening test.

4.1 Monte Carlo Simulation

The behaviour of the parameters R_{slope}^{max} , ΔEDC and EK when increasing the strength of the reflection P_{rel} were investigated by making many simulations of the same impulse responses. The mean value of the different parameters could be estimated and the simulation was performed until the confidence interval of the mean value and the variance was small enough. The reflection density was $G = 10$, 15 and 20 times the average reflection density at the time of the reflection and the results are plotted as a function of the strength of the reflection relative the direct

sound.

The results for R_{slope}^{max} can be seen in Figure 4.1 together with the confidence interval for the mean value (dashed line). The confidence intervals of the variance are of similar size, but not included in the plot to maintain clarity. The peak value increases as the strength of the reflection is increased, which is expected, but it also increases with volume. In the case of a larger volume, the average reflection density is smaller and a sudden increase in the reflection density therefore has a larger influence among its neighbouring reflections, comparing with the case of a smaller volume but the same size of the reflection. Therefore, the slope of the decay curve changes more suddenly for larger volumes. With increasing reverberation time, a similar effect is seen. A longer reverberation time masks the strong reflection, making it less dominant among its neighbours. The peak value of the slope ratio is therefore lower for longer reverberation times.

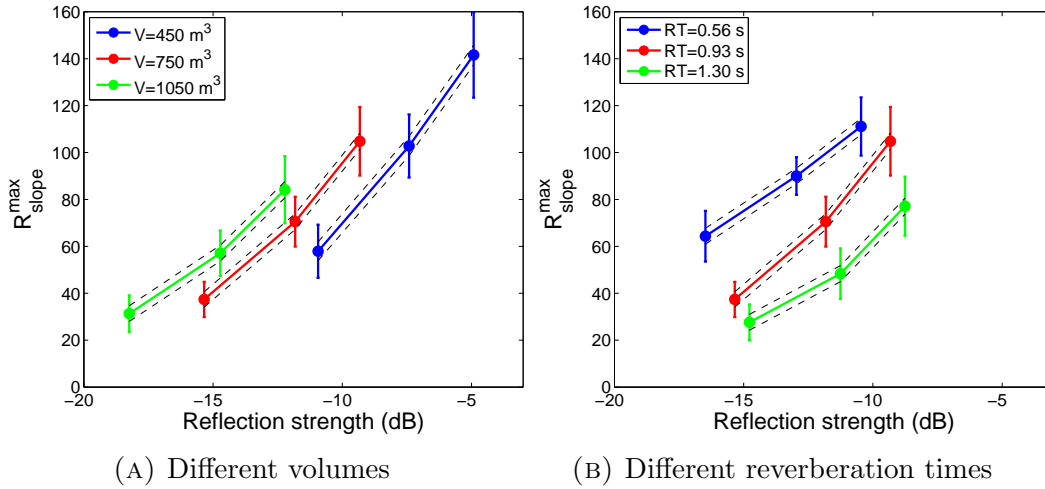


FIGURE 4.1: Peak value of R_{slope} as a function of the reflection strength relative the direct sound, P_{rel} .

The variance for R_{slope}^{max} is relatively large for the same value of the reflection strength, which is a problem because the error bars on the y-axis are almost overlapping for different strength of the reflection. The relative standard deviation can be seen in Table 4.1. A detection threshold in terms of reflection strength could then result in a peak-value that is arbitrary, but if also the subjective experience is different, i.e., the results of the listening test is differing for different runs (different realization of the Poisson process), this variation might not be a problem. This is

investigated further in Section 4.2.

The same simulation was run for the parameter ΔEDC and the result is shown in Figure 4.2. This parameter behaves similar to R_{slope}^{max} , but lower reverberation times results in very large values.

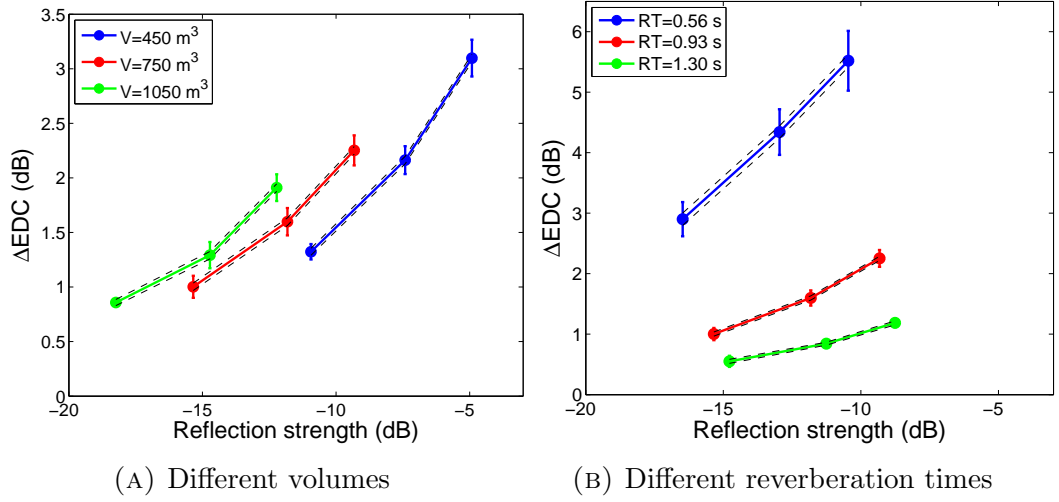
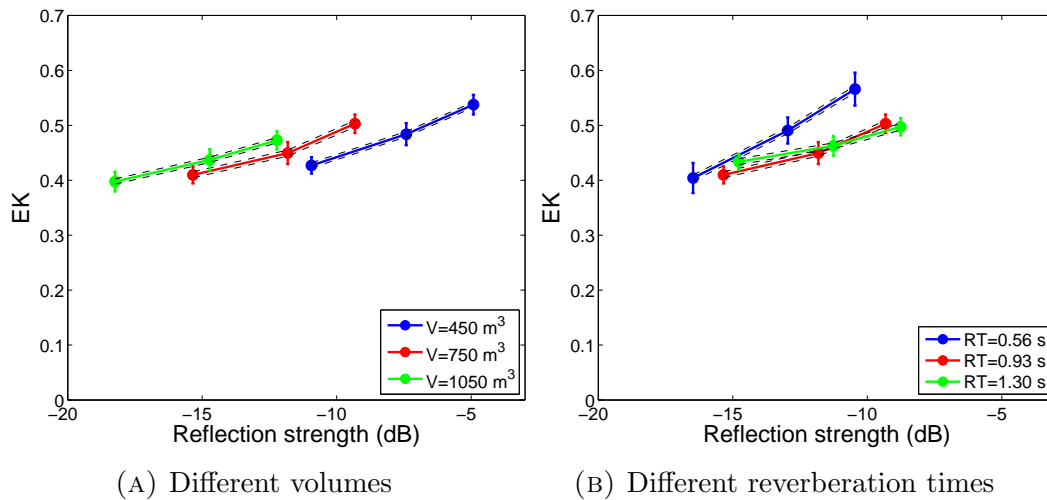


FIGURE 4.2: ΔEDC as a function of reflection strength. Notice the different scale of the y-axis.

The relative standard deviation is found in Table 4.1 and it can be seen that it is smaller than for R_{slope}^{max} . Since this measure is taking a larger interval of the slope ratio into account, it is probably more robust than R_{slope}^{max} , i.e., it does not take arbitrary values for different reflectograms.

Figure 4.3 shows the result for the criteria by Dietsch and Kraak and in Table 4.1 it can be read that this parameter has the lowest relative standard deviation. This parameter does not increase as much when increasing the strength of the reflection (compared to the other two measures), which can be interpreted as that the parameter EK is less sensitive. For increasing volume, the parameter shows larger results. It does not seem like there is a significant dependence on reverberation time, at least not for reverberation times longer than 0.93 s.

