Simulating Hearing Loss

Leo Koop

Literature Study

TU Delft Department of Applied Mathematics

Simulating Hearing Loss

Literature Study

Leo Koop

April 9, 2015

Faculty of Electrical Engineering, Mathematics, and Computer Science (EWI) · Delft University of Technology

The work in this thesis was done in conjunction with incas^3, An incas^3 project

Abstract

The purpose of this literature review is to give a concise overview of the hearing system and present phenomena that play a key role in hearing loss simulation. Approximately 10% of the global population is to some degree affected by hearing loss, in many cases causing significant disability. Advances in modeling the cochlea, the middle ear, and auditory deafferentation now present an opportunity to simulate what a hearing disabled person hears. For this, the output of a non-linear cochlea model must be synthesized to sound. Because the non-linear nature of the cochlea model presents difficulties that have not been previously explored a linear model is examined first. Finally, this literature review also defines the scope of the master thesis *Simulating Hearing Loss* to which it serves as precursor.

Table of Contents

List of Figures

List of Tables

Preface

On the afternoon of June 26, 2001 I lived through a traumatic experience which profoundly impacted the direction my life would take. In the thirty minutes, during which time the life of my sister seemed to hang in the balance and the memory of which is distorted by adrenaline, my life went from quite, to exceedingly chaotic, and then back to a promise of normality. The calm that the emergency response personnel brought to that scene is one of the clearest memories I have of that day. On that day I knew what a difference it could make when you receive help from others. Since then I myself have served as an EMT (Emergency Medical Technician). And on a few occasions I too have been a factor for calm in someone else's traumatic thirty minutes. If anything, these experiences thought me to value life, the quality of life, and to take the opportunity to make a difference when given the chance.

On an entirely different note, I have also had the pleasure of working as a sound engineer for various events and functions over the years. Sound engineers certainly know the value of a good auditory experience. Thus the choice to work on the topic of "Simulating Hearing Loss" was a relatively easy one.

Coming to grips with all that has already been developed in this field and the difficulties related to simulating hearing loss has certainly not been easy. But it was an enriching learning experience. The complexity of the human auditory system is most fascinating to study, but a serious challenge to model accurately. The realization that hearing loss once incurred will for the most part be permanent also shed light on some personal habits that I needed to alter. I would guess that 90% of my undergraduate studying was done with ear-buds in my ear. I probably have many weeks worth of noise overexposure and I am certainly not alone in this, given today's media flooded generation. Most significantly, however, I was largely oblivious to the risk I was taking. As such, I foresee an increased need for raising awareness for the risk of hearing loss. Something that the results of this project may help to facilitate.

My work on this project would not have succeeded if not for the help from many people. Most importantly I would like to thank Peter van Hengel for answering countless questions, pointing out helpful study material, and providing valuable feedback. Kees Vuik also helped in many ways, was always available to answer questions, and most importantly asked good questions in return.

Leo Koop

Chapter 1

Introduction

Hearing loss is a significant disability and can greatly decrease in the quality of life. Early this year the World Health Organization published a news release with this stark title "1.1 Billion People at Risk of Hearing Loss"[4]. The main reason being the use of personal audio devices. Of course hearing loss is not something new, but due to the recent increase in noise exposure due to the increased prevalence of music in popular culture, the relevance of caring for hearing disabled will be catapulted to new importance. This in turn increases the importance of proper modeling of hearing and hearing damage. Accurate simulation of hearing loss could be used to aid in prevention, education, as well as faster development of improved hearing aids. Fortunately, with the advent of the computer, attempts have been made to simulate hearing loss but most fall short of truly simulating what a hearing disabled person hears. To better understand the important factors in hearing loss this paper presents the mechanisms in the auditory system, common hearing impairing damages, and modern models of the cochlea as they relate to synthesising sound from such models.

This paper is divided into sections as follows. In Chapter 2 the reader is guided through the anatomy of the human ear and explains the functions of the different components of the auditory system. Common physiological damages that lead to hearing loss are also presented. Because the models of the cochlea play such a central role in hearing simulation Chapter 3 examines both linear and nonlinear models of the cochlea. All of this information is clearly summarized in Chapter 4 which ushers the reader into past attempts at hearing loss simulation and the potential obstacles of sound synthesising in Chapter 5. In chapter 6 the research questions are given including an outlook to possible solution methods.

Chapter 2

Hearing and Hearing Loss

Of the five senses hearing is perhaps the one that contributes the most to one's situational awareness as well as our complex social interaction skills. It is therefore no surprise that any deterioration of hearing directly impacts the quality of life.

2-1 Components of the Hearing System

The main components of the auditory system are the outer ear, the middle ear, and the cochlea. The outer ear functions to "funnel" sound into the ear and is composed of the pinna and the ear canal. The middle ear transfers sound vibrations from the eardrum to the oval window of the cochlea. Within the cochlea sound is neurally encoded to be transmitted to the brain via the cochlear nerve. See Figure 2-1.

2-1-1 The Outer and Middle Ear

The outer ear, consisting of the pinna which we see and the ear canal funnel sound into our ear. The outer ear ends at the ear drum. Internal to the eardrum is the middle ear. Within the hollow cavity of the middle ear are three bones the, malleus, incus, and stapes which are also known as the hammer, anvil and stirrup respectively. Together these bones are known as the ossicles. Of these bones the malleus is connected to the eardrum and the stapes is connected to or rests on the oval window of the inner ear. These bones allow for accurate sound wave transfer from the air filled outer ear to the fluid filled inner ear. If it were not for this mechanism then much of the sound would bounce away from the ear without making the air to fluid transition. The one remaining thing that we should point out here is that there are two muscles connected to these bones the stapedius muscle and the tympani muscle. These muscles allow loud sound transfer to the inner ear to be damped when they contract in response to loud sounds. An interesting piece of trivia is that the stapes is the smallest named bone in the body and the stapedius muscle is the smallest skeletal muscle in the body.

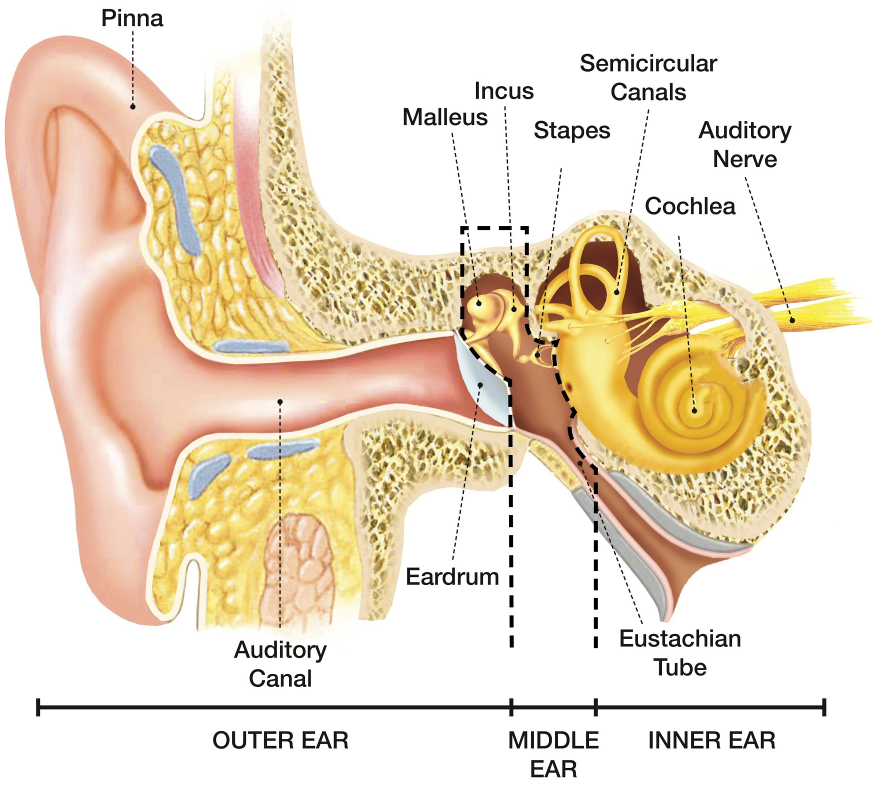

Figure 2-1: The human ear. Image source: [1]

