Fast Measurement of Hearing Sensitivity

Jussi Eskelinen

School of Electrical Engineering

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

Espoo 26.8.2021

Supervisor

Prof. Vesa Välimäki

Advisors

D.Sc. (Tech) Jussi Rämö

D.Sc. (Tech) Jyri Pakarinen

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 Jussi Eskelinen	
Title Fast Measurement of Hearing Sensitivity	
Degree programme Electronics and electrical engineering	
Major Acoustics and Audio Technology	**Code of major** ELEC3030
Supervisor Prof. Vesa Välimäki	
Advisors D.Sc. (Tech) Jussi Rämö, D.Sc. (Tech) Jyri Pakarinen	
Date 26.8.2021	**Number of pages** 52+8 — **Language** English

Abstract

Typically, headphones are calibrated according to the hearing characteristics of an average user. This creates a sub–optimal listening experience. The subjective sound quality of the headphones can be improved if user's individual hearing characteristics are taken into consideration. This is traditionally done by performing a pure–tone hearing test. This research attempted to find out if an equal–loudness test could have advantages over the standard pure–tone test. A standard equal–loudness test, based on test signals with 1/3 octave bandpass filtered pink noise, was adapted to a self–administered test format. Its performance in noise was compared to a naive implementation of a basic pure–tone audiometry test. It was discovered the mean of the difference between the equal–loudness test results was 2.1 dB and the mean of the difference between pure–tone audiometry test results was 4.5 dB. The standard deviation was 1.8 dB for the equal–loudness test and 4.0 dB for the pure–tone audiometry test. While these results are not conclusive due to the small sample size and differences in testing environments, they do suggest that the equal–loudness test could be more accurate than pure–tone audiometry in a noisy environment.

Keywords headphones, hearing test, hearing threshold, loudness

Aalto University
School of Electrical Engineering

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi
Diplomityön tiivistelmä

Tekijä Jussi Eskelinen

Työn nimi Kuulon herkkyyden nopea mittaus

Koulutusohjelma Tietokone-, kommunikaatio- ja informaatiotieteet

Pääaine Akustiikka ja audioteknologia **Pääaineen koodi** ELEC3030

Työn valvoja Prof. Vesa Välimäki

Työn ohjaajat TkT Jussi Rämö, TkT Jyri Pakarinen

Päivämäärä 26.8.2021 **Sivumäärä** 52+8 **Kieli** Englanti

Tiivistelmä

Tyypillisesti kuulokkeet kalibroidaan keskimääräisen ihmisen kuulon ominaisuuksien tai preferenssien mukaan. Täten kuulokkeiden tuottama kuuntelukokemus ei ole paras mahdollinen. Yksilön kokemaa kuulokkeiden äänenlaatua voidaan parantaa, jos yksilölliset vaihtelut kuulossa otetaan huomioon. Perinteisesti tähän on käytetty kuulokynnyksen määrittävää kuulotestiä, jossa käytetään sinisignaaleita. Tämän tutkimuksen tarkoitus oli selvittää, mitä hyötyjä voisi saada siitä, että tämä testi tehtäisiin mittaamalla henkilön kuulon äänekkyyskäyrästö. Perinteinen äänekkyystasotesti sovitettiin käyttäjän itsensä tehtäväksi testiksi. Testisignaaleina käytettiin 1/3-oktaavikaistanpäästösuodatettua vaaleanpunaista kohinaa. Tätä testiä verrattiin perinteiseen käyttäjän itsensä sinisignaaleilla tekemään kuulotestiin. Kummatkin testit tehtiin sekä melussa että hiljaisuudessa, ja näitä arvoja verrattiin toisiinsa. Havaittiin, että äänekkyystasotestin erot olivat keskimäärin 2.1 dB ja perinteisen kuulotestin 4.5 dB. Keskihajonta oli äänekkyystestille 1.8 dB ja perinteiselle kuulotestille 4.0 dB. Pienestä otannasta johtuen tulokset ovat vasta alustavia, mutta niiden valossa on mahdollista, että äänekkyystasotesti soveltuu paremmin testiksi meluisiin ympäristöihin kuin perinteiset kuulotestit.

Avainsanat kuulokkeet, kuulokäyrä, kuulotesti, äänekkyys

Preface

I would like to thank Professor Vesa Välimäki and my instructors D.Sc. Jussi Rämö and D.Sc. Jyri Pakarinen for their guidance. Special thanks to D.Sc. Petteri Hyvärinen for his help with laboratory equipment, to M.Sc. Lauri Vapalahti for collaborating with the project that led to this thesis, and to D.Sc. Juho Liski for being my de-facto third advisor. The biggest thank you goes to my parents, my father Professor Pekka Eskelinen in particular, for their support over the years.

Additionally, I would like to thank *Helsinki Central Library Oodi* for providing an air conditioned refuge during the July 2021 heatwave, my friends in Finland for unsolicited advice, tea and long walks, *Mr. Smirnoff* for his demanding stare, everyone on the IRC channel #kalatikka, and my friends in Germany for the good times.

Helsinki, 26.8.2021

Jussi Eskelinen

Contents

Symbols and abbreviations

Symbols

f	frequency
f_s	sampling frequency
T	period of a wave

Abbreviations

avg	average
ccm	cubic centimetres (cm^3)
dB	decibel
EQ	equalization
ERB	equivalent rectangular bandwidth
HL	hearing level
MAF	minimum audible field
med	median
sd	standard deviation
SPL	sound pressure level

1 Introduction

Typically, headphones are calibrated according to the hearing characteristics, or estimated preferences, of an average user. This creates a sub–optimal listening experience, especially for those with prominent hearing loss. It has been suggested [1][2][3] that subjective sound quality of the headphones can be improved if user's individual hearing characteristics are taken into consideration.

Traditionally, hearing parameters have been estimated by performing a hearing test. The most common hearing test is *pure–tone audiometry*. In this test, sinusoids of different standardized frequencies are played to the user, individually for each ear. The smallest sound pressure level at which the user can still hear a sound, the *absolute threshold of hearing*, is recorded at each frequency. This results in two hearing threshold curves, one for each ear, which are also known as *audiograms*. Traditionally, audiograms are used as a basis for assessing *hearing loss* and setting the parameters of a *hearing aid*[4].

Recently, a new use case for audiograms has emerged: *personalized sound* [5][6][7][8][9], also known as *personal sound*. It refers to technologies that use hearing data to modify the output of an audio device to match the user's hearing characteristics. These technologies can be used e.g. in mobile phones, laptops and other portable devices to improve the quality of music and speech or to compensate for hearing loss.

In the past, hearing tests were always performed by a qualified medical professional. However, recent research has shown no significant differences in test-retest reliability and accuracy between a hearing test performed by a medical professional and an automated hearing test performed by a test subject on their own, without professional aid[10]. In other words, professional skills and equipment are not needed in performing a hearing test. This has led to pure–tone audiometry being used in personalized sound applications to assess the hearing sensitivity of the user.

However, pure–tone audiometry has several drawbacks in this use case. Pure–tone audiometry is meant to be performed in a space that is as quiet as possible in order to obtain an accurate audiogram. Unfortunately, a mobile application user might take the test in an ordinary apartment, a noisy office, or even outdoors, where several noise sources can disturb the measurement and thus degrade measurement accuracy[11][12]. A pure–tone audiometry test can also take several minutes to perform, which might discourage users from taking one. Finally, it can be stressful to listen to the extremely quiet test signals used in pure–tone audiometry.

However, it is also possible to define hearing sensitivity by measuring the perceived *loudness* of a signal. This can be achieved by prompting the user to adjust the loudness of one test signal until the user perceives it to be as loud as a fixed-volume reference signal. This value is then recorded, and the test is repeated for a test signal with a different frequency[2]. This results in an *equal–loudness* contour[13]. This type of test avoids the aforementioned problems associated with pure-tone audiometry.

The aim of this thesis is to develop a fast equal–loudness hearing test to measure hearing sensitivity. We will compare the speed and accuracy of different traditional ways to measure hearing to a novel equal–loudness method, and develop a user interface and a testing method which enables a fast, accurate and comfortable hearing sensitivity measurement.

The remaining part of this thesis is divided into four chapters. Chapter 2 describes the basic functioning of hearing, the concepts of hearing threshold, equal–loudness and theory related to headphones. Chapter 3 describes the measurements that were performed for this thesis. Chapter 4 presents measurement results. Chapter 5 presents conclusions. Appendix A describes the Matlab implementation of the test program. Appendix B presents the flowchart of the equal–loudness test. Appendix C explains the box plot.

2 Theory and previous research

As mentioned in the introduction chapter, the topic of this thesis involves the concept of *personal sound*; technologies that modify the output of an audio device to improve the subjective audio quality, and which are used e.g. in mobile phones, laptops and other portable devices. Since a personal sound application could be used to compensate for hearing loss, as a hi–fi audio enhancement tool, or both, this research topic lies at the intersection of hearing aids, headphones, hearing tests and personal sound applications. Therefore, this chapter describes the relevant theory behind human hearing, hearing loss and audiometry, and takes a look at previous research regarding automated hearing tests, hearing tests for headphone calibration, and personalized sound applications.

2.1 Sound and hearing

One of the most basic question in acoustics is *"What is sound?"* As it turns out, there are two different ways to answer this question. Sound can mean both a perceived sensation, and vibration in a medium, e.g. the movement of air molecules[13]. Therefore, a distinction is made between these two phenomena. Vibrations in a medium are called *sound events*, and sound events experienced by a listener are called *auditory events*[13].

In order for a sound to be heard, its frequency needs to be within the *effective hearing area*, which is generally considered to be between 20 Hz and 20 kHz[14]. Lower and higher frequencies of very high intensity can also be heard — these are known as *infrasound* and *ultrasound*, respectively[13]. The sound pressure level also needs to be above the *absolute threshold of hearing*, also known as the *hearing threshold*, in order to be audible. The hearing threshold is shown in Figure 1. The hearing threshold curve has been obtained by averaging the hearing test results of several individuals, and therefore represents the hearing characteristics of an average listener. The topic of hearing tests will be covered later in chapter 2.5.1.

As can be seen from Figure 1, for most human individuals, hearing is at its most sensitive between 1–3 kHz, and that, in general, the lower the frequency, the higher the sound pressure level has to be for the sound to be audible. Figure 1 also shows that sound pressure levels above 130 dB are experienced as physically painful. This is known as the *pain threshold*[4]. The amplitude resolution of human hearing is approximately 1 dB, depending on the individual[14].

2.1.1 Sound pressure, loudness and loudness perception

The *sound pressure level* (SPL) of a sound event can be defined with the equation

$$L_p = 20 \log \frac{p}{p_0} \tag{1}$$

where p is the sound pressure and p_0 is the reference pressure level[14]. Usually, $p_0 = 20\,\mu\text{Pa}$. The unit of SPL is *decibels*, dB.

Knowing the sound pressure level of a sound event is not sufficient enough to describe how loud the auditory event is, because human perception of sound pressure level is dependent not only on the intensity of the signal, but also on the frequency, spectral tone, time, and background of the sound event[15]. Therefore, the term *loudness* has been chosen to describe how loud the auditory event is[14]. The unit of loudness is called a *sone*[14]. 1 sone is the equivalent of 40 phons[16]. Phon is a logarithmic unit of loudness, and 40 phons is the equivalent of a 1 kHz tone with a sound pressure level of 40 dB[14].

As stated earlier, the human ear is relatively insensitive to sounds of low frequency ($< 500\,\mathrm{Hz}$), and relatively more sensitive at around 1–3 kHz. The frequency-dependence of loudness perception of pure–tones can be visualized with equal–loudness contours, shown in Figure 1. The first equal–loudness contours were defined by Fletcher and Munson[17] in 1933, and they have been updated since then[18].

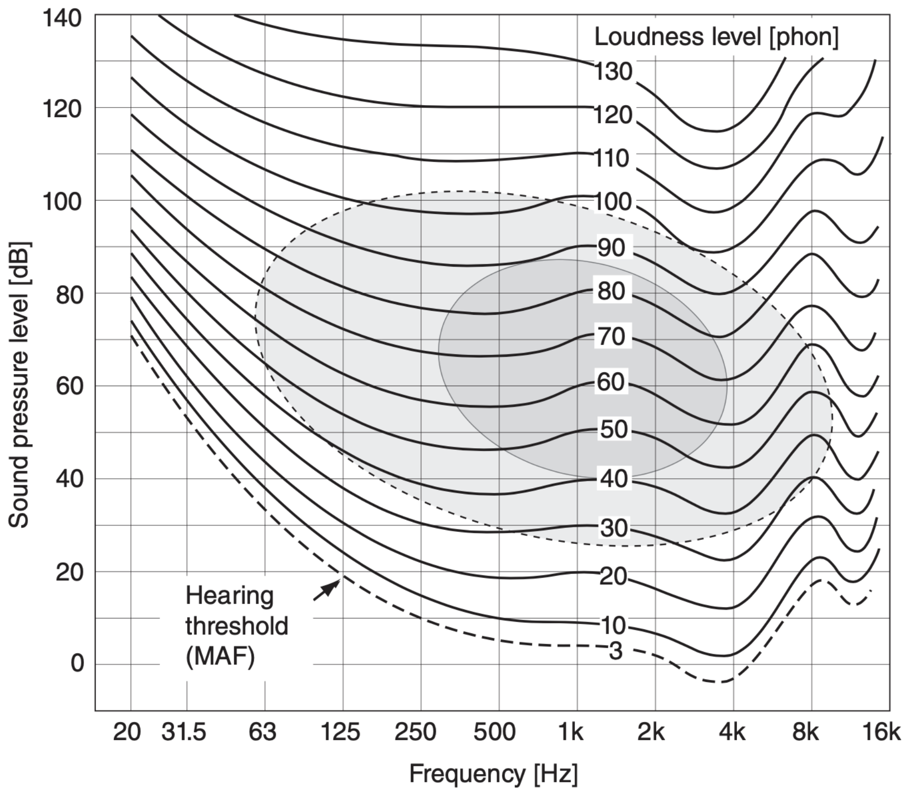

Figure 1: The threshold of hearing and the equal–loudness contours of hearing on the frequency – SPL plane. The lowest curve specifies the *minimum audible field* (MAF) or the hearing threshold. The uppermost curve specifies the approximate threshold of pain. The shaded light gray area shows the approximate range of frequencies and sound pressure levels for acoustic music. The shaded dark gray area shows the approximate range for speech. Adapted from [13].

2.2 Structure of the human ear

The structure of the human ear is shown in Figure 2. As can be seen from the figure, the ear can be divided into the *outer ear*, *middle ear* and *inner ear*. The outer ear consists of the *pinna*, the *concha*, and the *ear canal*.[13] These are mainly responsible for directing the sound waves towards the middle ear. The middle ear contains the *eardum*, which vibrates according to the pressure changes of air, and the *ossicles*, the three small bones (*malleus*, *incus* and *stapes*) which amplify the vibrations of the eardrum and transfer them into the inner ear. The middle ear also contains the *eustachian tube*, which joins the middle ear with the nasal cavity. This enables equalization of ear pressure between the middle ear and the throat.[14]

The inner ear contains the *vestibular system*, which is the organ responsible for the sense of balance, and the fuild-filled *cochlea*, which is connected to the ossicles via the *oval window*. The oval window transfers the vibrations of the eardrum into the cochlea, more precisely to the *scala vestibuli*.[13] The ossicles amplify the very small pressure of the sound waves 30 fold on the oval window[14]. Most of the amplification comes from the differences in the areas of the eardrum and the round window. The cochlea also has a second opening, the *round window*, which allows the fluid inside to move freely.[13]

The vibrations are transmitted via the oval window to the basilar membrane. On top of the basilar membrane sits the *organ of Corti*, a complex organ of gelatinous mass and hair cells. The hair cells are attached to nerve fibers. The vibration of the stapes against the oval window causes pressure variations in the cochlea, which

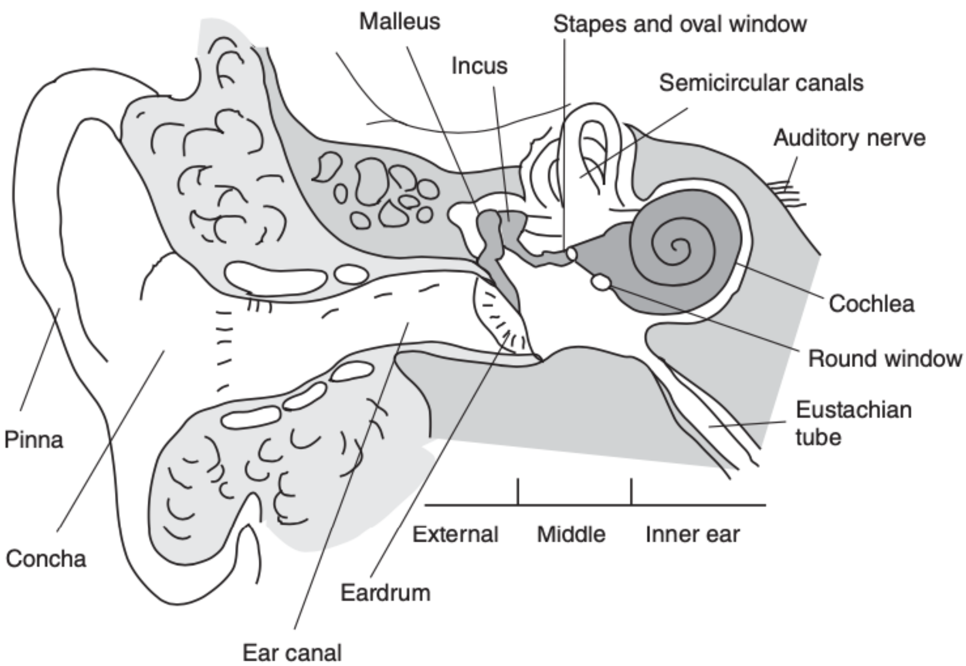

Figure 2: A schematic diagram of the human ear. Not to scale. Adapted from [13].

in turn excites mechanical vibrations in the basilar membrane. In turn, this makes the small hairs on the organ of Corti bend, which generates nerve impulses. The auditory nerve transmits these impulses to the auditory cortex of the brain, where they are then interpreted as auditory events.[13]

The cochlea is shaped like a spiral. Its cross section is shown in Figure 3. Inside the cochlea are the basilar membrane and the bony shelf, which divide the cochlea into two halves. Above the basilar membrane is the *scala vestibuli* and below it the *scala tympani*. A thin membrane known as the *Reissner's membrane* separates the cochlea into one more section called the *scala media*. The organ of Corti is located on top of the basilar membrane. The *tectorial membrane* is a gelatinous structure that sits right above the hair cells.[13]

The basilar membrane plays a part in frequency analysis. High frequencies cause vibrations in the region near the oval window, where the basilar membrane is narrow and stiff. Conversely, the basilar membrane reacts to low frequencies near its loose far end.[13] Next, we will discuss how the basilar membrane distinguishes between different frequencies and some of the phenomena associated with this auditory filtering.

