Comparative approaches to otoacoustic emissions:
Towards an understanding of why the ear emits sound

by
Christopher Bergevin

Submitted to the Speech and Hearing Bioscience and Technology
Harvard-MIT Division of Health Sciences and Technology
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
May 2007

© Christopher Bergevin, MMVII. All rights reserved.
The author hereby grants to MIT permission to reproduce and
distribute publicly paper and electronic copies of this thesis document
in whole or in part.

Author
Speech and Hearing Bioscience and Technology
Harvard-MIT Division of Health Sciences and Technology
May 25, 2007

Certified by
Christopher A. Shera, Ph.D.
Professor of Otology and Laryngology, Harvard Medical School
Thesis Supervisor

Certified by
Dennis M. Freeman, Ph.D.
Professor of Electrical Engineering, Massachusetts Institute of Technology
Thesis Supervisor

Accepted by
Martha L. Gray, Ph.D.
Edward Hood Taplin Professor of Medical and Electrical Engineering
Director, Harvard-MIT Division of Health Sciences and Technology
Comparative approaches to otoacoustic emissions:
Towards an understanding of why the ear emits sound

by

Christopher Bergevin

Submitted to the Speech and Hearing Bioscience and Technology
Harvard-MIT Division of Health Sciences and Technology
on May 25, 2007, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The ear represents a remarkable achievement in sensory physiology. It is very fast (timescales on the order of 1-100 kHz), has a large bandwidth (~10 octaves in the human), highly sensitive (threshold is ultimately determined by thermal noise), operates over an enormous dynamic range (factor of a trillion in terms of energy input), capable of sharp stimulus selectivity (e.g. frequency and intensity discrimination) and exhibits robust nonlinear behavior. As a sensor designed to detect acoustic sound pressure, surprisingly, the ear also emits sound as well. These otoacoustic emissions (OAEs) have been developed extensively for clinical applications (healthy ears emits while impaired ones do not), though their full potential has yet to be realized. Much of the effort gone into understanding OAEs has been developed within the context of mammals, where specific anatomical and physiological features (e.g. traveling waves and somatic motility) are thought to play an integral role in generation. This thesis approaches OAEs comparatively, systematically characterizing emissions in humans and an array of non-mammals (chickens, geckos and frogs) who lack these mammalian features and exhibit a diverse range of morphologies. First, our results show that for a fixed set of stimulus conditions (employing moderate intensities), emissions are relatively largest in the gecko and frog (the two species with the fewest number of sensory cells) and smallest in the human and chicken for frequencies below ~2 kHz. At higher frequencies (3-5 kHz), emissions descend toward the noise floor for the non-mammals but remain relatively constant in human. Second, OAE phase behavior indicates that emissions are generated by multiple mechanisms in the human and chicken (and possibly gecko in certain stimulus conditions), but not the frog. OAEs in all species exhibit significant delays (~1 ms or longer), those being largest in humans. Tuning can explain these delays in all species except the frog, where some additional source of delay is present. Lastly, non-monotonic growth (relative to stimulus intensity) was found in all species, suggesting the presence of multiple level-dependent sources. We interpret the observed similarities and differences in emission properties across species within the context of anatomical/physiological comparisons.
Thesis Supervisor: Christopher A. Shera, Ph.D.
Title: Professor of Otology and Laryngology, Harvard Medical School

Thesis Supervisor: Dennis M. Freeman, Ph.D.
Title: Professor of Electrical Engineering, Massachusetts Institute of Technology
What we observe is not nature in itself, but nature exposed to our method of questioning.

— Werner Heisenberg
Acknowledgments

The work described in this thesis represents the culmination of integrating inputs from many, many individuals. I could attempt to provide an exhaustive list, but surely many important people would be left out. So I will compromise: if you are reading this and we crossed paths in some form or another at least once before 5/25/07, thank you. There is (at least) a little bit of you in this document......

Particular thanks goes to the members of my thesis committee: Chris Shera and Denny Freeman (my two advisors), John Guinan, John Rosowski, Paul Fahey and AJ Aranyosi. Their input/feedback/criticisms have been invaluable.
Contents

Acknowledgments 7

1 Basis for a Comparative Study 15
  1.1 Overview ........................................ 15
  1.2 Human Peripheral Auditory Anatomy ............ 17
  1.3 Cochlear Amplification? ......................... 22
  1.4 Mammalian Otoacoustic Emissions ............... 26
  1.5 Non-mammalian Ears ................................ 32
  1.6 Non-mammalian Otoacoustic Emissions ........... 33
  1.7 Summary of Approach ............................. 34

2 eOAE Methods 37
  2.1 Measurement System .............................. 37
  2.2 Stimulus Paradigms and Analysis ............... 38
    2.2.1 I/O Delays .................................. 40
    2.2.2 System Distortion ............................ 41
  2.3 Subjects .......................................... 42
  2.4 Data Visualization ............................... 47

3 Comparison of SFOAEs Across Species 49
  3.1 Basis for a Comparative Study of OAEs ........... 49
  3.2 Overview of Anatomical Differences ............. 51
    3.2.1 Chickens ..................................... 52
Are Otoacoustic Emissions Generated by Multiple Mechanisms in Non-mammals?

4.1 Approaches Towards Classifying OAEs

4.2 Definition of an eOAE Generation Mechanism

4.3 eOAE Results

4.3.1 DPOAEs

4.3.2 SFOAE vs. DPOAE Phase Gradient Comparison

4.3.3 DPOAE Dependence Upon Primary Ratio

4.4 Discussion

4.4.1 Comparison with Previous DPOAE Studies

4.4.2 Effect of Stimulus Parameter Choice

4.4.3 Do Non-Mammals Show Evidence for Multiple Mechanisms?

4.4.4 Differences at Lower Frequencies

4.4.5 Differences in Size
6.5.2 Comparison of eOAEs and Single-Source Model ............ 157
6.5.3 Further Evidence for Multiple Sources ......................... 162
6.5.4 Non-Monotonic Behavior In ANF Responses ..................... 165
6.6 Summary ............................................. 167

7 Potpourri ............................................. 169

7.1 Other OAE Properties ........................................ 169
  7.1.1 Effect of Death ........................................ 169
  7.1.2 Ossicular Interruption ................................. 169

7.2 OAEs a By-product of Cochlear Amplification? .................. 170

7.3 Effects of Physical Dimensions ................................ 172

7.4 Future Work ........................................... 173

7.5 Recapitulation ........................................... 175

A Harmonic Oscillator ........................................ 177

A.0.1 Undamped, Undriven Case - Simple Harmonic Oscillator ..... 178
A.0.2 Driven Harmonic Oscillator ................................ 179
A.0.3 Damped Driven Oscillator ................................ 181
A.0.4 Quality Factor ........................................... 183
A.0.5 Build-up Time .......................................... 184
A.0.6 Impedance ............................................. 184
A.0.7 Analogy to Acoustics and Electronics ..................... 185
A.0.8 Applying the Harmonic Oscillator to Problems in Hearing .. 186

B Traveling Waves in the Mammalian Cochlea - A Simple Model 187

B.0.9 Transmission Line ....................................... 187
B.0.10 Passive Linear Model ..................................... 190
B.0.11 Model Comparison to Experimental Data .................... 201
B.0.12 Conclusions ........................................... 205

C Using Phase Gradients to Estimate Time Delays .................. 207

C.1 Gradients in the Steady-State Condition ....................... 207
C.1.1 Delays Associated With Tuning ................. 210

D Correcting for I/O Delays in Phase Measurements of eOAEs 213
  D.1 Overview ........................................ 213
  D.2 Phase Correction .................................. 215
    D.2.1 Total System Delay ............................. 215
    D.2.2 SFOAEs ....................................... 216
    D.2.3 DPOAEs ....................................... 217
    D.2.4 Summary ..................................... 220
Chapter 1

Basis for a Comparative Study

1.1 Overview

"Cochlea wave propagation ....". So starts the seminal paper that first reported the existence of evoked otoacoustic emissions [Kemp, 1978]. Much effort since that time has gone into both characterizing OAEs experimentally and developing theory to explain their origin. Much of this theory has been largely based upon mammalian morphology, or as Kemp put it, a ‘cochlea’ that supports a traveling ‘wave’ along its length. This approach has had much success in providing a framework for our understanding of emission generation as well as their clinical interpretation. The purpose of this research is to characterize emission properties in a number of species with different inner ear anatomies and examine comparatively how emissions may be generated in different types of ears.

