Design of a Multi-Channel Device for Conveying Vibrotactile Information and Vibrotactile Threshold Experiment

Elvar Atli Ævarsson

Faculty of Electrical and Computer Engineering
University of Iceland
2020
DESIGN OF A MULTI-CHANNEL DEVICE FOR CONVEYING VIBROTACTILE INFORMATION AND VIBROTACTILE THRESHOLD EXPERIMENT

Elvar Atli Ævarsson

60 ECTS thesis submitted in partial fulfillment of a Magister Scientiarum degree in Electrical and Computer Engineering

Advisors
Kristinn Andersen
Rúnar Unnþórsson

Faculty Representative
Kristinn Andersen

M.Sc. committee
Kristinn Andersen
Rúnar Unnþórsson

Faculty of Electrical and Computer Engineering
School of Engineering and Natural Sciences
University of Iceland
Reykjavik, October 2020
Design of a Multi-Channel Device for Conveying Vibrotactile Information and Vibrotactile Threshold Experiment
Designing a device for conveying tactile information
60 ECTS thesis submitted in partial fulfillment of a M.Sc. degree in Electrical and Computer Engineering

Copyright © 2020 Elvar Atli Ævarsson
All rights reserved

Faculty of Electrical and Computer Engineering
School of Engineering and Natural Sciences
University of Iceland
Dunhagi 5
107, Reykjavik, Reykjavik
Iceland

Telephone: 525 4000

Bibliographic information:

Printing: Háskólaprent, Fálkagata 2, 107 Reykjavik
Reykjavik, Iceland, October 2020
Abstract

Listening to music is an experience that can generate great pleasure and enhance the mood. While some people view music listening as a pivotal part in their everyday existence, others are not so fortunate as to be able to enjoy music due to severe hearing impairment. The cochlear implant is a great technological and biomedical achievement that has restored hearing in hundreds of thousands of deaf and hard-of-hearing individuals over the past decades. However, the limited spectral resolution that the implant delivers to the brain results in many recipients describing their music listening experience as disappointing.

This thesis describes work done for the ACUTE: ACoUstic and Tactile Engineering group at the University of Iceland, as part of their long-term objective of augmenting the musical enjoyment of cochlear implant recipients by applying an additional information stream to the skin. A multi-channel hardware system for high definition vibrotactile feedback was designed and built. It is intended for in-lab psychophysical experiments in the development of a multi-channel vibrotactile display, to be used alongside musical playback. A vibrotactile threshold experiment was designed and conducted, the purpose of which was to map the vibrotactile threshold of the wrist using the newly designed hardware system and the parallel vibration actuators of choice. Results were compared to older studies and it was found that when vibration is applied parallel to the skin, the perceived frequency is twice that of the signal frequency.
Útdráttur


Í þessari ritgerð verður vinnuframlagi til ACUTE: ACoUstic and Tactile Engineering hópsins við Háskóla Íslands gerð skil. Langtímanmarkmið hópsins er að bæta tónlistaruupplifun fólks sem hlotið hefur kuðungsígræðslu með því að veita viðbótar upplýsingastreymi í gegnum húð. Fjölrása vélbúnaður var hannaður og smíðaður. Búnaðurinn er ætlaður til sáleðlisfæðilegra tilrauna á rannsóknarstofu og mun vera notaður í próun á fjölrása titringsskjá (e. vibrotactile display) sem ætlaður er til notkunar með afspilun tónlistar. Athugun á titringsnæmti samsíðsins var framkvæmd með búnaðinum og titringsvaka sem gefur samþéða titring á húð. Niðurstöður voru bornar saman við lýrri rannsóknir og sýna að því að veita samþéða titring á húð, þá er skynjuð tíðni tvöföld tíðni merkisins.
Contents

List of Figures ix
List of Tables xi
Abbreviations xiii
Acknowledgments xv

1 Introduction 1

2 Background 3
  2.1 The ear .......................................................... 3
    2.1.1 Inner-, middle- and outer ear .......................... 3
    2.1.2 The cochlea .............................................. 4
    2.1.3 Hearing and hearing loss ............................... 6
  2.2 The cochlear implant ......................................... 7
    2.2.1 Historical overview ..................................... 7
    2.2.2 How CI works ........................................... 8
    2.2.3 Benefits and limitations ............................... 9
    2.2.4 CI and music perception ............................... 10
    2.2.5 CI and music listening ................................ 12
  2.3 The somatosensory system ................................... 13
    2.3.1 Skin anatomy ............................................ 13
    2.3.2 Sensory substitution .................................... 15
  2.4 Vibrotactile feedback ........................................ 16
    2.4.1 Wearable subwoofers .................................. 16
    2.4.2 Tactile displays ........................................ 16
    2.4.3 The ACUTE group ....................................... 18

3 Design 21
  3.1 The design problem and project objective .................. 21
  3.2 Actuator ...................................................... 23
    3.2.1 Vibrotactile actuators ................................. 24
    3.2.2 L5 ....................................................... 28
  3.3 Hardware system overview .................................. 30
## Contents

3.4 Audio interface .................................................. 30  
3.4.1 MADI .......................................................... 31  
3.4.2 RME MADIface XT .......................................... 31  
3.5 Digital-to-analog converter ..................................... 33  
3.5.1 Ferrofish A32 ................................................ 34  
3.6 Amplifier .......................................................... 34  
3.6.1 RC filters .................................................... 35  
3.6.2 Pulse-width modulation ..................................... 36  
3.6.3 Boost converters ............................................ 37  
3.6.4 Amplifier classes ........................................... 39  
3.6.5 TPA2025D1 .................................................. 42  
3.6.6 KiCad ........................................................ 44  
3.6.7 Amplifier design ............................................. 45  
3.7 Signal flow ........................................................ 51  

4 Vibrotactile Threshold Experiment .................................. 53  
4.1 Introduction ...................................................... 53  
4.2 Fundaments ....................................................... 54  
4.2.1 Decibel ....................................................... 54  
4.2.2 Window functions .......................................... 54  
4.2.3 Absolute threshold ......................................... 55  
4.2.4 Maximum likelihood procedure ........................... 57  
4.2.5 Masking ..................................................... 58  
4.3 Method ............................................................ 59  
4.4 Results ........................................................... 62  
4.5 Discussion ....................................................... 64  

5 Conclusion and Future Work ......................................... 67  

Bibliography .......................................................... 69
List of Figures

2.1 Diagram of the human ear ........................................... 4
2.2 The uncurled cochlea and basilar membrane ...................... 5
2.3 Auditory filters .......................................................... 6
2.4 Diagram of the cochlear implant ..................................... 8
2.5 CI signal processing ..................................................... 9
2.6 Results from a survey on the musical habits of CI recipients . . 12
2.7 Mechanoreceptors of the skin ....................................... 14
2.8 Vibrotactile threshold at the fingertip ............................... 15
2.9 The SoV Tactile Belt .................................................... 18

3.1 Black box depiction of the multi-channel hardware system ...... 22
3.2 Electromagnets ........................................................... 26
3.3 L5 voice coil actuator .................................................. 28
3.4 Acceleration response of the L5 actuator ......................... 29
3.5 RME MADIface XT rear view ....................................... 32
3.6 TotalMix GUI layout .................................................... 32
3.7 Ferrofish A32 AD/DA converter rear view ....................... 34
3.8 Pulse-width modulation ................................................. 37
LIST OF FIGURES

3.9 Boost converter ................................................. 38
3.10 Block diagram of a class D amplifier ......................... 42
3.11 TPA2025D1 block diagram .................................... 43
3.12 TPA2025D1 current consumption .............................. 44
3.13 Multi-channel amplifier rear- and front view .................. 46
3.14 Amplifier channel schematic .................................... 47
3.15 Amplifier channel PCB layout ................................... 49
3.16 Multi-channel amplifier insides .................................. 50
3.17 Signal flowchart .................................................. 52
4.1 Tukey Window ..................................................... 56
4.2 Tukey window spectral leakage ..................................... 56
4.3 Psychometric function .............................................. 58
4.4 Wristband prototype ................................................. 60
4.5 Experimental setup ................................................ 61
4.6 Experimental actuator placements ................................. 61
4.7 Average vibrotactile threshold at the wrist ...................... 62
4.8 Threshold standard deviation ..................................... 63
4.9 Signal frequency vs. acceleration ................................. 64
List of Tables

2.1 The different modalities and optimal frequencies of the four major mechanoreceptors .............................................. 14

3.1 Comparing the viability of the main actuator alternatives for vibrotactile feedback ................................................. 28

3.2 Comparing the performances of the main amplifier classes ............... 42
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>Alternating Current</td>
</tr>
<tr>
<td>AD/DA</td>
<td>Analog-to-Digital/Digital-to-Analog</td>
</tr>
<tr>
<td>ACUTE</td>
<td>ACoUstic and Tactile Engineering</td>
</tr>
<tr>
<td>AES</td>
<td>Audio Engineering Society</td>
</tr>
<tr>
<td>AGC</td>
<td>Automatic Gain Control</td>
</tr>
<tr>
<td>ASIO</td>
<td>Audio Stream Input/Output</td>
</tr>
<tr>
<td>APT</td>
<td>Auto Pass Through</td>
</tr>
<tr>
<td>BGA</td>
<td>Ball grid array</td>
</tr>
<tr>
<td>BM</td>
<td>Basilar Membrane</td>
</tr>
<tr>
<td>CI</td>
<td>Cochlear Implant</td>
</tr>
<tr>
<td>DAC</td>
<td>Digital-to-Analog Converter</td>
</tr>
<tr>
<td>DC</td>
<td>Direct Current</td>
</tr>
<tr>
<td>DSP</td>
<td>Digital Signal Processor</td>
</tr>
<tr>
<td>DTU</td>
<td>Danmarks Tekniske Universitet / Technical University of Denmark</td>
</tr>
<tr>
<td>EDA</td>
<td>Electronic Design Automation</td>
</tr>
<tr>
<td>ERM</td>
<td>Eccentric Rotating Mass</td>
</tr>
<tr>
<td>FFT</td>
<td>Fast Fourier Transform</td>
</tr>
</tbody>
</table>
**LIST OF TABLES**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPF</td>
<td>High-Pass Filter</td>
</tr>
<tr>
<td>IC</td>
<td>Integrated Circuit</td>
</tr>
<tr>
<td>IHC</td>
<td>Inner Hair Cell</td>
</tr>
<tr>
<td>I/O</td>
<td>Input/Output</td>
</tr>
<tr>
<td>LPF</td>
<td>Low-Pass Filter</td>
</tr>
<tr>
<td>MADI</td>
<td>Multi-channel Audio Digital Interface</td>
</tr>
<tr>
<td>MHC</td>
<td>Model Human Cochlea</td>
</tr>
<tr>
<td>MLP</td>
<td>Maximum Likelihood Procedure</td>
</tr>
<tr>
<td>NH</td>
<td>Normal Hearing</td>
</tr>
<tr>
<td>OHC</td>
<td>Outer Hair Cell</td>
</tr>
<tr>
<td>OS</td>
<td>Operating System</td>
</tr>
<tr>
<td>PC</td>
<td>Personal Computer</td>
</tr>
<tr>
<td>PCB</td>
<td>Printed Circuit Board</td>
</tr>
<tr>
<td>PWM</td>
<td>Pulse-Width Modulator</td>
</tr>
<tr>
<td>REA</td>
<td>Rotary Electromagnetic Actuator</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency</td>
</tr>
<tr>
<td>SoV</td>
<td>Sound of Vision</td>
</tr>
<tr>
<td>THD</td>
<td>Total Harmonic Distortion</td>
</tr>
<tr>
<td>VCA</td>
<td>Voice Coil Actuator</td>
</tr>
<tr>
<td>VR</td>
<td>Virtual Reality</td>
</tr>
</tbody>
</table>
Acknowledgments

The work described in this thesis was funded by RANNÍS (Technology Development Fund grant no. 176713-0612).

I would like to thank my advisors Rūnar Unnþórsson, for providing me the opportunity to work on a project that I deeply care about, and Kristinn Andersen, both of whom have assisted me in any way they can. ACUTE team members Árni Kristjánsson, for his insight in perceptual matters, and Pórhildur Ásgeirsdóttir, whose contribution to the experimental process was essential.

My wife Rut Guðnadóttir, my son Frosti and my daughter Svala, for inspiration and encouragement, and for all the sacrifices made over the past five years and past few months especially, allowing me the time I needed. My mother Steinunn Lindbergsdóttir for her endless support.

Ben and Andrew at Lofelt, for taking the time to provide thorough answers to my questions, in friendly communication. Helgi Hafnar Gestsson, for offering his debugging advice, and likewise Sveinn Kjartansson and Kristinn Sturluson at Stúdíó Sýrland. Jeremy Marozeau at DTU for sharing his knowledge on the musical habits of cochlear implant recipients.
1 Introduction

Cochlear implants (CI) have improved the quality of life of hundreds of thousands of people with severe hearing impairment. Being able to hear and understand speech is an untold luxury to people who have spent years or decades in silence. In a place coding strategy that mimics the behaviour of the basilar membrane (BM) inside the cochlea, CI bypasses the faulty inner ear and directly triggers auditory nerve ends by applying electrical pulses via implanted electrodes. With the vast improvements in both technology and surgical techniques over the past few decades, CI is gradually improving. However, it is still a long way off from providing normal hearing. While the implant has shown great results in terms of one-on-one communication in quiet environments, complex sounds such as noisy environments and music still suffer greatly.