2.3 Critical bands and masking

When two frequencies are close to each other on the basilar membrane, they are said to lie within the same *critical band*[14]. Each critical band is like a data collecting

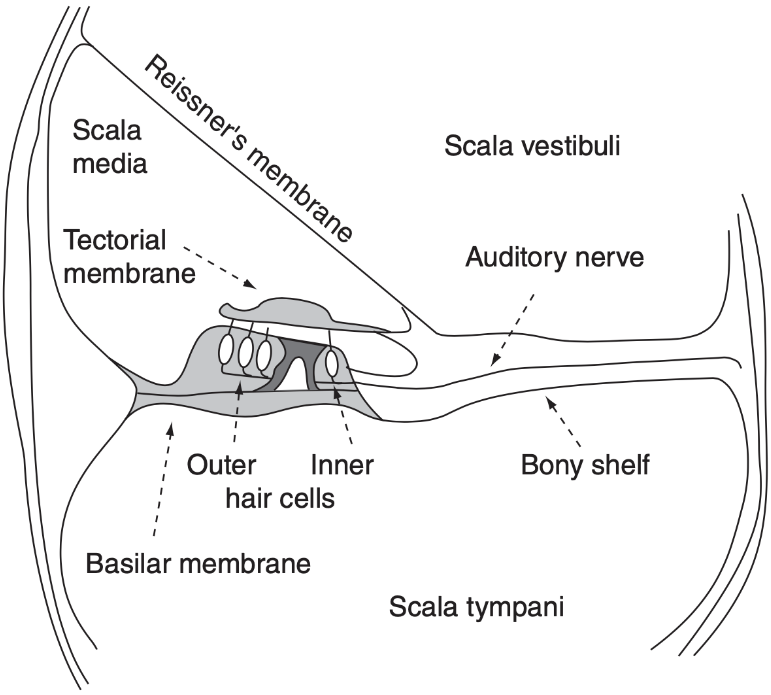

Figure 3: Cross–section of the cochlea. Adapted from [13].

unit. The critical bandwidth varies with centre frequency, so that the bandwidth is nearly constant in low frequencies and roughly proportional to frequency at high frequencies[14]. For most of the audible range, critical bandwidths are approximately 1/3 of an octave in width[14]. Since 1/3 octave filters correspond to the characteristics of human hearing, they are common in sound analysis and when generating signals for listening tests[13].

An octave means the doubling of frequency, in other words

$$f_2 = 2f_1 \tag{2}$$

where f_1 is the lower band edge and f_2 is the upper band edge[14]. 1/3 octave band filters are a more narrow, and defined with the equation

$$f_4 = 2^{\frac{1}{3}} f_3 \tag{3}$$

where f_3 is the lower band edge and f_4 is the upper band edge[19]. Examples of octave and 1/3 octave filters are shown in Figure 4.

Another way to model the critical bands in human hearing is to use the *equivalent rectangular bandwidth*, ERB, which gives an approximation of the critical bandwidths of human hearing, where the critical bands are approximated by using rectangular band-pass filters[20][13]. ERB bandwidths can be calculated with the function

$$\Delta f_{ERB} = 24.7 + 0.108 f_c \tag{4}$$

where f_c is the center frequency of the critical band[20].

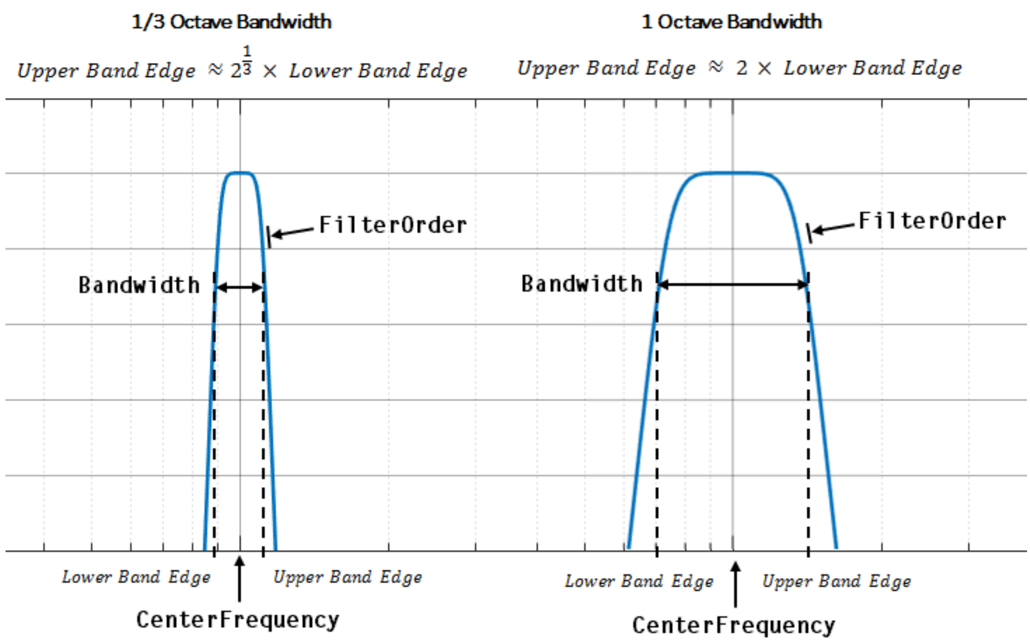

Figure 4: Octave filters and 1/3 octave filters. Adapted from [19].

The concept of critical bands is closely linked to the phenomenon of *masking*. Masking occurs when one sound raises the threshold for hearing another sound[14], either making it seemingly more quiet or preventing it from being heard altogether. In other words, masking leads to reduction of the perceived loudness of a sound[15]. This happens if the spectrum of both the masking sound and the masker sound reside within the same critical band[13]. Masking can occur with pure tones, complex sounds and both narrow and broad band noise[13]. Figure 5 shows how the hearing threshold curve changes in the presence of a narrow band noise signal at different sound pressure levels. Figure 6 shows how loud tones in quiet need to be in order to be as loud as they would be in masking noise with different sound pressure levels[15].

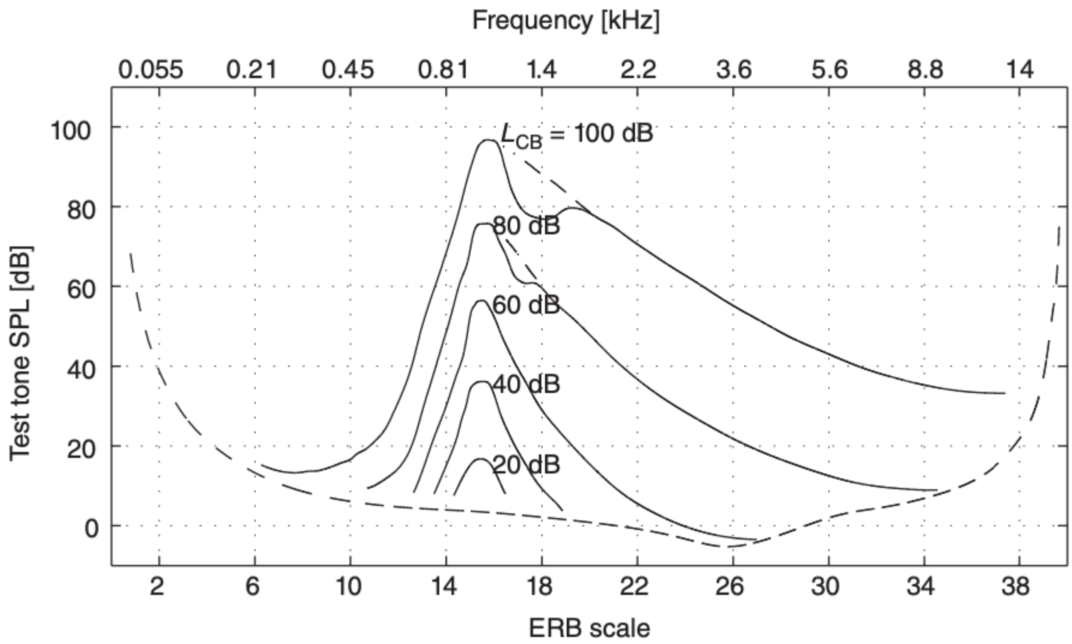

Figure 5: Masking threshold curves caused by a narrow band noise signal with a centre frequency of 1 kHz and different sound pressure levels $L_C B$. Adapted from [13].

Masking by noise is also used in pure–tone audiometry[4]. For certain patients with hearing loss, the sound pressure levels for one ear need to be so loud that they can hear the sound with the other ear[4]. A noise signal can be played for for the non–tested ear to prevent the test signal from being audible and causing a false positive result[4].

2.4 Hearing loss

Hearing loss means the partial or total inability to hear, either with one or both ears. There are numerous possible causes for hearing loss: it can be genetic, or related to birth complications, diseases, infections, ototoxic drugs (e.g. drugs for cancer treatment), head trauma or exposure to excessive noise or loud music. People are also

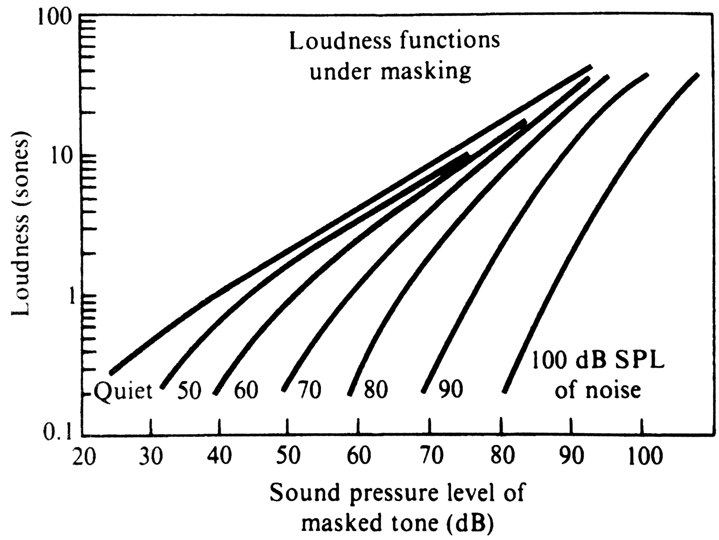

Figure 6: This figure shows how the presence of white noise reduces the apparent loudness of a 1000 Hz tone. These loudness functions were generated by asking listeners to adjust the loudness of a tone in quiet until it was as loud as the tone in noise. Adapted from [14].

living longer than ever before, and therefore ageing–related hearing loss is becoming increasingly common.[4] All in all, according to the World Health Organization, the number of people with disabling hearing loss has been increasing, and an estimated 1.5 billion people are currently living with hearing loss[21]. Table 1 shows the sound pressure levels of some example sounds in decibels, and describes their effect on human beings.

Hearing loss can either be *conductive* or *sensorineural*. Conductive hearing loss means that there is a problem in transferring the sound waves in the ear canal, eardrum or middle ear. As an extreme example, some individuals are born without the ear canal. Sensorineural hearing loss means that the cause of hearing loss is either in the inner ear, sensory organ, or the auditory nerve.[4]

Hearing loss can be either temporary or permanent.[14] One of the most common causes for hearing loss is exposure to noise.[14] The ear can recover from short periods of noise exposure, but being exposed to loud noises frequently can cause a permanent shift in the hearing threshold. Excessive exposure to noise destroys the hair cells in the organ of Corti, which leads to hearing loss.[22] Different stages of damage to the organ of Corti are shown in Figure 7a–d.

Hearing loss has a number of effects on the person's hearing, including reduced audibility and dynamic range, problems in being able to localize sounds and understanding speech in background noise[4]. One of the biggest problems is that it makes communication difficult or impossible, which leads to social problems[14]. In children, hearing loss delays spoken language development[4]. Preventing, identifying and addressing hearing loss is cost-effective for the society[14].

In the case of a loud (>85 dB) sound event, muscles attached to the eardrum

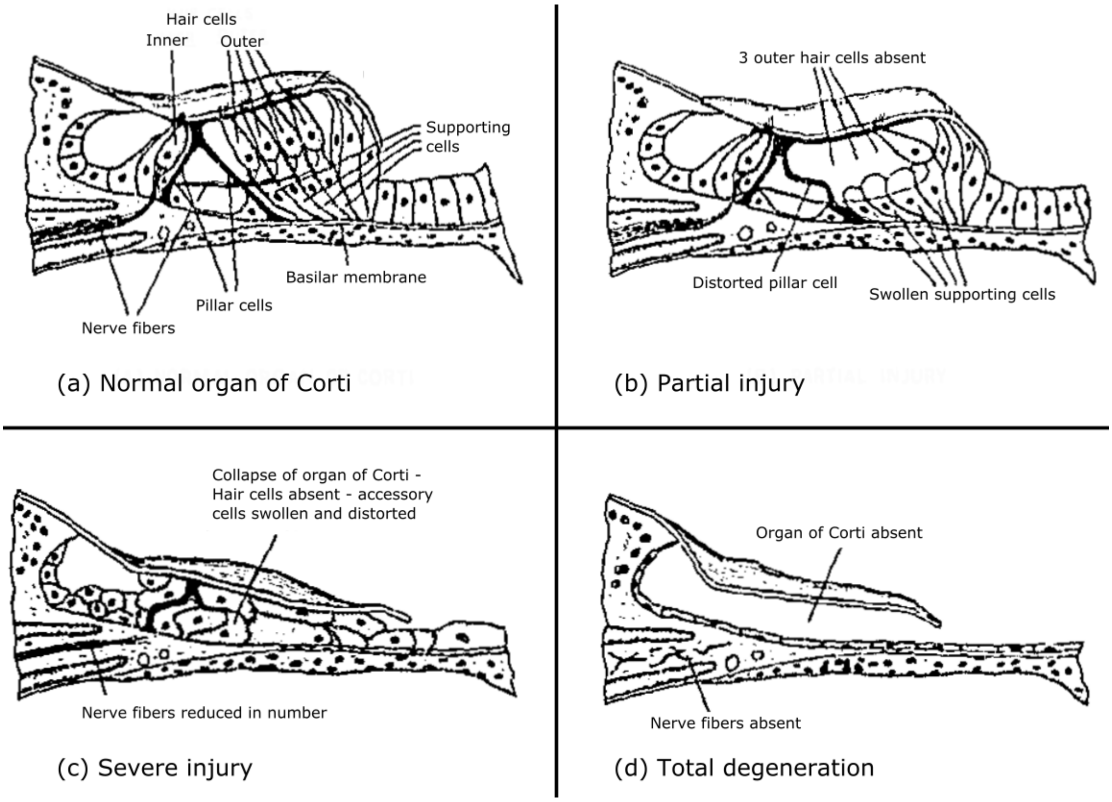

Figure 7: Drawings that illustrate the noise–induced permanent damage in the organ of Corti. Adapted from [22].

and the ossicles pull the stapes away from the oval window of the cochlea, which results in less sound energy being transmitted into the cochlea. This is known as the *acoustic reflex* and it protects the hairs inside the organ of Corti. Unfortunately, it takes up to 1.5 seconds for full protection to occur.[14]

2.4.1 Hearing aids

The purpose of a hearing aid is to compensate for hearing loss, the most important goal being to improve intelligibility of speech[4]. The first hearing aids were horn-shaped devices called ear trumpets, which were directed at the desired sound source, thus eliminating unwanted background noise. Modern hearing aids were developed starting from the 1950s. They utilized the most cutting edge technology at their time, since they were the first consumer devices to use miniaturized components, due to the desire to make the hearing aid portable and inconspicuous, and therefore as small as possible.[24]

A modern hearing aid consists of a microphone, an amplifier, a loudspeaker (known as a *receiver*), a battery and a supporting structure in which all of these components are housed. In modern hearing aids, amplifiers use digital signal processing to process the signal. This processing includes amplification, equalization, and dynamics

Everyday sounds and noises	SPL [dB]	Typical effect after prolonged exposure
Hearing threshold	0	
Normal breathing	10	
Ticking watch	20	
Soft whisper	30	Typically no hearing damage
Refrigerator hum	40	
Normal conversation, air conditioner	60	
Washing machine	70	You may feel annoyed by the noise
City traffic (inside the car)	80–85	You may feel very annoyed
Leaf blower	80–85	Damage to hearing possible after 2 hours of exposure
Motorcycle	95	Damage to hearing possible after about 50 minutes of exposure
Subway train, sporting events	100	Hearing loss possible after 15 minutes
Nightclub	105–110	Hearing loss possible in less than 5 minutes
Shouting	110	Hearing loss possible in less than 2 minutes
Standing beside or near sirens	120	Pain and ear injury
Firecrackers	140–150	Pain and ear injury

Table 1: Sound pressure levels of everyday sounds and their effect on human beings. Adapted from [23].

processing in the form of compression: in practice this is implemented as a *multiband compressor*[4]. Hearing aids are usually adjustable, so that they can be used in different situations. For example, a gain setting that is best suited for a conversation in a quiet home might not be the best for another environment.

