An Analytic Model of the Cochlea and Functional Interpretations

by

Samiya A Alkhairy

Submitted to the Harvard-MIT Program in Health Sciences and Technology
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Biomedical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Sep 2017

Author .....
Harvard-MIT Program in Health Sciences and Technology
Aug 28, 2017

Certified by .....
Christopher A. Shera, PhD
Professor of Otolaryngology, Harvard Medical School
Thesis Supervisor

Accepted by .....
Emery N. Brown, MD, PhD
Director, Harvard-MIT Program in Health Sciences and Technology
Professor of Computational Neuroscience and Health Sciences and Technology

An Analytic Model of the Cochlea and Functional Interpretations

by

Samiya A Alkhairy

Submitted to the Harvard-MIT Program in Health Sciences and Technology
on Aug 28, 2017, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Biomedical Engineering

Abstract

The cochlea is part of the peripheral auditory system that has unique and intriguing features - for example it acts as a wave-based frequency analyzer and amplifies traveling waves. The human cochlea is particularly interesting due to its critical role in our ability to process speech. To better understand how the cochlea works, we develop a model of the mammalian cochlea. We develop the model using a mixed physical-phenomenological approach. Specifically, we utilize existing work on the physics of classical box-representations of the cochlea, as well as the behavior of recent data-derived wavenumber estimates. We provide closed-form expressions for macromechanical responses - the pressure difference across the Organ of Corti (OoC), and the OoC velocity, as well as the response characteristics - such as bandwidth and group delay. We also provide expressions for the wavenumber of the pressure traveling wave and the impedance of the OoC that underlie these macromechanical responses and are particularly important variables which provide us with information regarding how the cochlea works; they are a window to properties such as effective stiffness, positive and negative damping or amplifier profile, incremental wavelengths, gain and decay, phase and group velocities, and dispersivity. The expressions are in terms of three model constants, which can be reduced to two constants for most applications. Spatial variation is implicitly incorporated through an assumption of scaling symmetry, which relates space and frequency, and reduces the problem to a single independent dimension. We perform and discuss various tests of the model. We then exemplify a model application by determining the wavenumber and impedance from observable response characteristics. To do so, we determine closed-form expressions for the model constants in terms of the response characteristics. Then, using these expressions, along with values for human response characteristics that are available from psychoacoustic measurements or otoacoustic emissions, we determine the human wavenumber and impedance. In addition, we determine the difference in the wavenumber and impedance in the human base (where the OoC velocity responds maximally to high frequencies), and the human apex (where the OoC velocity responds maximally to low frequencies) and discuss their interpretations. The model

is primarily valid near the peak region of the traveling wave, and is linear - therefore the model, as is, does not account for cochlear nonlinearity, and hence is primarily suitable for low stimulus levels. Finally, we discuss other scientific and engineering model applications which we can pursue, as well as potential modifications to the model, including suggestions regarding incorporating nonlinearity.

Thesis Supervisor: Christopher A. Shera, PhD
Title: Professor of Otolaryngology, Harvard Medical School

Thesis Committee Chair: John J. Guinan, PhD
Title: Professor of Otolaryngology, Harvard Medical School

Thesis Reader: Dennis M. Freeman, PhD
Title: Professor of Electrical Engineering, MIT

Thesis Reader: John J. Rowoski, PhD
Title: Gudrun Larsen Eliasen and Nels Kristian Eliasen Professor of Otology and Laryngology, Harvard Medical School

Thesis Reader: Alan J. Grodzinsky, PhD
Title: Professor of Biological, Electrical, and Mechanical Engineering, MIT

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this chapter, we provide the reader with background about the mammalian cochlea. Note that this chapter and the remainder of the thesis can mostly be read independently of each other, but this chapter should be useful for providing the reader with a bridge between the real cochlea and model assumptions.

1.1 Cochlear Structure

The peripheral auditory system consists of the outer ear, middle ear ending with the stapes, and cochlea in the inner ear. The cochlea is where the auditory signal is spectrally decomposed and converted into neural signals that are then projected to the brain for further processing. The cochlea consists of fluid compartments, membranes and cells. Longitudinally, it extends from the base to the apex; its other dimensions are referred to as radial (horizontal) and transverse (vertical). It is generally stimulated by stapes motion at its base. The most important elements are as follows: The major membranes are the basilar membrane (BM), reticular lamina, and the tectorial membrane. There are fluid compartments that extend between the various membranes. Two fluid compartments between the membranes and the bony boundaries of the cochlea are referred to as the scala vestibuli and scala tympani. The compliant outer hair cells are of interest as they separate the basilar membrane and reticular lamina. They are also linked to tectorial membrane motion. These cells are also important for signal gain and nonlinearity. Note that the membranes, cells, and fluids that lie between the scala vestibuli and scala tympani fluid compartments collectively form the organ of

Corti (OoC).

1.2 Wave phenomenon

The cochlea consists of two fluid filled compartments (the scala vestibuli and tympani) separated by a multi-partition interface called the organ of Corti, which in turn contains multiple fluid compartments, membranes like the basilar and tectorial membranes, and cells like the outer hair cells and nerve cells. In the longitudinal direction, the interface is coupled *directly* through its own properties (e.g. viscoelastic properties of the membranes, inertia and viscosity of the fluid in the tunnel of Corti or subtectorial space, geometric and elastic properties of Deiter cells), as well as *indirectly*, through the surrounding fluid (e.g. inertia and viscosity of the scalae vestibuli and tympani fluid). This coupling allows for the phenomenon of traveling waves and hence propagation of energy or signals in the longitudinal directions. The cochlea lies in the inner ear, takes vibrational inputs from the middle ear, and transforms it to outputs or responses - e.g. mechanical responses such as basilar membrane velocity, as well as neural responses, such as auditory nerve fiber (ANF) firing patterns. In this chapter only, we shall use the term response for both mechanical responses, and neural responses unless explicitly stated otherwise.

More specifically, the signal from the base of the cochlea nearest the middle ear propagates longitudinally to the apex, as a result of the direct and indirect longitudinal coupling. The traveling wave propagates in dispersive media, which produces a frequency dependence in the responses (more specifically, the dispersion is such that it results in tuned responses). Furthermore, the spatial variation in the anatomic structures in the cochlea produce a spatial dependence in the responses. This, combined with the tuning, leads to the frequency-space (ω, x) dependence of the responses and the characteristic frequency map $\mathrm{CF}(x)$. Most output data measurements, are made from a specific point along the longitudinal axis, as a function of frequency. This suggests not only studying the system from its natural traveling wave perspective, but also from a transfer function (TF) view.

Longitudinally propagating waves

Radially and vertically propagating waves are generally neglected by modelers and only longitudinally propagating ones are considered. In spite of the cochlea being a very inhomogeneous system,

the internal media reflections are negligible relative to the dominant forward traveling wave which propagates from the base (where the stapes is) to the apex [1]. The primary coupling mechanism is generally thought to be contributed by the fluids in the major cochlear compartments. It is the hydrodynamic (bulk fluid flow known as 'slow wave') rather than acoustic fluid motion (compressional waves known as 'fast wave') that is of general interest to those studying cochlear mechanics.

1.3 Cochlear Motion

In response to stapes motion, the auditory signal propagates through the scalae fluid as hydrodynamic traveling waves that propagate in the longitudinal direction. This propagation is dispersive with the OoC acting as a spectral prism operating on the signal passing through the scalae fluid. In addition to dispersivity, the interplay between fluid and its flexible boundary, the OoC, also results in motion of the OoC. The interplay and dispersive nature culminate in traveling waves with higher frequencies peaking at the base, and those with lower frequencies peaking at the apex. Furthermore, for most frequencies of interest, the signal is thought to decay sufficiently before reaching the far apex, such that the reverse-traveling fluid motion (and associated OoC motion) is negligible.

Measurements of Mechanical and Neural Responses

In some mammalian species, though not human, both mechanical and neural responses can be measured in the cochlea. While measurements at different cochlear locations would be relevant, experiments are generally only able to provide results from a few longitudinal locations along the cochlea [2]. The motion or behavior at such locations along the length of the cochlea is usually measured in response to an impulse (click stimulus) or sinusoidal stimuli (tones). With stapes velocity as an input, the auditory nerve fibers transfer functions (ANF TFs) serve as the final sensory output of the cochlea and these TFs are generally assumed to approximate the cochlear contribution to psychoacoustic measurements. Transfer functions of the basilar membrane vertical

[1] Note that there are internal media reflections that form reverse traveling waves, and it is relevant to consider those when studying otoacoustic emissions (sounds generated by the ear). However, the magnitude of these reverse traveling waves is much smaller than that of the dominant forward traveling wave. The reverse traveling waves do not contribute much to 'forward phenomenon' such as tuning and transfer function measurements described below. Therefore, reverse traveling waves are generally neglected when studying cochlear mechanics.

[2] Exceptions to this include [51] and the more recent [38] that cater their measurements towards the traveling wave view

vibrations in response to stapes motion (BM TFs) are similar to these ANF TFs in terms of the general features, and are the most commonly measured mechanical responses of the cochlea. The hallmarks of hearing are frequency selectivity and sensitivity, and, indeed these two characteristics are seen in the BM TFs.

The experimentally measured responses like BM TFs, ANF TFs are usually described by quantities such as peak frequency, maximum group delay, and sharpness of tuning. We refer to these characteristics of the frequency response as 'response characteristics'.

1.4 Behavior of the Cochlea

1.4.1 Amplification

High gain is a distinct feature of the normal cochlea. From the perspective of the BM, literature suggests that the motion of the BM is amplified. This view is largely due to observations of *local* amplification. Specifically, isolated outer hair cell hair bundles force-displacement curves show regions of negative-stiffness; and more importantly, isolated outer hair cells (OHCs) show an ability to change length through electrical stimuli which suggests that its phase of motion in the cochlea *may* be consistent with positive feedback and lead to amplification of the recorded BM and RL responses [15].

Note that it is generally assumed that the key mechanical difference between alive and dead cochlea is that the OHC does not produce an active force in the dead cochlea's case. Because of this assumption regarding the absence of OHC motility, the dead cochlea is usually referred to as the passive cochlea.

1.4.2 Nonlinearity and Level Dependence

Measured mechanical responses depend on stimulus level. Specifically, as the level is increased, the TF pattern changes and the pattern of interference between waves with different properties changes [15]. Note that at very high levels, the general structure of the TFs in a *normal* alive cochlea approximates those of a dead cochlea. While this makes it appealing to equate a passive cochlea with active cochlea in response to high levels, it does not hold in interference phenomena. Specifically, waxing and waning of a click response of the BM or ANF is seen at high levels in an

alive cochlea, but this is not the case for the dead (passive) cochlea - perhaps indicating that factors other than OHC motility change upon death.

1.4.3 Apex and base

There are some differences between the apex and base both in terms of mechanics, and response characteristics. For example: studies have shown structural differences including membrane orientation; the membrane velocity response magnitude curves as a function of frequency are more symmetric about the peak in the apex, as compared to the base, in which the high frequency decay is much faster than the low frequency growth; another difference is only observable when measuring responses are multiple levels - specifically nonlinearity is much less prominent in the apex than it is in the base; the cochlear apex is important for speech processing as it analyzes low frequencies, yet it has not been studied much compared to the cochlear base (experimentally, this is the case, because of experimental difficulties in accessing the apex); lastly, the peak frequency varies more slowly as a function of distance from the stapes when compared with the base (in other words, the CF-map, which is an exponentially decaying function, has a local characteristic length that is greater for the apex than the base).

Chapter 2

Cochlear Model and Human Function

2.1 Overview

The peripheral auditory system consists of the outer ear, the middle ear, and the cochlea which is a coiled structure in the inner ear, as illustrated in figure 2.1.

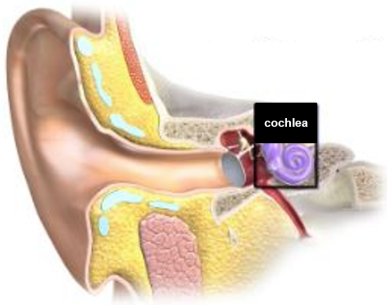

Figure 2.1: The peripheral auditory system consists of the outer ear, middle ear, and cochlea in the inner ear which is our focus. Modified from [60].

In this thesis, we focus on the cochlea. The cochlea, in which signals propagate as traveling waves, is a particularly interesting system. For example, it acts as a frequency analyzer and separates frequency components of an incoming sound along the length of its coil. This is quite relevant as signals are transmitted to the brain based on the spatial map of the frequency analyzer. Hence, we hear and perceive sounds based on the properties of this frequency analyzer. Another interesting aspect of the cochlea is a power source or amplifier: there is power supplied to the traveling wave at various frequencies and locations, and absorbed at others.

The cochlea is therefore a particularly intriguing (and useful) system, and of particular interest

is the human cochlea due to our ability to process and comprehend speech. In order to understand the cochlea - or any system for that matter, we need to develop a model. **Specifically, in this thesis, we develop an analytic model of the cochlea, and then use this model, along with observations, in order to contribute to our understanding of how the cochlea works.**

The model is for most mammals including humans. The reason this generalization is possible is because all mammals, with the exclusion of bats, have cochleae that seem to be structurally similar and operate in similar ways at a basic level [1]. Therefore, we can make interpretations regarding the function of the cochlea for any of these mammals with the exception of bats. In addition - if our entire focus is on the *human* cochlea, we can utilize data from other mammals. Mechanical measurements cannot be performed in humans, because they require surgery (and placing probes in the cochlea), but they can be performed with other mammals which provides us with the ability to test the model - an integral part of any model building process.

Over the course of this chapter, we shall take the reader through: (1) model development; (2) model test; (3) translating the model so that it can be applied; and finally (4) applying it in order for us to use the model, along with observations, to contribute to our understanding of how the cochlea works.

[1] Note, however, that there are differences between the various mammals in terms of the details - for example: the size of the cochlea, number and of hair cells and stereocilia, etc. Our note regarding similarity should be considered at a general structural and functional level. For example, all mammalian cochlea have basilar membranes, but the details (material and geometric properties of these basilar membranes) depend on the species.

2.2 Physics-Component of Model

In this chapter, we begin with the model development. First, we uncoil the cochlea that was shown in figure, 2.1. This is a classical assumption that doesn't seem to make much of a difference for the peak region of the traveling wave [47] [2].

Structure

Classical models, or classical box representations of the cochlea, have the transverse direction vertically, and the longitudinal direction, along x, as shown in figure 2.2. Longitudinally, the cochlea extends in x from the base to the apex. For a description of the derivation of the classical box representations, we refer the reader to [52, 35].

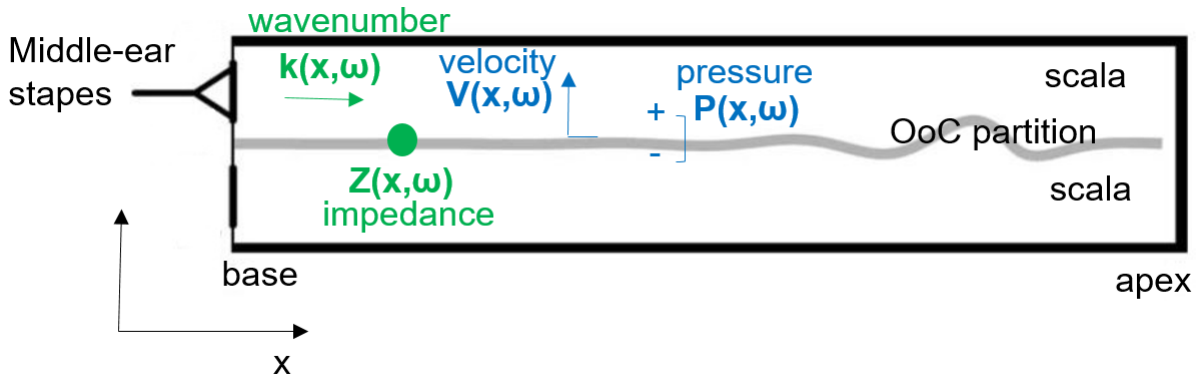

Figure 2.2: The figure illustrates the classical box representation of the cochlea. The scalae fluid compartments are separated by a single partition (with unknown properties) representing the OoC (in gray). We have annotated the figure with the transverse and longitudinal dimensions - y and x respectively, as well as the response variables for transpartition pressure P (pressure difference right across the OoC partition), and the partition velocity, V, which we have chosen to be positive in the upwards direction. The stimulus to the cochlea is specified by the stapes of the middle ear. In response, a pressure signal propagates longitudinally from the base to the apex according to the complex wavenumber, k_x, in a frequency dependent manner. The effective impedance of the OoC, Z, determines how the OoC moves in the transverse direction in response to the pressure difference across it. The figure is modified from [41].

When we hear a sound, the stapes - which is a bone in the middle ear, vibrates, and that vibration is the input to the cochlea. The cochlea itself consists of two fluid filled compartments, called the scalae, which are separated by the organ of Corti (OoC) which can be treated as a single

[2] Also, see [1, 31] for alternative views and emphasis

partition [3]. We discuss important assumptions in section J.

Motion

When we hear a sound, the stapes vibrates, and that sets into motion a pressure wave that propagates along the length of the cochlea from the base to the apex. The pressure wave interacts with the nonrigid OoC partition which has properties that depend on space and frequency. Here we will describe variables in space-frequency instead of space-time for simplicity. This implicitly requires linearity. The linearity assumption is presumed to be valid at low sound intensities [4].

This pressure propagation can be described by the following equation [5],

$$\frac{dP}{dx}(x,\omega) + ik_x(x,\omega)P(x,\omega) = 0, \text{ equivalently, } P(x,\omega) = P_m(\omega)e^{-i\int_{x_m}^{x} k_x(\tilde{x},\omega)d\tilde{x}} . \tag{2.1}$$

The main point of this equation is that the pressure is simply a traveling wave, and its properties vary with space and frequency based on the wavenumber, k_x. The wavenumber specifies how the pressure wave changes as it propagates along the length of the cochlea. In this sense, the pressure wave can be thought of as a water wave in an ocean whose amplitude and properties change depending on water depth - and consequently, on the distance from the shore.

In addition, the wavenumber dictates where the pressure waves peak. High frequencies peak in the base, and lower frequencies peak in the apex. This is relevant because the cochlea transmits information about the peaks, and where they are, to the brain and hence they are important for how we process sounds. As the pressure wave propagates along the length of the cochlea, it encounters the OoC partition which is nonrigid. Therefore, any pressure difference across the OoC partition will set it into motion. Specifically, the OoC partition will vibrate with a velocity, V, that we have defined as positive upwards in figure 2.2.

The response of the OoC partition - the velocity at which it will move at in response to a pressure

[3] The OoC partition is an effective partition, and should *not* be confused with a particular membrane - e.g. basilar membrane

[4] which are also the primary levels of interest as they concern cochlear amplification - the power source in the cochlea, which seems to have the greatest effects at low sound intensity levels [15]

[5] The pressure, P in chapter 2 is the short-wave pressure difference between the two scala right across the OoC partition - see section J for details

difference across it, is dictated by the OoC impedance, Z [6],

$$Z(x,\omega) \triangleq \frac{P(x,\omega)}{V(x,\omega)} \ . \tag{2.2}$$

In the region of the peak, which is what we are most interested in, this impedance is related to the wavenumber through the following equation, which depends on the density of the fluid and frequency.

$$k_x(x,\omega) = \frac{-2i\rho\omega}{Z(x,\omega)} \ . \tag{2.3}$$

This is true for the short-wave approximation, which is valid near the peak (which holds the most important information for the brain at the low sound levels considered here) [7].

So far, we have presented the physics relating the variables - k_x, Z, P, V together. As mentioned earlier, such equations have been previously derived - e.g. see [52, 35].

Scaling symmetry

Here, it is interesting to note an empirical observation: if we consider measurements of the velocity, V, at various points as a function of frequency, then plot them as a function of *normalized* frequency, these velocity curves will line up on top of each other to a certain extent. This is called scaling symmetry.

This normalized frequency, β,

$$\beta(x,\omega) \triangleq \frac{f}{\mathrm{CF}(x)} \ , \tag{2.4}$$

contains information regarding frequency, but also contains information regarding space, in the function $\mathrm{CF}(x)$, which is the characteristic frequency, or peak frequency at a particular location x [8].

From a modeling perspective, this empirical observation is quite useful since it allows us to consider β alone instead of having to consider both x and ω simultaneously. So, we replace our

[6] The common term for this is local or point impedance however, we avoid using this term as the impedance in this equation is truly effective - it is a result of not only local contributions but also those of neighboring regions. Therefore, we shall refer to the ratio of P to V simply as impedance.

[7] See section J for an explanation of the short-wave approximation

[8] More rigorously, the CF and the peak frequency are different, however, they are approximately the same at low intensities, and hence not of concern to us here. Empirically, the CF has been shown to follow, $\mathrm{CF}(x) = f_{max}e^{-\frac{x}{l}}$, for most regions. The value of $\mathrm{CF}(x)$, or equivalently, f_{max} and l is known for many species

wavenumber, that depends on both space and frequency, with a dependence only on this normalized frequency, β, in an attempt to make the model development process far easier.

Variables

Before we proceed to the next section, we differentiate between two groups of these variables:

1. Observable variables, or macromechanical response variables - the pressure and velocity, which are variables we can measure and observe in responses to certain stimuli
2. The wavenumber and impedance cannot be observed directly. Instead, they must be determined from observable variables.

The main points we want the reader to take from this section is that: (1) we summarized the purely physics-based relationship between the variables k_x, Z, P, V; and (2) the wavenumber and impedance fully encode the function of the cochlea in such a box representation model - in other words, the wavenumber and impedance encode how the cochlea works, which is what we are trying to contribute towards in this thesis. Our objective then, is to develop a model and determine the wavenumber and impedance of the cochlea.

2.3 Functional Aspects of Wavenumber and Impedance

Before continuing with the model development and its application, we will further discuss the specific importance of the wavenumber and impedance in this section.

Impedance

The impedance is a complex function that has a real part and an imaginary part. The real part of the impedance, $\mathrm{Re}\{Z\}$, encodes effective positive and negative damping. If $\mathrm{Re}\{Z\} > 0$, we have positive damping which means that power is absorbed in the model. Conversely, $\mathrm{Re}\{Z\} < 0$ (or negative damping) means that power is supplied to the traveling wave in the model [9]. This is particularly interesting for those studying cochlear amplification [10]. The profiles of the amplification - the frequencies and locations at which there is power amplification and absorption, is of particular interest. The imaginary part of the impedance, $\mathrm{Im}\{Z\}$, on the other hand, encodes aspects such as effective stiffness (if negative) or mass (if positive) of the OoC.

The impedance can also be used to construct a dynamic system representation of the OoC which would be quite useful [11]. This dynamic system representation could be a feedback loop, or the configuration in which the power source must occur in within this feedback loop.

Wavenumber

The wavenumber also has a real and imaginary part which contain information about how the cochlea works by specifying the properties of the pressure wave. A pressure traveling wave is illustrated in figure 2.3 as it propagates along the length of the cochlea, and we use this illustration to describe the function of the wavenumber.

The $\mathrm{Re}\{k_x\}$ encodes propagation aspects: it provides us with information such as the incremental wavelength, as well as the phase and group velocities which encode how fast the pressure wave travels as it propagates from base to apex [12]. The $\mathrm{Re}\{k_x\}$ also encodes dispersivity. In other words, the

[9] Note, however, the the *magnitude* of the real part of the impedance does not, alone, dictate the magnitude of the power inputted or absorbed

[10] There are cells within the OoC called outer hair cells, which amplify the motion of the OoC partition as it vibrates in response to a pressure difference across the OoC

[11] See section 3

[12] The reader should note that the incremental wavelength, and phase and group velocities of the cochlear pressure traveling wave take on a different meaning than the typical properties in a homogeneous system. This is due to the

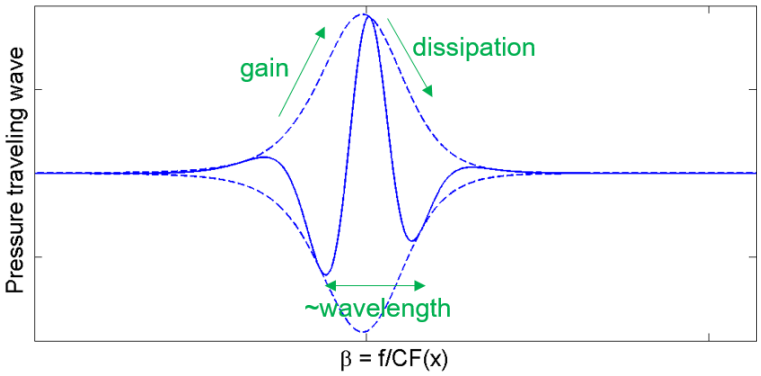

Figure 2.3: Illustration of wavenumber's functional significance: The figure shows a pressure traveling wave generated from our model as a function of β (which should be thought of here as a transformed spatial variable - for a fixed frequency). In the main text, we use this traveling wave to illustrate how the wavenumber encodes information regarding how the cochlea works. As the figure is illustrative of propagation in β rather than x, the annotated properties are of a wavenumber $k_\beta \triangleq k_x \frac{l}{\beta}$ that describes propagation in β, but is qualitatively similar to x except at extreme values of β.

wavenumber dictates how pressure waves separate based on their frequencies as they propagate along the length of the cochlea. The $\text{Im}\{k_x\}$ provides us with the gain aspects: information regarding the pressure wave growth and decay.

Given the importance of the wavenumber and impedance in encoding how the cochlea works, our goal is to determine them. More specifically, we determine k_x, Z from observations.

fact that the pressure wave changes in amplitude as it propagates along the length of the cochlea and measures such as the wavelength and phase and group velocities that we are most familiar with are defined when the amplitude is constant with respect to location (or at least the amplitude varies slowly relative to the variation of the wavelength in space)

2.4 Observations to Determine Wavenumber and Impedance

In this section, we discuss observations we can use towards our application of determining the wavenumber and impedance. As we previously mentioned in section 2.2, the macromechanical responses - the pressure, P, and velocity, V, can be measured or observed. Figure 2.4 illustrates a macromechanical response.

