Auditory localisation of low-frequency sound sources

Mădălina Anamaria Năstasă

School of Electrical Engineering

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo 29.07.2022

Supervisor and advisor

Prof. Ville Pulkki

Aalto University
School of Electrical Engineering

Aalto University
School of Electrical Engineering

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of the master's thesis

Author Mădălina Anamaria Năstasă

Title Auditory localisation of low-frequency sound sources

Degree programme Computer, Communication and Information Sciences

Major Acoustics and Audio Technology **Code of major** ELEC3030

Supervisor and advisor Prof. Ville Pulkki

Date 29.07.2022 **Number of pages** 59+1 **Language** English

Abstract

This thesis investigates the extent of the human ability to perceive changes in the direction of low-frequency sound sources, within the range of 31.5 and 100 Hz, in a large, highly damped environment. As it stands, this investigation is the only attempt made at determining the MAA for the low-frequency spectrum. In addition, the binaural cues utilised by the auditory system that enable this localisation ability have been studied. These objectives have been reached through two experiments, a relative localisation and a lateralisation task, conducted using a 2AFC method. The first psychoacoustic experiment, with the subwoofers positioned at 0°, 10°, 20° and 45° in the left azimuth plane, disprove the general held belief that there is little directional information or that localisation is limited to a left-right discrimination in the low-frequency spectrum. Results indicate that humans can reliably detect a change as small as 10° in the direction of pure tones with a frequency of over 63.5 Hz, and octave bands of pink noise with a centre frequency as low as 31.5 Hz. For sinusoids under 63.5 Hz, a separation angle higher than 45° is needed for reaching the MAA threshold. Furthermore, using an auditory model to extract binaural cues from the experimental setup, the concept of MAITD, described as the minimum ITD difference between two subwoofer positions that results in 75% of correct responses, has been introduced. The second experiment utilised variations in these extracted binaural cues in a psychoacoustic test conducted over headphones. Results show that ITD is the primary cue utilised by the auditory system to determine changes in the location of a low-frequency sound source. Moreover, manipulation of the ITD cues in the lateralisation task according to the computed MAITD of the previous experiment result in similar percentages of correct responses across both experiments and a displacement of the auditory events. It has also been demonstrated that the presence of modal behaviour in the listening environment hinders the localisation ability, as it leads to erroneous ITD cues. This fact could have contributed to the belief that, as a rule, low-frequency sound sources are not localisable, a view which this thesis aims to refute.

Keywords Auditory localisation, low-frequency, MAA, ITD

Preface

First and foremost, I am grateful to my supervisor Prof. Ville Pulkki for supporting and mentoring me alongside the journey of this thesis, and for bringing this topic to my attention to begin with.

I am also grateful to Entropy and its beautiful people for providing the loud bass, a place to call home and being a source of life inspiration.

I am grateful to my mom Gabriela and Neluțu for their trust in me and support alongside my studies, and all the love from so many kilometers away.

It goes without saying that I am grateful to all my friends for being such a strong presence in my life.

Vilppu, I am grateful to you for our life of sound and love.

Lastly, I am grateful for the silence within, as it is the empty canvas upon which the colours of sound can be painted.

Otaniemi, 29.7.2022

Mădălina A. Năstasă

Contents

Abstract 3

Preface 4

Contents 5

Symbols and Abbreviations 7

1 Introduction 8

2 Spatial hearing 10
- 2.1 Localisation cues . . . 10
  - 2.1.1 Interaural time difference . . . 11
  - 2.1.2 Interaural level difference . . . 12
- 2.2 Localisation accuracy . . . 13
  - 2.2.1 Localisation in the horizontal plane . . . 13
- 2.3 Modelling of spatial hearing . . . 14
  - 2.3.1 Model structure of auditory periphery . . . 14
  - 2.3.2 Model structure of binaural hearing . . . 14
- 2.4 Spatial hearing at low frequencies . . . 16
- 2.5 Summary . . . 17

3 Low-frequency localisation: A current understanding 18
- 3.1 Studies claiming that there is little directional information in the low-frequency spectrum . . . 18
- 3.2 Studies claiming that there is directional information in the low-frequency spectrum . . . 21
- 3.3 Summary . . . 23

4 Methodology 24
- 4.1 Methods of evaluating localisation abilities . . . 24
  - 4.1.1 Relative localisation and the MAA method . . . 24
  - 4.1.2 Lateralisation . . . 25
- 4.2 Experiment 1: Sound localisation ability in the anechoic chamber . . . 26
  - 4.2.1 Test Setup . . . 26
  - 4.2.2 Acoustical measurements of the test setup . . . 29
  - 4.2.3 Apparatus and Stimuli . . . 30
  - 4.2.4 Subjects . . . 30
  - 4.2.5 Procedure . . . 31
  - 4.2.6 Results . . . 32
- 4.3 Experiment 2: Lateralisation with binaural cue attributes manipulation 34
  - 4.3.1 Extraction of ITD and ILD cues . . . 34
  - 4.3.2 Test Setup . . . 37
  - 4.3.3 Apparatus and Stimuli . . . 37

4.3.4 Subjects 38
4.3.5 Procedure 38
4.3.6 Results 39

5 Discussion **41**
5.1 First experiment 41
5.1.1 Determination of MAA 43
5.2 Second experiment 49

6 Conclusion and Future Work **54**
6.1 Future Work 55

References **57**

Symbols and Abbreviations

Abbreviations

ITD	Interaural Time Difference
ILD	Interaural Level Difference
IPD	Interaural Phase Difference
2AFC	Two Alternative Forced Choice
IACC	Interaural Cross-Correlation
MAA	Minimum Audible Angle
MAITD	Minimum Audible Interaural Time Difference
MAILD	Minimum Audible Interaural Level Difference
DOA	Direction of Arrival
SPL	Sound Pressure Level
dB(C)	Decibel C Weighted
IR	Impulse Response
KEMAR	Knowles Electronics Manikin for Acoustic Research
BRIR	Binaural Room Impulse Response

1 Introduction

For several decades, the capacity of humans to localise sound sources, also known as auditory localisation, has been the subject of intensive study. What is referred to here as sound source localisation can be described as the perception of the source direction, distance and spatial extent [21]. From a biological point of view, the most crucial aspect of a sound source is its location, as it has been key to our survival by allowing us to detect family, nourishment, as well as predators. Apart from offering a clear evolutionary advantage to its possessors, spatial hearing stands as the pillar for understanding and being in one's environment.

The mechanisms behind sound source localisation have been extensively studied through auditory experiments since the early 1900s [23]. Auditory experiments are typically performed through a series of subjectively based tests of physical and psychological phenomena that play a role in spatial hearing.

However, unravelling the mechanism enabling spatial hearing is a rather complex task that is dependent upon numerous individual factors, including the size and form of the head, geometry of the ear, the nature of the sounds involved, and the attributes of the listening environment. Due to these challenges, discrepancies arise between our understanding of different spatial hearing areas. One of these areas is the topic of low-frequency sound source localisation. In this study, a low-frequency sound source will be considered to be below 100 Hz. Although little work has focused on this topic, low-frequency sound source localisation has succeeded in dividing the audio world into two opposing camps: some argue that it is beyond humans' capacity to localise a source below a certain frequency, while others are firmly convinced that this task is indeed possible down to the lowest audible frequency. A review of the existing studies on the matter [11], suggests that the lack of consensus occurs as a result of differences in experimental conditions.

For instance, studies supporting the notion that low frequencies are not localisable [3, 25, 28, 30] have been conducted in small listening rooms, using little dampening with the subwoofers placed in positions that can excite room modes. Room modes occur as an interaction between the radiating source and its reflections within the room, resulting in two types of interference: constructive interference (nodes) and destructive interference (nulls). Constructive interference amplifies the sound, while destructive interference reduces the sound, resulting in a pattern of higher and lower sound pressure levels within the room. However, room-mode nodes can hinder localisation due to large phase differences, thereby creating conflicting binaural cues at the ears of the listener. Thus, it is not a surprise that the studies supporting low-frequency localisation [4, 24, 15, 8] have been conducted in rooms that either have a less strong influence on reverberation or are very highly damped.

Although these studies have provided insights into low-frequency localisation, no work has yet attempted to determine whether humans can detect low-frequency directionality in an environment where the effects of the room are highly attenuated. Such a study might have been difficult to conduct in the past due to the challenges posed by the low frequency range on the experimental setup. For instance, the experimental environment needs to fulfill the requirements for anechoic conditions down to the lowest audible frequencies. However, up until recently, anechoic room conditions at low frequencies have rarely been encountered in acoustical laboratories because such a room would need a considerable amount of absorbent material for long-wavelength sound sources. Moreover, the experimental environment would also require multiple loudspeakers capable of reproducing low-frequency signals at high pressure levels with as little distortion as possible.

Recent advancements in the field have made it possible for these conditions to be more prevalent in acoustical laboratories and have enabled this low-frequency localisation topic to be pursued further. Thus, this thesis sets out to investigate the extent of low-frequency localisation using an experimental environment that possesses minimal modal behaviour in the low frequency range of interest.

The aim of this work is to determine the extent of the human ability to localise the direction of a low-frequency sound source, as well as to identify the cues used by the auditory system for enabling this localisation in the low-frequency spectrum. These aims will be accomplished through two experiments. The first experiment tests the ability of the listeners to perceive changes in the direction of the low-frequency sound events in a nearly anechoic chamber. From the listening setup of this experiment, the binaural attributes of the ear input signals are extracted and used in the second experiment. In the second experiment, variations in these attributes were presented over headphones to human subjects in order to identify the cues causing changes in the perceived direction of the sound event.

The remainder of this thesis is divided into five chapters. Chapter 2 provides an overview of spatial hearing, with an emphasis on the main binaural cues for auditory localisation in the horizontal plane; presents a simple auditory model to extract ITD and ILD cues from the ear input signals; and discusses plausible cues for low-frequency sound localisation. Chapter 3 reviews the literature on the localisation perception of low-frequency sound sources. Chapter 4 describes the methodology and the experiments, as well as presents the results. Chapter 5 discusses the results by drawing conclusions from the results and comparing these results to previous research. Chapter 6 concludes the thesis by describing the limitations of this study and proposing further work.

2 Spatial hearing

Spatial hearing allows one to orientate in its own environment, and through the properties of reflected and reverberated sounds, it can also lead to deciphering the attributes of the listening environment and its acoustic properties [21]. Blauert (1997) describes localisation as the law or rule by which the location of an auditory event is related to a specific attribute of a sound event in an acoustic environment. Strictly speaking, the task of localisation is concerned with how well the perceived direction of a sound source corresponds to its actual direction. Localisation blur, which is a property of localisation, is the smallest change of a specific sound event attribute that results in a change in the localisation of the auditory event. By studying the localisation blur, we can measure the resolution of the auditory system. Research has shown that the region of most accuracy lies within the forward direction, where the lowest limit for localisation blur is about 1°, and as the sound source moves towards right or left, this value increases. Localisation in the angles of elevation is less precise.

Spatial hearing is the result of the processing of two types of cues, which are referred to as monaural and binaural hearing. While monaural hearing is concerned with listening conditions where no differences between both ear signals are present, binaural hearing is the case in which the ear input signals are not identical and interaural differences arise.

In this chapter, the focus will be placed on binaural hearing cues with an emphasis on localisation in the horizontal plane. Furthermore, localisation accuracy will be discussed, and a simple model to predict localisation cues will be introduced. Lastly, the discussion will move onto localisation in the context of low frequencies.

2.1 Localisation cues

In order to better describe the mechanisms of sound source localisation, it is important to firstly define the three planes in which spatial hearing operates in a spherical coordinate system. The median plane, which lies vertically, divides the head equally; the horizontal plane is the plane that passes through the head at ear level; and the frontal plane divides the space in the front-back direction. Figure 1 shows a graphical representation of these planes. The intersection point of all three planes is at the centre of the head, considered the origin of the coordinate system which allows for specifying the angles of a sound source relative to the head. The direction of a sound source on this coordinate system is described by: its azimuth angle φ, which is produced by projection onto the horizontal plane; its elevation angle δ, formed by projection onto the median plane; and radius distance r. The area of focus in this thesis is in the horizontal plane, where the azimuth angle of the sound source varies from -180° to 180°. In this particular plane, it is the binaural cues that have been found to be dominant in localisation. This is because when the ear input signals are no longer identical, the differences derived from these two signals are characteristic of the direction of the sound source and its distance from the listener.

The two binaural cues are *interaural time difference* and *interaural level difference*.

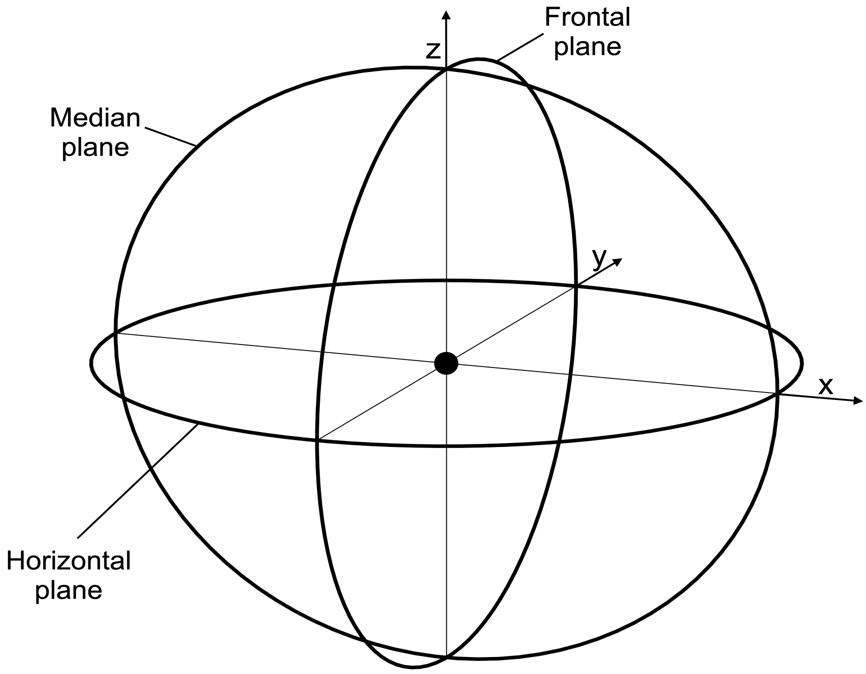

Figure 1: The spherical coordinate system and its three main planes

2.1.1 Interaural time difference

Interaural time difference (ITD) occurs as a result of the arrival time differences between the two ear signals due to the differing distances the source has to travel to the ears. This means that the signal arriving at the contralateral ear is temporarily displaced compared to the ipsilateral one when the sound source is placed closer to the ipsilateral ear. The degree of time shift can be easily computed using a spherical head model that takes into account the diameter of the head and the azimuth direction of the source. This approximation rule goes as follows:

$$\tau = \frac{D}{2c}(\varphi + \sin\varphi), \tag{1}$$

where φ is the azimuth angle in radians, D is the diameter of the head, and c is the velocity of sound. Figure 2 shows how ITD results from the propagation path difference. According to this model, for sound sources whose azimuth varies from 0° to 90°, the ITD varies from 0μs to 690μs.

The auditory system decodes ITD cues differently based on the frequency of the sound source. For low-frequency signals, up to 1.5 kHz, ITD is decoded as the phase difference between the ear input signals which provide effective information about the location of a sound source. At high frequencies, where the size of the head becomes comparable with half the wavelength of the sound source, making estimations of the direction based on phase differences becomes ambiguous. However, the auditory system extracts ITD cues at these frequencies from the temporal differences between

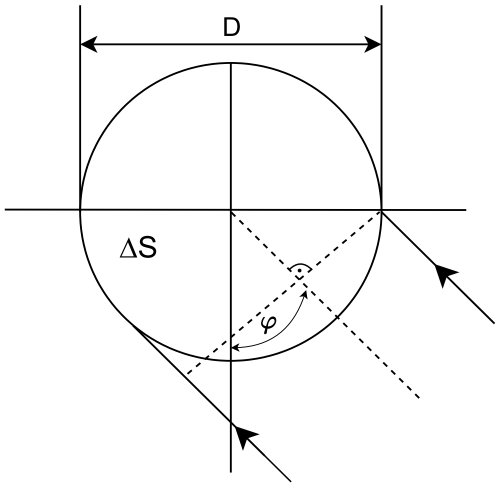

Figure 2: Diagram of the ITD caused by the propagation path difference ΔS

the envelopes of the waveforms rather than the fine structure of the signals themselves [2, 10]. Although ITD cues are existent for high frequencies, directional information is best decoded in this frequency range when ILD cues are also present.

2.1.2 Interaural level difference

Interaural level difference (ILD) is the level difference between the two ear canal signals, primarily caused by the interaction of the sound waves with the head. A sound wave from a distant source arriving at the head causes reflection at the ipsilateral ear, which results in a pressure level increase, and shadowing at the contralateral ear, resulting in a decrease of pressure level [21]. The phenomena of scattering is frequency dependent, thus the level differences between the ears increase as the frequency increases. This is because low-frequency sounds have longer wavelengths compared with the size of the head and the process of diffraction occurs, which results in little or no shadowing effect. On the other hand, at high frequencies where the wavelengths are short compared to the dimensions of the head, the phenomena of diffraction is minimum and scattering effects are more prevalent.

Kietz (1953) has observed that, for signals of any frequency, the auditory event moves fully to one side when ILD is about 15-20 dB. Due to the frequency dependent nature of scattering for plane waves, ILD is a dominant directional cue for high frequencies. However, it has been shown that humans are still sensitive to ILDs at all frequencies, but for the low frequency range, ILD remains a weak cue.

2.2 Localisation accuracy

The previous section has demonstrated that the availability of cues offered to human listeners depends highly on the nature of the sound involved. Multiple studies have shown the dependency of the signal content on the degree of localisation [2]. For instance, accuracy is best with sounds that have a broadband spectrum and strong transient information, such as noise bursts or speech, and in the case of sinusoids or narrow band signals, the performance decreases. Localisation accuracy is generally measured in free field conditions, where the effects of room reflections and reverberation are minimised, as they can negatively impact the localisation ability.

2.2.1 Localisation in the horizontal plane

Generally speaking, broadband sounds in free field conditions are located accurately, with the most precise region being located directly in front of the subject, a slight decrease of accuracy at the back and the lowest performance at the sides [2, 21].