2-1-2 The Cochlea

The inner ear consists of the vestibular system and the cochlea. The vestibular system enhances our balancing ability and is not related to hearing. Because of this little attention will be given to it. The cochlea on the other hand will take center stage for much of this research. The cochlea is a coiled tube that resembles the shape of a snail. Within the cochlea are three chambers. The superior (superior versus inferior is medical terminology for a part of the body that is above or below another part of the body when the body is in a natural standing position) chamber is the scala vestibule and the inferior chamber is the scala tympani. In-between these two is the scala media in which is the organ of Corti in which are the hair cells that sense the sound vibrations coming into the cochlea which the organ of Corti then transduces pressure waves to action potentials. These action potentials travel to the brain as electrical stimulus via the cochlear nerve. The two ends of the cochlea are referred to as the base near the middle ear and the apex at the tip of the spiral. At the base are two membranes the oval window (or fenestra ovalis) and the round window. The stapes bone from the middle ear rests on the oval window which is how sound comes to the cochlea. The scala vestibuli and the scala tympani are connected at the apex. All three chambers of the cochlea are fluid filled. Both the scala vestibuli and the scala tympani are filled with perilymph and the scala media is filled with endolymph. These fluids are mechanically equivalent and not significantly different from water. They differ however in chemical composition, creating a potential difference over the hair cells, which is essential in the creation of nerve signals from motion of the organ of Corti.

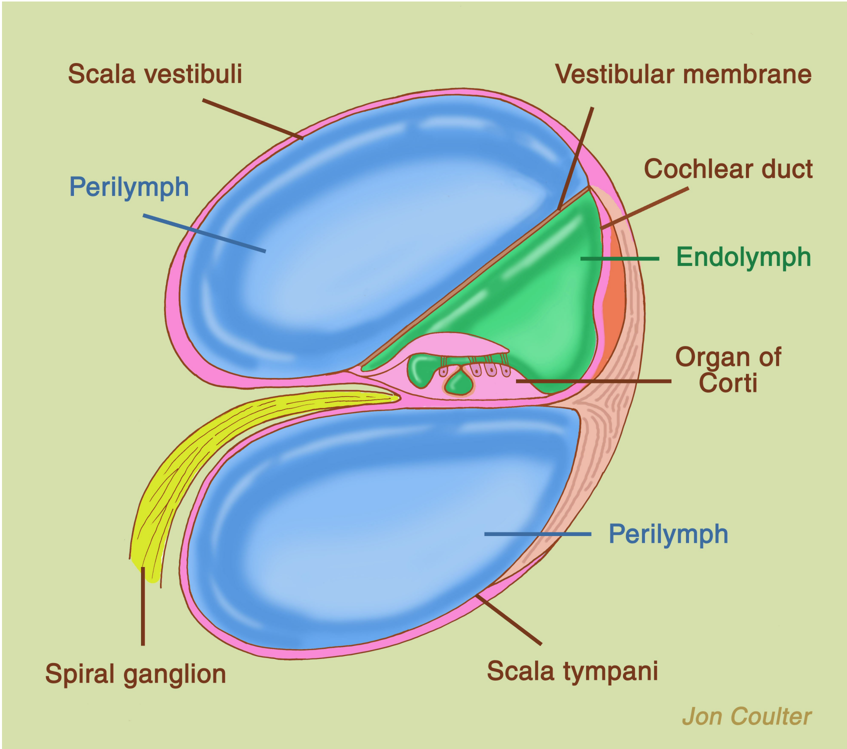

Figure 2-2: A cross section of the cochlea, Image source: [2]

The scala media is separated from the scala vestibule with the Reissner's membrane. And the basilar membrane separates the scala media from the scala tympani. These two membranes and the scala media are often referred to simply as the cochlear partition in models of the cochlea. The basilar membrane is also more frequently mentioned in the literature. When the literature cites measurements taken of the vibrations of the cochlear partition the basilar membrane is likely to be what was measured.

One should note that in the organ of corti there are a few different rows of hair cells. The row of inner hair cells are responsible for the actual electrical encoding. Three rows of outer hair cells form a feedback system that most likely allows for amplifications of certain sounds.

2-1-3 The Cochlear Nerve

There are just a few things that need to be mentioned related to the cochlear nerve. Obviously the sound signal is transferred from the cochlea to the brain via this nerve. Lopez and Barrios[5] showed that deafferentation of the auditory nerve can also be a source of hearing loss. An accurate hearing loss simulator would therefore need to take this into account.

2-2 Cochlear Emissions

From what is described until this point one may be tempted to believe that the cochlea is a relatively straight forward sound encoding organ. In truth however it is an exceedingly intricate and very finely tuned encoding organ capable of amplifying some sounds and damping others. One indicator of this are various otoaucustic emissions that were first experimentally verified by Kemp in the late 70s[6]. In essence an otoaucustic emissions is a sound that is generated from withing the ear. There are two basic types of emissions the first being evoked emissions and the second being spontaneous emissions. Interestingly when the cochlea is damaged these emissions tend to decrease or entirely disappear. This is why measuring cochlear emissions has become a technique for measuring the health of the inner ear. And this is also why cochlear emissions are important in relation to a hearing loss simulator as one would want to use the information from an emissions test to set the parameters of the hearing loss simulator. Listed below are a few Emissions that will play a role in the current study. This is however not an exhaustive list. For further information I recommend Chapter 4 of Cochlear Mechanics by Duifhuis[7].

2-2-1 Distortion Product Otoacoustic Emissions (DPOAEs)

One of the more important of the various emissions is the distortion product otoaucustic emission (DPOAE). This emission is evoked using two tones at frequencies f_1 and f_2 with $f_1 < f_2$. The result is that there are a number of emissions at different frequencies also referred to as combination tones (CTs) at frequencies $mf_1 - nf_2$ where m and n are integers and $mf_1 > nf_2$. The CT is referred to as even or odd if $m + n$ is even or odd respectively [7, p. 68]. Although there are a number of CTs the ones that are understood the best are $2f_1 - f_2$ called the "cubic" distortion tone, and $f_2 - f_1$ which is referred to as the quadratic combination tone. These are likely to play a key role in parameter adjusting in the future.

2-2-2 Spontanious Otoacoustic Emissions (SOAEs)

Unlike the case of DPOAEs some cochlear emission occur without being evoked. Hence the term Spontaneous Otoacoustic Emission (SOAE). This implies the generation of sound in the cochlea without there being an input sound. This in turn is strong evidence for a cochlear amplifier.