Concurrently, the use of tactile displays has been increasing in both entertainment and assistive technology. Tactile displays can be used to replicate texture, shape or other tactile parameters of an item, but can also be used as an additional information stream to compensate for a loss of sense such as sight or hearing. A recent example of the latter is the Tactile Belt developed as part of the Sound of Vision project (SoV), where a high resolution tactile display conveys information on the surrounding environment to blind individuals, by mapping images from a 3D camera to localized vibrations on the display. Building on the idea of the Tactile Belt, a new international joint effort seeks to investigate if the musical features lost to CI recipients can be reproduced by high resolution tactile displays.

The work described in this thesis is part of the ongoing project introduced above, where the long term objective is to enhance the music enjoyment of CI recipients by circumventing the limitations of the implant with supplemental vibrotactile information stream. A multi-channel device for conveying this information was designed and built. Although the primary focus is to convey music, and each design decision is made in that regard, the system can be used for any kind of vibrotactile information stream from computer to user. A vibrotactile threshold experiment was implemented and conducted using the newly designed device. The experiment provided some results and information that are crucial for the development of multi-channel vibrotactile encoding of music, for software to accompany the device.
1 Introduction

In chapter 2, the project background and the fundamental concepts are introduced. Chapter 3 outlines the design problem and project objective, as well as describing the hardware design. The experimental process is then covered in chapter 4, followed by general discussion on the project conclusion and future work in chapter 5.
2 Background

In this chapter the matters that are crucial for understanding the problem at hand are introduced. It begins with a brief discussion on the anatomy of the human ear and how the cochlea in the inner ear responds to sounds of different frequencies. The cochlear implant will then be discussed and how it restores hearing but also distorts the perception of music. Moving on to the somatosensory system, the skin’s mechanoreceptors are considered and how they respond to vibrations. The chapter will then focus on tactile displays and how vibrotactile feedback can be used as sensory substitution. A few devices and projects dealing with vibrotactile feedback will be looked into, concluding with the Sound of Vision (SoV) project and how the vibrotactile part of the SoV project has been continued by the ACUTE (ACoUstic and Tactile Engineering) group, for which this work is done.

2.1 The ear

In mammals, the ears are the brain’s sensors for monitoring acoustic pressure fluctuations in the environment. The following sections will go through the basic function of the human ear and the act of hearing, with the operation of the cochlea and the case of sensorineural hearing loss being of special significance.

2.1.1 Inner-, middle- and outer ear

The human ear, illustrated in Fig. 2.1, can be divided into four parts; the outer ear, the middle ear, the inner ear and the auditory nerve that connects to the brain. The outer ear consists of the pinna and the ear canal. The special shape of the pinna results in diffraction and reflection of the sound so the signal that reaches the ear canal will be direction dependent. In that way, the pinna plays an integral part in sound localization.

At the end of the ear canal is the tympanic membrane (eardrum). Attached to the tympanic membrane is a chain of three bones, hammer (L. Malleus), anvil (L.
2 Background

Fig. 2.1. Diagram of the human ear. The outer ear consists of the pinna and the ear canal. The tympanic membrane (ear drum) makes the connection into the middle ear. The middle ear consists of the ossicular chain (Malleus, Incus, Stapes). The oval window connects to the inner ear, which consists of the semicircular canals and the cochlea [1].

Incus) and stirrup (L. Stapes). These are the smallest bones in the human body and together they form the ossicular chain of the middle ear. The footplate of the stirrup makes the connection between the middle ear and the inner ear, referred to as the oval window.

The inner ear consists of the semicircular canals (related to the sense of balance) and the cochlea (related to the sense of hearing). While the outer- and middle ears are filled with air, the cochlea in the inner ear is filled with lymph fluid. The middle ear thus serves the purpose of impedance matching between the air filled outer ear and fluid filled inner ear. This is achieved by means of a lever function of the ossicular chain together with the area ratio between the tympanic membrane and the oval window. Without this function, the majority of the acoustic energy would be reflected back off the boundary between the air and liquid.

2.1.2 The cochlea

The cochlea is a spiral-shaped bone, approximately 2.75 turns and 32 mm in length from base to apex. It is divided into three channels (L. scala), scala vestibuli, scala tympani and scala media. Fig. 2.2 illustrates this with a hypothetical uncurled cochlea. The oval window connects to the scala vestibuli while another connection between the middle- and inner ears, the round window, connects to the scala tym-
2.1 The ear

Fig. 2.2. Top: Side view of an uncurled cochlea. The scala media and scala tympani are separated by the basilar membrane. Bottom: Top view diagram of the BM and its special structure. The membrane is narrow and stiff at the base and gets gradually wider and less stiff, resulting in varying resonance frequencies along its length [2].

pani. Since the fluid in the cochlea is incompressible, an inward movement of the oval window will result in a corresponding outward movement of the round window to equalize the pressure.

The two channels, scala vestibuli and scala tympani can be thought of as two tubes that lead from the base to the apex. They are connected at the top through a hole, the helicotrema, but otherwise separated by the third channel, scala media, which is covered by a membranous wall. Scala vestibuli and scala media are separated by Reissner's membrane while scala tympani and scala media are separated by the basilar membrane (BM). Inside the scala media, on top of the BM, is the organ of Corti which contains approx. 3500 inner hair cells (IHCs) and 12000 outer hair cells (OHCs) that connect to nerve fibers leading to the brain. The IHCs are the main sensory cells while the OHCs are amplification cells.

**Basilar membrane and auditory filters**

The BM is very important for the function of hearing and essentially determines the way we perceive sound. It varies in width and stiffness along its length, being narrowest and stiffest at the base of the cochlea but widest and least stiff at the apex. Its mechanical characteristics is what ultimately determines the range of
2 Background

Fig. 2.3. An illustration of how a signal containing frequencies 400 Hz, 1.6 kHz and 6.4 kHz (a) will resonate at three different locations along the BM (b). High frequencies resonate near the base and low frequencies near the apex [3].

Audible frequencies, about 20 Hz to 20 kHz. The special shape and structure also results in different resonance frequencies at different points along the membrane, which in turn leads to varying sensitivity to frequency ranges.

When exposed to a pure tone, the BM will vibrate with a certain pattern at a certain location. If the frequency of the pure tone is changed, the vibration pattern will stay the same but move to a different point on the membrane. Fig. 2.3 illustrates the case of a signal containing three pure tones. High frequencies will generate a response close to the base, where the membrane is narrower and stiffer, while low frequencies resonate near the apex, where the membrane is wider and looser. This function can be thought of as resulting from filter operation by an array of band-pass filters applied along the length of the membrane, known as the auditory filters. These filters are overlapping, level-dependent and non-linear, with decreasing bandwidth from base to apex. The bandwidth is called critical bandwidth and is closely related to the phenomenon of auditory masking, further explained later in section 4.2.5.

2.1.3 Hearing and hearing loss

Normally, hearing occurs when acoustic waves enter the ear canal and produce vibrations of the tympanic membrane. The middle ear ensures a corresponding wave propagation through the lymph fluid in the cochlea. This causes the BM to vibrate at specific locations, depending on the frequency content of the signal, which affects the IHCs similarly placed in the organ of Corti. For low-level signals, the OHCs will amplify the vibration of the BM so that it triggers the IHCs. The hairs on top of the inner cells will then bend back and forth and produce spatiotemporal nerve impulses which the brain receives and interprets as sound.
Hearing loss is generally sorted into three categories; conductive-, sensorineural- and mixed hearing loss. Conductive hearing loss can be a temporary or permanent condition, depending on the cause. In this case the conduction of the acoustic signal to the inner ear is partially or fully blocked due to obstructions or damage in the outer- or middle ear. Sensorineural hearing loss however is always permanent. It is caused by damage to either hair cells or the auditory nerve which results in weak or no signal reaching the brain. In the case of OHCs being damaged, which is often the cause of age-related hearing loss, hearing aids can often provide enough amplification to the BM oscillation so that it will activate the auditory nerve. However, if the IHCs are damaged, normal hearing aids are unable to restore hearing as amplification is of little use if the sensor is dead. Fortunately, people who suffer from this condition can be considered candidates for cochlear implants (CI).

2.2 The cochlear implant

CI provides a way to bypass the eardrum, ossicular chain, BM and IHC and stimulate the auditory nerve directly. It is the only technological solution to date that can restore sensorineural hearing loss. The basic structure is outlined in figure 2.4. The device is composed of two parts; an outer part (containing a microphone, digital signal processor (DSP) and transmitter) and a surgically implanted inner part (containing a receiver and an electrode array). The receiver is aligned with the transmitter on the other side of the skull by a pair of magnets while the electrode array is implanted inside the cochlea.

2.2.1 Historical overview

The history of CI can be traced back to 1957 when French doctors André Djourno and Charles Eyries implanted a device, designed to produce the sensation of hearing by stimulation of the auditory nerve with electricity, into a deaf patient of theirs [5]. Initially a single channel device with only one pair of electrodes, it was unable to stimulate different nerve fibers along the cochlea with different information and was of limited use as such. Further experiments with single channel devices were carried out by various parties throughout the 1960s and ’70s.

In the late ’70s, two separate teams led by Ingeborg Hochmair (Austria) and Graeme M. Clarke (Australia) began experimenting with multi-channel devices [6]. The addition of multiple channels and place coding strategies allowed for different frequencies activating different electrodes along the cochlea and thus simulating the auditory-filter function of the BM. The devices developed by Hochmair and Clarke were first
2 Background

Fig. 2.4. Diagram of the CI. Outer part consists of microphone, processor and transmitter. Inner part consists of a receiver and an electrode array [4].

implanted into patients in 1977 and 1978, respectively. These early multi-channel devices vastly improved speech perception compared to single-channel devices and became blueprints for future iterations of the cochlear implant technology.

Currently, there are six manufacturers of cochlear implant devices worldwide, with Cochlear Ltd. from Australia being the largest, with over 379,000 registered implants at the time of writing [7]. This company was launched in 1981 to handle mass-production and further development of the Nucleus Implant device designed by dr. Graeme M. Clarke. Another company, MED-EL, was formed in Austria around Hochmair’s work. More companies have been designing and manufacturing their own implant devices in later years, namely Advanced Bionics (USA), Oticon Medical (Denmark), Microson (Spain) and Nurotron (China).

2.2.2 How CI works

Despite several systems available from different manufacturers, the basic function of all CI devices is generally the same. A microphone is placed at the outer ear to pick up acoustic signals from the environment. A DSP filters the signal into 12-22 frequency bands, depending on the manufacturer, with the output of each frequency band dedicated to a single electrode in the array inside the cochlea. High frequency bands are allocated to electrodes near the base and low frequency bands to electrodes near the apex. The information from the DSP is transmitted through the skull via radio frequencies (RF). The implanted receiver decodes the signal and converts to electrical pulses to be delivered to the electrode array. The electrodes thus trigger specific nerve fibers depending on the frequency content of the incoming
Fig. 2.5. The basic functioning of a CI DSP. The input signal is filtered via band-pass filters. The envelope of each filter output is calculated and used to modulate a square wave pulse-train to be transmitted to the electrode array [5].

acoustic signal.

CI signal processing

Fig. 2.5 shows the basic processing algorithm with a simple example of a 4-channel device processing the speech sound "sa" [5]. After filtering the input signal to four frequency bands, the DSP extracts the envelope of the signal at the output of each band-pass filter. By integrating the area under the envelope curves, the amount of energy at each frequency band can be estimated for each point in time. In that way, each filter envelope is used to modulate the amplitude of a train of square wave pulses at a fixed frequency. The modulated pulse-train is then transmitted to the electrode assigned to the given frequency band.

2.2.3 Benefits and limitations

As with most signals, sound can be described in terms of both its temporal- and spectral aspects. After receiving the implant, CI recipients gain good perception
of the temporal elements of sound. This provides many profits in terms of speech recognition. For example, the sound duration helps with vowel identification (with words like "had" typically longer than "hid") while gaps, or stops, help with identifying the plosives (p, b, k, t). However, many factors contribute to a rather poor frequency perception. First and foremost, the maximum amount of 22 electrodes will never match the resolution provided by 3500 IHC. In addition, the relatively large distance between the electrodes and the auditory nerve results in a current spread inside the cochlea. This in turn results in an unwanted overlap between zones, with neighboring electrodes partially stimulating the same group of nerve fibers.

Despite these drawbacks, CI has done wonders for many people, especially in terms of speech perception. Even with the low spectral resolution, the qualities are good enough for the device to successfully mimic crucial elements of speech, such as the spectral peaks known as formants which characterize the vowels. Research has shown that the perception of speech improves significantly with a higher number of electrodes [8]. On average, adult users of the 22-channel Nucleus 24 device are about 78% correct in identifying words in sentences [9]. However, this is only the case for spoken words in a quiet environment, not for more complex acoustic signals, such as speech-in-noise or music.