Numerous studies have measured the spatial resolution of the auditory system by tracking the degree of change necessary to shift the position of an auditory event. The results of such studies represent, what Blauert (1997) has termed, the localisation blur. Mills (1958) measured the spatial resolution using the minimum audible angle (MAA), defined as the smallest detectable difference between the azimuths of two identical sources of sound. While the term MAA can be misleadingly thought to signify only changes in angle, Mills notes that it can also be expressed as changes in time and amplitude of the sound at the ears of the listeners, making it equivalent to the term used by Blauert. For ease of understanding, for the remainder of this thesis, the degree of change in ITD and ILD necessary to shift the position of an auditory event will be expressed as minimum audible ITD (MAITD) and minimum audible ILD (MAILD).

The measurement procedure of MAA is through a two alternative forced choice (2AFC) test. In such an experiment, the subject is presented with a signal from a reference position, followed by one from a position to one side of the reference. The task of the subject is to state whether the second auditory event lay to the right or to the left of the reference. A lateral displacement is accepted when 75% of the answers are in agreement [2]. Mills (1958) has conducted such a study using sinusoidal test signals. He has shown that the best resolution is about 1° for a reference azimuth of 0° for frequencies below 1000 Hz, where the auditory system mostly relies on ITD as a cue. At the frequency point of 1800 Hz, where both ITD and ILD offer ambiguous cues, the MAA increases slightly, and becomes impossible to detect if the reference azimuth is 60° or 75°. The resolution improves at higher frequencies because ILD can provide useful cues to sound location. It is important to note that the results observed in this test are mostly predominant for sinusoidal signals. The lateralisation of complex, broadband, impulse-like sounds is easier to detect as more cues are present. For instance, the presence of onsets and offsets, and a change in intensity or spectral structure as a function of time can help provide more reliable ITD and ILD cues.

2.3 Modelling of spatial hearing

Perceptual models, also known as psychoacoustic models, represent a theory of how the auditory system works, and are designed to explain the results of psychoacoustic experiments by matching quantitative predictions with physiological data. Such models are achieved by computational processing of digitised signals that extract metrics related to aspects of hearing [21]. For the purpose of this thesis, the focus will be placed on a model that emulates the auditory periphery, and a model for binaural hearing.

2.3.1 Model structure of auditory periphery

A simple simulation of the auditory periphery can be implemented by modelling the response of the basilar membrane, the inner hair cells and the auditory nerve fibers. These processes can be achieved with a filter bank of multiple band-pass filters, followed by signal envelope detection and a low-pass filter [21]. Figure 3 shows such a model. In the following, a detailed overview of each step in the simulation will be given.

The basilar membrane reacts tonotopically to sound, meaning that the response of each basilar membrane portion is strongest for a particular frequency [16]. As a result, each basilar membrane portion has the function of a frequency filter, with the whole basilar membrane being a bank of overlapping filters. The most commonly used method to emulate the frequency resolution of the basilar membrane is the gammatone filter bank [20]. The gammatone filter simulates the impulse response of auditory nerve fibers as estimated by reverse correlation techniques [16]. The impulse response of this filter consists of the product of two components: a carrier tone of a frequency equal to the best frequency of a particular basilar membrane site and a statistical gamma-distribution function that shapes the impulse response envelope. The gammatone filterbank has the advantage of being a relatively simple and computationally efficient way to model data pertaining to auditory frequency selectivity. Typically, an auditory model covers the frequency of 30 Hz to 18 kHz, utilising about 42 bands in the gammatone filter banks.

The inner hair cells are responsible for transforming the mechanical vibrations of the basilar membrane in electrical impulses of the auditory nerve. The dependency of the rate of impulses on the displacement of the basilar membrane can be characterised by the half-wave rectification [21]. The half-wave rectified filter-bank outputs are then fed into first or second order low-pass filters. The role of such a filter is to emulate the resistor-capacitor filtering of the inner hair cell membrane, where the synchrony between the excitation and the firing rate disappears at frequencies around 1 kHz.

2.3.2 Model structure of binaural hearing

The coincidence detection model proposed by Jeffress (1948) is the basis for the majority of binaural processing algorithms. In this model, axons which are represented by delay lines connect the ear canal inputs to coincidence-detector (CD) neurons.

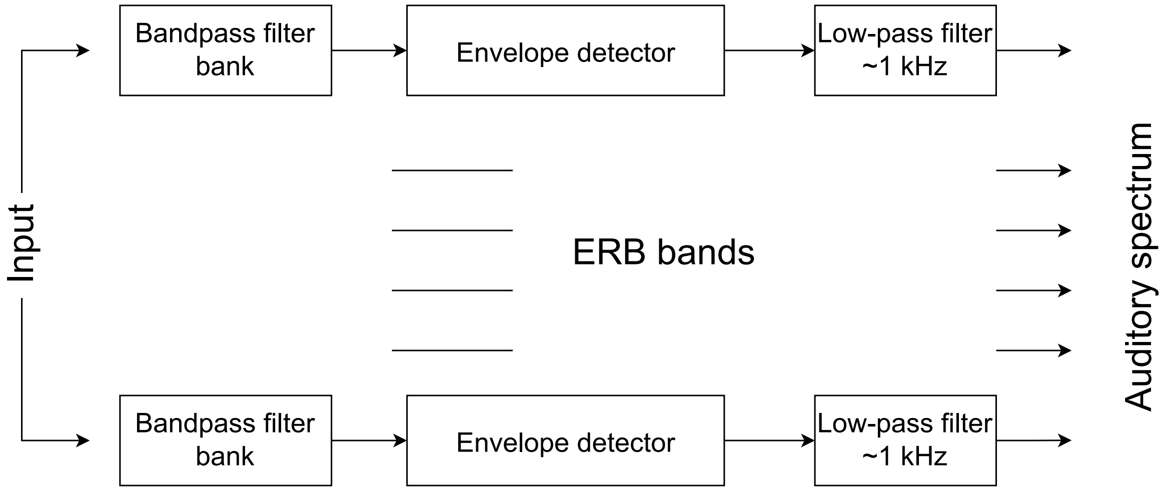

Figure 3: A diagram of the structure of an auditory model

The outputs of the CD neurons are compared with each other, and the ITD is extracted from the highest activity coincidence-detector neuron [21]. This model can be emulated computationally through the normalised interaural cross-correlation function (IACC):

$$\gamma(t,\tau) = \frac{\int_t^{t+\Delta t} x_1(T-\tau/2)x_r(T+\tau/2)dT}{\sqrt{\int_t^{t+\Delta t} x_1^2(T)dT + \int_t^{t+\Delta t} x_r^2(T)dT}}, \tag{2}$$

where t is the time, τ is the interaural delay, Δt the length of the integration window, and x_r and x_l are the ear input signals.

Extracting binaural attributes from a chosen input begins with feeding the signals into the model of auditory periphery. The result is a filter-bank of a number of bands that are half-wave rectified and low-pass filtered. In order to calculate the ITD, the IACC is computed for each band of the filterbank. The ITD is thus estimated as the interaural delay corresponding to the maximum value of the IACC function, tuned in the range of -1 ms to 1 ms (which is correspondent to the maximum ITD value between the ear canal signals). ILD can be computed by comparing the energy of the left and the right-ear signals for each frequency band of the filterbank and for each time frame of the signal. The calculation is performed on time frames to maximise accuracy. The selection of the time frame length is dependent on the application.

2.4 Spatial hearing at low frequencies

Auditory localisation at low frequencies has been studied in the context of binaural hearing attributes and auditory modelling in [4] and [18]. In this section, the possible binaural cues utilised by the auditory system will be highlighted based on the auditory modelling process performed by Braasch et al. (2004), and the psychoacoustic experiment performed by Mohamed & Cabrera (2008).

In their study, Braasch et al. (2004) investigated the ability to localise low frequencies in small listening rooms and, using a binaural model, to analyse which cues could influence this localisation. In the psychoacoustic experiment, the placement of the subwoofers was in a standard 5 channel surround setup. The stimuli were band-passed white noise bursts with three centre frequencies (31.5 Hz, 63 Hz, 125 Hz). The results of the psychoacoustic experiment show that listeners have a good discernment of the left, right and centre locations of the subwoofers, and no front to back discrimination. The authors set out to explain the results of this experiment by using a model based on the cross-correlation algorithm to estimate the target positions of the sources in the experiment. The conclusion is that it is very unlikely that the auditory system utilises monaural or ILD cues as the stimuli are only an octave wide and very low in frequency. The lack of front/back discrimination can be explained by the absence of these cues. The measured ILDs had a maximum magnitude value of approximately 2 dB and showed no distinct separation between the different speaker positions. When it comes to ITD, the cross-correlation peak varied with the speaker position in such a way that the lateral displacement of the peak always matched the speaker position tested. Thus, the authors suggest that ITD could be accounted as the main cue used by the auditory system to detect changes in the location of auditory events even for the very low frequency range. It has been further observed that the presence of reverberation can distort the cue and create conflicting information.

Mohamed & Cabrera (2008) examines the hypothesis that ITD can provide a viable cue for lateralisation in the very low frequency range of human hearing (20 Hz to 100 Hz) in an experiment that is an extension of Braasch's work. The term lateralisation is used in this context to denote left right pseudo-localisation that occurs for sources that are experienced as inside the head. A subjective test is conducted by examining the human sensitivity to lateralisation induced by interaural time differences with headphone-based production of auditory stimuli. Eight centre frequencies (comprising 20, 25, 31.5, 40, 50, 63, 80 and 100 Hz) and seven interaural time delays, in the range of -650 μs and 650 μs, were used. The results confirm that lateralisation based on ITD can occur in the very low frequency range, with the strength of the effect increasing as the frequency increases. The authors conclude that while localisation acuity is poor at very low frequencies, some lateralisation is at least possible at frequencies as low as 31.5 Hz (albeit marginal) and 40 Hz, and the acuity increases with frequency. The findings of this study support the findings of Braasch et al. (2004) in that the interaural time difference is the mechanism for lateralisation of low frequencies.

2.5 Summary

In this chapter, a general overview of spatial hearing was presented, with a broad focus on localisation in the horizontal plane and the role of binaural cues. It is clear that ITD and ILD cues pay an important part in localisation, and the localisation accuracy is both dependant on the nature of the signal and the characteristics of the environment.

Modelling of spatial hearing was introduced with a simple model of the auditory periphery that simulates human hearing, and a cross-correlation model that extracts binaural cues information. Such models will be used in the context of the second experiment of this thesis.

Furthermore, the auditory cues available for localisation in the very low frequencies were discussed in the context of the limited existing literature. The authors of these studies support the theory that low-frequency lateralisation is based solely on the decomposition of ITD cues. In the following chapter, an extensive review of the low-frequency localisation studies will be presented, with an emphasis on the methodology used by different authors and their contrasting results.

3 Low-frequency localisation: A current understanding

The studies focusing on the low frequency range have mostly questioned the role of multiple low-frequency signals in the context of modern surround sound systems. As it stands, in a standard 5.1 surround setup, the very low information is conveyed through a single subwoofer, mainly so that the main loudspeakers can be made smaller and placed more conveniently in a living room environment. Authors have questioned whether the use of multiple subwoofers offer any perceptual differences on the auditory impression. While some studies suggest that the low-frequency spectrum offers marginal directional information, other studies claim that low frequencies are localisable, and hence, the use of multiple subwoofer systems offer various benefits. Such benefits could be an increase of the auditory image width and better control of the room mode behaviour.

While this thesis does not focus on loudspeaker reproduction systems, it is important to take into consideration the results of the earlier studies as they make assumptions about the human ability to localise low-frequency sound sources. Hence, in this chapter a thorough review of the existing literature will be conducted in order to find out what has caused the lack of consensus and what is an appropriate methodology to investigate localisation in the low-frequency spectrum.

3.1 Studies claiming that there is little directional information in the low-frequency spectrum

In the first study on this topic, Borenius (1985) has investigated the effect of different crossover frequency and time delay settings on a stereophonic subwoofer system, in order to find out at what frequency and delay time the location of the subwoofer can be determined. Fifteen participants took part in the test, amongst which half were classed as "more or less experienced" listeners, while the other half were "ordinary" people. The listening space was a heavily damped small sized room (7.0 m x 5.2 m x 2.7 m), with a reverberation time of 0.25s for frequencies higher than 125 Hz. The task of the listeners was to adjust both the crossover frequency and the time delay between the loudspeakers until they found the threshold at which the sound image becomes "messy and disturbing". The author concludes that, based on the crossover frequency threshold values, there is little directional information contained in the sound signal below 200 Hz, and none at all, below 100 Hz. Furthermore, it is noted that speech signals are highly sensitive to delay errors, even with crossover frequencies below 100 Hz, and that music signals on average tolerate much higher delay errors. This shows that that relative time differences between the subwoofer and main source are more easily perceivable, than the sense of direction. For practical applications, the author recommends the use of a crossover frequency between 100 Hz and 200 Hz. Therefore, Borenius is one of first authors to support the theory that there is little to be gained from multiple sound reproduction systems in the low frequency range as there is little perception of direction.

Based on the assumption that lower frequency components do not contain relevant stereophonic information, and as a continuation of Borenius' work, Theile and Kügler (1992) set out to determine the absolute threshold of audible differences in regard to crossover frequency and subwoofer location. To achieve this, a full range stereo loudspeaker system was compared to a subwoofer placed at three different positions: centre front wall, front right corner and 90° right of the listener. As in the previous test, the characteristics of the room are that of a living room, with poor absorption at low frequencies. An AB test methodology was used to find the absolute threshold of perceptibility, with the parameters being crossover frequency and subwoofer position. As in Borenius' test, the stimuli were speech and music, but also pink noise. In total, twelve participants took part in the test. In the discussion, the authors mention that the determination of crossover frequency in relation to subwoofer placement is very difficult to investigate due to the spectral colouration caused by room modes. However, they conclude that the permissible crossover frequency of a subwoofer should be no higher than 100 Hz if one subwoofer is used. While the authors believe that one subwoofer is enough, they mention the benefits of using several subwoofers in order to maximise the uniformity of the pressure distribution that leads to a flatter amplitude response across a bigger area. While this study does not directly conclude on the directionality of low-frequency sound sources, it introduces the idea that the presence of standing waves influences the outcome of localisation, especially in non-free field conditions.

Zacharov et al. (1997) investigated the number of subwoofers needed and their position in the context of a 5.1 domestic sound system in a typically sized listening room (5.03m x 6.03 x 2.63m). The subwoofers were placed at the centre front, centre rear and 90° left of the listener, against the wall. In contrast to the sound material used in the previous research conducted on the matter, the program material used in this work was film soundtracks. The only participants in the subjective evaluation were the authors themselves, who functioned as informed expert listeners. Their conclusion was that, in all situations, the low energy frequency could not be localised. They have also argued that with a crossover frequency of 85Hz, the number, placement and delay alignment of the low-frequency sources appear to be noncritical. These results are somewhat not aligned with Borenius' work in that the delay is not a source of critical issues, although the test signals were different in nature, and the crossover frequency in the latter test was of 100 Hz. It is important to note this experiment was conducted with the intention of analysing the context of audiovisual material in a domestic environment, and that the presence of visual stimuli could interact with the auditory analysis.

Welti (2004) has set out to compare the audibility of two channel versus single channel bass reproduction in a small room environment (6.4 m x 7.3 m. x 2.7 m). Four different subwoofer configurations were tested, namely: one mono subwoofer placed at centre front; two summed mono subwoofers located at left and right front corners; two discrete left and right channel subwoofers located at the same position as the previous configuration; and two channel subwoofers located at +/- 90° relative to the listener. The stimuli used comprised of four short program loops and three music programs. The test was conducted using an ABX triangle method and in total, five trained listeners participated. The author concluded that the only noticeable difference occurred when comparing the centre mono subwoofer and the stereo configuration located to the left and right of the listener. Contrarily, no difference could be heard when comparing the front mono or stereo configurations. While Welti's test partly dismissed that audible differences would be heard across the different subwoofer configurations, the author did suggest that this test does not prove in any way that there are no merits to stereo subwoofer reproduction. In a later work [27], the author argues that the use of two subwoofers can be beneficial as it provides a more flat frequency response to a larger listening area, a claim similar to Theile and Kügler's. Tests have shown that two subwoofers placed on opposite edges of a rectangular space significantly reduce the effect of room modes over a large listening area.

In a similar work to Borenius, Kelloniemi et al. (2005) investigate the effects of subwoofer crossover frequencies in the context of a small listening room (6.35×5.58×2.71m). Four pairs of symmetrically arranged speakers at the angles 15°, 30°, 50° and 80° are used, and two different stimuli: full-frequency range pink noise and a frequency slide played on an fretless electric bass. The test was conducted using a version of the 2AFC adaptive method. The study confirms that the subwoofer location becomes detectable for a crossover frequency over 120 Hz and in certain positions in the room, but generally the direction of low-frequency sound cannot be judged. In a similar trend to Borenius, the authors argue that the audible differences are dependant upon the nature of the sound sample, but also that the subwoofer could be localised, if its level is high enough compared to other sources. In addition to this, the effect of room modes on the outcome of the results are also mentioned, as they have more influence over the results than the spatial angle between the speakers when the angle was larger than 30°.

In a general study of localisation performance of intensity stereophony, Benjamin (2006) points out that localisation below 200 Hz is negatively affected by room effects as room reflections change the phase of the summed signals at the ears. Although his findings suggest that there is little directional information below 200 Hz, in contrast to all the studies presented above, he argues that a strong sense of localisation is still experienced, although it is frequently incorrect. As a result, the author implies not only that ITD cues still operate at this low frequency, but also that localisation is strongly experienced because of them.

In a work addressing the finding of Kelloniemi et al. that a subwoofer could be localised if its level is high enough compared to other sources, Rämö et al. (2012) target to measure the effects of level and crossover frequency of broadband loudspeakers on the detection of one and two subwoofers. Three full range loudspeakers are used, with two acting as subwoofers placed at -70° and 110°, while the third is a full range high frequency source placed at 30°. The test signals used range from steady harmonic tones to impulsive and noisy sounds. Similar to Kelloniemi's methodology, the test was conducted using an adaptive 2AFC method. The experiment took place in a room conforming to the listening room standard environment and twelve subjects participated. Consistent with the works of Borenius and Kelloniemi, the authors conclude that the ease of detection of the direction of the subwoofer increases with crossover frequency, which means that when the crossover is below 100 Hz, the location of the subwoofer is not critical. Furthermore, the two subwoofer configuration reduces the perception of low frequencies as originating from a separate source. In accordance to Welti, this suggests that that dual subwoofer setups may be preferable in terms of their spatial perception qualities, in addition to their advantages related to room modes.