2-2-3 Cochlear Amplifier

Although the best method to model a cochlear amplifier is not yet too well defined, the existence of the physical amplifier and its workings are widely accepted although definitive experimental proof is still lacking. On the organ of Corti the outer hair cells form a positive electromechanical feedback mechanism that allows it to increase movement of the basilar membrane and thereby amplifying the sound.

2-3 Hearing Loss

Although there are some cases of hearing loss that have relatively simple solutions like an obstructed ear canal others are much more serious and do not have an easy fix. The severity of hearing loss is classified as mild, moderate, severe, profound, and totally deaf. Causes of hearing loss include aging, noise, chemicals, medications, illnesses, genetics, and of course trauma. Treatment depends, of course, on the condition. Some conditions benefit from surgery for example to remove scar tissue on the eardrum. Many conditions however cannot be directly treated and are alternately relieved by the use of hearing aids or even a cochlear implant. Although they may relieve the severity of the disability hearing aids and or cochlear implants hardly restore the full function of the auditory system.

2-3-1 Age

Age is perhaps the most common source of hearing loss. This is most likely due to damage or wear on the delicate structures within the cochlea such as the inner and outer hair cells that do not regenerate when damaged. This type of hearing loss may start in the late twenties to early thirties and is typically progressive. Hearing loss due to age tends affects the sensitivity to the higher frequencies.

2-3-2 Noise Exposure

Hearing loss due to noise exposure is dependent on the dB level of the noise of music and the length of time that an individual is exposed to the noise. A 3 dB increase in level constitutes a doubling of the intensity of the sound and thus will do as much damage in half the exposure time. According to the Exposure Action Value, an eight hour exposure of noise at the 85 dB level is considered safe. However a conservative estimate for a safe exposure time at a 91 dB noise level is only two hours.

Oishi and Schacht[8] place noise related hearing loss at 5% of the global population. Unlike hearing loss due to age, loss due to noise exposure affects mainly the frequencies in the 3,000 - 6,000 Hz range. Also unlike loss due to age, loss due to noise exposure is something over which an individual has much control. Proper choices related to using ear protection or limiting exposure to loud music would significantly reduce later hearing loss.

Today there is an increased awareness of this type of hearing loss as well as better education related to it. One potential use for a high quality hearing loss simulator would be in educating young people of the damages of noise exposure. Physicians, parents, or even teachers could use such a tool to give young people a good impression of what it is like to live with hearing loss, and thereby prevent such loss in the first place.

2-3-3 Inner Hair Loss

Related to the above causes of hearing loss, what really happens within the ear is that there is damage to the hair cell structures. Depending on where the damage is, there is a differing loss of function. Should there be damage to the inner hair cells then there is a direct loss to the electrical encoding mechanism. Although the BM may vibrate at the location of the given hair cell, it is unable to prompt a resulting electrical stimulus.

2-3-4 Outer Hair Loss

However in cases where the damage is related to the outer hair cells then there is a loss of functionality of the cochlear amplifier. The positive feedback brought about by the outer hair cells is reduced and this in turn decreases the sensitivity of the ear.

2-3-5 Auditory Deafferentation

Auditory deafferentation or the loss of nerve fibers is another area of hearing loss that is being studied. One very troublesome affect of hearing loss is decreased speech perception in noisy environments. A possible cause for this loss of perception is due to auditory deafferentation [9] [10] [5]. Furthermore, Lopez-Poveda and Barrios showed that this can be modeled by stochastically undersampling the sound waveform. As mentioned before, any hearing loss simulator must take this phenomenon into account.

2-3-6 Other Hearing Problems

Although simulating hearing loss is the main focus of this review and of the following thesis, hearing problems are not limited solely to hearing loss. If time permits or perhaps a future add on could be to also simulate other hearing problems.

Tinnitus

One such hearing problem is tinnitus, or ringing of the ear. In essence a person perceives a sound despite the sound not really being there. Although it has many potential causes a common cause is noise induced hearing loss. There are two classifications of tinnitus one being subjective and the other being objective. Subjective tinnitus might be quite hard to accurately simulate as this is purely subjective, and therefore there is no way to measure or test for it. Although objective tinnitus can be tested for similarly to otoaucustic emissions the frequency and level cannot be measured directly. So any simulations will be qualitative.

Hyperacusis

Hyperacusis is an over sensitivity to sound. It is possible that for an individual it will only affect certain frequencies or volumes of sounds. Some theories place the fault of hyperacusis in the brain's perception of sound. Others link it to a fault in the tensor muscle in the middle ear. Because both are not directly related to a model of the cochlea this would require additional consideration and will therefore not be given precedence in this study.

Chapter 3

Modeling the Cochlea

The main equation of the cochlear model is:

$$\frac{\partial^2 p}{\partial x^2}(x,t) - \frac{2\rho\partial^2 y}{h\partial t^2}(x,t) = 0, \quad 0 \leq x \leq L, \quad t \geq 0 \tag{3-1}$$

In this equation $y(x,t)$ is the excitation of the oscillator, ρ represents the density of the cochlear fluid, h is the height of the scala and the pressure on the cochlear partition $p(x,t)$ can be written as:

$$p(x,t) = m\ddot{y}(x,t) + d(x)\dot{y}(x,t) + s(x)y(x,t) \tag{3-2}$$

Here m is the mass per unit area. And s and d are stiffness and damping terms respectively. It should be noted here that both the stiffness and damping terms depend on the position x. Nonlinearities in the model are introduced via the parameters s and d.

Next we introduce a new term as the sum of the stiffness and damping terms as follows:

$$g(x,t) = d(x)\dot{y}(x,t) + s(x)y(x,t) \tag{3-3}$$

With that $m\ddot{y}(x,t)$ can be rewritten as:

$$m\ddot{y}(x,t) = p(x,t) - g(x,t) \tag{3-4}$$

This allows us to then rewrite Equation (3-1) as:

$$\frac{\partial^2 p}{\partial x^2}(x,t) - \kappa p(x,t) = -\kappa g(x,t) \tag{3-5}$$

Here $\kappa = \frac{2\rho}{hm}$.

Parameter	Value	Description
	Middle Ear:	
A_{st}	$3 \times 10^{-6}\text{m}^2$	Stapes area
A_{tm}	$60 \times 10^{-6}\text{m}^2$	Tympanic membrane area
ME_Q	0.4	Middle ear quality factor
ME_N	30	Middle ear transformer
	Cochlea:	
x	$0 \leq x \leq 35 \times 10^{-3}\text{m}$	Longitudinal coordinate
$\tilde{x}$	$0 \leq \tilde{x} \leq 1$	Normalized length
m	$0.5\ kg\ \text{m}^{-2}$	Membrane mass per unit area
ϵ	5×10^{-2}	Modulation factor for the impulse resonator
s_0	$10^{10}\frac{Pa}{\text{m}}$	First stiffness term
λ	300m^{-1}	Relationship between frequency and location x

Table 3-1: Parameter set of the linear cochlea model

The Equation (3-5) at a single time point is an ODE (ordinary differential equation). The approach to solve the model of the cochlea is to discretize this ODE in space and solve the system in each time step. Furthermore the Equation (3-2) can be transformed into two first order equations in time. That is we let $v(x,t) = \partial y(x,t)/\partial t$. And then introduce the following two equations:

$$\frac{\partial y}{\partial t}(x,t) = v(x,t) \tag{3-6}$$

$$\frac{\partial v}{\partial t}(x,t) = \frac{p(x,t) - g(x,t)}{m} \tag{3-7}$$

Then Equations (3-6) and (3-7) can be solved with RK4 (Classical 4th order Runge Kutta). Of course the actual solving of this system would require the inclusion of boundary conditions and some finer details of discretization.