Usually, a hearing aid has multiple fitting sessions where its parameters are adjusted, and its gain is increased gradually. Hearing loss makes the ear more sensitive to the frequencies where hearing loss has occurred, resulting in the sound feeling uncomfortable or painful at a much lower sound pressure level. For example, a 20 dB hearing loss at a certain frequency should not be compensated with a 20 dB boost at that frequency. Instead, the required gain is based on the listener's *most comfortable level* (MCL).[4] In 1944, Lybarger [25] obsereved that on average, people chose to amplify the signal at a specific frequency by approximately half the amount of hearing threshold loss at that frequency. This is known as the *half-gain rule*. A hearing aid does not make the hearing perfect, and e.g. distinguishing between spectral differences becomes more difficult when sound is amplified.[4]

2.5 Psychoacoustics and listening tests

Psychoacoustics is the scientific study of how humans perceive sound This includes loudness of sound, sound localization and many other facets. Psychoacoustics can be studied by performing various *listening tests*, also known as *psychoacoustic tests*. When designing a psychoacoustic test, the sounds chosen should enable the test subject to reliably report the characteristics of the auditory event. Some of the most common stimuli include pure–tones, amplitude– or frequency modulated tones, sine-wave sweeps, chirp signals, white noise and pink noise, just to mention a few. The auditory system is a very complex non-linear, time-varying system, which makes it difficult to measure.

Listening tests are typically performed either on headphones or loudspeakers. In some cases, it can be beneficial if the test space reflects the situation where the application is used. For example, listening tests for loudspeakers might give more accurate results when performed in a listening test room with acoustics resembling an ordinary living room, not in an anechoic chamber.[13] It might also be useful to perform tests in other real-life environments, e.g. in noisy environments[14]. Recently, self-made hearing threshold assessments and other hearing tests have become more common in personal sound applications[5][7][9] and hearing aids[10][11].

2.5.1 Hearing threshold and equal–loudness measurements

The hearing threshold measured in audiometry is an *absolute threshold*: a limit of hearing. A *difference threshold* means the smallest possible audible change in one of the attributes of the sound event, for example in an equal–loudness measurement.[13]

Psychoacoustic tests can also be classified according to their method[13]. For example, *Békésy audiometry* is a test done by the *method of tracking*, where the test subject's actions influence how the sound is changed. In Békésy audiometry, the test subject is instructed to press a button as long as they are able to hear the test sound, and release the button when they can no longer hear it. The level of the signal is decreased until the button is released. The test signal is a slow frequency sweep.[26]. The equal–loudness test, however, is an *adjustment test* (also known as the *method of adjustment*), where the user is asked to adjust the value of an attribute until it has reached a desired value, in this case the same perceived loudness as the reference signal has[13].

An *audiogram* is a graph that shows the individual's hearing threshold in comparison to a standardized hearing threshold. The audiogram can be established with the aid of hearing test (*audiometry*), most often pure-tone audiometry. In pure–tone audiometry, a listener listens to a sine wave test signal with headphones at different volumes, and the lowest volume at which the listener reports to having heard the signal is recorded as their hearing threshold for that frequency. The test is performed separately for each ear and each test frequency. To guarantee the accuracy of the test result, the tests should be done by professionals with calibrated equipment and test facilities according to international standards.[4] The ISO standard for audiometry [27] gives guidelines for the sound pressure levels, air temperature and other details of the testing environment.

There are also other ways to test hearing. For example, since speech intelligibility is often the most important goal in hearing tests, hearing can also be tested by playing speech sounds that are mixed with noise[4]. Hearing tests can also be automated; for example, the aforementioned Békésy audiometry and equal–loudness tests are automated[26]. Finally, traditional hearing tests have been recently improved with the help of new technologies, such as *machine learning*[28].

When performing hearing threshold measurements, the personal character and confidence of the listener affects the results. Some subjects tend to perceive auditory events even without a sound event, and some only report hearing very clear auditory events. This means that e.g. in hearing threshold measurements the audibility of a signal should be tested multiple times to verify the result.[13]

2.6 Head-related transfer function

A *head-related transfer function* (HRTF) is a transfer function that is used to describe how the shape and material properties of the human head and upper body affect hearing. This transfer function can be defined separately for different angles of incidence and also for different points in the ear canal. HRTFs have many uses, including generating binaural signals for virtual reality and auralization applications, and determining the headphone transfer function.[29]

Since every individual has a uniquely shaped head, the transfer function is also personal and changes as the shape of the head changes, usually through ageing. For example, the shape of the ear canal can alter the way different frequencies are perceived between individuals.[29]. This can be seen e.g. in Figure 8 which shows the sound transmission to the eardrum from the entrance of the ear canal for 12 test subjects. It can be seen that for certain frequencies, the differences can be as much as 20 dB[30].

The simplest model of the ear canal is a straight 3 cm tube with a hard termination. Hammershøi and Møller [30] question the validity of this simplification, since amplitude peaks for e.g. test subjects AH, FC and RM do not follow a 3:1 ratio, as can be seen from Figure 8.

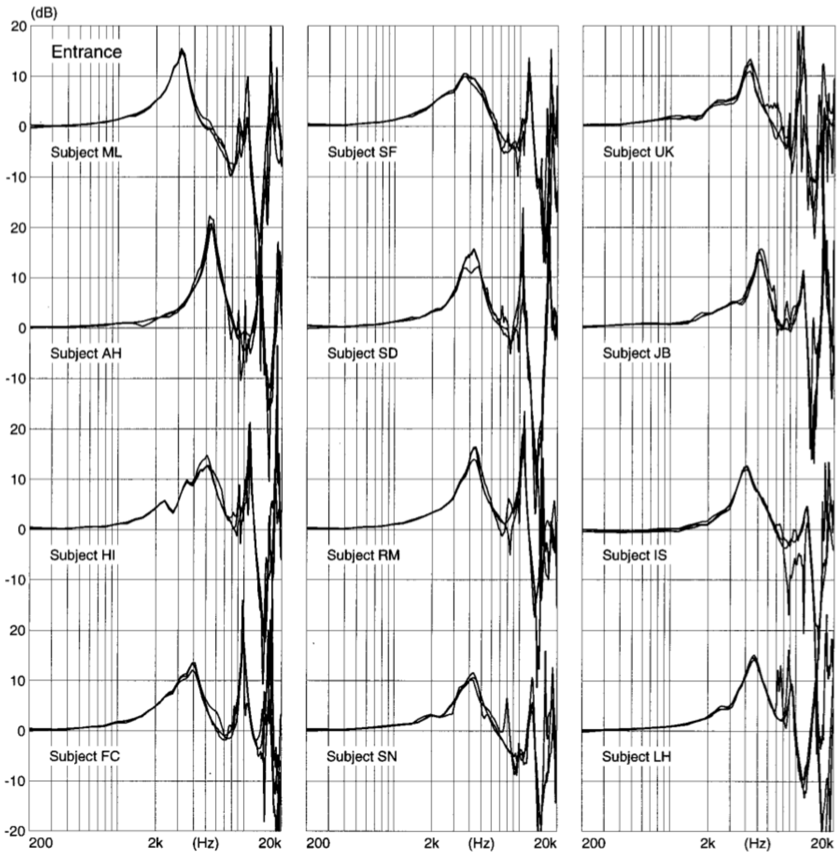

Figure 8: The frequency magnitude response of sound transmitted to the eardrum from the entrance of the ear canal for the left ear of 12 subjects. Each frame contains measurements for three different directions. Adapted from [30].

2.7 Headphones

Headphones were first used by telephone switch operators, and the first telephone headsets were headphones[31]. The Sony Walkman in 1979 popularized the headphone for consumer use[31], and they have become increasingly popular ever since[32].

Compared to loudspeakers, headphones are more portable and have better isolation and separation between channels[33]. They are also capable of producing excellent sound quality regardless of the acoustic properties of the listening environment, attenuation of ambient noise, and offer increased privacy as well as being potentially less disturbing to people in the same or adjacent physical spaces[34].

Dangers and possible drawbacks of headphone listening include the lack of sociability and dangers arising from not being able to hear your surroundings. When listening with headphones, the effects of the listening environment and the person's head are not present, which is a difference to when listening to audio with loudspeakers. Therefore, the listener experiences the sounds inside their head.[33]

There are different types of headphones: *circumaural*, which press against the head and cover the entire ear; *supra-aural*, which press against the ears; *intra-concha*, which are located right outside the ear canal; and *in-ear*, which are pushed into the ear canal[35]. Different headphone types are illustrated in Figure 9.

2.7.1 Frequency response of headphones

Usually, loudspeakers are designed to have a flat frequency response[14]. While headphones are essentially small loudspeakers, in practice they are different from loudspeakers from them in several ways.

When the magnitude response of a loudspeaker with a flat magnitude response is measured from inside the ear, the shape of the ear changes the response by boosting frequencies around 3 kHz[35]. This is shown in Figure 10. Therefore, unlike a loudspeaker, a good headphone should not have a flat frequency response[33]. However, an industry standard target response has not been determined, and different headphone manufacturers have different views of what they desire from the target response.

Toole [36] outlined problems in the design and evaluation of stereophonic headphones already in 1984, especially the problem with individual variation between ears. Toole also predicts that while some problems can be solved in time, some will remain.[36] Theile [34] presented and compared different methods of determining the frequency response of headphones already in 1986. The two methods mentioned are *loudness–comparison* and *diffuse field response*. Voinier [37] asked test subjects to grade the quality of different headphone transfer functions. Recently [1], it has been determined that due to the large amount of variation in the HRTFs between different people, it is not possible to find a standard headphone equalization that would offer the best possible audio fidelity for most individuals. Finally, recent research has found no correlation between headphone frequency response and their retail price[38].

Different headphone types also have different frequency responses as indicated by Figure 11. When using in-ear headphones, the resonance of the ear canal changes

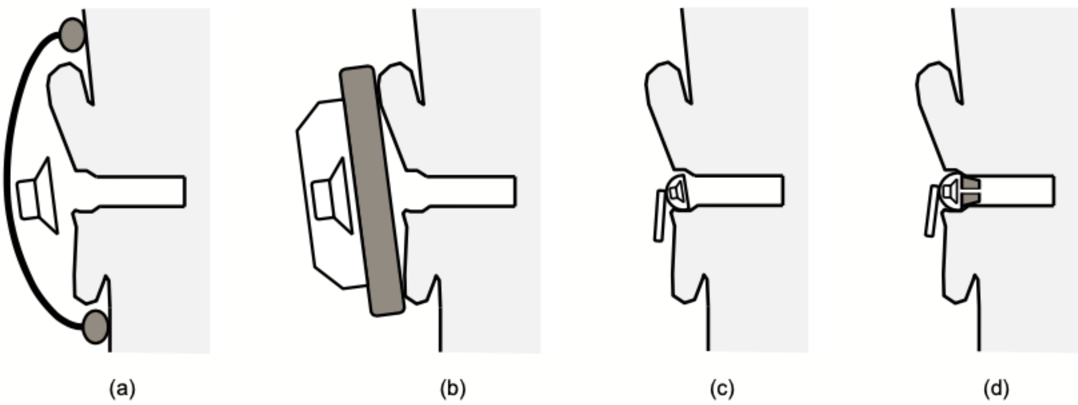

Figure 9: Headphones according to ITU-T Recommendation P.57: a) circumaural, b) supra-aural, c) intra-concha and d) in-ear. Adapted from [35].

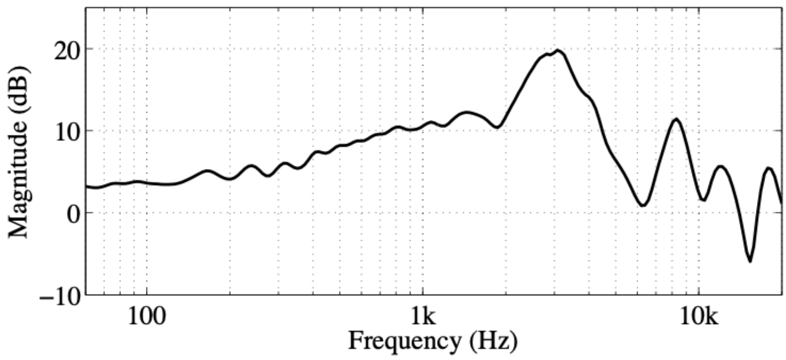

Figure 10: Effect of the torso, outer ear, and ear canal on the frequency response measured inside the ear canal of a human subject. The peak at around 3 kHz is created by the quarter-wavelength resonance of the open ear canal. Adapted from [35].

because it turns from a quarter wavelength resonator into a half-wave resonator. This removes the quarter-wavelength resonance peak and adds a half-wave resonance peak, which would make the sound unnatural.[33] This has to be taken into account in the design of the headphone. Additionally, some in-ear headphones introduce an *occlusion effect*, where the speaker's own voice sounds unpleasant (or at least strange) because the headphones prevent them from hearing airborne sound[39]. There are ways to mitigate this by adding active or passive *leakage* to the headphone design[39].

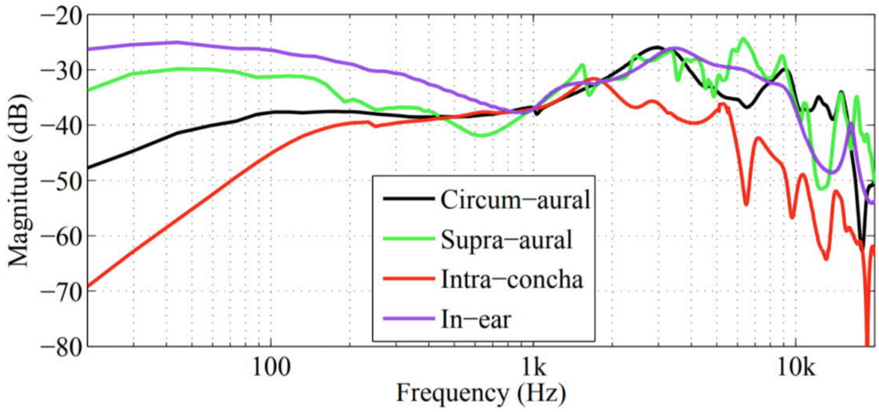

Figure 11: Magnitude reponses of different headphone types. Adapted from [35].

The quality of headphones is measured by using a dummy head or an ear canal simulator. Since headphones will always sound a little bit different depending on their position, several measurements are performed with the headphones slightly differently positioned each time. If the fit is not correct, the magnitude response is drastically different from the ideal.[33]

2.7.2 Headphone equalization and personal sound

The differences in headphone magnitude responses, as shown in Figure 11, and the variation between individuals, as indicated by Figure 8, have led to research in different headphone equalization systems which are often referred to as personal sound technologies. Some examples include the self–calibrating earphone by Hefio[3], the Nuraphone by Nura[6], Audiodo[5] and SoundID[7]. They all use hearing data to modify the output of an audio device, and therefore fall within the definition of personalized sound. SoundID and Audiodo are based on hearing tests, while the Nuraphone claims to obtain hearing data by measuring *otoacoustic emissions*, very quiet sounds that the cochlea makes when it is exposed to sound[13]. The self–calibrating earphone by Hefio estimates the acoustical transfer function from sound pressure at the entrance of the ear canal to sound pressure at the eardrum by measuring test sounds with a microphone[3].

Many systems in the field of personal sound have already been patented. Raz and Aronson [40] describe a method and device for playing modified audio signals in accordance with hearing capabilities of an individual. Swartz [41] describes a listener specific audio reproduction system that corrects distortion caused by the system as well as hearing impairment suffered by the listener. Raz [42] describes a system and method for improved audio perception for optimal listening of music, which include a hearing test, audio signal compensation, and a signal processing system, all implemented in an audio device. Rader [9] describes a system for sound enhancement for mobile phones and other products producing personalized audio for users, which includes a hearing profile, a hearing loss profile and a personal choice profile.

3 Measurements

This section describes the various measurements that were performed during this research. First, measurements with diagnostic audiometry are described. Then, measurements with the novel equal–loudness system and the basic pure–tone audiometry test system are described.

3.1 Diagnostic audiometry

In order to have a reliable point of comparison for the equal–loudness and the basic pure–tone audiometry tests, the audiograms of two test subjects were obtained by performing a standardized diagnostic pure–tone hearing test. This also helped in becoming familiar with the standard way in which hearing threshold measurements are usually performed. The audiometer that was used in these measurements was the Interacoustics Diagnostic Audiometer Model AD28, shown in Figure 12.