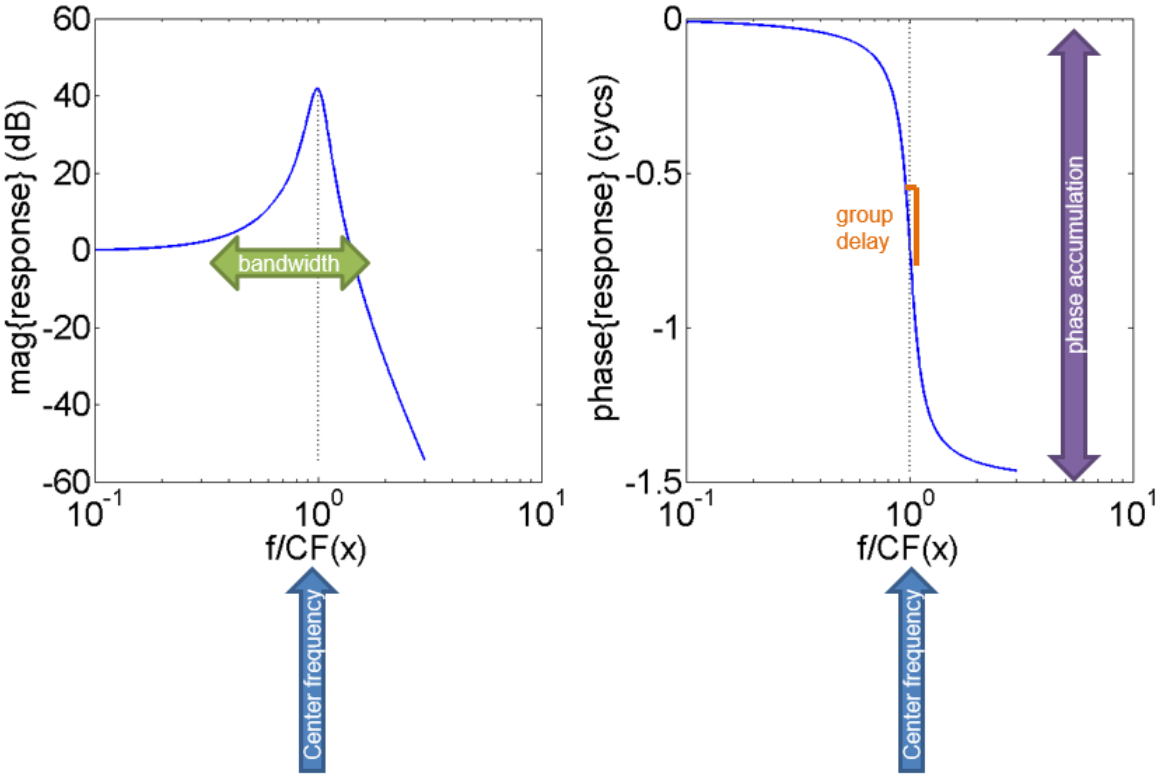

Figure 2.4: Macromechanical Responses and Response Characteristics: The figure shows a macromechanical response generated from the model that is used purely for illustrative purposes to (1) show the general shape, and (2) introduce response characteristics [13]. On the left is magnitude in dB, and on the right is phase in cycles, as a function of the normalized frequency, β. We have annotated response characteristics which we will use along with our model in order to determine the wavenumber and impedance which encode how the cochlea works.

Macromechanical Responses

Such macromechanical responses, or portions of such responses, can be determined from mechanical measurements in nonhuman mammals (these mechanical measurements cannot be performed in humans as they require surgery and probes). Certain aspects of the macromechanical responses

can be estimated using psychoacoustic studies (these fall into a broad class of studies which include the commonly used hearing tests used in the clinic, but are more sophisticated). Otoacoustic emissions, sounds that originate from the ear, can also provide us with some information about the macromechanical responses.

Issues with Macromechanical Responses

We want to use observations, along with our model, in order to determine the wavenumber and impedance which encode how the cochlea works. However, there are some noisiness and incompleteness issues with using macromechanical responses directly. Specifically, mechanical measurements are usually noisy, and more importantly cannot be performed in humans using today's technologies. Psychoacoustic measurements only provide us with half the information - the magnitude of the response [14]. Hence, they cannot provide us with a full set of information from which we can determine the wavenumber and impedance. A similar problem arises with otoacoustic emissions, as they only provide us with information regarding the maximum slope of the phase curve.

Response Characteristics

Therefore, in place of using macromechical response observations, we shall instead use features of these observations that we call response characteristics. Using the response characteristics (as opposed to the macromechanical responses) towards our goal allows us to avoid problems associated with incompleteness of information from psychoacoustic responses and otoacoustic emissions, and also enables us to determine the wavenumber and impedance in humans - from whom we cannot obtain mechanical measurements. This use of response characteristics also allows us to side-step issues related to noisiness to a certain extent, due to the fact that certain response characteristics (e.g. equivalent rectangular bandwidth, phase accumulation) can be computed in a manner that integrates information over the entire data available from the response curve, rather than using particular points.

The response characteristics from which we determine the wavenumber and impedance are listed below. These response characteristics are in the normalized frequency, β, domain [15]. The response

[14] Note that phase information can be determined from magnitude information if certain assumptions are made. For example, the magnitude and phase are related via the Hilber transform under the assumption of minimum phase.

[15] They can be very easily converted to the equivalent response characteristics in the frequency domain (which are

characteristics are also illustrated in figure 2.4 [16].

- Center frequency, β_{center}: the normalized frequency, β, at which the macromechanical responses (pressure or velocity) approximately peak
- Bandwidth, BW_n: a measure of the width of the magnitude curve of the macromechanical responses by determining where it drops from its peak value by n dB
- Maximum group delay, N: a measure of the maximum slope of the phase curve
- Phase accumulation, ϕ_{accum}: the difference between the maximum and minimum phase

typically used) as discussed in section B.1

[16] Note that the listed response characteristics are (normalized) frequency domain response characteristics related to the macromechanical responses in the frequency domain, which is what we limit ourselves to in this section. A discussion of the time domain macromechanical responses that can alternatively be used is left to section D.3.

2.5 Goal and Modeling Strategy

Our goal is to develop an analytic model of the cochlea, then use it along with response characteristics to determine the wavenumber and impedance, which encode how the cochlea works. Strictly speaking, we do not need to require that the model be analytic for us to determine the wavenumber and impedance. However, requiring that the model be analytic allows us to readily determine the wavenumber and impedance from response characteristics.

For the same reasons of achievability, we will also impose that the model expressions are not only analytic, but also closed-form expressions in space, x, and frequency, ω, or in normalized frequency, β. By closed-form expressions, we mean that we can have expressions such as algebraic, logarithmic and trigonometric functions, but we cannot have unevaluated integrals and derivatives, or special functions, such as the hypergeometric function.

We require closed-form expressions for three types of variables:

1. The wavenumber, k_x, and impedance, Z: which encode how the cochlea works, and are what we want to determine

2. The response characteristics: which we use in order to determine the wavenumber and impedance

3. Macromechanical responses - pressure, P, and velocity, V: which we do not *directly* need towards the particular application of determining the wavenumber and impedance from response characteristics. However, once we have expressions for macromechanical responses, we can then derive our expressions for response characteristics. Hence, it enables the application of determining the wavenumber and impedance from response characteristics, but is not used directly towards it [17].

We first focus on developing the model expressions for the wavenumber, impedance, and macromechanical responses in section 2.7. Then, we translate the expressions for pressure to determine our expressions for response characteristics in section 2.8, which we shall then use towards our application in section 2.9.

[17] The macromechanical expressions are, however, directly useful for other model applications such as determining the wavenumber and impedance from the macromechanical responses themselves as mentioned in section 3, and for use as auditory filters in cochlear implants, hearing aids, and compressing audio files, as explained in section I.

2.6 Phenomenological-Component of Model

In this section, we continue with the model development and determine our expressions for wavenumber, impedance, pressure and velocity. We first refer the reader back to the physical relationships between these four variables in equations 2.1 - 2.3 of section 2.2.

The equations 2.1 - 2.3 provide us with the physical relationships between P, V, k_x, Z, but do *not* provide us with expressions for any of these variables. Hence, we need an additional component that we can use along with the physical relationships between the variables in order to determine our model expressions for P, V, k_x, Z.

Specifically, we choose an expression for one of these four variables, then use it along with physical relations in order to determine expressions for the remaining variables. Previously, a wide class of classical models assumed a particular form for the impedance, Z - specifically, resonance simple harmonic oscillators (e.g. [45]). The issue with this approach is that the derived model expressions for the pressure and velocity were not closed-form expressions, and hence, it would be difficult to use to determine the wavenumber and impedance from response characteristics. More fundamentally, the velocity generated from these models is inconsistent with the velocity measured from data.

We instead construct an expression for the wavenumber, k_x, then use it along with the physical inter-relations to determine expressions for all the remaining variables. Our approach is therefore a mixed physical-phenomenological approach as it uses the physical relations and has a phenomenological component since we construct the wavenumber based on certain phenomenon.

The remainder of this section is dedicated to explaining our approach for constructing the wavenumber, k_x. We utilized published observations derived from chinchilla data, which shows the estimated real and imaginary parts of the wavenumber in figure 2.5, as a function of the normalized frequency, β [41].

To use this data-derived wavenumber in order to construct the model expressions for the wavenumber, we assume that the general features of the wavenumbers are qualitatively consistent across various species and regions we are interested in such as the base and apex. In addition to the general features derived from data, we also impose physical and mathematical constraints in constructing our wavenumber expression. For example, we want the pressure to travel purely in the

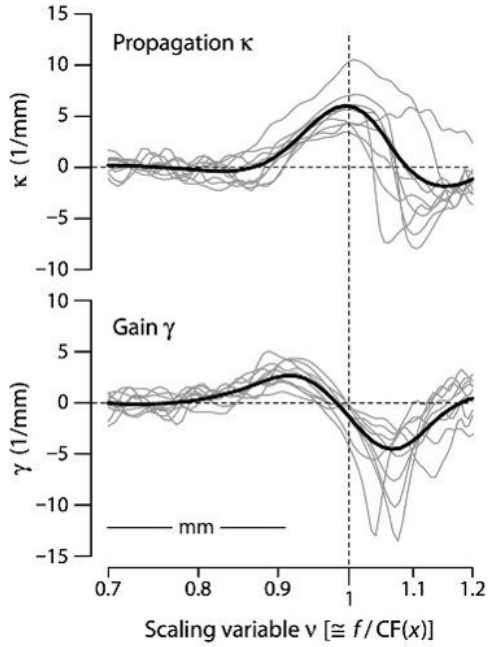

Figure 2.5: Loess trends of real ($\kappa = \mathrm{Re}\{k_x\}$) and imaginary ($\gamma = \mathrm{Im}\{k_x\}$) parts of the estimated wavenumber of [41] are in black, and individual curves from locations with CF$= 8 - 10$ kHz are in grey. The individual curves were estimated from Wiener Kernel data from auditory nerve fibers of chinchilla. The wavenumber parts are shown as a function of the scaling variable, which is equivalent to our β in the basal half of the cochlea. Notice that the zero crossing of γ and the maximum of κ occur at approximately $\beta = 1$. The figure is from [41]. The black curve (the trend line) captures the general shape and we use it to guide us in constructing an expression for the wavenumber. Curves from the apex follow a similar trend but differ quantitatively.

forward direction - from the base to the apex, and we want all our expressions to be closed-form expressions.

To construct an expression for the wavenumber, we investigated a variety of expression families, and found that our constructed wavenumber in equation 2.5 of section 2.7 is suitable as a simple expression consistent with our constraints. The next section provides the reader with our model expressions for the wavenumber, impedance, pressure and velocity, as well as comparisons with available literature.

2.7 Model Developments and Tests

2.7.1 Wavenumber

The model expression we construct for the wavenumber, k_x, is as follows. The equation below is written such that both sides are dimensionless,

$$k_x \frac{l}{\beta} = 2B_u \frac{i\beta + A_p}{(i\beta - p)(i\beta - \overline{p})} \ . \tag{2.5}$$

Notice that this expression for k_x is closed form and has a dependence on ω, x in terms of β alone, as we have assumed, which couples the dependence on x and ω [18]. For compactness, and defining poles and zeros, we shall define an independent variable, s,

$$s \triangleq i\beta \ . \tag{2.6}$$

The expression for k is a rational transfer function that has a pair of complex conjugate poles, $p = ib_p - A_p$ and $\overline{p} = -ib_p - A_p$, as well as a real zero at $s = -A_p$. The three model constants, A_p, b_p, B_u take on positive real values. The constant, l is the space constant of the cochlear map, $\mathrm{CF}(x)$, that is empirically known for a variety of species, including humans.

We note that a recent model [56], independently derived by Zweig, ended up with a wavenumber that is similar to our above expression in the sense of having a pair of complex conjugate poles, which is particularly encouraging for our model and his, as the models were developed using very different methods.

The wavenumber expression can be written using the partial fraction expansion as,

$$k_x \frac{l}{\beta} = \frac{B_u}{s - p} + \frac{B_u}{s - \overline{p}} \xrightarrow{\text{sharp-filter approx.}} \frac{B_u}{s - p} \ . \tag{2.7}$$

In the equation above, we have also provided expressions for the sharp-filter approximation of k_x, which holds when $|s - \overline{p}| \gg |s - p|$ or equivalently where A_p is small, $A_p \ll \beta + b_p$. Near the peak of V (i.e. near $\beta \approx b_p$), the sharp-filter condition is $A_p \ll 2b_p$. The sharp-filter approximation

[18] Function-wise, this intriguingly couples the inhomogeneity (spatial variation of material properties) and dispersivity in the cochlea

provides us with useful intuition due to the simple nature of the expressions.

In addition to constructing an expression for the wavenumber, we also determined expressions for the real and imaginary parts of k_x separately. When β is real, then,

$$\mathrm{Im}\{k_x\frac{l}{\beta}\} = \frac{-B_u(\beta-b_p)}{A_p^2+(\beta-b_p)^2} + \frac{-B_u(\beta+b_p)}{A_p^2+(\beta+b_p)^2} \xrightarrow{\text{sharp-filter approx.}} \frac{-B_u(\beta-b_p)}{A_p^2+(\beta-b_p)^2}\ , \tag{2.8a}$$

$$\mathrm{Re}\{k_x\frac{l}{\beta}\} = \frac{B_uA_p}{A_p^2+(\beta-b_p)^2} + \frac{B_uA_p}{A_p^2+(\beta+b_p)^2} \xrightarrow{\text{sharp-filter approx.}} \frac{B_uA_p}{A_p^2+(\beta-b_p)^2}\ . \tag{2.8b}$$

The $\mathrm{Re}\{k_x\}, \mathrm{Im}\{k_x\}$ encode certain aspects of how the cochlea works (see section 2.3). As previously mentioned, the real part of the wavenumber encodes aspects such as the wavelengths and phase and group velocities (how fast the pressure wave propagates along the length of the cochlea), as well as dispersivity. The imaginary part of the wavenumber encodes gain and dissipation.

Note that in constructing our model expression for k_x, we have used the data described in 2.6 and subjected our construction to mathematical and physical constraints. We bring the reader's attention to two constraints we imposed that intentionally result in a discrepancy between the behavior of our model wavenumber and that of the data:

- Constraint 1: The model $\mathrm{Re}\{k_x\}$ is positive everywhere to have a purely forward traveling pressure wave
- Constraint 2: The model $\mathrm{Im}\{k_x\} > 0$ prior to the peak of the traveling wave, then $\mathrm{Im}\{k_x\} < 0$ beyond the peak in order that the amplitude of the pressure traveling wave monotonically increases then monotonically decreases

Model Test

Here, we discuss the model test using the model expression for the wavenumber. We use the data-derived wavenumbers of [41] from figure 2.5 for our model test using our wavenumber expressions, as shown in figure 2.6.

We first direct the reader's attention to the real part of the wavenumber in figure 2.6 which shows that the model has a peak in $\mathrm{Re}\{k_x\}$ that occurs when the normalized frequency $\beta = 1$. This peak indicates the minimum incremental wavelength, and is a feature of the data that the model captures. The model deviates from the data in other aspects. For example, $\mathrm{Re}\{k_x\}$ in the model is

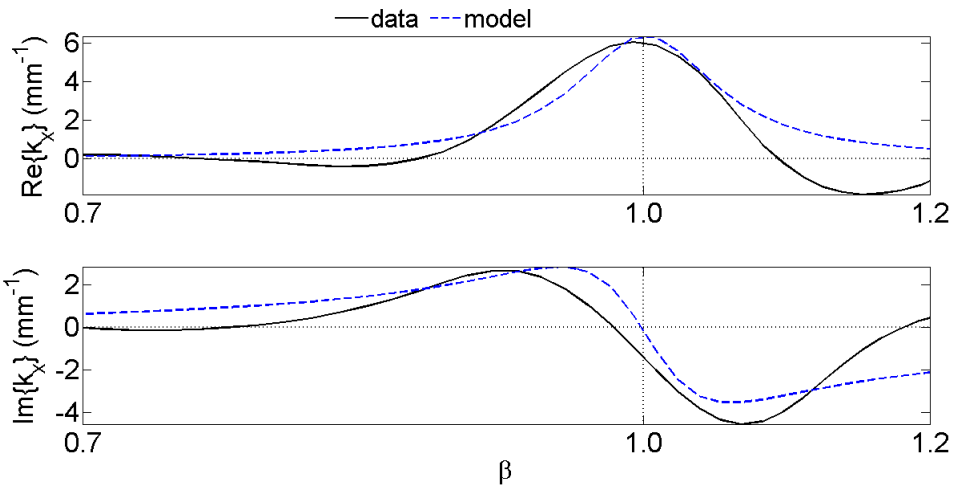

Figure 2.6: Model test using wavenumber: The solid black line reproduces data from figure 2.5 which is from [41]. The dashed blue line shows the model fits for chinchilla which have a cochlear map space constant $l = 3.8$ mm (see [41]). The top panel shows the real part of the wavenumber, which encodes propagation properties, and the bottom shows the imaginary part of the wavenumber which encodes gain properties. The $\mathrm{Re}\{k_x\}, \mathrm{Im}\{k_x\}$ are plotted as a function of the normalized frequency, β. The best-fit model constants used to construct our model curves are: $A_p = 0.05, B_u = 1.3$, with $b_p = 1$ held as a fixed constant (for a discussion regarding the value of these constants, see note K.7).

always positive, as a result of one of our constraints, whereas this is not the case with the data at higher values of β [19].

The imaginary part of the wavenumber of both the model and data is first positive. The fact that $\mathrm{Im}\{k_x\} > 0$ here indicates that the pressure wave grows as it propagates along the length of the cochlea in this region. Then $\mathrm{Im}\{k_x\}$ has a zero crossing indicating that the pressure reaches its peak. Finally, $\mathrm{Im}\{k_x\}$ becomes negative indicating that the pressure decays [20]. There is a deviation between the model and the data for larger values of β, where the model does not return to zero as quickly as the data. The model curve also does not become positive - a result of another constraint we imposed during model construction [21].

Our objective here was to show that the model captures general features of the data-derived

[19] Our constraint that the model $\mathrm{Re}\{k_x\} > 0$ everywhere is due to our assumption that the pressure wave propagates purely in the forward direction, as well as the coupling between x, ω imposed by our assumption that $k_x = k_x(\beta)$. We expect that the discrepancy in the sign of k_x (with that of [41] having a region of negative values) is due to their using transfer functions and assuming traveling wave scaling symmetry of velocity. As a result, the sign of their estimated $\mathrm{Re}\{k_x\}$ should *not* be interpreted as a reversal in traveling wave direction.

[20] The peak and the troughs of $\mathrm{Im}\{k_x\}$ indicate where the maximum slopes are in the pressure magnitude, and hence (due to our assumption that P is scaling symmetric), related to the sharpness of tuning

[21] Specifically, we imposed that the pressure grows ($\mathrm{Im}\{k_x\} > 0$), then has a peak ($\mathrm{Im}\{k_x\} = 0$), then decays ($\mathrm{Im}\{k_x\} < 0$) and does not grow again (remains $\mathrm{Im}\{k_x\} < 0$).

wavenumbers, which is critical towards using our model for the purposes of determining the wavenumber and impedance, which encode how the cochlea works, from response characteristics. For the remainder of this section, we shall introduce our model expressions for Z, P, V, and interleave those expressions with model tests, to gradually build the reader's familiarity and confidence in the model for the exemplifying application of determining the wavenumber and impedance from response characteristics.

2.7.2 Impedance

We use our constructed expression for the wavenumber in section 2.7.1, along with the physical inter-relations of section 2.2 in order to derive our expression for the impedance, which also encodes how the model works.

Our expression for the normalized impedance,

$$\frac{Z}{2\pi\rho l \mathrm{CF}(x)} = \frac{-i}{B_u}\frac{(s-p)(s-\overline{p})}{s+A_p} \xrightarrow{\text{sharp-filter approx.}} \frac{-2i}{B_u}(s-p)\ , \tag{2.9}$$

is also a closed-form expression, and is in terms of the same three model constants, A_p, b_p, B_u as the wavenumber - which is critical for our application as the model must behave as a single unit. The sharp-filter approximation for Z is derived by using the sharp-filter approximation of k_x [22].

The real and the imaginary parts of the impedance are expressed separately as,

$$\mathrm{Re}\{\frac{Z}{2\pi\rho l \mathrm{CF}(x)}\} = \frac{\beta}{B_u}\frac{\beta^2 + (A_p^2 - b_p^2)}{A_p^2+\beta^2} \xrightarrow{\text{sharp-filter approx.}} \frac{2}{B_u}(\beta - b_p) \tag{2.10a}$$

$$\mathrm{Im}\{\frac{Z}{2\pi\rho l \mathrm{CF}(x)}\} = -\frac{A_p}{B_u}\frac{\beta^2 + (A_p^2 + b_p^2)}{A_p^2+\beta^2} \xrightarrow{\text{sharp-filter approx.}} -\frac{2A_p}{B_u}\ . \tag{2.10b}$$

Near the peak of the macromechanical responses (where the sharp-filter approximation is valid), the impedance real and imaginary parts at a particular point, x, have particularly simple expressions: the real part is a line with a zero crossing and a positive slope - it is first negative then positive; the imaginary part is a constant independent of frequency, and more specifically a negative constant.

The simplicity of the above expressions near the peak provides us with a certain degree of immediate intuition regarding how the model works. Specifically, as we previously mentioned that the real part of the impedance indicates positive or negative effective damping. When it is positive, there is power absorption, and where it is negative, there is power amplification - i.e. power is inputted to the pressure traveling wave at those frequencies. The fact that the real part of the impedance near the peak is a straight line that crosses zero, means that we have a region of both, separated by the $\beta = b_p$ (which is close to the peak of the macromechanical responses), that can be visualized simply. The imaginary part of the normalized impedance is negative, which means that

[22] as opposed to taking the limit of the full expression for Z

the OoC acts as an effective stiffness rather than an effective mass, and there is no 'resonance' [23].

Model Test

Here we test the model using the impedance expressions. Figure 2.7 shows impedance estimated from data [12] from guinea pigs, as well as impedances from our model for comparison.

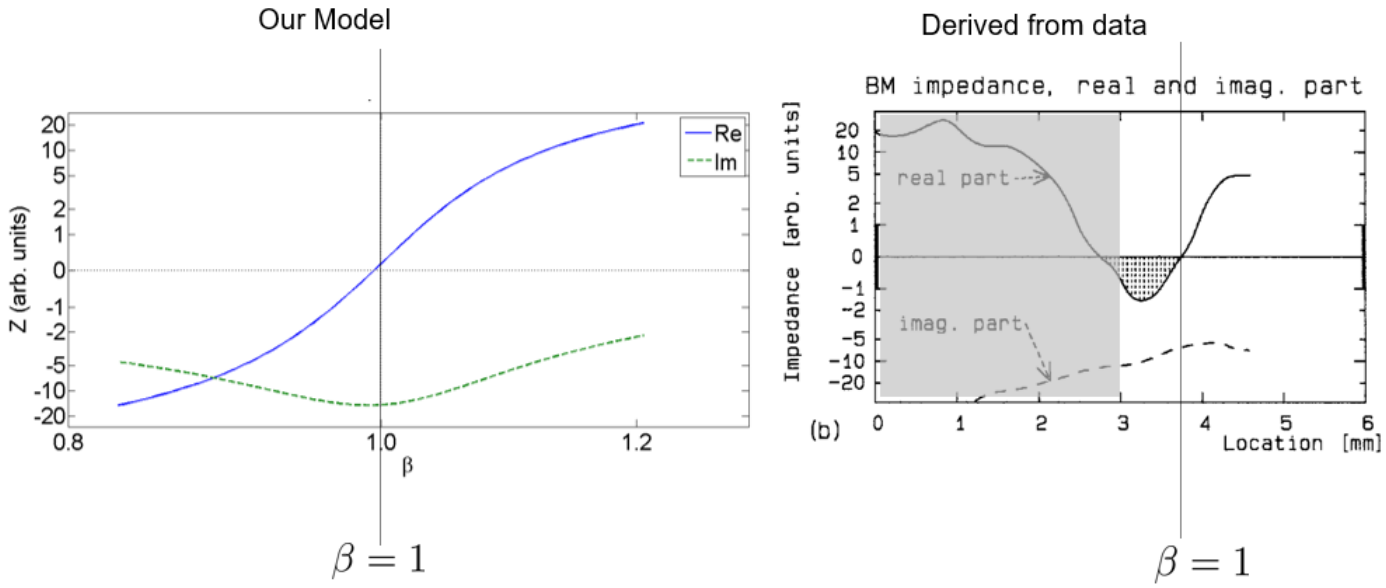

Figure 2.7: Model test using impedance expression: The right panel shows impedances derived from guinea pigs - adapted from [12]. The solid line is the real part of the impedance and the dashed line is the imaginary part. The authors have plotted the curves as a function of location, but it is more appropriate to consider them as a function of normalized frequency, β. We have grayed-out the region where the model used by [12] is not valid, and have annotated where $\beta = 1$ occurs. The left panel shows our model's real and imaginary parts of the impedance as a function of normalized frequency, β. The model constants used to construct our model curves are: $A_p = 0.1, B_u = 3, b_p = 1$ (note that our model is *not* fitted to the curves on the right). The arbitrary units are defined by [12] such that 1 in arbitrary units corresponds to 2000 $\frac{kg}{m^2.s}$.

The real parts from both our model and the referenced estimates are both first negative, indicating negative damping (power supplied to the traveling wave at those frequencies [24]). Then the curves cross zero, and become positive, indicating power absorption at those frequencies.

The imaginary parts is always negative indicating a stiffness dominated, as opposed to a mass dominated, OoC. This is the case for both our model, and the curve derived from data. There are

[23] The fact that it is a constant means that we can treat the effective stiffness as a true stiffness that is independent of frequency. However, the reader must note that the stiffness is still effective in the sense that it a sum of various properties of the OoC

[24] assuming the physics of section 2.2

differences in the shape of the curves, but the model captures the general features [25]. It is also appropriate to compare our model impedance, as well as the impedance derived in [12] with other studies, such as [14] and [56] - these share most of the general features.

So, far, we have provided the reader with expressions for the wavenumber and impedance which encode how the model works. Next, we provide the reader with the model expressions for the macromechanical responses - velocity and pressure, which we later use to determine expressions for response characteristics which, in turn, we then use towards our application (as it is using the response characteristics that we determine the values of the wavenumber and impedance).

[25] The reader should note that other estimates of the impedance yield different shapes, including [14] and [56]

2.7.3 Pressure

Here, we use the expression we constructed for the wavenumber, along with the physical interrelations in order to derive the model expression for pressure, P. We first rewrite equation 2.1 as,

$$\frac{d\log(P)}{d\beta} + i\frac{k_x l}{\beta} = 0 \,, \tag{2.11}$$

in order to simplify the problem [26]. The equation has an associated boundary condition or an integration constant discussed in section A [27].

This yields,

$$P = C\bigg((s-p)(s-\overline{p})\bigg)^{-B_u} \xrightarrow{\text{sharp-filter approx.}} \tilde{C}(s-p)^{-B_u} \,, \tag{2.12}$$

which is a closed-form expression in terms of the same model constants, A_p, b_p, B_u. The constants, $C, \tilde{C}$ are not determinable in our model, and hence the model pressure must always be defined in reference to some value.