2.2.4 CI and music perception

Music can be defined as an ensemble of audible acoustic frequencies carefully organized over a period of time. As described in section 2.2.2, the main processing strategy of CI is based on extracting the envelope while removing the fine structure of the incoming signal. This has the effect that some important spectral information has already been removed from the complex music signal when entering the cochlea. Further emphasised by the limited spectral accuracy that the electrode array delivers, this results in a highly distorted perception of music. Some of the key features of music and how they are perceptually affected by CI will be outlined in the following sections.

Pitch

Pitch is directly related to the frequency of vibration and is a measure of how high or low a given note is perceived by the listener. With the limited number of electrodes that modern CI technology allows, pitch perception is dramatically reduced. On average, CI recipients need a 25% increase or decrease in pitch to be able to identify the direction of pitch change [10]. To put this into perspective, the relative pitch percentage between the notes A and C# (a major third in the 12 tone system)
is +25.99%. This suggests that CI users will have considerably more difficulties recognizing melodies and discriminating between pitches than normal hearing (NH) listeners, as confirmed by studies [11].

Harmony

Closely related to the pitch phenomenon is harmony. Harmony relates to the relationship of multiple pitches presented simultaneously, i.e. chords or harmonized melodies. Weak pitch perception greatly affects the perception of harmony. Harmonic sounds can be categorized as either consonant or dissonant, with consonant sounds generally described by NH listeners as pleasant while dissonance is normally associated with unpleasantness. However, CI recipients show little to no distinction between the two [12] with consonant and dissonant chords rated on average as equally pleasant.

Timbre

Another important feature in music is timbre. Sometimes also referred to as tone quality, it is essentially what differentiates how we perceive tones of the same fundamental frequency but from separate sound sources. For instance, a violin and a piano, both playing a middle C, will not sound the same to a NH listener because of the different timbres of the instruments. While pitch is characterized by the fundamental frequency, timbre is largely characterized by the harmonics (integer multiples of the fundamental frequency) as well as the temporal envelope. The harmonic content describes perceptual features such as brightness, richness and warmth while the temporal content describes the impulsiveness, attack-time and decay-time. CI users generally show limited ability in perceiving timbre [13] although they are able to distinguish impulsive instruments from non-impulsive instruments [14].

Rhythm

Rhythm in music can be defined as either periodic or aperiodic temporal patterns of strong and weak elements that relate to the forward movement of the music in time. Comprised of tempo, beats, pulses, meters, accents and other elements, rhythm can be straight and steady, highly complex and everything in between. In the case of simple rhythmnical tasks such as pattern reproduction, CI users perform at a similar level as NH listeners [15]. This is a positive thing and should perhaps be expected since the temporal amplitudes are preserved in CI coding. However, it has been argued that while CI users do well at those simple tasks, real music often contains a
2 Background

![Bar charts showing musical habits](image)

Fig. 2.6. Results from a survey on the musical habits of more than 100 CI recipients, conducted by Looi and She [17]. Top: CI users spend more time listening to music than they did just before receiving the implant, but less time than before suffering from hearing loss. Bottom: CI users describe more enjoyment from listening to music than just before receiving the implant, but less enjoyment than before suffering from hearing loss.

much more complex rhythmical structure. Perceiving rhythm in music often relies on the segregation of multiple audio streams, an ability which is greatly affected by the CI coding strategy. How CI users perceive polyphonic rhythms and "groove" remains unclear [16].

2.2.5 CI and music listening

The previous sections list a number of reasons for one to assume that people with CIs do not listen to music at all. However, that’s not entirely the case. A survey conducted by Looi and She [17] shows that CI users spend more time listening to music overall than they did just before receiving the implant. Likewise, CI users proclaim more listening enjoyment on average than just before the implant. Fig. 2.6 shows this comparison. As also shown in the figure, the amount of time spent listening, as well as the enjoyment level, is significantly less for those same individuals than before they started to suffer from hearing loss. This confirms that the musical satisfaction of CI users does not compare to the experience of their NH counterparts. With the average CI user clearly showing interest in music but also describing the music listening enjoyment level in the neutral region, a mechanism to help enhance the experience is duly needed.
2.3 The somatosensory system

Somatosensation relates to information carried to the brain from either the inside of the body or the body’s surface, with the main sources of information being load-bearing structures such as tendons and ligaments or body tissues such as muscles or skin. This applies to the perception of touch, temperature, pain, pressure, movement, and of special relevance to this project; vibration.

2.3.1 Skin anatomy

In order to understand how we perceive physical vibration, a brief summary of the anatomy of the skin is required. The skin is the largest organ of the human body. A soft and flexible tissue, it serves multiple purposes, mainly body temperature regulation, body protection and sensory stimulation. It is divided into three primary types of layers, the Epidermis (which is the outermost layer), the Dermis and the Hypodermis (which is the innermost layer). These layers vary in thickness, with the Epidermis being thinnest and the Dermis thickest. The thickness of the layers also varies for various body-parts. For instance, the Epidermis has a thickness of about 0.5 mm on the eyelids and 1.5 mm on the palm of the hand, while the Dermis has a thickness of 0.6 mm and 3 mm on the same locations, respectively [18].

Skin receptors and modalities

Tactile stimulation is presented to the brain by means of mechanoreceptors based in the outer two layers of the skin. These receptors are nerve endings of different shapes and modalities and with different rates of adaptation. The four major types of mechanoreceptors are Merkel’s disk receptors, Meissner’s corpuscles, Pacinian corpuscles and Ruffini endings. As shown in Fig. 2.7, each receptor is at a specific depth inside the skin. In the palm, the shallowest receptor (Meissner’s corpuscle) is at an approx. 0.7 mm depth below the surface while the deepest (Pacinian corpuscle) is at 2 mm [19].

Table 2.1 shows the major mechanoreceptors, the type of stimulation to which they are most sensitive and the range of frequencies to which they respond [21; 22]. Of those receptors, two are of special interest to us. The rapidly adaptive Meissner’s corpuscle is responsible for the sensation of light touch or stroking but also senses vibration, especially at low frequencies in the range of 10 – 50 Hz. When sensing higher frequency vibrations, the main source of information is the Pacinian corpuscle. It is a very rapidly adaptive receptor and is most sensitive to frequencies in the range...
2 Background

Fig. 2.7. Vertical section of the skin, showing the major types of mechanoreceptors (Merkel’s disk receptor, Meissner’s corpuscle, Pacinian corpuscle and Ruffini ending) based in the outer two layers of the skin (Epidermis and Dermis) [20].

Table 2.1: The different modalities and optimal frequencies of the four major mechanoreceptors [21; 22].

<table>
<thead>
<tr>
<th>Receptor</th>
<th>Main sensory modality</th>
<th>Frequency range [Hz]</th>
<th>Rate of adaptation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Merkel’s disk receptor</td>
<td>Pressure</td>
<td>5 - 15</td>
<td>Slowly adaptive</td>
</tr>
<tr>
<td>Meissner’s corpuscle</td>
<td>Stroking</td>
<td>10 - 50</td>
<td>Rapidly adaptive</td>
</tr>
<tr>
<td>Pacinian corpuscle</td>
<td>Vibration</td>
<td>20 - 1000</td>
<td>Rapidly adaptive</td>
</tr>
<tr>
<td>Ruffini ending</td>
<td>Stretching</td>
<td>n/a</td>
<td>Slowly adaptive</td>
</tr>
</tbody>
</table>

of 50 – 600 Hz with a highest sensitivity at around 250 Hz [19], although it does respond to frequencies down to 20 Hz and up to 1000 Hz.

Skin sensitivity

The mechanoreceptors are more densely distributed in glabrous (hairless) skin than in hairy skin. Thus, the glabrous parts (fingertips, palms, soles, etc.) are more sensitive to tactile stimulation through the skin than other parts of the body [23]. R.T. Verrillo measured the vibrotactile thresholds across frequencies 25 – 700 Hz at the tip of the middle finger and found it to be a U-shaped curve, with maximum sensitivity around 250 Hz as shown in Fig. 2.8 [24]. Furthermore, he demonstrated the impact of the contactor size, with his results showing a 3 dB threshold decrease per doubling of the contactor area. A higher threshold should be expected for hairy skin due to the more sparse distribution of receptors, particularly the Pacinian corpuscles.
2.3 The somatosensory system

Fig. 2.8. The vibrotactile threshold at the fingertip, as measured by R.T Verrillo (1971). The plot shows a maximum sensitivity at around 250 Hz [24].

2.3.2 Sensory substitution

The brain has a way of adapting to sensory loss by merging the functionalities of different sensory systems. This remarkable ability is called cross-modal plasticity, or sensory substitution, and is a type of neural plasticity where new neural connections are formed in brain lobes that stop receiving sensory input. This results in the strengthening of other sensory systems, with inactive areas of the brain essentially reassigned new roles. For instance, blind people tend to learn how to use auditory cues to a greater effect than sighted individuals. In this case the visual cortex of the brain is exploited to enhance the sense of hearing. Deaf people also tend to have better peripheral vision than NH people [25] which is a case of the opposite. These are some examples of reorganizations that happen naturally in the brain over time to compensate for sensory depletion. But sensory substitution can also be achieved by means of training.

In later years, the use of the somatosensory system to aid the visually impaired and the hard of hearing has been the subject of much research. Recent results suggest that the use of vibrotactile stimulation can improve the speech-in-noise intelligibility of CI users [26], provided with the appropriate amount of training. This indicates that the auditory cortex of the brain can be trained to enhance the somatosensory system, which gives a reason for optimism that vibration induced somatosensation can be of benefit to CI users when processing complex acoustic signals.
2 Background

2.4 Vibrotactile feedback

Humans have five senses; sight, hearing, taste, smell and touch. But when communicating with electronic systems through the years we have mostly only needed two; sight and hearing. In recent years however, the use of touch simulation in human/electronics interaction has been growing. Most people know this from using vibration mode on their mobile phones or from tapping on touch screens. This touch simulation is known as tactile feedback (or more generally; haptic feedback or haptics) and the term vibrotactile feedback refers to vibration induced tactile feedback. This has provided many exciting new possibilities in the entertainment industry but also new methods for assisting the visually- and/or hearing impaired.

2.4.1 Wearable subwoofers

A popular use of vibrotactile feedback is the conductance of bass frequencies to the body through wearable subwoofers. This provides a more complete immersion experience in both music and cinema, as well as gaming and virtual reality (VR). These devices come in many shapes and forms. Lofelt produces the Basslet, a wearable subwoofer in the form of a wristband, BassMe makes a subwoofer that suspends from the shoulder and vibrates against the chest and SUBPAC offers vibrotactile solutions in the form of backpacks and seatbacks, to name a few. The aforementioned products all serve the purpose of vibrating against the skin along to audio playback from either music, movies or video games, especially effective in a VR environment.

2.4.2 Tactile displays

Tactile displays are devices designed to enable the use of tactile feedback when exchanging information with computers and other electronic systems. While wearable subwoofers offer high definition tactile feedback with frequencies in the range of approx. $10 - 250$ Hz, tactile displays often present an increase in resolution by augmenting the number of stimulators. This is according to the exposition where resolution is the amount of pixels on a display and definition is the amount of information that can be delivered by each pixel. Tactile displays exploit the modalities of the skin’s sensors by applying the type of stimulation that they respond to, whether it be thermal, electrical or mechanical stimulation such as pressure or vibration. A few successful tactile display devices will now be discussed and, for the sake of relevance, focusing on wearable devices that primarily exploit the vibration modality of the Pacinian corpuscle.
2.4 Vibrotactile feedback

Integrated clothing

In 2003, Toney et al. claimed to be the first to successfully implement a socially covert tactile display into wearable clothing by integrating it to shoulder pad inserts [27]. The inserts could then be inserted into standard business suits. Using a pancake electric motor for tactile stimulation, it was intended as mobile access to information such as communication alerts. Since then, a breakthrough has been made in the field of integrated clothing with the term "smart clothing" now an established expression. Vibrotactile feedback is utilized in wearables such as the Nadi X yoga pants, which monitors your posture using sensors and gives adjustment guidance by sending vibrations to relevant parts on the body.

The Emoti-Chair

Music has been presented to the deaf and hard-of-hearing using tactile displays. One such device was presented in 2009 when the Emoti-Chair was demonstrated for the first time [28]. The Emoti-Chair is a sensory substitution system designed to give deaf people a musical experience by vibrating along to audio signals. The sensory substitution controller that the Emoti-Chair employs is the Model Human Cochlea (MHC). Designed by Karam et al., the MHC is an 8-channel vibrotactile stimuli system driving a tactile display of an $8 \times 2$ array of voice coils arranged in the back of the chair [29]. In a place coding strategy similar to that of the human cochlea as described in section 2.1.2, the MHC splits the audio signal into frequency bands and presents them with regards to common perception in pitch height, with low frequencies stimulating the lower back and high frequencies stimulating the upper back.