A classification scheme for mammalian OAEs has been proposed which attributes emissions as arising from two fundamentally different mechanisms [Shera and Guinan, 1999]. The argument focuses upon observed properties of emission phase and uses an anatomical framework based upon mammalian morphology to provide a physical basis into how the two mechanisms differ. One mechanism involves generation at the peak of the traveling wave which, due to scaling symmetry, yields a relatively frequency-independent emission phase. The other mechanism involves a coherent reflection of forward-propagating energy due to place-fixed impedance irregularities and accounts
for emissions with a rotating phase [Zweig and Shera, 1995]. Taken together, the proposed taxonomy accounts for all different characterized types of emissions, both spontaneous (SOAE) and evoked (eOAE). Due to the relatively generic anatomy of the mammalian cochlea, it is generally assumed that this framework applies across a broad class of mammalian ears (including humans). Explanation aside, different emission types have very different frequency dependencies of their phases, indicating that there are at least two different mechanisms for emission generation in the mammalian ear.

Numerous studies have examined OAEs in a wide array of non-mammalian species such as lizards [Rosowski et al., 1984; Manley et al., 1993], birds [Taschenberger and Manley, 1997] and frogs [van Dijk et al., 1996; Meenderink, 2005]. It has been shown that OAEs, arising either spontaneously or via an evoking stimulus, are not at all unique to the mammalian ear and are clearly present in a wide range of vertebrate ears. Thus the emission of sound is a fairly generic process of most ears [Koppl, 1995]. In many of these studies, OAEs have been used as an assay towards understanding some other peripheral feature, such as hair cell regeneration in the chick due to acoustic overexposure [Ipakchi et al., 2005]. However, it is not clear if emissions are generated similarly in avian and mammalian ears or in fundamentally different ways. Furthermore, there is much debate as to what similarities and differences between mammalian and non-mammalian emissions imply about current OAE and cochlear models. The connection is complicated by how the so-called cochlear amplifier may differ between mammals and non-mammals [Manley, 2001].

Our choice of species to examine here comprises several distinct inner ear anatomies. We specifically chose animals which are known (directly or indirectly) to lack basilar membrane (BM) traveling waves and/or outer hair cells (*i.e.* cell body somatic motility). Differences among the species are schematized in Fig. 3-1. For a given ear, two different views are shown to facilitate comparison. Two of the most obvious differences are the number of hair cells in a radial cross-section (as well as the

---

1Ch.4 provides a discussion of what the definition of a mechanism is.
2Significant eOAEs have also been observed in invertebrates such as grasshoppers and locusts. They have been shown to be physiologically sensitive and highly dependent upon the state of the neurons synapsed to the peripheral auditory organs [Mockel et al., 2006].
orientation of the bundles) and the shape/size of the overlying tectorium (TM) in addition to the means by which it couples to the bundles and the organ as a whole.

The purpose of this first chapter is to set a basis for what we might gain from a comparative study and how we can use that knowledge to further our understanding about how the ear emits sound. The chapter is intended to provide the reader with the background motivating Chapters 3-6. First, we provide a general overview of the anatomy and physiology of the auditory periphery and how they vary across species. This is followed by a discussion about the cochlear amplifier, providing a basis for the subsequent discussion about current thought regarding mammalian OAEs. Lastly, we pose the main questions we address in this thesis, with subsequent chapters being devoted to each question.

1.2 Human Peripheral Auditory Anatomy

The ear can be roughly distinguished into three different parts: the outer, middle and inner ear (Fig. 1-1). Sound enters the external auditory meatus, or ear canal

\[\text{Figure 1-1: Overview of the human auditory periphery [A. Greene].}\]

---

\(^3\)Detailed description of the non-mammalian anatomy is described in Ch.3.
Figure 1-2: Schematic showing BM traveling wave along a straightened (uncoiled) cochlea [Geisler, 1990].

(outer ear), and sets the tympanic membrane (TyM) into motion. On the other side coupled to the TyM, are the three ossicles which span the middle ear. To a first degree, these bones serve as impedance matchers between air-filled outer world and the fluid-filled inner ear. The final bone, the stapes, is coupled to the oval window, entrance to the snail-shaped cochlea (inner ear). For our purposes, the mammalian auditory periphery is considered relatively generic (at least relative to non-mammals) and the description below is considered to apply to all mammals (not just humans)

The cochlea is a long coiled tube comprised of three different chambers. A flexible partition called the basilar membrane (BM) separates the top two chambers (scala vestibuli and scala media) from the bottom (scala tympani), except at the most apical region of the cochlea where there is a hole called the helicotrema. The width and thickness of the BM change along the cochlear length, as does the cross-sectional area of the chambers. As the stapes footplate moves in and out in response to sound, pressure variations in the fluid will be setup inside the cochlea.

A flexible membrane called the round window allows for the volume displacements outwards as the stapes pushes inwards (since the cochlear fluid is largely incompressible).

4 Obviously there are significant differences across mammalian species. These can range from general (such as overall dimensions) to highly specific (e.g. the structure and shape of the supporting cells or special high frequency adaptations found in bats). Considering the importance of using animal models with the intention of clinical benefit back to humans, comparative OAE studies specifically across mammals are of high value and numerous researchers have explored this path [for example, see Souter, 1995].
ible). This creates a pressure difference across the basilar membrane. As a result, a traveling wave propagates along the BM as shown in Fig. 1-2\(^5\). Each point along the BM resonates at a particular frequency due to its graded mass and stiffness. It is this tonotopic organization that allows the cochlea to function as a spectrum analyzer. Higher frequencies excite basal regions (near the stapes) while lower frequencies stimulate more apically. Fig. 1-3 shows data from the seminal work of von Bekesy, showing BM displacement magnitude and phase in a human cadaver ear. Because he measured in dead ears, Bekesy's measurements fail to capture features observed in a living ear thought to arise from physiological amplification mechanisms (discussed in a subsequent section below).

Sitting on top of the BM in the scala media is a remarkable structure called the organ of corti (OoC), as shown in Fig. 1-4. The OoC contains hair cells (HCs) which effectively act as mechano-electro transducers, mapping BM motion to action potentials in the auditory nerve fibers (ANFs) innervating the HCs. A stereociliary bundle extends out of the epithelial surface of the HC that contains a unique set of transduction channels. As the BM is displaced and moves upwards in the transverse direction, there is a shearing between the BM and the overlying tectorial membrane (TM) in the radial direction\(^6\). This shearing causes a deflection of the stereociliary bundle, thereby stimulating the transduction channels [Corey and Hudspeth, 1979]\(^7\). The scala media is unique in that its fluid composition creates an extremely high potential of about +70 mV (the largest resting potential in the entire body) due to the pumping action of the stria vascularis which also causes a large \(K^+\) concentration. This potential is quite high compared to the resting potential of the hair cell, which is at about -60 mV. As a result of this large potential difference, small bundle deflections

\(^5\)See Appendix B for further discussion.

\(^6\)The TM is an anisotropic, extracellular gelatinous matrix. It has unique structural, electrical and mechanical properties that likely play a role in energy propagation throughout the cochlea. However, its exact role is still not currently well understood due to the difficulty associated with TM-based measurements.

\(^7\)Everything up to the level of BM displacement is generally referred to as macro-mechanics. However, motions at the level of the HC bundle and TM-recticular lamina displacements are referred to as micro-mechanics. Motions on these smaller length scales are currently not well understood and thus the exact mechanism by which HC transduction channels are excited is not known.
Fig. 3.7 Travelling waves in the cochlea were first shown by von Bekesy. The full lines show the pattern of the deflection of the cochlear partition at successive instants, as numbered. The waves are contained within an envelope which is static (dotted lines). Stimulus frequency: 200 Hz. From von Bekesy (1960, Fig. 12.17).