The magnitude and phase can individually be expressed as,

$$\begin{aligned}\frac{\log(10)}{20}\text{mag}\{\frac{P}{P_{ref}}\} &= -\frac{B_u}{2}\log\left(A_p^2 + (\beta - b_p)^2\right) - \frac{B_u}{2}\log\left(A_p^2 + (\beta + b_p)^2\right) \\ &\xrightarrow[\text{approx.}]{\text{sharp-filter}} -\frac{B_u}{2}\log\left(A_p^2 + (\beta - b_p)^2\right) \,, \end{aligned} \tag{2.13a}$$

$$\begin{aligned}\text{phase}\{\frac{P}{P_{ref}}\} &= -iB_u\tan^{-1}\left(\frac{\beta - b_p}{A_p}\right) - iB_u\tan^{-1}\left(\frac{\beta + b_p}{A_p}\right) \\ &\xrightarrow[\text{approx.}]{\text{sharp-filter}} -iB_u\tan^{-1}\left(\frac{\beta - b_p}{A_p}\right) \,. \end{aligned} \tag{2.13b}$$

The individual magnitude and phase expressions for the pressure are of utmost importance. These

[26] It is very relevant to consider $k_\beta \triangleq \frac{k_x l}{\beta}$ here which is the wavenumber that determines the properties of the pressure wave as it propagates along, β (a transformed space), rather than k_x itself: The $\log(P)$ has a real part which is proportional to the magnitude of P in decibles, and can be determined separately by integrating $\text{Re}\{k_\beta\}$; $\text{Im}\{\log(P)\}$ is the phase in radians, and can be determined separately by integrating $\text{Re}\{k_\beta\}$. Equivalently, the real and imaginary parts of the wavenumber can be determined from the magnitude and phase of P. For example $\text{Im}\{k_\beta\} = \frac{d}{d\beta}\text{Im}\{\log(P)\}$ provides us with another phenomenological route through which we can impose the general shape of $\text{Im}\{k_x\}$. Indeed, we can utilize these relationship between the real and imaginary parts of the wavenumber and the derivatives of the magnitude and the phase of the pressure respectively, along with the qualitative behavior of recent pressure measurements [14], to construct the model wavenumber expression based on an alternative phenomenological approach to section 2.6. Both the aforementioned approaches and the one we explained in section section 2.6 lead to the same expected behavior of, and constraints on, k_β (or k_x), and hence, our constructed expression for k_x is suitable in either case. The fact that the model is the same regardless of our approach for developing it is a strong support for the model. As a side note, this fact also provides support for the [41] study. We shall use k_β in the appendices

[27] Also, alternative expressions for P, V using an alternative boundary condition are left to section A.5

two expressions are particularly useful because we need them to determine our expressions for the response characteristics (which we directly need towards our goal of determining the wavenumber and impedance from response characteristics) [28]. For example, we can determine our expression for the bandwidth using mag$\{P\}$, and we can determine our expressions for phase accumulation and group delay using our expression for phase$\{P\}$.

Additional information regarding P, including its time-domain representation is left to section D. We will not test our model using the pressure expressions here, nor is it necessary. This is because the model variables are all inter-related through the physics and hence a test of any of the model expressions is a test of the entire model. Additional model tests are left to section A.

[28] Strictly speaking, we could use these two expressions or alternatively the real and imaginary parts of k_β. This is due to the direct and individual relation between the magnitude and phase of P and the real and imaginary parts of k_β (or k_x) respectively. See section B.1 for details on the derivation of the response characteristics using the model expressions

2.7.4 Velocity

Here we use the expression we constructed for the wavenumber, along with the physical inter-relations in order to derive the model expression for velocity, V,

$$\frac{V}{P_{ref}} = \frac{iB_u}{\rho l \mathrm{CF}(x)} \frac{s + A_p}{\left((s-p)(s-\overline{p})\right)^{B_u+1}} \ . \tag{2.14}$$

The velocity normalized to its CF value (approximately its peak value) at a particular location from a transfer function perspective (which is the most commonly performed mechanical measurement) is,

$$\mathcal{V}(x,\omega) = \frac{s + A_p}{i + A_p} \left(\frac{(s-p)(s-\overline{p})}{(i-p)(i-\overline{p})} \right)^{-(B_u+1)} \ . \tag{2.15}$$

The above equations show closed-form expressions in terms of the same set of model constants, A_p, b_p, B_u.

Additional notes regarding V are left to section D.

Note on the Form

It is interesting to note that the velocity expression is similar in terms of having a pair of complex conjugate poles raised to a power to a class of auditory filter models - the gammatone filter family [19] [29]. This is encouraging and also quite useful since the gammatone filters are widely used, and they can be applied towards cochlear implants, hearing aids, and audio-engineering purposes. Hence, it is possible to use our model for the same applications as the gammatone filters, while noting that our model has an additional component to it - that of representing the associated wavenumber and impedance which encode functional aspects that underlie the macromechanical responses.

[29] Details regarding this comparison and leveraging it towards applying our model for other applications are left to section H

Model Test

Mechanical Data Here we test our model using the velocity expression. Figure 2.8 shows the model fit to data from mechanical measurements [30]. The method of fitting is left to section A.4 [31].

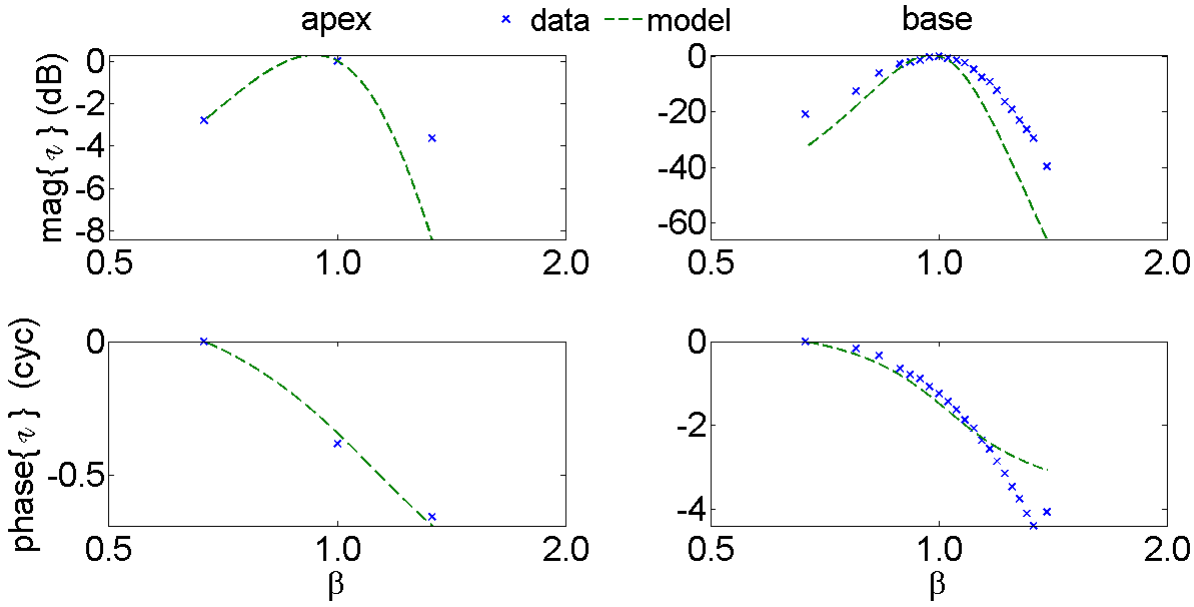

Figure 2.8: Model test using velocity expression I: The model (green dashed) expressions for velocity response is fit to measured data (blue crosses) in the transfer function domain for datasets from [6] (apex, left) and [39] (base, right) in the region of the peak. The magnitude (top) and phase (bottom) are plotted as a function of normalized frequency β. The model constant b_p is fixed at $b_p = 1$ for both fits. For the fit to [6] data, the estimated model constant values for A_p, and B_u are 0.5, and 2.7 respectively, and the objective function value is 0.23. For the fit to [39] data, the estimated model constant values for A_p, and B_u are 0.24, and 8.7 respectively, and the objective function value is 0.25. See K.6 for notes on these values for model constants

We find that the model fits the data best near the peak of the magnitude curve [32]. The model

[30] The measurements are of velocities of membranes in the OoC relative to the stimuli at a particular location, as a function of frequency around the peak. The data from [6] is of velocity of the lateral margin of the tectorial membrane within the OoC from a region in the apex of a chinchilla which has a CF of approximately 0.3 kHz - note that [6] lists 0.5 kHz as the CF instead, but we find the peak to occur closer to 0.3 kHz. The data from [39] is of the basilar membrane within the OoC from a region in the base in a different chinchilla which has a CF of about 9.5 kHz. Note that, to first approximation, we have assumed that the various membranes within the OoC move as a single partition, based on classical assumptions.

[31] Qualitatively, the reported data generally have the same behavior as our P, V. An exception to this is the existence of a positive phase gradient at frequencies beyond the peak frequency in the reported data. We expect that this is due to our model construction process which assumes purely forward traveling waves for P (which dominates the shape of V), and fundamentally ties together the traveling wave and transfer function perspectives through using β. Hence in our model, the gradient of the pressure phase with respect to frequency must have the same sign as its gradient with respect to location, whereas this is not necessary true when we go beyond our assumption that $P = P(\beta)$.

[32] This is due to our focus on the peak during model construction, but it is also partly due to our choice of objective function that we minimize - see section A

captures the slope of the phase data which is an important feature. The fact that the fit is better near the peak even in the apex is despite the sparsity of data there.

Neural Data Given the sparsity in mechanical data, it is relevant to use an additional dataset to test the model. In figure 2.9 we fit the model expression for velocity to data recorded from chinchilla auditory nerve fibers [39]. [33]. These are not mechanical measurements, but are assumed to approximate the velocity of the OoC [34]. This figure also shows that the model fits to the peak region is better, and that the model is able to predict the slope of the phase curves.

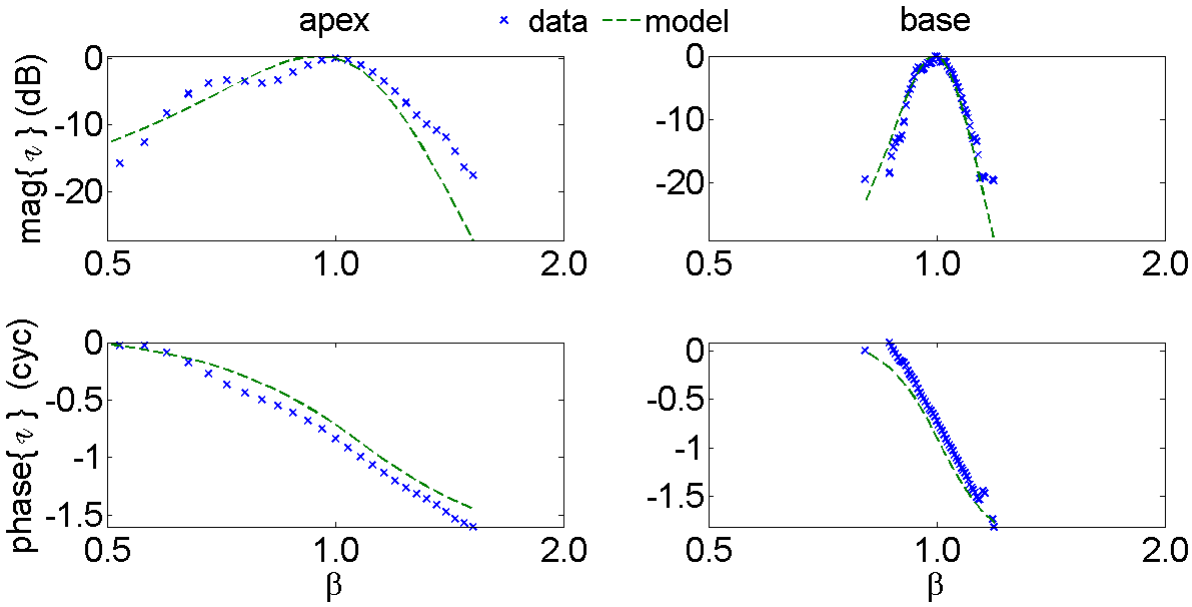

Figure 2.9: Model test using velocity expression II: The model (green dashed) expressions for velocity response is fit to measured data (blue crosses) from Wiener Kernels of chinchilla neural data from the apex (left) and base (right) in the region of the peak [37]. The magnitude (top) and phase (bottom) are plotted as a function of normalized frequency β. The model constant b_p is fixed at $b_p = 1$ for both fits. For the fit to the point in the apex, the estimated model constant values for A_p, and B_u are 0.5, and 2.7 respectively, and the objective function value is 0.16. For the fit to the point in the base, the estimated model constant values for A_p, and B_u are 0.4, and 3.9 respectively, and the objective function value is 0.09

We must note that a limitation of this particular validation (due to lack of available data) is that it is from single locations, and hence cannot validate the model comprehensively.

[33] This data set is Wiener-Kernels of auditory nerve fiber recordings and hence provides both magnitude and phase information

[34] Mechanical motion in the cochlea is transmitted to the auditory nerve fibers and from there to the brain

2.8 Translation for Application: Model Response Characteristics and Constants

In previous sections, we discussed the model development and tests for various variables. Specifically, in section 2.7, we provided the reader with expressions for the wavenumber and impedance, as well as the macromechanical responses - pressure and velocity. In this section, we translate the model so that we can use it towards the application of determining the wavenumber and impedance from response characteristics.

2.8.1 Response Characteristics

In table 2.1, we present the expressions for the response characteristics (introduced in figure 2.4) in terms of the model constants. We derived these expressions from the sharp-filter approximation of the magnitude and phase expressions for pressure. We leave the definitions and derivation of these response characteristics to section B.1.

Characteristics in normalized frequency, β	**Sharp-Filter Approximation**
Center frequency (β_{center})	b_p
Maximum (normalized) group delay (N) in periods of the CF(x)	$\frac{b_p B_u}{2\pi A_p}$
Phase accumulation (ϕ_{accum}) in cycles	$\frac{B_u}{2}$
n-dB Bandwidth (BW$_n$)	$2A_p\left(e^{\frac{\log(10)n}{10B_u}}-1\right)^{\frac{1}{2}}$

Table 2.1: Response characteristics in terms of model constants

As can be seen from table 2.1, the center frequency is in terms of one of the model constants, b_p. The phase accumulation is in terms of another, B_u. The group delay is a combination of the three model constants, and the bandwidth is in terms of A_p, B_u.

2.8.2 Determining Model Constants

In the previous section, we provided the reader closed-form expressions for the response characteristics in terms of the model constants. However, to order to apply the model we need to follow these steps:

1. We are given values of response characteristics (observables)
2. From these, we must determine values for the model constants - A_p, b_p, B_u
3. Then, we use these values for model constants along with our expressions to determine estimates of the wavenumber, k_x, and impedance, Z

In order to follow the above steps, we need to have expressions for the model constants in terms of the response characteristics. We list these in table 2.2. The left column contains the three model constants, A_p, b_p, B_u, and the right column contains a set of expressions for determining the model constants from response characteristics. Note that this is one possible set of expressions, since there are three model constants, but multiple potential response characteristics that we can use to determine the model constants.

Model Constant	**Estimate**	**In Terms of**
b_p	$\beta_{center} \approx 1$	normalized peak frequency = 1
B_u	$2\phi_{accum}$	phase accumulation
A_p	$\frac{b_p B_u}{2\pi N}$	mixture including maximum group delay

Table 2.2: Model constants in terms of response characteristics

The model constant b_p can be written in terms of the normalized peak frequency, β_{center}, which we can set to a fixed value of 1 in most cases. This fixed-value reduced the model to a 2-constant model, which is quite useful, and which we use for most of the remainder of this chapter. The model constant B_u is determined here by the phase accumulation, and A_p is determined by a mixture of response characteristics, including maximum group delay.

Due to the fact that the bandwidth is one of the more commonly reported response characteristics, it is relevant to provide an additional expression for one of the model constants in

terms of the bandwidth. Specifically, we can determine B_u by solving the implicit equation: $\mathrm{BW}_n N - \frac{B_u}{\pi}\left(e^{\frac{\log(10)n}{10B_u}} - 1\right)^{\frac{1}{2}} = 0$ for B_u given the response characteristics, BW_n, N. Whereas this equation is implicit, and hence inherently more difficult to solve than those in table 2.2, we make use of it in determining the wavenumber and impedance for humans because phase accumulations are not observable in humans.

2.9 Application: Determining Wavenumber and Impedance Using Response Characteristics in Humans

In this section, we discuss the method and results of applying our model towards our goal. In both the following subsections, we are given, or can determine, response characteristics from measurements or observations. From these response characteristic values, we determine values for the model constants based on section 2.8, which we then use in our expressions to determine the model wavenumber and impedance. Section 2.9.1 demonstrates the method and results of the application with a simple example that prepares the reader for section 2.9.2 which contains discussions specific to the human cochlea.

2.9.1 Application - Example

The primary purpose of this section, is to solidify our discussion regarding applications with a straightforward example. The approach to determine the wavenumber and impedance is as follows [35]:

1. We use the following values for response charactertics: $\beta_{center} = 1, N = 4.8, \phi_{accum} = 1.5, \mathrm{BW}_{3dB} = 0.1$. These values are from generic data sets, and are inter-consistent - in other words, using any combination of response characteristics, provides us with the same values for the three model constants

2. Using the above values for the response characteristics, we determine the following values for the model constants: $b_p = 1, A_p = 0.1, B_u = 3$

3. Finally, plugging the estimates for the model constants into our expressions for the wavenumber and impedance, provides us with the results shown in figure 2.10.

Our focus in this section is not to describe the resultant figure, but rather to demonstrate that we can use observations, and specifically response characteristics, to determine the wavenumber and impedance.

[35] For a discussion regarding determining the *variation* in the wavenumber and impedance based on differences or changes in model constants or response characteristics, see section C.2

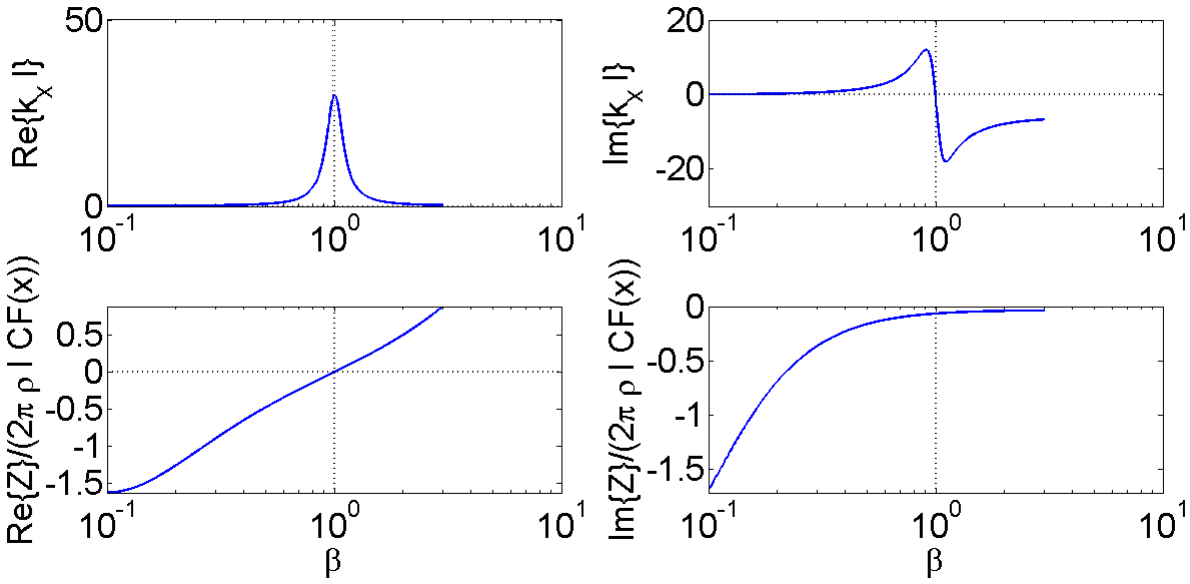

Figure 2.10: The figure illustrates the model application. The top row shows the wavenumber, and the bottom row the impedance, with the real parts on the left and the imaginary parts on the right. The model wavenumber and impedance were determined using the following values for model constants: $b_p = 1, A_p = 0.1, B_u = 3$, which were derived from the following values for response characteristics: $\beta_{center} = 1, N = 4.8, \phi_{accum} = 1.5, \mathrm{BW}_{3dB} = 0.1$

2.9.2 Application - Human Cochlea

Of particular interest is to determine the wavenumber and impedance of *humans*. The reason humans are our focus in this section is because the human cochlea is scientifically particularly interesting, and no previous estimates of the wavenumber and impedance exist for humans.

On the other hand, we *do* have reported values of human response characteristics which we can use towards our goal. Specifically,

- We have values for the sharpness of tuning in humans from psychoacoustic measurements [43]. The sharpness of tuning is directly related to our bandwidth response characteristic [36]. The greater the sharpness of tuning, Q_{erb}, the smaller the bandwidth, BW_{3db}
- We also have data of the ratio of two of our response characteristics - the sharpness of tuning and maximum group delay, $\frac{Q}{N}$ [44]. This ratio is constant as determined from chinchilla, and is thought to be species independent. Hence, we assume that this ratio takes on the same constant value in humans as well [37].

In figure 2.11, we have replotted the reported values for the two response characteristics, Q, $\frac{Q}{N}$ [38]. In order to use these two response characteristics to determine the wavenumber and impedance, we first determine the model constants, which we also show in figure 2.11. Specifically, we determine the model constant B_u from the ratio, $\frac{Q}{N}$ [39], then determine the model constant A_p from the estimated B_u and the sharpness of tuning, Q. For details on how we estimate the model constants from response characteristics, see section B.2.

We can immediately use these model constants to determine the wavenumber and impedance, but first we note that the model constants (as is the case with the response characteristics we used to

[36] For details on how to convert the equivalent rectangular bandwidth sharpness of tuning, Q_{erb}, see section B.2

[37] More specifically, this ratio $\frac{Q}{N}$ has been found to be constant in chinchilla [44]. Other mammals, such as guinea pigs have empirically been shown to have a constant ratio of the sharpness of tuning to the maximum group delay of stimulus frequency otoacoustic emissions, SFOAE (as opposed to a maximum group delay of the OoC partition velocity, N, as we have used here). Assuming the ratio of maximum SFOAE group delay to N is constant (≈ 2) across mammalian species and locations [44], we can extend this to assume that $\frac{Q}{N}$ is constant in humans as well.

[38] We should note that Q, N are both domain-independent - they take on the same values whether they are computed in the f-domain (as is the case with reported data), or in the β-domain (as in our model), and hence we do not need to convert from one domain to the other in order to use the reported values

[39] In fact, we chose to use $\frac{Q}{N}$ as one of the response characteristics (rather than, for example, N itself) to determine the model because a constant value of $\frac{Q}{N}$ corresponds to a constant value of B_u. This correspondence is assuming the pressure sharp-filter approximation

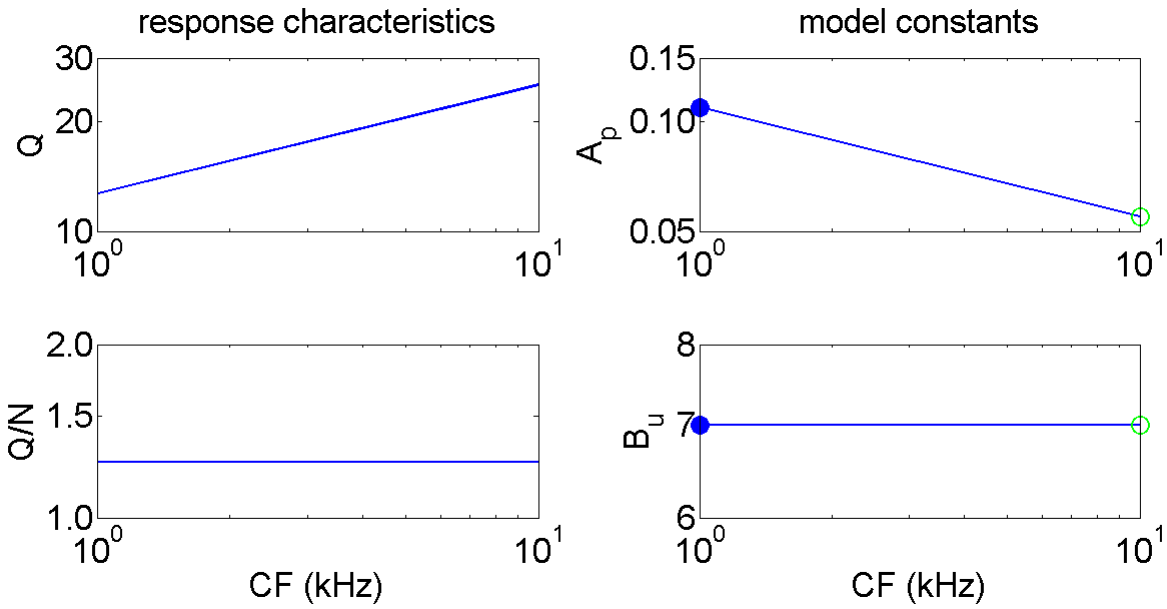

Figure 2.11: Reported human response characteristics and estimated model constants: The left panels show the reported response characteristics data from humans - as a function of location (through CF(x)), used to generate the human model. The top left panel shows that human sharpness of tuning, Q from psychoacoustic studies [43], as a function of location via the characteristic frequency, CF(x), which is the peak frequency of the velocity at each location. The bottom left panel shows data of the ratio of two response characteristics, $\frac{Q}{N}$ from [44], as a function of location or CF. The right panel shows the model constants estimated from these response characteristic values: the top panel shows A_p, and the bottom shows B_u (b_p is fixed to 1) as they vary as a function of location. In the main text, we discuss the variation of the wavenumber and impedance from base to apex by focusing in two points - one in the apex with a lower CF encircled in filled blue, and another in the base with a higher CF encircled with the unfilled green line.

estimate them), vary from base to apex [40]. Indeed, figure 2.11 shows not only the model constants estimated from human response characteristics, but also their variation with location. Hence, we can not only (1) determine the wavenumber and impedance from the human response characteristics, but we can also (2) differential between their profiles in the human base and apex (which inherently provides us with information regarding how these two regions of the cochlea function differently). We discuss both these applications in this section by demonstrating the wavenumber and impedance of two points along the length of the cochlea - one from the apex with a low CF, and another from the base with a higher CF [41].

In figure 2.12, we have generated the model wavenumber and impedance for the two points in the human apex and base (using the model constants $A_p = 0.11, B_u = 7$, and $A_p = 0.055, B_u = 7$ respectively). For the remainder of this section, we shall limit our discussion to information we can very directly describe based on the wavenumber and impedance, as well as their differences between the base and apex, from figure 2.12.

The real part of the wavenumber has a peak, the value of which determines the minimum incremental wavelength of the pressure traveling wave as it propagates along the length of the cochlea. The peak is greater in the base than the apex, indicating that the incremental wavelength is much smaller in the base than in the apex near the peak of the wave (though this is not the case outside the region of the peak). The imaginary part of the wavenumber has a peak and trough that have larger values in the base than the apex which indicates that the gain and dissipation are faster in the base than in the apex (we make this observation only for the region close to the peak of the traveling wave - i.e. around $\beta = 1$).

The imaginary part of the impedance at a particular location, x, normalized to the peak frequency at that location, CF(x), is negative (indicating an effective stiffness rather than mass), and is similar in the base and the apex. Removing this normalization, reveals that the effective stiffness is greater in the base than in the apex, as expected based on experimental data. The real part of the impedance (effective positive and negative damping) normalized to the CF at a particular point is the same in the base and the apex. It is only when we remove the normalization by CF do we see

[40] This variation is an extrapolation of our model which is generally valid when the variation of the model constants with respect to location is slow

[41] From the two sets of model constants, ($A_p = 0.11, B_u = 7$ for the apex, and $A_p = 0.055, B_u = 7$ for the base), we can determine the macromechanical responses, pressure and velocity, for the base and the apex, discussed in section E. Our main objective in this section, though, is the wavenumber and impedance as well as their variation.