3.2 Studies claiming that there is directional information in the low-frequency spectrum

In one of the first pieces of work to support the ability to localise low-frequencies, Braasch et al. (2004) investigate the lateral displacement left, right and centre of a subwoofer system placed at each of the five standard surround sound locations. The test stimuli used were white noise bursts at different centre frequencies (31.5, 63 and 125 Hz). In contrast to the methodology used in the previous tests, in this work the participants were asked to indicate the speaker closest to the direction of the auditory event, which is in essence a source identification test. As in the previous studies, the test environment was a small listening room with a reverberation time (T20) of 0.2s at 63 Hz. Four listeners participated in the test, two out of which were the authors. The outcome of the psychoacoustic test shows that listeners showed no difficulty in determining whether the sound arrived from left, centre or right, while the surround-left and surround-right posed more challenges. The model simulation completed by the authors backs the finding of the psychoacoustic test in that ITD offers viable cues for low-frequency localisation. The findings of this work are in opposition to the works of Theile & Kügler and Zacharov, and the authors pinpoint that the differences in the listening environment and the test methodology might be a cause of the different results. Braasch et al. support the use of two or more subwoofers, advantageously placed at the side of the listener for larger ITDs values.

In a study with a primary focus on the extension of auditory imagery, Martens (2002) compares a mono with a stereo two subwoofer system in an anechoic environment (not strictly anechoic at the lowest frequencies) in order to answer the question whether reproduction of two subwoofer signals enable increased variation in auditory spatial imagery. The test took the form of a 2AFC test, and four listeners participated in the experiment. The results of all listening experiments reported in the paper indicate that extended control over auditory spatial imagery was provided by the use of two subwoofers to deliver decorrelated low-frequency signals. These findings raise the idea that the use of multiple subwoofers does not only benefit from control of modal behaviour in the listening space, but by presenting decorrelated signals through the system, the auditory image suffers from a positive increase in spatial image quality. In a later study conducted in two reverberant reproduction environments, characterised as "home" and "lab", Martens et al. (2005) tested the listeners' ability to discriminate between correlated and decorrelated low-frequency audio signals, emanating from multiple subwoofers. The home environment was a very small, highly reverberant home theatre space, while the lab environment was larger and less reverberant than the first space. Results show that the discrimination between the subwoofers positioned at the left and right of the listener was perfectly correct under the lab conditions for frequencies as low as 40 Hz. On the other hand, the decorrelated signals were not so easily distinguished from those that were correlated when reproduced in the home environment. In this latter case, discrimination performance was very good for octave-band noise with centre frequency of 100 Hz, but it decreased to chance levels as the centre frequency of the stimulus went below 50 Hz. This observation can provide an insight onto the claim that directional information is lost at the very low frequencies, but this is only the case when the listening environment is highly reverberant. Through measurements of the subwoofer responses at the position of the listener, the authors conclude that the discrimination was enabled by the interaural phase differences that were presented at the listener's ears, a finding in accordance to Braasch et al. (2004) and Mohamed & Cabrera (2008).

Subkey et al. (2005) argues for the potential benefits of the use of multiple subwoofers, and tried to extend the studies of Braasch et al. and Martens in the context of image size and localisation of low frequency sound. The study conducts the listening test in an anechoic environment, except that the anechoicity is not satisfied below 250 Hz, hence for the frequency range of interest, the room can be described as a small highly damped environment. Four subwoofers were used, positioned 45 degrees to the left and right, in front of and behind the subject. For the test stimuli, 1/3 octave bands of pink noise, in the frequency range of 25 to 100 Hz, were used. Fifteen subjects participated in the test and head movement was not restricted. The results yield evidence that localisation in the left-right dimension is possible down to the lowest frequencies of human hearing, while front-back discrimination is poor. The author hypothesise that the only plausible cue is the interaural time difference, and as there are no prospects for source-related spectral cues, front-back discrimination can be only incidental. Furthermore, consistent with the studies of Martens and Welti, the authors note that the image width shows an expansion through decorrelation.

Griesinger (1986) comes to a similar conclusion to Braasch et al., Subkey and Martens, concluding that humans can easily localise low frequency sources in a listening room with 20 degrees of accuracy down to 63 Hz. The author also argues that inaccuracies in localisation are caused by room mode nodes, which create conflicting ITD cues at the listener's ears.

3.3 Summary

The analysis of the previously published works on the low-frequency topic shows that the claim that there is little directional information in frequencies below 100 Hz is erroneous and only valid under certain conditions. For instance, studies supporting this claim have only been conducted in small rooms with highly reverberant conditions, very different methodologies and with a limited number of participants or inexperienced listeners. A number of authors note, however, that the localisation ability might be impeached by the presence of room modes that negatively impact the binaural cues presented at the ear of the listener. This claim holds true as the studies that support the ability to localise low-frequencies have been conducted in environments where reverberation was more controlled. Such studies have shown that discrimination between left and right sources is possible with high accuracy down to the lowest audible frequencies of human hearing. The effect of the environment has been clearly demonstrated by Martens et al. (2005) where the discrimination accuracy of subwoofer sources in large, less reverberant spaces was highly correct while small, reverberant spaces produced inaccurate results. Thus, it remains to be said that low-frequency localisation is dependent upon the conditions of the environment, but not impossible. This view is in accordance to that of Hill et al. (2012), who propose the idea that low-frequency sound sources are localisable under the right conditions. This is the first work to investigate the influence of room size, subwoofer and listener position on the outcomes of localisation ability. Both simulations and subjective results show that larger rooms allow for localisation of lower frequencies, and listeners located closer to a subwoofer exhibit less directional confusion.

Therefore, the current understanding this thesis is based upon is that a lack of localisation accuracy in the low-frequency spectrum is not necessarily due to the human lack of ability to localise these sources, but rather a result of the poor localisation cues the auditory system is offered. The behaviour of low-frequency localisation could be more in depth studied in an environment with a low degree of reverberation as the auditory system could benefit from more uncorrupted cues. Thus, the question that this thesis addresses is what is the degree of spatial resolution when it comes to determining the direction of a low-frequency sound source in a highly damped, large environment. This question has not been addressed in any of the previous work so far as the only degree of accuracy studied did not go beyond left and right discrimination. Furthermore, alongside this question, the binaural cues used by the auditory system will be also studied. Many authors noted that this localisation ability is enabled by ITD cues, but thus far, the impact of ITD and ILD values on the localisation of low-frequency sound sources has not been investigated.

In the following chapter, the methodology followed by the two experiments and their results will be introduced.

4 Methodology

Psychoacoustics is a branch of psychophysics that involves the scientific study of auditory functions through systematic experimentation. While it is not directly related to physiological research, the goal of psychoacoustics is to correlate its phenomena with the human physiology of hearing.

This thesis utilises psychoacoustic experiments to determine the human ability to localise sound sources in the low-frequency spectrum, as well as to identify the cues that enable it. In this chapter, an overview of psychoacoustic methods related to sound source localisation will be given in the first section, with an emphasis on the methodology applied in this thesis. Section 3 and 4 present the two experiments of this study: the first experiment, conducted in an anechoic chamber (slight modal behaviour is present in the low-frequency range of interest), tests the human ability to perceive changes in the direction of low-frequency sound events; and the second experiment, conducted over headphones, aims to identify which cues causes changes in the perceived direction of the low-frequency sound event. In the final section, the results of both experiments are presented.

4.1 Methods of evaluating localisation abilities

Localisation abilities are generally measured through two types of psychophysical investigations: absolute and relative localisation studies. The aim of absolute localisation studies is to determine the accuracy with which human listeners can localise the direction of a sound, by asking subject to indicate the perceived location of the source. On the other hand, relative studies focus on the spatial resolution of the auditory system by measuring the smallest change in source direction that the listener can reliably discern. The results of a relative localisation generate, what Blauert (1997) termed, localisation blur or MAA [17].

When localisation is studied in the context of headphone reproduction, unless special techniques are utilised, the listening results in inside-the-head localisation [21]. In lateralisation experiments, auditory events are controlled in the right-left direction by modifying the binaural attributes presented to the ears.

In the following, the methodology of evaluating the relative localisation and lateralisation will be discussed separately.

4.1.1 Relative localisation and the MAA method

Relative localisation experiments allow for the measurement of the human spatial localisation abilities. The minimum audible angle (MAA) gives a good indication of the spatial resolution of the auditory system as the auditory space is less differentiated than the space in which sound sources exist. MAA is defined as the smallest detectable change in angular position relative to the subject when the spacial resolution is studied using stimuli presented over loudspeakers [19]. The MAILD and MAITD, which will be also used in this thesis, refer to the smallest detectable change in ITD and ILD respectively. The MAA can be determined using the psychophysical method of a two-alternative forced-choice test procedure. In such an experimental test, the subject is presented with one sound which is considered the reference position for that trial, followed by another sound from a position to one side of the reference. The subject is then required to state whether the second auditory event lay to the right or the left of the reference. The MAA is then calculated as the angle where 75% of the responses are correct. This percentage has been chosen as it is the halfway point between 50% (the point where no subject perceives a displacement and all are guessing randomly) and 100% (the point where all subjects perceive a lateral displacement) [2].

A 2AFC procedure also benefits from the possibility of analysing the overall performance of the listeners using signal detection theory [9]. This method introduced by Wood and Bizley (2015), proposes that the sensitivity factor (*d'*) and listener bias (*Bias*) at different separation angles could be used as a measure of the spatial localisation abilities. A hit can be arbitrarily defined as leftwards choices for leftwards moving stimuli, and false alarms as leftwards choices for rightwards moving stimuli. As a result, the sensitivity index can be calculated as:

$$d' = Z_{\text{Hit}} - Z_{\text{F}_\text{a}}, \tag{3}$$

where Z(p) is the inverse cumulative distribution function of the Gaussian distribution [7].

The listener bias, with negative numbers indicating a bias towards leftwards choices, can be determined as such [7]:

$$Z_{\text{Bias}} = -\frac{Z_{\text{Hit}} - Z_{\text{F}_\text{a}}}{2} \tag{4}$$

4.1.2 Lateralisation

In a lateralisation experiment, sound is presented dichotically to subjects using headphones so that the monaural cues are strongly attenuated. Dichotic listening is the arrangement where the sounds presented to the left and right ear are processed differently. While in natural conditions, monaural and binaural attributes work together in specific combinations, this type of experiment isolates the binaural cues and varies them, while the monaural cues remain constant. This allows the experimenter to find out which attributes of the ear input signal leads to a lateral displacement of the auditory event. Because the lateralised reproduction often results in very artificial sounds and the lateral displacement is only within one dimension of the auditory space, there is no correlation between the perceived lateralisation and a specific azimuth of the auditory event. However, despite this, lateralisation studies allow for conclusions to be drawn about how the auditory system evaluates interaural signal differences.

The lateral displacement of auditory events can be measured with the same techniques used to measure the position of auditory events. These can be, in principal, absolute or relative tasks. In an absolute task, the lateral displacement can be measured by using a predefined rating scale in which the subject marks the lateral position of the auditory image. By using a relative task, the lateralisation blur can be measured. The lateralisation blur, similar to the localisation blur, represents the degree of change in the attributes of the ear input signal that leads to a just noticeable difference in the lateral displacement. Essentially, the lateralisation blur is the equivalent of MAA in the context of a lateralisation experiment.

The measurement of lateralisation blur is well suited to the two-alternative forced-choice procedure. Such a procedure benefits from several advantages. Firstly, judging the absolute position of an auditory event that is localised inside the head is a rather difficult task. It could be argued that offering the listener an acoustical pointer which, in the context of a 2AFC task is the first sound presented, could ease the task of the subject. Moreover, lateralisation is perceived differently by each subject and differences arise between how the participants use a predefined rating scale. This occurs especially if the rating scale is a slider varying from an absolute left to an absolute right. A 2AFC task requires the judgement of a lateral position in relation to the other, and hence, the inter-subject variability caused by the use of a scale is removed. In this way, the focus is not placed onto the absolute lateral position of the auditory image, but rather onto the displacement of the event in relation to the other.

In a lateralisation experiment, the interaural time differences can be manipulated by passing the signal through a delay unit on its way to the two ears. If the delay for both units is equal, it is expected that the auditory image will be positioned in the median plane of the subject. When the delay is changed at one ear, the auditory image will shift towards the ear at which the sound arrives earlier. Manipulation of the interaural level differences creates the same effect. These cues can be altered by modifying the amplitude of each signal fed to the two ears. While an equal amplitude will result in an auditory image located in the median plane, with different amplitudes between the ear signals, the image will migrate towards the ear in which the signal is louder. Thus, interaural time and level differences can lead to lateral displacements of the auditory event.

4.2 Experiment 1: Sound localisation ability in the anechoic chamber

The first experiment was designed with the aim of finding the spatial resolution of low-frequency sound sources in an environment where the influence of room modes is diminished.

4.2.1 Test Setup

The test was conducted in an anechoic environment that fulfills free field conditions, in accordance to ISO 3745 [26], down to 50 Hz with a few exceptions. The exceptions mean that there is a slight modal behaviour in the low-frequency spectrum, which will be investigated for the experimental configuration chosen. Nonetheless, for the frequency range of interest, the room can be considered a large, highly damped environment. The anechoic chamber is fitted with absorbing wedges with a length of 750 mm, together with panel resonators installed behind the wedges to control the low-frequency behaviour of the room. The background noise levels inside the chamber is extremely low, substantially below the human threshold of hearing. The dimensions of the room are wedge tip-to-tip 8.2 m x 8.2 m (width x length) and the height is 7.2 m. The room has a raised floor net structure with sound absorbing material below it.

Four Genelec 7380A active subwoofers were used in the test, with the lower cutoff frequency point at 16 Hz and the upper cutoff at 120 Hz. In the methodology of previous studies, the subwoofers have been usually placed in the standard surround setup, at 0°, ± 30° and ± 110°, or near the edges of the room. These studies have offered insight into the discrimination ability of low-frequency sound sources in small rooms that are damped, but the aim of this thesis is to investigate the spatial resolution in the low-frequency spectrum beyond left-right and front-back discrimination in a larger, more damped environment.

In order to determine the best configuration of the listening room for the experiment, a preliminary study was conducted. The four subwoofers were placed at 0°, 30°, 45° and 65° in the left azimuth plane, with the subject facing the space between the second and third subwoofer. Octave bands of pink noise bursts and sine tones ranging from the frequency of 31.5 Hz and 100 Hz were used. In total, ten subjects participated in this preliminary study that used a 2AFC test procedure. The results have shown that the smallest separation angle of this configuration yields good results even in the lower frequency bands, and a separation angle of 65° is easily distinguishable by subjects in almost all frequency bands. Thus, for the actual test configuration, it was decided that smaller separation angles should be used in order to determine a more accurate representation of the human resolution of low-frequency localisation.

Due to the large dimensions of the subwoofers, it is difficult to obtain a small separation angle between the sound sources when they are conventionally placed. Therefore, in order to achieve a separation angle of 10°, one of the subwoofers was stacked on top of two subwoofers placed on the floor. The subwoofers that were placed on the raised floor net structure had their drivers pointing downwards, and the one on top had its driver pointing upwards. The test setup configuration had the radiating sound sources placed roughly at 0°, 10°, 20° and 45°. A diagram of the placement of the subwoofers within the anechoic room can be seen in figure 4 and an actual picture of the setup in figure 5. For the remainder of the thesis, the subwoofers will be addressed by the numbers shown in the diagram 4.

The listening test setup also included a pair Genelec 8020B loudspeakers that played decorrelated pink noise with the purpose of masking any high frequency distortion components caused by the subwoofers.

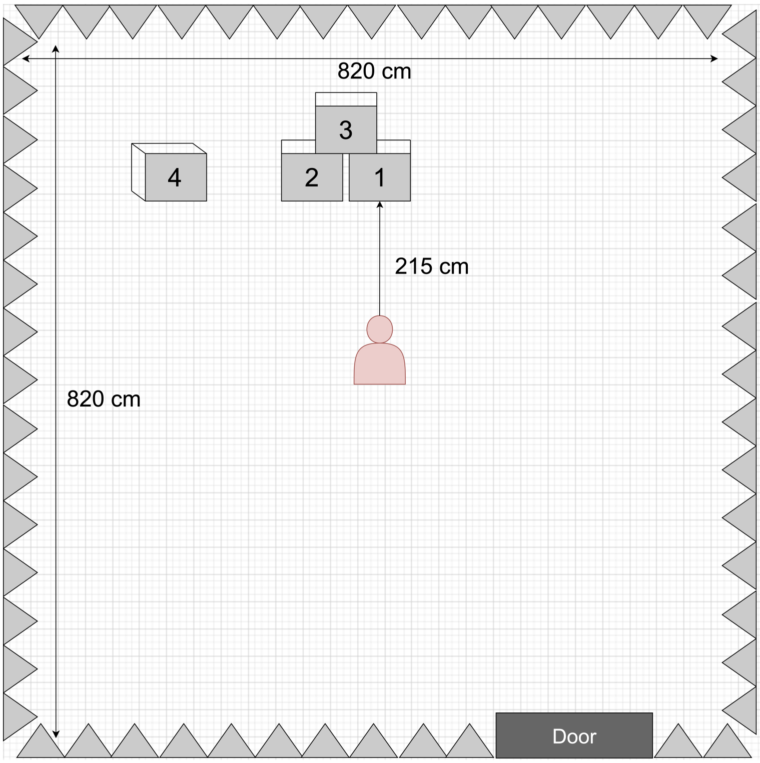

Figure 4: Subwoofer configuration within the anechoic chamber

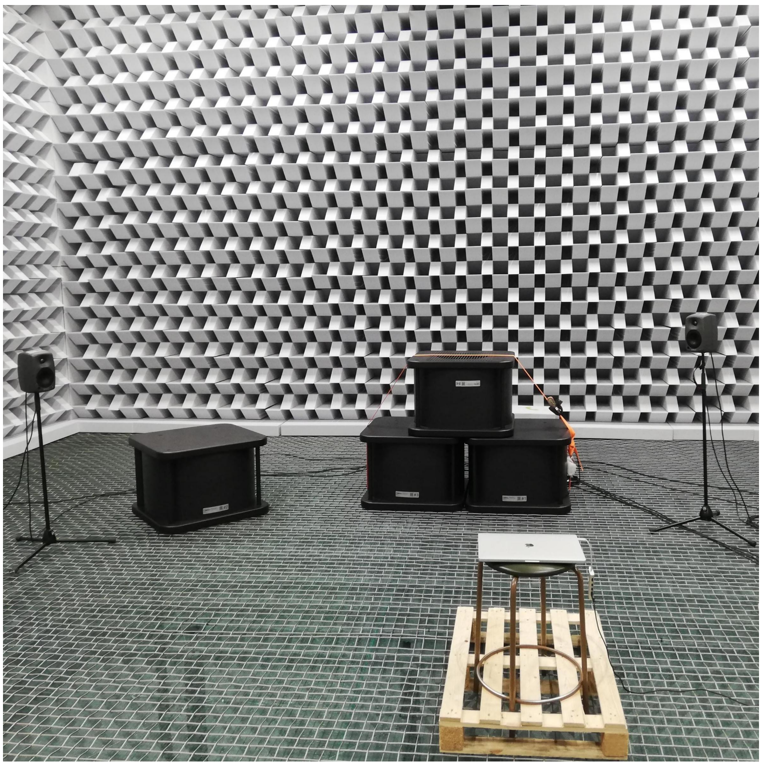

Figure 5: The listening test setup displaying the position of the subwoofers, the listener's chair and the pair of loudspeakers for decorrelated noise

4.2.2 Acoustical measurements of the test setup

Since the subwoofers were placed rather unconventionally, an evaluation of the actual position of the radiating sound sources was performed. For this purpose, the Microflown USP 3D sound intensity probe was used in order to determine the direction of the sound emitted by the subwoofers. The probe operates in the range of 20 Hz to 10000 Hz, so it could be relied upon in this low-frequency investigation. The output of the probe consists of four elements: the sound pressure P, and the three-dimensional components of the particle velocity vector U_x, U_y, U_z. The intensity vector expresses the net flow of sound energy as a 3D vector and can be computed as such [21]:

$$I_x = P^* U_x;$$
$$I_y = P^* U_y;$$
$$I_z = P^* U_z,$$

where P and the 3D particle velocity components U are expressed as frequency domain signals.