Fig. 3.8 Displacement envelopes on the cochlear partition are shown for tones of different frequency. The lower plot shows the relative phase angle of the displacement. From von Bekesy (1960, Fig. 11.58).

Figure 1-3: Data obtained by von Bekesy in human cadaver ears [Bekesy, 1960].
cause appreciable changes in HC membrane potential, triggering a synaptic release of neurotransmitter to the innervating nerves.

There are two distinct types of HCs in the mammalian inner ear: inner hair cells (IHCs) and outer hair cells (OHCs). There are roughly three OHCs for every IHC. IHCs receive the bulk of the afferent innervation (going to the brain) while OHCs receive the bulk of the efferent innervation (coming back down from the brain). Mammalian OHCs appear unique in that they exhibit somatic cell motility, a process by which the cell changes its length in response to mechanical or electrical stimulation [Brownell et al., 1985]. It is commonly believed somatic motility plays an important role in cochlear mechanics (further described in subsequent sections).
1.3 Cochlear Amplification?

Most animals are able to hear over a very large range of sound pressure amplitudes (dynamic range)\(^8\) and frequencies (10 octaves for humans). The lower end of the dynamic range constitutes very small amounts of energy transmitted into the inner ear, comparable to the thermal energy present in the fluids. It was suggested [Gold, 1948] that in order for the ear to be able to detect such small inputs, there must be some sort of active mechanism wherein the ear adds energy to the signal in order to boost its detection (i.e. amplification). The idea is that an active contribution is necessary in light of the input being comparable to the noise floor and the presence of mechanisms for energy loss in the sound conduction pathway (such as the viscous damping of the cochlear fluids).

The notion of the ear being capable of generating power to boost its ability to detect a signal was not taken very seriously until the 1970's when two important

---

\(^8\)The dynamic range is about 120 dB for humans, or a factor of a million. Sound intensity goes as the square of pressure, so the human dynamic range spans a million million fold range in energy!
developments took place. One was the discovery of OAEs in a normal healthy human ear [Kemp, 1978]. OAEs were discovered in the course of trying to understand the micro-structure in audiograms, which can vary significantly from person to person. While the initial measurements were made using an evoking stimulus, it was soon found that the ear can spontaneously emit sound (i.e. in the absence of any stimulus) that can be measured in the ear canal using a low-noise microphone. The other development stemmed from the measurements of BM displacement in a living squirrel monkey [Rhode, 1971]. Rhode demonstrated *nonlinear growth of the BM velocity* with respect to stimulus intensity (*compressive growth* at CF). The measurements of von Bekesy shown in Fig. 1-3 taken in the apical end of the human cadaver temporal bone do not show nonlinear behavior such as that seen by Rhode. This is in agreement with the observation by Rhode that the nonlinear response was highly physiologically vulnerable. Dead ears behave linearly while healthy ones exhibit robust nonlinear behavior.

It should be emphasized here that compressive growth does not directly imply that an active mechanism is present. Rhode's observation of the high degree of physiological vulnerability is more striking in that regard. Compressive growth can be seen as beneficial, extending the range of intensities one is sensitive to when considering the constraint of the limited dynamic range of inputs to the central nervous system (see Fig. 1-5).

The presence of OAEs and compressive growth led to the incorporation of an *active* element into cochlear models [Kim *et al.* 1980, Neely and Kim, 1983] and the term *cochlear amplifier* was eventually coined [Davis, 1983]. These models argued that an active contribution was necessary in order to account for the sensitivity and sharp frequency selectivity that the ear was known to be capable of. Furthermore, the knowledge that a healthy ear and a dead ear yielded two very different types of responses (such as whether it exhibited OAEs or compressive BM growth) served to elucidate the notion of a *second filter*.

---

9 This 'second filter' was devised as the means to explain the discrepancy between the behavior of the BM (insensitive and broadly tuned) and ANF responses (highly sensitive and sharply tuned).
Rhode's data were later re-examined using a cochlear model that took an inverse approach [Zweig, 1992; also see de Boer, 1996]. Here, the form of the model was assumed (based upon simple physical and anatomical assumptions) and a 'backward' approach was taken to solve for the model's parameters using the measured data as the model solution\(^{10}\). It was found that in order for the model to predict the measured data, it must exhibit a region where the real part of the cochlear input impedance was negative (i.e. active energy contribution).

All these observations described above provided a foundation for the notion that an amplification mechanism(s) was present in the ear that serves to boost detection of low-level signals. These mechanisms would thus play a pivotal role in extending the dynamic range of the ear (Fig. 1-5) and the overall nonlinear level growth response. Furthermore, active elements inside the inner ear could account for OAEs, particularly spontaneous emissions. The fact that responses were highly physiologically vulnerable only strengthened this argument.

The physical origin of the active mechanism is currently an issue of intense debate. There are two prominent sides which argue as to what is the basis for cochlear amplification:

- **OHC motility** - Upon bundle deflection, experiments have shown that OHCs undergo a conformational change in cell length (with evidence pointing towards a strong role played by the membrane protein *prestin*). It is thought that these cell length changes mechanically couple back to BM motion, amplifying it [Brownell *et al.*, 1985]\(^{11}\). OHC motility appears to be unique to mammalian ears.

- **Transduction channel dynamics** - The mechanically gated ion channels at the tips of the HC bundles have been shown to exhibit appreciable forces when stimulated and even oscillate spontaneously [Le Goff *et al.* 2005]. While the

---

\(^{10}\)Inverse approaches are difficult, particularly in the regard that they rarely provide a unique solution.

\(^{11}\)It has also been proposed that changes in OHC stiffness (due to prestin conformational changes) are an important factor [He *et al.*, 2003]. Current research has cast some doubt on this notion however, suggesting stiffness changes may be artifactual [Hallworth, 2007].
elements constituting these channels are not well understood, it is thought that they can exert forces in a physically beneficial way to effectively boost the bundle’s motion\textsuperscript{12}.

While neither of these two approaches has convincingly addressed the issue, they have provided much insight into what the underlying mechanism could be and how various components could work together to form it. It is plausible that both the above processes are present in the mammalian ear and work together. Furthermore, it remains to be seen how possible amplification mechanisms differ in mammals and non-mammals (evidence in the latter has indicated a lack of hair cell somatic motility), one of the underlying themes motivating this thesis.

Based upon the observation of spontaneous otoacoustic emissions alone, it is clear that there is some sort of active mechanical mechanism which allows for a flow of energy out of the ear. Statistical analyses have indicated that SOAEs are not merely filtered noise (which might derive from thermal energy present in the inner ear fluids), but result from a stable, self-sustained oscillator\textsuperscript{13}. Furthermore, both the cochlear responses indicative of a healthy ear (such sharp tuning derived from ANF measurements) and OAEs show a similar time course in physiological vulnerability, indicating that the two are intimately linked\textsuperscript{14}. So while the basis (and existence) of cochlear amplification is still hotly debated, OAEs present a fruitful means towards understanding a possible role of active mechanisms\textsuperscript{15}.

\textsuperscript{12}The ion channels at the tips of the stereocilia have a strong nonlinear dependence upon Ca\textsuperscript{2+}. This observation sets the basis for a model where the active mechanical mechanisms at the bundle tip rest on the boundary of an instability [Choe et al., 1998]. As a result, a small deflection can push them into an unstable region where amplification can occur.

\textsuperscript{13}Some descriptions have proposed that SOAE generation is highly analogous to the amplification process taking place in a laser cavity [Shera, 2003].

\textsuperscript{14}This link has been extensively exploited with the development of OAEs for clinical use, as described in the next section.