Figure 12: Close-up of the front panel of the Interacoustics Diagnostic Audiometer Model AD 28.

3.1.1 Calibration of the audiometer

In order to use the audiometer, it was necessary to determine how much the sound pressure levels it produces have drifted since the last time it was calibrated. Instructions for calibration were provided by the service manual[43] and the ISO

389-1:2017 standard[44]. The calibration procedure includes both hardware and software calibration, but software calibration was deemed sufficient.

According to the manual, the calibration requires the following instruments: a sound level meter (Figure 14), artificial ear (Figure 15), a 6ccm coupler, an artificial mastoid and a sine wave generator. The mastoid is a bone in the human skull, which is a part of the temporal bone. It houses the structure of the ear and is therefore important for bone conduction. However, the artificial mastoid was not required in this case, because the bone conducting test is not relevant for this research.

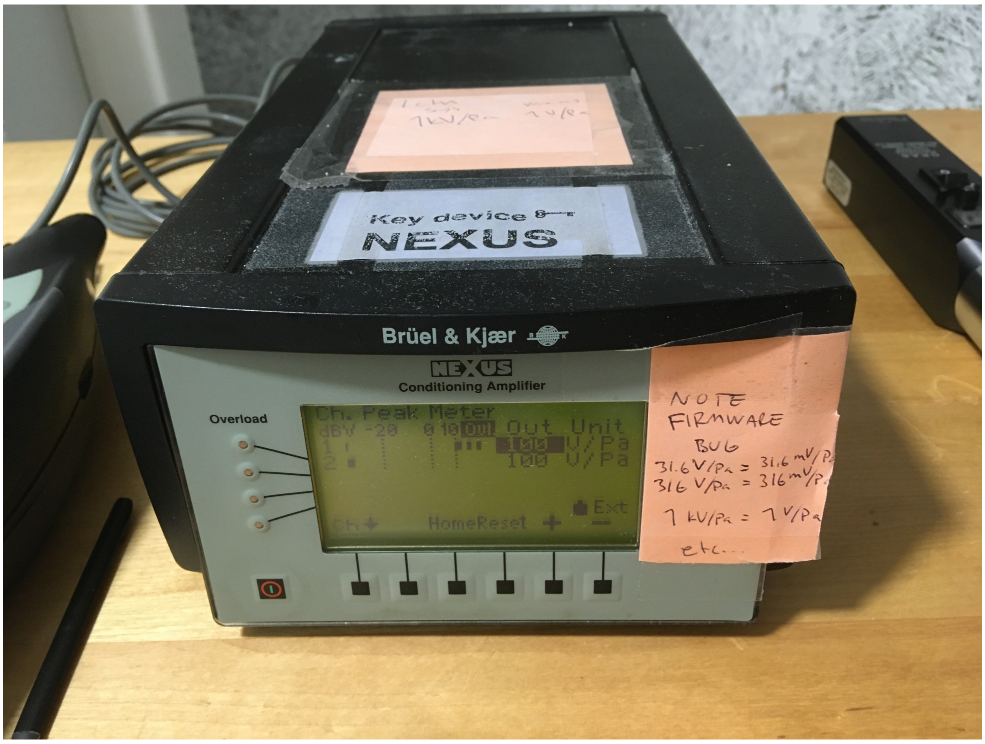

Figure 13: Brüel & Kjær Nexus Conditioning Amplifier.

Before calibrating the audiometer, the sound pressure level meter needed to be calibrated according to the instructions in its user manual[45]. It was calibrated by using the GRAS 42AP Intelligent Pistonphone, which produces a stable 250 Hz signal at 114 dB[46]. Also, the artificial ear was to be attached to a microphone. The artificial ear used was a Brüel & Kjær Artificial Ear 4153 and the microphone was a Brüel & Kjær Type 4192.

The artificial ear fulfils the requirements of IEC 60318-1 standard for an ear simulator for the measurement of supra–aural and circumaural earphones, meaning that it has an acoustical impedance similar to that of the human ear[47][48] and it also includes the 6ccm coupler.

The setup was as follows:

1. The desired test signal from the audiometer was fed into the Telephonics TDH-39P Audiometric Earphones, shown in Figure 16.

2. The output of the headphone was detected by the Brüel & Kjær microphone which was attached to the artificial ear.

3. The signal of the measurement microphone was fed into the Brüel & Kjær Nexus Conditioning Amplifier for amplification, shown in Figure 13.

4. The sound pressure level meter measured the sound pressure level of the test signal, which was recorded.

Steps 1–4 were then repeated separately for each signal and for both channels. The IEC 60318-1 standard[48] contains a table which shows the reference sound pressure levels for each frequency; the difference between this value and the value of the sound pressure level meter is the error of the measurement. This error is shown in Figure 17 for both channels separately. It can be seen that for most frequencies, the error is less than or equal to 1 dB. The accuracy was considered to be sufficient for the purposes of this measurement, and that therefore no further calibration was necessary.

Figure 14: Brüel & Kjær Hand Held Analyzer 2250. It was used as the sound pressure level meter when calibrating the audiometer.

Figure 15: Brüel & Kjær Artificial Ear 4153 with IEC coupler.

3.1.2 Measurements with the audiometer

The front panel of the audiometer is shown in Figure 12. The front panel has numerous controls and a screen that shows the sound pressure level (relative to hearing level) and the frequency of the test signal. There are two LEDs, one indicating when the test tone is being played, and one that lights when the test subject has responded to a test signal by pressing a button. There are two buttons for increasing and decreasing the test tone frequency, a button for choosing if the test tone level is adjusted in 1 dB or 5 dB intervals, and a rotary jog wheel for changing the level of the test frequency. The panel has many other additional controls, but they were not relevant for this measurement.

The audiometric measurements were performed according to the ISO 8253-1 standard[27]. The test was carried out using manual audiometry. There were 11 test frequencies: 125 Hz, 250 Hz, 500 Hz, 750 Hz, 1000 Hz, 1500 Hz, 2000 Hz, 3000 Hz, 4000 Hz, 6000 Hz and 8000 Hz. The test tones were presented from 1000 Hz upwards, and then followed by the lower frequency range (125 Hz – 750 Hz) in descending order. The test tone duration was between 1–2 s.

There are two different ways to test hearing, the *ascending method* and the *bracketing method*. In this case, the ascending method was used. During the test the

Figure 16: Telephonics TDH-39P Audiometric Earphones, positioned on top of the Brüel & Kjær Artificial Ear 4153 shown in Figure 15.

test subject was seated in a quiet audiometric test room, wearing the headphones shown in Figure 16. The test subject was instructed to press the button whenever they heard a sound.

As instructed by the ISO 8253-1 standard[27], the test subjects were first familiarized with the task by presenting a tone of 1000 Hz at a clearly audible hearing level (e.g. 40 dB). Then, the level of the tone was reduced in steps of 20 dB until no response occurred, and then increased in steps of 10 dB until a response occurred. The response was verified by presenting the test tone again. This level was recorded for future reference, because it was used as the initial test sound level of the hearing test. According to the ISO 8253-1 standard, the hearing test according to the ascending method can be performed using the following four steps:

Step 1:

- The first test tone, whose level is 10 dB below the lowest level of the response of the test subject during the familiarization session, is presented. If the test subject does not respond to the test tone, its level is increased in steps of 5 dB until a response occurs.

Step 2:

- After the response, the level is decreased in steps of 10 dB until there is no response. Then, another ascent is started with 5 dB steps. This is continued, until three responses occur at the same level out of a maximum of five ascents. This level is recorded as the hearing threshold level for that particular frequency.
- If there has not been three or more positive responses, a test tone 10 dB louder than the level of the last response is presented to the test subject. Then, the test procedure is repeated: the level is decreased 10 dB after a response, and increased 5 dB until a response occurs.

Step 3:

- The previous responses are used as an estimated audible level, which is the starting point for the level of the next test frequency. Step 2 is repeated until all test frequencies for one ear are finished.
- Finally, the measurement at 1000 Hz is repeated. If the results at 1000 Hz agree to 5 dB or less with those of the first measurements for the same ear, the

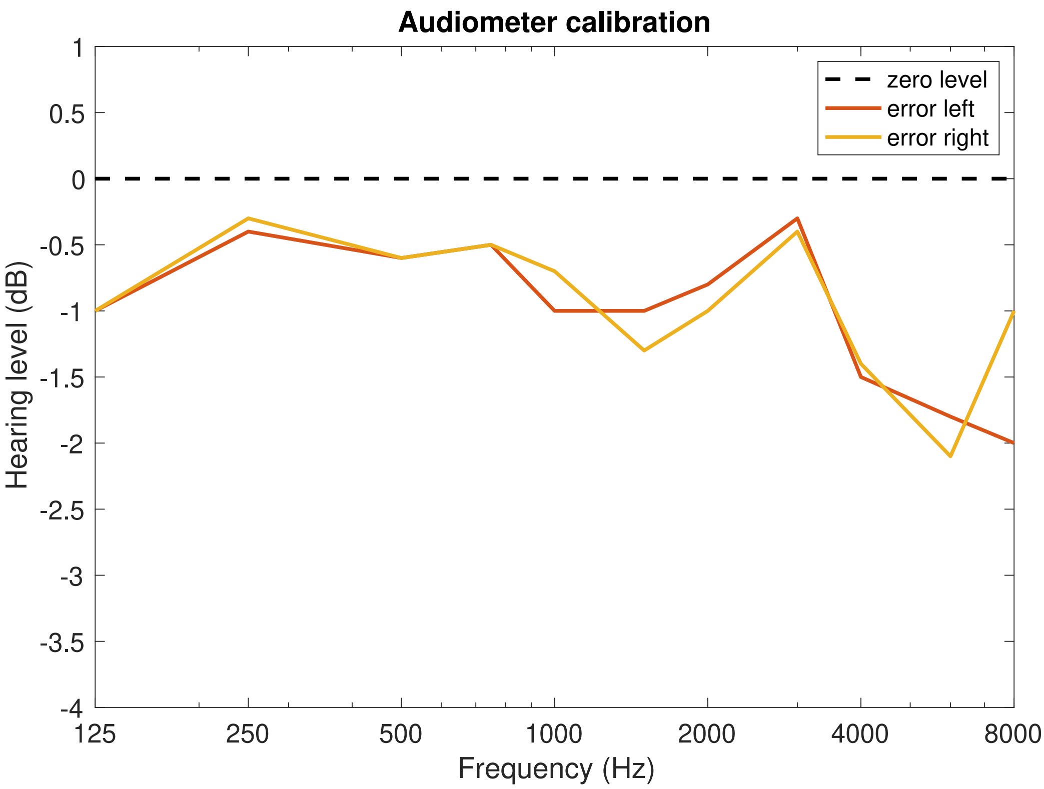

Figure 17: Error of the audiometer. Dotted line shows the zero level. The orange line shows the error of the left channel and the yellow line shows the error of the right channel.

other ear can be tested. If the difference is 10 dB or more, the test is to be started over until agreement to 5 dB or less has been obtained.

Step 4:

- Steps 1–3 are repeated until both ears have been tested.

The results of the audiometric measurement for two test subjects can be seen in orange in Figures 23 and 24.

3.2 Equal–loudness test program

A novel equal–loudness test program was developed as a part of this research and tested with five test subjects. The motivation behind using an equal–loudness test instead of a traditional pure–tone based test signal was to determine if it is more usable in a noisy everyday environment, and additionally, if comparing the loudness levels of two signals would be faster than trying to assess whether or not a signal is audible, which is how most hearing tests work.

When using this equal–loudness test, the situation differs from diagnostic audiometry, where the frequency response of the headphone is known. Let us consider a situation where the ear of a certain test subject has the transfer function H_E and a certain headphone has a transfer function H_H. If these systems are connected to each other (in other words, that the test subject wears the headphones), their combined transfer function H_T – the transfer function of the system consisting of the ear and the headphone – can be defined as a multiplication[49]; in other words

$$H_T = H_H H_E. \tag{5}$$

When performing hearing characteristic measurements with headphones, it is not actually the ear that is being measured but the transfer function H_T. When performing traditional audiometry, the headphone transfer function of audiometric headphones is known, which makes it possible to determine the transfer function H_E of the ear, or

$$H_E = \frac{H_T}{H_H}. \tag{6}$$

When developing an equal–loudness test that will be used with various different headphones whose H_H is not known, H_E cannot be calculated unless the transfer function of that particular headphone is measured. However, since the headphone and the ear are inseparable in this usage anyway, the separation of these transfer functions is not necessary, and only the total system H_T comprising both the headphone and the ear can be considered.

3.2.1 Generating the test signals

As explained earlier in chapter 2, fractional octave bandpass filters can be used to approximate how human ears perceive loudness. Therefore, noise signals filtered with 1/3 octave bandpass filters were used. The center frequencies for the filters were 63 Hz, 125.9 Hz, 251.2 Hz, 501.2 Hz, 1000 Hz, 1995.3 Hz, 3981.1 Hz and 7943.3 Hz. The frequency magnitude response of the filters is shown in Figure 18. In the test, the 1000 Hz signal was used as the reference signal to which other signals were compared.

At first, the test signals were considered to be filtered from white noise. The power spectral density of white noise is shown in Figure 19. However, it was discovered that filtering the white noise with the 1/3 octave filters, shown in 18, resulted in test signals with different sound pressure levels, which made the test signals incomparable with each other.

It was important to be able to generate test signals that would have the same sound pressure levels. Griesinger [1] used 27 band 1/3 octave filtered pink noise test signals and a 500 Hz 1/3 octave filtered reference signal in his research into headphone measurement through loudness comparisons of noise bands. Horbach [2] used 23 pre-equalized, equal–loudness burst signals that were derived from an auditory ERB filter bank, and multiple reference signals to avoid large differences between the test signals.

It was discovered that filtering pink noise with the 1/3 octave filters resulted in test signals with equal sound pressure levels. It was also discovered that pink noise sounded more pleasant as a test signal than white noise. It would have been possible

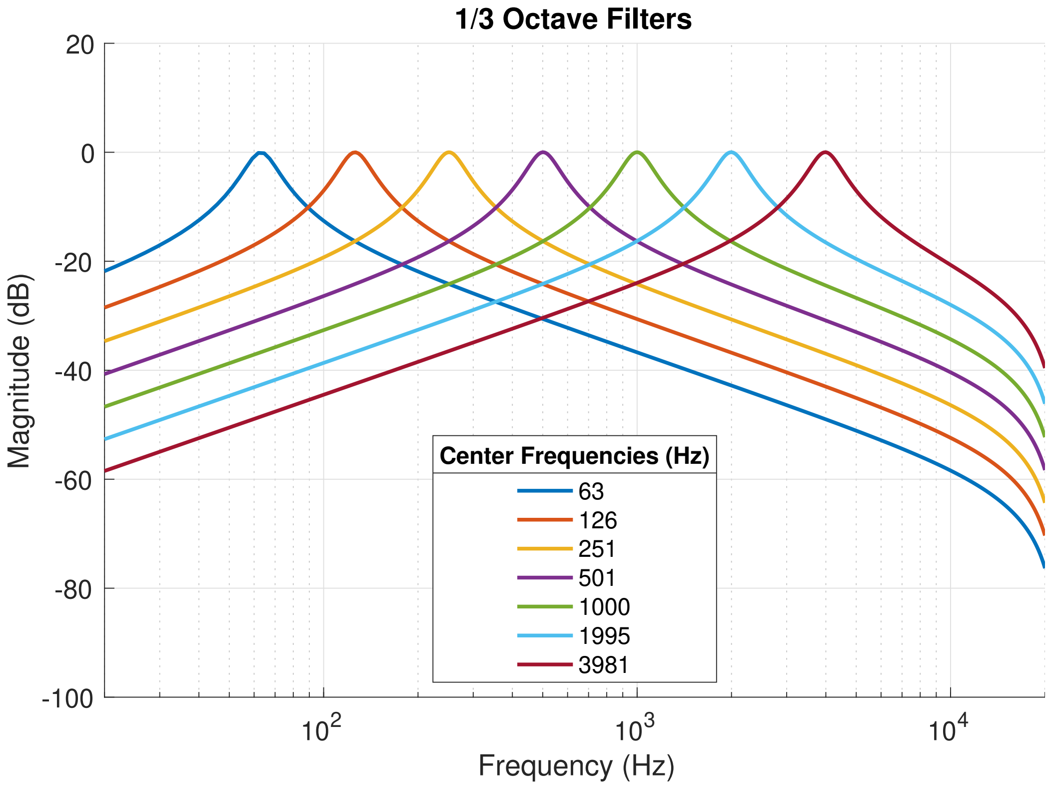

Figure 18: The frequency magnitude response of the 1/3 octave bandpass filters that were used to create the test signals for the equal–loudness test.