From here, the direction of arrival (DOA) of sound can be determined as the opposite of that of the intensity vector. The DOA of the sound source can be expressed in the spherical coordinate system by its azimuth and elevation angle. For the measurement, the 3D probe was placed facing the centre of subwoofer 3 and a recording of one minute of pink noise was taken for each subwoofer. The intensity vector and DOA was calculated from these recordings. Figure 6 shows the results of the calculation in the azimuth plane, which is the only plane we are concerned with in this investigation.

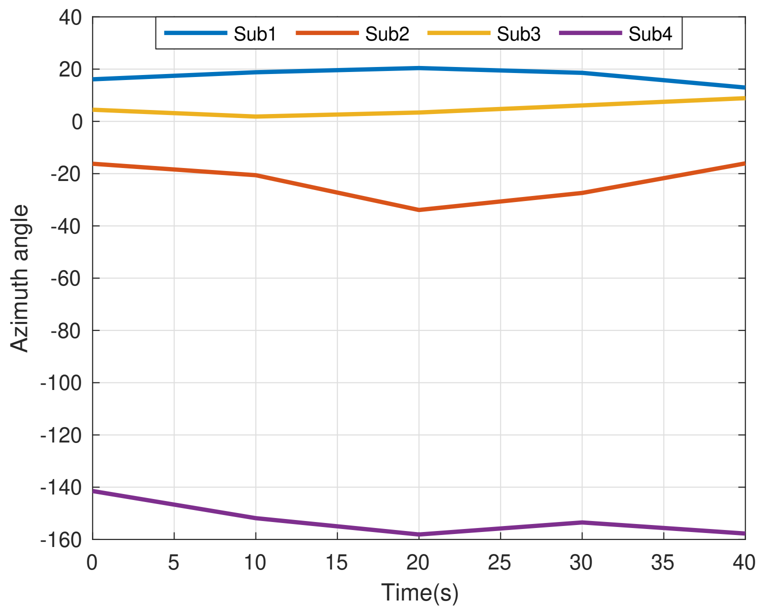

Figure 6: Azimuth position of the measured DOA of sound radiating from the subwoofers

The measurement shows that the direction of the sound sources match the positions of the subwoofers, and thus that the direction of subwoofer 3 is located between subwoofer 1 and 2 as required.

4.2.3 Apparatus and Stimuli

Two types of audio signals were used in the test: pure sinusoids and bursts of pink noise, generated using Max MSP at a sampling frequency of 48 kHz and 24-bit resolution. In order to avoid any high-frequency content as a results of clicks and to prevent the listener's use of onset effects (spectral widening) as cues, all stimuli were gated using 100ms onset and offset ramps.

The stimuli were 1000ms long and loudness matched at a level of 80 dB(C)SPL measured at the listener position. This playback level was chosen as the sound components were clearly audible and comfortable at all frequencies. The pink noise bursts were 1 octave width band-passed, with the centre frequency of 31.5, 40, 50, 63, 80 and 100 Hz. The sine tones had the same frequency range as the pink noise, except that the 31.5 Hz frequency was removed.

An important matter to consider, especially when playing sinusoids, is the presence of subwoofer non-linear high-frequency distortion products that could offer unwanted localisation cues. To avoid this, decorrelated full-spectrum pink noise was played through two separate loudspeakers at a level of 58 dB(C)SPL. The 31.5 Hz sine tone was removed from the test as it presented two loud harmonic distortion components. Masking these components would have resulted into an uncomfortably loud level of the pink noise for the duration of the test, which would have fatigued the listeners. The loudness level of 58 dB(C)SPL was chosen as to mask the loudest harmonic distortion components.

4.2.4 Subjects

The listener was positioned in the centre of the room, roughly 2m away from the subwoofers which were aligned equidistantly from the subject for an equal loudness balance. The subject was instructed to face the centre of subwoofer 1 and no head movements were allowed in the test. The chair of the subject was placed on a wooden structure for the subject to place their feet on in order to minimise the vibration of the wire net floor structure. However, minimal haptic feedback could still be felt in the knees area, but this was not considered a problem as it is a natural cue.

In total, 18 subjects took part in the test, two out of which were female, with a total average age of around 30. All subjects reported normal hearing. The participants were either researchers at the Acoustics Lab of Aalto university, students of the Acoustics and Audio Technology major or people with a musical background. Most subjects had previous experience with localisation listening tests.

4.2.5 Procedure

The listening test was a 2AFC test, with no adaptive staircase method as the number of separation angles available were rather small for such a method. In total, 6 subwoofer pairs were formed for the test, with the naming as follows:

Pair 1:	Subwoofer 3	Subwoofer 2
Pair 2:	Subwoofer 1	Subwoofer 3
Pair 3:	Subwoofer 1	Subwoofer 2
Pair 4:	Subwoofer 2	Subwoofer 4
Pair 5:	Subwoofer 3	Subwoofer 4
Pair 6:	Subwoofer 1	Subwoofer 4

The order of the pairs was formed according to the separation angle, from smallest to largest. Although the separation angle between the first and second pair is equal, they are considered different pairs as pair 1 tests the localisation ability 10° away from the median plane and should yield different results. In the test, the position of the reference was changed equally between the first and the second subwoofer in the pair. As a result, the sound sources travelled equally to the left and to the right.

The total number of conditions were: 6 pairs $*$ 2 references (either first subwoofer or second subwoofer of the pair) $*$ 6 pink noise stimulus $*$ 5 pure tones $=$ 132 conditions. In order to obtain reliable results, each condition was repeated 5 times, which results in a total of 660 trials.

The test interface was programmed in Max MSP, and for each trial the order of the subwoofer pair, the reference and stimuli were randomised. At the start of the test, a pair was randomly chosen, the reference subwoofer played followed by the second subwoofer and the subject was asked: *In which direction is the sound moving?* The subject could indicate their response by pressing either the left or right arrow on the keyboard. Before indicating their response, the subject was forced to wait for the whole duration of the stimuli (2 seconds) to play. After indicating their response, the test would progress to the next trial.

Before starting the test, to insure that the subjects are familiar with the task, they were instructed both in writing and orally. Furthermore, a practice test with a few trials was run so that they could get accustomed to the testing method, the interface and the stimuli. Subsequently, they could begin the test. In average, the subjects took about 30 minutes to complete the test.

4.2.6 Results

The results of the listening test were grouped on stimuli type and frequency, and have been plotted as a boxplot in order to show the inter-subject variability across different conditions, in figure 7 and 8. The central red mark indicates the median, the top and bottom edges are the 25^{th} and 75^{th} percentile, and the outliers are plotted using the '+' symbol. The separation angle of 10° marked with an asterisk indicates pair 1, formed by subwoofer 3 and 2.

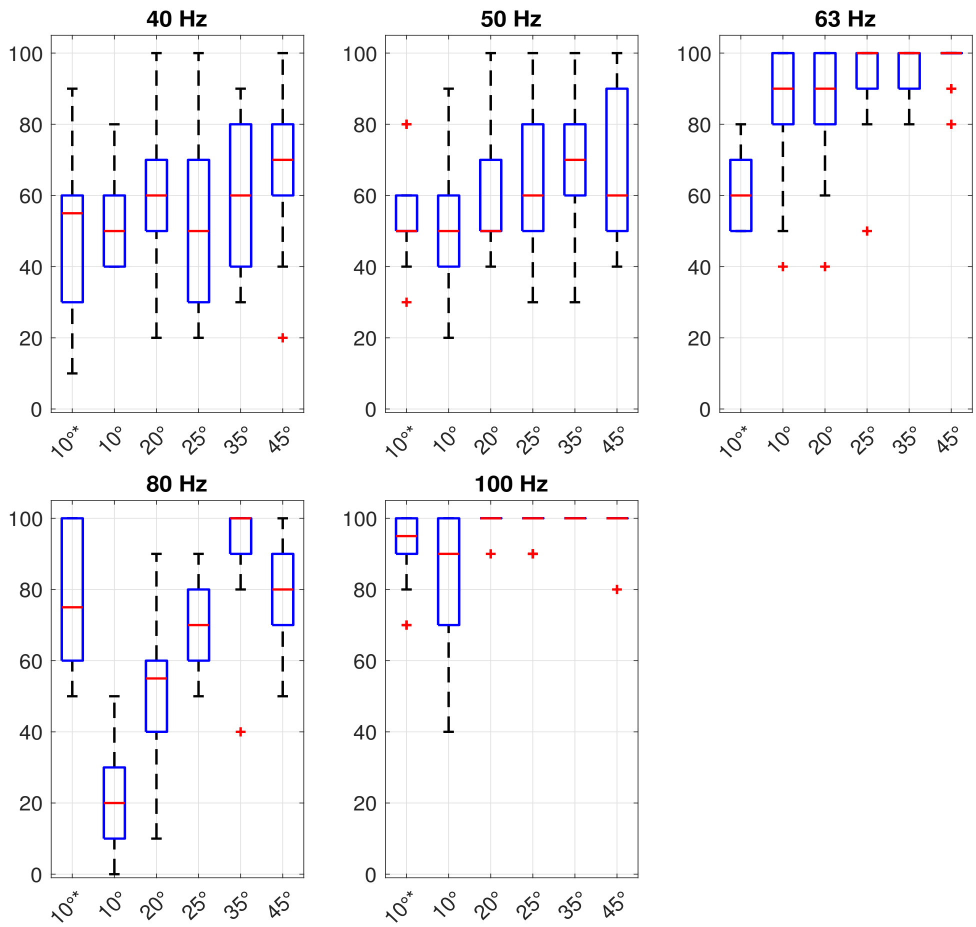

Figure 7: Boxplots of the correct percentages of experiment 1 for the sine tone stimuli

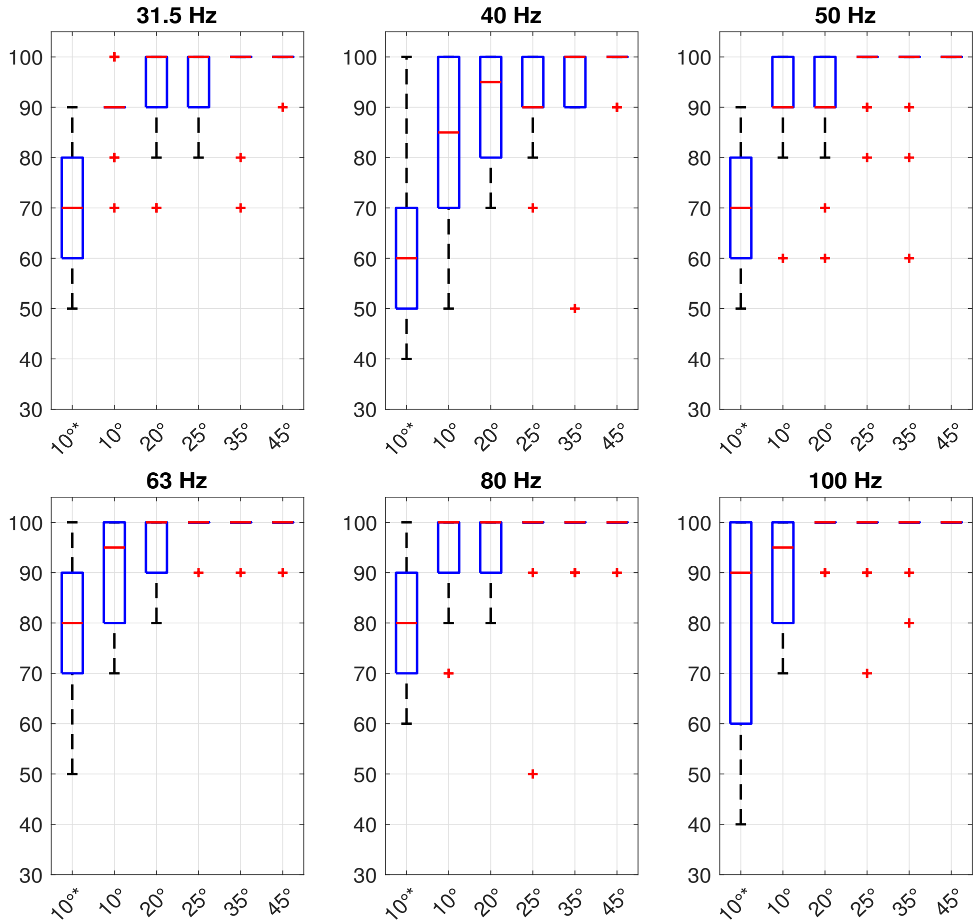

Figure 8: Boxplots of the correct percentages of experiment 1 for the pink noise stimuli

4.3 Experiment 2: Lateralisation with binaural cue attributes manipulation

The second experiment aims to identify the binaural cues that enable low-frequency localisation. For this purpose, a lateralisation experiment was designed using the ITD and ILD cues extracted from the first experimental setup. By replicating these cues, it was hoped to find whether manipulations of the same binaural attributes generate similar results to the first experiment.

4.3.1 Extraction of ITD and ILD cues

In order to extract binaural cues from the first experiment, a psychoacoustic model was implemented. The model consists of an auditory periphery part, as described in chapter 2 section 2.3, and a binaural processing cross-correlation algorithm. The auditory periphery model takes into account the behaviour of the basilar membrane, the inner hair cells and the auditory nerve fibers. These processes have been achieved with a gammatone filter bank of 7 bands (20 to 100 Hz), a signal envelope detection and a low-pass filter.

The objective was to apply the model on the same test signals as in the psychoacoustic experiment. For this purpose, KEMAR, a dummy head and torso simulator equipped with microphones in its ears, was used to record signals from the experimental setup. The sampling rate of all recorded signals was of 48 kHz. The measurements consist of binaural room impulse responses (BRIR) and binaural recordings of the sinusoids of different frequencies used in the listening test for each subwoofer location. The dummy head was placed at the same position as the subject in the experiment. Since dummy head recordings were not taken for the pink noise stimuli, the BRIRs were convolved with these stimuli in order to confer them the directional information. Therefore, the binaural model was fed with two different types of stimuli: the sine tones and the pink noise at all frequencies of the psychoacoustic test.

The ITD is determined using the IACC function between the left and right ear signals, tuned to a time lag between -1 and 1ms. The time lag was chosen as it corresponds to the maximum ITD between the ear canal signals. The ITD value for a specific auditory channel is calculated as the time lag corresponding to the most prominent peak in the normalised IACC function. A negative ITD value indicates a source perceived in the left azimuth plane, while a positive ITD indicates a source perceived to the right. The ILD can be estimated for each individual frequency channel by comparing the frame-based energy of the left and the right-ear signal representations. The temporal resolution was controlled by choosing a frame size of 20ms and a step size of 10ms. These values have been chosen as they yield accurate results for free-field conditions, and are not susceptible to interaural coherence [21]. The resulting ILD is expressed in dB, with negative values indicating a sound source positioned at the right-hand side, whereas a positive ILD value corresponds to a sound source located at the left-hand side.