3-1 A Linear Model

In this research two cochlear models will be used. The first being a linear and the second being a non-linear model. The main equation to be solved still remains Equation (3-1). However the specific terms and arguments change.

3-1-1 A Linear Parameter Set

For the linear model the following are the parameters to be used. The mass m is as found in Table 3-1. The stiffness term is defined as follows:

$$s(x) = s_0 e^{-\lambda x} \tag{3-8}$$

And the damping term to be used is:

$$d(x) = \epsilon\sqrt{m\ s(x)} \tag{3-9}$$

3-1-2 Shortcomings of a Linear Model

From a modeling perspective it would be nice if a linear model could capture all the necessary processes of the cochlea. If this were the case then many classical techniques of DSP (see Section 5) would be available to us. However there are a number of phenomenon that simply cannot be captured with a linear model.

Among the things that a linear model cannot account for are [7, p. 60-61]:

- Dynamic Range, that is the compression of sound at the 100 - 120 dB level to and output of 30 - 50 dB in the auditory nerve.
- Combination tones
- Cochlear Emissions

Interestingly, with damage to the cochlea the above non-linear processes tend to linearize.

Because of the shortcomings of the linear model a nonlinear model will also be considered and is in fact the main interest of this research. Solving the linear case may serve primarily as a stepping stone to solving the nonlinear case.

3-2 A Non-Linear Model

A nonlinear model then, tries to account for the phenomenon evident in the cochlea that a linear model cannot capture. Over the years a number of models have been proposed. One particular one that is regarded as a very accurate one is the one set forth here.

3-2-1 The Parameter Set used by Epp et al.

The specific nonlinear model that we will consider is the one used by Bastian Epp and Jesko L. Verhey[11], and based on work by van den Raadt and Duifhuis (1990); van Hengel et al. (1996)[12]; Mauermann et al. (1999)[13] and Talmadge et al. (1998)[14].

Unlike the linear model the damping term is not as straight forward for this model. The reason for this is that they used negative damping for low velocity /dB. This is what allows the modeling of the cochlear amplifier as well as SOAEs. The nonlinear damping term looks like this:

Parameter	Value	Description
	Middle Ear:	
A_{st}	$3 \times 10^{-6}\mathrm{m}^2$	Stapes area
A_{tm}	$60 \times 10^{-6}\mathrm{m}^2$	Tympanic membrane area
ME_Q	0.4	Middle ear quality factor
ME_N	30	Middle ear transformer
	Cochlea:	
x	$0 \leq x \leq 35 \times 10^{-3}\mathrm{m}$	Longitudinal coordinate
$\tilde{x}$	$0 \leq \tilde{x} \leq 1$	Normalized length
m	$0.375\ kg\ \mathrm{m}^{-2}$	Membrane mass per unit area
$\tau(x)$	$1.742/f_R(x)$	Delay of feedback stiffness
γ	0.12	Amplitude plateau
δ_{sat}	$0.2 \times 10^{0.52\tilde{x}}$	Damping saturation
δ_{neg}	$-0.12 \times 10^{-0.17\tilde{x}}$	Negative damping
α	40	Middle ear turning point velocity
μ_α	6	Slope at lower turning point
β	-10	Lower turning point velocity
μ_β	5	Slope at lower turning point
A	$20\ 832\ s^{-1}$	x-f map coefficient
λ	$60\ m^{-1}$	x-f map length constant
κ	$145.5\ s^{-1}$	x-f map correction
σ_{zweig}	$0.1416 \times 10^{-0.17\tilde{x}}$	Feedback stiffness amplitude
ϵ	0.01	Roughness scaling coefficient

Table 3-2: Parameter set of the nonlinear cochlea model

$$d(x,\dot{y}) = \Bigg\{(1-\gamma)(\delta_{sat}-\delta_{neg})\left[1-\frac{1}{1+e^{(\Lambda-\alpha)/\mu_\alpha}}\right] \\ +\gamma(\delta_{sat}-\delta_{neg})\left[1-\frac{1}{1+e^{\Lambda-\beta/\mu_\beta}}\right]+\delta_{neg}\Bigg\}\sqrt{m\ s(x)} \tag{3-10}$$

Where Λ is the velocity level in dB and defined by:

$$\Lambda = 20\log_{10}\left(\frac{|\dot{y}|}{y_0}\right), \quad y_0 = 10^{-6}\frac{m}{s} \tag{3-11}$$

Considering that we have negative damping below a certain threshold should immediately cause one to question stability. And this indeed would be a problem because it would make the model unstable. Therefore a delayed feedback stiffness term c is needed, one that vanishes at higher levels:

$$c(x,\dot{y}) = \Bigg\{(1-\gamma)(\sigma_{zweig})\left[\frac{1}{1+e^{(\Lambda-\alpha)/\mu_\alpha}}\right] \\ +\gamma(\sigma_{zweig})\left[\frac{1}{1+e^{\Lambda-\beta/\mu_\beta}}\right]\Bigg\} \tag{3-12}$$

With the inclusion of this c term then Equation (3-2) for a specific time point changes to:

$$p(x) = m\ddot{y}(x)+d(x,\dot{y})\dot{y}(x)+s(x)[y(x)+c(\dot{y})y(t)|_{t-\tau}] \tag{3-13}$$

It is important to note here that for the c term there is a time shift, or a delay ramp.

The reader should also realize that the information given here, related to modeling the cochlea, is limited to the information currently needed to understand the coming research. Those needing more information should consult the book *Cochlear Mechanics* by Duifhuis[7].

3-3 Model Outputs

In modeling the cochlea one can use a very simple sound like the one in Figure 3-1. Because this is a tone with one constant frequency the response within the cochlea should be seen mostly in a particular section of the CP. This can be seen in Figure 3-2. It should be noted that the response here is shown only for a particular timestep, in this case the last timestep.

A more intricate sound like one of a group of people singing would of course have a more interesting cochlear response as seen in Figure 3-3.