\textsuperscript{15}As a caveat, we should emphasize the point that the term cochlear amplifier is not well defined and tends to be used in a number of highly variable contexts.
1.4 Mammalian Otoacoustic Emissions

Much effort has gone into understanding OAEs and how they are generated. This is an important step in furthering our understanding of how the underlying cochlear amplification processes function. A key motivation of this thesis stems from the notion that there are many popular misconceptions and varied consensus about the nature of the *cochlear amplifier* as noted above. One advantage of the OAE approach (as opposed to more direct physiological measurements) is that OAEs can be measured non-invasively. This provides a window into a healthy, normal-functioning ear (direct physiological visualization of the inner ear requires extensive surgery that can damage the sensitive structures inside). However, as a disadvantage, the ear is composed of many parts and OAEs provide effectively only a 'scalar' measurement coming from this complex system\(^\text{16}\).

In the mammalian inner ear, pressure differences propagate along the length of the cochlear partition, setting up a traveling wave (Fig. 1-2). Along its length, there is an impedance gradient of the BM due to changes in width/thickness/.... (downward trend seen in Fig. 1-6). This gradient gives rise to a tonotopic distribution, where different locations resonate at different frequencies. From here on out with respect to notation, I will use BM to refer to the cochlear partition for simplicity. However, the actual partition itself is composed of multiple structures (the basilar membrane, organ of Corti and tectorial membrane).

One theory of mammalian OAEs postulates that certain types of emissions arise due to a back-scattering of energy propagating along the length of the BM \cite{Zweig1995}. The basic idea is that in addition to the gross changes in BM impedance, there is a continuous distribution of random impedance irregularities along the length of the BM ('roughness' as shown in Fig. 1-6). As the traveling wave propagates along the cochlear length \cite{Bekesy1960}, the response will grow at the characteristic place for a given frequency component of the stimulus (Fig. 1-7). This

\(^{16}\)Specifically, one needs to tread carefully when equating OAE properties back to specific components of the inner ear, where the whole is more than the sum of the parts ('*categorial novum*' as outlined by the German polymath Nicolai Hartmann; also see Gestalt theory for parallels as well as *The Dialectical Biologist* by R. Levins and R. Lewontin).
Effective stiffness of cochlear partition changes along its length due to varying thickness, mass, etc. Additionally, it is hypothesized that there is a uniform distribution of random impedance irregularities. These can arise from a number of factors such as sudden changes in cell density or membrane stiffness, micro-variations in the fibrillar structure of the BM or TM, or differences in the coupling of the micro-mechanical structures (e.g., HC-TM connections). Adapted from Shera and Guinan, 2006.

will result in a region of scattered energy (due to the impedance irregularities) that can interfere constructively to form a backward-propagating signal. Back-scattered energy is thought to be the basis for reflection-based OAEs (as described below). The active mechanism(s) is thought to play an important role here (particularly at lower levels) in order to help boost the response amplitude at the characteristic place. This model has shown consistent agreement with physiological data [Shera and Guinan, 2006].

This reflection-based framework has been further developed to explain both spontaneous emissions [Shera, 2003] and aspects of other emission types [Kalluri and Shera, 2001]. It is most beneficial at this point to develop some of the background on cochlear mechanics necessary to describe the model's implications/predictions, which are important aspects of the proposed thesis.

Fig. 1-7 shows a snapshot of the BM for two different single tone inputs (dark curve for frequency $f_1$ and grey curve for frequency $f_2$, where $f_2 > f_1$). The ‘snapshot’ is taken at the same point in the stimulus phase for the given frequency. The purpose of the figure is to demonstrate two different principles:
Wave shifts when frequency increases $f_2 > f_1$

KEY POINTS:
1. Wave pattern shifts with frequency
2. # of wavelengths is preserved

Figure 1-7: Snapshot of basilar membrane traveling wave for two different single tonal inputs. Goal of figure is to demonstrate concept of scaling symmetry. See text for description.

1. **Tonotopy** - The cochlear partition’s width and thickness increase along the length (longitudinally) of the mammalian cochlea and that forms the basis for a tonotopic gradient. The frequency at which a given location along the BM has a maximum amplitude response is called the characteristic frequency (CF) for that spot. Basal locations have higher CFs and they decrease monotonically as you move towards the apex. Thus, frequency maps to space (e.g. the characteristic place $\leftrightarrow$ characteristic frequency).

2. **Scaling Symmetry** - Measurements indicate that the shape of the traveling wave vibration pattern remains fairly constant (at least locally) as you move along the length of the BM with respect to the CF at that given location. Qualitatively, the number of wavelengths accumulated between the stapes and the peak of the response will be approximately constant for different frequencies/locations. In more quantitative terms, the BM response is said to *scale*. The response at a given point is no longer dependent upon both the stimulus frequency $f$ and the CF at that spot, but just the ratio $f/\text{CF}$ of the two. The place (or frequency) scale is effectively *normalized* by the characteristic place (CF).
These two properties have consequences for the frequency dependence of OAE phase as shown in Fig. 1-8. The purpose of this figure is to visually show how the phase of the emission can have a different frequency dependence based upon the underlying mechanism responsible for generation. OAE data have shown emissions evoked using one stimulus paradigm result in emissions whose phase is relatively invariant with stimulus frequency (DPOAEs), while a different stimulus paradigm evokes an emission whose phase depends strongly upon frequency (SFOAEs).

As described earlier, there is a gradual variation in BM impedance along its length giving rise to a tonotopic distribution. Additionally, superimposed on top of this there is a uniform distribution of small impedance irregularities (Fig. 1-6). Due to sudden jumps in impedance between adjacent locations, these irregularities will affect the forward-propagating energy, reflecting some back towards the stapes. This will be more prevalent in regions where the response amplitude is maximal. For simplicity,
we conceptualize to one spot of irregularity as shown shown on the left of Fig. 1-8. Since this spot is scattering energy backwards, it is thus acting as an emission source. As the frequency of the stimulus changes, the phase at this source location will change. We refer to the phase response of an emission generated by such a reflection mechanism as place-fixed. Conversely, if the emission arises at the peak of the traveling wave (perhaps due to distortion arising from cochlear non-linearities), scaling symmetry constrains the phase to vary little as frequency changes. This is shown on the right and we refer to this phase behavior as wave-fixed17.

The place-fixed versus wave-fixed distinction was used as a basis for the OAE classification scheme shown in Fig. 1-9. It describes mammalian emissions as arising from two fundamentally different mechanisms: linear reflection (place-fixed) and non-linear distortion (wave-fixed) [Shera and Guinan, 1999]. Emission data are classified by their phase behavior: stimulus frequency emissions (SFOAEs) have a rapidly rotating phase response (place-fixed) while \(2f_1 - f_2\) distortion product emissions (DPOAEs) have a relatively flat phase (wave-fixed). Furthermore, OAEs measured at the ear canal typically contain a mixture of both components. For example, this is reflected in the phase response for different components of \(2f_2 - f_1\) upon un-mixing [Kalluri and Shera, 2001]. It is worth noting here that the nature of cochlear non-linearity(s) that gives rise to the DPOAEs is not yet well understood. One general train of thought is that they arise at the point of transduction in the hair cells.

As mentioned earlier, the theory of reflection-based evoked emissions has also been further developed to explain the generation of SOAEs [Shera, 2003]. In this model, waves can be reflected not only from coherent reflection via BM irregularities (forming backward-traveling waves), but from the stapes as well, thereby effectively recreating forward-traveling waves. This process thereby sets up standing waves in the cochlea where loss of energy (from both viscous forces and only partial reflection of acoustic energy at the stapes, where the 'leaked' energy then appears in the ear canal as an SOAE) is counteracted by the active gain mechanism inside the cochlea. Thus, the

17As a note, I will use the terms phase-gradient and delay interchangeably to refer to how the phase changes with frequency, though these terms are not necessarily physically the same.
Mechanism-Based Taxonomy for OAEs

Otoacoustic Emissions

OAEs that arise by
Linear Reflection

Reflection Emissions
Due to coherent reflection from 'random' impedance perturbations
Examples: Echo emissions (SFOAEs and TEOAEs) at low levels

Spontaneous Emissions
Due to standing waves caused by 'run-away' multiple internal coherent reflection
(from 'random' perturbations and stapes) stabilized by cochlear nonlinearities

OAEs that arise by
Nonlinear Distortion

Distortion Emissions
Due to nonlinearities acting as 'sources' of cochlear traveling waves
Examples: DPOAEs when coherent reflection from the DF place is negligible

Evoked Emissions
Typically a mixture of emissions produced by both mechanisms

Figure 1-9: Schematic showing proposed classification scheme for mammalian OAEs. [adapted from Shera and Guinan, 1999]

model proposes that the mammalian cochlea is acting as a hydro-mechanical analog to a laser cavity\textsuperscript{18}.