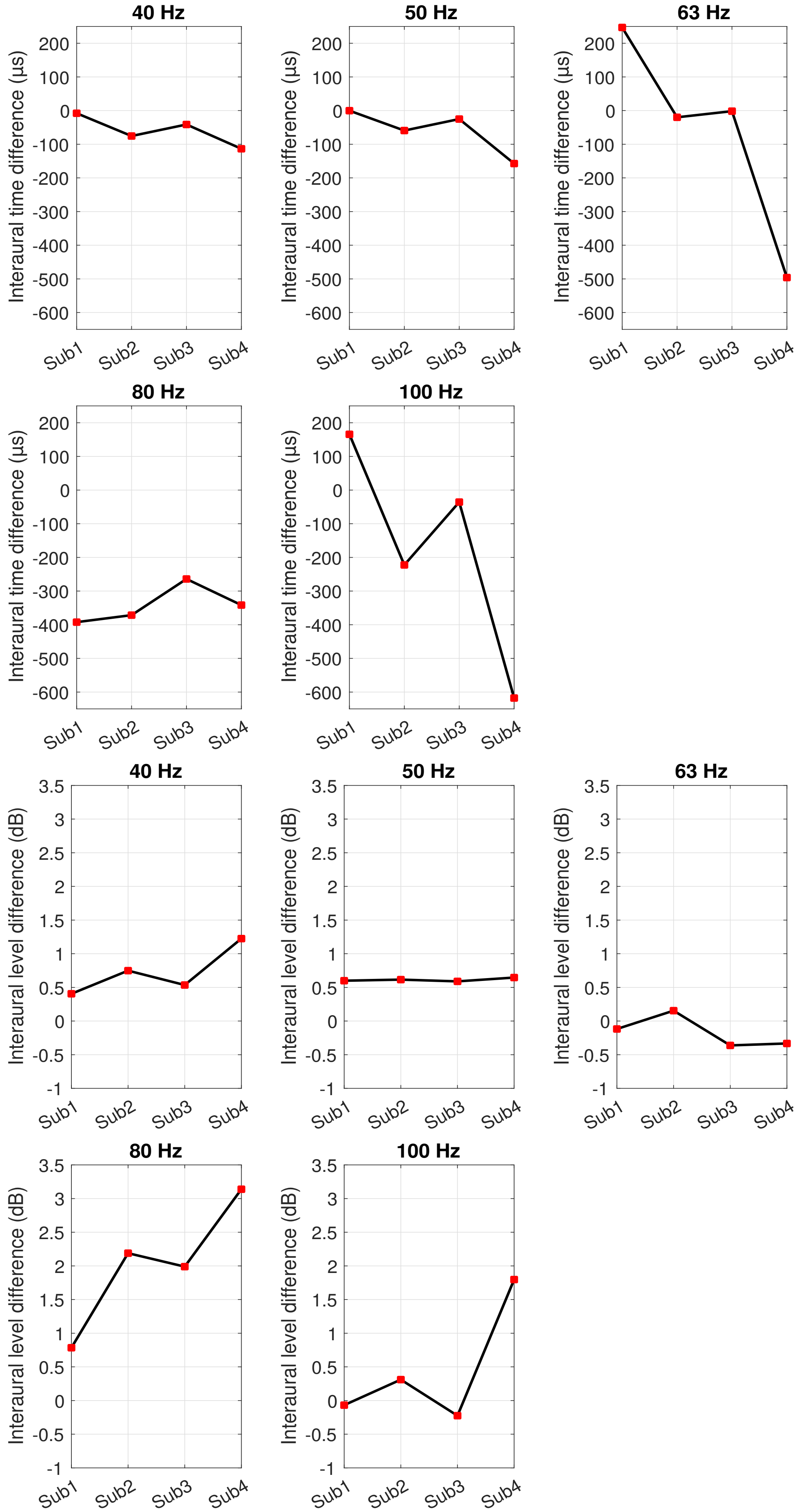

Figure 9: The computed ITD and ILD cues of the psychoacoustic model for the sine tone stimuli

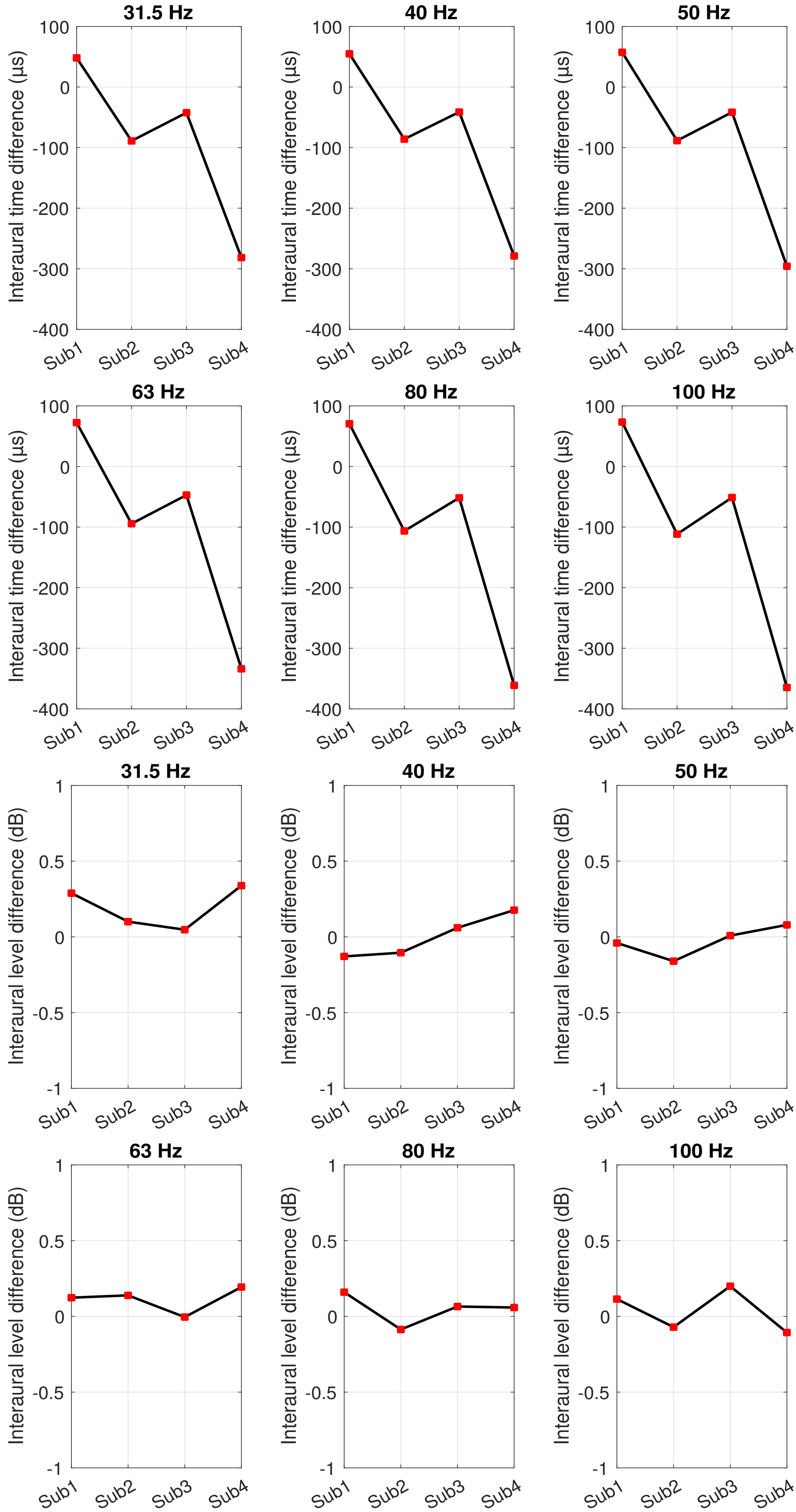

Figure 10: The computed ITD and ILD cues of the psychoacoustic model for the pink noise stimuli

The values of the computed ITD and ILD for each stimuli type and subwoofer position can be seen in figure 9, 10. The values are an average of all frequency bands of the filterbank for ease of display, and because the values did not fluctuate vastly across the different bands. For the most part, these values follow the expected trend, in that the ITD match the position of the subwoofers and the ILD values are minimal. An interesting case is the ITD and ILD values for the 80 Hz sine tone, where it appears that subwoofer 3 is the right-most position and ILD differences become significant. As these cues could have been corrupted by the presence of room modes, the significance of correct ITD and ILD cues was studied in the context of the second psychoacoustic test.

4.3.2 Test Setup

The experiment took place in the listening booths of the Acoustics lab that are meant for conducting listening tests over headphones. The listening environment has a low background noise that prevents masking and distraction from the stimuli presented. The headphones used for reproducing the stimuli were Sennheiser HD650 with an Objective 2+ ODAC amplifier.

4.3.3 Apparatus and Stimuli

The stimuli for this experiment were created by adding the extracted ITD and ILD cues to the original stimuli used in the first psychoacoustic test. The stimuli were not convolved with any HRIRs from the first listening test configuration. Instead, the ITD condition was created by time-shifting either the left or right channel according to the values measured from the psychoacoustic model. A negative ITD resulted in the right-ear channel being delayed by that time difference, while a positive ITD resulted in the left-ear channel being delayed. For the ILD condition, a positive ILD value lead to an amplification of the left-ear channel, while a negative value lead to a decrease in the left-ear channel. The third condition of the test had both the ITD and ILD cues added to the stimuli.

For the time delay and level cues, it was decided that only the values of subwoofer 1, 2 and 4 would be applied to the audio signals, with the exception of the 80 Hz sine tone. This is because the differences between subwoofer 2 and 3, while not negligible, would be difficult to identify in a lateralisation experiment, when localisation happens inside the head. This was validated through a pilot test completed by the author. For the case of the 80 Hz sine tone case, the values of subwoofer 1, 3 and 4 were chosen, and as the ITD and ILD cues from the pink noise stimuli are not corrupted by modal behaviour, an extra condition was added. This condition consists of applying the cues of the 80 Hz pink noise stimuli of subwoofer 1, 2 and 4 to the 80 Hz sine tone. This was done in order to validate whether an uncorrupted set of cues for this frequency band changes the outcome of the results. The ITD and ILD values used for each type of stimuli and frequency band have been summarised in table 1 and 2. Pairs were formed for each condition, stimulus type and frequency, so that each subset of a stimulus of a certain frequency had three different pairs.

Sine tones ITD				Sine tones ILD		
40 Hz	-8 μs	-76 μs	-113 μs	0.4 dB	0.7 dB	1.2 dB
50 Hz	-0.2 μs	-59 μs	-157 μs	0.6 dB	0.6 dB	0.6 dB
63.5 Hz	247 μs	-20 μs	-496 μs	-0.1 dB	0.2 dB	-0.3 dB
80 Hz	-392 μs	-264 μs	-342 μs	0.8 dB	2 dB	3 dB
80 Hz (PN)	71 μs	-106 μs	-361 μs	0.2 dB	-0.09 dB	0.06 dB
100 Hz	166 μs	-222 μs	-618 μs	-0.06 dB	0.3 dB	1.8 dB

Table 1: ITD and ILD applied to the sine tones, where each column represents the values of a certain subwoofer position

Pink noise ITD				Pink noise ILD		
31.5 Hz	48 μs	-89 μs	-281 μs	0.3 dB	0.1 dB	0.3 dB
40 Hz	54 μs	-86 μs	-279 μs	-0.1 dB	-0.1 dB	0.2 dB
50 Hz	57 μs	-89 μs	-296 μs	-0.04 dB	-0.2 dB	0.07 dB
63.5 Hz	72 μs	-94 μs	-334 μs	0.1 dB	0.1 dB	0.2 dB
80 Hz	71 μs	-106 μs	-361 μs	0.2 dB	-0.09 dB	0.06 dB
100 Hz	73 μs	-112 μs	-365 μs	0.11 dB	-0.07 dB	-0.1 dB

Table 2: ITD and ILD values applied to the pink noise, where each column represents the values of a certain subwoofer position

To reduce the possibility of any high-frequency content that can provide unwanted lateralisation cues, all the stimuli were low-pass filtered with a 6^{th} order Chebychev with a cutoff frequency of 200 Hz. Furthermore, all stimuli had a smooth fade-in and fade-out of 150 ms. All stimuli were loudness matched and presented at a comfortable level in the test, where they were clearly audible and created no distortion effects in the headphones. There were in total: (6 sine tones $*$ 3 subwoofers + 6 pink noise stimuli $*$ 3 subwoofers) $*$ 3 conditions (ITD, ILD, Both) = 108 stimuli. In order to obtain reliable results, the stimuli were repeated 4 times, which results in 432 trials.

4.3.4 Subjects

In total, 15 subjects participated in the test, nine out of which had taken part in the first psychoacoustic experiment. All subjects reported normal hearing.

4.3.5 Procedure

The listening test followed the same procedure as the first experiment, namely a 2AFC test. The same graphical interface was used for the purpose of this test. In practice, the second experiment is a lateralisation task of the first experiment, with the difference that only three subwoofer positions were used. The reference position was changed equally between the first and the second element of the pair. Throughout the test, the conditions presented were randomised. A practice session was conducted before each test, so that the subjects would become familiar with the nature of a lateralisation task, the interface and the stimuli. In average, the participants took about 20 minutes to complete the test.

4.3.6 Results

Boxplots of the results for each stimuli type can be seen in figure 11 and 12. The central red mark indicates the median, the top and bottom edges are the 25^{th} and 75^{th} percentile, and the outliers are plotted using the '+' symbol. For all frequencies, beside the 80 Hz sine, pair 1 is formed by the binaural cues of subwoofer 1 and 2, pair 2 is formed by the binaural cues of subwoofer 2 and 4, and pair 3 is formed by the binaural cues of subwoofer 1 and 4. In the case of the 80 Hz sinusoid, subwoofer 2 is replaced by subwoofer 3. The condition marked with 80 Hz (PN) represents the case where the binaural cues of the pink noise stimuli were applied to the 80 Hz sine.

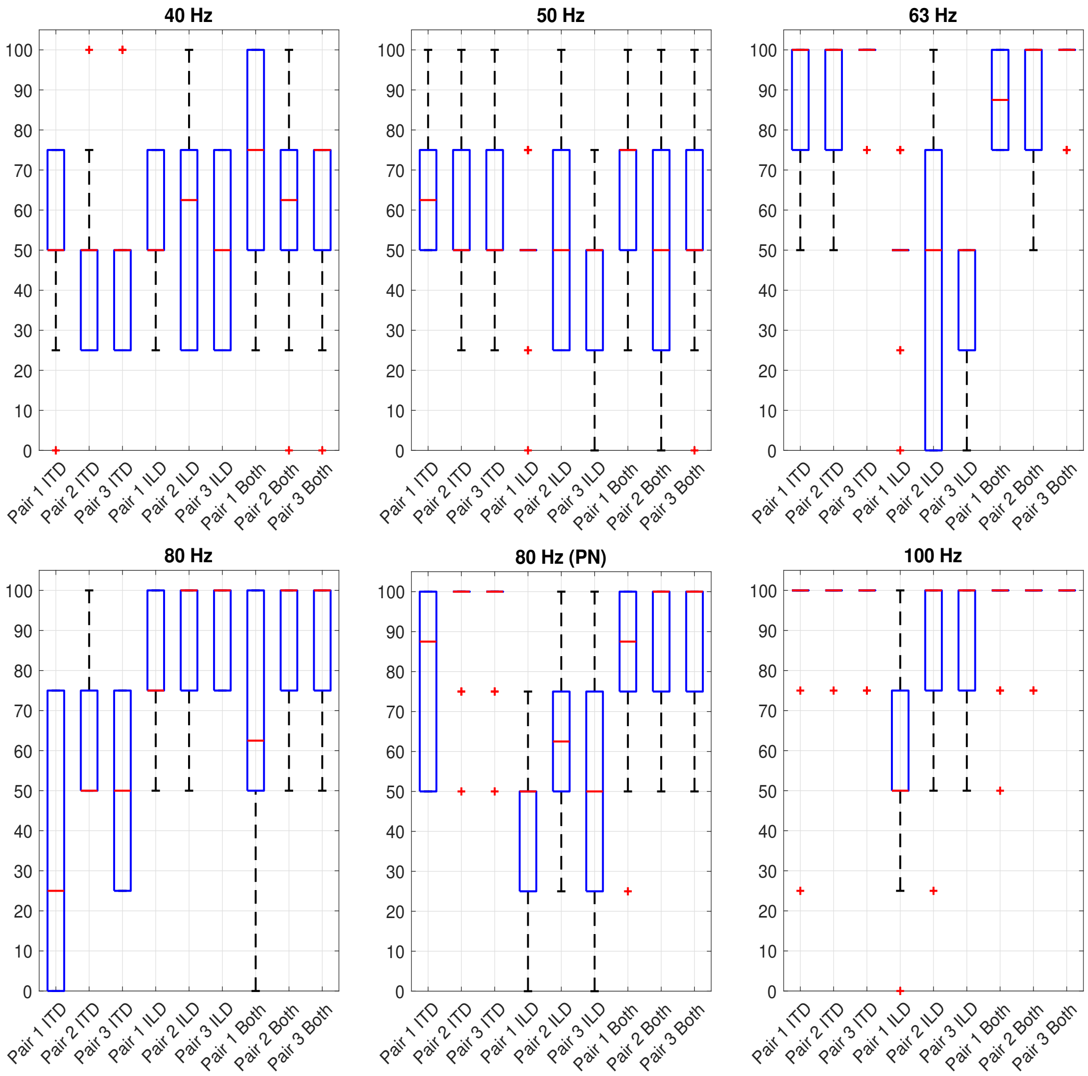

Figure 11: Boxplots of the results of experiment 2 for the sine tone stimuli

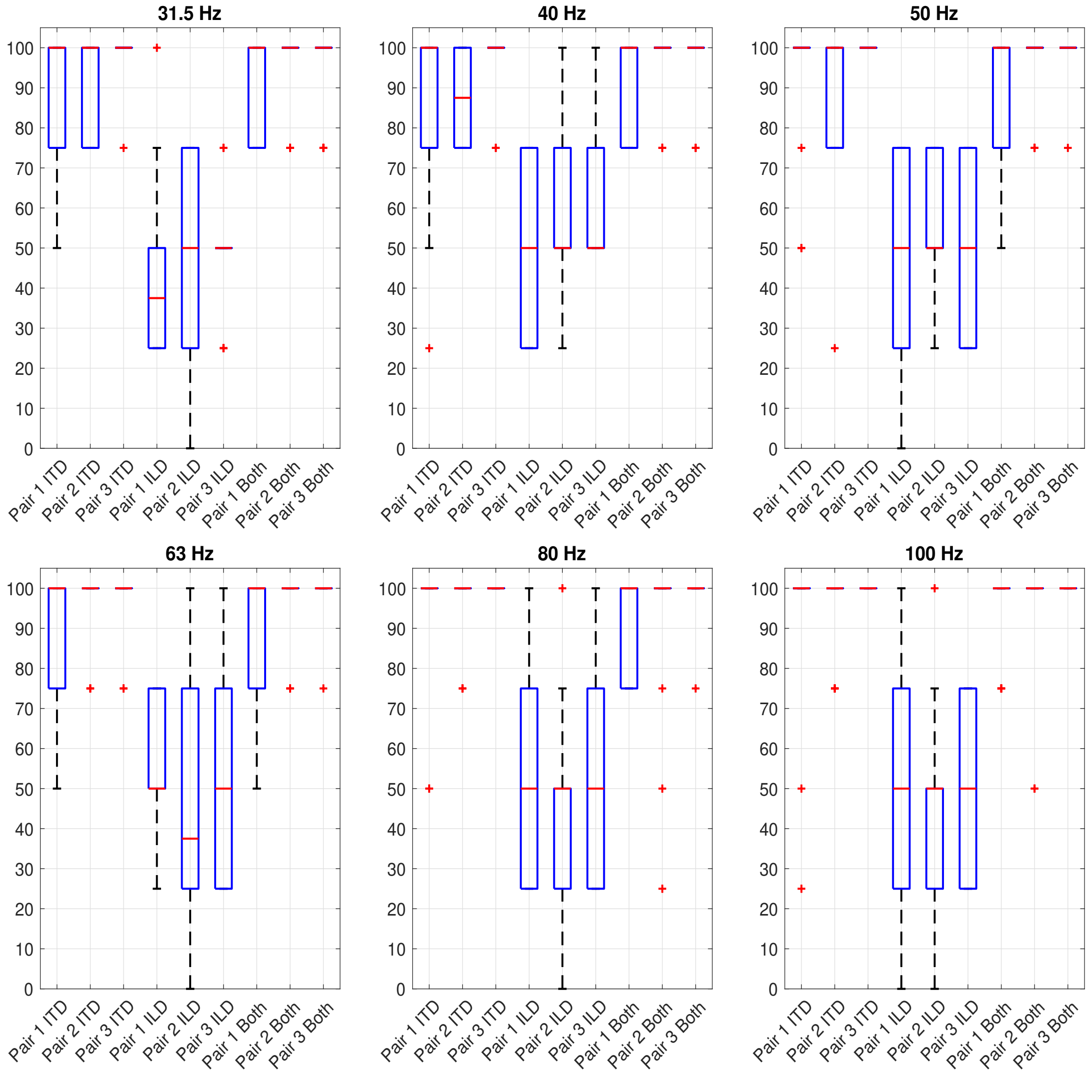

Figure 12: Boxplots of the results of experiment 2 for the pink noise stimuli

5 Discussion

In this chapter, the results of the first and second experiment presented in the previous chapter will be discussed. Furthermore, the MAA and MAITD will be calculated for the different stimuli type and frequencies.