FIGURE 4.3: Parameter EK as a function of reflection strength

| Reflection strength (dB) | | -15 | -12 | -9 |
|--------------------------|-------------------|-----|-----|----|
| %RSD | R_{slope}^{max} | 11 | 13 | 15 |
| | ΔEDC | 9 | 9 | 7 |
| | EK | 5 | 5 | 4 |

TABLE 4.1: Relative standard deviation for different strength of the reflection and parameters. The volume is $V = 750 \text{ m}^3$ and the reverberation time $RT = 0.93 \text{ s}$. Other volumes and reverberation times shows similar deviations.

4.2 Repeatability

In test 1, each subject repeated the experiment 4 times for each delay time and with a fixed volume and reverberation time. The detection thresholds in terms of reflection strength is plotted in Figure 4.4 for all 16 subjects and the different markers and colours corresponds to a different realization of the impulse response. The results was analysed with a two-way ANOVA and presented in Table 4.2.

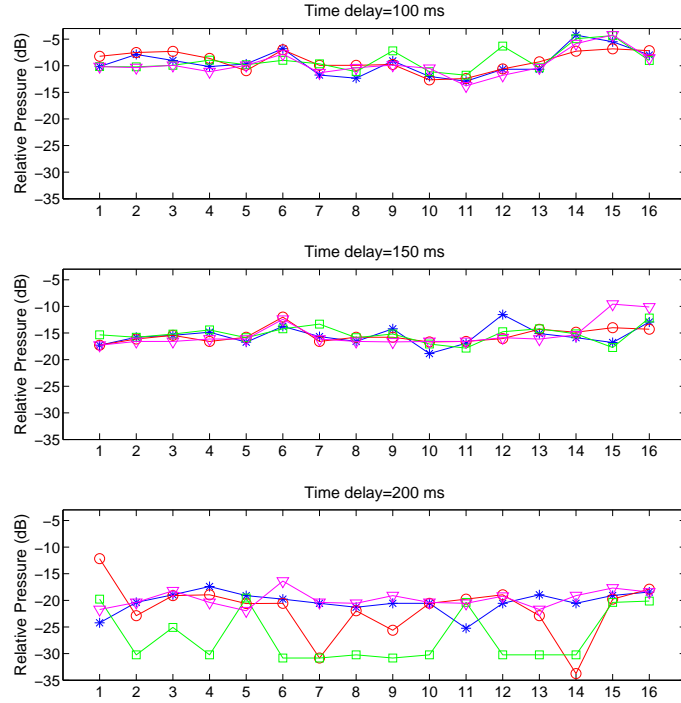


FIGURE 4.4: Individual results for test 1. The volume is $V = 750 \text{ m}^3$ and reverberation time $RT = 0.93 \text{ s}$

| Delay time (ms) | Subject | | Realization | |
|-----------------|---------|----------------------|-------------|------------|
| | F | p -value | F | p -value |
| 100 | 5.27 | $6.87 \cdot 10^{-6}$ | 1.88 | 0.15 |
| 150 | 2.84 | 0.0035 | 0.24 | 0.87 |
| 200 | 0.71 | 0.76 | 0.85 | 0.47 |

TABLE 4.2: Results of the 2-way ANOVA for the parameters *Realizations* and *Subjects*

The p -value for the factor *Subject* is smaller than the value for the factor *Realization* for delay times 100 and 150 ms. The null-Hypothesis " H_{01} - The data categorized as different *Realizations* all belong to the same distribution" can not be rejected, meaning that the variance is not significant for this factor and the repeatability of the test is sufficient. On the other hand, the null hypothesis is rejected for different subjects, i.e., the subjects has different thresholds.

For delay times of 200 ms, the null hypothesis is neither rejected for factor realizations, meaning that the repeatability is not sufficient for this delay time. It might be that this variance is due to one particular run (green) that has a lower threshold

than the other runs for a time delay of 200 ms. It could also be that the delay time is so long that it is the threshold of the reflection alone that is being measured.

Correlation between the thresholds

For delay times of 100 ms and 150 ms it can be concluded that the realization of the Poisson process does not have a large influence on the detection threshold. For further investigations of the detection threshold in terms of the parameters, the parameters R_{slope}^{max} , ΔEDC and EK should preferably follow this threshold. This is investigated by finding the correlation coefficient between the parameters and the strength of the reflection for all subjects and runs (in total 64 runs) using equation 3.13. If the parameter takes a random value, or if it is not sensitive enough, this shows a large deviation from threshold for reflection strength. Table 4.3 summarizes the results, and it can be concluded that ΔEDC tracks the results found in relative strength. Probably, R_{slope}^{max} has a lower correlation due to its randomness, and EK because it is not sensitive enough.

| Parameter | R_{slope}^{max} | ΔEDC | EK |
|-------------------------|-------------------|--------------|--------|
| Correlation coefficient | 0.8397 | 0.9080 | 0.7097 |

TABLE 4.3: Correlation between the three parameters and the relative strength of the reflection at thresholds for all runs

4.3 Dependence of delay time

Test 1 was run for three different delay times, $t_{delay} = 100, 150$ and 200 ms. The threshold for the strength of the reflection relative the direct sound (P_{rel}) and the three parameters (R_{slope}^{max} , ΔEDC and EK) is presented as box-plots in Figure 4.5. The red bar in the box-plot is the median value of the data, the borders of the box are the 75% and 25% percentile and the whiskers are the maximum and minimum values of the data. Some data point are regarded as outliers and these are marked separately as crosses. It can be seen that there is a clear trend for time delay of the reflection, as expected.

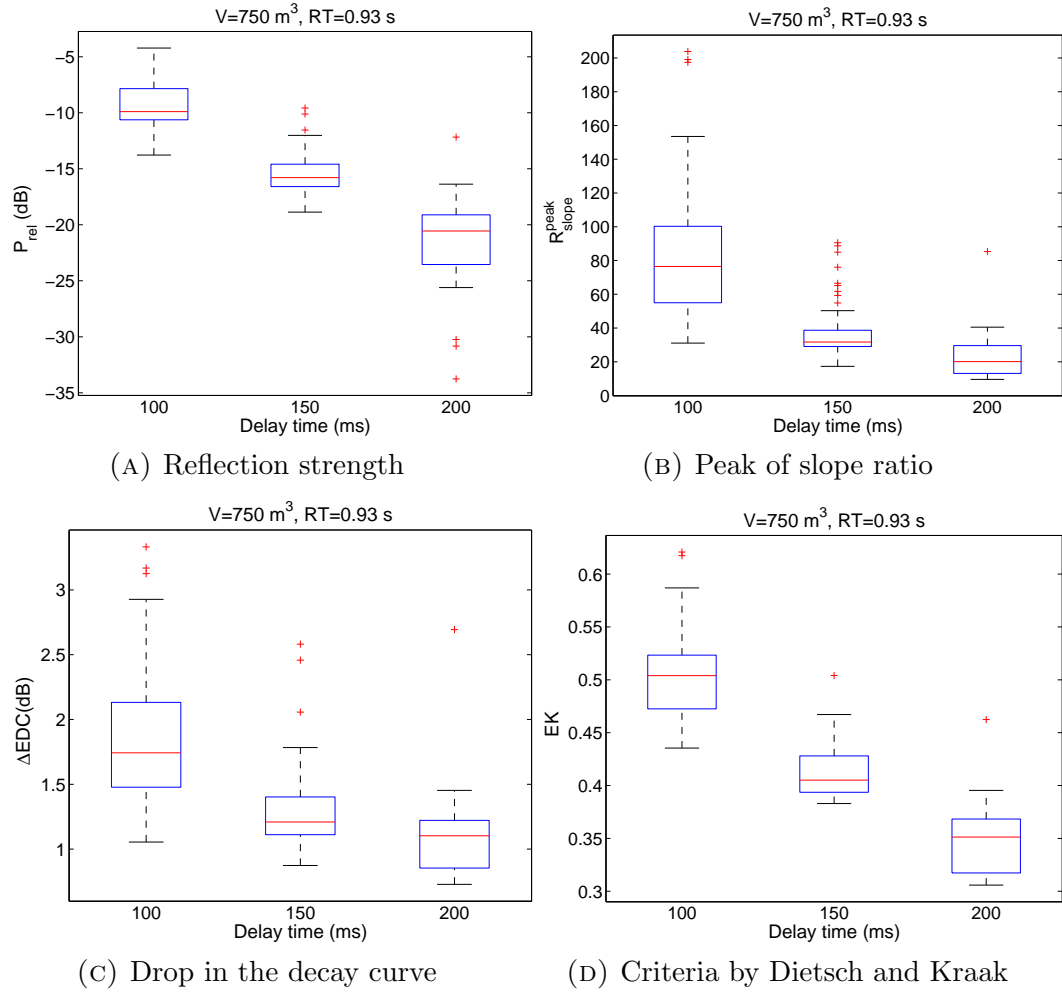


FIGURE 4.5: The detection threshold for the reflection strength and the different parameters as a function of delay time

4.3.1 Cumulative distribution

Figure 4.6 shows the results for different delay times from test 1 as a cumulative distribution, where the y-axis shows the percentage of subjects that detected the reflection.