In addition to elucidating the role of a possible amplification processes, another motivation for better understanding OAE generation stems from the fact that audiologists make extensive use of their measurement in the clinic for diagnostic purposes. Emissions are generated in the normal healthy ear, so an absence or abnormality in their characteristics can indicate possible pathologies to an audiologist. Furthering our knowledge of how emissions are generated potentially extends their usefulness in the clinic. For example, if we can better understand the correlation between emission phase gradients and cochlear tuning, OAEs could provide a direct and objective measure of an individual's frequency selectivity [Shera, Guinan and Oxenham, 2002]\textsuperscript{19}. Also, by clarifying misconceptions which abound for OAEs, clinicians can avoid misinterpretation and incorrect diagnoses.

\textsuperscript{18}A strength of this specific model for SOAEs is that it makes testable predictions based upon the relation between different emission types (that have been shown to be consistent with human emission data).

\textsuperscript{19}This notion is further addressed in Ch.3 and Ch.7.
1.5 Non-mammalian Ears

The general morphology of the non-mammalian auditory periphery is similar to that of mammals in the sense that there are distinct outer, middle and inner ears. Similar to the human, all the non-mammalian species examined here (chicken, gecko and frog) have a tympanic middle ear. However, outer ear anatomy differs greatly. While the chickens have an external auditory meatus, the frog does not (their TyM is flush with the side of their head). Geckos have a shallow recess in which their TyM sits. We did not examine closely how outer ear properties (such as TyM thickness or effective area) compare across species.\(^{20}\)

For the middle ear, all three species have a similar ossicular structure. This comprises a single connecting bone called the *columella*, which couples the TyM to the stapedial footplate (in contrast to three bones in the human ear). One point of view is that a single ossicular system is not an efficient transmitter of higher frequency information and thus has led non-mammalian inner ears to evolve towards optimizing detection of lower frequency sounds [Manley, 1990]. All three species show some degree of direct acoustic coupling between both tympanic membranes via coupled middle ear airspaces (again, in contrast to humans). This is achieved via the inter-aural tube in the chicken, and by the oral cavity in the gecko and frog. This coupling is believed to provide a mechanism for increased binaural directional sensitivity [Fletcher and Thwaites, 1979]. Differences aside, the middle ear plays a very similar mechanical role in these non-mammalian species as it does in mammals. Specifically, it acts as an impedance transformer, transmitting sound pressure from a low impedance medium (air) to a high impedance one (the fluids of the inner ear). The role of the middle ear is also important to consider within the context of this study, as it is the means by which energy generated by the inner ear (associated with the OAE generation mechanisms) is subsequently transmitted outwards and appears at the outer ear as an emission.

The greatest diversity across these species is apparent upon examination of the

\(^{20}\)It is worth noting that a tympanic outer ear is not necessary for the detection of eOAEs [van Dijk et al., 2002].
inner ear anatomy. A schematic providing an overview of the inner ear anatomy in the different species examined in this study is shown in Fig. 3-1. Differences both in terms of anatomy and hearing perception across the non-mammalian species are described in detail in Ch.3.

1.6 Non-mammalian Otoacoustic Emissions

As mentioned previously, numerous studies have examined non-mammalian OAEs. Spontaneous emissions have been seen in most groups of vertebrates tested. These include birds [Taschenberger and Manley, 1997], lizards [Koppl and Manley, 1993] and frogs [van Dijk et al., 1996]. However, in a given class, there can be much diversity across species with regards to the presence of SOAEs. For example, SOAEs have not been observed in chicken ears while they have been reported to be readily apparent in owls. Furthermore, in a given species, ears from a limited percentage of all individuals may readily emit spontaneously while the majority of ears do not.

Reports of OAEs in non-mammals have typically focused upon evoked emissions. Various aspects of these studies are discussed in more detail in Ch.3-6. It is also worthwhile to note that OAEs have been examined in non-vertebrate species such as moths and grasshoppers. Non-vertebrate studies are described in more detail in Ch.4.

What is clear from these studies is that most types of ears are capable of emitting sound. So in spite of significant morphological differences across species, OAEs are a common feature and appear to be a relatively generic property of the ear [Koppl, 1995]. This observation motivates further study of emission properties to better understand how the underlying generation mechanisms are similar and different across species. Knowledge of how emission properties compare coupled with what we know about how the underlying anatomy/physiology differs across species will serve to reveal fundamental processes that are common to all ears.

---

21 This has been readily apparent over the course of study described in this thesis. Some individual geckos readily show consistent SOAE activity, though the majority of the animals did not exhibit any SOAEs at all.
1.7 Summary of Approach

Human anatomy and OAEs have been well characterized and serve here as a reference for our interpretation across species. To summarize, we are looking at four different species, only one of which (based upon present evidence) has hair cells exhibiting somatic motility (humans) and two which are thought to lack a traveling wave propagating along the length of the auditory organ via a flexible BM (geckos and frogs). Perceptual aspects of auditory function in the non-mammalian species examined in this study are summarized briefly in Ch.3

The primary approach of the current study is to examine OAE magnitude and phase behavior across a number of different species in a systematic way. For this study, we use a single measurement system and stimulus paradigms to avoid systematic errors which might confound interpretation. Many previous OAE studies have failed to report phase information, which as alluded to above, can provide additional insights into emission generation. We look across a variety of species focusing upon two different types of emissions: distortion products (DPOAEs) and stimulus frequency emissions (SFOAEs). As there are a large number of different stimulus conditions which can conceivably be tested, we will limit ourselves here to those which will address the following four groups of questions (while maintaining the comparative approach throughout):

- How do SFOAE phase gradient delays compare across species? What type of information do these phase gradients reveal?

- Does the mammalian OAE classification scheme apply to non-mammalian ears? Specifically, do non-mammalian ears exhibit evidence for multiple mechanisms of emission generation?

- How does emission frequency dependence vary with respect to stimulus intensity? In particular, how do eOAE phase gradients vary with level?

- At a fixed frequency, how do emission properties vary with respect to stimulus intensity? Is there evidence for multiple level-dependent OAE gen-
eration mechanisms in the ear (which could potentially explain non-monotonic growth)? Does dependence upon stimulus intensity differ significantly across species?

The subsequent chapters in the thesis further motivate each one of these four groups of questions and present data to address each.
Chapter 2

eOAE Methods

2.1 Measurement System

All measurements reported in this thesis were obtained using the same stimulus paradigms, acquisition codes, and ER-10C for all species. A desktop computer housed a 24-bit soundcard (Lynx TWO-A, Lynx Studio Technology), whose synchronous I/O was controlled using a custom data-acquisition system written in C. Experiments performed on chicken ears were done at University of Pennsylvania using a different computer, soundcard (same model) and isolation booth, but all other aspects were identical. A sample rate of 44.1 kHz was used to transduce signals to/from an Etymotic ER-10C (whose gain was set at +40 dB). The microphone signal was filtered using a high-pass filter (consisting simply of a resistor and capacitor) with a cut-off frequency of 410 Hz.