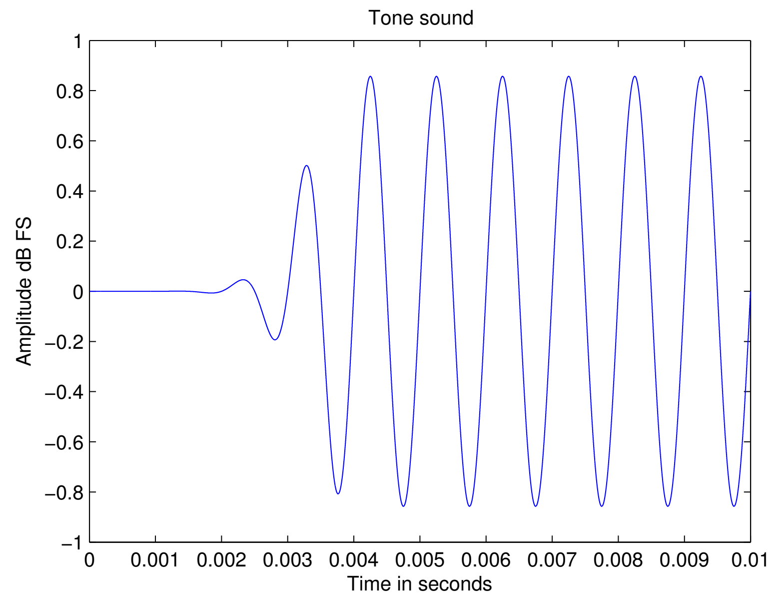

Figure 3-1: A simple tone sound wave with a gradual onset

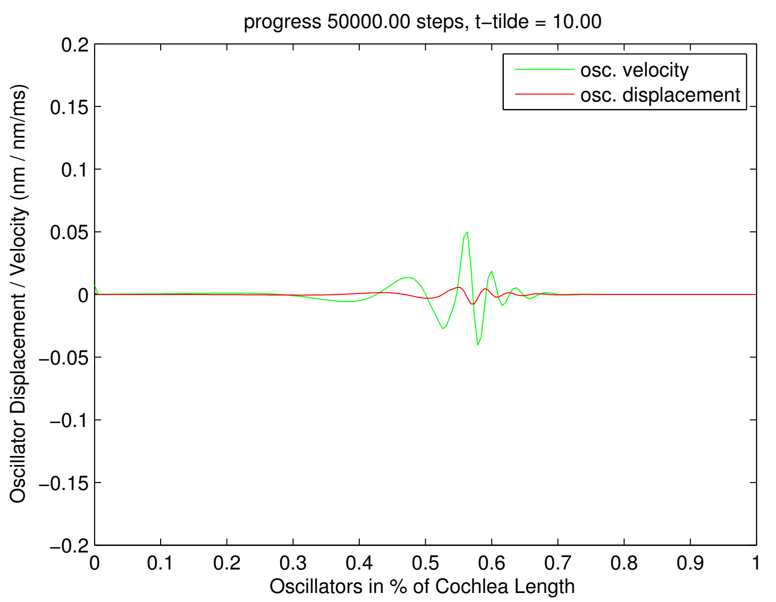

Figure 3-2: Modeling the response of the basilar membrane to the tone in Figure 3-1

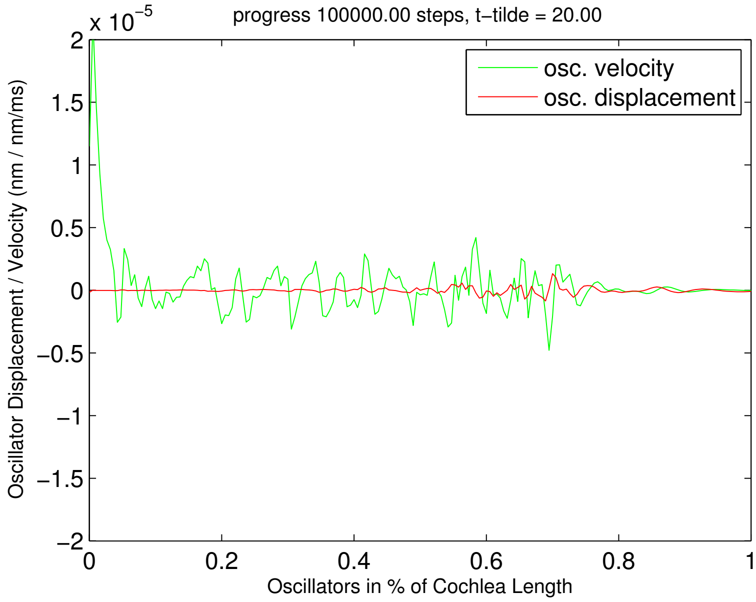

Figure 3-3: Modeling the response of the basilar membrane to a vocal sound clip

Chapter 4

An Overview

At this point the reader should have a reasonable understanding of the components of the ear, mechanisms of hearing, and some common disabilities. Models of the cochlea have also been explored as well as some of the challenges that relate to proper modeling.

In brief what happens in the physical realm is that sound is "funneled" into the the ear in the outer ear. The outer ear in an air medium ends at the eardrum. The middle ear via three bones transfers the vibrations from the air filled outer ear to the fluid filled inner ear. If not for the middle ear and the mechanical transfer between the two mediums much of the sound would not make the air to fluid transition.

The sound waves enter the cochlea via the stapes, which rests on the oval window of the cochlea. These sound waves travel into the cochlea and eventually dissipate across the CP causing it to vibrate. The tapering of the cochlea determines how far into the cochlea the wave travels before it dissipates. A high frequency wave dissipates almost immediately whereas a low frequency wave is able to travel much farther into more "shallow" fluid before it dissipates. For an illustration of the energy flow and the resulting vibrations of the CP see Figure 4-1.

In the physical ear these vibrations cause an encoding of the sound which is then transferred to and interpreted in the brain.

The model of the ear that is used takes information related to the middle ear and the pressure waves in the cochlea to simulate the movement of the BM. The encoding process is at this time not included in the model. In essence the current problem to solve is: given movement of the BM can we synthesize the (sound) signal that caused the movement.

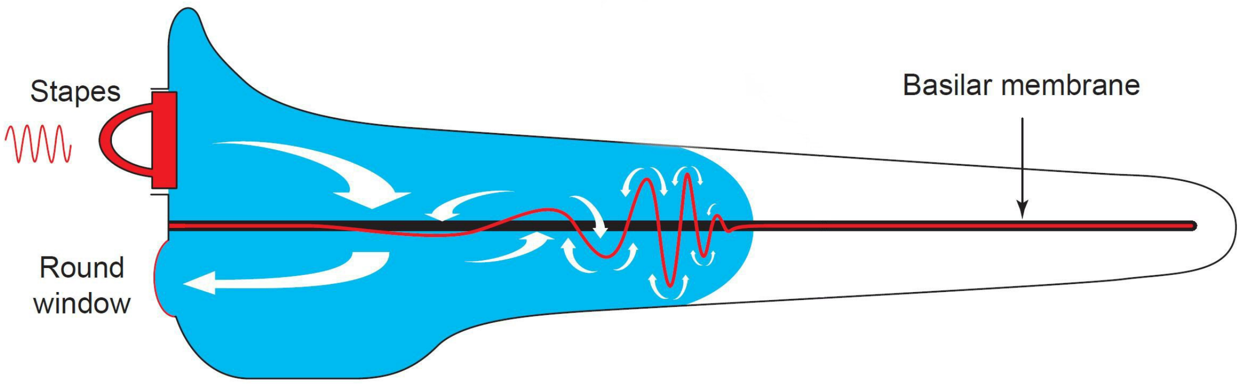

Figure 4-1: Energy flow in the cochlea and movement of the BM. Image source: [3]

Chapter 5

Digital Signal Processing as Used for Current Hearing Loss Simulation

Modifying sound, adding effects, filtering sound and many other sound processing techniques have a long and rich history. Today's music industry is build on this. Frequency filters are such a common sound processing tool that it should be no surprise that the first approach to simulate hearing loss would be to think of the frequency range that is affected and implement a simple filter to filter out or reduce those frequencies. If some frequency smearing is added then one obtains an even better simulation of what damaged hearing sounds like.