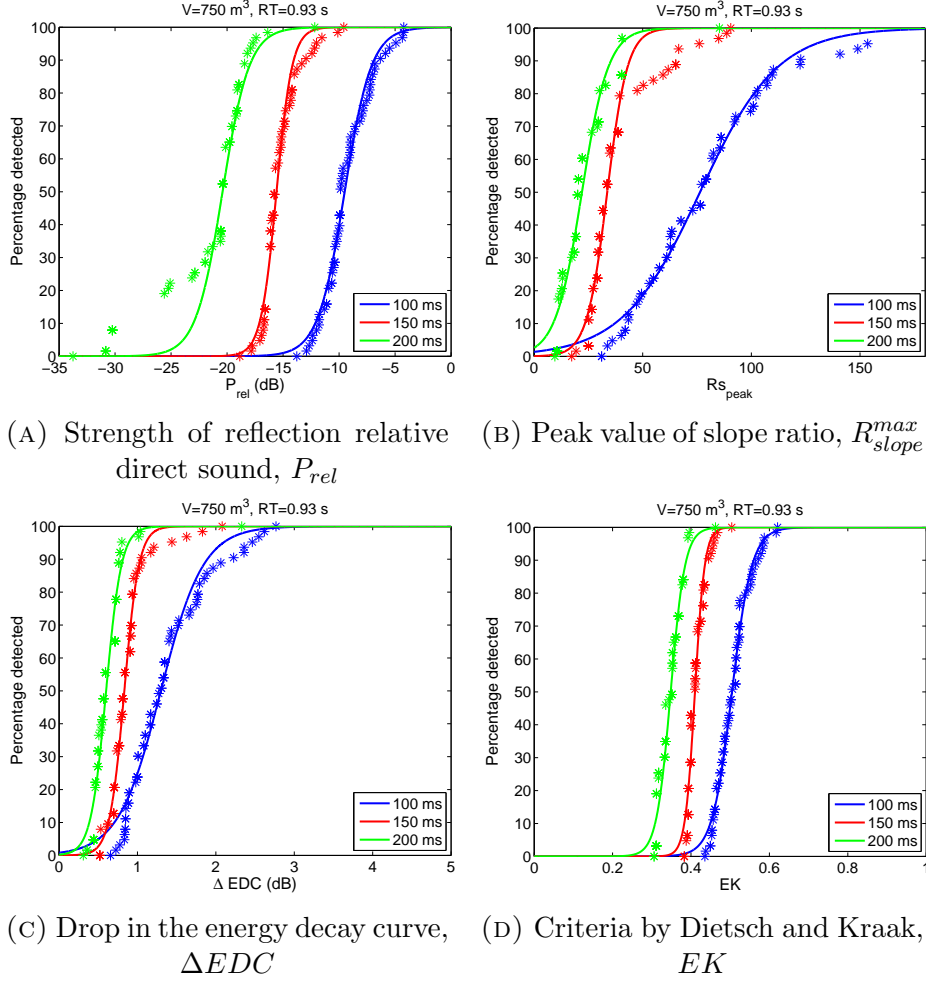


FIGURE 4.6: Cumulative distribution for different time delays together with the fitted logistic function

These results can be compared with the results from Haas presented in Section 1.2, with the difference that here the x-axis is the strength of the reflection instead of delay time. The threshold in this experiment is -10 dB at $t_{delay} = 100$ ms and compared with the results from Haas experiment, the same level of the reflection is not disturbing for delay times up to at least $t_{delay} = 160$ ms. it should be kept in mind that the signal used in the experiment by Haas was a speech signal, which should result in higher thresholds. Furthermore, the thresholds in the experiments by Haas is the threshold of disturbance, which is also supposed to be larger than the detection threshold.

From the cumulative distribution the value at which 50% of the subjects detected¹

¹Due to the design of the listening test, the detection threshold for each subject corresponds to the 70%-value of the psychometric function of the test subject

the reflection can be found and from these points a regression curve was fitted to investigate the trend due to delay time (see Figure 4.7).

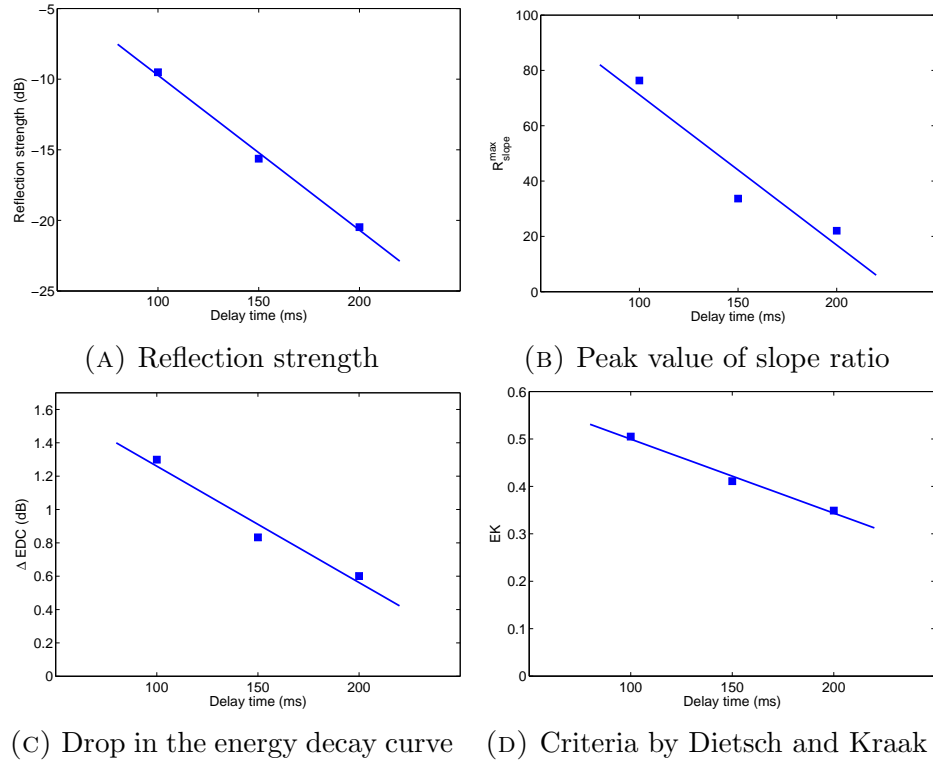


FIGURE 4.7: Regression line fitted to the 50% points in Figure 4.5

The slope β_1 of the regression lines together with F , and p -values is found in Table 4.4. The slope for reflection strength is significant, but non of the parameters show a significant trend on a level of $\alpha = 0.05$. Although, it should be kept in mind that only three points have been used in the regression analysis. When all data points is used for the regression, the slope is significant in all cases.

| Parameter | β_1 | F | p -value |
|--------------------------|-----------|--------|------------|
| P_{rel} [dB/ms] | -0.110 | -14.96 | 0.042 |
| R_{slope}^{max} [1/ms] | -0.543 | -3.03 | 0.203 |
| ΔEDC [dB/ms] | -0.007 | -5.17 | 0.122 |
| EK [1/ms] | -0.002 | -8.44 | 0.075 |

TABLE 4.4: Results of the ANOVA of the regression fit in Figure 4.7.