The probe earphones were calibrated in-situ using flat-spectrum, random-phase noise. These calibrations were repeatedly verified throughout the duration of the experiment for a given subject. Re-calibration was performed if the actual level presented differed by more than 3 dB from the specified value. The microphone calibration was tested using a pistonphone (which had a 94 dB SPL output at 1007 Hz) and found to conform well to the shipped specifications. The microphone frequency response was flat (within ±1-2 dB) across frequency range examined in this study.
2.2 Stimulus Paradigms and Analysis

For SFOAEs, we used a suppression paradigm similar to that previously employed by Shera and Guinan (1999). A schematic of this method is shown in Fig. 2-2. The first earphone (EP1) produces a signal consisting of a 464 ms sinusoidal buffer at the probe frequency $f_p$, ramped on/off over 12 ms at the ends of the buffer. The second earphone (EP2) also produces a 464 ms signal, but at the suppressor frequency $f_s$, which was ramped on only during the latter half of the buffer (the first half was silence). The microphone response was extracted from two 186 ms (8192 pts.) segments from the total buffer, one from the probe alone and one with the probe+suppressor. These segments were taken at least 20 ms after the end of the ramping-on periods to allow any transient behavior to decay. Thus, the measurements reported here are for the steady-state condition.

The Fourier transform of each segment was taken and the complex difference of the two Fourier coefficients at $f_p$ revealed the emission magnitude and phase. Unless

---

$^1$Phase shifts introduced with these delays were accounted for and described in further detail in Appendix D.
Step 1. Present Probe Alone (emission is present)

FFT reveals magnitude and phase AT Probe Freq.

- Since the system is non-linear, the presence of the suppressor tone will affect what is going on at the probe frequency (Step 2)

Step 2. Present both Probe & Suppressor tones (emission not present)

FFT reveals magnitude and phase AT Probe freq.

Subtraction of the phasors in the complex plane reveals both the magnitude and phase of the OAE in the suppression paradigm

- Subtraction of the phasors in the complex plane reveals both the magnitude and phase of the OAE in the suppression paradigm

Step 3. Subtract phasors

Figure 2-2: SFOAE suppression paradigm schematized.

otherwise noted, the suppressor level (\( L_s \)) was 15 dB above the probe level (\( L_p \)) and \( f_s \) was 40 Hz above \( f_p \). When measuring level growth functions, stimulus levels were randomized. For DPOAEs, each earphone produced a single frequency (\( EP1 - f_1 \), \( EP2 - f_2 \)) over a 244 ms buffer, each tone being ramped on/off at the ends. Similar to the SFOAEs, a 186 ms segment was extracted from the microphone response.

For both SFOAEs and DPOAEs, 35 buffers were averaged, excluding any flagged by an artifact-rejection paradigm (similar to that of Shera and Guinan, 1999)\(^2\). Furthermore, all stimulus frequencies were quantized such that an integral number of stimulus periods fit in the response segment. This means that the nominal values of quantities such as \( f_s - f_p \) and \( f_2/f_1 \) (which are specified to be constant for a given frequency sweep) could vary a small amount between successive steps in a sweep. These variations were always less than 2% (and usually much smaller). Frequency step-size during sweeps was small enough to avoid ambiguity during the phase unwrapping.

\(^2\)The artifact rejection was done in the time domain by subtracting two buffers and applying a threshold criterion to the largest difference. The presence of the high-pass filter for the microphone signal significantly increased the usefulness of the artifact rejection paradigm (noise was largest at frequencies below ~200 Hz).
The response from the microphone was Fourier transformed and depending upon the emission type, either the emission magnitude and phase (SFOAEs) or the whole spectrum (DPOAEs) saved to file. Further analysis was done using Matlab (phase correction, phase unwrapping, etc.). Unless otherwise noted, all error bars indicate the standard error of the averages for a given stimulus set (after standard propagation of error among the various measured quantities). The noise floor was found by averaging the ±3 frequency bins adjacent to the one of interest for both SFOAEs and DPOAEs. In the case where multiple emissions are plotted together, the noise floor curve shown is typical of what is seen for any given measurement.

2.2.1 I/O Delays

In addition to determining the necessary voltages to drive the earphones, the calibration characterized the delay associated with the D/A (signal from computer going to the earphones) and A/D (response signal coming from microphone to computer) converters. These delays, totaling ~2.5 ms, account for the electronic delay associated with I/O delays from the 24 bit A/D and D/A converters. The total system delay is described in more detail in Appendix D. It was important to quantify this delay as it must be accounted for in the emission phase.

For DPOAEs, the relative contribution (to the total phase shift) between the shifts associated with the D/A and A/D also needs to be measured. Basically, the relative electronic delays were measured by passing the signal across a diode (which acts as an instantaneous nonlinearity), providing a reference phase for any given distortion product that can be used to correct the DPOAE phase. Corrections to account for phase shifts associated with measurement system I/O delays are described in more detail in Appendix D. OAE phase results using this method were consistent with those previously reported [Knight and Kemp, 2000; Shera and Guinan, 1999].
2.2.2 System Distortion

Distortion associated with nonlinearities in the measurement system was present and introduced some degree of artifact. However, as described below, artifactual distortion was relatively small compared to the eOAEs and could be clearly distinguished from emissions. The physical basis for this distortion most likely derives from nonlinear characteristics of the ER-10C earphones. The probe contained two earphones, due to the fact that a significant amount of intermodulation distortion was present when a single earphone was driven with two sinusoids (harmonic distortion is described below). However, the output of the earphones was linear with respect to driving voltage.

System distortion could be measured in a number of different ways (i.e. probe coupling configurations). In the simplest case, artifactual distortion was measured using small cavities (~0.5 cc), with both rigid and non-rigid terminations. Emission measurements were also made in a profoundly deaf subject (sensorineural hearing loss with a unilateral cochlear implant; presumably lacking any emissions) to provide a measure with a comparable acoustic coupling to that for human subjects (where the volume is larger, ~3 cc).\(^3\)

For DPOAEs, distortion occurred at both harmonic and intermodulation frequencies. Quadratic and cubic harmonic distortions (with the probe coupled in various configurations) are shown in Fig. 2-3 and Fig. 2-4. Several observations are worth noting. First, the amplitude of the harmonic distortion depends strongly upon the earphone’s acoustic load impedance. Second, both \(f_1\) and \(f_2\) harmonics appear to be ‘filtered’ in a similar way (this is more apparent in the cubic case). Third, emission phase (which is corrected for I/O delays) does exhibit a complex frequency dependence (the phase of the primaries is constant and zero). The source of this frequency dependence is unknown. The bottom-line here is that the ER-10C is a poor choice

\(^3\)Though this subject had a cochlear implant in one ear, both ears were tested. SFOAE \((L_p = 40\ dB\ SPL)\) and DPOAE \((L_1 = L_2 = 65\ dB\ SPL, \ f_2/f_1 = 1.22)\) measures were found to be completely in the noise floor, the noise being comparable to that present in most human subjects. So no artifact was apparent at all in either ear for this subject.
for studying harmonic DPOAEs due to the nonlinearity of its earphones\textsuperscript{4}.

Some degree of intermodulation distortion artifact was present when using $L_1 = L_2 = 65$ dB SPL, although it was only apparent in $2f_1 - f_2$ when $f_{dp}$ was \textasciitilde 2-3 kHz and when the coupling volume was small\textsuperscript{5}. Evidence indicated that the artifactual distortion arises from electrical cross-talk between the two earphones leads in the probe housing. Intermodulation artifacts were typically masked by the noise floor.

SFOAE artifacts were also present and most likely derived from the same mechanisms responsible for DP intermodulation distortion. These artifacts could be observed as residuals of the suppression paradigm and were only apparent above the noise floor for high stimulus levels ($L_p > 60$ dB SPL).

As indicated above, the magnitude of artifactual distortion depended upon the actual acoustic coupling of the probe. However, when coupled to the ear of any given species, both DP intermodulation distortion artifacts and SFOAE artifacts were $\approx 70$-80 dB below the evoking stimuli levels and typically beneath the acoustic noise floor. Furthermore, artifacts lacked the steep downward trends in their phase curves (as a function of frequency) readily apparent in the phase response of emissions. Thus, it was straightforward to determine when artifacts were present and distinguish them from actual emissions. Based upon characterization of the measurement system's artifactual distortion, we believe that all SFOAE and DPOAE intermodulation data presented in Ch.3-6 derive solely from sources inside the inner ear.