5.1 First experiment

For the first psychoacoustic experiment, figure 7 and 8 show that, for all stimuli type and frequency bands, except the 80 Hz band of the sine, the percentage of correct responses were above the 50% threshold of chance, which indicates the point where subjects are guessing randomly. This suggests that participants were perceiving a change in the direction of the sound sources for almost all cases. Regarding the stimulus type, the pink noise had a better localisation performance than the pure sinusoids, as expected. The reason for this is that the waveform of noise fluctuates randomly as a function of time, and thus, provides transient information all throughout the presentation of the stimulus, while the pure tones provide only information relating to ongoing phase differences [19]. Moreover, sine tones are more susceptible to modal effects, as a narrow-band signal presents more frequencies so that the modal effect is not as large. The lowest median percentage recorded for the sine tones was of 20%, while the lowest median percentage recorded for the pink noise was of 60%. Overall, the sine tones presented more inter-subject variability than the pink noise.

For all frequency bands of the pink noise stimuli, a separation angle between 20° and 45° resulted in over 80% of correct responses. This could be a result of the large bandwidth (1 octave) of the band-pass filter as larger bandwidth sound sources have been shown to be less susceptible to errors in the judgement of localisation [5]. For the pure tone stimuli of frequencies under 63 Hz, the changes in separation angle had little effect on the percentages of the correct responses. For frequencies of 63 Hz and over, the lowest percentages were recorded for the separation angle of 10°, with the exception of 80 Hz. The pair with a separation angle of 10° away from the median plane, had an overall worse performance for both stimuli types than the pair with the reference located on the median plane. This result is in accordance to the findings of Mills (1958), in that the localisation abilities decrease when the reference azimuth moves away from the median plane.

The 80 Hz sine tone band displayed results that do not follow the expected trend. The ITD and ILD cues of this frequency, displayed in figure 9, show a possible corruption of the phase cues due to the presence of room modes. The presence of modal behaviour results in subwoofer 3 appearing to be the right-most sound source, while subwoofer 1 is the left-most source. The ILD cues of this frequency are the only ones to have a level difference of more than 3 dB, but interestingly, they match the position of the subwoofers. The modal behaviour of the room has been analysed by taking impulse responses of each subwoofer position using the G.R.A.S. 46AF 1/2 free field microphone. The frequency response of all the measurements is shown in figure 13. The room presents a pattern of nodes and nulls, where the

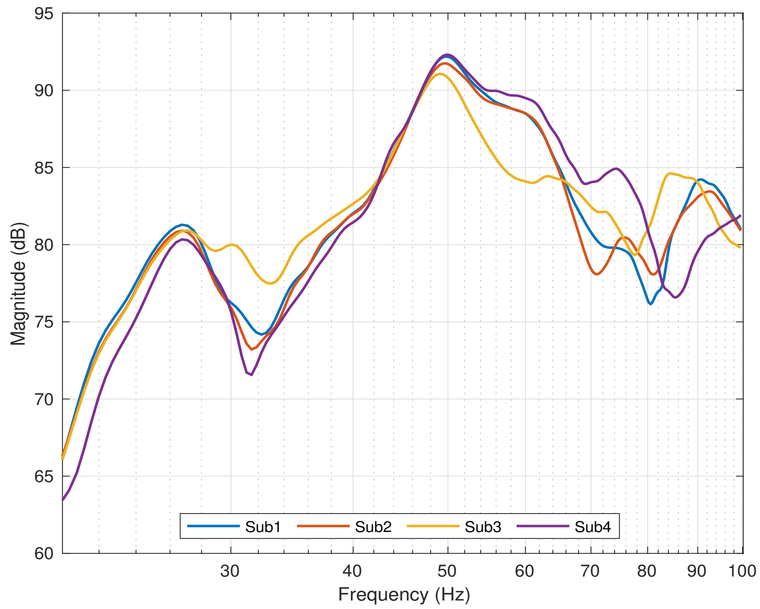

Figure 13: Frequency response of the subwoofers in the anechoic chamber showing modal behaviour

highest magnitude difference between the different subwoofer positions is of 10 dB. For most of the frequencies used in the experiment, the modal behaviour did not affect the results as the response differences between the subwoofer positions are not substantial. The most problematic frequencies are those of 31.5 and 80 Hz. Only the pink noise stimulus was used for the 31.5 Hz frequency, but as the computed ITD and ILD from 7 and 8 show, the noise stimulus is not significantly affected by modal behaviour. This could be due to the large bandwidth of the signal that does not excite only one frequency.

On the other hand, sinusoids are affected by room modes, and the different sound pressure gradients they create can negatively impact the binaural cue attributes. In the case of the 80 Hz frequency region, the magnitude difference between subwoofer 1 and 4 was of 5 dB. Furthermore, the phase cues of subwoofer 1 and 2 could have been affected as they are positioned in a null point. Thus the negative ITD values of this frequency range could be a result of the negative pressure gradient. The separation angles most affected by this were 10° and 20°, both being pairs containing subwoofer 1 and 2. A more in-depth analysis was performed in order to explain why four pairs resulted in correct percentages over 70. For this, the delta ITD and ILD of each pair was calculated, in which delta signifies the difference in time and level between the ITD and ILD of the two subwoofers in a pair. A positive delta ITD value or a negative delta ILD indicates a leftwards travelling auditory image for a leftwards moving stimulus, while a negative delta ITD or a positive delta ILD indicates a rightwards perceived auditory image for a leftwards moving stimulus. These values have been summarised in table 3 in order to observe the impact of the delta ITD and ILD on the correct percentages for the 80 Hz sine tone stimulus.

The data presented in the table clearly shows that the pairs with a positive ΔITD, thus a perceived source moving in the correct direction, have the highest correct percentages. Since the ΔILD positions the perceived source in the correct direction, a trade-off between the binaural cues used by auditory system appears to happen. As a result, when the negative ΔITD is smaller than 51μs, the ΔILD positively influences the outcome of the results. This is best observed for the separation angle of 45° (formed by subwoofer 1 and 4), where a level difference of 2.4 dB positions the source in the right direction, even though the ITD cues are erroneous. This trade-off between the ITD and ILD cues appears to explain why, even though this frequency band was affected by room modes, some of the directions of movement have still been perceived correctly.

Separation angle	10°*	10°	20°	25°	35°	45°
ΔITD	107 μs	-128 μs	-21 μs	-30 μs	77 μs	-51 μs
ΔILD	-0.12 dB	-1.28 dB	-1.4 dB	-1 dB	-1.15 dB	-2.4 dB
Correct Percentage	80%	20%	51%	71%	94%	78%

Table 3: Delta ITD and ILD values for each pair of the 80 Hz sine tone and their correct percentages

5.1.1 Determination of MAA

In order to determine the MAA, the results of all participants have been averaged, and categorised on stimulus type and frequency. These results can be seen in figure 14 and 16. Furthermore, the correct percentages for each frequency of stimulus type have been plotted as a function of the ΔITD in figure 15 and 17. This has been done in order to investigate the minimum time difference between two sources needed in order to produce 75% of correct responses. A ΔITD value that produces 75% of correct responses represents the MAITD. The ΔILD has not been plotted, as the ITD is the primary cue that enables localisation in the very low-frequency spectrum, when room modes do not affect the results. In all of these figures, the solid black line represents the 50% threshold of chance, and the dashed line represents the MAA threshold of 75% of correct responses. The red squares represent the percentages of correct responses for a specific angle or ITD value, and the thick black line which represents the psychometric function is an interpolation of these values.

The case of the pure tone stimulus will be discussed first. For the frequencies of 40 and 50 Hz, the MAA and MAITD can not be determined as the correct responses do not go above the 75% threshold. The results as a function of ΔITD show that the maximum time difference for 40 Hz is 105 μs and 157 μs for 50 Hz, which is equivalent to a delta interaural phase difference (ΔIPD) of 1.5° and 2.8° respectively. For this two frequencies, a separation angle greater than 45° is needed in order to determine the MAA.

For the 63.5 Hz frequency, the MAA lies between the pairs with the same opening angle. Figure 15 clearly shows that the pair located away from the median plane creates a lower ΔITD, hence why localisation performance is lower. The MAITD for 63 Hz is located at a ΔITD of 158μs, equivalent to a ΔIPD of 3.6°.

The results of the 80 Hz frequency band were corrupted by the presence of room modes and a reliable MAA cannot be determined from these results. However it is important to note that the separation angle of 10°, away from the median plane, presented a correctly moving source with a ΔITD of 107μs (ΔIPD = 3°) and a ΔILD below 0.2 dB, which means that the level differences did not influence the outcome of the results. This value of ΔITD recorded 80% of correct responses so it could be hypothesised that the MAITD for 80 Hz lies somewhere around 107μs.

The frequency of 100 Hz has gathered correct percentages that are over 80%. The smallest value of ΔITD is of 187 μs (ΔIPD = 6.7°), and results in a correct percentage of 92. Thus, in order to determine the MAA for a pure tone of 100 Hz, a separation angle smaller than 10° is needed.

While the MAITD could only be determined for the frequency of 63.5 Hz, a hypothesis can be constructed by evaluating the values of ΔIPD for all frequency bands. Values of ΔIPD below 3° have resulted only in correct percentages below 70, while values above 3.6° have passed the threshold of MAA. Thus, it could be argued that the MAITD of low-frequency pure tones lies at a minimum ΔIPD of 3.6°. The equivalent ΔITD values for such a phase difference have been summed up in table 4.

Frequency	40 Hz	50 Hz	63.5 Hz	80 Hz	100 Hz
MAITD	250μs	200μs	158μs	125μs	100μs

Table 4: Hypothesised MAITD values for the sine tone stimuli

For the pink noise stimuli, the MAITD can be estimated for the frequencies of 31.5, 40 and 50 Hz. Similar to the case of the 63.5 Hz pure tone, the MAA for these frequencies has its threshold between pairs of the same separation angle of 10°. The MAITD of 31.5 Hz is 60μs; for 40 Hz it is 77μs; and for 50 Hz it is 60μs.

For the frequencies above 50 Hz, a ΔITD between 50 and 60μs result in a correct percentage of 80 and over. Thus for determining the MAITD for these frequencies, a smaller separation angle than 10° between the subwoofers is needed.

The pink noise stimuli does not display the same relationship of MAITD with frequency as does the sine tone stimuli. In the case of the pure tones, the MAITD seems to decrease with frequency, while for the pink noise, it stays constant. This phenomena cue be a result of the nature of the pink noise stimulus. A lower ITD might be sufficient because pink noise changes its intensity and spectral structure as a function of time, and thus continually provides cues that are not subject to phase ambiguities. On the other hand, the sinusoid can only provide cycle by cycle, axis-crossing-related temporal information, hence why larger ITD values might be needed. Moreover, sounds with larger bandwidths provide less ambiguous ITD cues [19] because phase ambiguities can be resolved by comparisons across frequency. This could account for the lower values of MAITD as the pink noise stimulus has a bandwidth of one octave.

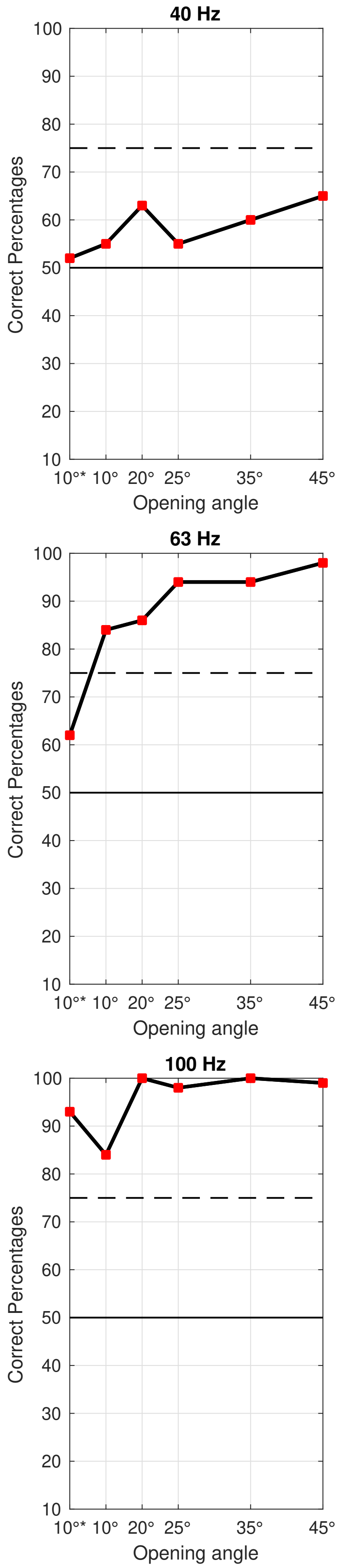

Figure 14: Average of the results of experiment 1 for the sine tone stimuli

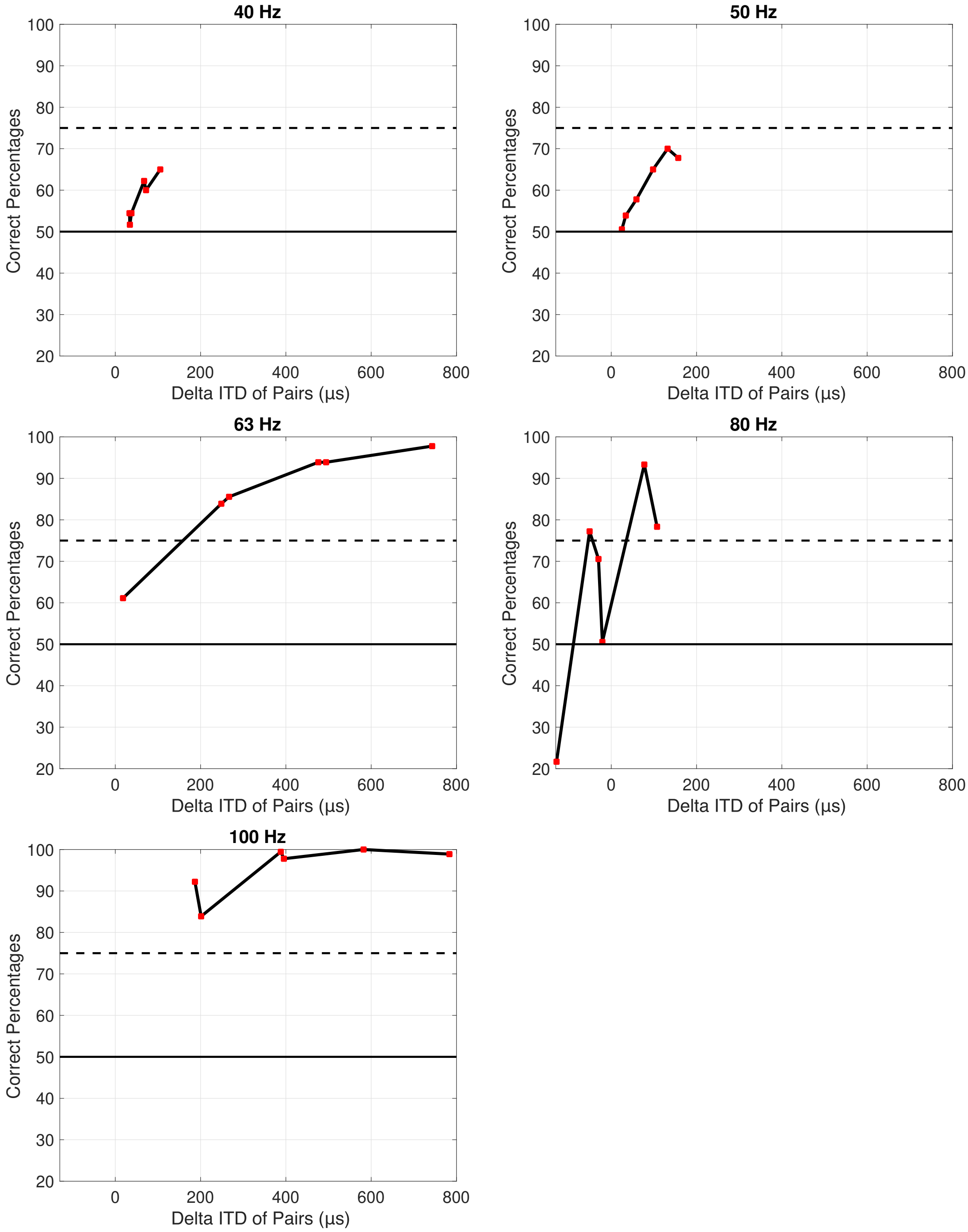

Figure 15: Average of the results of experiment 1 for the sine tone stimuli plotted as a function of ΔITD