5-1 Common DSP Techniques

In Digital Signal Processing there are two common domains in which sound is processed. The first domain is the time domain. In the time domain a simple tone would look like a sine wave in time. A more complex sound with multiple tones would be simply be the sum of each tone in time.

In the Frequency Domain there is not the element of time. Rather in the frequency domain we have a representation of how much of a given sound signal lies within a given frequency.

As explained below it is often easier to work with sound in the frequency domain than it is in the time domain. Another advantage of thinking of sound in the frequency domain is that we tell the difference between sounds by their different frequency content.

5-2 Convolution and Deconvolution vs. Multiplication and Division

In many ways it is easier to edit sound in the frequency domain. Take for example a high frequency filter. If one chooses to implement this in the time domain one would use a moving average filter that iterates over all the sample points and averages each sample in comparison to the neighboring sample points. As a result of such an operation the high frequency sound waves would be smoothed or filtered out.

An alternate way to implement the same filter would be to transfer the sound signal to the frequency domain and then multiply the signal with an appropriate vector such that the desired frequency range was set to zero. The result of the multiplication can then be transferred back to the time domain where the targeted frequency will no longer be in the signal.

Note that of the two operations that were described above, the operation in the time domain falls into the category of a convolution whereas the operation in the frequency domain is a multiplication. Performing an inverse operation of the filter would be extremely difficult in the time domain as it is not so easy to find an inverse of a convolution operation.

However the inverse of a multiplication operation is simply division, something that is easy to find. This makes the frequency domain the domain of choice when an inverse operation is needed, in the upcoming project an inverse cochlea model is what we are after.

For example, if we have some convolution h of a signal x resulting in the signal y then in the time domain we have the case Equation 5-1 where the convolution h is known, but finding the inverse convolution h^{-1} if it exists is entirely dependent on the nature of the convolution

$$y(t) = x(t) * h(t) \tag{5-1}$$

However this same operation if carried out in the frequency domain where X, Y, and H are the frequency domain variants of x, y, and h respectively, might look as follows:

$$Y(f) = X(f) \cdot H(f) \tag{5-2}$$

Which allows us to easily find the definition of H and H^{-1} as:

$$H(f) = \frac{Y(f)}{X(f)} \tag{5-3}$$

$$H^{-1}(f) = \frac{1}{H(f)} \tag{5-4}$$

In our case we are interested in finding $x(t)$ given $y(t)$ and $h(t)$. One approach then is to work in the frequency domain as in Equation 5-3 in which case it is easy to find $X(f)$ as follows:

$$X(f) = \frac{Y(f)}{H(f)} \tag{5-5}$$

and then transfering X back to the time domain gives x which we are interested in. Of course in these examples it must hold that $X(f) \neq 0$ and $H(f) \neq 0$.

5-3 Current Simulators

It is not that simulating hearing loss has not been attempted before. Swiss hearing and wireless systems manufacturer Phonak[15] has a browser based simulator on their website. Phonak also founded the non-profit *Hear the World Foundation* who also have a web-based simulator[16]. The Office of Mine Safety and Health Research promotes a Windows based simulator aiming to avert hearing loss by prevention through proper education. One of the better simulators that is readily available is one done by Mark Huckvale at the University College London[17].

All of these simulators allow the user to select prerecorded background noise, and music or speech to be played in various combinations. The user can then select the level of hearing loss to be simulated. For the first three simulators mentioned, there is not as much documentation on what technique they use, but it is probably a frequency filter with potentially also a varying of amplitude sensitivity.

Of these simulators perhaps the best one is the one done by Huckvale. In this simulator the sound is edited related to amplitude sensitivity, frequency range and spectral detail. The fact that spectral detail, or frequency smearing is included in this simulator is what sets this one apart from the others that are mentioned.

5-4 The Problems with Current Simulators and DSP Techniques

The reasons why these simulators fall short of being a true scientific tool in relation to simulating hearing loss is because of the following. Firstly, they are limited to specific prerecorded sound files. Second, their purely signal processing approach cannot account for many of the phenomenon characteristic of hearing loss such as loss of cochlear emissions and the affects of the cochlear amplifier.

One reason why those phenomenon are so difficult to simulate effectively is because it is these very phenomenon that are responsible for the nonlinearity of the model. As was mentioned previously the Fourier and inverse Fourier transformation play a key role in such signal processing techniques and because the Fourier transform is suspect in a nonlinear application and alternate approach is needed for the nonlinear case.

Chapter 6

Research Question

The goal of the research project Simulating Hearing Loss is to synthesis sound from a non-linear cochlea model of a damaged cochlea. The first objective will be to synthesis sound from a linear model and then to move on to also do this for a non-linear model. Cochlear damage types that will be looked at are overexposure related damage, auditory deafferentation, and possibly also tinnitus and hyperacusis. Specific attention will be given to the minimization of sound artifacts. Furthermore, if a solution to the stated problem is found then the latency of the simulation will also be examined.

In the case of the linear model there are a number of things that will need to be considered. One approach to getting the sound from the model is to use a so called inverse filter method (see Section 6-1). For this method the CM is viewed as a sound filter. Then in a very similar way as described in Section 5-2 an inverse filter can be found and the original sound is reconstructed. With this method some things that need to be considered are the effect that the:

- Choice of oscillator(s)
- Choice of window
- Breadth of window
- Length of impulse response
- Number of oscillators in the CM

have on the quality of the sound reconstruction.

6-1 Previous Work by incas^3

Some time ago incas^3 developed methods to find which parts of the cochleogram were most likely dominated by the sound of a single (target) source, and which were not. They created

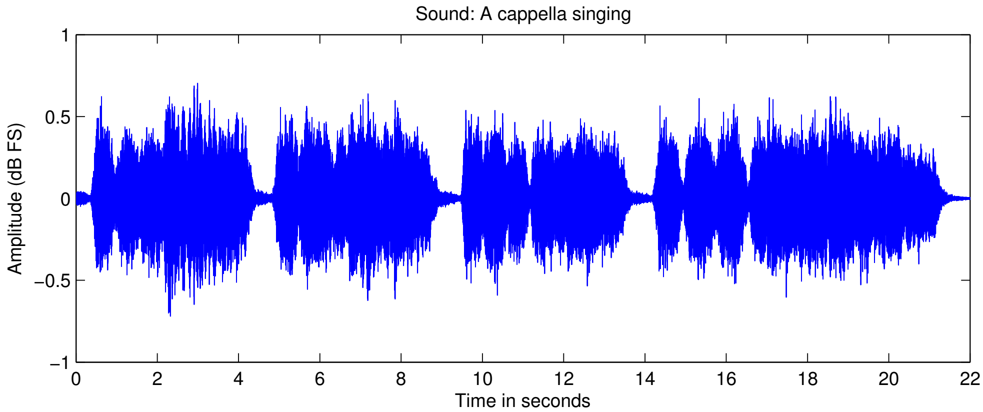

Figure 6-1: Sound profile of an A cappella sound clip

masks based on that information to be used in a speech recognition system. Because they wanted to have a way of telling how good these masks were, they tried to reconstruct an audio signal from a cochleogram multiplied by the mask. There was a problem in that they calculated the cochleogram and masks at a lower sampling frequency than the cochlea model was using. (The cochleogram gives a sort of envelope of the cochlear response at each position, which can be described with a much lower sampling rate.)