\section{2.3 Subjects}

Experiments involving humans, geckos, and frogs were all performed at MIT. Human experiments were approved by both the MIT COUHES (Committee On the Use of Humans as Experimental Subjects) and Human-Studies Committee at the Mas-
Figure 2-3: Quadratic harmonic distortion (i.e. \( f_{dB} = 2f_1 \) or \( 2f_2 \) for the given distortion plotted) when probe is coupled in various configurations. Stimulus conditions are the same for each case. Dashed line indicates noise floor. \([L_1 = L_2 = 65\text{ dB SPL}, f_1 = 1\text{ kHz}, f_2/f_1 = 1.22]\)
Figure 2-4: Same as Fig. 2-3, except that the cubic harmonic distortions are plotted.
The experimental protocol involving both the geckos and frogs was approved by the MIT CAC (Committee on Animal Care). Experiments involving chickens were performed at the University of Pennsylvania and were approved by the IACUC (Institutional Animal Care and Use Committee).

For the lizards, 9 different leopard geckos (*Eublepharis macularius*) and 4 Tokay geckos (*Gekko gecko*)\(^6\) were used. Animals were obtained from at least two different vendors for a given species and ranged in size from 20-55 g (leopards) and 40-65 g (tokays). Based upon behavior and external appearance, these two species are very different, but have fairly similar auditory anatomical morphology [Wever, 1978]. Lizards were kept in aquarium tanks with a 12-hour light cycle and fed crickets dusted with vitamins twice a week. All were healthy and active. Prior to each experiment, an animal was dosed with 25 mg/kg Nembutal i.p. to reduce movement; this dose was effective for \(\approx\)4-5 hours. The animal recovered completely within a few hours and was subsequently used multiple times over the course of two years (always with at least one month recovery time between experiments). During the experiment, lizards were placed on a vibration isolation table in a noise-reduction chamber. Body temperature was monitored using a calibrated thermocouple placed in the mouth, propping it open. Temperature was kept in the range of \(\approx 81 - 86^\circ\) F, depending upon the placement depth of the thermocouple in the mouth, using a regulated heating blanket (Harvard Apparatus). Preliminary experiments indicated that emission phase gradients were not very sensitive to body temperature, though temperature dependence was not thoroughly explored in this study. The probe was coupled to the external ear using a short tube attached to the foam tip and sealed to the head using copious amounts of vaseline. The probe was held in place using a flexible holder. This ensured a tight (closed) acoustic coupling and minimized low-frequency losses.

For the frogs, 6 different leopard frogs (*Rana pipiens pipiens*) ranging in size of 40-80 g were used in this study. Except for one ear in one animal (which yielded no detectable emissions above the distortion noise floor), all ears showed similar results.

\(^6\)Tokays are unique in the lizard family in that they use vocalizations to express both aggressiveness and sexual interest. Two species of gecko were used because leopard gecko are significantly easier to work with/handle, while there is a wealth of study that has been published on tokay ears.
Frogs were kept in an aquarium that allowed them to be either in or out of water and were routinely fed vitamin-dusted crickets twice a week. For the experiments, frogs were given an i.p. injection of Nembutal at a dose of 45 mg/kg. This was sufficient to keep them sedated for 5-6 hours, after which they recovered completely (and were subsequently used for future experiments). The protocol was then identical to that for the lizard, except that the frogs were not placed on a heating pad. Rather, they were wrapped in wet paper towels to facilitate respiration. As with the gecko, a thermocouple was inserted into the mouth to monitor temperature and had the added effect of propping the mouth open. Body temperature was in the range of 66 – 69°F and remained relatively stable over the course of the experiment.

Nine different White Leghorn chickens (Gallus domesticus), aged P14-P28 (days post-hatching) and weighing 150-300 g, were used in the study. The anesthesia paradigm for the chickens was similar to that used in Ipakchi et al., 2003. Briefly, chickens were injected i.m. with urethane (ethyl carbamate) at a dose of 2.5 g/kg. Due to the anesthesia’s effect on the bird’s lungs, a tracheostomy was performed to assure free breathing. Feathers were removed around the external ear opening, but no further surgery was done as a clear path to the tympanic membrane was visible. Chicks were placed on a vibration-isolation table inside a sound-attenuation booth and body temperature was maintained at \( \approx 106°F \) via a regulated heating pad. Similar to the gecko and frog, the probe was coupled to the outer ear via vaseline and a small tube attached to the foam tip. In some cases, the beak was held open using a head-holder plate, but in others the beak was closed; no systematic differences in emissions between the two conditions were observed. In some experiments, data were obtained from both ears in one animal. Upon termination of the experiment, the chicken was euthanized. Due to respiration and cardiac-related artifacts, the chickens typically exhibited higher acoustic noise floors than all the other species (including humans).

Human subjects were seated comfortably in a quiet room and were awake for the duration of the experiment. Eight subjects (three males, aged 28, 31, and 64; five females, ages between 26 and 32) with normal hearing participated. Both ears were
tested in some of the subjects. The probe fit snugly into the ear canal using a foam coupling tip.

2.4 Data Visualization

Throughout this thesis, OAEs are characterized by their magnitudes (expressed in dB) and phases. To simplify the interpretation of emission phase in terms of a time delay, phase information is typically characterized by phase gradients, which are derivatives of the phase function with respect to emission frequency. There are many approaches that can be taken to quantify this value, particularly in light of the sometimes irregular nature of the phase curves\(^7\). We took the simplest approach with the goal of preserving potentially relevant information (a loss of information is a danger associated with smoothing paradigms). Thus, the phase slope was computed as a local derivative using Matlab's `gradient.m` function. It should be noted that the phase unwrapping (also done in Matlab using the built-in function `unwrap.m`) is an important step when computing phase gradients that requires careful consideration\(^8\). Our choice for the unwrapping method derives from approaches taken in previous analyses [Shera and Guinan, 1999] that have proven fruitful.

In some figures, trend lines are included to guide visualization\(^9\). These fits were computed using the `loess` technique [Cleveland, 1993]\(^10\). Briefly, loess is a locally weighted polynomial regression fit. We use an unweighted (with regards to the y-axis), first-order polynomial to fit. At any given x-value\(^11\), typically 20 – 40% of the total number of points (centered about the point the fit is being computed at)

---

\(^7\)See Schairer et al. (2006) for further discussion.

\(^8\)Unwrapping is necessary due to the ambiguity associated with traversing around a circle. It is imperative that adjacent measurements be finely spaced enough that the unwrapping algorithm can keep track of how the phasor (representing the emission magnitude and phase in the complex plane) is rotating. Problems arise when there is a sudden phase jump. For example, the unwrapping algorithm needs to decide if a sudden jump of \(1/3\) cycle is indeed a shift \(1/3\) ahead or \(2/3\) backwards. These jumps are apparent in the phase data and tend to have little effect on overall trends.

\(^9\)All plotting was ultimately done using Matlab in addition to Adobe Illustrator.

\(^10\)Actual matlab code used was 'exquisitely crafted' by C. Shera. See the 'Downloads' tab at http://web.mit.edu/aps/

\(^11\)To avoid confusion, x-values refer to the independent variable (such as frequency) and the y-values the dependent variables (such as eOAE magnitude).
are used. When multiple data points occur at the same x-value, a small amount of
gaussian noise was used to offset data points slightly along the x-axis when computing
the fits (helps make plots more readable).