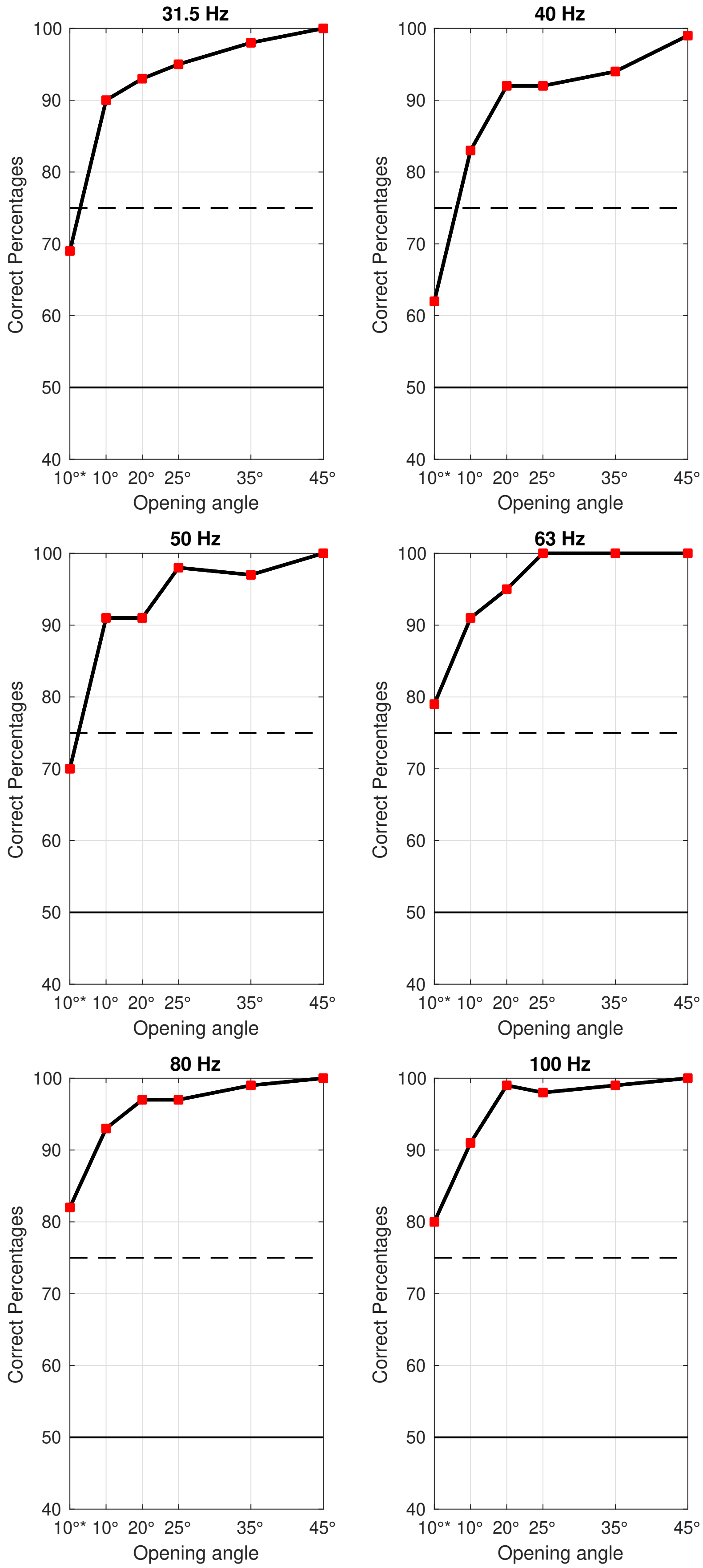

Figure 16: Average of the results of experiment 1 for the pink noise stimuli

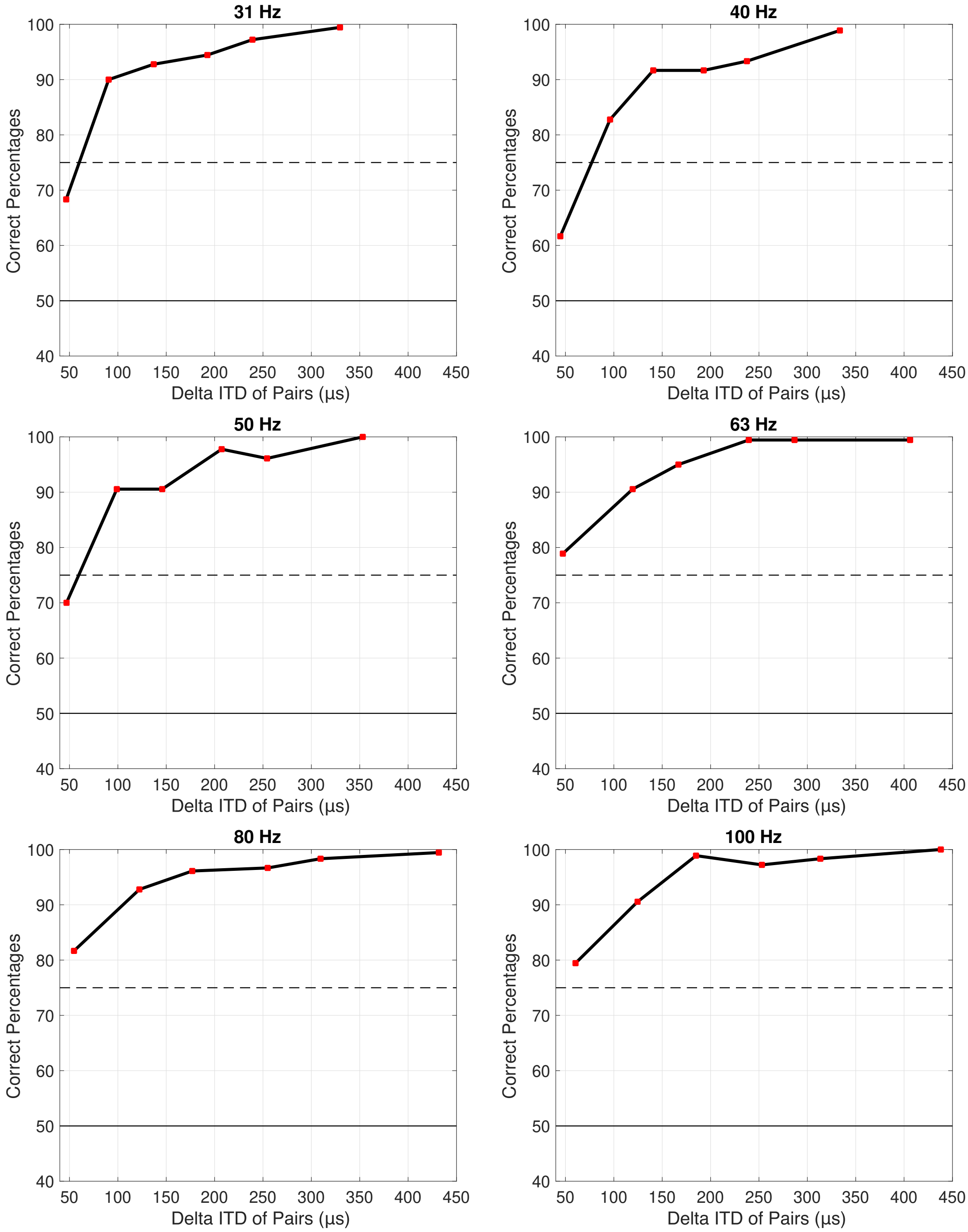

Figure 17: Average of the results of experiment 1 for the pink noise stimuli plotted as a function of ΔITD

5.2 Second experiment

The boxplot results of the lateralisation task can be seen in figure 11 and 12. An average of the results of all participants, categorised on stimulus type and frequency is presented in figure 18 and 19. Moreover, similar to the presentation of the results of experiment 1, the correct percentages for each frequency of stimulus type have been plotted as a function of the ΔITD in figure 20 and 21. In all the figures, the solid line represents the 50% threshold of chance and the dashed line represents the 75% threshold of MAA. The angle of separation increases with pair number, so that pair 1 has the smallest degree of separation while pair 3 has the highest.

Overall the results of the lateralisation experiment confirm that the ITD is the primary cue used by the auditory system to determine the location of a low-frequency sound source. This result is in accordance to the finding of Mohamed & Cabrera (2008), who performed a similar lateralisation task which consisted of only the ITD condition and a method of absolute localisation. Furthermore, the results seem to be in close agreement to the results of the first experiment, in that the ITD cues for a given stimulus and frequency resulted in a similar percentage of correct responses.

For the pink noise stimuli, as the ΔILD of all pairs is not larger than 0.5 dB, the number of correct responses for the ILD condition did not go above 65%. In the case of the sine tones, a variation of up to 2.3 dB has been observed in the 80 Hz condition, and this has lead to ΔILD to be the main cue used in this case. As it has been noted previously, this is the condition where the phase cues have been corrupted by room modes, and it can be clearly seen how it affects pair 1 and pair 3 of the ITD condition. Figure 20 demonstrates how a negative ΔITD, which indicates an auditory image perceived in the wrong direction, leads to results under the threshold of chance, while a positive ΔITD leads to an increase in the number of correct responses. The 80 Hz (PN) condition, where the uncorrupted ITD and ILD cues of the pink noise stimulus have been applied to the 80 Hz sinusoid, proves that ITD is the primary cue used, and shows a performance of over 80% of correct responses. In the case of the 100 Hz sinusoid, the ILD condition shows an increase in the number of correct responses, and it could be hypothesised that ILD becomes a secondary cue as the frequency increases. Pair 2 and pair 3 of this condition perform above the threshold of MAA as ΔILD varies between 1.5 and 1.9 dB, while for pair 1 it stays below 0.5 dB, and thus, below the threshold of MAA.

Similar to the first experiment, the results of the pure tones under the frequency of 63.5 Hz do not go above the threshold of MAA. For this frequency range, ΔITD does not exceed 200 μs, while the frequencies above this range exceed the threshold of MAA with ΔITD above 200 μs. Thus, it could be argued that in order to reach the threshold of MAA for pure sinusoidal sounds, a ΔITD of at least 200 μs is needed.

As the binaural cues used in the test were extracted from the position of subwoofer 1, 2 and 4, except the 80 Hz condition of the sine tone, the number of correct percentages of the ITD condition for the pink noise stimuli all exceed the threshold of MAA. This is in accordance to the results of the first experiment, where the pairs formed by these subwoofers recorded a number of correct percentages over 80%.

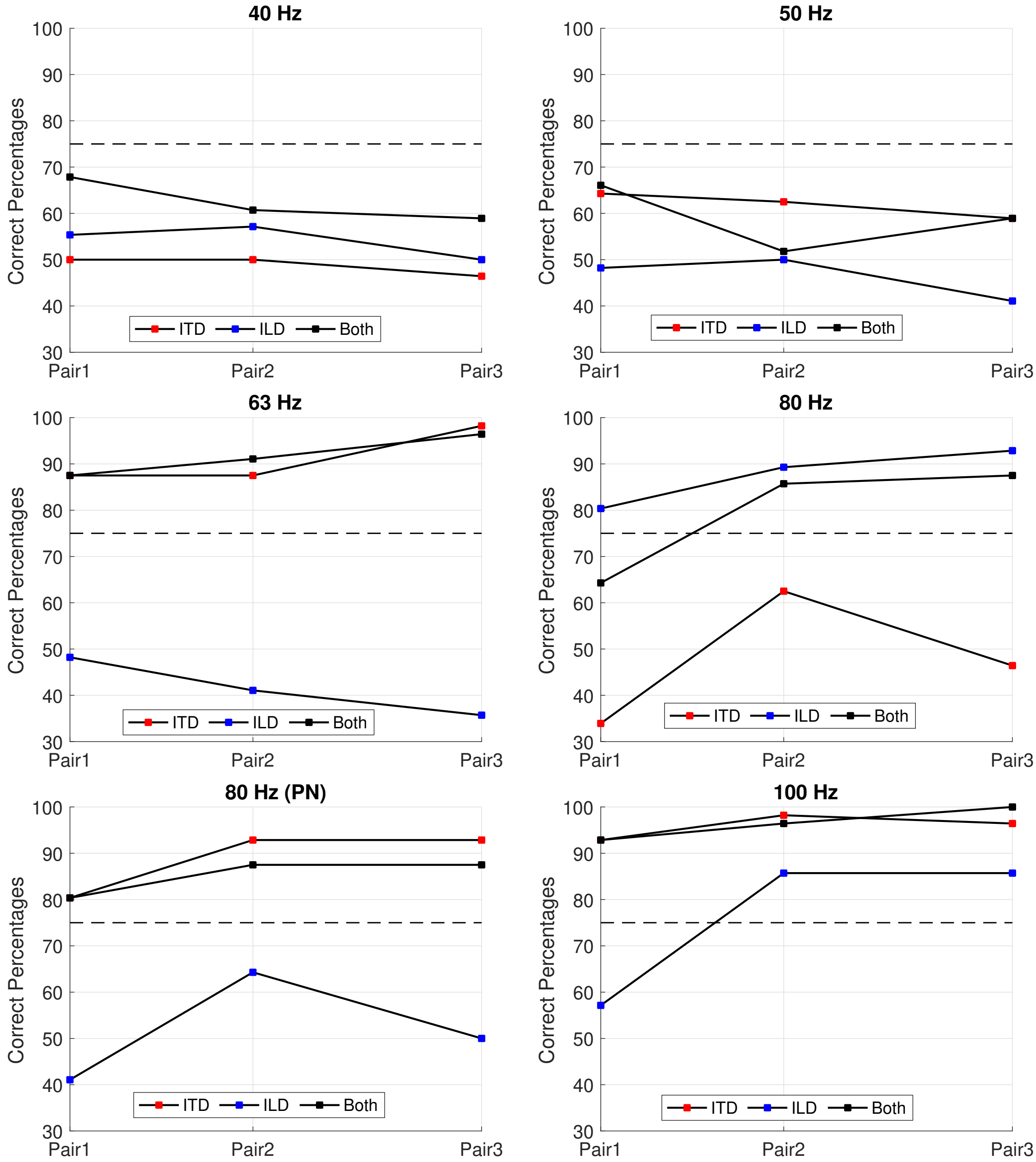

Figure 18: Average of the results of experiment 2 for the sine tone stimuli for ITD, ILD and Both condition

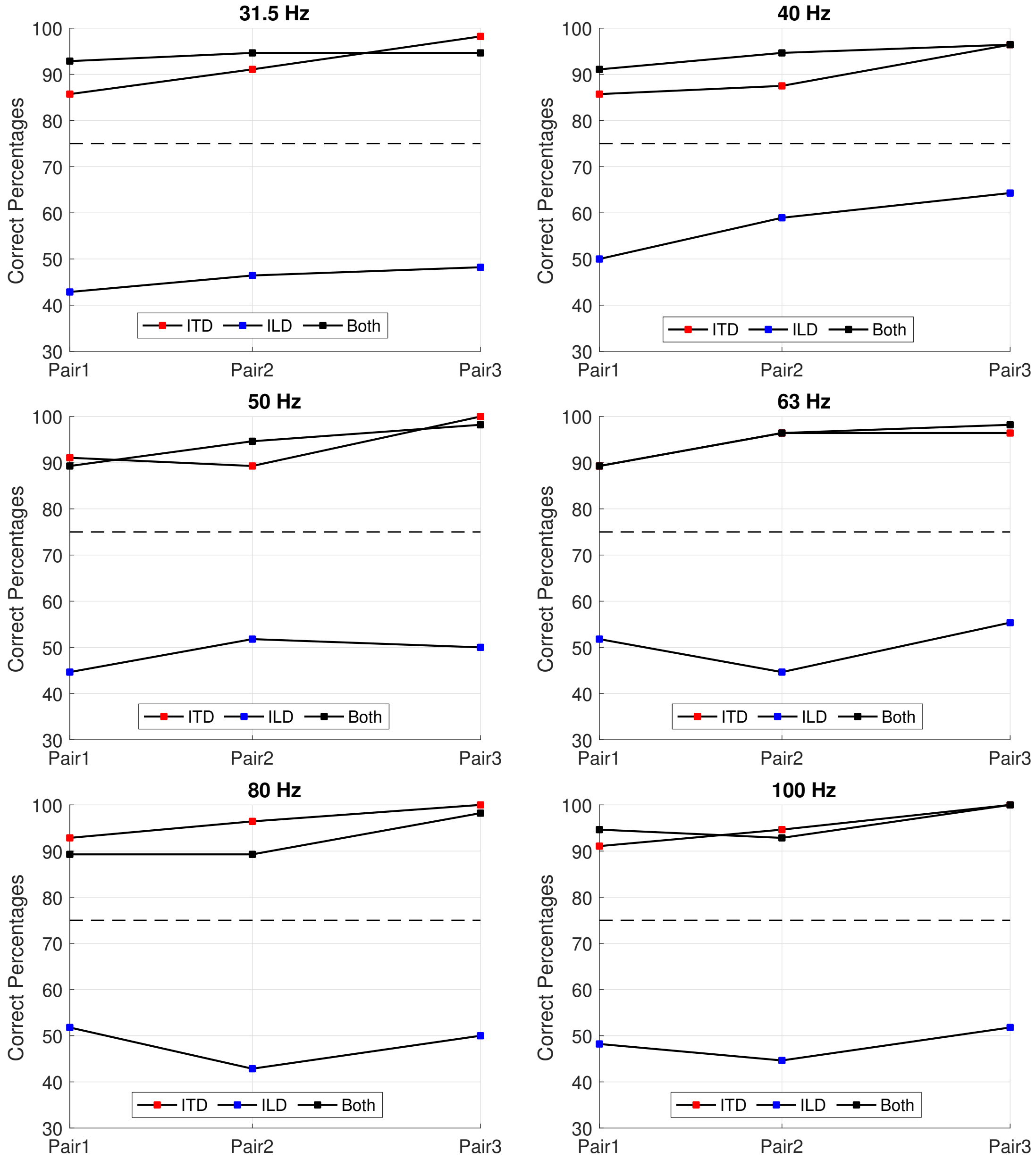

Figure 19: Average of the results of experiment 2 for the pink noise stimuli for ITD, ILD and Both condition