The slope for the reflection strength as a function of delay time is -0.1dB/ms , which is a small slope compared to the results found by Burgtorf et al. (Section

1.2), who found the slope to be -0.57 dB/ms. The detection threshold according to this experiment is lower than the results found in this thesis, even though the test signal is speech instead of an impulse response. In the experiments by Burgtorf, the impulse response did only consist of two reflections, the direct sound and the reflection, which means that the reflection was not masked by any surrounding reflections and this might be an explanation to this deviation.

4.4 Dependence on volume and reverberation time

4.4.1 Thresholds for the reflection strength

In test two the detection threshold was found for different room properties, i.e. for different reverberation time and volume. The detection thresholds in terms of reflection strength for different volumes is plotted as a box plot in figure 4.8a ($RT = 0.93$ s) and for different reverberation times in figure 4.8b ($V = 750$ m³).

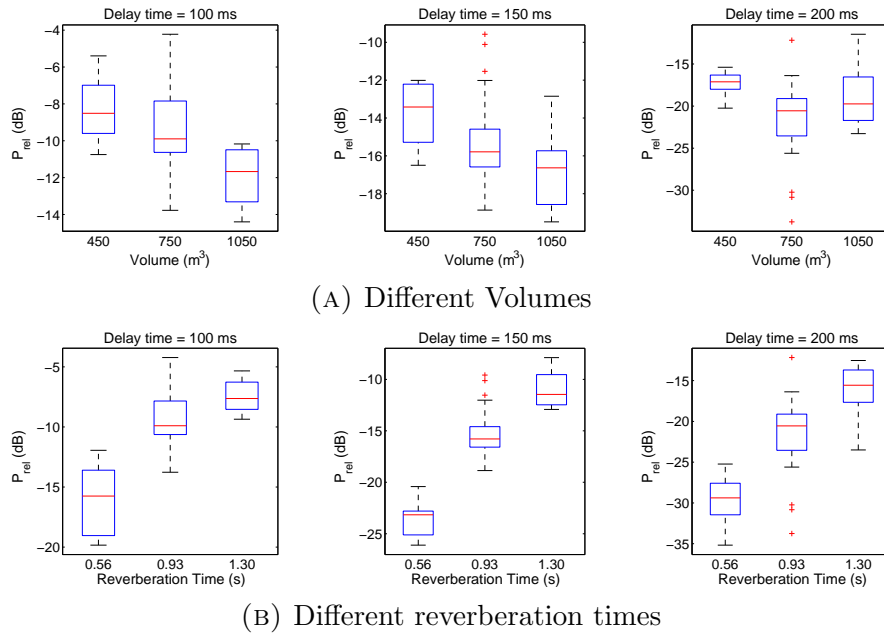


FIGURE 4.8: Detection threshold in reflection strength

The volume seems to have an influence on the detection threshold and this is confirmed by the ANOVA seen in Table 4.5 although, not all comparisons between the volumes shows a significant difference. The reflection density is larger for smaller volumes, for a specific time, which could explain this relation. If there are

more reflections before and after the strong reflection, as for a small volume, the detection threshold is higher due to a masking effect.

| Reflection strength - Volume | | | |
|------------------------------|------------------------------------|--------|------------|
| Delay time (ms) | Compared volumes (m ³) | F | p -value |
| 100 | 450-750 | 1.603 | 0.209 |
| | 750-1050 | 10.738 | 0.001 |
| 150 | 450-750 | 5.928 | 0.017 |
| | 750-1050 | 4.156 | 0.045 |
| 200 | 450-750 | 8.568 | 0.004 |
| | 750-1050 | 3.766 | 0.056 |

TABLE 4.5: Summary of the ANOVA comparing different volumes for different delay times.

The threshold does also differ significantly for different reverberation times, as seen in figure 4.8b and confirmed by the ANOVA summarized in Table 4.6. The reflection density is the same within the plots, since delay time and volume is constant. But the strong reflection is masked due to the long reverberation time. The reflections before and after the strong reflection has therefore larger amplitude for longer reverberation times, for a given delay time.

| Reflection strength - Reverberation Time | | | |
|--|-----------------|---------|-----------------------|
| Delay time (ms) | Compared RT (s) | F | p -value |
| 100 | 0.56-0.93 | 62.448 | $2.76 \cdot 10^{-11}$ |
| | 0.93-1.30 | 5.509 | 0.0218 |
| 150 | 0.56-0.93 | 153.103 | 0.0218 |
| | 0.93-1.30 | 44.376 | $5.14 \cdot 10^{-9}$ |
| 200 | 0.56-0.93 | 20.132 | $3.03 \cdot 10^{-5}$ |
| | 0.93-1.30 | 12.142 | $8.54 \cdot 10^{-4}$ |

TABLE 4.6: Results of ANOVA for different reverberation times

4.4.2 Thresholds for the criteria

Volume

The detection thresholds are now plotted in terms of the different parameters R_{slope}^{max} , ΔEDC and EK . The results for various volumes are plotted in Figure 4.9. The dependence of volume is no longer as clear as for the threshold expressed in terms of the reflection strength.

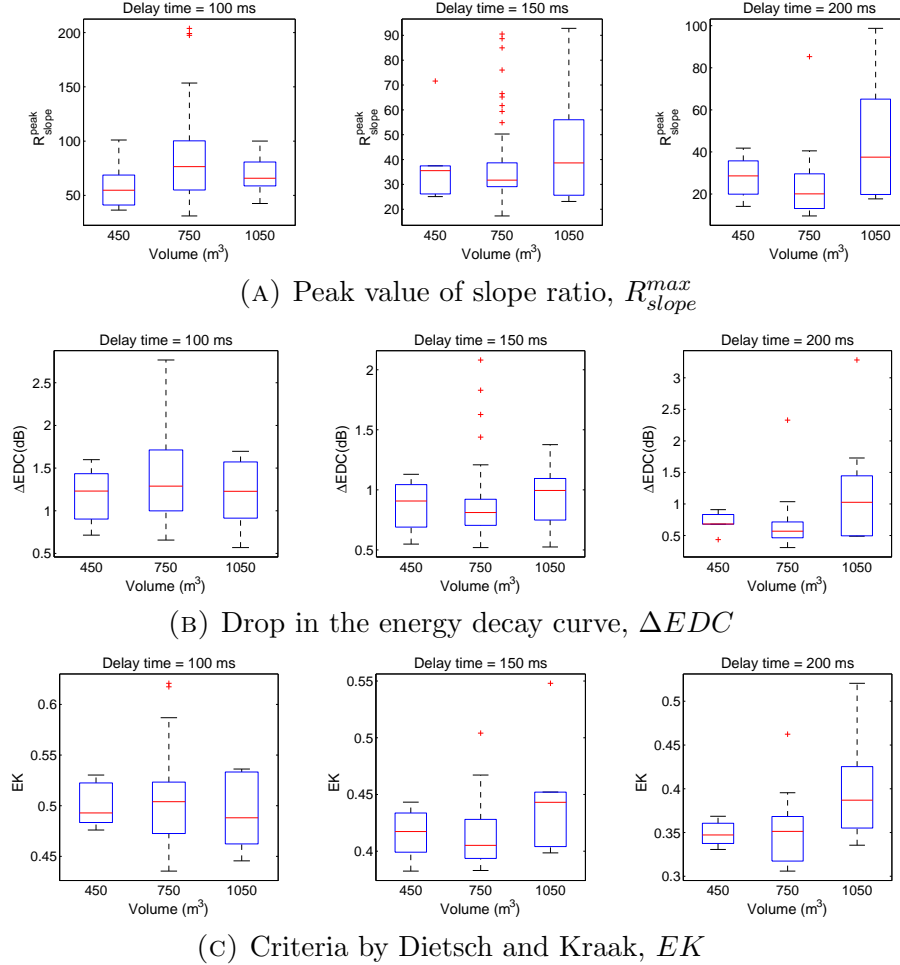


FIGURE 4.9: Thresholds in terms of the different criteria for different volumes

The F-test was performed by comparing all three groups in one test. If no significance is shown, none of the groups are significantly different from the others. For delay times $t_{delay} = 100$ ms and 150ms, the F-test, presented in Table 4.7, shows that there is no significant difference between any of the three volume-groups. In the case of $t_{delay} = 200$ ms, the F-test shows significance due to difference between the rooms with volume $V = 750$ m³ and $V = 1050$ m³ while the difference between $V = 750$ m³ and $V = 450$ m³ is not significant (not shown here).