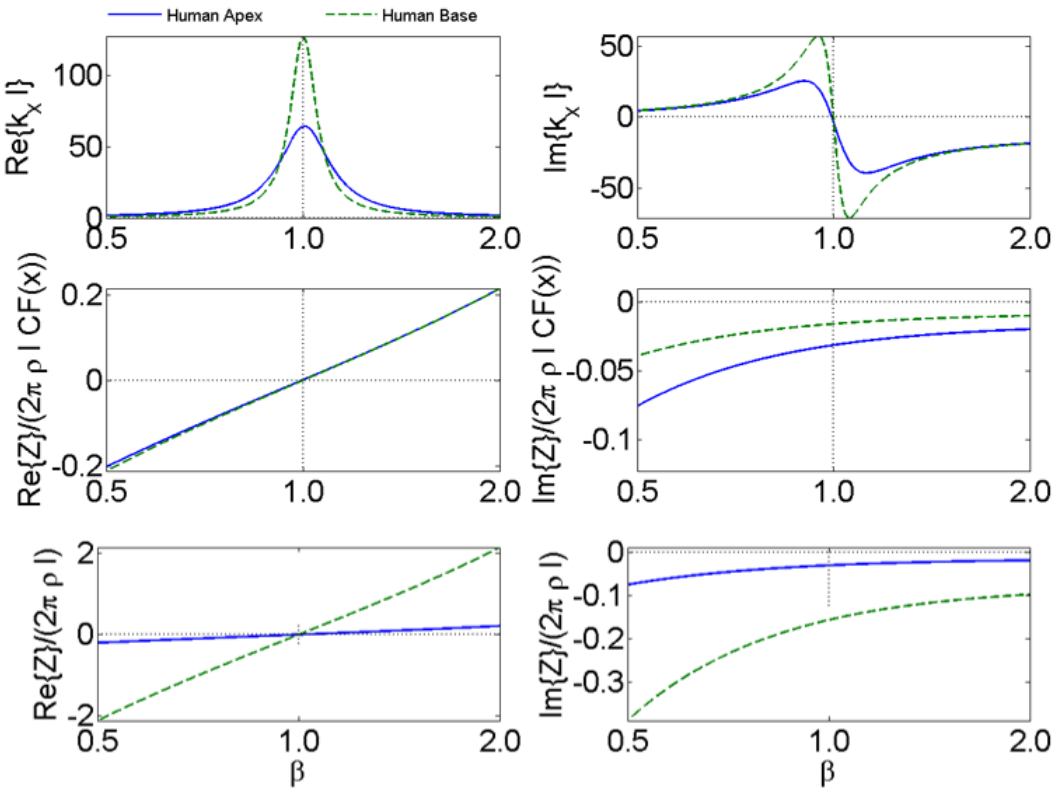

Figure 2.12: Estimated human wavenumber and impedance: The top panel shows the real and imaginary parts of the model generated human wavenumber, the middle panel shows the real and imaginary parts of the impedance (normalized to CF), and the bottom panel shows the real and imaginary parts of the impedance with the CF normalization taken out. The blue solid lines are for a point in the apex which has model constants $A_p = 0.11, B_u = 7$, and the dashed green lines are for a point in the base which has model constants $A_p = 0.055, B_u = 7$. The values for A_p, B_u are estimated from response charactersitics.

that the effective positive and negative damping is much greater in the base than the apex. The fact that the impedance normalized by CF (and particularly its real part) is very similar in the base and apex is an intriguing curiosity.

In this section, we have determined the human wavenumber and impedances, for which we have no previous estimates, and discussed certain differences between the base and the apex. From these estimates, it is possible to derive other functional information, such as phase velocity, dispersivity, power flux, effective stiffness, etc, that we briefly described in section 2.3, as well as their differences between the base and the apex. More generally, these differences in how the cochlea operates may be between various species, regions, or pathologies that have different response characteristics but the same underlying model structure.

2.10 Summary

In this section, we present the reader with a summary of this chapter. First, we developed our analytic model of the cochlea. Specifically, we presented the reader with closed-form expressions in x, ω for the wavenumber, impedance, as well as pressure and velocity. We tested the model in various ways and found it to works satisfactorily. We then translated the model so that it can be applied towards our goal by determining closed-form expressions for the model constants, A_p, b_p, B_u, in terms of response characteristics such as peak normalized frequency and group delay. We then used available values of response characteristics from humans in order to determine values for the model constants. We then substituted the values for these model constants in our expressions for the wavenumber and impedance. In other words, we derived estimates for the human wavenumber and impedance (unobservables, which encode how the cochlea works) from response characteristics (observable available data). In the process, we also determined a parameterized model structure of the cochlea for most mammals, and provided model constants' values for humans.

Chapter 3

Conclusion: Alternative Uses and Future Work

In the previous chapter, we presented a model of the mammalian cochlea, and determined the values of model constants for humans. The model encodes various types of variables: macromechanical responses (pressure and velocity), response characteristics (e.g. bandwidth, group delay), and the wavenumber and impedance. Hence, we can use the model for a variety of purposes some of which are the basis for our future work. In this section, we briefly outline such applications as well as our thoughts on modifying the model. Lastly, we discuss model limitations.

3.1 Using a Model Variable to Determine Another

The model encodes the aforementioned variables as (1) closed-form expressions, that are (2) a function of the *same* set of three model constants. Therefore, the model can be used to determine any one of these variables from another, as illustrated in figure 3.1.

Observables to Wavenumber and Impedance

For example, in the chapter 2, we used observable response characteristics from noninvasive experiments to determine the wavenumber and impedance to help understand how the cochlea functions from the perspective of functional aspects of the wavenumber and impedance - described in section 2.3. Alternatively, we may use macromechanical responses to determine the wavenumber and impedance in cases where these responses are available, or at least measurable (which is the case with nonhuman mammals where invasive experiments can be performed), and do not suffer from

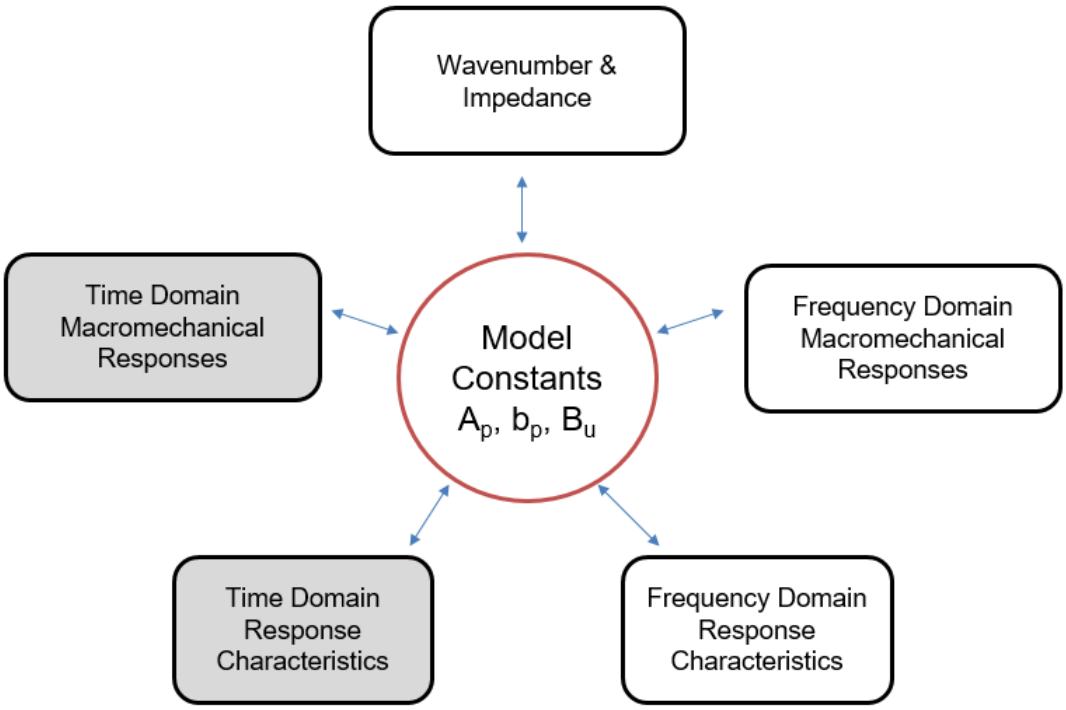

Figure 3.1: Model components to illustrate applications: The model has closed form expressions for the variables parameterized by a single set of model constants. The components in the clear boxes are described in the main chapter, and the components of the shaded boxes are in section D.3 or can be directly derived from it. The double arrows indicate that paths that we can follow to apply the model. Examples of such applications are discussed in the main text.

low signal to noise ratios.

Differences in Pathologies, Regions, and Species

In addition, the model can be used to determine how changes in one variable correspond to changes in another through differences in the model constants - see sections C.2, D.1.2. This can be used, for example, to map changes in the underlying function encoded in the wavenumber and impedance to corresponding changes in observable such as macromechanical responses or their response characteristics. Such a mapping would be particularly informative in understanding what underlying changes must occur in various pathologies, regions, or species, for them to behave differently.

Data Acquisition and Processing

Where the model is valid, the expressions for pressure and velocity can be used to aid in the quality of data acquisition. Simultaneous mechanical measurements of pressure and velocity are quite noisy because the method for measuring pressure is relatively recent (e.g. [14]). Therefore, using the model as a constraint relating velocity and pressure can, in theory, aid in the quality of such data.

We can use the model to determine response characteristics from macromechanical responses. Consider, for example, a noisy data set for pressure, for which we may want to determine a response characteristic that uses single data points or derivatives - e.g. group delay. Noisy data sets are unreliable for computing such information directly. We can instead determine the model constants by fitting the model pressure expression to data, then use these model constants to determine the value of the group delay. This proposed approach utilizes information from along the entire data set rather than taking derivatives or using a small number of data points. Note that, for this application, we suggest using an alternative objective function than that of section A.4 in order to distribute the fitness weight beyond the peak.

For Auditory Filter Applications

Conversely, we may determine macromechanical responses from desired response characteristics. In addition, the model allows for mapping between the frequency domain response characteristics (e.g.

bandwidth), time domain response characteristics (e.g. rise-time), frequency domain macromechanical responses, and time domain macromechanical responses (see section D.3). These mappings are quite useful for auditory filter applications described in the following section.

3.2 Using Macromechanical Response Expressions for Auditory Filters

The model has representations for macromechanical response variables which can be used as auditory filters. Such filters are useful for the design of cochlear implants, speech processing, compressing audio files, etc. The model can also be used in perceptual/behavioral studies that require integrating a model of the cochlea with models of higher neural centers, and can be used to determine what aspects of the wavenumber, impedance, pressure or velocity are important for determining higher neural function such as localization. Similarly, the model can be used to understand features of these variables that are important for cochlear implant users (e.g. fine timing is not important for understanding speech). We leave the details of these applications, as well as the model features that enable them, to section I.

We also note that the model has certain similarities to other models - see section H. Hence, our model can be used towards similar applications. Also as a result of similarities, we can leverage current software and hardware implementations for applying our model for auditory filter purposes thereby reducing the additional implementation efforts associated with adopting new models.

3.3 Deeper Interpretation of Wavenumber and Impedance

As mentioned previously, the wavenumber and impedance encode a wealth of information regarding how the cochlea works. Therefore, it is relevant to extract these specific features from the wavenumber and impedance such as those listed in section 2.3, as well to determine their dependence on species and location.

Of these features, the power flux, its profile, and their variation from base to apex are of particular interest to those studying cochlear amplification. In addition, we can use the model (effective) impedance and determine the nature of the corresponding local impedance given various

forms of longitudinal coupling internal to the OoC partition itself (whether it be viscoelastic or hydrodynamic, etc).

It would be useful to construct a control system representation that is consistent with our expression for the impedance in the form of - for example, a feedback loop with an op-amp, which can be better used to describe how the OoC responds to a particular pressure difference across it in a step-by-step manner. Such a description would not contain elements such as negative damping, or frequency dependent stiffnesses which are difficult to understand. A good starting point by [56] that should be pursued is described in section H. Such a dynamic system representation can also help bridge between macromechanical models (such as ours), and micromechanical models that contain more detail in the OoC but cannot be solved analytically.

3.4 Modifications to Model

Passive Counterpart

The model we presented in this thesis is for the *active* cochlea. We suggest two potential methods to determine the equivalent model for the passive cochlea, where the 'cochlear amplifier' is inactive - the passive state of the cochlea corresponds to dead animals, and in response to high stimulus levels:

- Determine a dynamic system representation of the OoC impedance of our model that contains an amplifier. Removing the amplifier from the circuit should result in the passive counterpart of the model.

- Use the approach for model construction developed in chapter 2 to construct a passive version of the model. Specifically, the same physical component may be used (provided that the short-wave approximation holds). The phenomenological component may be determined by determining the *qualitative* behavior of the wavenumber from measurements of pressure that are available from passive cochlea (e.g. [14]), along with the relationship between wavenumber and pressure (specifically, equation 2.11).

Generalization

In the course of model development, we have assumed scaling symmetry of the wavenumber (and pressure). We may instead be interested in developing a more general model, or determining the degree of asymmetry in the cochlea. To do this, we may determine an expression for the the dependence of the model constants on location, x as we found in 2.9, then incorporate this dependence in the model expressions to develop a more generalized model. This approach is subject to the model constants varying slowly enough for the closed-form model expressions to still hold.

Nonlinearity

We have also assumed linearity, which is appropriate at low stimulus levels and allowed us to work in the frequency domain. One way of introducing 'effective' nonlinearity in the frequency domain is to determine an expression for the dependence of the model constants on stimulus level, then incorporate this dependence in the model expressions. To make the model truly nonlinear, we must incorporate the nonlinearity in the time domain of section D.3.

3.5 Further Tests

As mentioned in section 3.1, the model is a single unit that provides expressions for a number of related variables that are illustrated in figure 3.1. Therefore, an important feature of the model is that by determining the three model constants from any variable, we can determine all other variables. It is desirable, then, to test the model as a whole, and determine if the relationships between *simultaneous* experimental measurements of these variables are as predicted by our model.

In the absence of suitable experimental data (as is the case using current technology), the model can alternatively be tested by using outputs of computational models that incorporate micromechanics - e.g. [32]. An advantage of using such computational models for testing our model as a whole, is that we can make changes to the computational models to test whether our model holds when varying assumptions (for example, assumptions regarding longitudinal coupling through the tectorial membrane). On the other hand, a disadvantage of using such computational models, is that these models often assume a particular set of values for model constants that are rarely derived from cochlear measurements of material and geometric properties, and therefore, the outputs, and

in particular, their relations, are not necessarily reflective of the cochlea. Given this disadvantage, we would highly recommend using multiple computational models for such a test.

3.6 Limitations

In this section, we discuss limitations of the model and caveats to applying it that arise from the physical and phenomenological components.

Physical Component

As the relationships between the model variables are based on the physics of classical box representations, the model is limited by the assumed physics - see sections 2.2 and J for assumptions. We discuss some of these limitations here. For instance, the concept of impedance in this model assumes a single partition for the OoC (the reader must note that the model single OoC partition is *not* assumed to correspond to a particular anatomical membrane but rather can be thought of as some imaginary 'effective membrane'). Recent data [18] appear to support this single-partition assumption in passive cochleas as the absolute pressure measurements made in two fluid compartments at points close to the OoC are, to first approximation, anti-symmetric (neglecting geometric considerations, the single partition assumption is equivalent to anti-symmetric pressure). On the other hand, recent measurements of the various membranes using optical coherent tomography in the active cochlea seem to indicate a large differential motion between the top and bottom of the OoC, which means that the various membranes in the OoC should be considered separately. To what degree this effects the assumptions of the model (and at what frequencies and locations) is not entirely clear - if the differential motion is not very large, then the complexity can be considered a second order effect and assuming an 'effective membrane' is still suitable. We suggest performing absolute pressure measurements in the active cochlea in order to quantitatively determine how much deviation there is from pressure anti-symmetry, and therefore, from the single partition assumption. If the deviation is large, then we suggest re-deriving the physics from basic principles in a three-fluid compartment system with two partitions. The constitutive equations of the classical box representation (e.g. see section J) can be utilized in re-deriving the physics.

It is important to note that the relationship between the wavenumber and the (dominant com-

ponent of) pressure - equation 2.1, approximately holds in any situation where pressure traveling waves exist, which is certainly the case in the cochlea [1]. Hence, if either the model wavenumber or the model pressure is 'correct' (i.e. reflective of the actual cochlea), then so is the other variable. Therefore, this aspect of the model can generally be retained, and the wavenumber can still be determined from response characteristics of the pressure - regardless, of, for example, the validity of the single partition assumption, or other future experimental work.

We note that any direct forms of longitudinal coupling *within a single partition* (e.g. viscoelastic coupling) do *not* effect the physical equations, and hence the model does not change. What does change is determining local contributions from the (effective) impedance, Z. On the other hand, if we wish to incorporate non-inertial coupling in the fluid itself (e.g. for investigating the effects of fluid viscosity), then the constitutive equations change drastically, and the model cannot hold.

Phenomenological Component

Other limitations of the model come from our phenomenological component, and specifically the fact that we *constructed* (rather than derived) one of the variables - the wavenumber. In principle, any number of expressions can be used to fulfill the desired qualitative behavior of the wavenumber, and the constraints we imposed (such as rational transfer functions) are partly for convenience. This phenomenological aspect is a strength of the model, allowing for its simplicity (and therefore determining the wavenumber and impedance from response characteristics), but the reader must bear in mind that an alternative set of expressions for k_x, Z, P, V may also be appropriate.

[1] This relationship between pressure and wavenumber exactly true for a first order wave equation as is the case with the short-wave approximation of classical physics. It is approximately true for a second-order two wave equation such as the two-way long-wave approximation of classical physics, as it is equivalent to the WKB$_o$ approximate solution, which is valid provided that the spatial gradient of the wavenumber is sufficiently small (we have established that it is). It is similarly approximately true for higher order wave equations

Appendix A

Further Model Tests

In the following sections, we discuss additional model tests in order to determine the suitability of our model and our approach for determining the model wavenumber and impedance given response characteristics.

A.1 Reconstructability and Consistency

In order to properly test the model, we must investigate its consistency. We do so in this section by checking if the values of response characteristics used to estimate the model constants are well approximated by reconstructed response characteristics values from the model.

To address this issue, we start with the same set of values for response characteristics in section 2.9.1, which gives us the same set of values for the model constants. We then substitute these values into our expressions for the pressure and velocity which are shown in figure A.1. From these model macromechanical responses, we can then compute the model response characteristics. We then compute the model response characteristics, and compare these with the response characteristics we used to generate the model. The relative errors between the two sets of response characteristics are in table A.1.

Response Characteristic	**Relative Error in P**	**Relative Error in V**
$\epsilon_{\beta_{center}}$	0.5%	0.3%
ϵ_N	0.3%	−33%
$\epsilon_{\phi_{accum}}$	3%	−24%

$\epsilon_{\mathrm{BW}_{3dB}}$	-0.6%	-15%

Table A.1: Relative errors in reconstructability: The table shows the relative errors between the given response characteristic values used to generate the model, and those computed once the model was constructed. The relative errors in the response characteristics are defined as $\epsilon = \frac{z_{input} - z_{model}}{z_{input}}$ for each of the response characteristics, where z_{input} is the response characteristics, z used to construct the model, and z_{model} is the response characteristics computed from the macromechanical responses generated from the constructed model. In constructing the model, z_{input} is assumed to be (1) of the pressure P, and (2) within the sharp-filter approximation. The z_{model} in ϵ is computed from the constructed model pressure in the left column, and from the constructed model velocity in the right column.

The relative errors are generally smaller for the response characteristics computed from P compared to those computed from V. This is expected as the model expressions for constants in terms of response characteristics (table 2.2) is based on the characteristics of P. The relative errors in P characteristics are small, especially considering that the model expressions for constants in terms of characteristics is based on the sharp filter approximation of P rather than the full expressions. This suggests that the expressions based on the sharp-filter approximations are sufficient, and there is no need for adding complexity by using response characteristics based on the full expressions.

Note that the relative errors we have reported are very slightly higher than their true values due to computational error in numerically computing characteristics from macromechanical responses (more specifically, finite resolution of β). We expect that it is for this reason that the relative error for β_{center} is artificially lower using V compared to P.

We should also note that, if we use the model to determine *changes* in the wavenumber and impedance that underlie changes (or differences) in response characteristics, we need to ensure that these changes are less than the relative errors. Otherwise, the inferred wavenumber and impedance changes cannot be considered reliable.

Also, it is clear from table A.1, that the error in certain response characteristic is smaller than others, and hence it is reasonable to try to use the lower-error response characteristics as much as possible to construct the model. Lastly, the reader should note that the expression that we used to determine the model constants in terms of response characteristics are best when the macromechanical responses are sharp, or in other words, the bandwidth is small. Hence, the approach is best for certain species or regions where this is true (humans fall into this category).

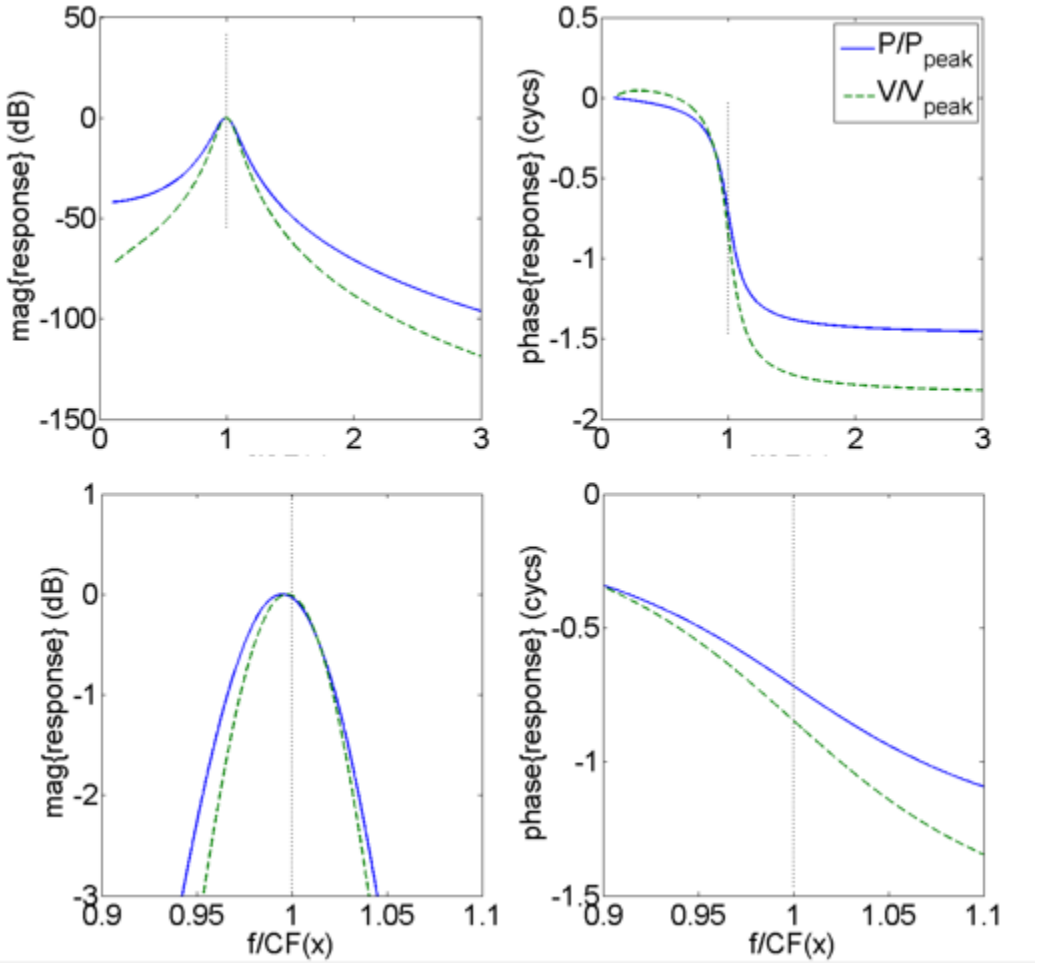

Figure A.1: Model consistency: We estimate model constant values based on a set of response characteristics, $\beta_{center} = 1$, $N = 4.8$, $\phi_{accum} = 1.5$, $\text{BW}_{3dB} = 0.1$ using table 2.2. From the resultant model constant values, we then determine the model macromechanical responses, P (blue solid), V (green dashed). From the plotted macromechanical responses, we compute the response characteristics numerically in order to determine relative errors in the characteristics. The pressure and velocity are plotted in reference to their peak values at a particular point, and as a function of the normalized frequency, β. The bottom row is a zoomed-in version of the top row to show a magnitude range of 3dB. The bottom left panel shows that the 3dB bandwidth is close to 0.1 which is what we used to construct these model curves, which is encouraging. The relative errors between the reconstructed response characteristics computed from these curves and those starting response characteristics used to generate the model are in table A.1.

A.2 Sharp-Filter Approximation

In this section, we discuss the validity of the sharp-filter criteria for our application as it is an essential component of determining the model from response characteristics. In this section, we also define a criteria by which the reader can determine the suitability of the sharp-filter approximation based on their intended applications.

As mentioned in section 2.7.1, the sharp-filter approximation holds for values of β at which the following measure,

$$\alpha \triangleq \frac{A_p}{\beta + b_p} \,, \tag{A.1}$$

takes on values $\ll 1$. At the peak, this is simply $\frac{A_p}{2}$. Figure A.2 shows that the sharp-filter approximation is valid in humans close to the peak across the various regions, which justifies section 2.9.2.

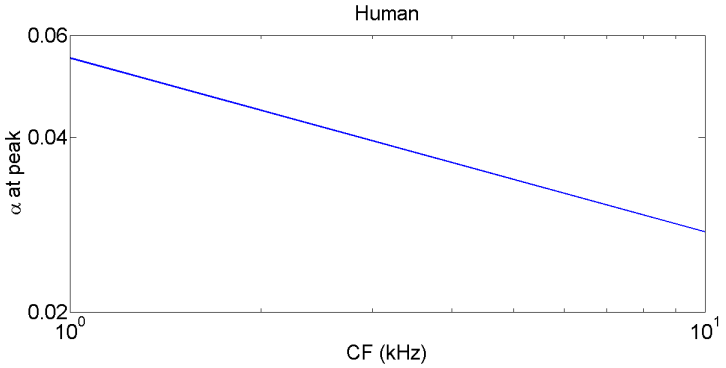

Figure A.2: Validity of sharp-filter approximation - humans: The figure shows that the sharp-filter criteria of equation A.1 holds in humans across the entire cochlea, and particularly so for the base.

Consider, alternatively, using $Q, \frac{Q}{N}$ response characteristics to determine the model constants for chinchilla. Based on the estimated A_p, as seen in figure A.3, the sharp-filter approximation criteria is met in the base of the chinchilla (which, intriguingly, is like the human apex) but not the chinchilla apex.

Indeed, if we attempt to construct the model wavenumber and impedance from the estimated chinchilla apex model constants, there is a clear discrepancy between the behavior of the full expressions and that of the sharp-filters. To illustrate this point, we have included figure, A.4.

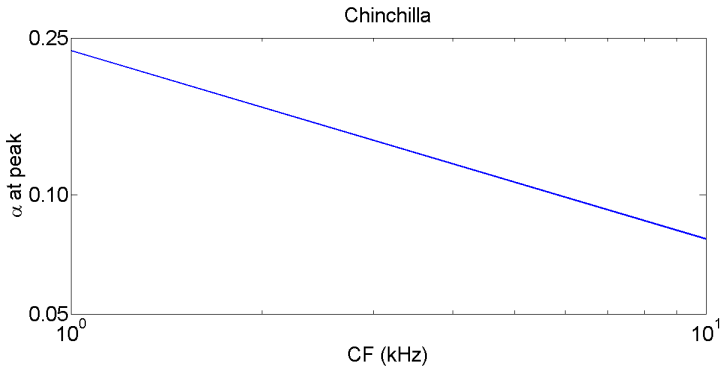

Figure A.3: Validity of sharp-filter approximation - chinchilla: The figure shows that the sharp-filter criteria of equation A.1 holds in base of the chinchilla but not the apex.