Sound of Vision

In the Horizon 2020 funded Sound of Vision (SoV) project (No: 643636) a cross-disciplinary research group from several countries (including Iceland, Romania, Hungary, Italy and Poland) developed a system for assisting visually impaired people with orientation and navigation [30–33]. The system is comprised of two depth cameras, one structured infrared light camera for indoor use and one stereo camera for outdoor use, worn on the subject’s forehead, a computer and software worn on the subject’s back, headphones delivering 3D audio cues and the SoV Tactile Belt shown in Fig. 2.9.

Worn around the waist and delivering tactile stimulation to the abdomen, the Tactile Belt has proved to be an important part of the system. Consisting of a $6 \times 10$ array
2 Background

Fig. 2.9. The SoV Tactile Belt comprised of a $6 \times 10$ array of ERM motors.

of eccentric rotating mass (ERM) motors forming a tactile display, the SoV Tactile Belt receives signals from an Arduino microcontroller. The Arduino connects to the computer which maps images from the 3D camera to localized vibrations on the tactile display, creating a visual of the near environment and positioning potential obstructions. The project was considered a success, with blind participants able to perform tasks in challenging environments without the use of the white cane. It received first prize in the “Tech for Society” category at the 2018 Innovation Radar Prize awards [34].

2.4.3 The ACUTE group

With some of the blind participants in the SoV project also having hearing problems and thus being unable to rely on 3D audio, the Tactile Belt proved to be an especially effective addition to the system (R. Unnþórsson, personal communication, July 31, 2020). The Tactile Belt was designed and built by the SoV team at the University of Iceland, supervised by Rúnar Unnþórsson, professor and head of faculty at the faculty of Industrial Engineering, Mechanical Engineering and Computer Science.

Speaking of how the ACUTE group came about, Rúnar Unnþórsson (personal communication, July 31, 2020) recounts how people at Eriksholm Research Centre (part of Oticon) in Denmark contacted him and expressed an interest in finding out if the Tactile Belt could also be used as a sensory substitution device for the hard-of-hearing. With Oticon Medical being one of the six manufacturers of CI devices, special emphasis was given to the possibility of using the Tactile Belt as an assistive device in conveying music to CI recipients to address the issues discussed in section 2.2.3. These talks led to a collaboration being launched between Oticon Medical, University of Iceland, Technical University of Denmark (DTU) and University of Southampton with the objective of researching auditory-tactile solutions.
for CI users, the goal of the Icelandic branch (the ACUTE group) being to modify and improve the SoV tactile solution to fit the concept.
3 Design

In the previous chapter the fundamental concepts behind this project were discussed and section 2.4.3 mentioned the goal of the ACUTE group to further develop the SoV vibrotactile solution. This chapter focuses on the design process of this project, fundamental work carried out for the ACUTE group as part of their long term objective. The chapter starts by defining and outlining the design problem and project objective and explores some of the actuator options available for tactile stimulation. The choice of voice coil actuators (VCA) will be argued after which the L5 actuator chosen for the project is introduced. The equipment deemed necessary for a multi-channel VCA vibrotactile feedback system is listed and the selection of equipment detailed. Finally, the electronic concepts related to the making of a custom-made amplifier, specially designed for this project, will be detailed as well as the amplifier design.

3.1 The design problem and project objective

Early on in the process it was decided that the Tactile Belt should be redesigned. A pivotal part in representing music with vibrotactile feedback was considered to be the complete control of both frequency and amplitude of the vibrations. As later discussed in detail in section 3.2.1, the use of ERM motors such as the ones used in the SoV solution was therefore unfitting.

When the work on this project started, it became clear that the important thing to address first was the type of actuator used. The actuator is the end device of the system, the stimulator which ultimately determines how the vibrations are delivered and perceived. The whole arrangement of both software and hardware is determined with the type of actuator in mind. Thus, replacing the actuators requires a revision of the whole system approach.

Another thing considered of importance was the conjunction of the tactile display design and psychophysical experiments. Each design decision should be made with the user perception in mind. Careful consideration should be given to the placement of the actuators, the number of actuators, audio-tactile mapping strategies, which
parts of the body to receive the stimuli with, etc. It was decided that in order to conduct these psychophysical experiments and tactile display iterations, an easily accessible and reliable in-lab hardware system to drive the experiments was needed.

The new hardware system should have the ability to deliver multiple independent channels of vibration signals at any given frequency and amplitude to be used. Although the exact number of actuators used was still to be decided on, the SoV actuator array of 60 motors was used as a reference. Thus, 60 channels was considered a minimum. The vibration frequency and amplitude of each channel should be controllable by the software in development and the signals delivered to the tactile display in development. Fig. 3.1 shows the hardware system depicted as a black box in the chain of operation. The black box system would be treated as an in-lab proof-of-concept system. Although a wearable and consumer-friendly system is the long-term objective, low-cost and wearability are not demanded for the black box at this stage. Rather, the system should provide easy accessibility to further development.

Fig. 3.1. Black box depiction of the multi-channel hardware system. The system is controlled by the software in development and delivers multiple channels of high definition vibration signals to the tactile display in development.
Having the knowledge of the non-uniform sensitivity of the skin across frequencies as previously shown in Fig. 2.8, it was also desired to acquire the vibrotactile threshold curve from the stimulation received by the actuators of choice. With the notion in mind that different channels could be assigned to different frequencies, the data from these measurements would then be used to compensate across channels in the software development by adjusting the amplitude according to the measured threshold at a given frequency.

In addition, the data from the tactile threshold measurements should be presented as to represent the perceived amplitude of stimuli. A transfer function from the hardware system signal output to the input of perceptual amplitude should therefore be derived.

To summarize, this project’s objective is to:

- Find the appropriate actuator to work with in the development of a vibrotactile assistive device for enhancing the music listening enjoyment of CI recipients.
- Design and build an in-lab proof-of-concept hardware system of 60 channels minimum that will enable psychophysical experiments with multi-channel vibrotactile feedback and different mapping.
- Implement and conduct a vibrotactile threshold experiment using the new hardware system and actuators.
- Present the data in a concise way such that it depicts the user experience.

### 3.2 Actuator

Considering the spectral limitations offered by modern CI technology and the nature of the stimulus (music) that the actuator is transporting in our case, high definition vibrations (as defined in section 2.4.2) were considered necessary for a comprehensive sensory substitution. Furthermore, the feasibility of implementing the actuators into wearable solutions later on should be taken into consideration. In the following section, some of the actuator options available are reviewed.
3 Design

3.2.1 Vibrotactile actuators

There is a wide variety of actuators that can be used to provide vibrotactile stimulation to the skin but they are generally divided into three main categories; rotary electromagnetic actuators, linear electromagnetic actuators and non-electromagnetic actuators. The actuators within these categories all convert electrical energy into mechanical energy but operate in different ways.

Rotary electromagnetic actuators

Rotary electromagnetic actuators (REA) are direct current (DC) motors designed to rotate when a steady voltage is applied to the leads. A common internal structure of the DC motor is the ERM structure, an off-center mass fixed to an armature that rotates in a magnetic field inside a stationary part called the stator. When the eccentric mass \( m \) at a distance \( r \) from the center of an axis rotates around the axis with an angular velocity \( \omega \), it creates a centrifugal force

\[
F = mr\omega^2
\]  

(3.1)

which is perceived as the vibration amplitude. This perceived amplitude is however also dependent on an external mass attached to the actuator as later discussed in section 3.2.2.

The torque \( T \) is determined by

\[
T = K_a \phi I_a
\]  

(3.2)

where \( \phi \) is the total flux determined by the field winding and field current, \( I_a \) is the current in the armature, determined by the back electromotive force and the armature winding resistance \( R_a \), and \( K_a \) is a constant determined by design parameters such as the physical dimensions of the motor.

With a supply voltage \( V \) the angular velocity of the ERM motor will be

\[
\omega = \frac{V}{K_a \phi} - \frac{T}{(K_a \phi)^2 R_a}
\]  

(3.3)

Thus, the speed of rotation can really only be controlled by adjusting the supply
3.2 Actuator

voltage, with all other parameters of equation (3.3) fixed and dependent on the internal structure of the motor. Inevitably, adjusting the supply voltage will at the same time have an effect on the amplitude since the angular velocity and the amplitude are related as seen by equation (3.1). Of course this will also affect the measured frequency, since

\[ f = \frac{\omega}{2\pi} \quad (3.4) \]

The frequency and amplitude of the vibration are therefore unavoidably linked when using the ERM motors. Attaining a high amplitude vibration at a low frequency, for instance, is impossible because an increase in voltage results in an increase in both frequency and amplitude.

**Linear electromagnetic actuators**

A linear electromagnetic actuator is essentially an electromagnet, an electrically conductive wire, most commonly copper, which is wrapped into a coil. The wires are covered with an electrically insulating material and are tightly wound. When a current \( I \) passes through the coil a magnetic field \( B \) is created which is strongest in the middle of the coil as represented by the combined field lines in Fig. 3.2. When a permanent magnet is placed in the center of the coil an interaction occurs between the two magnetic fields and the wire will experience a magnetic force (derived from the Lorentz force law) known as the Laplace force,

\[ F = I \int dl \times B \quad (3.5) \]

where \( l \) denotes the length and direction of the wire. This will result in the permanent magnet being either attracted to the magnetic field produced by the coil or repelled by it, depending on the orientation of the magnet poles and the direction of the current in the loop. A constant current will result in a steady force while an oscillating current will create a correspondingly oscillating force.

This is the function of the voice coil actuator (VCA). With a fixed coil and a free permanent magnet, the force will cause the magnet to move back and forth at an amplitude and frequency determined by the current, essentially behaving as a loudspeaker without the cone attached. Controlling the amplitude and frequency of the oscillation can be done by varying the current. Using the same example as with the REA, a high amplitude vibration at a low frequency can be achieved by
Fig. 3.2. Diagram of an electromagnet. A current passing through the loop of wire will produce a magnetic field in the direction represented by the lines. The magnetic field will be strongest in the center of the loop. Placing a permanent magnet (ferromagnetic material which produces its own magnetic field) in the center will result in it being either attracted to or repelled by the magnetic field produced by the coil. Changing the orientation of the current flow will change the direction of the magnetic field [35].
3.2 Actuator

producing a high amplitude and low frequency AC signal. Therefore, voice coils fulfill the design requirements of independent control of frequency and amplitude.

Non-electromagnetic actuators

Non-electromagnetic actuators take advantage of the piezoelectric effect, which is the ability of certain materials to respond to deformation caused by mechanical force by producing an outer electric field. This response is called the direct piezoelectric effect. But the effect is also reversible in the form of the inverse piezoelectric effect, where the piezoelectric material becomes strained and changes shape when put inside an electric field. In that way, piezoelectric materials can be used as sensors as well as actuators.

The inverse piezoelectric effect can be described mathematically in terms of the strain charge by

$$ S = s_E T + d^T E $$

or in matrix form

$$ \begin{pmatrix} S_{xx} \\ S_{yy} \\ S_{zz} \\ S_{yz} \\ S_{xz} \\ S_{xy} \end{pmatrix} = \begin{pmatrix} s_{E11} & s_{E12} & s_{E13} & s_{E14} & s_{E15} & s_{E16} \\ s_{E21} & s_{E22} & s_{E23} & s_{E24} & s_{E25} & s_{E26} \\ s_{E31} & s_{E32} & s_{E33} & s_{E34} & s_{E35} & s_{E36} \\ s_{E41} & s_{E42} & s_{E43} & s_{E44} & s_{E45} & s_{E46} \\ s_{E51} & s_{E52} & s_{E53} & s_{E54} & s_{E55} & s_{E56} \\ s_{E61} & s_{E62} & s_{E63} & s_{E64} & s_{E65} & s_{E66} \end{pmatrix} \begin{pmatrix} T_{xx} \\ T_{yy} \\ T_{zz} \\ T_{yz} \\ T_{xz} \\ T_{xy} \end{pmatrix} + \begin{pmatrix} d_{11} & d_{21} & d_{31} \\ d_{12} & d_{22} & d_{32} \\ d_{13} & d_{23} & d_{33} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} $$

where $ S $ is the strain, $ T $ is the stress, $ E $ is the electric field, $ s_E $ is the material compliance and $ d $ represents the material coupling properties. Therefore, piezoelectric materials do not use magnetism in its operation. Instead it’s the change in an outer electric field that causes the material to deform or change shape. Piezoelectric materials respond to AC signals and do offer the option of independent control of frequency and amplitude. However, they also require high voltage inputs (typically 100 V or more) which makes them challenging to work with.

Table 3.1 shows a summary of the properties of the main actuator types. Considering the design requirements of high definition vibration and practical integration, voice coils were considered the best option to work with.


Table 3.1: Comparing the viability of the main actuator alternatives for vibrotactile feedback.