Their work attacked this problem from two slightly different angles. In one approach movement of the basilar membrane was used to calculate an impulse response and from that an inverse filter. The other approach multiplied an envelope of the cochleogram with sin functions having a frequency that matches the center frequency of the cochleogram segments. The inverse filter method in this case has a longer delay and start-up time than does the cochleogram version.

6-2 Solution Method

Because the first model is linear, traditional methods will be used where the cochleogram will be used instead of a spectrogram as a frequency representation of sound. This should allow for the inverse operation to be determined. Special care will need to be taken to circumvent the introduction of error in the areas where the cochleogram has values close to zero.

In the case of the inverse filtering method introduced in Sections 6 and 6-1 what is done is that one uses a trace from an oscillator in the CM that is a response to some sound. For example, in response to the sound clip pictured in Figure 6-1, the oscillator trace for the 3$^{\text{rd}}$ oscillator is seen in Figure 6-2. As mentioned before, this trace can be viewed as filtered sound. To get the inverse filter, one gets the trace of the same oscillator but for an impulse (see Figure 6-3) instead of a sound clip. This trace then becomes the means to find $H(f)$ and eventually $H^{-1}(f)$ and $h(t)$ or more correctly $h^{-1}(t)$ in Equations 5-1 to 5-5. Something that is not mentioned in Section 5-2 is that before converting the inverse filter $H^{-1}(f)$ to $h^{-1}(t)$ it is multiplied by a tapering window, for example a "Hanning" or a "Gaussian" window.

Initial experiments suggest that the oscillator, length of impulse response, type of window, length of window, and the number of oscillators in the CM are all interrelated in the quality of resynthesized sound. Firstly, the type of window used seems to affect both the amount

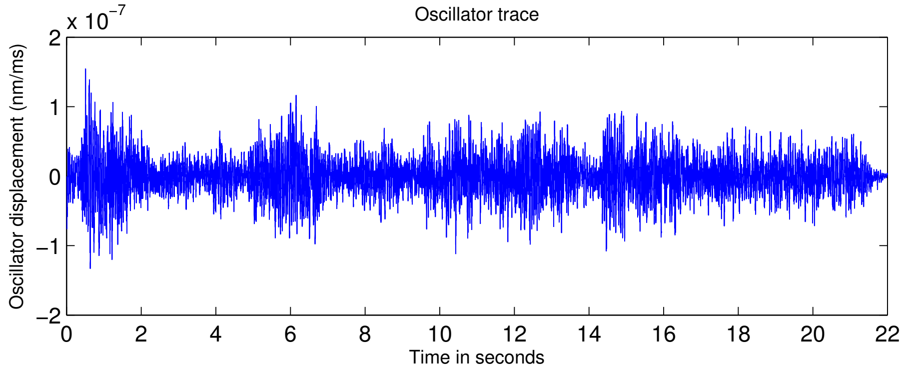

Figure 6-2: Trace of the 3rd oscillator in response to the sound pictured in Figure 6-1

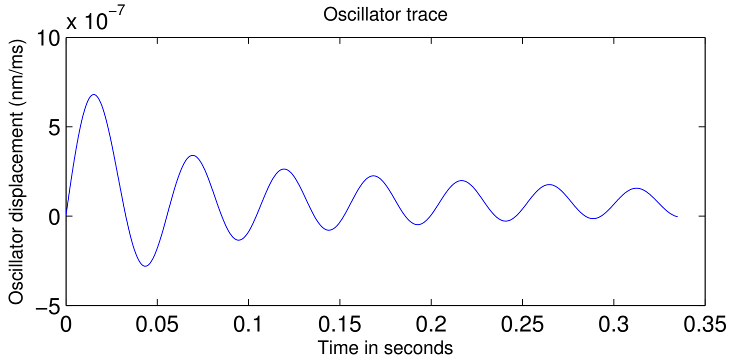

Figure 6-3: Trace of the 3rd oscillator in response to an impulse

of artifacts introduced into the reconstructed sound and also other unwanted effects. The Hanning window, for example, introduces undesired effects like reverberation. According to some preliminary testing a Gaussian window significantly reduces both unwanted effects and artifacts.

Also, for a different combination of, the number of oscillators in the model and the oscillator chosen for the trace, differing lengths of impulse responses are needed. In fact this choice seems to be very sensitive. For example, when using the 3rd oscillator and a model using 190 oscillators an impulse response, and consequently also window, of length of 1676 (approximately 0.34 seconds) seems to be a good choice. Using these parameters one gets the reconstructed or resynthesized sound seen in Figure 6-4. This looks remarkably similar to the original sound as seen in Figure 6-1. However, when using an impulse response that is only 10 samples shorter with length of 1666 samples we get a much inferior resynthesized sound (see Figure 6-5, and compare with Figure 6-4).

When using the singing sound clip reasonable sound reconstruction was possible using single oscillators up to ± 23, although there is already a noticeable decrease in quality at this point. What remains to be seen is how increasing the number of oscillators in the model would affect this. Another consideration is to use a sound clip with significantly lower frequency spectrum to see if there would then be an advantage to using later oscillators.

Initial tests using multiple oscillators vs. using a single oscillator did not increase the quality of the resynthesized sound. However, these tests were performed before the realization that choosing the right length of impulse response is so critical. This should therefore be tested

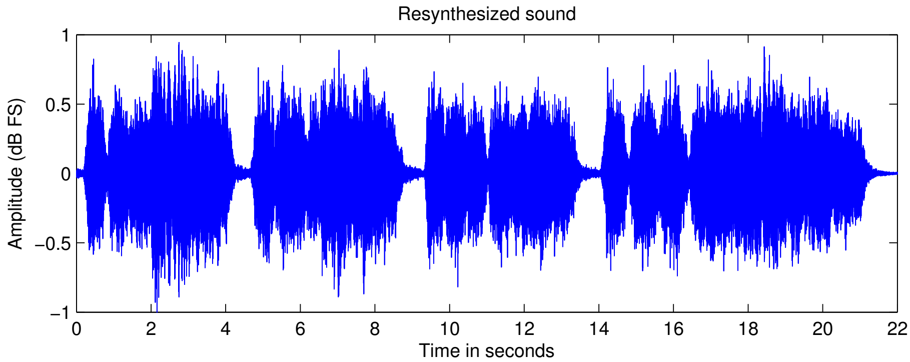

Figure 6-4: Resynthesized sound using an impulse response of length 1676

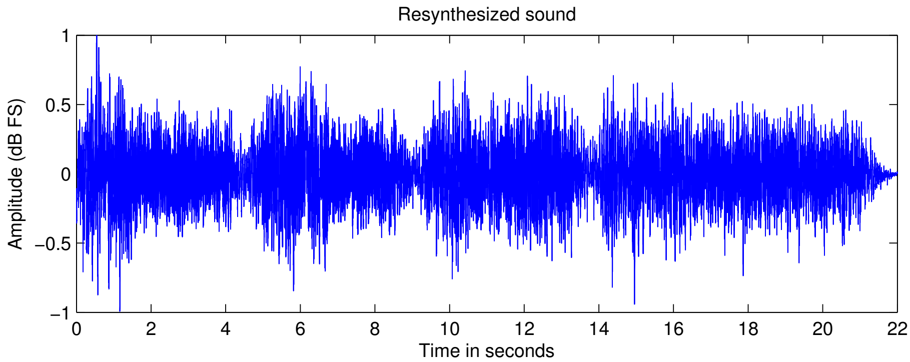

Figure 6-5: Resynthesized sound using an impulse response of length 1666

again using optimal impulse response length.