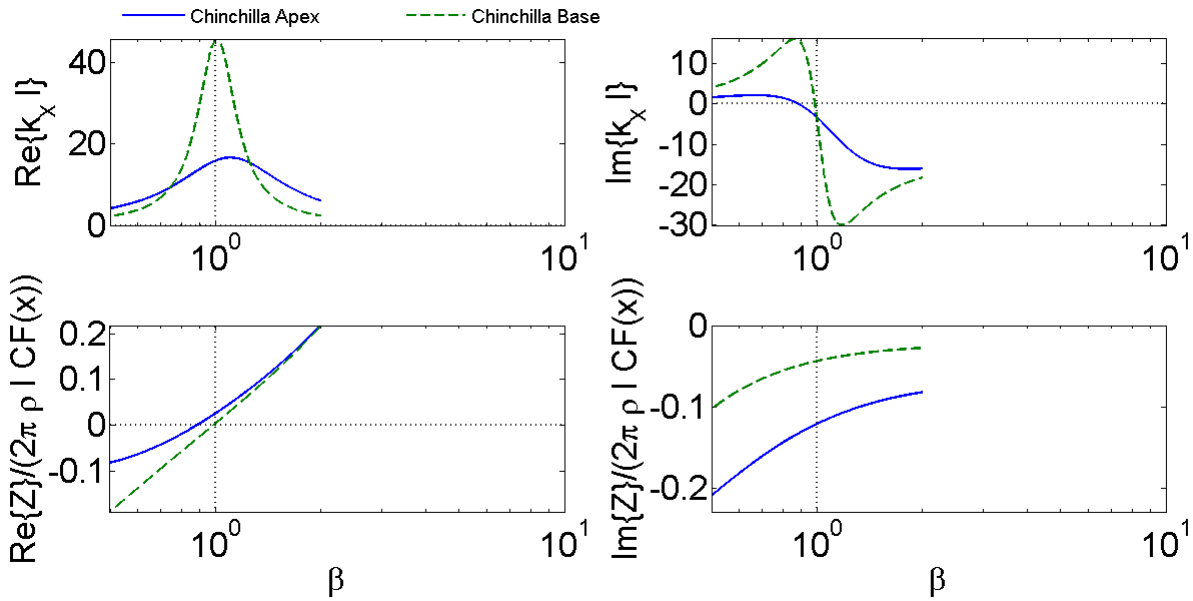

Figure A.4: Limitation of approach: The figure shows the real and imaginary parts of the wavenumber and impedance. The plots are generated using model constants estimated from response characteristics assuming our sharp-filter expressions for the chinchilla apex (solid blue line), and base (dashed green line). For the chinchilla apex, the peak of $\mathrm{Re}\{k_x l\}$, the zero crossing of $\mathrm{Re}\{Z\}$, and the zero crossing of $\mathrm{Im}\{k_x\}$ are shifted away from $\beta = 1$. More importantly, in the chinchilla apex, the positive values for $\mathrm{Im}\{k_x\}$ are small and hence there is negligible growth of the model pressure before the peak (relative to the decay after the peak) - which is inconsistent with experimental measurements. These observations reflect the fact that the model is not suitable for the chinchilla apex (which we have previously shown does not satisfy the sharp-filter approximation).

Our above discussion elucidates a limitation of our particular approach of using response characteristics to generate the model (or more specifically, determine its wavenumber and impedance). In such cases, alternative methods are necessary such as using macromechanical response information directly to determine the model constants, and from them, the wavenumber an impedance.

A.3 Pressure Response

In this section, we provide an additional test for the model by comparing our model (differential) pressure P with measurements of absolute pressure from the scala tympani close to the OoC of [14]. While The data is not of the differential transpartition pressure across the OoC, we assume that it approximates a scaled version of the transpartition pressure, based on classical model single-partition assumptions. The data P magnitude and phase are from a single location, and reported as a function of frequency around the peak relative to the stimulus. Both our model and this data show a peak in the magnitude, and a decaying phase of P, with a maximum group delay at the peak. The data shows a plateau in the phase for low frequencies as is the case with our model, but the behavior of the data at high frequencies beyond the peak is unclear.

A.4 Method for Fitting Model to Velocity Data

In this section, we do not provide a new test for the model, but instead explain the method by which we fit the model to velocity data in section 2.7.4. We also use a similar method for fitting to the data-derived wavenumber of section 2.7.1.

We obtain the model constant estimates for A_p, B_u by finding the minimum of an objective function on an A_p x B_u grid while fixing $b_p = 1$. It is appropriate to fix b_p since the CF of the single points from which the datasets were collected can be easily approximated from the measured magnitude curves. The simple brute force approach we employ avoids issues of local minima. We define the objective function to be $\sqrt{\frac{\sum_{i=1}^{m} |\Delta|^2}{m}}$ where m is the number of data points, and Δ is the complex residual, $\mathcal{V}_{data} - \mathcal{V}_{model}$. Note that this form for the objective function weighs the real and imaginary parts equally, and gives greater weight to regions where $\mathcal{V}$ is large and hence emphasizes the peak region. The A_p x B_u grid is constructed such that each of A_p and B_u has a range of 100 logarithmically-spaced values from 0.1 to 10. The most important aspect of this method is that the

grid values were chosen such that they incorporate expected model constant values that we derived from our model expressions for model constants as a function of response characteristics - as in table 2.2.

A.5 Boundary Condition Dependence

The expressions for the macromechanical responses, pressure and velocity (and hence the response characteristics), depend on our choice of effective boundary condition or integration constant for equation 2.11. For the physical aspect of the model to hold, the model pressure and velocity should be approximately independent of boundary conditions for the range of model constants we are interested in [1]. Therefore, in this section, we show that the model expressions are relatively independent of the chosen boundary condition for most cases, and derive expressions for the macromechanical responses with an alternate choice of boundary condition.

We break this section down as follows:

- We first provide the relationship between pressure and wavenumber in section A.5.1
- Then, we derive the parametric model expressions for the macromechanical responses using two choices of effective boundary conditions in section A.5.2
- Lastly, we discuss the dependence of the model expressions on the choice of boundary condition in section A.5.3 and show that, for most cases, the model is relatively independent of the choice of boundary condition, and hence our expressions in the main chapter hold regardless of this choice

A.5.1 Dependence of Relationship Between Pressure and Wavenumber on Boundary Condition

For simplicity of derivation, we define a rotated version of the wavenumber,

$$r \triangleq -ik_\beta = -ik_x \frac{l}{\beta} \,, \tag{A.2}$$

[1] Note that, for pure auditory filter purposes, this boundary condition independence does not matter

in order to rewrite equation 2.11as,

$$\frac{dP}{d\beta} - rP = 0 \ , \tag{A.3}$$

or equivalently as,

$$\frac{d\log(P)}{d\beta} - r = 0 \ . \tag{A.4}$$

The last equation is a linear first order separable ordinary differential equation and hence we arrive at its solution simply by direct integration,

$$\log(P) = \int r d\beta + c(\omega) \ , \tag{A.5}$$

where $c = c_{\propto \text{mag db}} + ic_{\text{phase}}; (c_{\propto \text{mag db}}, c_{\text{phase}}) \in \mathbb{R}$ is determined by an effective boundary condition. This formulation in terms of $\log(P)$ is appropriate for cases where it is desirable to have expressions for the magnitude and phase of the pressure individually (using the real and imaginary parts of r respectively), as we have shown in section 2.7.3. The boundary condition is encoded in the integration constant.

For determining an expression for the complex P (rather than its parts individually), we rewrite the relationship between r and P, as

$$P = C(\omega) e^{\int^{\beta} r_{\beta}(\tilde{\beta}) d\tilde{\beta}} \ , \tag{A.6}$$

or equivalently, as,

$$P = \hat{C}(\omega) e^{\int_{\Omega}^{\beta} r_{\beta}(\tilde{\beta}) d\tilde{\beta}} \ , \tag{A.7}$$

where,

$$\Omega \triangleq \frac{\omega}{\omega_{max}} \ . \tag{A.8}$$

We have provided both of these equivalent equations (A.6 and A.7) relating P to r as the derivation of the expression for P assuming the boundary condition used in section 2.7.3 is simpler using equation A.6, whereas the derivation of an expression for P is simpler using the alternate boundary condition detailed later in this section is simpler using equation A.7.

The integration constant is generally a function of ω (it is a constant with respect to x) where this constant is determined based on the basal boundary condition. However, as we are focused only

on the short-wave region, this presents us with a complication about how to define the boundary condition. In the main chapter, we assumed that the integration constant in constant with respect to ω, $C(\omega) = C$, and this is discussed further below.

A.5.2 Expressions for Macromechanical Responses For Various Boundary Conditions

In this section we derive our model expressions for the macromechanical responses from r and the equations for P in terms of r above - for two choices of boundary conditions. Hence it is relevant to provide the reader our model expression for r, which we get from equation 2.5:

$$r(s) = -2B_u i \frac{s + A_p}{(s-p)(s-\overline{p})} \ , \tag{A.9}$$

or equivalently (and easier for integration),

$$r(s) = -iB_u \left(\frac{1}{s-p} + \frac{1}{s-\overline{p}} \right) . \tag{A.10}$$

Note that the boundary condition is *effective*. This is because the physics component that we outlined in section 2.2 is already reduced to the short-wave approximation, and hence limited to the peak region of the pressure traveling wave.

Expression for Macromechanical Responses Assuming Scaling Symmetric Pressure

In this section, we discuss the boundary condition used for deriving the pressure in the main chapter - section 2.7.3, as well as detail the resultant derivation of the model expressions for pressure and velocity. The assumption here regarding the boundary condition is that the resultant expression for the pressure is scaling symmetric.

We first rewrite equation A.6 as follows to make the integration easier,

$$P(x, \omega) = C(\omega) e^{F(s)} \ . \tag{A.11}$$

Note that from equation, 2.11, we have defined,

$$F(s) \triangleq \int^{\beta} r(\tilde{\beta}) d\tilde{\beta} = -i \int^{s} r(\tilde{s}) d\tilde{s} \ . \tag{A.12}$$

Substituting with our model expression for r, A.10, we get,

$$F(s) = -B_u \bigg(log(s-p) + log(s-\overline{p}) \bigg) \ . \tag{A.13}$$

Taking the exponential, gives,

$$e^{F(s)} = \bigg((s-p)(s-\overline{p}) \bigg)^{-B_u} \ . \tag{A.14}$$

Substituting in equation, A.11, we get an expression for P,

$$P(x,w) = C(\omega) \bigg((s-p)(s-\overline{p}) \bigg)^{-B_u} \ . \tag{A.15}$$

Now we make an assumption that the effective boundary condition is such that P is approximately scaling symmetric. Therefore, we impose,

$$C(\omega) = C \ , \tag{A.16}$$

which gives us equation 2.12, the scaling symmetric P expression of section 2.7.3, and leads to the expressions for velocity, velocity normalize to the CF value, and response characteristics in the main chapter.

Expression for Macromechanical Responses Assuming Effective Stapes Boundary Condition

Here, we describe an alternative choice for the boundary condition and detail the derivation of the macromechanical expressions. In this case, we assume that $C(\omega)$ is not constant, but instead determined by the stapes boundary condition,

$$\frac{\partial}{\partial x} P(x,\omega)|_{x=o} = 2j\omega\rho v_{stapes} \ , \tag{A.17}$$

or equivalently,

$$\frac{\beta}{l}\frac{\partial}{\partial\beta}P(x,\omega)|_{\beta=\Omega} = 2j\omega\rho v_{stapes} \ . \tag{A.18}$$

This is an *effective* boundary condition as the short-wave region (which is where the model is developed for) does not truly begin at the stapes. Therefore, in the following equations, $v_{stapes}(\omega)$ not only includes the stapes effects, but also implicitly incorporates the effect of the long-wave region. If we assume that the stapes velocity is an impulse we can drop the ω dependence. We expect that the long-wave region contribution is largely flat in the ω domain as most of the tuning arises in the short wave region. However, we note that the long-wave region may contribute some spatial dependence.

Using equation A.7, and a derivation process similar to that outlined for the above boundary condition case, we arrive at an expression for P using the effective stapes boundary condition,

$$\frac{P(x,\omega)}{\omega_{max}\rho l v_{stapes}} = -\frac{1}{B_u}\frac{1}{s_o+A_p}\Big((s_o-p)(s_o-\overline{p})\Big)^{B_u+1}\Big((s-p)(s-\overline{p})\Big)^{-B_u} , \tag{A.19}$$

where we have defined,

$$s_o \triangleq s\Big|_{x=0} = i\Omega \ . \tag{A.20}$$

Indeed, the reader can see that this satisfies equation A.18.

By using the above expression for P, as well as the expression for Z in equation 2.9, we arrive at the following expression for velocity,

$$V(x,\omega) = v_{stapes}\frac{1}{s_o}\left(\frac{s+A_p}{s_o+A_p}\right)\left(\frac{(s_o-p)(s_o-\overline{p})}{(s-p)(s-\overline{p})}\right)^{B_u+1} . \tag{A.21}$$

If we focus on the transfer function view, as opposed to the traveling wave view, it appropriate to consider the velocity normalized to its value at CF,

$$\mathcal{V}(x,\omega) \triangleq \frac{V(x,\omega)}{V(x,\omega_{peak})} = \frac{i}{s}\frac{s+A_p}{i+A_p}\left(\frac{(s-p)(s-\overline{p})}{(i-p)(i-\overline{p})}\right)^{-(B_u+1)}\frac{s_{o,c}+A_p}{s_o+A_p}\left(\frac{(s_{o,c}-p)(s_{o,c}-\overline{p})}{(s_o-p)(s_o-\overline{p})}\right)^{-(B_u+1)} , \tag{A.22}$$

where we have used the notation,

$$s_{o,c} = i\frac{2\pi \mathrm{CF}(x)}{\omega_{max}} \ . \tag{A.23}$$

A.5.3 Dependence of Model Expressions on Boundary Conditions

Here, we compare equations 2.15, and A.22 and show the velocity normalized to its CF value is similar for both choices of effective boundary conditions, and therefore, the set of model expressions are valid under most conditions. The solutions diverge and show a dependence on the boundary condition very close to $x = 0$ for frequencies beyond the peak. The peak region itself is least affected by the choice of boundary condition [2]. We illustrate this with two sets of values for the model constants - derived in section 2.9.2, in figures A.5 and A.6.

Note that the figures are inherently a transfer function view. When using the effective stapes boundary condition, we note a reversal in phase at higher frequencies in the most basal locations which has been observed in experimental data. This suggests that the experimental observation regarding the phase reversal may be a result of proximity to the stapes basal boundary condition.

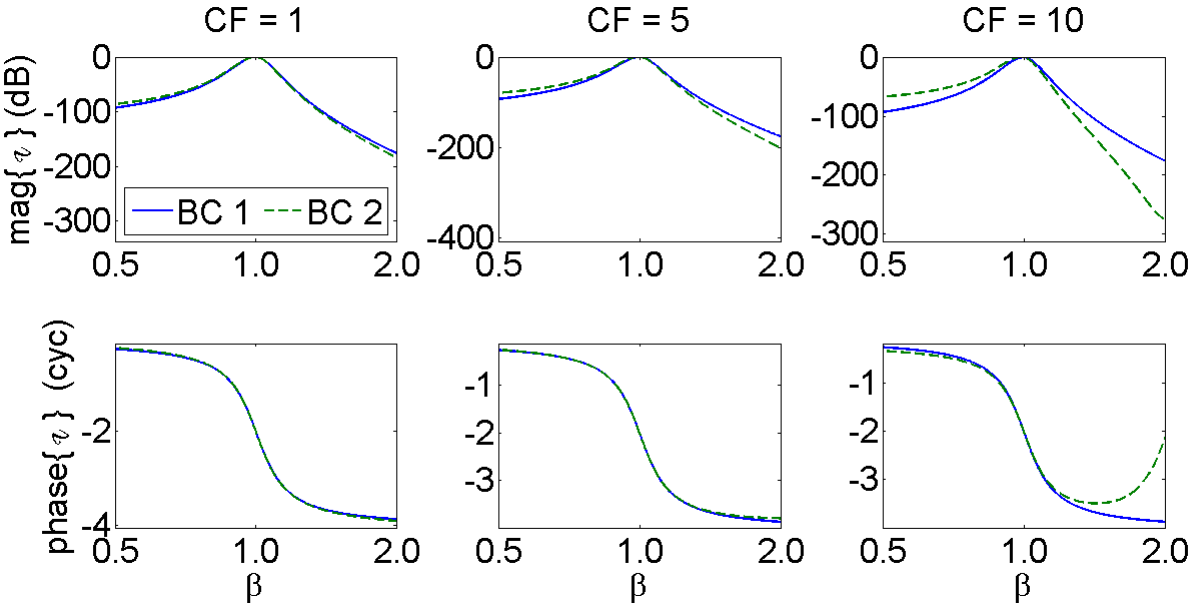

Figure A.5: The figure shows velocity normalized to its CF value at various locations along the length of the cochlea denoted by their CF. The solid blue line shows the model curves for the scaling symmetric pressure assumption and the dashed green line shows the effective stapes assumption for the boundary condition. As can be seen from the curves, the responses are relatively independent of the choice of boundary condition at the peak. The deviation between the two boundary conditions is greatest closest to $x = 0$ (when CF is closest to f_{max}), and at frequencies greater than CF. The model constants used are $b_p = 1, A_p = 0.1, B_u = 7$.

[2] Hence, we suggest using peak-centric response characteristics

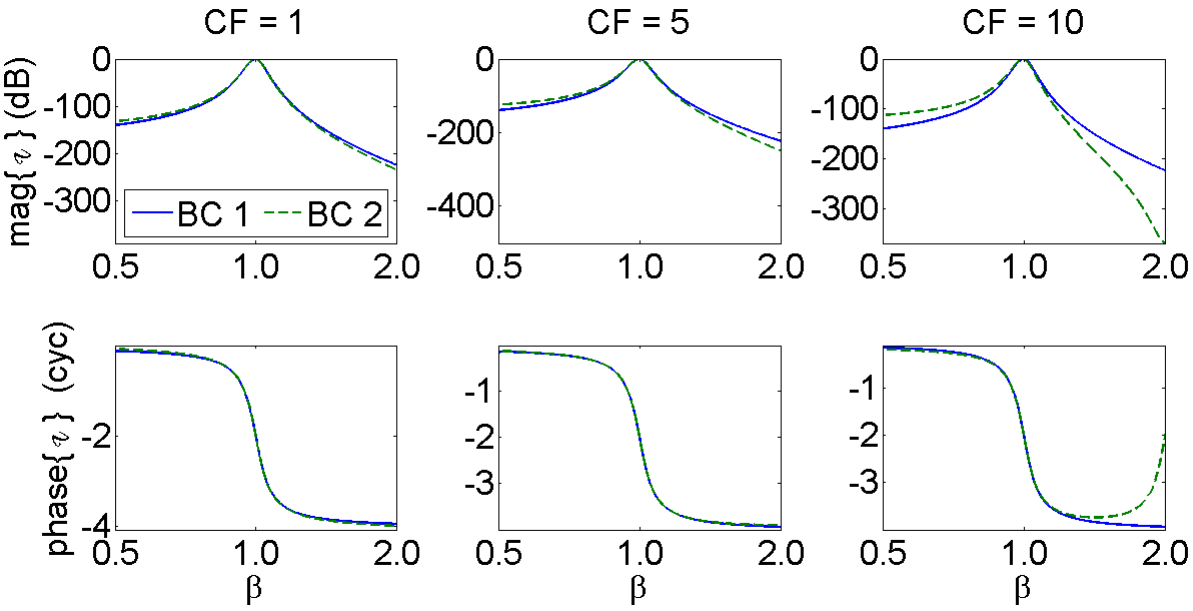

Figure A.6: The figure shows velocity normalized to its CF value at various locations along the length of the cochlea denoted by their CF. The solid blue line shows the model curves for the scaling symmetric pressure assumption and the dashed green line shows the effective stapes assumption for the boundary condition. As can be seen from the curves, the responses are relatively independent of the choice of boundary condition at the peak. The deviation between the two boundary conditions is greatest closest to $x = 0$ (when CF is closest to f_{max}), and at frequencies greater than CF. The model constants used are $b_p = 1, A_p = 0.055, B_u = 7$.

Appendix B

Derivation and Use of Response Characteristics

B.1 Derivation of Model Expressions for Response Characteristics

In this section, we provide the reader with details of our derivation of the response characteristics in table 2.1.

Derivation and Definitions

We first refer the reader back to table 2.1, where we have summarized our model's sharp-filter approximation expressions for certain characteristics of the response P [1]. We derived these directly from the expressions for P magnitude and phase (equation 2.13), and their direct individual (gradient) relations to the imaginary and real parts of $k_\beta \triangleq k_x \frac{l}{\beta}$ respectively. While we have chosen the macromechanical response P to derive the characteristics' expressions, we should note that near the peak, P and V are very similar, and hence it does not matter which macromechanical response we choose. We have chosen P because we have simple expressions for its magnitude and phase.

Also, recall that the response characteristics in table 2.1 are derived in the normalized frequency domain, β, rather than the traditional frequency domain, in order to provide response characteristics in a domain independent of location. The conversion from β to f is trivial for all response characteristics (as is important in order to easily use the model, since current work is reported in

[1] The more general case is also tractable, but certainly not as simple as the sharp-filter case

f-characteristics). We explain this conversion as part of this section.

In table 2.1, β_{center} could be β at which the magnitude peaks ($\beta_{center} \triangleq \beta|_{\max \text{mag}\{P\}} = \beta|_{\text{Im}\{k_\beta\}=0}$), or be the β at which the group delay is maximum ($\beta_{center} \triangleq \beta|_{\max(\tau_g)} = \beta|_{\max(\text{Re}\{k_\beta\})}$). In our case, these are equivalent - both give $\beta_{center} = b_p$. We define, $N \triangleq \frac{\beta_{center}}{2\pi} \max(-\frac{d\text{phase}\{P\}}{d\beta}) = \frac{\beta_{center}}{2\pi} \max(\text{Re}\{k_\beta\})$, and $\phi_{accum} \triangleq \frac{\max(\text{phase}\{P\}) - \min(\text{phase}\{P\})}{2\pi} = \frac{\int_{-\infty}^{\infty} \text{Re}\{k_\beta\} d\beta}{2\pi}$. Recall that phase$\{P\}$ is in radians and ϕ_{accum} is in cycles.

Phase Accumulation

One note we should make regarding phase accumulation is that, unlike all other response characteristics in the table, it is not a peak-centric characteristics. Hence, we expect that the phase accumulation may very well be inherently less accurate than the remaining characteristics whenever the contribution of the long-wave region of the data to the phase accumulation cannot be neglected - we expect this to be more the case in the apex, which has longer wavelengths, than the base. However, without sufficient data extending on either side of the peak, it is difficult to make absolute statements. Here, we assume that the predominant contribution to the phase accumulation comes from the short-wave region of the macromechanical response.

Bandwidth

The bandwidth may be defined in multiple ways. Common ones include n dB bandwidths. We define an n dB left/right bandwidth as the differential value $y_n = \beta_n - b_p$ at which the magnitude is n dB less than its value at b_p (the peak). Hence, our expression for y_n is obtained by solving, $-\frac{B_u}{2} \log(A_p^2) + \frac{B_u}{2} \log(A_p^2 + y_n^2) = \frac{\log(10)}{20} n$ for y_n. In the sharp-filter approximation, the magnitude is symmetric about the peak, at $\beta = b_p$, and hence the left and right bandwidths are equivalent, and the full bandwidth, BW_n, in table 2.1, is simply twice the right or left y_n. The n can be any number, but is typically 10 or 30.

The bandwidth is dictated by both A_p and B_u. More specifically, the bandwidth is linear in terms of A_p, and decreases with B_u. As to be expected, the bandwidth is larger for larger n - for example, the bandwidth 10 dB is larger than a BW 3 dB. For any set value of model constants A_p, B_u, the bandwidth increases more quickly for smaller n, than for larger n (provided that n is near the peak which is where the sharp-filter approximation is valid).

The conversion to the frequency domain bandwidth is simple, $y_n = \frac{1}{2}\text{BW}_n = \frac{1}{2}\frac{\Delta_n f}{\text{CF}(x)} = \frac{1}{2Q_{f,n}}$ - where y_n is the n dB left/right half bandwidth in β; $\text{BW}_n, \Delta_n f$ are the full bandwidths in the β, f domains respectively, and $Q_{f,n}$ is the quality factor derived from the n dB bandwidth in the frequency domain. This translation simplifies using the model, since most studies report a quality factor in the frequency domain, as mentioned earlier.

Group Delay

We have chosen to formulate N in this way because its formulation in β is equivalent to their formulation in f $\left(N \triangleq \frac{\beta_{center}}{2\pi}\max(-\frac{d\text{phase}\{P\}}{d\beta}) = \frac{f_{center}}{2\pi}\max(-\frac{d\text{phase}\{P\}}{df}) \right)$, and hence, this expressions for group delay can immediately be translated into the traditional domain, f. This equivalence across domains is also the case with phase accumulation.

Relation to Meanings of Pressure

Notice that the expressions in table 2.1 are consistent with our discussion on model constant meanings in section D.1. For example, the expression for y_n suggest that a smaller A_p leads to a sharper response, as is consistent with our sharp-filter approximation requirement defined in section 2.7.1, and that decreasing B_u increases the n dB bandwidth, which is consistent with figure D.1 and appendix D.1.2 .

Linearity to Determine Model Constants

Lastly, we note that linear expressions for the response characteristics are particularly desirable as they allow for determining the three model constants (and therefore all the model variables) from multiple response characteristics using linear least squares. Our expression for BW_n is nonlinear in the model constants, and cannot be linearized easily (unlike N). Therefore, it is desirable to use an alternative response characteristic that contains similar information whenever possible. The equivalent rectangular bandwidth (ERB) is a hypergeometric function of the model constants, and therefore, it is also unsuitable. We propose using the curvature of the pressure macromechanical response at the peak. It utilizes information closest to the peak (as is our focus), and is related to the bandwidth - the greater the curvature, the smaller the bandwidth. We define the peak curvature

as $S \triangleq -\frac{d^2 \text{mag}\{P\}}{d\beta^2}|_{\beta=b_p} = -\frac{20}{\log(10)}\frac{d\text{Re}\{r\}}{d\beta}|_{\beta=b_p}$. For our model, this yields, $S = \frac{20}{\log(10)}\frac{B_u}{A_p^2}$ in dB, which is simply linearized by taking the logarithm.

B.2 Converting Available Response Characteristics to Ours

In this section, we discuss how we convert available response characteristics into ours, in order to use them, as we did in section 2.9.2 to determine the wavenumber and impedances. This section is also important for a closely related application of using the *relationship* between response characteristics to infer certain aspects of the wavenumber and impedance - specifically [44] has empirically shown that the ratio of the equivalent rectangular bandwidth quality factor, Q, to the group delay, N, is approximately constant (spatially-invariant) along the length of the chinchilla cochlea [2]. This is despite the fact that both these response characteristics do vary from base to apex. It is of interest to understand what this approximately spatially-invariant ratio means in terms of constancy in underlying wavenumbers and impedances.

By translating this ratio into our characteristics, we can determine which model constant (or function of model constants) is approximately spatially-invariant, and from there, which aspects of the wavenumber and impedance are approximately spatially-invariant across the length of the cochlea. To summarize our results, we find that the model constant B_u must vary slowly across the length of the cochlea, and this interestingly means that the normalized effective positive and negative damping also have small variations. However, we must note that this does *not* immediately correspond to an approximately spatially-invariant power amplification and absorption, as the effective stiffness also factors into the power equation [3].