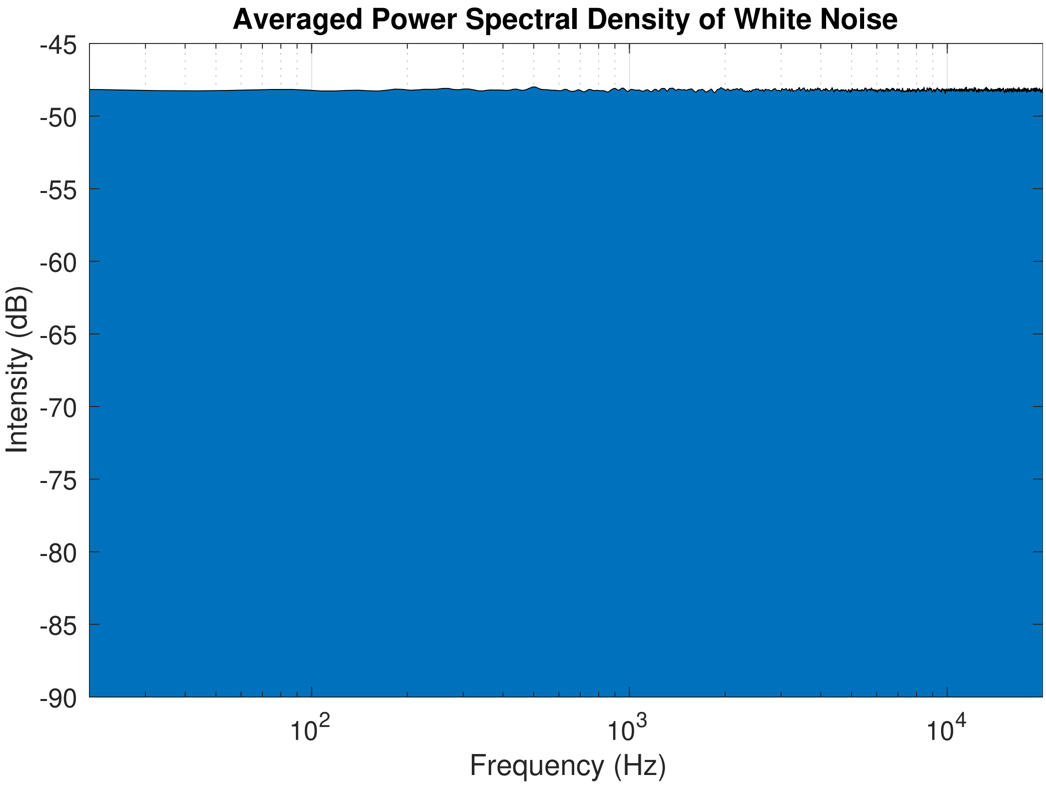

Figure 19: The averaged power spectral density of white noise.

to first generate test signals by filtering white noise and then adjust the gain levels of each signal separately – however, this would have resulted in differing sound pressure levels for the stop band. Wide–band noise signals are also less prone to effects of masking, unlike a single–frequency sine signal, and therefore they should perform better in a situation with background noise. The suitability of other test signals was not investigated.

The sound pressure level of white noise is shown in Figure 19 and the sound pressure level of pink noise in Figure 20. As can be seen, pink noise is noise whose power spectral density is inversely proportional to the frequency of the signal, or in other words

$$S(f) \propto \frac{1}{f} \tag{7}$$

whereas white noise has equal power spectral intensity per frequency interval. It can also be seen that the power spectral density decreases 10 dB per decade of frequency.[14]

The pink noise test signal is one second long. It is filtered with MATLAB's `octaveFilter` System object to generate seven 1/3 octave filtered test signals. The MATLAB implementation is explained in more detail in appendix A.

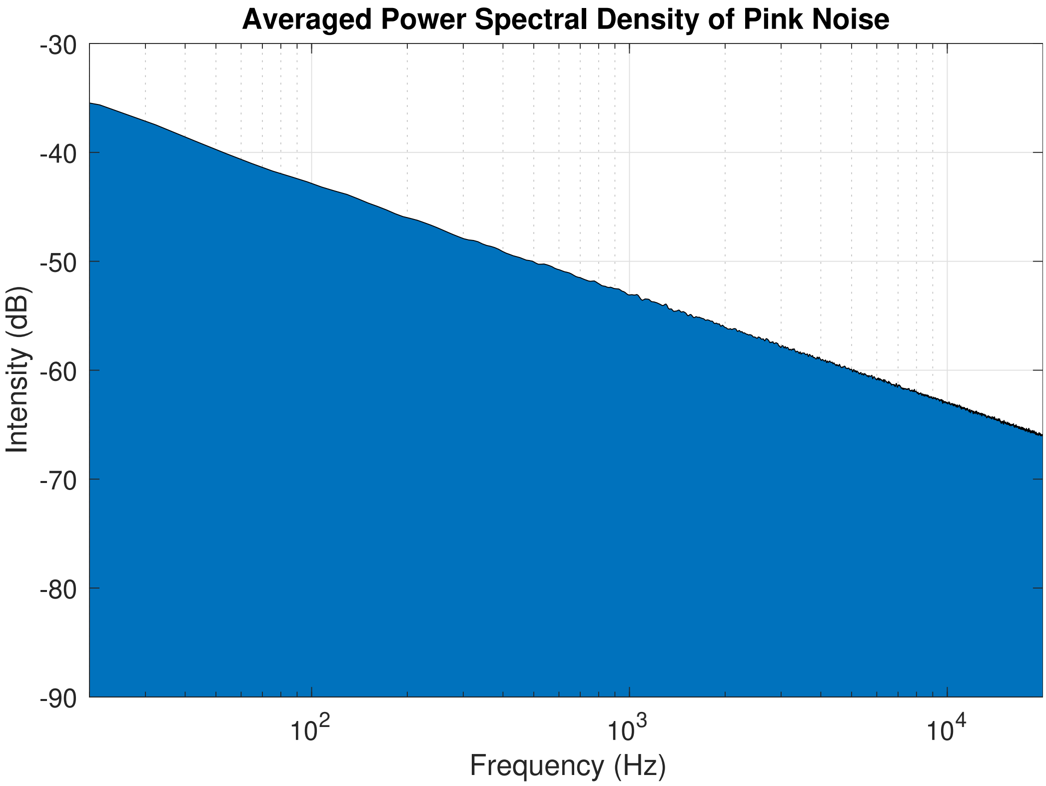

Figure 20: The averaged power spectral density of pink noise. The power spectral density decreases 10 dB per decade of frequency.

3.2.2 Measurements with the equal–loudness test program

The test program for the equal–loudness test is shown in Figure 21. The user interface shows a number of controls on the left side and a results graph on the right side. the *Comfortable level* slider adjusts the level of the reference signal in the beginning of the test. When the user is done, they can press *Set Comfortable Level*, and then start the test by pressing *Start*. The loudness of the test signal can be changed with the *Loudness* slider. There are two number indicators, one for the test frequency and one to indicate which channel is being tested (1 = Left, 2 = Right). There is also a light indicator that shows when the reference sound is being played, so that the user does not confuse the reference signal from the test signal. Pressing *Next* advances the test to the next test frequency. Test results are displayed on the graph. The blue curve indicates the results of the left ear and the red curve indicates the right ear. Additionally, a time indicator shows how long it took to finish the test. The results can be saved to a file by pressing *Save Results* and the test can be taken again by pressing the *Restart* button.

The test program has certain features that are useful for testing but unnecessary if used in most practical implementations: these include the channel and frequency indicators, the time indicator, the graph visualization, and the button for saving results to a file. It is important to note that these results do not tell much about the user's hearing, but instead about what the combined frequency response of the headphone and the ear is like, and that this data is mostly useful when analysed further or used as a basis for an equalization algorithm, like the one developed by Vapalahti[50].

3.2.3 Equal–loudness test process

The test is meant to be performed on headphones. The best results can be achieved in a quiet environment, but the test should perform relatively well in a noisy environment as well. It does not matter what the headphones are like, as long as they are stereophonic. Before starting the application, it should be ensured that the system volume is set at a comfortable listening level (usually 10–50%) and that the headphones have a good fit over or in the ears. System volume should not be adjusted during the test – if the test signals are too quiet, the test should be restarted with a louder initial volume.

The test consists of the following steps:

1. When the program is started, it immediately starts to play a constant signal of pink noise which has been filtered with a 1/3 octave bandpass filter with a 1000 Hz center frequency. This signal is meant for adjusting the volume of the reference signal. The loudness of the test signal is adjusted with the *Comfortable level* slider, so that the signal level is not too quiet or too loud. There is no right or wrong setting for this value, as long as the signal is audible and not painful or annoying to the listener. When the loudness level has been adjusted, it can be set by pressing *Set comfortable level*.

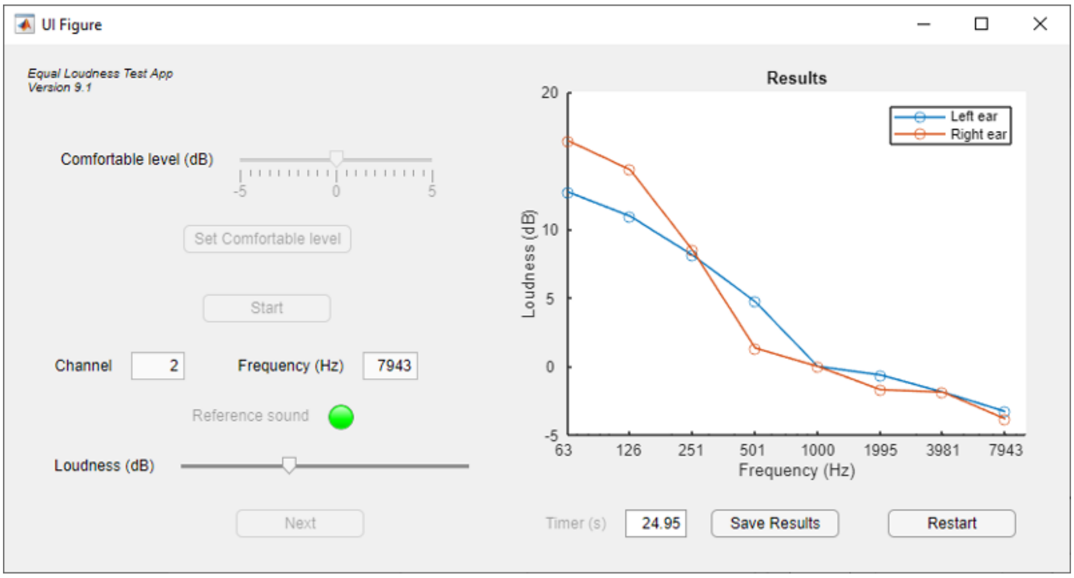

Figure 21: The user interface of the equal–loudness test after a test has been completed.

2. The test is continued by pressing the *Start* button. The test is performed for the left channel first. The application plays a loop consisting of one second long reference sound followed by one second of the test sound. An indicator marked *Reference signal* lights up every time the reference signal is being played. The *Channel* indicator shows the audio channel being tested (1 means left, 2 means right). The *Frequency* indicator shows which frequency band is being tested.

3. The *Loudness (dB)* slider is adjusted so that the test signal sounds as loud as the reference signal. When the user is done with adjusting the loudness, pressing *Next* advances the test to the next test frequency.

When the loudness level for all test signals has been adjusted, the test ends, and the results are displayed as a graph. The graph shows the amount of gain required to make different frequencies as loud as the 1 kHz 1/3 octave filtered pink noise test signal. The results can be saved to a *.mat* file by pressing *Save results*. This data can be used to equalized audio to a target response by opening the *.mat* file with the headphone equalization program developed by Vapalahti[50] or analyzed and processed further by other means. The test can be taken again by pressing the *Restart* button.

For the sake of clarity, this test process is shown in flowchart form in appendix B.

3.3 Basic pure–tone audiometry test program

Additionally, a naive implementation of a basic pure–tone audiometry test, similar to the ones used in personal sound applications, was developed by Vapalahti [50] for comparison purposes. The user interface is shown in Figure 22.

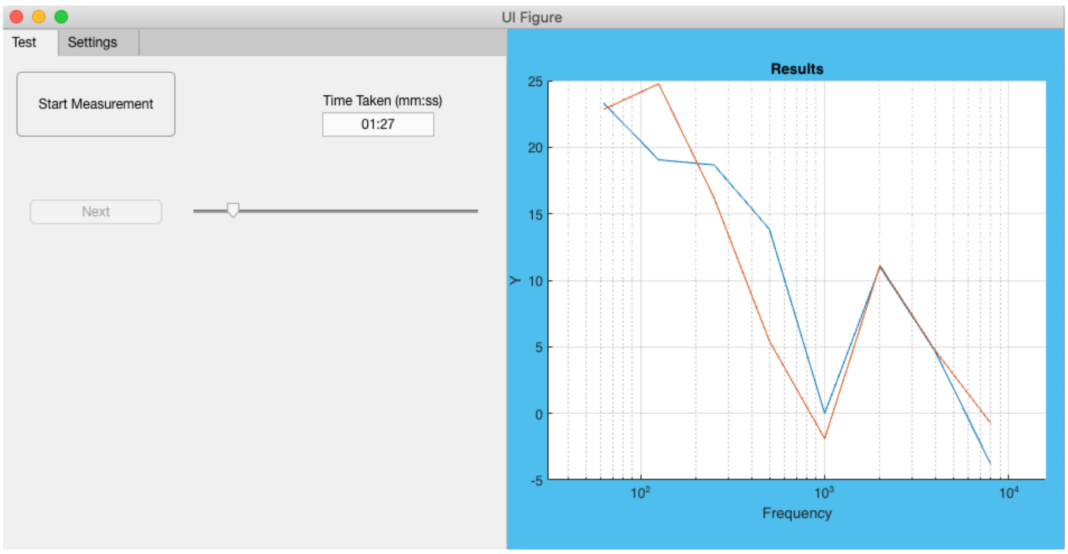

Figure 22: The user interface of the basic pure–tone audiometry test after the test has been finished. Developed by Vapalahti [50].

The basic pure–tone audiometry test works similarly to the ones used in the most primitive personal sound applications today. When the program is started, the user is presented with a screen with the choice to start the measurement or edit the settings. The settings are detailed and beyond the scope of this thesis, and they were not used by our test subjects. The user is asked to adjust the test tone of a pure–tone sine signal until it is not audible any longer at different frequencies and for both ears separately. The test starts at the lowest frequency and continues until the highest. The test frequencies are 62.5, 125, 250, 500, 1000, 2000, 4000 and 8000 Hz. After the test is finished, the results are displayed.

4 Results

This chapter presents the results of the naive implementation of a basic pure–tone audiometry test and the equal–loudness test. The naive pure–tone audiometry test application developed by Vapalahti [50], which is similar to the ones used in most smart phone applications for personal sound hearing assessment, was used as a reference test. This is a naive implemention of a basic pure–tone hearing threshold test as described in chapter 2.5.1.

Both the basic pure–tone audiometry and the equal–loudness test went through several iterations, as is common in any software development cycle. The results presented below are based on the final versions of both applications. For example, both programs initially suffered from the clipping of test signals due to oversights in gain staging, which made test data from early versions unusable later.

Test subjects were also asked to give free-form written feedback about the test. The comments were mostly informal in tone and related to their subjective experience, the hardware and software problems that they encountered and the characteristics of their headphones. These comments are referred to when necessary.

4.1 Comparison of the equal–loudness and basic pure–tone tests with diagnostic audiometry

Figures 23 and 24 show the comparison between the results for two different test subjects, respectively, for the equal–loudness test, the basic pure–tone audiometry test and the average of two separate tests performed with diagnostic audiometry. The equal–loudness and basic pure–tone audiometry results were converted to be hearing level equivalent with the functions provided by the IoSR Matlab Toolbox[51], so that they could be compared with the diagnostic audiometry results. Hearing level is measured in decibels with reference to the *audiometric zero*, or the hearing threshold[52].

Since the frequency response of the headphones that were used for the equal–loudness and basic pure–tone audiometry tests is not known, the basic pure–tone audiometry and equal–loudness results are not directly comparable to diagnostic audiometry. For example, the high and low frequency boost in both the basic pure–tone audiometry and equal–loudness tests might be due to the headphone frequency response. Moreover, the sound pressure levels produced by the headphones used in testing are not known. However, this comparison still illustrates that the tests produce results that are similar to diagnostic audiometry.

4.2 Comparison of the equal–loudness test and basic pure–tone test in a noisy environment

Pure–tone audiometry is meant to be performed in an environment that is as quiet as possible in order to obtain an accurate audiogram, but a mobile application user might take the test in an ordinary apartment, a noisy office, or even outdoors, where

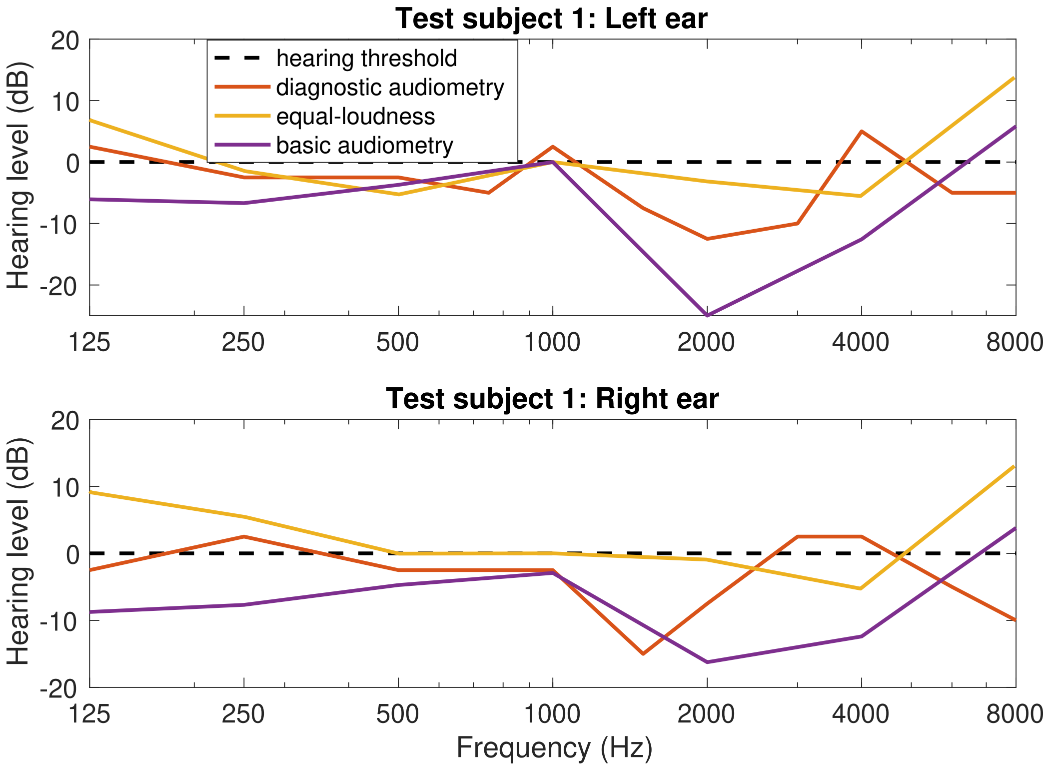

Figure 23: Comparison between the equal–loudness test, the basic pure–tone audiometry test and the average of two separate tests performed with diagnostic audiometry for test subject 1.

several noise sources can disturb the measurement and thus degrade the measurement accuracy.