<table>
<thead>
<tr>
<th>Actuator</th>
<th>High definition vibration?</th>
<th>Practical integration?</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC motor</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>VCA</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Piezo</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

Fig. 3.3. The L5 VCA. A voice coil packaged in a plastic case. A permanent magnet in the middle reacts to the magnetic field caused by the current flowing through the coil wire. AC signals will produce periodic upwards- and downwards movement of the magnet, according to the coordinate system on the image. Putting the flat 17.0 × 20.5 mm case surface up against the body results in a parallel vibration to the skin [36].

3.2.2 L5

The VCA that was chosen to provide vibrotactile stimulus in this project is the L5, shown in Fig. 3.3. It is manufactured by Lofelt GmbH and is the actuator used in their brand, the Basslet, previously introduced in section 2.4.1.

The L5 is a voice coil specially designed to give parallel vibrations to the skin. The wires are wound inside a plastic case of dimensions 17.0 × 20.5 × 6.2 mm (W × D × H). It can be driven by any low-power amplifier circuit as an 8 Ω speaker load, drawing an average current of 10 mA at medium volume and 57 mA at maximum volume. [37] By applying an audio signal to the L5 an AC current runs through the coil winding, producing a magnetic field as described in section 3.2.1. A permanent magnet in the middle responds by moving back and forth at the same frequency.
3.2 Actuator

Fig. 3.4. The acceleration response of the L5 actuator with three different masses attached (30 g, 60 g, 100 g). Measured by Lofelt using sine waves of maximal transient voltages (1.3 \( V_{\text{RMS}} \), 1.2 \( V_{\text{RMS}} \) and 1.1 \( V_{\text{RMS}} \), respectively). The graph shows a peak acceleration of 4.5 G at a resonance frequency 64 Hz using a 30 g mass. Acceleration decreases as external mass increases [37].

as the input signal, resulting in a parallel vibration stimulus when placed flat up against the skin. Metal plates are screwed to the edge of the plastic case and to the permanent magnet on both sides of the L5. The plates create a spring effect and return the magnet into starting position in the center of the case once the signal is turned off.

Acceleration is commonly used as a measure for vibration amplitude and is presented in the unit G, where G refers to the gravitational acceleration constant \( G = g = 9.81 \, \text{m/s}^2 \). What we sense as vibration is essentially an object being repeatedly displaced at certain frequencies. The actuator itself is always attached to an object, a mass, that vibrates. Looking at equation 3.1 we see that it does not take this external mass into account. A heavier object will take more force to generate the same acceleration than a lighter object. The amplitudes produced by the same actuator using the same force in displacing two objects of different masses will not be the same. Thus, the acceleration is a more accurate description of the perceived amplitude, usually measured with accelerometers.

Fig. 3.4 shows the L5 acceleration response as demonstrated in the L5 data sheet [37]. It was measured using sine waves of peak transient voltages for three different masses attached externally. The plot depicts the resonance frequency of the system (actuator and attached mass) at 64 Hz ±5%, observed as a peak on the acceleration axis. As also observed, the system offers a response of 1 G or more over the key tactile sensitivity range of 45 Hz to 250 Hz when attached to a 30 g mass. The
acceleration of frequencies outside this range is lower although the actuator does respond to any AC current applied.

The power handling is rated at 320 mW (RMS). Although the L5 can be driven by any audio amplifier, a low-power circuit of maximum 2 W output is advised (A. Lazdins, personal communication, August 2019). This is to prevent excessive current consumption from causing overheating and reducing the actuator lifetime.

### 3.3 Hardware system overview

Having decided on the L5 VCA as the tactile stimulator, the concept of the black box hardware system discussed in section 3.1 became more clear. The L5 should be driven by a low-power audio amplifier, thus a multi-channel audio playback system was needed.

The software in development should generate audio signals to multiple channels, the processing strategy to be determined later on. The signals should be handled by an external audio interface of a minimum 60 channels. With the amount of audio signals generated, a digital interface is required. Digital-to-analog conversion is therefore needed between the interface and the amplifying stage as the amplifier will not respond to digital signals. Since each L5 should be controlled individually, a minimum 60-channel low-power amplifier is desired.

For the sake of convenience, as the system is intended for in-lab experiments, rack mountable equipment is preferred. As discussed and listed in the following sections, the in-lab hardware rack system can be assembled from equipment already available on the market with the exception of the multi-channel amplifier.

### 3.4 Audio interface

Audio interfaces are professional sound cards designed to handle multi-channel recording and playback with low latency and high fidelity. They typically use drivers that follow the Audio Stream Input/Output (ASIO) protocol when working with Windows operating systems (OS). The ASIO drivers offer direct contact between the external audio interface and the software playback on the personal computer (PC) via serial connectors such as USB or Firewire (IEEE 1394), bypassing the internal audio signal path of the PC and reducing the latency.
3.4 Audio interface

3.4.1 MADI

In recent years, there’s been much growth in the use of digital audio equipment. Digital audio systems offer more compactness and lower-cost than analog systems, consisting more of software and plugins rather than physical hardware. It also makes a world of difference when it comes to signal processing and editing. Concurrently, there’s been an increase in demand for a greater number of channels to work with for large scale productions and recording, such as orchestral- or surround sound productions.

The serial Multi-channel Audio Digital Interface (MADI) is the pro industry standard for multi-channel audio and was defined by the Audio Engineering Society (AES). It builds on the AES3 digital audio standard, which offers two channels of 24-bit audio (32 bits total) per sample. Essentially, MADI signals contain 32 AES3 signals in series per frame, using time division multiplexing to fit 64 channels of real-time audio into one single cable. The transmission of MADI signals can either be made through 75 Ω BNC connectors, using standard Coax cables, or through SC input/output (I/O) connectors, using multimode optical fibre.

The protocol allows for a maximum transmission bit rate of $100 \text{ Mbit}_s$. With a bit depth of 32 bits and with 64 channels, then according to

$$\text{Bit rate} = \text{Sample rate} \times \text{Bit depth} \times \text{Channels}$$

the maximum sample rate is 48 kHz but can be increased to 96 kHz by reducing the number of channels to 32.

3.4.2 RME MADIface XT

The MADI protocol satisfies the design requirements for the number of channels. The audio interface chosen for the in-lab hardware system is the MADIface XT produced by RME. It was the first audio interface to offer USB 3.0 compatibility [38]. Fig. 3.5 shows the rear view of the device. With two SC I/O ports for optical use and one 75 Ω BNC I/O port for Coax, it offers up to 192 MADI channels, using sample rates 44.1 kHz or 48 kHz, or 96 MADI channels using a 96 kHz sample rate.
3 Design

Fig. 3.5. RME MADIface XT rear side. The device has two optical- and one coaxial MADI I/O ports, offering up to 192 channels of digital audio. Clock signals are distributed through BNC connectors, BNC In for an external clock source or BNC Out when using internal clock mode [38].

![Diagram of TotalMix GUI layout](image)

**Fig. 3.6.** Diagram of the TotalMix GUI layout. TotalMix offers manual routing from any hardware input (not used in our setup) or software playback input to any hardware output [38].

**TotalMix**

MADIface XT includes a digital real-time mixer, TotalMix. It allows for unlimited routing from any hardware- or software playback input to any hardware output. It also offers many options such as equalizing or signal splitting that are not necessary for the purpose of this project. What’s needed is simply to route software playback channels 1-64 to the hardware output channels belonging to the MADI port in use. Routing is done manually via the TotalMix graphic user interface (GUI), which depicts each channel as a fader and level meter. By placing each input- and output fader at 0 dB, the signal will go unaltered from the playback software to the hardware MADI outputs of the MADIface XT.
3.5 Digital-to-analog converter

Word clock

Digital audio systems consisting of two or more separate devices need a common word clock signal to synchronize all the equipment, securing that all inputs and outputs of the system are in phase ($0^\circ$). The term word clock refers to the fact that one sample contains one data word and each sample is clocked. MADIface XT has an internal clock mode and distributes the clock signal through additional BNC connectors, providing constant word clock output at a frequency set by the sampling rate. [38] By using the internal clock mode, the MADIface XT is defined as the master and all other equipment in the chain as slaves.

3.5 Digital-to-analog converter

As audio amplifiers only respond to analog signals, the signals coming from the MADIface XT need to be converted via a Digital-to-analog converter (DAC). DACs receive a discrete-time signal $x[n]$ and convert it to a continuous-time signal $x(t)$, typically by holding each sample value for a time interval $T$

$$T = \frac{1}{f_s} \quad (3.7)$$

where $f_s$ denotes the sampling frequency. This produces a sequence of rectangular functions of varying amplitudes determined by $x[n]$. The rectangular sequence then gets filtered to create a smooth continuous function replicating the original analog signal.

In accordance with the Nyquist-Shannon theorem, the original analog signal can be completely reconstructed by the DAC provided that the highest frequency component of the original signal (the Nyquist frequency) does not exceed one half of the sampling frequency ($B < \frac{f_s}{2}$, where $B$ is the Nyquist frequency). Failure to produce a sufficient sample rate will introduce a type of distortion called aliasing. In audio, the aliasing effect generates lower pitch sounds not present in the original signal. A minimum sampling frequency of 44.1 kHz is therefore used to secure an accurate reconstruction of all signals within the range of human hearing.
Fig. 3.7. Rear view of the Ferrofish A32 AD/DA converter. Each converter can transmit up to 32 channels of analog audio. By connecting a second A32 in series the number of analog output channels is increased to 64 [39].

3.5.1 Ferrofish A32

The Ferrofish A32 Analog-to-digital/Digital-to-analog (AD/DA) converter was chosen to handle the reconstruction of the signals from MADI to analog. Fig. 3.7 shows the rear view of the converter. It contains an optical SC I/O port as well as 75 Ω BNC port for MADI transmission and reception and 8 D-sub25 connectors for analog I/O (4 in and 4 out). Hence, each A32 is capable of transmitting 32 analog outputs, meaning that two devices are required for the full 64-channel system. By connecting the A32s in series and by internal routing within the first A32, MADI channels 33-64 can be routed to the MADI output of the first converter to be forwarded to the input of the second converter.

Each MADI I/O port induces a latency of 3 samples [40], which will cause the analog output signals of the two AD/DAs to be incoherent, although only slightly. This incoherence can be rectified by setting the two devices as primary and secondary, which will sync the two devices to the same delay of 3 samples. Further latency is to be expected from the conversion process depending on the sample rate (0.16 ms at 48 kHz, 0.06 ms at 96 kHz). Using the 48 kHz sampling frequency for the option of 64 channels, the total latency is therefore $3T + 0.16 \times 10^{-3} = 0.225 \text{ ms}$.

3.6 Amplifier

Audio electronic devices such as the Ferrofish A32 have analog outputs at a line level signal strength, called line out. A line out connection has an output impedance typically in the range of 100 - 600 Ω and is designed to drive a higher impedance line input (usually about 10 kΩ) of a cooperative audio device. The two impedances therefore form a voltage divider where most of the voltage submitted by the output is dropped across the line input. In that way, a signal path is formed between
separate devices with minimal loss of information.

VCAs however rely on the current in their performance. The serial impedance of the 8 $\Omega$ coil and the few hundred ohms of the line out will result in an insufficient current supply for the VCA causing the signal driving the L5 to be very weak. An amplifying circuit to lower the output impedance is required after the DAC stage to effectively drive the L5.

The audio interface and the DAC discussed in the previous sections are standard audio equipment already available on the market. However, our amplifier requirements are unorthodox. It should be low-power as to avoid excessive current consumption and overheating of the coils and it should offer easy connection options of up to 64 input- and 64 output channels in a compact, rack-mountable unit. The search for an amplifier fulfilling these requirements proved unsuccessful. After exploring the alternatives, the decision was made to build a low-power multi-channel amplifier precisely formatted to our needs.

In the following sections, some of the tools and theories used in the design process will be examined after which the multi-channel amplifier design is discussed.

### 3.6.1 RC filters

RC filters are first order filters made up of resistors and capacitors. They can serve the purpose of either low-pass filtering or high-pass filtering, or band-pass filtering by applying two filters of low- and high-pass function in series.

**Low-pass**

A low-pass filter (LPF) is a circuit that passes frequencies lower than a selected cutoff frequency $f_c$ and attenuates all higher frequencies. The cutoff frequency of RC filters is

$$f_c = \frac{1}{2\pi RC}$$  \hspace{1cm} (3.8)

where $R$ denotes the resistance value of the resistor and $C$ is the capacitance. The cutoff frequency is defined as the frequency where the total output power is half that of the input power, or -3 dB attenuation. For a low-pass function, the resistor is connected in series with the load and the capacitor in parallel. The capacitor
3 Design

exhibits a high impedance $Z_C$ at low frequencies as described by

$$Z_C = \frac{1}{j\omega C} \quad (3.9)$$

where $j$ is the imaginary unit. As the frequency increases, $Z_C$ decreases resulting in a lower voltage drop across the capacitor and the parallel load. At very high frequencies the capacitor is essentially seen as short circuiting, running the total current to ground.

High-pass

As opposed to LPFs, a high-pass filter (HPF) is a circuit that passes frequencies higher than the cutoff frequency described by equation (3.8). In this case the capacitor is connected in series with the load and the resistor in parallel. At very low frequencies the capacitor is essentially seen as an open circuit and the current is blocked. As the frequency increases, more and more current passes through the capacitor and to the resistor and load, resulting in more voltage drop across the load at high frequencies.