The remainder of this section is dedicated to methodological details on how we arrived at our result given the approximately spatially-invariant empirical ratio. In order to determine constant aspects of the wavenumber and impedance, we must first determine spatially-invariant aspects of the model constants. To do so, we must express the empirical ratio in terms of our response characteristics, and from there, in terms of our model constants. Let $g \triangleq \frac{Q_{f,erb}}{N}$, which is empirically

[2] Note that, by definition the equivalent rectangular bandwidth is the same regardless of whether it is computed in f or in β domains

[3] The pure dependence of power amplification and absorption on the effective damping is based on the sign only - a negative damping necessitates power amplification, and a positive damping necessitates power dissipation (from a classical physics perspective). However, we cannot determine the *amount* of power amplification and dissipation without having both the real and imaginary parts of the impedance (or any complete single variable)

found to be $g \approx 1.25 \pm 0.02$. We first express g in terms of our response characteristics listed in table 2.1. Using the empirical observation of [42] relating the equivalent rectangular bandwidth quality factor with the 10 dB quality factor (independent of species and location), we may write, $Q = \alpha Q_{f,10}$ where α is in the range $1.7-1.8$. In section B.1, we have expressed the relationship between $Q_{f,10}$ and y_{10} as $y_{10} = \frac{1}{2Q_{f,10}}$. From this, we can rewrite g in terms of our response characteristics: $\frac{2g}{\alpha} = \frac{1}{y_{10}N}$. Now, we have an expression for the empirical ratio in terms of our response characteristics. Next, we must express g in terms of our model constants. Using the expressions in table 2.1, we find that $\frac{\pi\alpha}{g} = b_p B_u \sqrt{e^{\frac{\log(10)}{B_u}} - 1}$. As $b_p \approx 1$, a spatially-invariant g necessitates a slowly varying B_u as mentioned above (more specifically, it is also centered around value of $B_u \approx 6 - 7$) - we use the term slowly-varying as opposed to spatially-invariant or constant due to the fact that g, α are not exactly spatially-invariant (and small variations in g are not negligible when solving for B_u). The translation of a slowly-varying B_u into a slowly-varying near-peak effective damping is based on section C.2.2.

Appendix C

Understanding the Model

C.1 Quick Intuition

In this section, we provides the reader with examples of properties that can be computed from response characteristics quickly, without the need for computing the full profile of the model wavenumber or impedance, and yet provide us with a good amount of intuition. Further details are left to section C.2.

We refer the reader back to the expressions for the model constants in terms of some of the response characteristics in table 2.2. We first introduced these model constants in the expression for the wavenumber - equation 2.5. From mapping table 2.2 to equation 2.5, it is evident that the wavenumber properties are quite tied to the response characteristics. For example, the peak normalized frequency, which is b_p, determines the imaginary part of the pole of k_x. The phase accumulation determines B_u and hence, the gain of the rational transfer function expression for the wavenumber. The model constant, A_p, which is determined by a mixture of response characteristics, gives us the negative of the real part of the pole, as well as the negative of the zero of the wavenumber. This means that we can create a quick picture of how the cochlea works based on response characteristics such group delay and phase accumulation - though we must note that phase accumulation is not easily measured experimentally, and cannot be measured in humans using current experimental methods.

We can also derive such quick intuition for the real and imaginary parts of the impedance. The model constant, B_u, which can be determined from phase accumulation, informs us about the slope

of the real part of the impedance. The group delay, which a ratio of A_p to B_u, gives us the effective stiffness.

C.2 Dependence of Wavenumber and Impedance on Model Constants

In the main chapter we have shown that we can determine the model wavenumber and impedance which encode how the cochlea works from response characteristics. In this section, we discuss the dependence of the wavenumber and impedance on model constants.

Figures C.1 and C.2 show the wavenumber and impedance generated using various values for the model constants. The figures, along with the following details, illustrate the dependence of the wavenumber and impedance on the model constants, and hence can be used to better understand the model. In addition, this section is particularly relevant for a related application for our model - determining how *changes* in response characteristics translate into changes in their underlying wavenumber and impedance (provided that the same model structure holds, and only the model constant values change). For example, if there is a different species, pathology or region, where the bandwidth and group delay are different, then we might want to know what that means in terms of how the underlying wavenumber or impedance are different.

C.2.1 Wavenumber

It is clear, from the sharp-filter formulations, that the model constants take on the following approximate meanings. These features and their model constants' dependencies are illustrated in figure C.1, and described for the wavenumber $k_\beta \triangleq k_x \frac{l}{\beta}$.

- b_p determines the β at which the maximum of $\mathrm{Re}\{k_\beta\}$, and the zero crossing of $\mathrm{Im}\{k_\beta\}$ occur. It is therefore appropriate to think in terms of, $\beta - b_p$, or deviation from a central value $\beta = b_p$.

- B_u, A_p both determine $|k_\beta|$ at any particular $\beta - b_p$. As the model constant A_p occurs as squared, the magnitude of the real and imaginary parts near the central value of β is far more sensitive to A_p than B_u. Particularly at the central value itself $\beta = b_p$, increasing $\frac{A_p}{B_u}$ has the effect of decreasing the magnitude of the peak of $\mathrm{Re}\{k_\beta\}$.

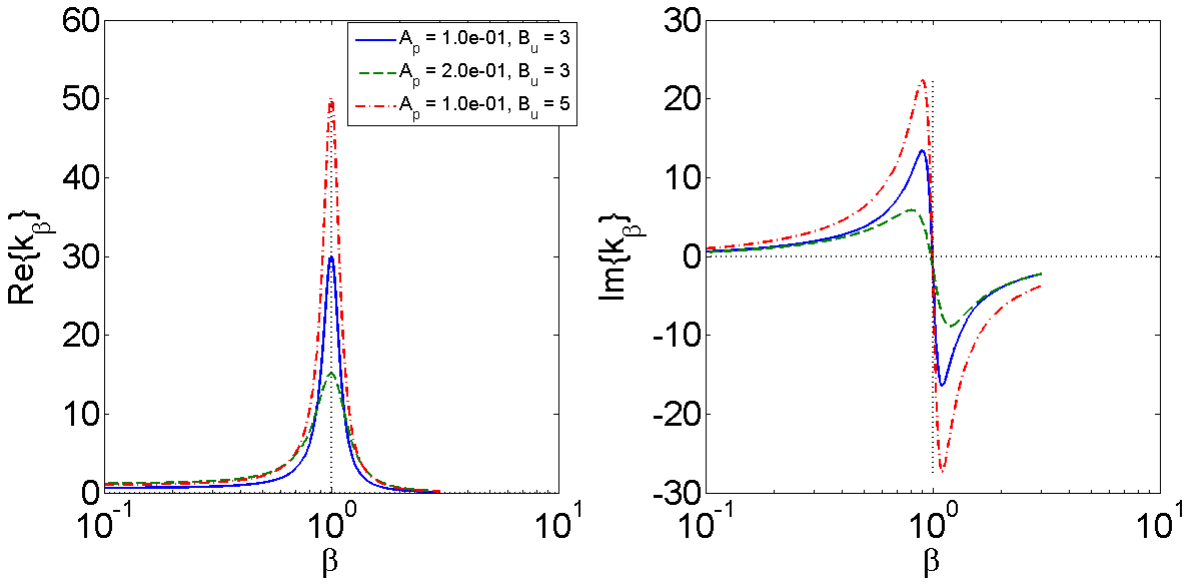

Figure C.1: Figure shows the effect of changing the model constants A_p, B_u on the real and imaginary parts of k_β (the wavenumber that describes the propagation of the pressure in the β, or the transformed space, domain). This is done by illustrating $\mathrm{Re}\{k_\beta\}, \mathrm{Im}\{k_\beta\}$ for a few choices of A_p, B_u. The effect of altering B_u is illustrated by going from the blue solid line to the red dotted-dashed line, and the effect of altering A_p is illustrated by going from the blue solid line to the green dashed line. As can be seen, and described further in the text, both A_p, B_u control the magnitude of k_β, and A_p controls the width of the peaks and trough of $\mathrm{Re}\{k_\beta\}, \mathrm{Im}\{k_\beta\}$.

- A_p determines the spread of $|k_\beta|$ in β, as can be seen from widths of the peaks and trough of the imaginary and real parts of k_β.

Therefore, A_p controls the shape of these functions and not just their magnitude, as opposed to B_u which simply only controls their magnitude. This can be seen from the fact that the $\max, \min \mathrm{Im}\{k_\beta\} = \pm\frac{B_u}{2A_p}$ (which shows that both A_p, B_u control the magnitudes), and the fact that this minimum and maximum occurs at $\beta = -b_p \pm A_p$ (which shows that only A_p predominantly controls the shape or width of the functions).

It is relevant, for certain purposes, to extend these understandings regarding model constant meanings to gain ($\propto \mathrm{Im}\{k_\beta\}$) and wavelengths ($\propto \frac{1}{\mathrm{Re}\{k_\beta\}}$).

C.2.2 Impedance

It is clear, from the sharp-filter formulations, that the model constants take on the following approximate meanings. Recall that the Z should be interpreted exclusively as an effective impedance at a fixed location as a function of frequency, and specifically in terms of the effective damping and stiffness. These features and their model constants' dependencies are illustrated in figure C.2.

- b_p determines the crossing frequency of negative to positive damping (which in turn are often associated with amplification and dissipation)
- B_u and A_p determines the amount of stiffness as $\frac{A_p}{B_u}$
- B_u determines the amount of negative and positive damping (and hence the amount of gain and dissipation [1])

The above are the exact meanings for the sharp filter approximation, and are approximate meanings for the general case. The sharp filter approximation is a good approximation around $\beta = b_p$, as can be seen from the figure and the expressions of section 2.7.2.

[1] The reader must note, however, that if they wish to translate these into power gain and loss, then other factors are involved as well

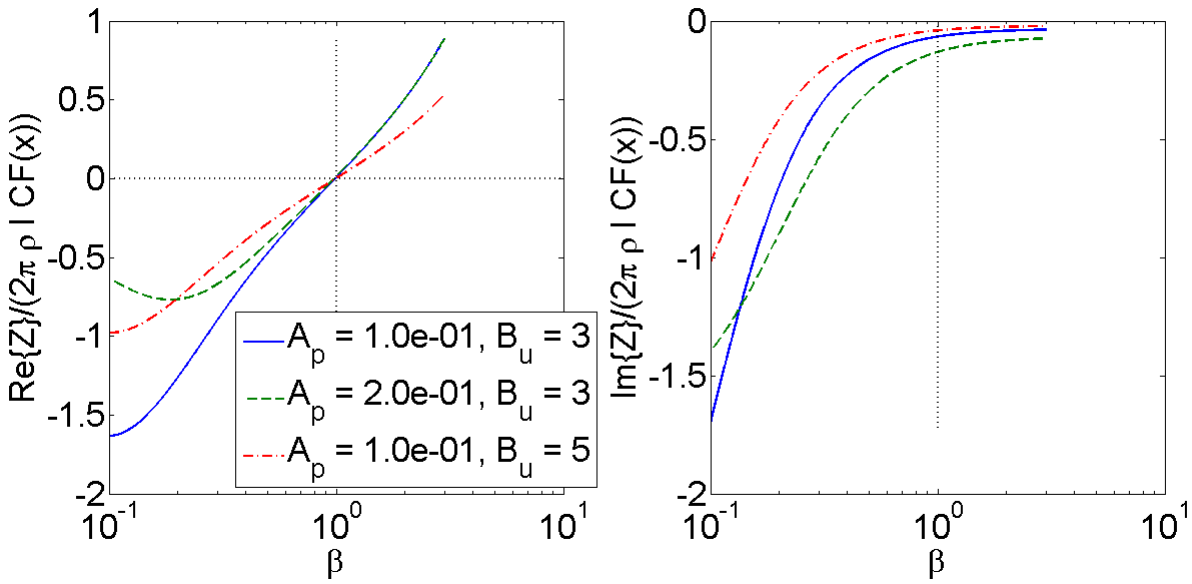

Figure C.2: Figure shows the effect of changing the model constants A_p, B_u on the real and imaginary parts of the normalized Z. This is done by illustrating $\mathrm{Re}\{Z\}, \mathrm{Im}\{Z\}$ for a few choices of A_p, B_u. The effect of altering B_u is illustrated by going from the blue solid line to the red dotted-dashed line, and the effect of altering A_p is illustrated by going from the blue solid line to the green dashed line. As can be seen, and described further in the text, both A_p, B_u control the amount of effective stiffness of $\mathrm{Im}\{Z\}$, and B_u controls the amount of positive and negative damping of $\mathrm{Re}\{Z\}$. We plot a normalized version of the impedance, as it is defined at any particular location. While the $\mathrm{Re}\{Z\}, \mathrm{Im}\{Z\}$ scale according to location, their qualitative behavior is the same regardless of region.

Appendix D

Further Model Macromechanical Response Details

D.1 Details Regarding Pressure Response

In this section, we provide the reader with additional information about the pressure macromechanical response. This is useful for engineering applications related to the design of auditory filters, as well as for scientific study such as relating observations regarding P to underlying wavenumber and impedance properties (through the dependence of P on model constants).

D.1.1 Description of Behavior

As can be inferred purely from equation 2.11 and the behavior of k_x, the pressure macromechanical response P in the traveling wave domain has a growth, peak and decay, and that the waves propagate purely in the forward direction which necessitates a positive group delay. As can be seen from the equations of section 2.7.3, P is scaling symmetric [1] and hence there is a direct correspondence between the traveling wave and transfer function properties of P. Hence, the fact that there is a growth, peak and decay in P in the traveling wave domain, translates into these specific behaviors in the transfer function domain as well.

[1] This is exactly true due to our chosen effective boundary condition. Using an alternative boundary condition somewhat decouples the traveling wave and transfer function domains and allows for negative group delay in the transfer function domain beyond the peak while retaining positive group delay throughout for the traveling wave domain as is required for forward traveling waves. See section A.5 for details on this and how altering the choice of boundary condition alters this and to what extent and where it does so.

D.1.2 Model Constants' Meanings

It is clear, from the equations of section 2.7.3, that the model constants take on the following approximate meanings from the pressure perspective. Note that these come from the fact that $\frac{d}{d\beta}\text{mag}\{P\} = \frac{20}{\log 10}\text{Im}\{k_\beta\}$ (in dB), and $\frac{d}{d\beta}\text{phase}\{P\} = -\text{Re}\{k_\beta\}$ (in radians) *individually* which is very powerful. These features and their model constant dependencies are illustrated in figure D.1.

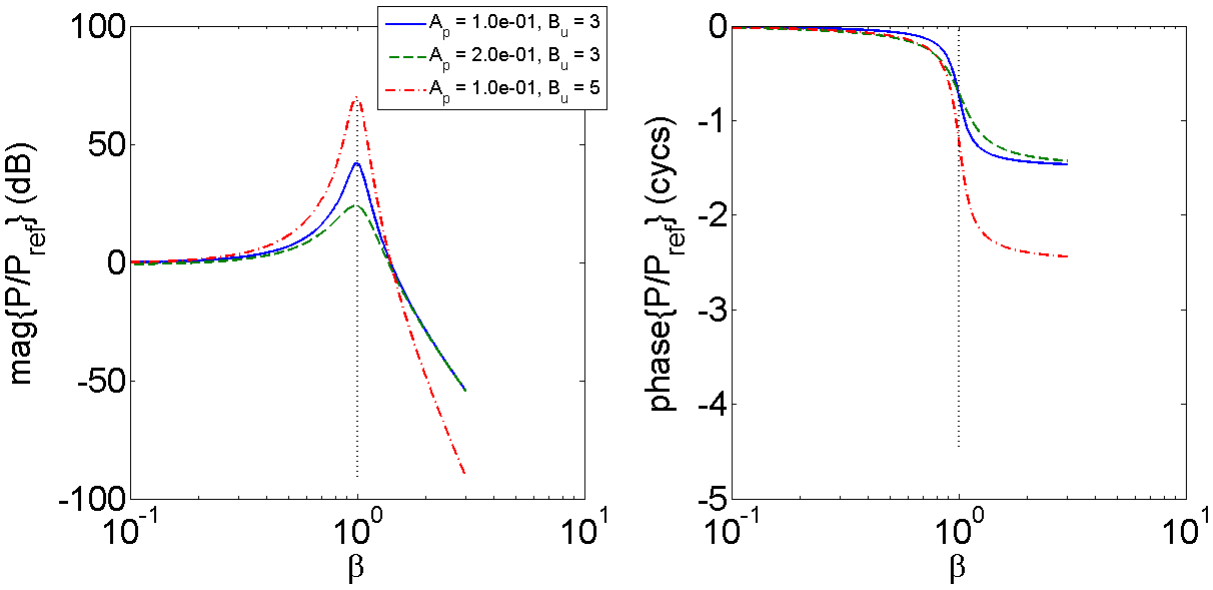

Figure D.1: Figure shows the meaning of model constants A_p, B_u as seen from the perspective of P. This is done by illustrating $\text{mag}\{P\}, \text{phase}\{P\}$ (up to a constant) for a few choices of A_p, B_u. The effect of altering B_u is illustrated by going from the blue solid line to the red dotted-dashed line, and the effect of altering A_p is illustrated by going from the blue solid line to the green dashed line. As can be seen, and described further in the text, B_u determines the phase accumulation, and the ratio of B_u, A_p, as described in table 2.1, determines the maximum group delay. B_u, A_p both determine the magnitude, with A_p determining the width of the peak and B_u scaling it.

- b_p determines the β at which the zero crossing of $\mathrm{Im}\{k_\beta\}$ and the maximum of $\mathrm{Re}\{k_\beta\}$ occur, which in turn determine the β at which the peak in the P magnitude and maximum group delay occur respectively

- B_u, A_p determine the shape (widths and slopes) and scale of the magnitude and phase. This is better seen by noting that B_u, A_p both determine the magnitude of $\mathrm{Im}\{k_\beta\}, \mathrm{Re}\{k_\beta\}$ at any particular β away from the peak (which corresponds to sharpness of magnitude of P and gradient of ϕ of P respectively)

- B_u determines the phase accumulation

- $\frac{B_u}{A_p}$ determines the peak group delay (which is determined by the value of $\mathrm{Re}\{k_\beta\}$ at the peak of the magnitude of P) - also see table 2.1

The above meanings are exact intended meanings for the sharp filter approximation which we constructed first, and are approximate meanings (especially true as we go further away from $\beta = -b_p$) for the general case.

D.2 Details Regarding Velocity Response

In this section, we provide the reader with additional details regarding the model velocity. We also make comments on the forms of P, V from an auditory filter perspective.

D.2.1 Description of Behavior and Model Constants' Meanings

The velocity V is determined by P and Z. For the model constants' values we deal with, it is dominated by P, as we expect based on our observations on data from [14]. Hence, to a first approximation, V magnitude and phase behavior can be described by those of P as can be seen from figure D.2, and the model constants have approximately the same meanings as they do for P.

D.2.2 On the Forms of Pressure and Velocity

Note that, if viewed from an auditory filter perspective, and if B_u is an integer, that P has B_u pairs of complex conjugate poles. It has no zeros. On the other hand, V has $B_u + 1$ pairs of complex conjugate poles in addition to a single real zero. Therefore, from a simplicity standpoint

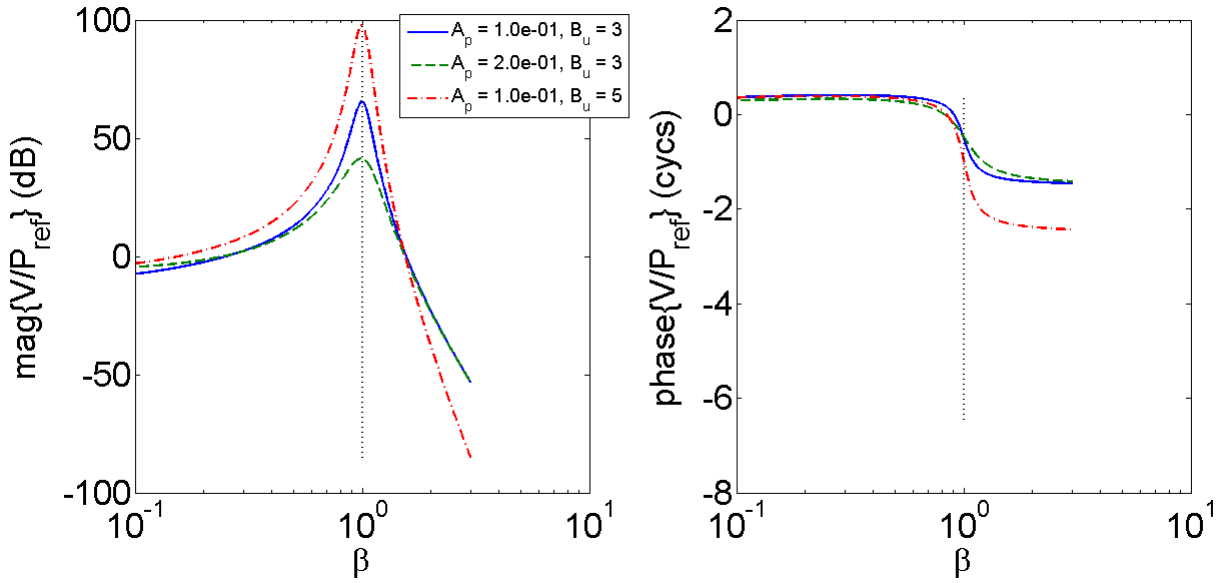

Figure D.2: Figure shows that the model behavior for V is well described by the behavior of P, and that the model constants have approximately the same meanings as they do for P.

alone, P is preferable to use for filter expressions. The expressions for P, V are very similar close to the peak for sharp filters as mentioned above, which allows for using one in place of another for applications that are *purely* related to auditory filters. Furthermore, note that knowing either P or V is entirely informative of the other macromechanical response. Therefore, having individual explicit expressions for *both* of these responses, as is the case in our model, may aid in cases where data acquisition from mechanical measurements is (1) incomplete, or (2) has low signal to noise ratio, as mentioned in 3.

D.3 Time Domain Pressure Response

In this section, we introduce the time-domain representation of the pressure, P. For integer values of B_u, our model has explicit time domain representations for pressure, which may be also be used as a filter representation. This closed-form nature of the time domain expression is useful for scientific study of pressure, and is fundamental to implementing auditory filters and studying or dictating their transient-characteristics. Specifically, we anticipate its primary uses by determining and using (see section 3): (1) the relationship between the wavenumber and impedance and time-domain response characteristics (e.g. rise time), (2) the relationship between time domain responses and frequency domain response characteristics, (3) the time domain response as an auditory filter in the time domain for applications that are preferably implemented in the time domain, (4) the correspondence between differences/changes in the time domain representation and underlying differences in the wavenumber and impedance, or alternatively, the frequency domain macromechanical responses.

We shall use the symbol q for the time domain representations to avoid confusion with the complex pole p.

D.3.1 Representation for Pressure in Time Domain

The pressure impulse response can be written as a function of normalized time (which encodes inherent location) [2],

$$\tilde{t} = \omega_{max} e^{-\chi} t = 2\pi \mathrm{CF}(x) t \; . \tag{D.1}$$

Notice that this normalized time is dimensionless. We also define, purely for the purposes of compactness,

$$\tilde{t_b} = b_p \tilde{t} \; . \tag{D.2}$$

As an example, we provide the expression for the impulse response for the case when $B_u = 2$,

[2] Time is normalized because our s can be thought of as a normalized version of the regular Laplace s which is $s' = i\omega$; our $s = \frac{s'}{\omega_{max} e^{-\chi}}$

$$\frac{q(\tilde{t})}{C2\pi\mathrm{CF}(x)} = \frac{e^{-A_p\tilde{t}}}{2b_p^3}\Big(\sin(\tilde{t}_b) - \tilde{t}_b\cos(\tilde{t}_b)\Big) \xrightarrow[\text{approx.}]{\text{highest-order}} \frac{e^{-A_p\tilde{t}}}{2b_p^3}\Big(-\tilde{t}_b\cos(\tilde{t}_b)\Big)\,. \tag{D.3}$$

And for $B_u = 3$, we have

$$\frac{q(\tilde{t})}{C2\pi\mathrm{CF}(x)} = \frac{e^{-A_p\tilde{t}}}{8b_p^5}\Big(3\sin(\tilde{t}_b) - 3\tilde{t}_b\cos(\tilde{t}_b) - \tilde{t}_b^2\sin(\tilde{t}_b))\Big) \xrightarrow[\text{approx.}]{\text{highest-order}} \frac{e^{-A_p\tilde{t}}}{8b_p^5}\Big(-\tilde{t}_b^2\sin(\tilde{t}_b))\Big)\,. \tag{D.4}$$

$B_u = 4$, yields

$$\frac{q(\tilde{t})}{C2\pi\mathrm{CF}(x)} = \frac{e^{-A_p\tilde{t}}}{48b_p^7}\Big(15\sin(\tilde{t}_b) - 15\tilde{t}_b\cos(\tilde{t}_b) - 6\tilde{t}_b^2\sin(\tilde{t}_b) + \tilde{t}_b^3\cos(\tilde{t}_b))\Big) \xrightarrow[\text{approx.}]{\text{highest-order}} \frac{e^{-A_p\tilde{t}}}{48b_p^7}\Big(\tilde{t}_b^3\cos(\tilde{t}_b))\Big)\,. \tag{D.5}$$

And $B_u = 5$, gives

$$\begin{aligned}\frac{q(\tilde{t})}{C2\pi\mathrm{CF}(x)} = \frac{e^{-A_p\tilde{t}}}{384b_p^9}\Big(105\sin(\tilde{t}_b) - 105\tilde{t}_b\cos(\tilde{t}_b) - 45\tilde{t}_b^2\sin(\tilde{t}_b) + 10\tilde{t}_b^3\cos(\tilde{t}_b) + \tilde{t}_b^4\sin(\tilde{t}_b))\Big) \\ \xrightarrow[\text{approx.}]{\text{highest-order}} \frac{e^{-A_p\tilde{t}}}{384b_p^9}\Big(\tilde{t}_b^4\sin(\tilde{t}_b))\Big)\end{aligned} \tag{D.6}$$

As we increase time, the polynomials in parentheses become increasingly dominated by the highest order terms. Therefore, for small A_p (which determines the decay rate), the expressions are well-approximated by their highest-order approximations. Practically, the approximations are very good for $A_p < 0.4$. There is no dependence on B_u as it does not occur in the exponential term. To generalize, expressions for $q(\tilde{t})$ with larger values for B_u can be simply represented using only the exponential factor and the highest order polynomial term. Hence, for $b_p = 1$, we can write,

$$\frac{q(\tilde{t})}{\mathrm{CF}(x)} \propto e^{-A_p\tilde{t}}\tilde{t}_b^{B_u-1}\cos(\tilde{t}_b + B_u\frac{\pi}{2})\,. \tag{D.7}$$

D.3.2 Description of Behavior

For the above formulations, and for any representation where $B_u \in \mathbb{Z} > 1$, the impulse response has a single frequency specified by the argument of the $\sin, \cos$ terms. The envelope of this response

grows when the polynomial portion dominates (at smaller times), then decays when the exponential decay factor $e^{-A_p\tilde{t}}$ dominates the behavior. This can be seen in figure D.3.