Also all the sound reconstructions up until this point seem to loose fidelity and have a quality about them as if a commercial audio processing software was used to remove noise from a recording. This may be due to one of two things, either the reduced sampling rate of the CM is insufficient to retain the fidelity, or there is a mismatch in the re-sampling procedure.

The above gives a good overview of what the solution for the linear model may look like. Solution methods for the nonlinear model have not yet been examined. It would stand to reason that the results of the linear inverse model might make for a good starting point for iterative improvement for the nonlinear model.

In their paper Lopez-Poveda and Barrios[5] simulate the neural encoding process or the neural transfer of the sound signal to the brain albeit in rather crude fashion. Physiologically this encoding happens in the cochlea. It would therefor stand to reason that perhaps the most accurate simulation of undersampled, or under encoded sound would be at the level where it happens withing the cochlea. No doubt in the simulator one wants to simulate both the affects of a damaged cochlea as it relates to the affected movement of the BM and also this potential lack of sound encoding. This implies that in the simulator one would want to simulate a stochastic under-encoding based on the results of a cochlea model of a damaged ear rather than on the original sound itself. The question would then be what the most effective or most accurate way of doing this would be. Does one first get the results from the cochlea model and then do a separate stochastic step on that data, or does one try to embed the stochastic process within the model of the cochlea.

6-3 Test Cases

The following will be taken into consideration as a test of how well sound was resynthesized. Both the results from the linear model and the nonlinear model must be evaluated according to sound accuracy evaluations. Furthermore the computation time will also be considered for both. Sound quality testing will start by trying to reproduce very basic sound(s) at first and then move to more complex sound.

Related to sound reproduction quality the following sounds may be used as test cases:

- A single tone
- A two tone signal
- A chirp
- Simple music (from a single instrument)
- Speech
- Singing
- Complex music
- Combinations of noise, music, and voice(s)

The original sound may then be compared to the re-synthesized sound by comparing the frequency spectrums. This would be especially helpful in the case of the chirp and could be used to correct re-sampling errors. Another rapid quality check is a visual inspection of the plotted sound. For example, a quick visual comparison of Figures 6-1, 6-4, and 6-5 knowing that the first is the original sound tells us that the reconstruction seen in Figure 6-4 is of much higher quality than than in Figure 6-5. Looking at Figures 6-4 and 6-5 and observing the increased error between sentences as seen after the 4^{th}, at the 9^{th}, and before the 14^{th} second suggests that summing the absolute value of the sound in such areas might provide one computational way to test the quality of the sound reconstruction. However the final quality test may need to be a subjective listening test. This may be the best way to detect effects like reverberation, artifacts, loss of fidelity, or a tonality shift.

Next, the developed algorithms will also be tested with respect to the delay, or computation time. If a hearing loss simulator was to ever be used in a live environment where pre-computing sound is not possible then the latency of the system would need to fall within certain thresholds. Any latency less than 20 ms is not detectable[18]. The average latency time were an individual will notice an audio-visual latency is 185.19 ms[18]. It should be noted that this value was an estimate for noticeable latency where audio preceded video. Because in a hearing loss simulator in a live environment sound could naturally not precede the visual stimulus the 185.18 ms latency is likely a safe estimate due to the fact that people are more tolerant to audio following visual than visual following audio.

Chapter 7

Conclusion

From this Literature review the reader should have gotten a very good overview of the aspects of the human auditory system that contribute to hearing and those that may be affected in the case of hearing loss. The components of a model of the cochlea and the simplifications that are made were also presented. The end goal of the research project is to synthesize sound from the model of a damaged cochlea. The target model to use is a nonlinear one, but the examination of a linear model may provide a stepping stone to the solution of the nonlinear case. For the linear model reasonable results have already been obtained. Aside from the quality of the resynthesized sound the computation time will also be of interest in the upcoming research. Good results from this project could greatly facilitate hearing aid testing and would also provide an invaluable tool for hearing loss prevention through education.

Bibliography

[1] G. R. Surgery, "Ear syringing."

[2] J. P. Coulter, "Sudden sensorineural hearing loss."

[3] R. Nobili, F. Mammano, and J. Ashmore, "How well do we understand the cochlea?," *Trends in Neurosciences*, vol. 21, no. 4, pp. 159 – 167, 1998.

[4] W. H. Organization, "1.1 billion people at risk of hearing loss."

[5] E. A. Lopez-Poveda and P. Barrios, "Perception of stochastically undersampled sound waveforms: A model of auditory deafferentation," *Frontiers in Neuroscience*, vol. 7, no. 124, 2013.

[6] D. T. Kemp, "Stimulated acoustic emissions from within the human auditory system," *The Journal of the Acoustical Society of America*, vol. 64, no. 5, 1978.

[7] H. Duifhuis, *Cochlear Mechanics, Introduction to a TIme Domain Analysis of the Nonlinear Cochlea.* Springer.

[8] N. Oishi and J. Schacht, "Emerging treatments for noise-induced hearing loss," *Expert Opinion on Emerging Drugs*, vol. 16, pp. 235–245, June 2011.

[9] C. A. Makary, J. Shin, S. G. Kujawa, M. C. Liberman, and S. N. Merchant, "Age-related primary cochlear neuronal degeneration in human temporal bones," *Journal of the Association for Research in Otolaryngology*, vol. 12, pp. 711–717, July 2011.

[10] S. G. Kujawa and M. C. Liberman, "Adding insult to injury: Cochlear nerve degeneration after âĂIJtemporaryâĂİ noise-induced hearing loss," *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 29, pp. 14077–14085, November 2009.

[11] B. Epp, J. L. Verhey, and M. Mauermann, "Modeling cochlear dynamics: Interrelation between cochlea mechanics and psychoacousticsa)," *The Journal of the Acoustical Society of America*, vol. 128, no. 4, 2010.

[12] P. W. J. van Hengel, H. Duifhuis, and M. P. M. G. van den Raadt, "Spatial periodicity in the cochlea: The result of interaction of spontaneous emissions?," *The Journal of the Acoustical Society of America*, vol. 99, no. 6, 1996.

[13] M. Mauermann, S. Uppenkamp, P. W. J. van Hengel, and B. Kollmeier, "Evidence for the distortion product frequency place as a source of distortion product otoacoustic emission (dpoae) fine structure in humans. i. fine structure and higher-order dpoae as a function of the frequency ratio f2/f1," *The Journal of the Acoustical Society of America*, vol. 106, no. 6, 1999.

[14] C. L. Talmadge, A. Tubis, G. R. Long, and P. Piskorski, "Modeling otoacoustic emission and hearing threshold fine structures," *The Journal of the Acoustical Society of America*, vol. 104, no. 3, 1998.

[15] Phonak, "Hearing loss simulator: Understanding mild and moderate hearing loss."

[16] H. the World Foundation, "What hearing loss sounds like."

[17] M. Huckvale, "Hearloss - hearing loss demonstrator."

[18] A. Younkin and P. Corriveau, "Determining the amount of audio-video synchronization errors perceptible to the average end-user," *Broadcasting, IEEE Transactions on*, vol. 54, pp. 623–627, Sept 2008.