In order to find out if the equal–loudness test performed better than the basic pure–tone test in a noisy real-life situation, test subjects were asked to perform both the equal–loudness and basic pure–tone audiometry tests once in a noisy outdoor environment, and once in a quiet indoor environment. All test subjects were male, of normal hearing, had participated in listening tests before, and were professionals or students working in the field of audio technology.

The absolute difference in decibels between the tests performed in noise and the test performed in a quiet environment were calculated for each test subject frequency and each ear. E.g. if the 2000 Hz result in noise was 20 dB and the result in a quiet environment was 15 dB, the difference is

$$|20\,\mathrm{dB} - 15\,\mathrm{dB}| = 5\,\mathrm{dB}. \tag{8}$$

The results are visualized in Figures 25–31 by using box plots, a data visualization method originally introduced by John Tukey[53]. Box plots are among the most frequently used statistical graphics[54]. On each box, the red central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th

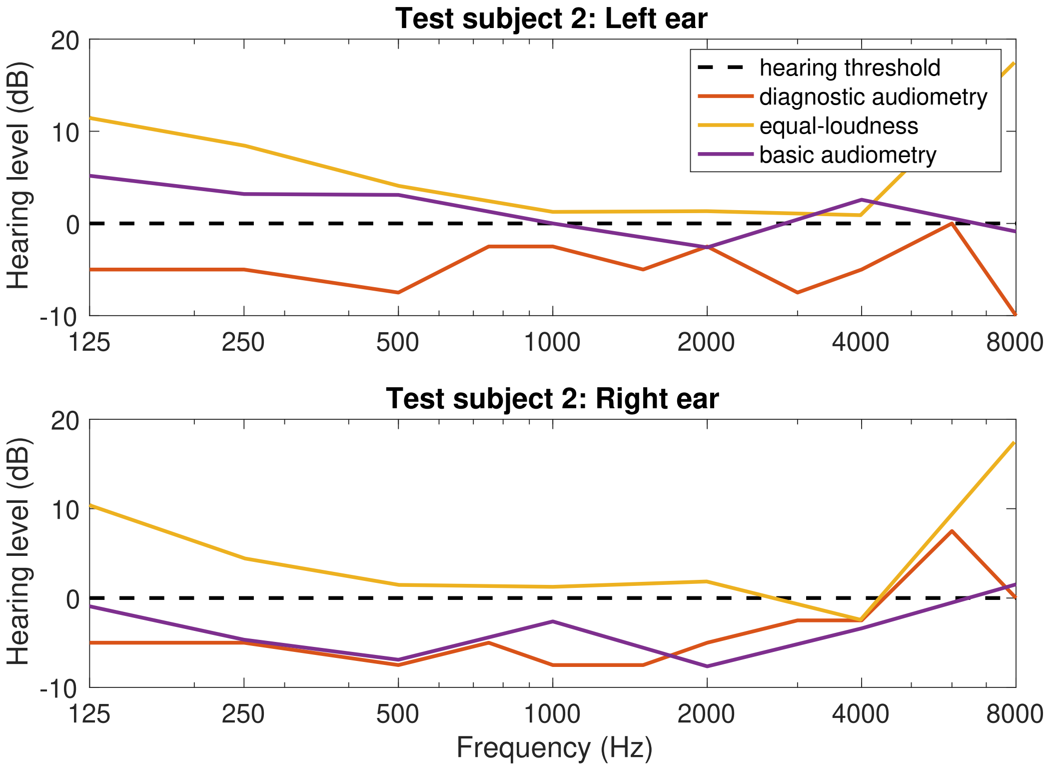

Figure 24: Comparison between the equal–loudness test, the basic pure–tone audiometry test and the average of two separate tests performed with diagnostic audiometry for test subject 2.

percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the '+' symbol. The boxplot is explained in more detail in appendix C. The outliers in these particular box plots might e.g. refer to a hastily performed test, or to a situation where the background noise characteristics have differed from other test situations (e.g. sudden louder noise level or background noise with a different frequency spectrum compared to ones experienced by other test subjects).

Additionally, the mean and the standard deviation of the differences in loudness levels between tests in noise and tests in a quiet environment were calculated. The mean of the difference between all equal–loudness tests was 2.1 dB and the mean of the difference between all basic pure–tone audiometry test results was 4.5 dB. The standard deviation was 1.8 dB for the equal–loudness test and 4.0 dB for the basic pure–tone audiometry test.

While these results are not conclusive due to the small number of test subjects and differences in testing environments, they do suggest that the equal–loudness test could be more accurate in a noisy environment. Comparison between the basic pure–tone audiometry and equal–loudness test is shown in Figure 25 and the mean, median and standard deviation of the tests are shown in Table 2.

	avg	med	sd
equal–loudness	2.1	1.3	1.8
pure–tone	4.5	3.5	4.0

Table 2: Mean, median and standard deviation of equal–loudness and basic pure–tone audiometry tests. All values are in decibels (dB). Measurements in noise were compared to measurements in a quiet location and their difference was calculated. The equal–loudness test performed consistently better in all cases.

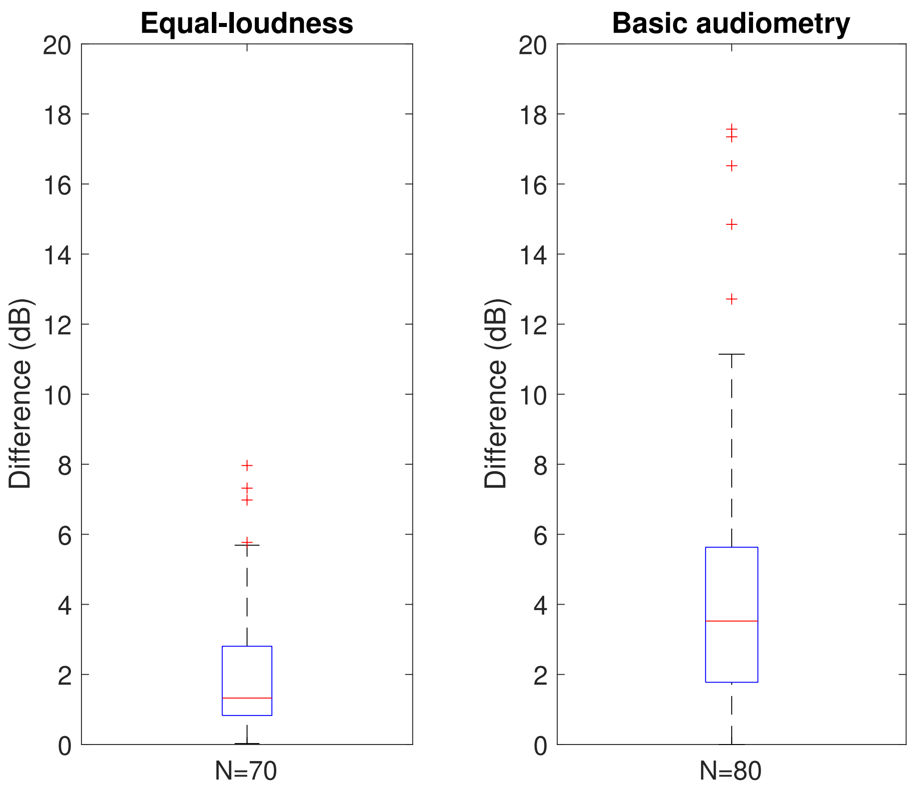

Figure 25: Box plot showing the differences in test results for equal–loudness and basic pure–tone audiometry. Tests performed in a noisy environment were compared to tests performed in a quiet environment. The numerical values for the mean, median, standard deviation and variance for both test types are shown in Table 2.

4.2.1 Variation between test subjects

Figure 26 shows the difference in loudness levels between basic–pure tone audiometry performed in noise and in a quiet environment. The larger the mean is, the bigger the average difference was between the results of the basic pure–tone audiometry test in noise and in a quiet environment, and therefore the more susceptible the test was to environmental noise. It is interesting to see that test subject 1 has widely different results from the rest of the test subjects. According to the test notes, test subject one performed the test in a very noisy place next to a loud fan, therefore making the test conditions different from the other test subjects. This, however, demonstrates that the basic pure–tone audiometry test is very unreliable in a very noisy environment.

In this light, the results of the equal–loudness test, shown in Figure 27, are very interesting. The performance for test subject 1 is much better than for the basic pure–tone audiometry. Therefore, the equal–loudness test seems to perform better even in an extremely noisy situation. This is possibly due to the test signal, 1/3 octave filtered noise, having a wider spectrum — it does not get masked under background noise as easily, unlike the pure–tone sine test signals of the basic pure–tone audiometry test.

4.2.2 Variation between frequencies

It was also of interest to find out if the results show any frequency dependence. Figure 28 shows the performance difference of the basic pure–tone audiometry test with different test frequencies for all test subjects and the Table 3 shows the mean and the median for all frequencies. Variation is high for 62.5 Hz, 2000 Hz, 4000 Hz and 8000 Hz.

The equal–loudness test performed significantly better in this regard, as can be seen from Figure 30 and the Table 4 which shows the mean and the median for all

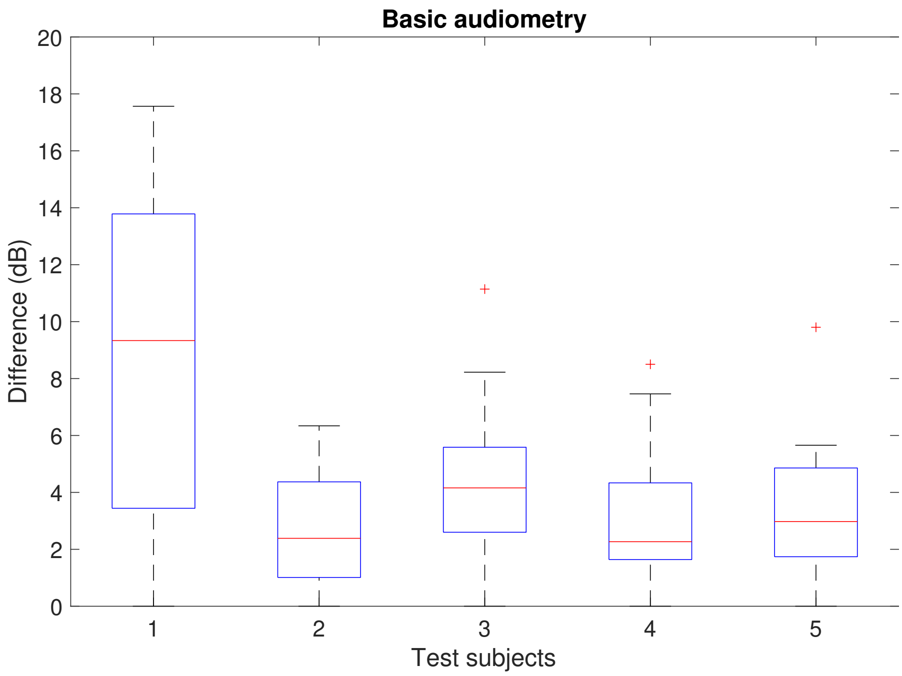

Figure 26: Box plot of the difference in loudness levels between the basic pure–tone audiometry test performed in noise and in a quiet environment for each of the five test subjects.

freq. (Hz)	62.5	125	250	500	1000	2000	4000	8000
avg (dB)	6.2156	3.4722	3.9120	3.6628	0.7639	5.8557	5.8693	6.3787
med (dB)	4.4736	3.0865	3.1831	3.1228	0.1259	3.7876	6.0097	5.8698

Table 3: Mean and median of basic pure–tone audiometry tests. Tests performed in a noisy environment were compared to tests performed in a quiet environment.

frequencies. The variance is a lot smaller, and interestingly, biggest differences seem to be in the 251 Hz and 500 Hz test signals.

Since test subject 1 had performed the test with a significantly louder background noise level, it is worth seeing what happens if the results are removed from the dataset. This is shown in Figures 29 and 31. As can be seen, the variance is reduced somewhat, and the outliers disappear, but the difference is not drastic, as most of the results in loud noise were interpreted as outliers in any case. This demonstrates the usefulness of the box plot for visualizing this type of data.

4.3 Test speed

It was discovered that it takes approximately 2.5 minutes to perform the equal–loudness test for both ears. However, since the testing time is not limited in any

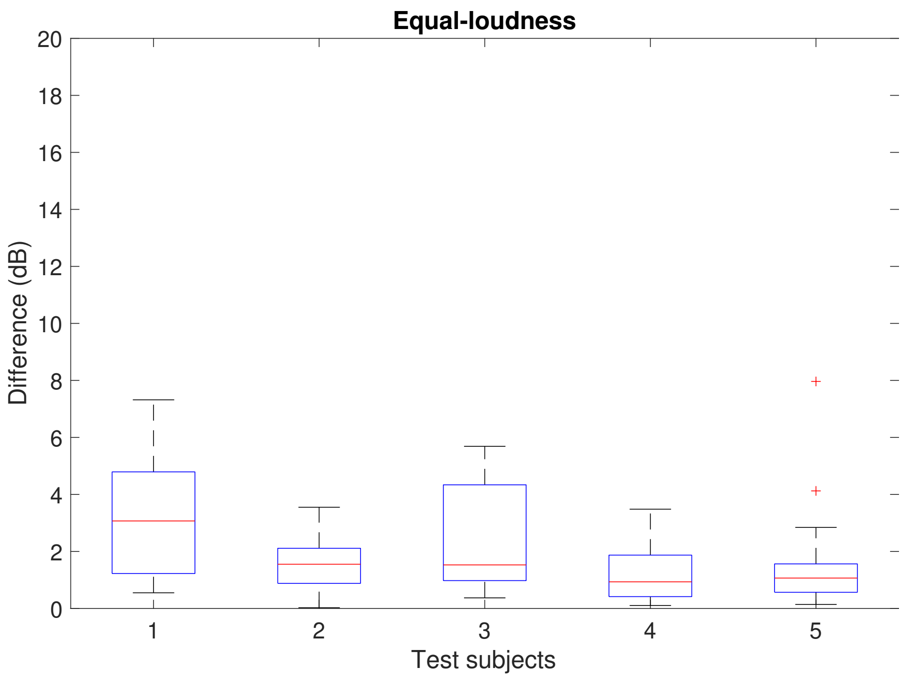

Figure 27: Box plot of the difference in loudness levels between the equal–loudness test performed in noise and in a quiet environment for each of the five test subjects.

freq. (Hz)	63	125	251	500	2000	3981	7943
mean (dB)	1.5254	2.1279	2.7788	3.2632	0.9392	1.7387	2.0391
median (dB)	1.0133	1.2356	2.1941	2.2256	0.5601	1.4566	1.5645

Table 4: Mean and median of equal–loudness tests. Tests performed in a noisy environment were compared to tests performed in a quiet environment.

way, some users might take longer to complete the test. The testing time can be shortened by choosing not to include certain test frequencies.

4.4 Discussion

Since the level of the test signals is set for both ears simultaneously, the possible difference in hearing between different ears on the test frequency is not considered. This could be rectified by adding a final step to the test where the user is asked to center the panning of a signal that has been filtered by an equalizer which has been adjusted according to the obtained equal–loudness curves.