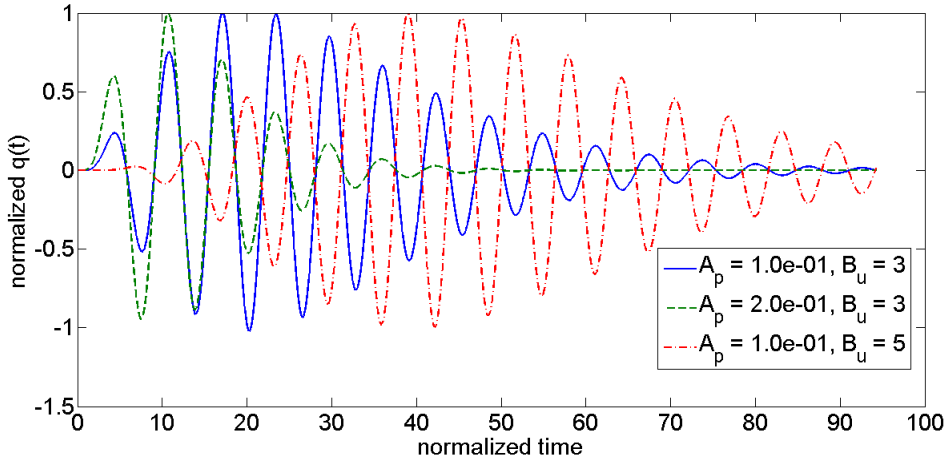

Figure D.3: The model impulse response for pressure shows that it has a ringing frequency modulated by an asymmetric envelope. Specifically, we plot $\frac{q(t)}{C2\pi\mathrm{CF}(x)}$ normalized to its maximum value as a function of $\tilde{t}$. This is exemplified for a fixed value of $b_p = 1$, for various integer values of B_u, A_p, to show the meaning of the model constants. As can be seen in $\tilde{t}$, A_p determines the decay time as we go from going from the blue solid line to the green dashed line. B_u determines the complexity of the response as well as the phase of the oscillatory (cos) factor at any given point in normalized time, and is illustrated by going from the blue solid line to the red dotted-dashed line.

D.3.3 Model Constant Meanings

It is clear, from the above formulations, that the model constants take on the following meanings. These features and their model constant dependencies are illustrated in figure D.3.

- b_p specifies the frequency of the response. Specifically, this is $f = \frac{\tilde{t}_b}{2\pi t} = \mathrm{CF}(x)b_p$ as can be seen from the arguments of the $\sin, \cos$. When $b_p = 1$, this is simply the characteristic frequency of any particular location, as according to the CF-map. This means that the frequency at which any location responds decrease with distance from the stapes.

- A_p determines the decay of the envelope. This can be determined from the later parts of the response. Specifically, the decay factor at large times is, $= e^{-A_p\tilde{t}} = e^{-A_p 2\pi\mathrm{CF}(x)t}$. This means that for larger A_p, the decay at larger times is faster. In addition, the spatial dependence, suggests that the decay at larger times is faster for locations closer to the stapes. When

expressed in terms of period of the local CF, the decay time is constant, as is reflective of the assumed local scaling

- B_u specifies the complexity of the response. Specifically, it increases the number of polynomial terms, $\tilde{t}_b^k$.
- B_u determines the phase in the cos factor at any point in normalized time, $\tilde{t}_b$.
- Both B_u and A_p specify the time at which the envelope of the response peaks which occurs at $\tilde{t} = \frac{B_u - 1}{A_p}$ (which is independent of location, as a result of scaling symmetry). Expressed in t, the time at which the envelope peaks is inversely proportional to $\mathrm{CF}(x)$. In other words, higher frequencies, which have maximal response in the base of the cochlear (closer to the stapes), peak faster than lower frequencies which have maximal response in the apex of the cochlear, further away from the stapes.

D.3.4 Comparison with Existing Work

There are currently no measurements of pressure impulse responses. However, we may compare our $q(t)$ to impulse responses of basilar membrane velocity [11] and reverse-correlation data of auditory nerve fibers [2]. We expect that these measures are qualitatively (though not quantitatively) similar based on the qualitative similarity of their frequency domain responses.

The model impulse response is qualitatively consistent with these measurements for $B_u \in \mathbb{Z} > 1$, for which the response envelop grows, then decays asymmetrically. However, while our model predicts a single frequency, the measurements [11, 2] suggest a glide settling at a dominant frequency being that predicted by our model. Note that $B_u = 1$ would provide an impulse response that will decay without a prior rise ($\frac{q(\tilde{t})}{C 2\pi \mathrm{CF}(x)} = \frac{e^{-A_p t} \sin(b_p \tilde{t})}{b_p}$), which would be inconsistent with these measurements and other instances of the model, and hence is not discussed.

In addition to experiments, we may also compare our representation with that of other models. The behavior of our $q(t)$ is qualitatively similar to the original Gamma-Tone Filter (GTF) which has a representation $at^{n-1}e^{-2\pi bt}\cos(2\pi f t + \phi)$ at each individual location [46]. Note that the GTF is purely a macromechanical response model and has no operational underpinnings (does not contain information regarding the wavenumber and impedance). It is this GTF from which many current auditory filters where designed.

Our model expressions for the pressure at a fixed location x_o (or single filter), is a summation of GTFs with different model constant values, and the highest-order approximation at a fixed x_o is equivalent to a GTF. Our model additionally provides (other than, of course, links to the other variable expressions) a dependence of model constants on $\mathrm{CF}(x)$ which is important for studying the impulse response in the cochlea and useful for design of filters centered around different frequencies, and, to our knowledge, is not an implicit feature of existing filters. Specifically, in terms of our model constants and implicit spatial variation, $n = B_u, \phi = \frac{\pi}{2}B_u, b = A_p\mathrm{CF}(x), f = \mathrm{CF}(x)$ for the case of $b_p = 1$. As the GTF is the foundation for many more recent filters, its relation to our model's q allows for leveraging existing GTF digital and analog implementations for implementing our model. Additional details regarding the GTF family is left to section H.

D.3.5 Estimating Model Constants

In chapter 2, we illustrated that the model can be determined from (frequency-domain) response characteristics. Similarly, we can determine the model from time domain responses or response characteristics. We illustrate this here by proposing a method for estimating the value of B_u from click responses.

In figure, D.3, we have illustrated $q(t)$ for two cases of B_u which controls the phase of the oscillatory factor in the response at any particular normalized time, $\tilde{t}_b$. These two cases have opposite polarity with respect to the other at the same $\tilde{t}_b$. In experimentally measured basilar membrane click responses, all the peaks measured from different locations along the cochlea approximately align in the same direction [16]. Therefore, this matching polarity (after accounting for normalized time) may also be used to determine B_u, as well its variation along the length of the cochlea.

Appendix E

Human Macromechanical Responses

In section 2.9.2, we determined two sets of model constants, ($A_p = 0.11, B_u = 7$ for the apex, and $A_p = 0.055, B_u = 7$ for the base), from available human response characteristics. Here, we determine the corresponding model macromechanical responses, pressure and velocity, for the base and the apex, shown in figure E.1. We also briefly discuss their consistency with the response characteristics used to generate them, as well as some potential applications that utilize them.

The top row shows the pressure macromechanical response generated by the model using response characteristics from humans (see section 2.9.2), and the bottom row shows the velocity. Differences in the apex (solid blue line), and the base (dashed green line) can be seen. The left panels show the magnitude which is sharper in the base, and the right panel shows the phase which has a larger slope in the base.

As can be seen in figure E.1, magnitudes of the macromechanical responses are sharper in the base than the apex, which is consistent with the data from the sharpness of tuning response characteristic that we used to generate the model curves, and increase with increasing characteristic frequency (decreasing distance from the middle ear stapes). The slopes of the phases in the figure are greater in the base than the apex, which is also consistent with the fact that we have a constant $\frac{Q}{N}$ and increasing Q with CF (or equivalently, a maximum group delay N that increases with CF). Hence, the model macromechanical responses are consistent with response characteristics used to generate the model.

These macromechanical responses, P, V have not been measured in humans as surgeries are involved in collecting mechanical data (recall that psychoacoustic measurements only provide us

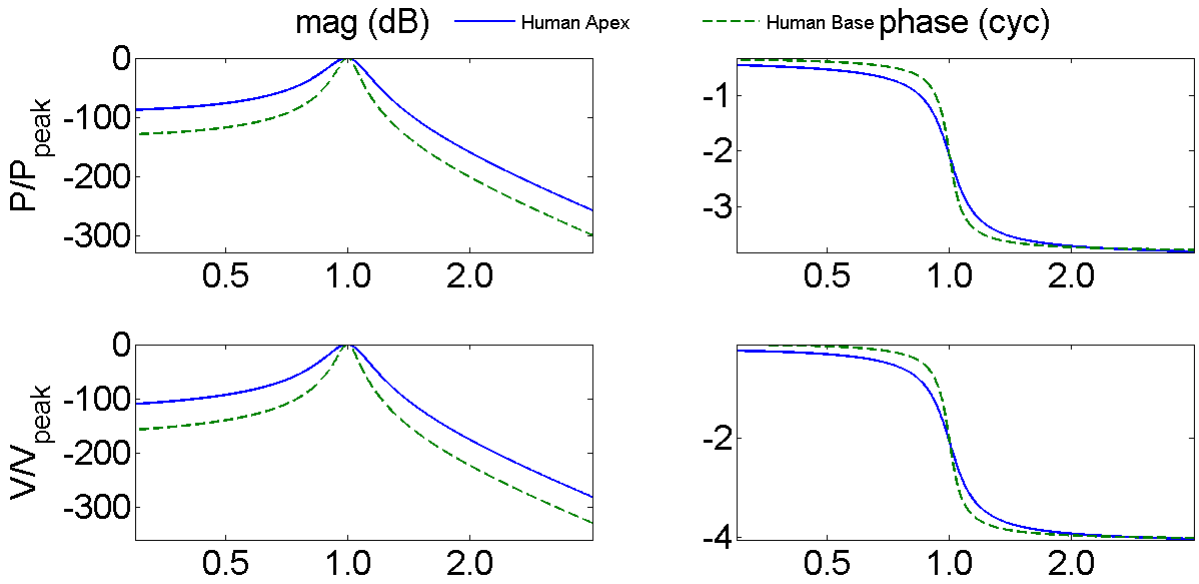

Figure E.1: Model macromechanical responses from available human response characteristics

with magnitude data). Furthermore, the model human P, V can be used towards auditory filter applications for humans (e.g. cochlear implants, hearing aids, and compressing audio files) as mentioned in section 3. The wavenumber and impedance that underlie these macromechanical responses are in section 2.9.2.

Appendix F

Terminology, Symbols and Model Components

In this section (table F.1), we provide a summary of the model components - our terminology and symbols.

Term	**Definition**	**Relevant Symbols**
Independent variables	In the frequency domain, the model dependent variables k, Z, P, V vary as a function of the independent variables, x, ω. The normalized frequency, β contains both x, ω. Each dependent variable is a function of either β, in which case the dependent variable is scaling symmetric, or a function of at most two of the independent variables, in which case the space and frequency are decoupled. The time-domain equivalent of β is $\tilde{t}$ which normalizes time, t by the period of the characteristic frequency	$x, \omega, \beta, t, \tilde{t}$

Macromechanical responses	Macromechanical responses of the cochlea or the model - e.g. velocity of the Organ of Corti (OoC) V, transpartition pressure P, may be represented in the time (e.g. impulse response) or frequency domains (e.g. as a complex construct or as separate magnitude and phase). Neural responses recorded from auditory nerve fibers are thought to be similar to the macromechanical responses	P, V
Auditory filter	A system (e.g. used in cochlear implants or audio engineering) through which an input signal is transformed to produce a desired output (e.g. one that has desired bandwidths)	(P, V)
Response characteristics	These may be response (see above definition) characteristics or filter (see above definition) characteristics. They are a description of the frequency domain representations of magnitude (e.g. bandwidth ERB, sharpness of tuning Q, skewness) and phase (e.g. peak group delay N, phase accumulation), or the time domain representation (e.g. rise time, fall time, skewness). The response characteristics can be determined from recordings of macromechanical responses, if available, or, if not (as is the case with humans), from psychoacoustic studies and otoacoustic emissions	$\beta_{center}, N,$ $\mathrm{BW}_n, y_n,$ ϕ_{accum}
Model constants	Model constants that parameterize the expressions for the model variables, k_x, Z, P, V developed in this paper	$\left(A_p, b_p, B_u\right)$ $\in \mathbb{R} > 0$

Wavenumber and impedance	Unobservable variables that encode how the cochlea works from a box-representation perspective. The wavenumber provides information from a propagation perspective - k_x describes the pressure traveling wave as it propagates along the longitudinal axis x of the cochlea, and k_β provides the same information but in the β (or transformed spatial) domain. The impedance is the ratio of the differential pressure to the partition velocity. Note that Z can only be defined in the box model as it allows for treating the OoC as a single partition. On the other hand, the k_x is the pressure wavenumber and its definition is not limited to a box model (it could, instead, represent the wavenumber of the dominant pressure traveling wave in a multi-wave model). However, its *interpretation* (e.g. with respect to power) does depend on the form of the model.	k_x, k_β, Z

Table F.1: Terminology and symbols of model components: Model variables and constants

Appendix G

General Relations Between Impedance, Wavenumber, and Pressure

In this section, we discuss the relationship between the wavenumber k_β, impedance Z and pressure P that we have observed based on the physical-component of the model alone - see section 2.2. This was useful in the model construction process. The relationship between the behaviors of these three variables is also particularly relevant, as information regarding any one of these variables also provides us with important information regarding the other two - as outlined in table G.1 and discussed below. In fact, given the similarity between P and V for the *active* cochlea, it also provides us with information regarding the fourth variable, V. A similar analysis can be performed where the long-wave approximation holds, which we expect to be relevant basal to the peak of the pressure traveling wave, and possibly for the passive case.

The behavior of $\mathrm{Re}\{Z\}, \mathrm{Im}\{Z\}$ is tightly linked to that of $\mathrm{Im}\{k_\beta\}, \mathrm{Re}\{k_\beta\}$ in our model. Any short-wave approximation of a two dimensional single partition box-representation model that contains inertial fluid coupling of the partition, has $Z = \frac{-i2\rho l \mathrm{CF}(x)}{k_\beta(x,\omega)}$. This means that, $sgn(\mathrm{Re}\{Z\}) = sgn(-\mathrm{Im}\{k_\beta\})$ and $sgn(\mathrm{Im}\{Z\}) = sgn(-\mathrm{Re}\{k_\beta\})$. In a model where $sgn(\mathrm{Im}\{k_\beta\})$ changes from positive to negative (also indicating a growth then decay in the P traveling wave based on equation 2.11) - as is the case in our model, Z must have negative then positive effective damping. Similarly, if a model has $sgn(-\mathrm{Re}\{k_\beta\})$ always negative (also indicating forward propagation of P in the traveling wave domain) - as is the case in our model, then Z must have an effective stiffness and not mass. Alternatively, if $|P|$ purely decays, then the Z would have positive damping and not negative

Variable	+	−
$sgn(\text{Im}\{Z\}) = sgn(\text{Re}\{-k_\beta\})$	• Effective mass rather than stiffness in OoC • Reverse traveling pressure wave	• Effective stiffness rather than mass in OoC • Forward traveling pressure wave
$sgn(\text{Re}\{Z\}) = sgn(\text{Im}\{-k_\beta\})$	• Effective positive rather than negative damping in OoC • Pressure traveling wave decays • (Power is absorbed)	• Effective negative rather than positive damping in OoC • Pressure traveling wave grows • (Power amplification of wave)

Table G.1: General physical relations

damping. Also, if P was a reverse traveling wave at some point, or equivalently, $sgn(-\text{Re}\{k_\beta\})$ had a region where it is negative, then Z would also have an effective mass component.

The peak(s) of the $\text{Re}\{k_\beta\}$ determines the minimum incremental wavelength (in β) of the pressure traveling wave. Peak(s) and trough(s) of $\text{Im}\{k_\beta\}$ indicate where the incremental gain or decay of the pressure wave is maximal.

The observations outlined above regarding the relationship between the wavenumber and the pressure traveling wave hold regardless of the assumptions on the boundary condition (or integration constant) needed to determine the pressure. The observations further extend to the transfer function perspective of the pressure provided that the choice of the boundary condition corresponds to P being approximately scaling symmetric $P \approx P(\beta)$.

As an aside, notice that if we assume a form for Z as a resonant simple harmonic oscillator, as was done in previous work (see section 2.6), then the real part of the impedance would always be positive, and therefore the pressure traveling wave would only decay which may very well be the case in a passive (but not active) cochlea. Furthermore, the imaginary part of the impedance would have a zero crossing from stiffness to mass dominated. Hence, the pressure phase gradient would be negative then positive, which also, may very well be the case in the cochlea. However, it would be unsuitable particularly if the pressure is scaling symmetric due to its interpretation in terms of the traveling wave direction.

Appendix H

Other Models

In this section, we shall repeatedly use the following terminology:

- Auditory filter models - models that contain response variable(s). For our model, these are macromechanical response variables, P, V. Such models can be used towards engineering applications such as designing filters for cochlear implants and compressing audio files.
- Mechanistic models - models that contain variables that encode how the cochlea works. For our model, these correspond to k, Z. Such models are primarily used for scientific study, but can also be extended, indirectly, to engineering applications.
- Controllability - the ability to control the model variables based on intended response characteristics. In section, 2.8, we described how to determine our model from response characteristics, such as the observed or (in the case of auditory filters) desired bandwidth and group delay. Here, we shall refer to this feature as 'controllability'.

H.1 Leveraging Other Models for Implementation and Further Study

In this section, we discuss how our model can benefit from our models in terms of implementations and scientific study.

H.1.1 For Implementation

Our model can be implemented or modified by benefiting from, and building on, existing work. This is quite useful as it decreases the need for spending time on creating new specialized algorithmic or implementation strategies that cater to the model. This leveraging of existing efforts is possible due to (1) our use of classical assumptions and (2) certain similarities between a subset of our derived expressions and independent previous work.

For example, the model's differential equation formulation (equation 2.11 rewritten in terms of P) allows for its modification to incorporate reverse traveling waves - which are necessary for transmission line implementations and the study of otoacoustic emissions. The former can be achieved by leveraging [10] to introduce a reverse traveling waves and hence increase the differential equation order to second order without much alteration to the forward traveling wave component. The second order equation can then then be implemented as a transmission line, which is how a good number of filters are implemented [1].

A second example is the model's relation to existing elegant decompositions of the gammatone filter - APGF (all-pole gammatone filter) or OZGF (one-zero gammatone filter) [19] through its expressions for P or V [2] allows it to utilize previous work towards time/frequency-domain analog implementations and algorithms for digital implementations that targeted silicon and power efficiency. These aspects we can utilize are particularly critical for machine hearing and auditory engineering as mentioned in figure I.1. In addition, the OZGF exhibits level dependent gain [19], and by similarity to the model, we expect that this may guide incorporating level-dependence into our model (and perhaps extending this level dependence into the wavenumber and impedance to

[1] The reverse traveling component is generally considered negligible (and hence the choice for the wavenumber for the reverse traveling waves is inconsequential) as the forward traveling wave decays sufficiently by the apex for relevant frequency components.

[2] The expression for the APGF, originally derived by finding an approximation for an algorithmic implementation for the original GTF, is $\frac{b^N}{(s^2+as+b)^N}$ where $s = i\omega$ and a, b are real parametric constants as can be inferred from equation 5 of [19]. This is quite similar to our expression for P (that only has pairs of complex conjugate poles) if our C is determined entirely from the other parameters, $C = b^N$, and only for integer (N) values of B_u. The expression for the OZGF, is $\frac{b^{N-0.5}(s+c)}{(s^2+as+b^2)^N}$ as can be inferred from equation 7 of [19]. This is quite similar to our V (that has a real zero in addition to pairs of complex conjugate poles) if our zero is not imposed to be the negative of the real part of our pole, our C is determined entirely from other parameters $\frac{iCB_u e^{\chi}}{\rho l \omega_{max}} = b^{N-0.5}$ and only for integer $(N-1)$ values of B_u . Note that these APGF and OZGF models are in a single independent variable ω, and the parameters are chosen for *each* location. This is as opposed to encoding (or benefiting from) an implicit spatial variation as in our model. Due to such differences, some degree of modifications will be required in order to use the implementations/algorithms for our model, but the fundamentals of implementation should be similar

ultimately understand what alterations within the cochlea underlie the level dependency of V data).

H.1.2 For Scientific Study

A third example, is the model's relation to the recent Zweig 2015 model [56] [3]. The Zweig model has translated its impedance into an elegant representation of the dynamics of the force balance on the OoC which we can cater towards our model and use it as a starting point for explaining the dynamics of the system behavior as mentioned in figure I.1 [4]. For instance, we may use a modified version of Zweig's dynamic representation along with our parameterizaion of model constants in terms of response charactersitics in order to promote deeper understanding when studying what aspects of the underlying wavenumber and impedance are important for cochlear implants users.

In addition to leveraging existing models for implementation and further study, the similarities between our model and those mentioned above strengthen the support for the expressions of these models and our own. This is especially true as the derivation methods are fundamentally different than ours and yet arrive at similar expressions for subsets of our variables.

H.2 General Comparison with Other Models

H.2.1 On Model Variables and Applications

Current models purely for auditory filter design have no physical basis, and therefore cannot benefit in its design from information regarding the true cochlear system, such as the inherent spatial variation of bandwidth that is encoded in some mechanistic models of the cochlea such as [55]. Nor can it be used for determining which aspects of the system are important for various filter functions or scientific phenomenon such as masking.

[3] The Zweig model is a 4-parameter model of the mechanistic model type that is centered around the impedance. The additional parameter of Zweig has the effect of intermixing our $\mathrm{Im}\{k_\beta\}$ with $\mathrm{Re}\{k_\beta\}$ or equivalently our $\mathrm{mag}\{P\}$ with $\mathrm{phase}\{P\}$. Since the form of the impedance is similar for a particular approximate case of Zweig's impedance (Zweig's impedance expression for this case is $\frac{1}{Z_{\hat{s}}^{\pm}} = \frac{s}{\mathcal{M}\omega_c\delta}(\frac{a_s^{\pm}}{s-\hat{s}} + \frac{a_s^{\pm *}}{s-\hat{s}^*})$ where $s = i\beta$, and $a_s^{\pm *}, \hat{s}$ are complex parametric constants and $\mathcal{M}, \delta$ are real constants of known values and $\omega_c = 2\pi\mathrm{CF}(x)$ is the known location dependent characteristic frequency - as can be seen in equation 80 of [56]), Zweig's formulation for the OoC dynamic representation of force balance can be extrapolated to our model.

[4] Note that the similarity to our model breaks down for the response variables P and V (and hence the reader must be cautious not to extrapolate between the models on that front). This is because Zweig's P is not scaling symmetric, his P and V are integral (not closed-form) expressions, and most importantly, they are in the traveling wave domains and cannot be transferred to the transfer function domain without specification of the role of the boundary condition (and therefore cannot be used as filter representations as it would require integrating to arrive at the traveling domain representation for each individual frequency, then sampling different frequencies).

On the other hand, current mechanistic models of the cochlea (that may contain representations for impedance and/or wavenumber) do not provide closed-form (or even analytic) expressions for transfer functions of macromechanical responses such as pressure and velocity [5]. Instead - at the very least, these need to be computed through iterating through the traveling wave integrations then sampling at various frequencies [6]. Therefore, it is difficult for such models to be used for filter design, or relating response characteristics (which are usually in the transfer function domain) to mechanisms. Examples include the nonparametric [12], and the parametric [56]. Conversely, our equations extend to *both* the transfer function and traveling wave domain representations (unlike current similar models that are limited to the traveling wave domain) and hence can be used to extract information on both fronts.

Note that the applications intended for our model do *not* extend to studies of detailed mechanics of the cochlea. Those are better performed with complex mechanical models of the cochlea (e.g. [32]) which offer an ability to vary model constants (e.g. tectorial membrane properties) and observe resultant changes in responses - *provided* an assumed set of values for model constants, and detailed structural properties. Naturally, as a result of their construction for their specific purposes, these complex mechanical models are inappropriate for our purposes - e.g. deriving insights regarding the wavenumber and impedance or their relations to response characteristics.

H.2.2 On Determining Model from Response Characteristics

In comparison to models purely for mechanistic studies - e.g. [12], [55], controllability is a unique feature of our model which allows for a mechanistic model to be appropriate for filter design and other applications that benefit from controllability and intuition-building [7]. In comparison to models purely for filter-design - e.g. all pole gammatone filter (APGF), one zero gammatone filter (OZGF) [30], our model requires parametrization of three model constants only for all filters as opposed to parameterizing each filter individually, while allowing for inherent spatial variation of response

[5] To our knowledge, the only model besides ours that gives closed form analytic expressions for the macromechanical responses (and their response characteristics) is Helmholtz [61]. However, that assumed an entirely uncoupled partition which was later established to be an inappropriate assumption as it does not allow for traveling waves

[6] Indeed, more complex models (e.g. [32, 26]) require complex computational simulations in the time domain then Fourier transforms. And for these, *all* controllability goes out the window

[7] The exception to this uniqueness of controllability is the Helmholtz model [61] which assumed an uncoupled partition with a resonant simple harmonic oscillator as an assumed form for impedance. However, the macromechanics of the Helmholtz model is inconsistent with the current classical view, which suggests that we must take longitudinal coupling (which allows for traveling waves) into consideration

characteristics. This is possible because of our assumption of scaling symmetry and working in the transformed independent variable, β [8]. Additionally, when compared to pure filter-design models, the proposed model offers expressions not only for filters (based on desired bandwidth, for example), but also for the wavenumber and impedance, and hence be used to extract information regarding the underlying function from the response characteristics.

[8] We have shown that our assumption of scaling symmetry of k is appropriate on local scales, as the model constants vary slowly with location - see section 2.9.2

Appendix I

Model Features and Applications

I.1 General Purposes and Criteria

Based on current literature, models of the cochlea are generally either used for scientific studies or for engineering applications. We may divide these intended applications into the following two categories. In this section, we only outline potential applications in a categorical sense - we provide specific examples in section 3. The main objective in this section is to explain desirable criteria, as well as a very broad perspective on uses, so that the reader is then independently able to come up with their own specific uses, and then determine if our model is suitable for their desired applications based on desirable model criteria.

1. **Design of auditory filters** such as is useful for machine-hearing (cochlear implants and hearing aids), speech recognition, perceptual studies, audio-engineering. Current ad-hoc models do not reflect physics of the cochlea (unlike category 2), and hence cannot benefit from realism, accuracy, intuition, and physical information that it can offer. Developing an accurate auditory filter model that reflects physics in cochlea will allow for furthering these designs and studies.

2. **Study of scientific phenomenon** such as amplification, power flow, dissipation and other information encoded in variables such as the wavenumber and impedance. Current models of cochlear physics are too complex (unlike category 1) to: derive much intuition from, be used for auditory filters, or, be parametrized based on intended response characteristics.

Note that we have made a division between (1) and (2) above whereas this distinction is not always appropriate. For example, in several instances of masking studies or studies of higher nervous centers that utilize cochlear models, we would still want to be able to understand what physical aspects of the cochlea are important for various aspects of masking, or various behaviors at the higher nervous centers. Indeed, it is more accurate to say that (1) and (2) are two different aspects that may apply alone or together depending on the purpose. Therefore, a model that bridges between these purposes is very useful.

The model we have developed is appropriate for a number of scientific and engineering purposes such that it satisfies desirable criteria for those individual purposes. In addition to individual purposes, it is quite clear, that attributes of each purpose category is also very beneficial for the other category - e.g. see figure I.1. Therefore, it is a natural desire to build a model that has combined beneficial criteria, as we have attempted to incorporate into our model.

We have listed certain criteria we believe are important for any cochlear model in figure I.1. Other perspectives on criteria that may be interesting to the reader are in [30] which has a good description of ten criteria for some of the auditory filter design purposes, and in [17] which gives a good overview of criteria specific to cochlear implants and constraints on their algorithms and implementations.

To a certain degree, our model (A) satisfies criteria for the application categories mentioned in the previous section (e.g. figure I.1), and (B) bridges between these multiple purposes. Currently, most applications use models that fulfill, some but not all criteria, and hence there is much room for improvement to fulfill (A). To our knowledge, no models currently exist for (B).