3.6.2 Pulse-width modulation

A pulse-width modulator (PWM) is a circuit that converts analog signals into logic interpretations of the original signals. They are commonly used as on/off controllers for power transistors in a switched-mode setup. Fig. 3.8 shows a simple PWM circuit and the relevant waveforms. An analog comparator circuit receives input voltage $v_c(t)$ from an analog source and compares it to an input signal $v_{saw}(t)$ generated by a sawtooth generator. The comparator produces a logic high output whenever $v_c(t)$ is greater than $v_{saw}(t)$, otherwise it produces a logic low.

The logic output signal $\delta(t)$ is periodic with a time interval $T_s$ and with a duty cycle $D$. The duty cycle is defined as the fraction of one period in which the system is active and it will be proportional to the analog input signal strength. If $v_c(t)$ is always 0 V then the duty cycle will be zero. If $v_c(t)$ is always greater than or equal to $v_{saw}(t)$ then the duty cycle is 1. Overall, the duty cycle will be a function of $v_c$,

$$D = \frac{v_c(t)}{V_M} \quad \text{for } 0 \leq v_c(t) \leq V_M \quad (3.10)$$
3.6 Amplifier

Fig. 3.8. Left: A simple PWM circuit. A comparator circuit compares an analog input signal $v_c(t)$ to a sawtooth wave $v_{saw}(t)$. Right: Input and output waveforms of the PWM circuit. Output is high whenever $v_c(t)$ is greater than or equal to $v_{saw}(t)$, otherwise low [41].

where $V_M$ is the maximum value of $v_{saw}(t)$.

3.6.3 Boost converters

A boost converter is a switched-mode power supply that steps up the voltage from its input. It usually consists of two or more semiconductors (MOSFET and diode) and two energy storage components (inductor and capacitor). The following explanation and equations follow the methodology presented by Erickson and Maksimovic [41].

Fig. 3.9 shows a diagram of a simple boost converter with input voltage $V_g$ and output voltage $V$. A PWM signal drives the MOSFET, turning it on over time interval $DT_s$, where $T_s$ is the interval of a switching period, and off at time interval $(1 - D)T_s$. The figure also shows the two semiconductors modeled as an ideal switch. A PWM logic high during $DT_s$ results in the equivalent circuit shown with switch in position 1. Assuming a steady state, then during this period of the cycle, inductor voltage is $v_L = V_g$. At a logic low during $(1 - D)T_s$ the MOSFET turns off and the diode conducts, resulting in the equivalent circuit with switch in position 2. Now the inductor voltage will be $v_L = V_g - V$. The conversion ratio $M$ can be derived with the knowledge that the average inductor voltage over one period in steady state is zero.

$$\int_0^{T_s} v_L(t) dt = V_g DT_s + (V_g - V)(1 - D)T_s = 0$$
which gives the conversion ratio

\[ M = \frac{V}{V_g} = \frac{1}{1 - D} \]  

which can be expressed in terms of gain level \( L_G \) as

\[ L_G = 20 \log \left( \frac{V}{V_g} \right) = 20 \log(M) \]

This is of course assuming an ideal switch. In practice, parameters such as semiconductor on-resistances and inductor copper loss reduce the conversion ratio slightly. It can however be observed that the output voltage is for the most part dependent on the input voltage and the duty cycle of the PWM signal driving the MOSFET.

The value of a boost converter inductor should be determined with regards to the
3.6 Amplifier

desired inductor ripple current $\Delta i_L$. This is the peak-to-peak variation of the average inductor current $I_L$,

$$I_L = \frac{IV}{V_g\eta}$$

(3.13)

where $I$ is the output current and $\eta$ is the efficiency. Lower ripple currents reduce core losses in the inductor and increase the device power capability but also help minimise the risk of electromagnetic interference. However, smaller ripples do demand higher inductance values $L$ as seen by

$$L = \frac{V_g \times (V - V_g)}{\Delta i_L \times f_{sw} \times V}$$

(3.14)

where $f_{sw}$ is the switching frequency.

The capacitor value is determined with the maximum allowed output voltage ripple $\Delta V$ in mind. The voltage ripple is the peak-to-peak variation in output voltage and should preferably be as small as possible as it’s a common cause for noise and distortion. The capacitance $C$ is given by

$$C = \frac{I \times (V - V_g)}{\Delta V \times f_{sw} \times V}$$

(3.15)

It can be seen that a smaller ripple results in higher capacitance which in turn leads to reduced response time. Choosing the value of the boost converter capacitor is therefore a trade-off between the suppression of noise and suppression of large output current transients.

3.6.4 Amplifier classes

An important function of an amplifier output stage is to provide low output impedance to prevent the loss of gain. Another concern is the ability of the amplifier to deliver the required power to the load efficiently. The efficiency $\eta$ describes the ratio of the power used to the power supplied,

$$\eta \equiv \frac{P_L}{P_S}$$

(3.16)
where $P_L$ is the load power and $P_S$ is the supply power. Ideally, the power dissipated should be as low as possible and $\eta$ as close to 1 as possible. Power dissipation in a transistor leads to a raised internal junction temperature which can ultimately destroy the component. Taking the ACUTE long term goal of a portable and wearable system into account, high efficiency is especially important as it prolongs the battery life. But more importantly, heat can become a decisive issue with regards to comfort.

Some common amplifier categories, or classes, will now be explored. These classes all provide the low output impedance needed to drive the load but use different methods. The suitability of each amplifier class depends on the goal that one is trying to accomplish. It will be argued that a class D amplifier is the appropriate choice for this design, providing the right balance between efficiency and quality.

**Class A**

Class A amplifiers are analog amplifiers with an output stage biased at a DC current that is greater than the amplitude of the incoming signal. As a result, the output transistor conducts for the entire cycle of the input signal ($360^\circ$). Since the transistor is always conducting there is no turn on delay, which guarantees a better performance at high frequencies and generates fewer harmonics in the output signal. Thus, the audio quality is generally very good. However, this type of amplifier can only attain a maximum efficiency of $\eta = \frac{1}{4}$, or 25%.

**Class B**

A more efficient version of an analog amplifier output stage is class B. It consists of two complementary transistors (nnp and pnp) that work in a push-pull fashion and has a maximum efficiency of $\eta = \frac{\pi}{4}$, or 78.5%. Class B amplifiers are biased at zero DC current and each transistor conducts for a half cycle ($180^\circ$). The nnp "pushes" current into the load when the input signal $v_I$ is positive and the pnp "pulls" current from the load when $v_I$ is negative. When $v_I$ is around zero both transistors are cut off, introducing a nonlinearity, a dead zone, which results in crossover distortion to the output signal that significantly reduces quality.

**Class AB**

Class AB amplifiers seek to eliminate crossover distortion by operating in the same push-pull fashion as class B amplifiers, but biasing the transistors at a small DC
3.6 Amplifier

current. Thus, both transistors conduct for more than half the cycle, overlapping the dead zone. Biasing is done by applying a voltage between the base junctions of the two transistors and passing a constant quiescent current when there is no input signal. This quiescent current however results in dissipated power which reduces the efficiency of the device. In practice, $\eta$ is in the range of 50 - 60\% [42].

Class C

Class C amplifiers are usually employed for RF signals and were therefore not considered for this design.

Class D

Whereas the previously discussed amplifier classes are purely analog, a class D amplifier uses analog and digital techniques. It operates with switching transistors that are switched on and off by a pulse format conversion of an audio input signal, converted by a PWM. The frequency of the PWM signal determines the switching frequency of the transistors and is usually at least 200 kHz or 10 times greater than the highest frequency component of the signal to be amplified (20 kHz, upper limit of audible frequencies). The amplitude and frequency of the PWM signal is constant but the width of each pulse carries the information on the magnitude of the audio signal at that point in time.

Fig. 3.10 shows a simple class D amplifier output stage with an input PWM, which is a logical inverse of the PWM signal. Class D uses two MOSFET transistors, a PMOS ($Q_P$) that conducts only during low input and NMOS ($Q_N$) which conducts when input is high. The transistors operate as switches, by switching between an external voltage source $V_{DD}$ and ground, allowing a constant current flow to the output when $Q_P$ is conducting and turning the current off when $Q_N$ conducts. The output signal is an amplified PWM signal (inverse of incoming signal) that passes through a LPF to derive the time average and reconstruct the original audio signal information as good as possible. The LPF cutoff frequency is just above the highest frequency of the audio signal so all switching harmonics will be filtered out. Theoretically, the efficiency of class D amplifiers is 100\% but due to practical issues such as transistor on-resistances, the maximum efficiency is about 90\%.

Table 3.2 shows a quick comparison of the amplifier class performances. It shows that when it comes to efficiency, class D amplifiers are the best solution and this is achieved without sacrificing too much quality. It can be argued that due to the far greater sensitivity of the ear compared to that of the skin, a slight dip in
Fig. 3.10. A simple class D amplifier. When PWM input signal is low, $Q_P$ is on and $Q_N$ is off, allowing constant current flow. When PWM input is high, $Q_P$ is off and $Q_N$ is on, switching the circuit connection to ground. The output is an amplified PWM signal that is passed through an LPF to recover the amplified audio signal [42].

Table 3.2: Comparing the performances of the main amplifier classes.

<table>
<thead>
<tr>
<th>Amplifier class</th>
<th>Output quality</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class A</td>
<td>Very good</td>
<td>Poor</td>
</tr>
<tr>
<td>Class B</td>
<td>Poor</td>
<td>Good</td>
</tr>
<tr>
<td>Class AB</td>
<td>Good</td>
<td>Average</td>
</tr>
<tr>
<td>Class D</td>
<td>Good</td>
<td>Very good</td>
</tr>
</tbody>
</table>

audio quality will not be an issue when it comes to tactile stimulation. Efficiency was deemed to be the most important factor resulting in class D as the amplifier category of choice.

3.6.5 TPA2025D1

The Texas Instruments TPA2025D1 integrated circuit (IC) is a class D audio amplifier that comes in a space saving $1.53 \times 1.98$ mm ball grid array (BGA) package (12-ball, 0.5 mm pitch). An automatic gain control (AGC) circuit for battery tracking limits the current consumption at low battery voltage to extend the battery life. These features are an advantage when it comes to portability. TPA2025D1
was chosen as the class D amplifier of use in the design of the amplifier channels driving the L5 VCAs in this project. A high input impedance (24 kΩ), low output impedance device, it delivers up to 1.9 W to an 8 Ω load, assuming total harmonic distortion (THD) to be 1%. The class D amplifier has a fixed gain of 20 dB and an efficiency of around 85% [43].

Figure 3.11 shows a block diagram of the IC. It powers up at a logic high applied to the enable pin EN and will work with a DC supply voltage in the range of 2.5 - 5.2 V (6 V absolute maximum) applied to VBAT. An audio signal (IN+, IN-) enters the AGC circuit before it is amplified. The voltage on AGC pin determines the AGC threshold and gain is reduced once VBAT drops below that threshold. The signal is then converted to pulse format via the PWM and transmitted to the class D amplifier stage, implemented in an H-bridge circuit. The amplified PWM signal then exits through the IC outputs (OUT+, OUT-).

A built-in boost converter operates from VBAT and generates a steady 5.75 V supply voltage PVDD for the class D amplifier stage. An internal oscillator generates a 1.2 MHz switching frequency for the boost converter MOSFET. The boost converter has an auto pass through (APT) mode where an internal bypass switch directly connects PVDD to VBAT. During low output signal strength, the class D amplifier operates with VBAT. When the output signal exceeds the APT threshold (2 V peak) then the boost converter is activated and the class D amplifier operates with PVDD. An external boost converter inductor and capacitor should connect to pins SW and
3 Design

The total supply current can be estimated from the plot shown in Fig. 3.12. Assuming a 5 V DC supply voltage, the supply current at full power (1.9 W) is approx. 0.45 A.

3.6.6 KiCad

KiCad is a free electronic design automation (EDA) software that enables the conversion of electronic circuit design schematics to printed circuit board (PCB) layouts and to Gerber formats for manufacturing. It features three main applications in an integrated environment:

- **Eeschema** - A schematic capture software for circuit design and drawing in a hierarchical schematic sheets environment. It includes a symbol library and a symbol editor, where component symbols can be created or edited. Each schematic symbol is manually coupled with a PCB footprint.

- **Pcbnew** - A PCB layout tool used in association with Eeschema. Here, all the PCB attributes (number of layers, component placements, track routing and track widths, vias, etc.) are defined. Pcbnew includes a footprint library and footprint editor, where new component footprints can be created and added to the existing library.

- **GerbView** - A Gerber file and drill file viewer. Gerber format is the standard
used in the PCB manufacturing industry to describe the PCB details (copper layers, solder mask, etc.) while drill files contain the information on through holes and vias.