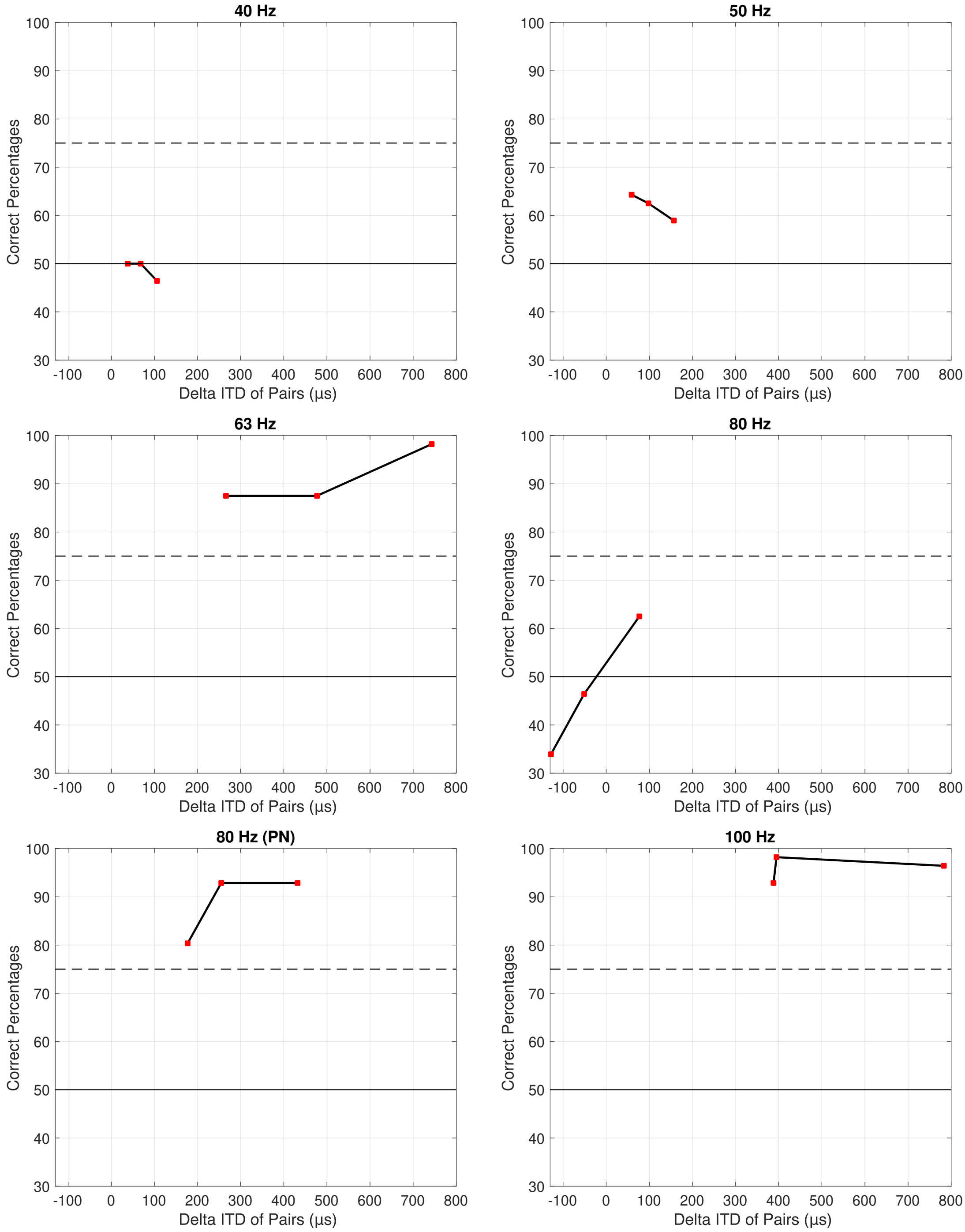

Figure 20: Average of the results of experiment 2 for ITD condition plotted as a function of ΔITD for the sine tone stimuli

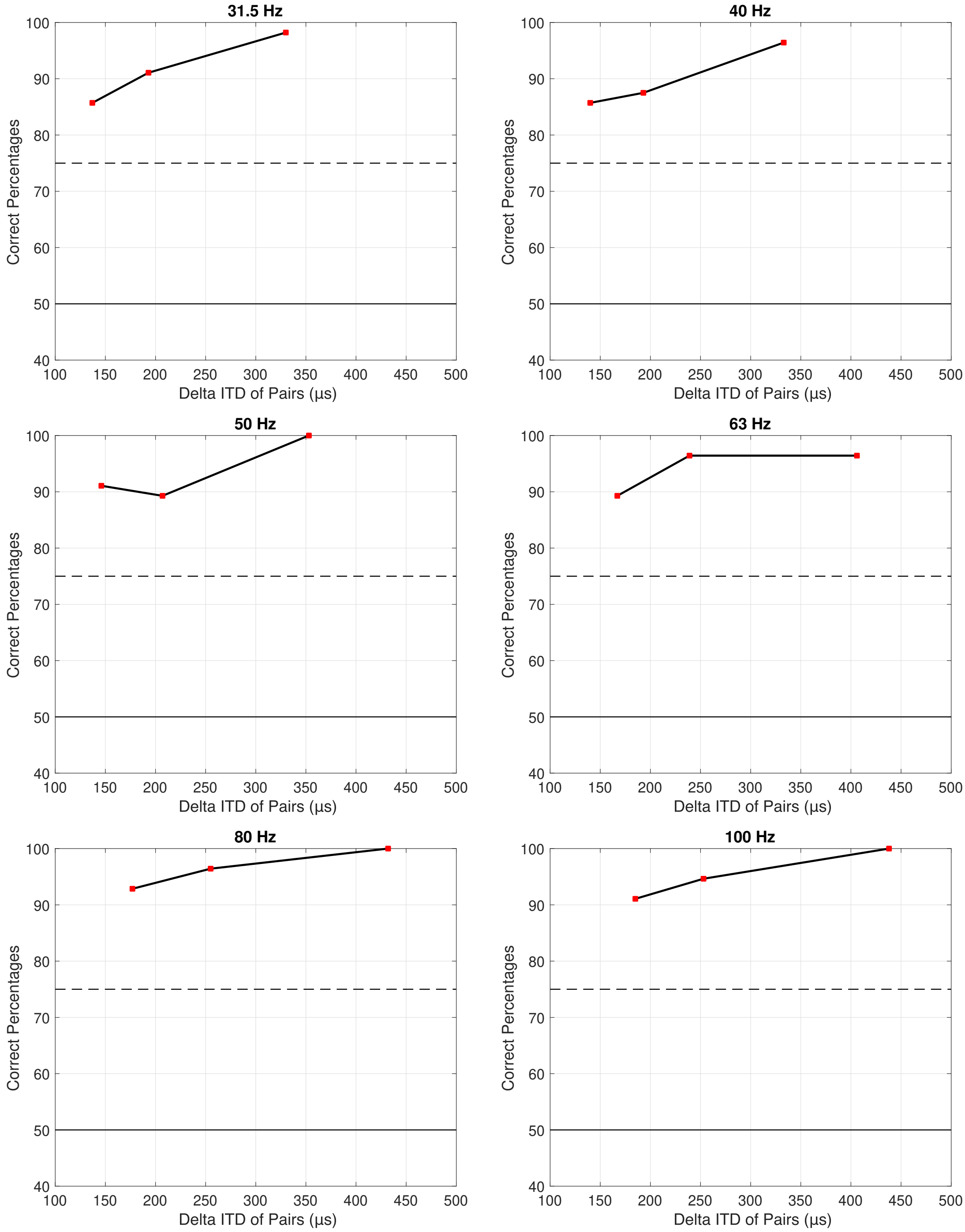

Figure 21: Average of the results of experiment 2 for ITD condition plotted as a function of ΔITD for the pink noise stimuli

6 Conclusion and Future Work

The aim of this thesis was to determine the extent of the human ability to perceive changes in the direction of low-frequency sound sources, within the range of 31.5 and 100 Hz, in a large, highly damped environment. Furthermore, the cues used by the auditory system that enable this localisation ability have also been studied. These objectives have been reached through two experiments.

The first psychoacoustic experiment that utilised a relative localisation method has proved that humans have the ability to localise sound sources in the very low-frequency spectrum with a high degree of accuracy when the listening environment is approximately anechoic. This experiment disproves the erroneous belief that low-frequency sound sources do not contain any directional information, which has been supported by previous studies conducted in small, reverberant spaces. Furthermore, the studies that support the localisability of the low-frequency spectrum, have only been limited to a capability of left-right discrimination. Dealing with a matter that has not been approached in any literature so far, the current study has set out to determine the MAA and MAITD for frequencies below 100 Hz. The MAITD represents the minimum ITD difference (ΔITD) between two subwoofer positions that results in 75% of correct responses. For this purpose, five separation angles of 10°, 20°, 25°, 35° and 45° have been studied.

Results show that the correct responses for the localisation of pure tones below 63.5 Hz are above chance level, but below the threshold of MAA, when the degree of separation is between 10° and 45°. The maximum ΔITD of this frequency range was of 157μs. Therefore, a determination of MAA and MAITD for pure tones of frequencies under 63.5 Hz would require a greater separation angle than 45° in order to create larger ITD values. For pure tones with a frequency above 63.5 Hz, the correct responses were much higher. This is due to the fact that much larger ITD values could be used as localisation cues. The MAITD for 63.5 Hz had a ΔITD of 158 μs and lies at a MAA of 10°. The case of the 80 Hz pure sinusoid shows that room modes can have a negative impact on the localisation ability, as the phase cues of four pairs of subwoofers have been corrupted. For this frequency, some of the directions of movement have still been perceived correctly as there appeared to be a trade-off between erroneous ITD cues and correct ILD cues. This case shows how room modes can hinder localisation, which has probably lead to the belief that low frequencies are not localisable. The MAA could not be determined for the 100 Hz sinusoid as all correct percentages were above 90, and the minimum value of ΔITD was of 187μs. For this frequency, a separation angle smaller than 10° would be required in order to calculate the MAA. By evaluating the values of ΔIPD for all frequency bands, it appears that a ΔIPD below 3° have resulted in correct percentages below 70, while values above 3.6° have passed the threshold of MAA. Thus, it was hypothesised that the MAITD of low-frequency pure tones lies at a minimum ΔIPD of 3.6°.

In the case of the pink noise stimuli, the localisation performance was of a better degree, as expected. This is due to the nature of the waveform of noise that provides more transient information all throughout its presentation, and because of the large bandwidth (1 octave) of the stimulus. The MAITD of the frequency bands of 31.5, 40 and 50Hz are 60μs, 77μs and 60μs respectively. For frequencies above this range, the MAITD could not be determined as the correct percentages were over 80. Thus, similar to the case of the pure 100 Hz sinusoid, the MAA of pink noise stimuli of frequencies above 50 Hz is smaller than 10°.

Overall, the results of the first experiment show that humans can reliably detect a change of 10° in the direction of pure tones with a frequency of over 63.5 Hz and octave bands of pink noise with a centre frequency as low as 31.5 Hz. It could be argued that the detection of the pink noise stimulus movement was effortless as the localisation of all separation angles, beside the 10° away from the median plane, for each frequency recorded correct percentages over the threshold of MAA. For pure tones under 63.5 Hz, while the number of correct percentages was as high as 70%, a higher separation angle than 45° would be needed for reaching the MAA threshold. Nonetheless, these results demonstrate that the auditory system has a capacity of detecting changes in the direction of low-frequency sound sources that goes beyond what has been described in previous literature.

The second psychoacoustic experiment, a lateralisation task of the first experiment, has shown that ITD is the primary cue utilised by the auditory system to determine changes in the location of a low-frequency sound source. This result is in close agreement to that of Braasch et al. (2004) and Mohamed & Cabrera (2008), who support the theory that low-frequency localisation is based on the decomposition of ITD cues. Moreover, the results of this experiment are in close agreement to the results of the first experiment, in that the ITD values for a given stimulus and frequency resulted in a similar percentage of correct responses. This demonstrates the importance of the MAITD, and how manipulation of the ITD in the lateralisation test according to the MAITD can successfully displace an auditory event.

It is hoped that by proving that humans have the ability to localise low-frequency sources in environments close to free-field conditions, the importance of low-frequency reproduction will be recognised by the audio community. As it stands, live sound engineers and electro-acoustic composers and performers have an intuitive knowledge about this importance as low-frequency reproduction plays a crucial part in their acts [6]. While localisation of the low-frequency spectrum is impeached in rooms that possess a strong modal behaviour, the benefits of using multiple subwoofers cannot be denied when the modal behaviour of the room is controlled or in environments close to free-field conditions, like large outdoor spaces.

After all, the low-frequency spectrum is a powerful tool for composers to fully exploit and for the listener to fully cherish.

6.1 Future Work

This thesis is one more step taken towards our understanding of the human capacity to localise low-frequency sound sources, which as it stands, is quite limited. This knowledge can be extended in a multitude of ways. For instance, higher separation angles between the subwoofers for frequencies under 63 Hz could lead to the determination of the MAA. Conversely, a separation angle smaller than 10° could help determine the MAA for frequencies over 63 Hz. The impact of sound source bandwidth could be more in-depth analysed by choosing noise sources with varying bandwidths.

While this work only dealt with free-field conditions, an interesting aspect to study is the influence of room modes on the localisation ability. This phenomena could be unravelled by conducting studies in rooms of different acoustic behaviour.

References

[1] E. Benjamin, *An experimental verification of localization in two-channel stereo*, in Audio Engineering Society Convention 121, Audio Engineering Society, 2006.

[2] J. Blauert, *Spatial hearing: the psychophysics of human sound localization*, MIT press, 1997.

[3] J. Borenius, *Perceptibility of direction and time delay errors in subwoofer reproduction*, in Audio Engineering Society Convention 79, Audio Engineering Society, 1985.

[4] J. Braasch, W. L. Martens, and W. Woszczyk, *Modeling auditory localization of subwoofer signals in multi-channel loudspeaker arrays*, in Audio Engineering Society Convention 117, Audio Engineering Society, 2004.

[5] R. A. Butler, *The bandwidth effect on monaural and binaural localization*, Hearing research, 21 (1986), pp. 67–73.

[6] R. T. Dean, *Low frequency spatialization in electroacoustic music and performance: Composition meets perception.*, Acoustics Australia, 42 (2014).

[7] L. C. Freeman, K. C. Wood, and J. K. Bizley, *Multisensory stimuli improve relative localisation judgments compared to unisensory auditory or visual stimuli*, The Journal of the Acoustical Society of America, 143 (2018), pp. EL516–EL522.

[8] D. Griesinger, *Spaciousness and localization in listening rooms and their effects on the recording technique*, Journal of the Audio Engineering Society, 34 (1986), pp. 255–268.

[9] M. J. Hautus, N. A. Macmillan, and C. D. Creelman, *Detection theory: A user's guide*, Routledge, 2021.

[10] G. B. Henning, *Detectability of interaural delay in high-frequency complex waveforms*, The Journal of the Acoustical Society of America, 55 (1974), pp. 84–90.

[11] A. J. Hill, S. P. Lewis, and M. O. Hawksford, *Towards a generalized theory of low-frequency sound source localization*, (2012).

[12] L. A. Jeffress, *A place theory of sound localization.*, Journal of comparative and physiological psychology, 41 (1948), p. 35.

[13] A. Kelloniemi, J. Ahonen, O. Paajanen, and V. Pulkki, *Detection of subwoofer depending on crossover frequency and spatial angle between subwoofer and main speaker*, in 118th Convention of the AES, paper, vol. 6431, Citeseer, 2005.

[14] H. Kietz and H. Zangemeister, *Special audiometry of the middle ear*, Zeitschrift fur Laryngologie, Rhinologie, Otologie und ihre Grenzgebiete, 32 (1953), pp. 58–61.

[15] W. L. Martens, *Subjective evaluation of auditory spatial imagery associated with decorrelated subwoofer signals*, Georgia Institute of Technology, 2002.

[16] R. Meddis, E. A. Lopez-Poveda, R. R. Fay, and A. N. Popper, *Computational models of the auditory system*, Springer, 2010.

[17] A. W. Mills, *On the minimum audible angle*, The Journal of the Acoustical Society of America, 30 (1958), pp. 237–246.

[18] M. Mohamed and D. Cabrera, *Human sensitivity to interaural phase difference for very low frequency sound*, in Acoustics 2008-Australian Acoustical Society Conference, 2008, p. 47.

[19] B. C. Moore, *An introduction to the psychology of hearing*, Brill, 2012.

[20] R. D. Patterson, M. Unoki, and T. Irino, *Extending the domain of center frequencies for the compressive gammachirp auditory filter*, The Journal of the Acoustical Society of America, 114 (2003), pp. 1529–1542.

[21] V. Pulkki and M. Karjalainen, *Communication acoustics: an introduction to speech, audio and psychoacoustics*, John Wiley & Sons, 2015.

[22] J. Rämö, S. Bergen, J. Parker, V.-M. Yli-Kätkä, and V. Pulkki, *Detection of two subwoofers: Effect of broad-band-channel level and crossover frequency*, in Audio Engineering Society Convention 132, Audio Engineering Society, 2012.

[23] J. W. Strutt, *On our perception of sound direction*, Philosophical Magazine, 13 (1907), pp. 214–32.

[24] A. Subkey, D. Cabrera, and S. Ferguson, *Localization and image size effects for low frequency sound*, in 118th Convention of the AES, paper, vol. 6325, 2005.

[25] G. Theile and C. Kügler, *Loudspeaker reproduction: Study on the subwoofer concept*, in Audio Engineering Society Convention 92, Audio Engineering Society, 1992.

[26] E. UNI, *3745: 2012*, Acoustics. Determination of Sound Power Levels of Noise Sources using Sound Pressure Precision Methods for Anechoic and Hemi-Anechoic Rooms. Available online: https://www. iso. org/standard/45362. html (accessed on 1 June 2019).

[27] T. Welti and A. Devantier, *Low-frequency optimization using multiple subwoofers*, Journal of the Audio Engineering Society, 54 (2006), pp. 347–364.

[28] T. S. Welti, *Subjective comparison of single channel versus two channel subwoofer reproduction*, in Audio Engineering Society Convention 117, Audio Engineering Society, 2004.

[29] K. C. Wood and J. K. Bizley, *Relative sound localisation abilities in human listeners*, The Journal of the Acoustical Society of America, 138 (2015), pp. 674–686.

[30] N. Zacharov, S. Bech, and D. J. Meares, *The use of subwoofers in the context of surround sound program reproduction*, in Audio Engineering Society Convention 102, Audio Engineering Society, 1997.