Different parameters - Volume

| Parameter | Delay time (ms) | F | p -value |
|-------------------|-----------------|-------|----------------------|
| R_{slope}^{max} | 100 | 1.741 | 0.182 |
| | 150 | 0.532 | 0.589 |
| | 200 | 8.220 | $5.80 \cdot 10^{-4}$ |
| ΔEDC | 100 | 0.852 | 0.430 |
| | 150 | 0.334 | 0.717 |
| | 200 | 9.115 | $2.80 \cdot 10^{-4}$ |
| EK | 100 | 0.432 | 0.650 |
| | 150 | 4.663 | 0.012 |
| | 200 | 9.065 | $2.91 \cdot 10^{-4}$ |

TABLE 4.7: Results of the ANOVA for different volumes

Reverberation time

The results is now plotted for different reverberation times in Figure 4.10. The reverberation time seem to have an influence on the parameter by Dietsch and Kraak, but not for ΔEDC . This can be seen as a advantage for this parameters, since the reverberation time does not have to be taken into account when calculating the detection threshold.

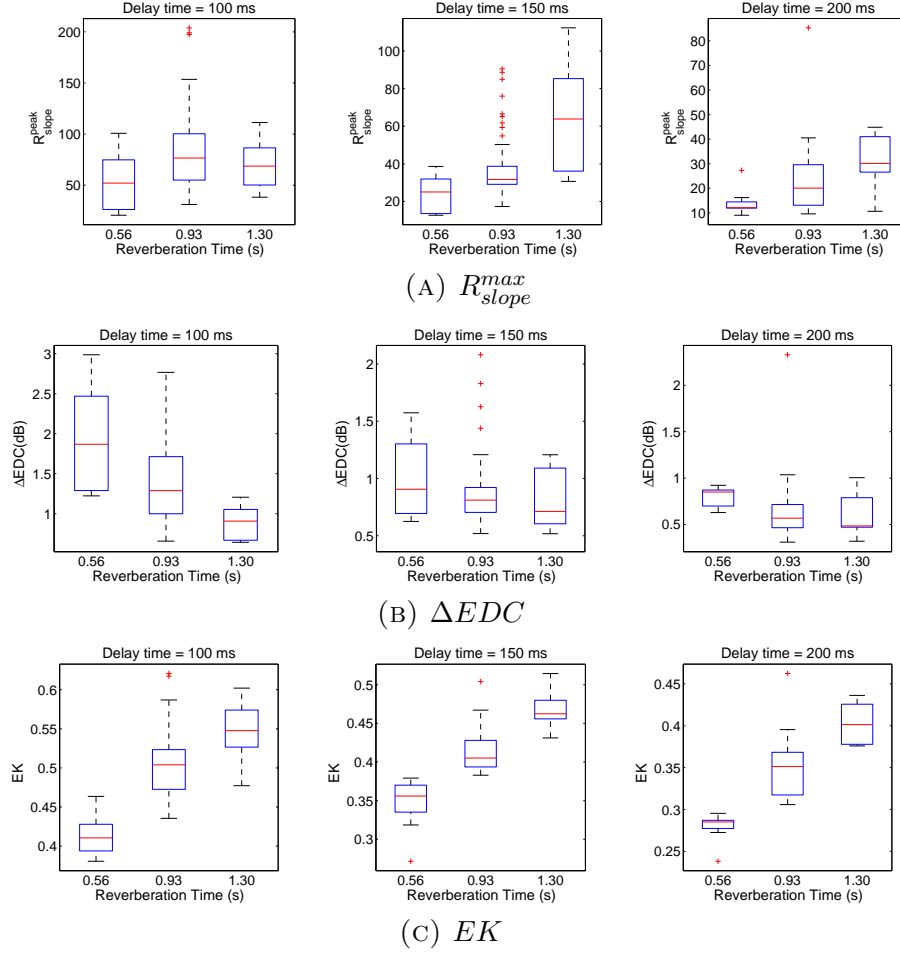


FIGURE 4.10: Thresholds in terms of the different criteria for different reverberation times

Here the ANOVA was made by comparing the set of data two by two, i.e., the thresholds for a reverberation time of $RT = 0.56$ and 0.93 s and then for $RT = 0.93$ and 1.30 s. The results is shown in Table 4.8. It can be seen that only for a delay time 100 ms the reverberation time has an influence for ΔEDC , but not for longer delay times. For parameter R_{slope}^{max} , the results looks more arbitrary, the reverberation time has a significant influence only for delay times of 150 ms. For EK , different reverberation time has a significant impact on the detection threshold. Comparing these thresholds in EK with the thresholds for $EK_{50\%}$ presented in section 2.5.1, shows that the threshold has a large dependence on the signal type. For impulse responses, the 50%-detection threshold is around $0.4 - 0.4$, while the threshold at which 50% of the subject is annoyed by the echo is 1.0 for a speech signal. This threshold is also larger because it is the annoyance threshold and not the detection threshold that is measured.

Different parameters - Reverberation time

| Parameter | Delay time (ms) | Compared RT (s) | F | p -value |
|-------------------|-----------------|-------------------|--------|----------------------|
| R_{slope}^{max} | 100 | 0.56-0.93 | 3.939 | 0.051 |
| | | 0.93-1.30 | 0.666 | 0.417 |
| | 150 | 0.56-0.93 | 5.466 | 0.022 |
| | | 0.93-1.30 | 14.092 | $3.56 \cdot 10^{-7}$ |
| | 200 | 0.56-0.93 | 3.841 | 0.053 |
| | | 0.93-1.30 | 3.167 | 0.079 |
| ΔEDC | 100 | 0.56-0.93 | 7.778 | 0.006 |
| | | 0.93-1.30 | 6.900 | 0.010 |
| | 150 | 0.56-0.93 | 1.601 | 0.209 |
| | | 0.93-1.30 | 0.190 | 0.664 |
| | 200 | 0.56-0.93 | 3.696 | 0.058 |
| | | 0.93-1.30 | 0.028 | 0.866 |
| EK | 100 | 0.56-0.93 | 34.834 | $1.16 \cdot 10^{-7}$ |
| | | 0.93-1.30 | 6.041 | 0.016 |
| | 150 | 0.56-0.93 | 48.844 | $1.30 \cdot 10^{-9}$ |
| | | 0.93-1.30 | 38.317 | $3.60 \cdot 10^{-8}$ |
| | 200 | 0.56-0.93 | 38.265 | $3.67 \cdot 10^{-8}$ |
| | | 0.93-1.30 | 26.121 | $2.66 \cdot 10^{-6}$ |

TABLE 4.8: Results of the ANOVA for different reverberation times

The variance in the data seems to be quite large compared to the range where the threshold was found. Table ?? below shows the relative standard deviation for the different parameters and time delays for a certain room ($V = 750 \text{ m}^3$ and $RT = 0.93 \text{ s}$). The $\%RSD$ for the reflection strength is due to variability in the experiment, i.e., different subjects and different runs. For R_{slope}^{max} the relative standard deviation has increased, and this is also the case for ΔEDC although not as much as for R_{slope}^{max} . The parameter EK has instead a smaller relative standard deviation than that of the threshold in terms of reflection strength, which could mean that this parameter is not that sensitive.

| | | | |
|--------------------------|-----|-----|-----|
| Delay time (ms) | 100 | 150 | 200 |
| Reflection strength (dB) | 23 | 11 | 21 |
| R_{slope}^{max} | 47 | 44 | 55 |
| ΔEDC | 28 | 24 | 25 |
| EK | 9 | 6 | 9 |

TABLE 4.9: Relative standard deviation for the different parameters.