However, when each test subject was asked to use their equal–loudness hearing test results as a basis for the equalization algorithm in the program developed by Vapalahti [50], and to listen to the end result and assess the sound quality, they never

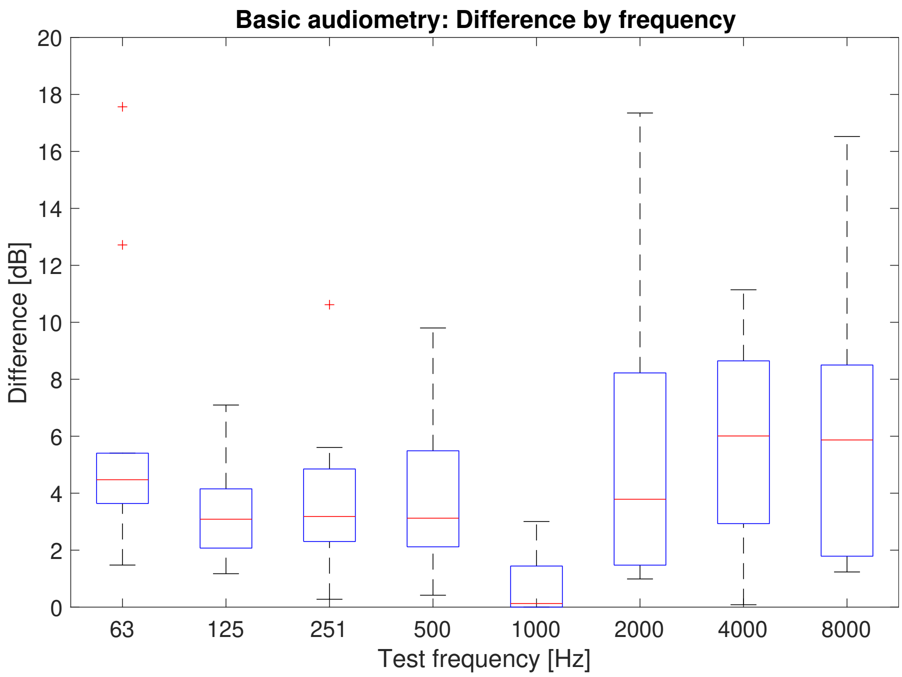

Figure 28: Performance of the basic pure–tone audiometry test on different test frequencies for all five test subjects.

commented that the stereo image produced by the results of the equal–loudness test would have been unnatural. On the contrary, many pointed out that the equalization based on the basic pure–tone test results often produced an unnatural stereo image.

The comparison between measurements in noise and in a quiet environment were not standardized, so noise levels varied between different test subjects and the background noise levels were not measured. This could have been solved by mixing constant background noise (e.g. a recording of traffic or public transport sounds) behind the test signal to mimic a real-life situation. However, performing the test in a real–life environment was considered an advantage, since this represented the situations in which the test was meant to be used.

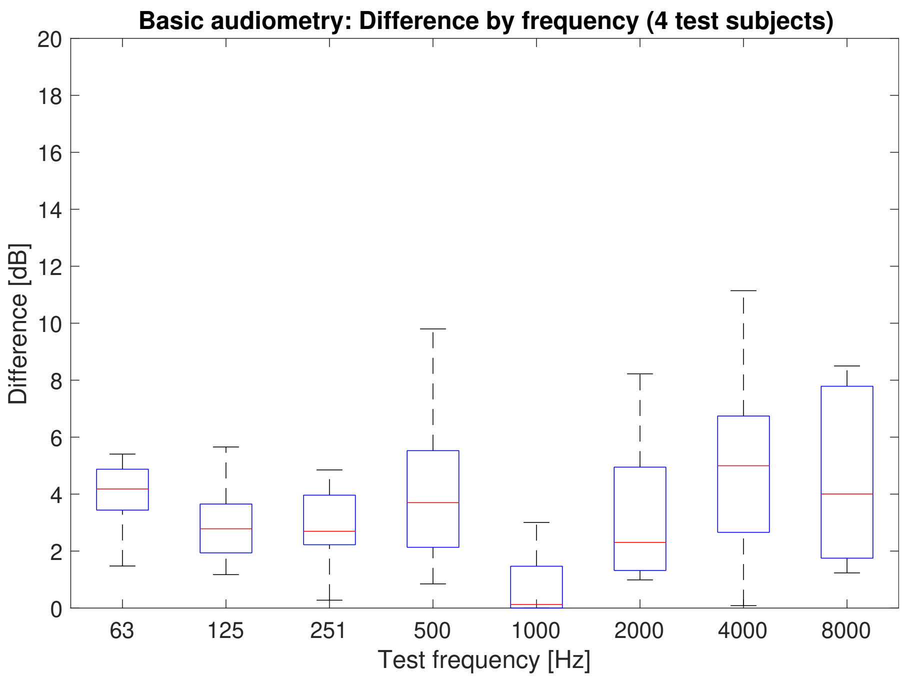

Figure 29: Performance of the basic pure–tone audiometry test on different test frequencies for four test subjects, test subject 1 omitted.

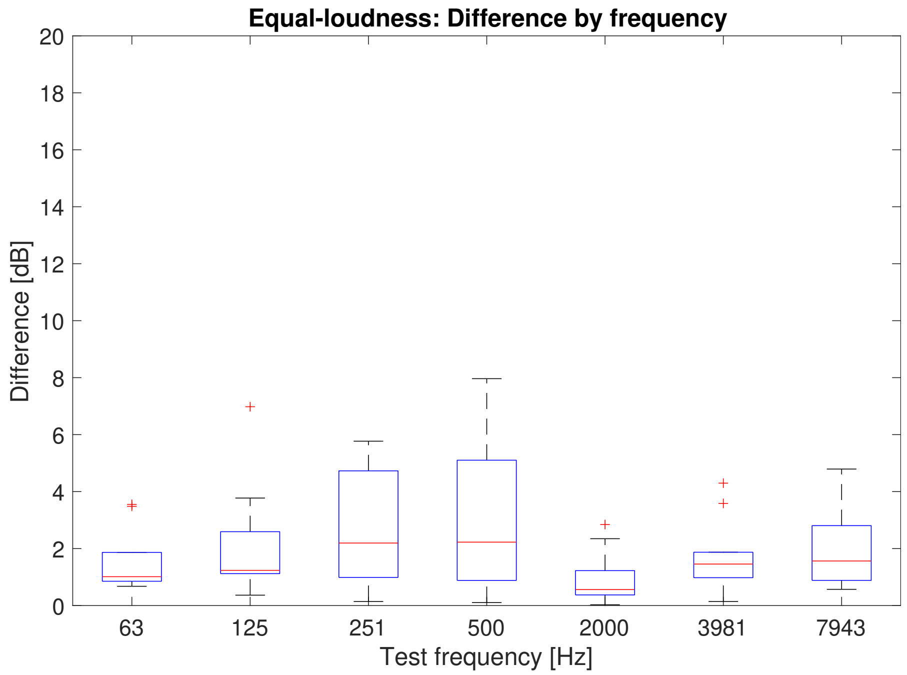

Figure 30: Performance of the equal–loudness test with different test frequencies for all five test subjects.

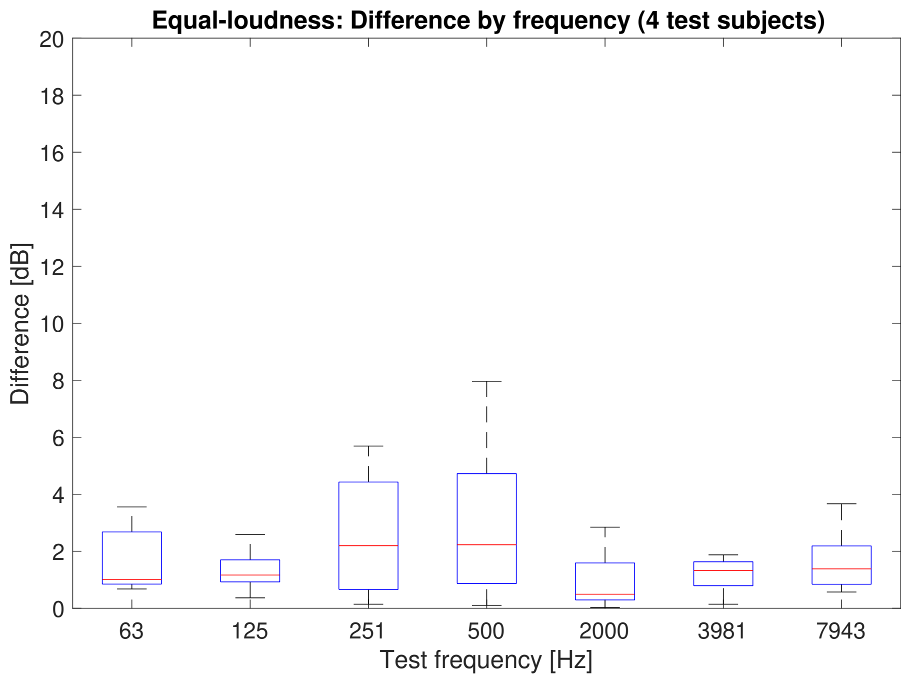

Figure 31: Performance difference of equal–loudness test with different test frequencies for four test subjects, test subject 1 omitted.

5 Conclusions

The aim of this thesis was to describe the implementation of a working prototype of an equal–loudness based hearing sensitivity measurement system for personal sound applications. First, relevant theory and previous research regarding hearing, hearing loss, hearing aids and hearing tests was covered, as well as the relevant theory and previous research concerning headphones.

The design of the equal–loudness program was based on the standard equal–loudness test, which was adapted to a self–administered test format. The structure of the test was explained, and its performance was compared to diagnostic audiometry and a basic pure–tone hearing threshold test, which was tested by five different test subjects.

Relevant results were shown as box plot interpretations of the data that was gathered during the measurements. The main result was that the equal–loudness test performed significantly better in a noisy environment than a basic pure–tone hearing threshold test, and it is therefore a promising candidate for personal sound applications and other mobile hearing tests.

References

[1] D. Griesinger, "Frequency response adaptation in binaural hearing," in *Audio Engineering Society Convention 126*. Audio Engineering Society, 2009.

[2] U. Horbach, "Characterizing the frequency response of headphones—a new paradigm," in *Audio Engineering Society Convention 138*. Audio Engineering Society, 2015.

[3] J. Backman, T. Campbell, J. Kleimola, and M. Hiipakka, "A self-calibrating earphone," *Journal of the Audio Engineering Society*, May 2017.

[4] H. Dillon, *Hearing Aids*, 2nd ed. Turramurra, Australia: Boomerang Press, 2012.

[5] *Personal Sound | Audiodo - Your Unique Hearing*, Audiodo AB, 2020. [Online]. Available: https://www.audiodo.com/personal-sound/

[6] *Nuraphone*, Nura Operations Pty. Ltd., 2020. [Online]. Available: https://www.nuraphone.com/pages/how-it-works

[7] *SoundID | Sonarworks*, Sonarworks Ltd, 2020. [Online]. Available: https://www.sonarworks.com/soundid

[8] *Personalize Your Audio Experience with Skullcandy's App*, Skullcandy Inc., 2020. [Online]. Available: https://www.skullcandy.com/personalize

[9] R. S. Rader, C. Menzel, B. W. Edwards, S. Puria, and B. B. Johansen, "Sound enhancement for mobile phones and other products producing personalized audio for users," U.S. Patent 6 944 474 B2, 2005.

[10] F. Mahomed, D. Swanepoel, R. Eikelboom, and M. Soer, "Validity of automated threshold audiometry: a systematic review and meta-analysis," *Ear and hearing*, vol. 34, no. 6, pp. 745–752, 2013.

[11] J. Saliba, M. Al-Reefi, J. Carriere, N. Verma, C. Provencal, and J. Rappaport, "Accuracy of mobile-based audiometry in the evaluation of hearing loss in quiet and noisy environments," *Otolaryngology–Head and Neck Surgery*, vol. 156, no. 4, pp. 706–711, 2017.

[12] D. Swanepoel, H. Myburgh, D. Howe, F. Mahomed, and R. Eikelboom, "Smartphone hearing screening with integrated quality control and data management," *International journal of audiology*, vol. 53, no. 12, pp. 841–849, 2014.

[13] V. Pulkki and M. Karjalainen, *Communication acoustics : an introduction to speech, audio, and psychoacoustics*, 1st ed. Chichester, West Sussex, United Kingdom: Wiley, 2015.

[14] T. D. Rossing, F. R. Moore, and P. A. Wheeler, *The Science of Sound*, 3rd ed. Reading, MA: Addison-Wesley, 2002.

[15] B. Scharf, "Loudness," *Handbook of perception*, vol. 4, pp. 187–242, 1978.

[16] S. S. Stevens, "The measurement of loudness," *The Journal of the Acoustical Society of America*, vol. 27, no. 5, pp. 815–829, 1955.

[17] H. Fletcher and W. A. Munson, "Loudness, its definition, measurement and calculation," *The Journal of the Acoustical Society of America*, vol. 5, no. 2, pp. 82–108, 1933.

[18] Y. Suzuki and H. Takeshima, "Equal-loudness-level contours for pure tones," *The Journal of the Acoustical Society of America*, vol. 116, no. 2, pp. 918–933, 2004.

[19] *Matlab R2020a Documentation*, The MathWorks, Natick, MA, USA, 2020. [Online]. Available: https://mathworks.com/help/matlab/

[20] B. R. Glasberg and B. C. Moore, "Derivation of auditory filter shapes from notched-noise data," *Hearing research*, vol. 47, no. 1-2, pp. 103–138, 1990.

[21] World Health Organization, "World report on hearing," Geneva, 2021.

[22] J. D. Miller, "Effects of noise on people," *The Journal of the Acoustical Society of America*, vol. 56, no. 3, pp. 729–764, 1974.

[23] U.S. Department of Health & Human Services, Centers for Disease Control and Prevention, "What Noises Cause Hearing Loss?" [Online]. Available: https://www.cdc.gov/nceh/hearing_loss/what_noises_cause_hearing_loss.html

[24] M. Mills, "Hearing aids and the history of electronics miniaturization," *IEEE Annals of the History of Computing*, vol. 33, no. 2, pp. 24–45, 2011.

[25] S. Lybarger, "U.S. Patent Application SN 532, 278," 1944.

[26] J. Jerger, "Bekesy audiometry in analysis of auditory disorders," *Journal of Speech, Language and Hearing Research*, vol. 3, no. 3, pp. 275–287, 1960.

[27] "ISO 8253-1: Acoustics – Audiometric test methods – Part 1: Pure-tone air and bone conduction audiometry," 2020.

[28] X. D. Song, B. M. Wallace, J. R. Gardner, N. M. Ledbetter, K. Q. Weinberger, and D. L. Barbour, "Fast, continuous audiogram estimation using machine learning," *Ear and hearing*, vol. 36, no. 6, p. e326, 2015.

[29] H. Møller, M. F. Sørensen, D. Hammershøi, and C. B. Jensen, "Head-related transfer functions of human subjects," *Journal of the Audio Engineering Society*, vol. 43, no. 5, pp. 300–321, 1995.

[30] D. Hammershøi and H. Møller, "Sound transmission to and within the human ear canal," *The Journal of the Acoustical Society of America*, vol. 100, no. 1, pp. 408–427, 1996.

[31] L. Nyre, "Urban Headphone Listening and the Situational Fit of Music, Radio and Podcasting," *Journal of Radio & Audio Media*, vol. 22, no. 2, pp. 279–298, 2015.

[32] G. Lorho, "Subjective Evaluation of Headphone Target Frequency Responses," *Journal of the Audio Engineering Society*, may 2009.

[33] J. Borwick, "Loudspeaker and headphone handbook," Oxford, 2001.

[34] G. Theile, "On the standardization of the frequency response of high-quality studio headphones," *Journal of the Audio Engineering Society*, vol. 34, no. 12, pp. 956–969, December 1986.

[35] J. Rämö, "Equalization techniques for headphone listening," Ph.D. dissertation, Aalto University, 2014.

[36] F. E. Toole, "The Acoustics and Psychoacoustics of Headphones," *Journal of the Audio Engineering Society*, 1984.

[37] T. Voinier and F. Briolle, "Transfer function and subjective quality of headphones: Part 1, transfer function measurements," *Journal of the Audio Engineering Society*, 1992.

[38] J. Breebaart, "No correlation between headphone frequency response and retail price," *The Journal of the Acoustical Society of America*, vol. 141, no. 6, pp. EL526–EL530, 2017.

[39] J. Mejia, H. Dillon, and M. Fisher, "Active cancellation of occlusion: An electronic vent for hearing aids and hearing protectors," *The Journal of the Acoustical Society of America*, vol. 124, no. 1, pp. 235–240, 2008. [Online]. Available: https://doi.org/10.1121/1.2908279

[40] O. N. Raz and D. Aronson, "Method and device for playing modified audio signals," U.S. Patent 9 680 438 B2, 2017.

[41] R. Swartz, "Listener specific audio reproduction system," U.S. Patent 7 564 979 B2, 2009.

[42] Y. Raz, "System and method for improved audio perception," U.S. Patent 9 943 253 B2, 2018.

[43] *AD28 Service Manual*, Interacoustics, Denmark.

[44] "Acoustics. Reference zero for the calibration of audiometric equipment. Part 1: Reference equivalent threshold sound pressure levels for pure tones and supra-aural earphones (ISO 389-1:2017)," Suomen Standardoimisliitto SFS ry, 2018.

[45] *User Manual: Hand Held Analyzer Types 2250 and 2270*, Brüel & Kjær, Denmark, 2016.

[46] "Product Data and Specifications: Pistonphone Type 42AP," G.R.A.S. Sound & Vibration, Denmark, 2011.

[47] "Product data: Artificial Ear Types 4152 and 4153," Brüel & Kjær, Denmark, 2020.

[48] "IEC 60318-1: Electroacoustics - Simulators of human head and ear - Part 1: Ear simulator for the measurement of supra-aural and circumaural earphones," International Electrotechnical Commission, 2009.

[49] S. K. Mitra, *Digital Signal Processing: A Computer-Based Approach*, 4th ed. New York, NY: McGraw-Hill Higher Education, 2011.

[50] L. Vapalahti, "Audio equalization based on personal hearing data," Master's thesis, Aalto University, 2020.