### 3.6.7 Amplifier design

In previous sections, some fundamental concepts regarding the design of the multi-channel amplifier have been discussed. These include PWM, boost converters and class D amplifiers, all of which are important in understanding the TPA2025D1 amplifier chosen for this project. The actual design of the multi-channel amplifier will now be detailed. As previously stated, the multi-channel amplifier is part of an in-lab rack system and as such, is housed in 19\(^{n}\) rack-mountable enclosures. However, steering towards a wearable solution, each amplifier channel occupies its own PCB.

#### Enclosure

The multi-channel amplifier audio line level in is via D-sub25 female connectors on back side of the enclosure as shown in Fig. 3.13 (a). The connectors are wired according to the TASCAM DB-25 pinout standard \[44\] and connected to each channel PCB. The channel PCB outputs connect to RCA-F connectors on the front side of the enclosure (Fig. 3.13 (b)). Shielded audio cables carry the signals to- and from each channel PCB.

Implementing RCA enclosure outputs instead of space saving multi-plugs is to enable the connection of 1-64 L5s, allowing single channel- as well as multi-channel experiments. In that way there’s no barrier in the amount of actuators used in the experiments, within the specified frame. Each 19\(^{n}\) rack unit space can fit 16 RCA connectors, resulting in 4U total. The multi-channel amplifier is therefore housed in two separate 2U enclosures.

#### Power supply

Current consumption is perhaps the main source of uncertainty in this design. The 0.45 A total supply current specified in the TPA2025D1 data sheet is for the case of full power use and 1% THD. Actual current consumption of the device in normal use can only be accurately predicted by measuring the performance of channel and load. This is among the things discussed later in chapter 5.
Following the data sheet specification, it has to be ensured that the DC supply can handle worst case scenario. The maximum current consumption of one amplifier enclosure unit is 14.4 A. This can be achieved by implementing two 5 V / 8 A DC power supplies in each enclosure, with each supply powering 16 channels. The leads from the power supply DC outputs connect to a Veroboard voltage distributor from where PCB DC inputs of channels 1-16 and 17-32 receive their 5 V supply in parallel. The power supply AC inputs are connected to an AC on/off switch on the back side of the enclosure. Earth wire grounds the enclosure by connecting to chassis.

**Channels**

As previously stated, each channel resides on its own PCB. This is done to make the transition process from rack-mount to wearable less demanding later on. The KiCad EDA was used for the circuit drawing and the design of the PCB layout.

Fig. 3.14 shows an Eeschema schematic of one channel. Each PCB has three 2-pin sockets (J1, J2, J3) for the wiring of DC in, audio in and audio out, respectively. Enable pin EN is connected to the 5 V DC supply through resistor R1 so that the IC is activated by the AC power switch on back side of the enclosure. Capacitor C2 acts as a charge reservoir in case of input voltage droop, while C3 and C4 are decoupling capacitors, shunting noise and voltage spikes of high frequencies to ground.

Due to uncertainties regarding the current consumption, the Texas Instruments suggestions were followed for the boost converter component values L1 and C1. As VCAs are non-linear loads with varying impedances over frequency, there’s a risk of small load current flowing back through the H-bridge and to the boost converter.
3.6 Amplifier

Fig. 3.14. Eeschema design schematic of one amplifier channel.
output stage. To prevent flow-back current from driving the PVDD voltage above the absolute maximum operational voltage, a 6.8 V Zener diode (D1) was connected to the PVDD output.

The audio input stage utilizes a LPF and HPF to create a pass-band operation focusing on the relevant frequencies for vibrotactile stimulation. The HPF components (R2, C5 and R3, C6), together with the IC input impedance, result in a lower cutoff frequency $f_{CL} \approx 10$ Hz, while the LPF (R4, C7) sets the upper cutoff frequency at $f_{CU} \approx 1000$ Hz. Pin 1 of J2 connects the IN- input to ground. ¹

The output filters (L2, C8 and L3, C9) act as electromagnetic interference suppressors. This is to prevent the class D amplifier switching frequency, generated by the PWM, and the resulting switching harmonics from reaching the PCB outputs. L2 and L3 are ferrite beads, commonly employed in class D output stages due to their wide range filtering and resistive behaviour at high frequencies.

PCB layout

Fig. 3.15 shows the PCB layout created using the KiCad tool Pcbnew. The layout can be divided to an outer- and inner frame of measures $30 \times 40$ mm and $16.5 \times 20.5$ mm ($W \times D$), respectively. The inner frame consists of the amplifier channel which can be broken off the outer frame to be mounted on the L5 VCA for the wearable solution.

The PCB layout is made of 4 copper layers. The front layer carries the components and track routing, while the back layer has a ground fill, indicated on the figure by a green barrier, to where components are grounded with through holes. An additional inner layer and microvia was required to access EN pin of the TPA2025D1 IC, which is positioned in the center of the BGA.

¹Connecting pin 1 of J2 to ground was found to be unnecessary and does in fact lead to unbalanced audio input. Since the Ferrofish A32 outputs are fully balanced (GND, OUT+, OUT-), the correct approach would have been to implement J2 as a 3-pin socket (GND, IN+, IN-), connect GND pin to ground and treat the IC audio inputs as balanced. Further discussed in chapter 5.
Fig. 3.15. Pcbnew PCB layout of one amplifier channel. Made up of 4 layers, track routing on front layer, ground fill on bottom layer and an inner layer to access EN pin of the TPA2025D1 IC.
Fig. 3.16. The 32-channel amplifier assembled in a rack mount enclosure. Two amplifiers were built for the 64-channel system.
3.7 Signal flow

Fig. 3.17 shows the flow of the signal generated by the software in development, through the in-lab multi-channel hardware system and to the tactile display in development. The PC connects to the RME MADIface XT via USB and the ASIO driver offers direct contact between the software and interface. The software can send a predefined number of audio signals (up to 64) to a matching number of software playback channels on the MADIface XT audio interface. Using a 48 kHz sampling frequency, TotalMix FX then routes the signals to a single MADI output port.

A Ferrofish A32 AD/DA converter receives up to 32 channels and forwards the excess MADI signals to a second A32. The two A32s then convert the MADI signals to line level analog audio and route them to analog D-sub25 outputs, up to 8 ports in total. The line level analog signals are transmitted to the D-sub25 multi-channel amplifier inputs and get boosted up by 20 dB. Any number of L5 VCAs matching the number of signals produced by the software in development can be connected via RCA connectors of the two multi-channel amplifiers. The VCAs are activated by amplified audio signals and deliver high definition, high resolution vibrotactile feedback to the skin.
Fig. 3.17. Signal flowchart. Software on PC transmits audio signals directly to the audio interface via USB. The interface sends MADI signals to the two DACs. The MADI signal gets converted to multi-channel line level audio and transmitted to the amplifying stage. The two amplifiers boost the signal up by 20 dB and deliver up to 64 channels of amplified analog audio signals to the tactile display.
4 Vibrotactile Threshold Experiment

Chapter 3 dealt with the design of a multi-channel hardware system, intended as an in-lab proof-of-concept system in the development of a music signal information emitting vibrotactile display. It is designed for the ACUTE group, a cross-disciplinary group at the University of Iceland with the long-term objective of enhancing the music listening enjoyment of CI recipients. In this section, a vibrotactile threshold experiment designed and carried out using the new hardware system will be covered. The goal of this experiment was to map the vibrotactile threshold curve of the wrist using the actuators of choice (L5 VCA). The chapter begins with the introduction of some of the concepts used in the experiment and then describes the experimental method, after which the results are presented and discussed.

4.1 Introduction

In psychophysics, the relationship between physical stimuli and the human perception is investigated. Such experiments treat the test subject as a black box where the stimulus is the input and the response is the output. The response will in most cases vary across individuals as countless factors, such as physical attributes or lapses of attention, can contribute to the perceived sensation. Thus, reliable results can only be accomplished by repeating the procedure with as many individuals as possible.

In order to create an effective vibrotactile stimulation device, basic psychophysical experiments are a necessary precursor. As described in section 2.3.1, the skin’s response to physical vibration varies across frequency. Furthermore, sensitivity varies across the different body parts, with the glabrous skin on the palm and fingertips being the most sensitive. In this experiment, the vibrotactile threshold of the wrist was measured using the parallel vibration L5 VCA and hardware system design described in the previous chapter. The wrist was chosen for this initial experiment due to the relatively easy implementation of a wristband prototype. In addition, measuring the threshold at both the glabrous skin on the inside of the wrist and hairy skin on the outside of the wrist will give a rough comparison of the two skin
types. Although the previous chapter describes a multi-channel device, offering up to 64 independent channels, for this experiment only one channel was needed.

4.2 Fundaments

The following sections discuss some fundamental concepts and how they relate to this experiment.

4.2.1 Decibel

The decibel is a relative unit measure expressed on a logarithmic scale. It is symboled by $dB$ and often followed by a suffix to indicate the reference value. For instance, equation (3.12) shows the boost converter gain where the reference value is the input voltage $V_g$. A reference value of 1 V will give a symbol and suffix of $dB V$, while sound pressure levels are expressed in terms of $dB SPL$ where the reference value is the acoustic pressure at average hearing threshold, or $20 \mu P$.

$dB HL$

Standard audiometry tests give the results in $dB HL$, where the reference value is the average hearing threshold of a NH individual in $dB SPL$. Thus, $0 dB HL$ is the average threshold across all audible frequencies.

$dB FS$

Digital systems use $dB FS$ as the unit of measure for amplitude level. It is an inverse scale where the reference value is the maximum amplitude level that the system can emit, thus $0 dB FS$ refers to full scale and any decrease in amplitude will result in negative measures, i.e. $-50 dB FS$.

4.2.2 Window functions

Window functions are mathematical functions widely used in signal processing. The shape of the window is usually symmetric around a maximum value in the middle,
4.2 Fundaments

tapering towards the specified interval edges and zero valued outside the interval. A signal may be windowed by multiplying it with a window function in the time domain or convolving the fast Fourier transforms (FFT) for isolating specific frequencies to analyze further. Windowing a signal introduces an artefact called spectral leakage, where the frequency content of the original signal is slightly changed.

Tukey Window

A Tukey window is very flat in the time domain, with a value of 1 for the majority of the window length. The taper length towards the interval edges is determined by the constant \( \alpha \) where \( 0 \leq \alpha \leq 1 \). At \( \alpha = 0 \) the window will be rectangular and at \( \alpha = 1 \) the window will be bell shaped. An \( N \) sample Tukey window of length \( L = N + 1 \) is described by

\[
w[n] = \begin{cases} 
\frac{1}{2} \left( 1 - \cos \left( \frac{2\pi n}{\alpha L} \right) \right) & \text{for} \quad 0 \leq n < \frac{\alpha L}{2} \\
1 & \text{for} \quad \frac{\alpha L}{2} \leq n \leq \frac{N}{2} \\
w[N - n] & \text{for} \quad 0 \leq \frac{N}{2}
\end{cases}
\]  

When preparing for this experiment, a transient was observed in the L5 response to pure tone signals, specifically at low frequencies. To counter this problem, the pure tone signals were windowed with a Tukey window as shown in Fig. 4.1. Selection of the \( \alpha \) constant was made based on subjective, qualitative analysis and found to be \( \alpha = 0.35 \) for maximum effect. Fig. 4.2 shows the FFT of the windowed signal, depicting the resulting spectral leakage for frequencies 25 - 1000 Hz. It will mirror around the 25 Hz pure tone frequency and was observed to be -4 dB at 24 and 26 Hz and -18 dB at 22 and 28 Hz. Spectral leakage was therefore deemed to be insignificant to the results of this experiment.

4.2.3 Absolute threshold

In threshold experiments two types of thresholds may be investigated; absolute thresholds (minimum detectable stimulus) and discrimination thresholds (minimum detectable difference between two stimuli). Absolute threshold can be determined via yes/no tasks or \( n \)-alternative forced choice tasks, the latter of which presents an \( n \) amount of stimuli with only one containing the target stimulus. In simple yes/no tasks however, only one stimulus is presented in each interval and the subject indicates whether it was perceived or not.

Yes/no tasks are used in normal pure-tone audiometry tests, where the subject
4 Vibrotactile Threshold Experiment

**Fig. 4.1.** A 25 Hz pure tone signal windowed with a Tukey window, $\alpha = 0.35$.

**Fig. 4.2.** Spectral leakage introduced to the 25 Hz pure tone by windowing with a Tukey window, observed as -4 dB at 25 ± 1 Hz and -18 dB at 25 ± 3 Hz and thus deemed insignificant.
4.2 Fundaments

is usually presented with tones of sound pressure levels spanning from below to above the subject’s threshold in a sequence determined by the preferred procedure. Two main classes can be used to determine the threshold, adaptive or non-adaptive procedures. Adaptive procedures are generally considered to be more concise as each successive stimulus is calculated and presented based on the previous response.