4.4.3 Cumulative distribution

Figure 4.11 shows the results in the threshold for reflection strength for different volumes and delay times from test 2 as a cumulative distribution and a logistic function has been fitted to the curve.

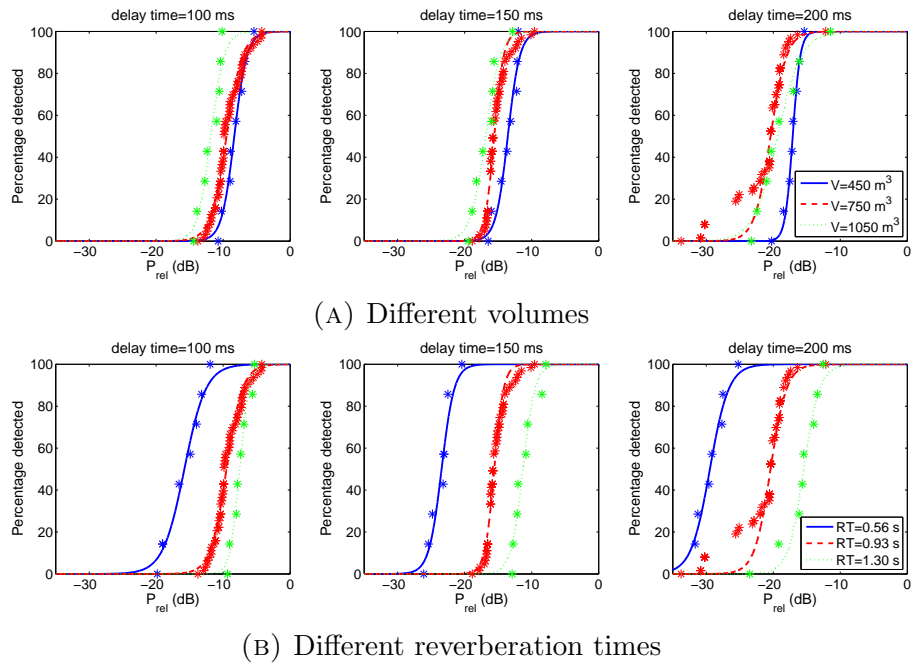


FIGURE 4.11: Cumulative distribution of the detection threshold in terms of the strength of reflection

From the fitted cumulative distribution function the 50% threshold can be found, meaning the level at which half of the number of the test subjects detects the reflection. From these thresholds a linear regression curve was fitted. The same procedure was done for the threshold in the different parameters and the result from the regression analysis can be seen in Figure 4.12.

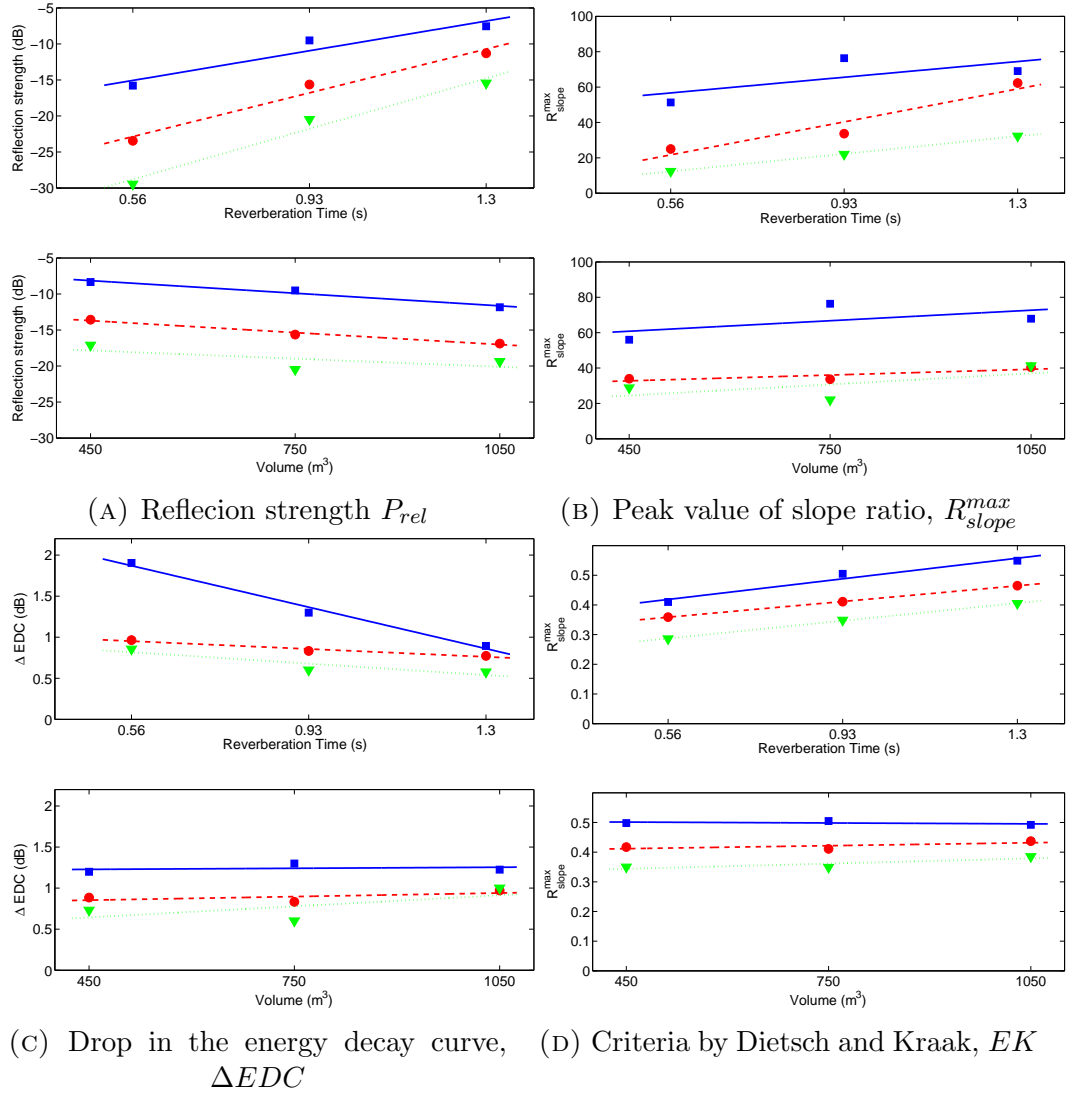


FIGURE 4.12: Fitted regression line to the 50% point of the cumulative distribution for different time delays 100 ms (blue line), 150 ms (red dashed line) and 200 ms (green dotted line)

For ΔEDC the slope is not significant (at a level of $\alpha = 0.05$) for any of the different delay times for volume and reverberation time. The reverberation time has an influence for EK , at least for longer delay times of the reflection, but the slope for volume time can be regarded as zero. Since there is only three points in the regression curve the p -values are in general quite large, even if it looks like there is a trend.

| Parameter | Delay time (ms) | β_1 | F | p -value |
|--------------------------|-----------------|-----------|-------|------------|
| P_{rel} [dB/ms] | 100 | -0.006 | -5.24 | 0.120 |
| | 150 | -0.005 | -7.13 | 0.089 |
| | 200 | -0.003 | -0.87 | 0.543 |
| R_{slope}^{max} [1/ms] | 100 | 0.187 | 4.720 | 0.132 |
| | 150 | 0.1432 | 91.79 | 0.006 |
| | 200 | 0.1608 | 29.44 | 0.021 |
| ΔEDC [dB/ms] | 100 | 0.0000 | 0.26 | 0.839 |
| | 150 | 0.0001 | 0.79 | 0.575 |
| | 200 | 0.0004 | 0.88 | 0.541 |
| EK [1/ms] | 100 | -0.0000 | -0.52 | 0.695 |
| | 150 | 0.0000 | 1.08 | 0.475 |
| | 200 | 0.0001 | 1.64 | 0.349 |