I.2 Criteria Met by the Model

Over the course of this thesis, we have implicitly shown that the model satisfies many of the criteria we listed in figure I.1: the simplicity of representations through having expressions in terms of 2-3 model constants; the ability to determine the entire model based on frequency domain characteristics in table 2.2; the computational efficiency of implementation through our explicit closed-form expressions (which are inherently the most computationally efficient form - e.g. they do not require integration or Fourier transforms); the accuracy and appropriateness of our model through compar-

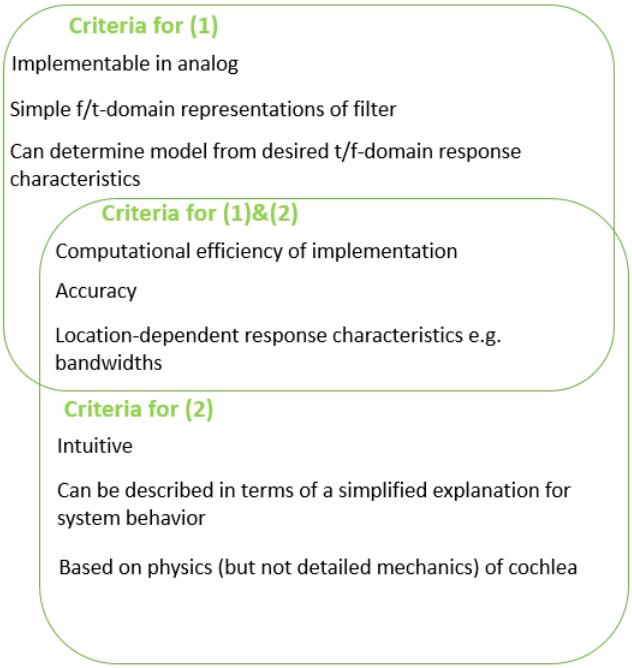

Figure I.1: Venn-Diagram: Constructing successful models for various purposes require particular criteria of which a subset is listed here. The criteria for various specific applications are lumped under (1) and (2) which are described in the main text.

ison with existing work and data; the location-dependence of response characteristics through using a transformed independent variable β; the intuitive nature through discussing meanings of model constants from the perspective of dependent variables, k_β, Z, P, V, and response characteristics; and the physical basis of our model by building on classical assumptions that relate the dependent variables to each other.

For the remainder of this section, we discuss some these satisfied criteria in more detail:

- The developed model bridges between cochlear models for a variety of purposes including scientific study and filter design for machine hearing or perceptual studies. It fits multiple criteria for these purposes, as mentioned in section I.1. An example of using this bridging is for behavioral models to understand how changes in impedance or wavenumber properties manifest as perceptual changes beyond higher neural centers.

- The fact that the model has both traveling wave and transfer function perspectives in its expressions allows for its implementation directly using various formulations: traveling wave, filter-bank, and (if modified to account for reverse traveling waves) transmission-line / filter-

cascade.

- The developed model offers a closed form expression for frequency-domain pressure and velocity representations. Closed form expressions are the most computationally efficient form compared to other representations, thereby easing studies that require iteration or multiple simulations. As an example, the fact that we have exact expressions for the time and frequency domains means that there is no need for taking FFTs, which makes matters both theoretically simpler, and computationally faster

- The developed model is simple in multiple domains and uses only two or three model constants (along with choosing location) that allow for bandwidth variation with location, without increasing the number of model constants [1]. The concept that, to first approximation only three model constants are needed to account for spatial variations is very powerful, and should certainly be tested if provided sufficient data.

- Relations to other models allows for its modification to build on previous works - see section H for details.

- The fact that the model is a single unit relating the variables, k_x, Z, P, V, and the fact that we performed tests of our model through comparing each of our model variables to *different* studies in the literature, provides independent support for the estimates of these various studies - e.g. the wavenumber estimates of [41], and the pressure of [14]. The converse is also true, with these various studies in the literature providing support to our entire model due to the fact that they each focus on a specific subsets of our variables, and are independent of one another.

- The ability to determine the model from response characteristics and to do so rather simplify is quite important as mentioned previously for several purposes. Below we list a few other examples of utilizing this:

  - If the expressions are simple enough, then the macromechanical responses/filters can be easily parameterizable using desirable level and phase characteristics. This is necessary

[1] Two if the third model constant, which specifies the peak of the magnitude of P in β coordinates is set to very close to $b_p = 1$. This is equivalent to assuming the P magnitude peak follows the CF map for a particular animal very closely. Allowing for $b_p \neq 1$ can be used to improve the fit to data, to account for data variability in CF maps.

for filter design, especially for machine hearing and perceptual studies, but also quite useful for the study of mechanisms and bridging purposes. Indeed, the importance of this property has been recognized by studies such as [28]. Short of using this parametizability or controllability to its full extent, the relationship between model constants and response characteristics can be used for initial guesses for model construction or fitting to data such as that of velocity.

– The ability to determine the impedance from response characteristics can help guide the construction of complex mechanical models of the cochlea which have a very large number of unknown (or uncertain) model constants (e.g. [32]) and generally encounter under-determined estimation problems. This use of our model to guide parameter estimation for complex models can be done by assuming, *to first approximation*, that the ratio of the scalae pressure difference to the velocity of a chosen membrane in their OoC corresponds to our Z. Parameterizing our Z based on desired response characteristics such as Q, N will lead to constructing a Z for our model, and therefore parametric constraints or initial estimates (that do *not* require entire simulations and inversion which are very computationally expensive) for the complex models. We suggest using our model in this way for initial estimates or non-rigid parametric constraints for complex models that have multiple partitions in the OoC.

Appendix J

Regarding the Physical Component of the Model

In this section, we provide further details regarding the physical component of the model beyond section 2.2. While the reader only requires the short-wave approximation of the classical physics for the main chapter, the details in this section are relevant for: understanding the model assumptions, and therefore, its limitations; and for developing modified box models - for example, by incorporating a third fluid compartment within a two-partition OoC. The symbols used in this section are listed in table J.1.

The process of derivation for the classical box model is not new and has been done before in a variety of ways [35, 45, 52, 48, 50], the fundamentals of which are similar. We discuss certain aspects here, but for a deeper discussion, the reader is encouraged to refer to these studies.

J.1 Constitutive Equations, Governing Equation, and Boundary Conditions

J.1.1 General Description

Most cochlear models treat the cochlea as an uncoiled rectangular box. All cochlear models treat the cochlea as two fluid filled compartments with an OoC interface. In classical models, the interface is treated as a single partition and only a single degree of freedom - vertical (transverse) motion, is considered. The fluid is assumed to be incompressible [35, 52]. Other assumptions are also listed in section 2.2. As described in previous work [35], the box approximation for the cochlea, usually treats it as a 2D system.

Symbol	Name
ooc	organ of Corti (synonymous with: partition, interface) also denoted as OoC
y	vertical (tranverse) distance from OoC also used to denote: vertical (transverse) component of a vector (in superscript)
x	longitudinal distance from stapes also used to denote: longitudinal component of a vector (in superscript)
sv	scala vestibuli
st	scala tympani
fl	scala vestibuli or scala tympani
u_{sv}	fluid particle velocity in scala vestibuli, vector
u_{st}	fluid particle velocity in scala tympani, vector
u_{fl}	fluid particle velocity in either scala, vector
u	differential fluid velocity $u_{sv} - u_{st}$, vector
$\mathcal{P}_{fl}$	fluid pressure in either scala
$\mathcal{P}$	differential pressure between the scala $\mathcal{P}_{sv} - \mathcal{P}_{st}$
P	differential pressure between the scala, $\mathcal{P}$, at the OoC
V	vertical velocity of OoC upwards into scala vestibuli, scalar
v_{stapes}	stapes velocity to the right into the cochlea, scalar
Z	impedance of OoC per unit area (width x length) defined as the mathematical ratio of P to V an effective impedance irrespective of any direct forms of longitudinal coupling within the OoC
ρ	volumetric density of fluid
H_{fl}	height of a scala, assumed to be equivalent for both scala

Table J.1: Symbols for physical component derivation. Note that the dependent variables such as $\mathcal{P}, P, u, V, ...$ are all in the same (frequency) domain, and we do not use capitalization to differential between domains. Also see table J.2.

Symbol	Value	Units	Name
x	$0 \to \infty$	m	distance from stapes along longitudinal axis
y	$0 \to H_{fl}$	m	distance from OoC along vertical (transverse) axis
ρ	$1e3$	kg/m^3	density of scalae fluid
H_{fl}		m	height of a scala

Table J.2: Geometric and material fluid properties: We list values for geometric and material properties of the fluid

Fluid pressure is related to fluid particle velocity by applying conservation of linear momentum to the fluid while assuming small motions, and inviscid fluid (even in the boundary layers which might not be a good approximation for some more complex models). The force balance across the interface (the pressure difference between the two fluid compartments) is model dependent.

The assumptions lead to Laplace's equation which is often treated in cartesian coordninates, and has an associated set of boundary conditions: While the longitudinal base (stapes) boundary condition and the vertical wall boundary condition are common across all models we have encountered and have not changed from the classical model times, the longitudinal apical (helicotrema) and vertical interface conditions are model-dependent.

J.1.2 Mathematical Description

Combining the constitutive equations (incompressible [1] fluid's equation of continuity, and force balance of an inviscid inertial fluid) yields Laplace's equation for the hydrodynamic pressure. Note that some of the studies start with a different partial differential equation assuming an irrotational flow (and hence deal with a velocity potential instead), but the concept is similar. Assuming the two scalae are symmetric about the partition, and have equal and opposite motion, gives us Laplace's equation for the differential pressure $\mathcal{P}(x,y) \triangleq \mathcal{P}_{sv-st}(x,y)$ as seen in table J.3 - note that the variables are implicitly a function of frequency. The use of the calligraphic $\mathcal{P}$ is to differentiate it from the $P = \mathcal{P}(x,0^+)$ we use in the main text, which is only informative of the longitudinal (x) spatial dependence at $y = 0^+$. The symbols used in this appendix are listed in table J.1. The vertical boundary conditions, as listed in table J.4 are at the rigid wall, $D_y\mathcal{P}(x,y)|_{y=H(x)} = 0$ and the partition, $\mathcal{P}(x,0) = \frac{1}{Z(x)}V(x)$. The next step is determining an analytic solution using the governing equation and the boundary conditions.

	Name	**In terms of**	**Definition/Equations**	**Explanation**
1	Incompressible fluid	u_{fl} $u \triangleq u_{sv-st}$	$\nabla.u_{fl} = 0$ $\nabla.u = 0$	Equation of continuity

[1] The compressibility contributes to the acoustic pressure component which is generally treated separately, if at all

2	Linearized inviscid fluid	$\mathcal{P}_{fl}$ $\mathcal{P} \triangleq \mathcal{P}_{sv-st}$	$\nabla\mathcal{P}_{fl} = j\omega\rho u_{fl}$ $\nabla\mathcal{P} = j\omega\rho u$	From Navier-Stokes
3	Linearized incompressible inviscid fluid	$\mathcal{P}$	$\nabla^2\mathcal{P} = \nabla.\nabla\mathcal{P} = j\omega\rho\nabla.u = 0$	Assumes uniform density; governing equation

Table J.3: Constitutive equations (1,2) to governing equation for fluid (3)

	Name	**In terms of**	**Definition/Equations**	**Explanation**
1	Rigid scalae walls	$u^y_{fl}(x, H_{fl})$	$u^y_{fl}(x, H_{fl}) = 0$ $\frac{\partial}{\partial y}\mathcal{P}(x, H_{fl}) = 0$	Vertical component of fluid velocity, assuming scalae height same
2	Force balance on OoC	$P \triangleq \mathcal{P}(x, 0)$	$P = Z(V)V$	Pressure difference across OoC (regardless of existence or type of direct coupling within the OoC itself)
3	Drive by stapes velocity and equivalence of volume velocities at oval and round windows due to assuming incompressibility	$u^x_{sv}(0, y) = v_{stapes}$ $u^x_{st}(0, y) = -v_{stapes}$	$u^x(0, y) = 2v_{stapes}$ $\frac{\partial}{\partial x}\mathcal{P}(0, y) = 2j\omega\rho v_{stapes}$	Longitudinal component of fluid velocity; Boundary condition needed to convert TW-TF
4	Semi-infinite	$x \to \infty$		Not needed (one-way)

Table J.4: Boundary conditions for 2D box version of governing equation

J.2 Partial Differential Equation to Ordinary Differential Equation

In the previous section, we described how we can arrive at Laplace's equation as well as the boundary conditions. Here, we outline the next step of providing a solution in two ways.

Note that while we can derive an exact analytic solution for Laplace's equation if we had spatially-invariant boundary conditions, the same cannot be said if the boundary conditions are spatially-dependent, as is the case above. Despite the spatial dependence on the boundary conditions, most of the literature [35, 52, 48] has assumed separation of variables, $p = X(x)Y(y)$ (where $X(x)$ specifies the longitudinal dependence, and $Y(y)$ specifies the vertical dependence of the solution), taken a single dominant component for $\mathcal{P}$, and imposed a *cosh* variation for $Y(y)$. Through this, the two-dimensional (x, y) version of the partial differential equation for pressure and the spatially-varying boundary conditions, is typically collapsed into short-wave (or long-wave) one dimensional (in x) approximations within certain conditions for validity, and with a focus on variables of the OoC motion, or immediately across it (at $y = 0^+$). This leads to the equations in section 2.2. Specifically, the idea is to take the assumed form for $Y(y)$, and formulate it to deal with an ordinary differential equation in x as opposed to a partial differential equation. In this way, we can take the short-wave and long-wave limits of the 2D derivation (based on the ratio of the wavelength to the height of the scala in *cosh*), and then reduce the equation to the one-way wave equation [2].

The resulting $x-$formulation of the short-wave equation is equivalent to other previous work [45] which derives it by working in the spatial-frequency domain (and may be more appealing than the assumed separation of variables approach outlined above). The long-wave equation is the same as previous work as well [35].

Note The reader should note that a 1D derivation (using the 1D versions of the equations in table J.3) is also in the literature. This approximates the problem in 1D, then finds a solution for the resulting Helmholtz equation (the 1D version of Laplace's equation). This is as opposed to the methods outlined above which attempt to find a 1D approximate solution to the 2D problem. If we

[2] In our case, we then translated the independent variable to β

derive the system starting from the 1D constitutive equations, then for the special case of tapering height[3], the 1D equation for P in β is the same as the long-wave approximation of the 2D equation.

[3]as opposed to, for example, a height that assumed to be constant from base to apex

Appendix K

Miscellaneous Notes

In this section, we make a few miscellaneous notes. Some of these are referred to in the rest of the thesis.

K.1 The terms OoC, interface, and partition are used interchangably here and all mean the Organ of Corti

K.2 Integration/differentiation with respect to β is done while holding ω constant (since it comes from the differential equation is set-up with respect to space and not frequency)

K.3 It is critical to differentiate between results that apply in the traveling wave (TW) domain and those that apply in the transfer function (TF) domain, and those that apply to both domains. We make notes for the equations accordingly. More specifically, some complications are created because the mechanistic equations (in which the model is constructed) are born from the TW domain, but the data and the desired responses are in the TF domain. Validation using TF-type data across multiple locations provides access to both these domains.

K.4 In many of the plots, we choose a logarithmic scale to plot β. This is because plotting in this way is similar to plotting as a function of ω (scaled to ω_{max}) on a log scale which is how transfer functions are usually best viewed. Furthermore, plotting against β on a logarithmic scala is similar to plotting as a function of x (normalized by l) which is usually how traveling waves are viewed

K.5 It is important to note which variable is assumed to be scaling symmetric for the model,

or for analysis. Specifically, in experimental measurements, the partition velocity is usually considered to be scaling symmetric, whereas here we assume that the wavenumber of the differential pressure wavenumber is scaling symmetric. This does *not* necessarily translate into scaling symmetry of other dependent variables. In other words, the term scaling symmetry should always be in conjunction with a specific variable.

K.6 In reference to the estimates for the model constants from fitting to velocity data in figure 2.8 of section 2.7.4, and the expected value for the model constant B_u and its direct relation to the ratio of response characteristics, $\frac{Q}{N}$ in section 2.9.2 and footnote [39]: Notice, in figure 2.8, that the base B_u 8.7 is indeed close to 7 which is the expectation drawn from section 2.9.2. However, this is not the case for the apex B_u because the sharp-filter approximation is severely violated.

K.7 In reference to the values for the model constants estimated from fitting to the wavenumber in figure 2.6 and to the velocity in figure 2.8: The value for B_u of figure figure 2.6 (1.3) is much less than the value we get from fitting macromechanical response data in the base (8.7). This should not be misleading regarding the sharp-filter approximation criteria which holds even for the wavenumber fits - the criteria measure here is 1/20 which is much less than 1. Instead, we expect that this discrepancy is because (1) the black line of figure 2.5 that we fit to is really a trend line and not the actual individual k, and more importantly (2) the real and imaginary parts of these lines are contradictory from the perspective of the model in which the the peak of its real part and the zero crossing of its imaginary part coincide and occur at the same beta. Also, the data sets used for fitting wavenumber and fitting velocity are different - it is in fact more appropriate to compare the wavenumber fits to neural data fits.

Bibliography

[1] Cai, H., Manoussaki, D., & Chadwick, R. (2005). Effects of coiling on the micromechanics of the mammalian cochlea. Journal of the Royal Society Interface, 2(4), 341-348.

[2] Carney, L. H., McDuffy, M. J., & Shekhter, I. (1999). Frequency glides in the impulse responses of auditory-nerve fibers. The Journal of the Acoustical Society of America, 105(4), 2384-2391.

[3] Cheng, H. (2007). Advanced analytic methods in applied mathematics, science, and engineering. LuBan Press.

[4] Crema, M. V. (2016). Can the phased array stimulation strategy be implemented using the advanced bionics cochlear implant? Doctoral dissertation, Massachusetts Institute of Technology.

[5] Cooper, N. P., & Rhode, W. S. (1995). Nonlinear mechanics at the apex of the guinea-pig cochlea. Hearing research, 82(2), 225-243.

[6] Cooper, N. P., & Rhode, W. S. (1997). Mechanical responses to two-tone distortion products in the apical and basal turns of the mammalian cochlea. Journal of neurophysiology, 78(1), 261-270.

[7] de Boer, E. (1979). Short-wave world revisited: Resonance in a two-dimensional cochlear model. Hearing Research, 1(3), 253-281.

[8] de Boer, E. (1995). The "inverse problem"solved for a three-dimensional model of the cochlea. II. Application to experimental data sets. The Journal of the Acoustical Society of America, 98(2), 904-910.

[9] de Boer, E. (1997). Connecting frequency selectivity and nonlinearity for models of the cochlea. Auditory neuroscience, 3(4), 377-388.

[10] de Boer, E. (2001). The short-wave model and waves in two directions. The Journal of the Acoustical Society of America, 109(1), 291-293.

[11] de Boer, E., & Nuttall, A. L. (1997). The mechanical waveform of the basilar membrane. I. Frequency modulations ("glides") in impulse responses and cross-correlation functions. The Journal of the Acoustical Society of America, 101(6), 3583-3592.

[12] de Boer, E., & Nuttall, A. L. (1999). The "inverse problem" solved for a three-dimensional model of the cochlea. III. Brushing-up the solution method. The Journal of the Acoustical Society of America, 105(6), 3410-3420.

[13] de Boer, E., & Nuttall, A. L. (2000). The mechanical waveform of the basilar membrane. II. From data to models—and back. The Journal of the Acoustical Society of America, 107(3), 1487-1496.

[14] Dong, W., & Olson, E. S. (2013). Detection of cochlear amplification and its activation. Biophysical journal, 105(4), 1067-1078.

[15] Geisler, C. D. (1998). From sound to synapse: physiology of the mammalian ear. Oxford University Press, USA.

[16] Guinan Jr, J. J. (2017). Personal communication.

[17] Hajiaghababa, F., Kermani, S., & Marateb, H. R. (2014). An Undecimated Wavelet-Based Method for Cochlear Implant Speech Processing. Journal of medical signals and sensors, 4(4), 247.

[18] Kale, S. S., & Olson, E. S. (2015). Intracochlear scala media pressure measurement: implications for models of cochlear mechanics. Biophysical journal, 109(12), 2678-2688.

[19] Katsiamis, A. G., Drakakis, E. M., & Lyon, R. F. (2007). Practical gammatone-like filters for auditory processing. EURASIP Journal on Audio, Speech, and Music Processing, 2007(1), 1-15.

[20] Lamb, J. S., & Chadwick, R. S. (2011). Dual traveling waves in an inner ear model with two degrees of freedom. Physical review letters, 107(8), 088101.

[21] Lee, H. Y., Raphael, P. D., Park, J., Ellerbee, A. K., Applegate, B. E., & Oghalai, J. S. (2015). Noninvasive in vivo imaging reveals differences between tectorial membrane and basilar membrane traveling waves in the mouse cochlea. Proceedings of the National Academy of Sciences, 112(10), 3128-3133.

[22] Lighthill, J. (1981). Energy flow in the cochlea. Journal of fluid mechanics, 106, 149-213.

[23] Lighthill, J. (1983). Advantages from describing cochlear mechanics in terms of energy flow. Mechanics of Hearing, 63-71.

[24] Lighthill, J. (1992). Acoustic streaming in the ear itself. Journal of Fluid Mechanics, 239, 551-606.

[25] Lighthill, J. (1991). Biomechanics of hearing sensitivity. Journal of Vibration and Acoustics, 113(1), 13.

[26] Liu, Y. W., & Neely, S. T. (2010). Distortion product emissions from a cochlear model with nonlinear mechanoelectrical transduction in outer hair cells. The Journal of the Acoustical Society of America, 127(4), 2420-2432.

[27] Loizou, P. C. (1999). Introduction to cochlear implants. IEEE Engineering in Medicine and Biology Magazine, 18(1), 32-42.

[28] Lopez-Poveda, E. A., & Meddis, R. (2001). A human nonlinear cochlear filterbank. The Journal of the Acoustical Society of America, 110(6), 3107-3118.

[29] Lyon, R. F. (1996). The all-pole gammatone filter and auditory models. Acustica.

[30] Lyon, R. F., Katsiamis, A. G., & Drakakis, E. M. (2010, May). History and future of auditory filter models. Proceedings of 2010 IEEE International Symposium on Circuits and Systems (pp. 3809-3812). IEEE.

[31] Manoussaki, D., & Chadwick, R. S. (2000). Effects of geometry on fluid loading in a coiled cochlea. SIAM Journal on Applied Mathematics, 61(2), 369-386.

[32] Meaud, J., & Grosh, K. (2010). The effect of tectorial membrane and basilar membrane longitudinal coupling in cochlear mechanics. The Journal of the Acoustical Society of America, 127(3), 1411-1421.

[33] Moore, B. C. (2003). Coding of sounds in the auditory system and its relevance to signal processing and coding in cochlear implants. Otology & Neurotology, 24(2), 243-254.

[34] Müller, M., Hoidis, S., & Smolders, J. W. (2010). A physiological frequency-position map of the chinchilla cochlea. Hearing research, 268(1), 184-193.

[35] Neely, S. T. (1978). Mathematical Models of the Mechanics of the Cochlea (Doctoral dissertation, California Institute of Technology).

[36] Peterson, L. C., & Bogert, B. P. (1950). A dynamical theory of the cochlea. The Journal of the Acoustical Society of America, 22(3), 369-381.

[37] Recio-Spinoso, A., Temchin, A. N., van Dijk, P., Fan, Y. H., & Ruggero, M. A. (2005). Wiener-kernel analysis of responses to noise of chinchilla auditory-nerve fibers. Journal of neurophysiology, 93(6), 3615-3634.

[38] Ren, T. (2002). Longitudinal pattern of basilar membrane vibration in the sensitive cochlea. Proceedings of the National Academy of Sciences, 99(26), 17101-17106.

[39] Robles, L., & Ruggero, M. A. (2001). Mechanics of the mammalian cochlea. Physiological reviews, 81(3), 1305-1352.

[40] Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. Science, 270(5234), 303.

[41] Shera, C. A. (2007). Laser amplification with a twist: traveling-wave propagation and gain functions from throughout the cochlea. The Journal of the Acoustical Society of America, 122(5), 2738-2758.

[42] Shera, C. A., & Guinan Jr, J. J. (2003). Stimulus-frequency-emission group delay: A test of coherent reflection filtering and a window on cochlear tuning. The Journal of the Acoustical Society of America, 113(5), 2762-2772.

[43] Shera, C. A., Guinan, J. J., & Oxenham, A. J. (2002). Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. Proceedings of the National Academy of Sciences, 99(5), 3318-3323.

[44] Shera, C. A., Guinan Jr, J. J., & Oxenham, A. J. (2010). Otoacoustic estimation of cochlear tuning: validation in the chinchilla. Journal of the Association for Research in Otolaryngology, 11(3), 343-365.

[45] Siebert, W. M. (1974). Ranke revisited - a simple short-wave cochlear model. The Journal of the Acoustical Society of America, 56(2), 594-600.

[46] Slaney, M. (1993). An efficient implementation of the Patterson-Holdsworth auditory filter bank. Apple Computer, Perception Group, Tech. Rep, 35, 8.

[47] Steele, C. R., & Zais, J. G. (1985). Effect of coiling in a cochlear model. The Journal of the Acoustical Society of America, 77(5), 1849-1852.

[48] Steele, C. R., & Taber, L. A. (1979). Comparison of WKB and finite difference calculations for a two-dimensional cochlear model. The Journal of the Acoustical Society of America, 65(4), 1001-1006.

[49] Strang, G. (2007). Computational science and engineering (Vol. 791). Wellesley: Wellesley-Cambridge Press.

[50] Viergever, M. A. (1980). Mechanics of the inner ear: a mathematical approach. PhD thesis, Technical University of Delft. 1980.

[51] von Békésy, G. (1948). On the elasticity of the cochlear partition. The Journal of the Acoustical Society of America, 20(3), 227-241.

[52] Watts, L. (1993). Cochlear Mechanics: Analysis and Analog VLSI. California Institute of Technology (Doctoral dissertation, Doctoral Thesis).

[53] Yoon, Y. J., Steele, C. R., & Puria, S. (2011). Feed-forward and feed-backward amplification model from cochlear cytoarchitecture: an interspecies comparison. Biophysical journal, 100(1), 1-10.

[54] Zhang, X., Heinz, M. G., Bruce, I. C., & Carney, L. H. (2001). A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. The Journal of the Acoustical Society of America, 109(2), 648-670.

[55] Zweig, G. (1991). Finding the impedance of the organ of Corti. The Journal of the Acoustical Society of America, 89(3), 1229-1254.

[56] Zweig, G. (2015). Linear cochlear mechanics a. The Journal of the Acoustical Society of America, 138(2), 1102-1121.

[57] Zweig, G., Lipes, R., & Pierce, J. R. (1976). The cochlear compromise. The Journal of the Acoustical Society of America, 59(4), 975-982.

[58] Zwislocki, J. (1953). Review of recent mathematical theories of cochlear dynamics. The Journal of the Acoustical Society of America, 25(4), 743-751.

[59] Zwislocki, J. (1953). Review of recent mathematical theories of cochlear dynamics. The Journal of the Acoustical Society of America, 25(4), 743-751.

[60] Blausen.com staff (2014). "Medical gallery of Blausen Medical 2014". Wikiversity Journal of Medicine 1 (2): 10. doi:10.15347/wjm/2014.010. ISSN 2002-4436.

[61] [Historical Overview of Cochlea]. Retrieved from http://147.162.36.50/cochlea/cochleapages /overview/history.htm