Psychometric function

By presenting the stimuli in a yes/no task across various stimulation levels, a function describing the probability \( P \) of receiving a "yes" answer can be constructed. The function is referred to as the psychometric function and is usually an S-shaped curve such as the one shown in Fig. 4.3. The psychometric function represents a subjective criterion on the performance level of an individual given a stimulus level \( x \) and can be described by a modified logistic function

\[
P(\text{yes}) = \alpha + (1 - \alpha) \left( \frac{1}{1 + e^{-k(x-m)}} \right)
\]

where \( \alpha \) is an estimated false-alarm proportion, \( k \) is the slope constant and \( m \) is the logistic midpoint. The threshold can then be defined as the level matching a predefined probability of receiving a "yes" response, i.e. \( P(\text{yes}) = 0.6 \) or 60%.

4.2.4 Maximum likelihood procedure

The maximum likelihood procedure (MLP) was proposed by David M. Green as a means for determining the absolute auditory threshold in an adaptive yes/no task. His approach utilizes a large number of hypothesized psychometric functions (hypotheses), all with the same slope and false-alarm rate but differing midpoints along the x-axis. Each tone is presented in an iterative manner. After each iteration, the response (yes or no) is used to calculate the likelihood of each hypothesis. The next presentation level is then calculated based on the hypothesis with the maximum likelihood. After the iteration process has run its course, the hypothesis with the maximum likelihood from the last iteration is used to calculate the threshold level. Further information on the calculations can be found in Green’s published work [45].

Green’s approach was embedded into the MLP MATLAB toolbox by Grassi et al. [46], a toolbox designed for auditory threshold experiments such as absolute threshold measurement. It features a GUI to set up the experiment and select the number of iterations and hypotheses to be used. The adaptive procedure runs
4 Vibrotactile Threshold Experiment

Fig. 4.3. An arbitrary psychometric function describing the probability of receiving a "yes" response in a yes/no task, given any level of stimulus [46].

through an iterative presentation of pure tones. After each iteration, the subject is asked to indicate if the tone was heard or not by pressing 1 for "yes" or 0 for "no", after which the next presentation level is calculated. When the specified number of iterations is reached, the results are automatically saved in .mat format and the process is repeated for the next tone.

In this experiment, the MLP toolbox was modified for the use of vibrotactile threshold measurements. Mainly, the standard audiometry frequency span of 125 - 8000 Hz was shifted down to 25 - 1000 Hz, the procedure run through of both left- and right channels (for both ears) was modified to only left channel, automatic conversion of dB FS to dB HL was removed and the generated pure tone signals were windowed with a Tukey window, \( \alpha = 0.35 \).

4.2.5 Masking

Masking is a phenomenon in the field of psychophysics where the presence of a stimulus affects the subject such that another present stimulus is imperceptible to the same individual. The perceptible stimulus (the masker) is then said to be masking the imperceptible stimulus. This is an especially well known occurrence in acoustics.

Masking of auditory signals has been observed in both the spectral- and temporal domains, with the masker and the target signal presented either simultaneously or at
different time intervals. If two pure tones, a masker and a target tone, are presented simultaneously, then the masking effect will only occur if the target tone lies within the critical bandwidth of the masker. Likewise, for complex signals, only the masker frequencies lying within shared critical bandwidths of the two signals will contribute to the masking of the target signal [47].

White noise

White noise is a broadband random signal of constant power spectral density. In audio, white noise signals contain all frequencies audible to the human ear, each presented with the same amount of energy. In terms of third-octave bands, where the upper-band edge frequency is the lower-band edge frequency times the cube root of two ($f(2) = \sqrt[3]{f(1)}$) and which are often used to model the critical bandwidth of the auditory filters discussed previously in section 2.1.2, this means that white noise gains 3 dB per third-octave band.

The L5 VCA used as the source of stimulus in this experiment can produce audible tones as it vibrates, especially at frequencies above 500 Hz. To prevent the audible tones from giving a false projection of somatosensation, subjects were presented with white noise through headphones over the course of the experiment to mask the audible tones. Although other noise types take the decreasing bandwidth of the auditory filters from high- to low frequencies into consideration in distributing the energy across third-octave bands, white noise was considered a good enough masker for the narrow range of frequencies used.

4.3 Method

To maximize the sensation of vibration, three different materials were tested before making the wristband prototype, a thick spongy material, soft polyester material and stiff elastic material. The stiff elastic material was chosen as it limits the material absorption of the vibration the most, based on subjective analysis. Figure 4.4 shows the wristband prototype. Velcro strips are sewn to the fabric for easy adjustment and fitting around the wrist. The L5 is attached to the center of the elastic band via double sided tape. 2 mm PETG copolyester plates are glued to each side of the L5 to prevent the moving permanent magnet and metal plates from rubbing against the skin and wristband material. Pin sockets are soldered to the coil wires and a shielded audio cable connects the L5 to the multi-channel amplifier RCA output.

Figures 4.5 and 4.6 show the experimental setup. A PC runs the experiment on
Windows OS using the modified MLP MATLAB toolbox and connects to the multi-channel system via USB. The TotalMix FX software playback- and hardware output faders were placed at 0 dB as to not provide any additional gain. Signal flow was as described in section 3.7, with MATLAB acting as the playback software.

A total of 30 participants took part in the experiment, 18 female and 12 male, aged 20 - 73. Participants sat on a chair across the table from the experimenter. All participants wore headphones and were presented with white noise to mask the audible frequencies from the L5 during the experiment. The wristband prototype was worn on participant’s dominant hand (all but one were right handed) with the arm resting on sponges to neutralize the table surface. The vibration was presented in 500 ms intervals of frequencies 25 - 1000 Hz in the following step size breakdown: 25:25:275, 300:50:450, 500:100:1000. The procedure used 20 iterations per frequency and 100 hypotheses for calculating the threshold. The first two frequencies tested served the purpose of familiarizing the participant with the process, to be presented again and the results overwritten later on in the experiment.

Participants were asked to indicate when the vibration was perceived, no matter how weak, either by verbal confirmation or by physical confirmation such as thumbs-up. Each participant went through the procedure in two trials, testing both sides of the wrist as shown in Fig. 4.6. Half the participants received the stimuli first on the inside of the wrist and the other half first on the outside. The participants were offered a short break in between trials as well as informed when each trial was
4.3 Method

Fig. 4.5. The experimental setup. A PC running the experiment connects to the multi-channel system via USB. A shielded audio cable carries the pure tone signal from the multi-channel amplifier output (a) to the L5 on the wristband prototype (b). Subjects wore headphones and were presented with white noise to mask the audible frequencies during the experiment.

(a) L5 on inside of wrist

(b) L5 on outside of wrist

Fig. 4.6. Actuator placements during the experiment. Vibrotactile threshold was measured for both the glabrous skin on inside of the wrist and hairy skin on outside of the wrist. Subjects rested their forearm on sponges, with the L5 facing up, during the procedure.
4 Vibrotactile Threshold Experiment

Fig. 4.7. Results from an experiment measuring the vibrotactile threshold at the wrist. Note that the y-axis is an inverse axis. Minimum threshold and thus maximum sensitivity is observed at 75 Hz and 125 Hz for the outside- and inside of the wrist, respectively. Optimal frequency range is observed as approx. 50 - 300 Hz.

half-way through. The experimental session lasted about an hour in total.

4.4 Results

Fig. 4.7 shows the signal amplitude at average threshold as a function of frequency for both the inside- and outside of the wrist. The graph shows that the thresholds for both sides essentially follow the same curve with sensitivity peaking at 75 Hz and 125 Hz for the outside- and inside of the wrist, respectively. As expected, the glabrous skin on the inside of the wrist shows more sensitivity to vibrotactile stimulation than the hairy skin on the outside. An approximately 10 dB FS difference in sensitivity is observed at peak and during roll-off towards higher frequencies. However, the results show little to no difference at lower frequencies (25 and 50 Hz).
4.4 Results

Fig. 4.8. Standard deviation of the threshold measurement. The blue and red curves show the average threshold values of the inside and outside of the wrist, respectively. The grey area represents the standard deviation in each case.

For the inside of the wrist, a -50 dB FS signal amplitude, or less, can be observed at what might be considered the optimal frequency range of 50 - 300 Hz. Frequencies above 500 Hz require close to full-scale signal strength to generate a response.

Fig. 4.8 shows the standard deviation of the threshold measurements. As observed, there is a significant difference in response across individuals. Some difference was observed in the results across age and gender, with older participants and males generally showing a bit higher threshold than females and younger participants. However, to be able to draw any conclusions from this statistic, the experiment would have to be repeated with a higher number of participants, more evenly distributed across age and gender.
4 Vibrotactile Threshold Experiment

Fig. 4.9. A 125 Hz signal will cause the parallel vibration L5 to accelerate both left and right during each cycle, resulting in perceived 250 Hz.

4.5 Discussion

Looking back at the previous discussion from section 2.3.1 on the skin’s mechanoreceptors and threshold experiment conducted by R.T. Verrillo [24], one might notice a factor of two relationship between the results from the two studies. Verrillo’s results showed the optimal frequency to be around 250 Hz while our results indicate that it’s closer to 125 Hz. This can be explained by referring to Fig. 4.9. Imagine the black rectangle depicted over one signal cycle as the top view boundary of the parallel vibration L5 VCA and the arrows as the direction in which it is moving. The voice coil displacement follows the AC signal applied, resulting in maximum displacement and maximum acceleration in both directions during each cycle. Maximum stimulation is therefore twice per signal cycle.

A noteworthy difference in Verrillo’s apparatus and this experimental setup is the orientation of the stimulus. While the L5 VCA vibrates left and right parallel to the skin, Verrillo’s actuators made perpendicular contact to the skin in a push and release fashion, giving a maximum push or release at polarizing signal peaks [48].
Maximum perceived stimulus is therefore once per cycle. This leads to the notion that when designing the software and tactile display, the emitted frequency from the software should always be half that of the perceptual target frequency.

Another point to consider in the software development, is that a roughly -10 dB FS attenuation might need to be applied to channels assigned to glabrous skin parts of the body. A good understanding of the amplitude dynamic range on offer was attained. The signal amplitude in dB FS at threshold shows the optimal frequency range at about 50 - 300 Hz, corresponding to perceived 100 - 600 Hz. Frequencies outside this range offer less in terms of dynamics.

Subjectivity plays a big part in the perceived vibration. Participants all share a similar threshold curve, steep towards a maximum sensitivity at around 75 - 150 Hz and a gradual reduction towards higher frequencies, but signal strength at threshold varies notably. When using the tactile display end product, users might need to go through a calibration process for maximum effect.
5 Conclusion and Future Work

A functional 64-channel hardware system for conveying vibrotactile information was designed and built. The system is intended for in-lab use in psychophysical experiments in the development of multi-channel vibrotactile encoding of music and a fully functional multi-channel vibrotactile display. However, the system is not limited for use alongside musical playback. It can be used for any kind of one way vibrotactile information stream from computer to user. The system consists of a digital audio interface, digital-to-analog converters (DAC) and custom-made multi-channel amplifiers and the information is rendered by parallel vibration voice coil actuators (VCA).

For the next version of the amplifier several design improvements are planned. It will have a fully balanced audio signal input stage to reduce noise and prevent ground loop. Building the new version will be followed by more thorough evaluation of electrical features such as THD + noise and total current consumption. Furthermore, the effect of noisy output signals on the perceived purity of the signal to the skin is subject to further investigation, in attaining the signal-to-noise ratio threshold of the skin.

The amplifier channels were designed such that they can either be placed inside the amplifier enclosure or broken off the outside frame to be placed on the L5 VCAs for wearables. Some challenging future steps will include implementing a wearable solution for the multi-channel signal handling, solved with a MADI interface and DACs in this stationary design.

The design of the tactile display is another future step. Planned follow-up experiments will be designed and carried out in order to determine the required number of actuators and their locations for maximum effect. These experiments will include discrimination tasks of both frequency and amplitude in terms of actuator spacing.

The vibrotactile threshold experiment described in chapter 4 has provided some vital information regarding future software development. Specifically, the amplitude dynamic range and optimal frequency range is now more clear. This data will be used to compensate the amplitude across channels, with each channel assigned to a specific frequency or frequency range. However, an additional step in the data presentation will be to convert the threshold values from dB FS to G, by measuring
5 Conclusion and Future Work

the acceleration of the VCA at the measured dB FS threshold values across frequency to more accurately project the user experience. This work is in preparation at the time of writing.

Standard auditory procedures can be altered to project the response of the skin. Such was the case in this experiment, where a standard audiometry test was manipulated to map the vibrotactile threshold of the wrist. More psychophysical experiments will be conducted with the multi-channel hardware system design. Among the things wished to acquire is the knowledge of the effect of actuator phase difference, vibrotactile masking and equal loudness across frequency, all of which will be used in the design of an effective assistive device for CI recipients to further enjoy their music.
Bibliography


[22] J. C. Quindlen, V. K. Lai, and V. H. Barocas, “Multiscale mechanical model of the pacinian corpuscle shows depth and anisotropy contribute to the receptor’s characteristic response to indentation,” PLOS Computational


[38] *MADIface XT*, RME, 2016, user’s guide, version 1.4.


[43] “Tpa2025d1 2-w constant output power class-d audio amplifier with class-g boost converter and battery tracking agc,” Texas Instruments, August 2011, tpa 2025D1 Datasheet, Revised December 2014.