TABLE 4.10: Results of the ANOVA for different reverberation volumes

| Parameter | Delay time (ms) | β_1 | F | p -value |
|-------------------------|-----------------|-----------|--------|------------|
| P_{rel} [dB/s] | 100 | 11.138 | 3.31 | 0.187 |
| | 150 | 16.415 | 6.04 | 0.104 |
| | 200 | 18.955 | 6.20 | 0.102 |
| R_{slope}^{max} [1/s] | 100 | -0.0000 | -0.519 | 0.694 |
| | 150 | 0.0000 | 1.082 | 0.474 |
| | 200 | 0.0001 | 1.638 | 0.348 |
| ΔEDC [dB/s] | 100 | -1.365 | -8.75 | 0.072 |
| | 150 | -0.258 | -4.66 | 0.135 |
| | 200 | -0.376 | -2.08 | 0.286 |
| EK [1/s] | 100 | 0.188 | 4.72 | 0.133 |
| | 150 | 0.143 | 91.80 | 0.007 |
| | 200 | 0.161 | 29.44 | 0.022 |

TABLE 4.11: Results of the ANOVA for different reverberation times

Chapter 5

Discussion

The results of the listening tests analysed in the previous chapter shows that the parameters behave quite differently. The question is now which properties are desirable for a good echo-parameter. One important fact is that the thresholds has only been investigated for the pure impulse response and not with a signal, such as speech or music, convolved with the signal. It should be kept in mind that the thresholds presented here is the lowest threshold that can be found if changing the signal. Thresholds for speech signal using these impulse responses has a higher detection threshold and the threshold for music is most likely even larger.

5.1 Summary

Monte Carlo simulation

The parameters behaved quite differently when increasing the strength of the reflection, the volume and reverberation time. The peak value of the slope ratio had a larger relative standard deviation than ΔEDC , and the reason to this can be explained from its way of calculation. Since the slope ratio was calculated by the instantaneous slope using a very short differentiation interval of $\Delta t = 2.5$ ms, the peak value is very sensitive, which is confirmed by the Monte Carlo simulation. In case of a strong reflection, the slope ratio is not only changed in terms of its peak value, but also the values around the peak value takes higher values (see Figure 2.6a). In an interval of 5 ms, the slope ratio is significantly larger than at

other points. The parameter ΔEDC is proportional to the area under the peaks¹, and therefore it is more robust than R_{slope}^{max} .

Low reverberation time had a large influence on ΔEDC . This is probably due to the calculation of the parameter, which is based on a time interval of 5 ms for all reverberation times. Lower reverberation times has a larger slope, and if this interval is too large, the average slope of the decay curve will have an influence.

A better approach might be to first calculate the time interval of the increase in the slope and use this interval in the calculation. The time of the drop in the decay curve, or the interval of increasing slope ratio, was found to be around 5 ms in this simulation, but this is probably not always the case in reality.

Repeatability of the test

The repeatability was shown to be sufficient for delay times of 100 and 150 ms with a volume of $V = 750 \text{ m}^3$ and reverberation time $RT = 0.93 \text{ s}$. Due to a lack of time, the repeatability was only found for one volume and reverberation time. The problem with repeatability for longer delay times will probably also have an effect for shorter reverberation times since the reflection at 200 ms is then even more separated from the rest of the reverberation tail. The results for a reverberation time of $RT = 0.56 \text{ s}$ should therefore also be taken with caution for longer delay times.

Evaluation of the parameters

One way of evaluating the parameters is to investigate how well they represent the strength of the reflection P_{rel} . The strength of the reflection is controlled in the simulation of the impulse response. The threshold expressed in P_{rel} only takes different values due to the different behaviour of the test subjects. On the other hand, different realizations of the impulse responses, the parameters R_{slope}^{max} , ΔEDC and EK will be slightly different also because of the different statistics. P_{rel} is the parameter for reflection strength in the simulation of the impulse response. It is therefore intuitive that a comparison between this threshold and the threshold of the parameters is a reasonable way of evaluating the parameters. The thresholds for the different parameters from test 1 showed that they all follow the threshold

¹This is because the slope ratio is the derivative of the energy decay curve. The integral over an interval in slope ratio therefore corresponds to the difference in the energy decay curve for this interval

for the strength of the reflection. By looking at the correlation coefficient between the parameters and the threshold in P_{rel} it could be seen that ΔEDC shows the best correlation with P_{rel} . The criteria by Dietsch and Kraak was not that sensitive to an increase in reflection strength, which was found from the Monte Carlo simulation and this is probably the reason to why this parameter has a lower correlation with P_{rel} . The peak value of the slope ratio on the other hand, was sensitive but still not well correlated with P_{rel} . This might have to do with its arbitrariness found from the Monte Carlo simulation, with other words it had a larger relative standard deviation.

Although parameter EK is shown to be insensitive, it should be reminded that the threshold is for detecting the reflection by listening to the impulse response only. In a real case scenario with a signal of speech or music, it might be sufficiently good in tracking the reflection strength.

Dependence on delay time, volume and reverberation time

The parameters showed a dependence on the delay time for all parameters and the detection threshold was independent on volume for all three parameters in contrast to a dependence for the reflection strength. This is a good feature of the parameters, since the volume does not have to be taken into consideration when finding the detection threshold for a measured impulse response. For ΔEDC this was also the case for the reverberation time at least for longer delay times. However, only three different volumes and reverberation times have been used in this experiment and even smaller, or larger, volumes might have an effect on the threshold.

The independence of parameter ΔEDC of volume can be described from figures 4.2a and 4.8a. The detection threshold in terms of relative pressure decreases as the volume increases (Figure 4.8a), but larger volumes results in a larger ΔEDC (Figure 4.2a) for a constant relative pressure. These two results compensates to make the threshold in terms of ΔEDC independent of volume. The same relation can be seen between the relative pressure and parameters R_{slope}^{max} and EK , and therefore these parameters are also independent of volume. The behaviour of the parameters from the Monte Carlo simulation shows that EK is less sensitive to reverberation time compared to R_{slope}^{max} and ΔEDC which is probably the reason why this parameter still has a dependence on reverberation time.

As discussed above, the detection threshold for measure R_{slope}^{peak} shows a larger relative variation than the other two measures. Still, R_{slope}^{max} is useful in detection of strong reflections since it is very sensitive to a increase in energy at the receiver.

5.2 Conclusions

The simulation method

Using statistical methods resulted in a trustfully impulse response for the purpose of this project. The convolution with a speech signal sounded realistic. Furthermore, the reverberation time calculated from the impulse response and the expected reverberation time defined when designing the room parameters coincided very well.

However, it should be kept in mind that much information have been lost using this simulation approach. The directional information was not taken into account, and this is of large importance for the subjective experience of a room since the human ear is sensitive to direction. This also has an influence on the detection of strong reflections [20]. The modified impulse responses which included a strong reflection also sounded realistic but it is at the same time hard to compare this experience with a real situation, with the same reverberation time, volume and echo delay.

Performance of the parameters

The instantaneous slope of the decay curve is an indicator of an increase in energy and this can be used to detect strong reflections. As the strength of an reflection increases, the subjective experience is affected and the parameters R_{slopes}^{max} , ΔEDC and EK follows this experience. For R_{slope}^{max} the result was more arbitrary, since the realization of the Poisson process has a large influence on this parameter. The other two parameters were more stable in this sense.

None of the three parameters showed significant difference for different volumes, which is a good property. All three showed a significant difference for different delay times. ΔEDC was also independent on reverberation time. With this in mind, ΔEDC is a good parameter for detecting reflections that can be perceived as an echo since there is no dependence on either volume or reverberation time.

By comparing the threshold for the parameters with the threshold in reflection strength it was found that ΔEDC showed the best correspondence. EK was not sensitive enough for this purpose, and R_{slope}^{max} was too sensitive and not robust enough but this parameter could still work as an indicator.

From these results a suggested method of echo detection is to first identify the echo by calculating the slope ratio and see if R_{slope}^{max} is above a threshold, and then find the drop in energy around this point to compare with detection threshold.