[51] C. Hummersone, "ISO 226:2003 Normal equal-loudness-level contours." [Online]. Available: https://github.com/IoSR-Surrey/MatlabToolbox

[52] P. L. Michael, "The pure-tone audiometer standard reference zero controversy," *American Association of Industrial Nurses Journal*, vol. 13, no. 4, pp. 7–10, 1965.

[53] J. W. Tukey, *Exploratory Data Analysis.* Addison–Wesley, 1977.

[54] H. Wickham and L. Stryjewski, "40 years of boxplots." [Online]. Available: https://vita.had.co.nz/papers/boxplots.pdf

[55] "ANSI S1.11-2004 standard: Octave-Band And Fractional-Octave-Band Analog And Digital Filters," Acoustical Society of America, 2004.

A Matlab implementation of the test program

The software prototype for testing and development purposes was built with MATLAB, a multi-paradigm programming language and numeric computing environment developed by MathWorks[19].

One of the most important functionalities provided by MATLAB was the Audio Toolbox, which is a set of functions that cover audio processing, speech analysis and acoustic measurements[19]. It includes algorithms for processing and generating audio signals, such as noise signals and different filters, which were very useful for this project.

Some of the algorithms were provided as System objects. MATLAB's System objects are designed specifically for implementing and simulating dynamic systems with inputs that change over time, and they can be used similarly to functions[19]. They are used similarly to functions: system objects are created with optional parameters (e.g. `octFilt = octaveFilter(parameters)`), after which data can be run through the object (e.g. `octFilt(x)`).

Another important functionality of MATLAB was its AppDesigner. The AppDesigner is an application development environment which enables creating graphical user interfaces and compiling MATLAB code into cross–platform applications.

A.1 The generation of the test signal

As mentioned in chapter 3.2.1, the test signals are 1/3 octave bandpass filtered pink noise. They are generated with the help of standard MATLAB functions as well as with algorithms from MATLAB's AudioToolbox library, especially with the `octaveFilter` system object. This code is shown in Listing 1.

First, equally spaced center frequencies are calculated for the eight 1/3 octave filters with the `getANSICenterFrequencies` function, which returns the center frequencies specified by the ANSI S1.11-2004 standard[55]. The standard provides performance requirements for analog, sampled data, and digital implementations of bandpass filters that comprise a filter set or spectrum analyzer for acoustical measurements.

As stated earlier, the octave filters are bandpass filters. The frequencies are created according to an octave filter design, designed with the `octaveFilter` system object. These frequencies are then used as parameters to generate 1/3 octave bandpass filters with the `octaveFilter` constructor that creates the `OctaveFilterBank` System object. It was chosen that filter order $N = 2$ for this octave filter. This is purely because it was the most pleasant sounding, and was preferred in comparison to other filter orders in the listening tests.

Next, a one second long pink noise signal is generated with the `pinknoise` function and normalized, and filtered with the `OctaveFilterBank`, which creates eight test signals. These signals are normalized and then attenuated by 20 dB, leaving room for boosting the test signals without introducing clipping. Since N dBs of attenuation can be described with the equation

$$N = 20 \log x, \tag{A1}$$

the equivalent linear attenuation can be calculated by dividing both sides of the equation by 20, and then raising both sides of equation A1 to the power of 10 and solving for x, or

$$\begin{aligned} \frac{N}{20} &= \log x \\ x &= 10^{\frac{N}{20}}. \end{aligned} \tag{A2}$$

The reference sound, `app.ReferenceSound`, is set to whatever the value of the global variable `app.ReferenceIndex` is — in this case, it refers to the array index containing the 1 kHz signal. A list of gain values, where the values of the Loudness slider will be stored, is also generated.

Finally, the test signals are generated. Since the signals are in stereo, one of the channels is always silent. A silent signal can be generated with the `zeros` command, since it generates a matrix full of zeros. Finally, the binaurally centred comfortable level test signal is generated and stored in the `ComfortableLevelTestSignal` variable.

Listing 1: MATLAB script for generating the test signals.

```
% create octave filtered noise
N = 2; % filter order
F0 = 1000;
Fs = app.Fs;
OctFilter = octaveFilter('FilterOrder', N, 'CenterFrequency', F0, 'Bandwidth', '1
     octave', 'SampleRate', Fs);

% get center frequenciees for filters
F0 = getANSICenterFrequencies(OctFilter);
% remove frequencies below 20 Hz and over 20 kHz
F0(F0<20) = [];
F0(F0>20e3) = [];
Nfc = length(F0);
app.ListOfCenterFrequencies = F0;
app.MaxFrequencies = Nfc;

% generate 1/3 octave filters
for i=1:Nfc
       OctaveFilterBank{i} = octaveFilter('FilterOrder', N, 'CenterFrequency', F0
           (i), 'Bandwidth', '1/3 octave', 'SampleRate', Fs);
end

% generate app.SignalLength seconds of pink noise
x = pinknoise(Fs*app.SignalLength);
x = x/max(abs(x)); % normalize

% filter the white noise
for i=1:Nfc
       OctaveFilteredNoise(:,i)=OctaveFilterBank{i}(x);
end

% normalize and leave enough room for boosting the signal without clipping it
```

```
coeff = (max(max(abs(OctaveFilteredNoise(:,2:8)))))^-1;
OctaveFilteredNoise = 10^(-20/20)*coeff*OctaveFilteredNoise;

app.ReferenceSound = OctaveFilteredNoise(:,app.ReferenceIndex);
app.ListOfGains = zeros(Nfc,2);

app.FrequencyHzEditField.Value = app.ListOfCenterFrequencies(app.StartIndex);

for i = 1:Nfc
       app.AllTestSignals{i,1} = [OctaveFilteredNoise(:,i),zeros(app.Fs*app.
           SignalLength,1)];
end
for i = 1:Nfc
       app.AllTestSignals{i,2} = [zeros(app.Fs*app.SignalLength,1),
           OctaveFilteredNoise(:,i)];
end

ComfortableLevelTestSignal = [OctaveFilteredNoise(:,app.ReferenceIndex),
    OctaveFilteredNoise(:,app.ReferenceIndex)];
```

A.2 Real–time audio in Matlab

Since the sound level is also adjusted in real–time, the audio has to be processed in real–time. How this is done is shown in Listing 2. This code is responsible for playing the test signals, and therefore it has to take the position of the loudness slider into account. This code was adapted partly from [50].

MATLAB can write audio samples to an audio output device by utilizing the `audioDeviceWriter` System object. First, the buffer size is defined as `BUFSIZE`. The `audioDeviceWriter` system object (in this instance called `Output`) is created. The output is initialized with the `setup` command by feeding a matrix full of zeros (in other words, silence) into the `Output`. Then, the desired data is written into the audio output inside a while loop, as long as the flag `app.ExitFlag` is true.

It is necessary to add a small pause inside the while loop to give time for background processes to execute. This can be done with a command `pause(0.001)`. However, MATLABalso has the `drawnow` function, which processes any pending callbacks. This allows e.g. for the variables holding information about which frequency is being tested (`app.TestIndex`) and which channel is being tested (`app.TestChannel`) to be changed via pressing the Next-button while the sounds are being played.

Audio is written into the output one frame at a time. The size of the frame is `BUFSIZE`. Then, the value of the loudness slider, `app.loudness`, adjusts the volume of the signal. As described in equation A2, an attenuation L in decibels can be turned into a linear coefficient k with the equation

$$k = 10^{L/20}. \tag{A3}$$

The command `Output(y)` writes the buffer `y` to the `Output`. After this, the variable `frame` is incremented by one so that during the next iteration, the next frame can be written into the buffer. After this, there is a similar loop section for playing the reference signal at reference level (`app.ComfortableLevel`).

After the loop is over, the line `app.TimerEditField.Value=toc` sets the time indicator to show the time that it took for the user to finish the test. Earlier in the code, when the test starts, there is a `tic` command, which starts a timer in MATLAB, and `toc` ends it. Finally, the `release(Output)` line releases the audio device.

Listing 2: MATLAB script for playing real-time audio.

```
% Button pushed function: StartButton
function StartButtonPushed(app, event)
       %% loop vector to output

       app.ChannelEditField.Value = 1;
       app.FrequencyHzEditField.Value = app.ListOfCenterFrequencies(app.TestIndex
           );

       app.StartButton.Enable = false;
       app.NextButton.Enable = true;
       app.ComfortableleveldBSlider.Enable = false;

       BUFSIZE = 512;

       Output = audioDeviceWriter('SampleRate',app.Fs);
       % initialize the output
       x = [zeros(app.Fs*app.SignalLength, 1),zeros(app.Fs*app.SignalLength, 1)];
       setup(Output,zeros( BUFSIZE, size(x,2) ))
       tic
       while true

           if(app.ExitFlag==true)
               break;
           end

           drawnow %pause

           frame = 0;
           while frame < floor(length(x)/BUFSIZE)-1

               % Read BUFSIZE samples from stere/mono file to Buffer
               y = app.AllTestSignals{app.TestIndex,app.TestChannel}( (1:BUFSIZE)+
                   frame*BUFSIZE , : );

               % Processing
               y = 10^(app.loudness/20)*y;

               % Write buffer to Output
               Output(y);

               frame = frame+1;
               drawnow %pause
           end

           drawnow % pause

           frame = 0;
```

```
        while frame < floor(length(x)/BUFSIZE)-1

            % Read BUFSIZE samples from stere/mono file to Buffer
            y = app.AllTestSignals{app.ReferenceIndex,app.TestChannel}( (1:
                BUFSIZE)+frame*BUFSIZE , : );

            % Processing
            y = 10^(app.ComfortableLevel/20)*y;

                    Output(y);

            frame = frame+1;
            drawnow %pause
        end

    end
    app.TimersEditField.Value=toc;
    % Release
    release(Output)

end
```

A.3 Adjusting parameters in real time

Real time slider adjustment in the Matlab App Designer requires using the `ValueChanging` callback function for the slider UI component. Otherwise, the slider value can is recorded only when the slider control is released. As an example, Listing 3 shows how the this is done for the comfortable level setting (variable `app.ComfortableLevel`). The `Value` property of the `event` object refers to the value of the slider.

Listing 3: Matlab script for real time slider adjustment.

```
% Value changing function: ComfortableleveldBSlider
function ComfortableleveldBSliderValueChanging(app, event)
        changingValue = event.Value;
        app.ComfortableLevel = changingValue;
end
```

A.4 Saving measurement data into a file

Measurement data can be saved into a text file that can then be opened with any application that can open text files. The code for this is shown in Listing 4. This is done by using the `save` function of Matlab. To get a valid file path, the function `uiputfile` is used.

Listing 4: Matlab script for saving a file.

```
% Button pushed function: SaveResultsButton
function SaveResultsButtonPushed(app, event)
        [filename, path] = uiputfile('*.mat','File Name','test.mat');
```

```
        if isequal(filename,0) || isequal(path,0)
                disp('User clicked Cancel.')
                else
                f = app.ListOfCenterFrequencies(app.StartIndex:app.EndIndex);
                L = app.ListOfGains(app.StartIndex:app.EndIndex,1)';
                R = app.ListOfGains(app.StartIndex:app.EndIndex,2)';

                save(fullfile(path,filename),'f', 'R', 'L', '-mat');
        end
end
```

B Flowchart of the equal–loudness test

This flowchart shows the equal–loudness test in a graphical format in Figure B. The test process is explained in more detail in chapter 3.2.3.

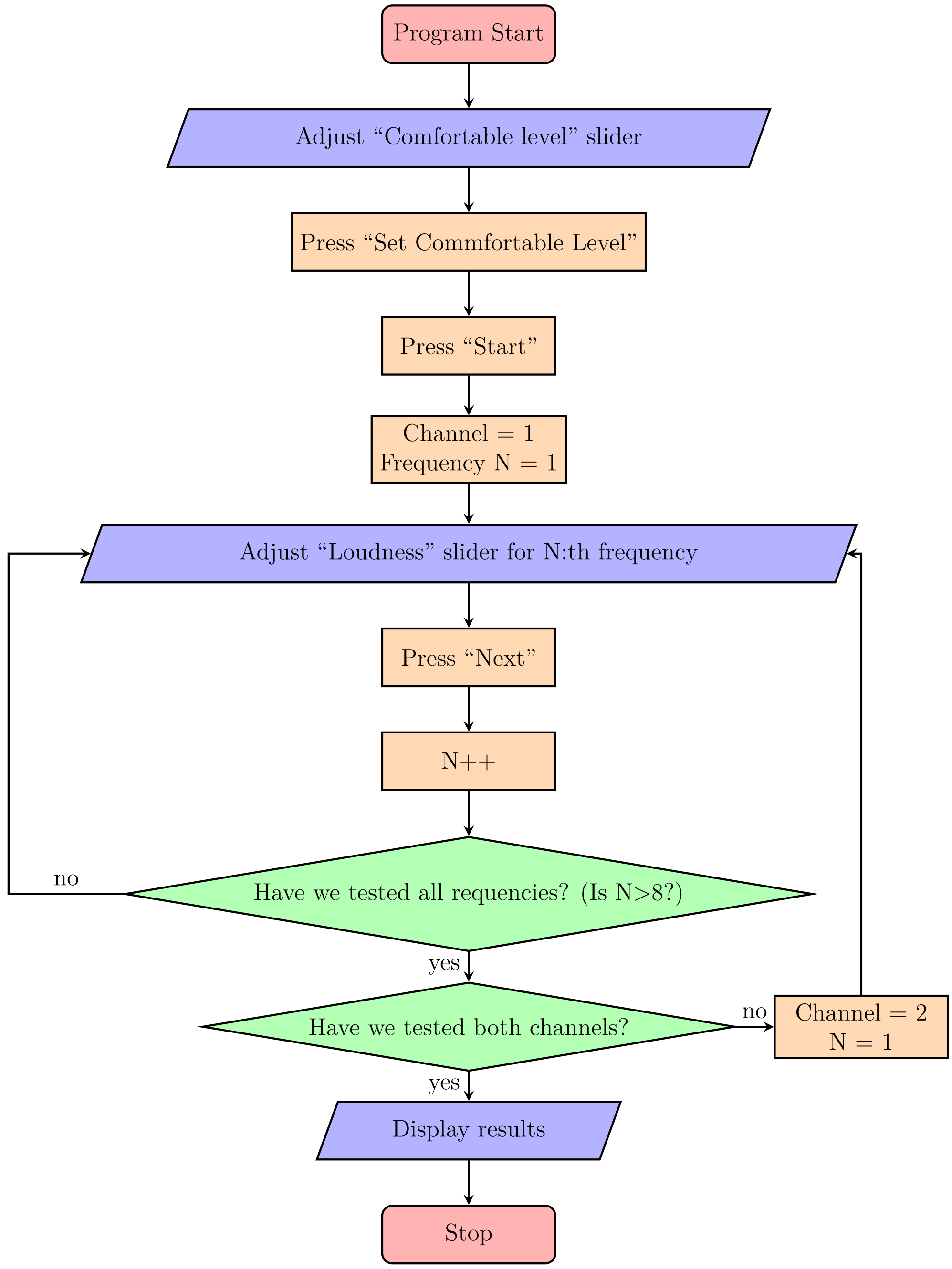

Figure B1: Flowchart of the equal–loudness test.

C Box plot

The box plot is a compact distributional summary of statistics that works for a wide range of data[54]. It consist of five components which give a robust summary of the distribution of a dataset:

- the median
- two hinges, the upper and lower fourths (quartiles),
- the data values adjacent to the upper and lower fences,
- two whiskers that connect the hinges to the fences, and
- outliers, individual points further away from the median than the extremes[54].

The box plot can be compared to a *normal distribution*, which has a median dividing the data in upper and lower halves, the upper and lower quartiles containing 50% of all data, and outliers in the far ends of the distribution.

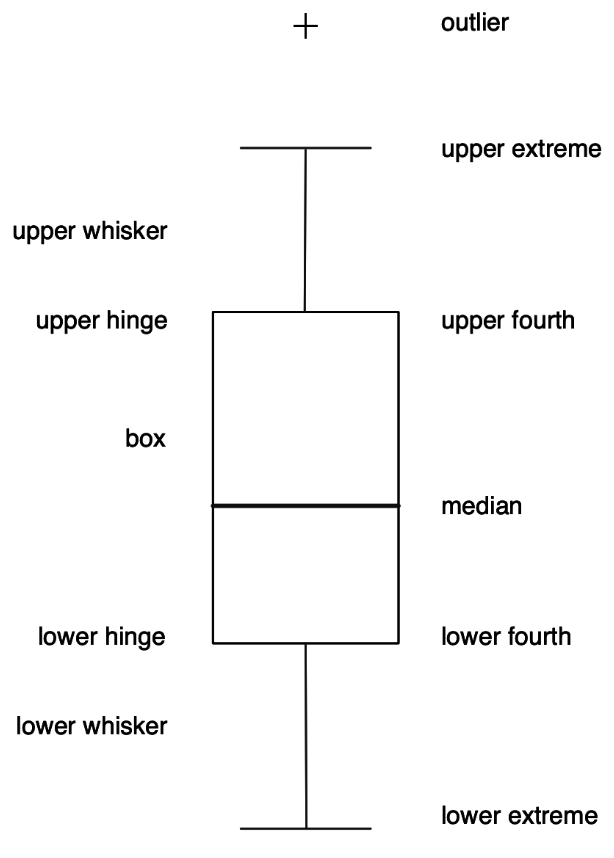

Figure C1: The elements of a box plot. The labels on the left name graphic elements, labels on the right describe the corresponding statistics information. Adapted from [54].