5.3 Further considerations

One major fact of this project and the set-up of the listening test is that only the detection thresholds by listening to the impulse responses have been found. This is of course not a realistic situation, but it still gives a impression of the behaviour of the parameters and the result can be regarded as the lowest threshold of detection. Next step would be to convolve the impulse responses with anechoic recordings of speech or music and then find the detection threshold or the threshold of annoyance.

Further, it would be interesting to see how the parameters perform in testing the annoyance level of the speaker oneself and this could be done with a real time convolution. As discussed in Section 1.1 the experience of an echo for the talker is a problem for the whole performance, and therefore the threshold with this experiment set-up would be preferable.

The results should then be compared with a real acoustic situation. The value of the parameters from a measure impulse response, compared with the subjective experience at this point will determine if the simulation method was sufficient and if the parameters have the same behaviour in the real world.

The parameters could also be improved further such that it is independent on reverberation time and delay time as well. The reverberation time could be accounted for if measured simultaneously.

Furthermore, the linear regression was a bit ambiguous since there were only three points used. By repeating the experiments for other reverberation times and volumes, a more reliable result can be accomplished. In this project the dependence on volume was investigated for only one reverberation time, but it is likely that the dependence on volume will show different results for other

reverberation times.

By modifying the reflection density at various positions in time a impulse composes a echo in a sufficient way, but further comparison with real echo-problematic situations is preferable.

Physical Constants

| | |
|--|-----------------|
| Absorption coefficient | α |
| Speed of Sound [m/s] | c |
| Frequency [Hz] | f |
| Sampling frequency[Hz] | f_s |
| Phase [rad] | φ |
| Wave number [rad/m] | k |
| Reflection strength relative average strength | G |
| Sound pressure level [dB] | $L(t)$ |
| Mean free path [m] | l_m |
| Number of reflections | N |
| Reflection order | N_r |
| Angular frequency [rad/s] | ω |
| Sound pressure [Pa] | p |
| Volume velocity [m ³ /s] | Q |
| Density of air [kg/m ³] | ρ |
| Time [s] | t |
| Delay time[ms] | t_{delay} |
| Reflection strength relative direct sound [dB] | P_{rel} |
| Distance [m] | r |
| Reflection density [reflections/s] | $\frac{dN}{dt}$ |
| Room surface [m ²] | S |
| Volume [m ³] | V |
| Estimated mean | μ |
| Estimated standard deviation | σ^2 |
| Estimated intersection with y-axis for a regression line | β_0 |
| Estimated slope, regression line | β_1 |

Symbols

| | |
|-------------------|------------------------------------|
| EK | Echo criteria by Dietsch and Kraak |
| EDC | Energy Decay Curve [dB] |
| ΔEDC | Drop in Energy Decay Curve [dB] |
| R_{slope} | Slope Ratio |
| R_{slope}^{max} | Max value of Slope Ratio |
| RT | Reverberation time [s] |
| $\%RSD$ | Relative variation |
| ANOVA | Analysis of Variance |
| MS | Mean Square |
| SS | Sum of Squares |
| df | degrees of freedom |

Bibliography

- [1] Kuttruff, H. *Room acoustics*, 5th edition. (Taylor & Francis, New York, 2009)
- [2] Wargert, S., Óskarsdóttir, G., Hansen, I. "Örebro Concert Hall - project in Architectural Acoustics" Acoustic Technology, DTU Electrical Engineering, 2013
- [3] DS-EN 3382-1 "Acoustics - Measurement of room acoustic parameters - Part 1: Performance spaces"
- [4] Morfey, C.L., *Dictionary of Acoustics* (Academic Press, 2001)
- [5] Moore, B. *An introduction to the psychology of hearing* (Academic press, 2003)
- [6] Plack, C. J. *The sense of hearing* (Taylor and Francis Group, New York, 2005)
- [7] Handel, S. *Listening - An Introduction to the Perception of Auditory Events* (The MIT press, 1991)
- [8] Haas, H. "Über den einfluss eines einfachechos auf die hörsamkeit von sprache", ACOUSTICA 1.2 (1951)
- [9] Muncey, RW., Nickson, AFB, Dubout, P. "The acceptability of speech and music with a single artificial echo" ACOUSTICA 3.3 (1953): 168.
- [10] Seraphim, H. P. "Über die wahrnehmbarkeit mehrerer rückwürfe von sprachschall", ACOUSTICA 11.1 (1961)
- [11] Bolt, Doak "A tentative criteria for the short-time transient response of auditoriums". JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA 22.4 (1950)

-
- [12] Dietsch L, Kraak W. "Ein objektives Kriterium zur Erfassung von Echostörungen bei Musik- und Sprachdarbietungen". ACOUSTICA 60 (1986): 205-216
- [13] Lovstad, A.(2003) "Evaluation of Objective Echo Criteria", M.Sc, Norwegian University of Science and Technology, Norway
- [14] Jeong, C.-H., Jacobsen, F. and Brunskog, J. "Thresholds For the Slope Ratio In Determining Transition Time and Quantifying Diffuser Performance In Situ". JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA 132.3 (2012): 1427-1435.
- [15] Jeong, C.-H., Brunskog, J. and Jacobsen, F. "Room Acoustic Transition Time Based On Reflection Overlap (L)". JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA 127 (2010): 2733-2736.
- [16] , Reichardt, W. et al "Raumeindruck als obergriff von räumlichkeit und Halligkeit, Erläuterungen des Raumeindrucksmaßes R" ACOUSTICA 40.2 (1978): 277-290,
- [17] Papoulis, A. *Probability, Random Variables and Stochastic Processes*, 2nd edition. (McGraw-Hill Book Company, 1984)
- [18] Schroeder. "New Method of Measuring Reverberation Time". JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA 37.3 (1965): 409-412.
- [19] Jacobsen, F. and Juhl, P. *Radiation of Sound*, Acoustic Technology, DTU Electrical Engineering
- [20] Burgtorf, W., Oehlschlägel, HK. "Untersuchungen über die richtungsabhängige wahrnehmbarkeit verzögerte schallsignale" ACOUSTICA 14.1 (1964)
- [21] ISO 9613-1 "Acoustics - Attenuation of sound during propagation outdoors - Part 1: Calculation of the sound absorption by the atmosphere"
- [22] Brüel & Kjær (2009). *Head and Torso Simulator Type 4128 - Product data*, Nærum, Denmark.
- [23] Brunskog, J. and Jeong, C.-H., (2013) "Shot noise Poisson process model of a RIR, autocorrelation and the transition time". Unpublished manuscript.

-
- [24] <http://www.hoertechnik-audiologie.de/psylab/>
 - [25] Paulson, T. (2007) *Psychoacoustic Measuring Methods*, Acoustic Technology, DTU Electrical Engineering: Lecture note 31230-08
 - [26] Freund, J. E. *Modern Elementary Statistics*, Third Edition. (Prentice-Hall, New Jersey, 1967)
 - [27] Terry Dielman, T. *Applied Regression Analysis*, 3rd edition (Thomson Learning)
 - [28] Bloom, G. *Sannolikhets teori och statistikteori med tillämpningar*, (Studentlitteratur, 2004)

Appendix A

Listening tests

A.1 Instruction for the listening test

Instruction to the subjects before listening test 2. The instructions were similar for test 1.

”You will now listen to impulse responses for rooms with various reverberation time and volume. The impulse response has been modified to include a strong reflection with a certain delay time, either 100 ms, 150 ms or 200 ms. The aim of the experiment is to investigate how strong this reflection has to be in order to detect it. For each repetition you are going to hear 3 different impulse responses, two without strong reflections and one with a strong reflection and your task is to identify which of the three samples that had a strong reflection. The parameters time delay of the reflection, reverberation time and volume are fixed for each run and the test variable is the strength of the reflection.

Before the test starts, you will hear one impulse response with different strength of the reflection. It will start with the impulse response free from reflections and then the strength of the reflection will decrease. That is to make you familiar with the test signal.

After that you will make one test-round for training, before the real test starts. When half of the test is done (6 runs) you are very welcome to take a break.”