Lastly, we used a threshold criterion in some figures (such as Fig. 3-5 and Fig.4-5)
in order to avoid inclusion of points close to the noise floor. Only emissions whose
magnitudes were at least 10 dB above the noise floor are included in those plots.
Given that the typical emission magnitude was relatively large, few points rarely
failed to meet this criteria. Thus, the overall trends were not strongly affected by
including or excluding points close to the noise floor.
Chapter 3

Comparison of SFOAEs Across Species

3.1 Basis for a Comparative Study of OAEs

Otoacoustic emissions (OAEs) are sounds generated and subsequently emitted by the inner ear. While it is generally thought that OAEs do not directly serve a purpose in hearing perception, they have nonetheless been developed as a useful clinical assay for the diagnosis of hearing impairment. Healthy ears emit, but impaired ones have reduced emissions or lack them altogether\(^1\). Better understanding the underlying mechanisms that give rise to OAEs can further their clinical potential. Much of our knowledge about OAEs has stemmed from mammalian observations, where specific morphological features have factored significantly into their interpretation. Our goal here is to examine OAEs both in humans and some representative non-mammalian species that have significant relative differences in their auditory anatomy and physiology. By characterizing emission properties in ears that lack features thought to play an important role in mammalian OAE production, we stand to gain further insight.

\(^1\)By 'healthy', we are referring to normal air conduction thresholds. In some cases, it is however possible to have a healthy ear that lacks measurable spontaneous and stimulus-frequency emissions. Additionally, some impaired ears readily show normal emission magnitudes (hearing loss in these individuals is likely neural in origin). These two exceptions tend to be rare, an observation that makes OAEs still be of significant clinical use.
into the underlying emission generation mechanisms.

OAEs are generally regarded as a by-product of an underlying amplification mechanism presumed to be present in the ear. The so-called *cochlear amplifier* has been proposed as the basis for the sensitivity and large dynamic range of hearing. In mammals, the foundation for an amplification mechanism is commonly thought to originate with the outer hair cells (OHCs), which exhibit a mechanical response to sound stimulation that can generate appreciable forces and/or changes in OHC stiffness [Dallos, 1996]. Non-mammals do not have this feature in their hair cells [He et al. 2003; Koppl et al. 2004] and lack numerous other inner ear features specific to mammals, such as basilar-membrane traveling waves. Considering the physiological differences between mammals and non-mammals, further study of OAEs will provide insight into the notion of an amplifier present in the inner ear.

We examine four different groups here. First, we use humans as the standard for mammals. Second, we look at the chicken ear. Many studies have examined the chicken auditory periphery, a chief motivation stemming from its unique ability to regenerate itself and return hearing function to near-normal even after significant damage. Third, we investigate geckos, a lizard species that uses vocalizations for a wide array of purposes such as mating and signifying aggression. Lastly, we examine the frog ear. Frogs are unique in that they have two anatomically distinct regions in their ears sensitive to acoustic stimuli (excluding the sacculus). An overview of the auditory anatomy of these species and their hearing acuity (i.e. thresholds, sharpness of tuning) is given in Sec.3.2.

Previous studies have examined OAEs in non-mammals, characterizing various properties of their emissions and discussing them relative to mammals [see Ch.1, Manley et al. 2007]. Our intention here is to make a systematic comparison across a wide array of species using the *same measurement system and stimulus paradigms*, so to minimize methodological differences that could cloud interpretation. This is par-

---

2 We look at two species of geckos in this study. The two species, the leopard and toka gecko, are highly different in both overall appearance and behavior. However, the anatomy of their outer, middle and inner ears is similar. The reason for using the two species here stems from the lack of ANF data in one, which will be important for later discussion.
particularly advantageous with regard to emission phase and the associated frequency
gradient, a useful measure indicative of underlying delays in inner-ear function. This
chapter focuses solely on stimulus-frequency emissions (SFOAEs)\(^3\). We start by pro-
viding an overview of how the anatomy and physiology differ among the species
examined here.

### 3.2 Overview of Anatomical Differences

This study examines SFOAEs in four different groups: human (*Homo sapien sapiens*),
chicken (*Gallus gallus domesticus*), gecko (two species: *Eublepharis macularius* and
*Gekko gecko*), and frog (*Rana pipiens pipiens*)\(^4\). While all the species examined have
a tympanic middle ear, outer ear anatomies differ greatly. While the chickens have
an external auditory meatus, the frog does not (their tympanic membrane, or TyM, is
flush with the side of their head). Geckos have a very shallow meatus in which
their TyM is recessed. We did not examine closely how middle ear properties (such
as TyM thickness or effective area) compare across species\(^5\).

The the non-mammalian species all have a similar ossicular structure. Their
middle ear is comprised of a single connecting bone called the *columella* that couples
the TyM to the stapedial footplate, in contrast to three bones spanning the human
middle ear. One point of view is that a single ossicular system is not an efficient
transmitter of higher frequency information and thus has led non-mammalian inner
ears to evolve towards optimizing detection of lower frequency sounds [Manley, 1990].
All three species show some degree of direct acoustic coupling between their bilateral
tympanic membranes via the middle-ear airspace (in contrast to humans). Bilateral
coupling is achieved in the chicken via the long interaural tube, and by the oral
cavity in the gecko and frog. This coupling is believed to provide a mechanism for

---

\(^3\)DPOAEs are subsequently discussed in Ch.4.

\(^4\)The common name for *Eublepharis macularius* is leopard gecko while the common name for
*Rana pipiens pipiens* is leopard frog. To avoid confusion in this regard, we simply refer to the
leopard frog as ‘frog’.

\(^5\)It is worth noting that the presence of a tympanic membrane is not necessary for the detection
of DPOAEs [van Dijk et al., 2002].
increased binaural directional sensitivity [Fletcher and Thwaites, 1979]. Differences aside, the middle ears of both mammals and non-mammals play a similar mechanical role. Specifically, the middle ear acts as an impedance transformer, transmitting sound pressure from a low-impedance medium (air) to a high-impedance one (the fluids of the inner ear). It does this by creating a mechanical pressure gain via a complex levering system and area ratio. The middle ear is also the means by which energy generated by the inner ear (associated with the OAE generation mechanisms) is subsequently transmitted outwards and appears at the outer ear as an emission.

The greatest amount of diversity across the species in question is apparent upon examination of the inner-ear anatomy (Fig. 3-1). Analogous to the organ of Corti in the mammalian cochlea, bird and lizard hair cells are situated in a structure called the basilar papilla. The frog has two morphologically distinct auditory papilla in the inner ear. Specifics of the inner ear structures of each species are summarized below. For brevity, we do not discuss human cochlear anatomy but do provide a brief description of objective measures of their hearing perception. We reserve the term cochlea for discussing the inner ear of mammals.

### 3.2.1 Chickens

Chickens have a short BM (≈5 mm) that curves gently by 90° over its length. The BM width and thickness change along its length (as well as hair-cell bundle properties such as height and number of stereovilli), correlating to the tonotopic gradient observed from ANF responses [Manley et al., 1987; Chen, Salvi and Shero, 1994]. Evidence from avian species suggests that a longitudinally traveling wave is present along the BM [Bekesy, 1960; Gummer et al., 1987]. There are ≈5000 hair cells situated in a hexagonal fashion [Tilney and Saunders, 1983]. In general, two distinct types of hair cells have been characterized: short hair cells (SHC) sitting directly atop the BM (receiving the bulk of the efferent innervation) and tall hair cells (THC, with the bulk of the afferent innervation) [Tanaka and Smith, 1978]. Chick hair cells do

---

An overview for human auditory anatomy is given in Ch.1. For a more detailed review, see Dallos et al. 1996]
Figure 3-1: Simplified schematic of different species’ inner ear anatomy. Two perspectives are provided for each species: a cross-sectional view (top) and a top-down view of the sensory epithelium (bottom). Except for the frog, the arrows in the top-down view represent an individual hair cell (HC), the direction indicating the bundle’s polarization (pointing from shortest to tallest row). Dashed lines in human and chicken indicate separation between laying over bone and flexible BM. For the frog, the entire longitudinal length of only the amphibian papilla (AP) is shown and arrows indicate gross trends of the HCs (the finely dashed bounding box corresponds to where the cross-section would lay and the coarsely dashed line represents where the sensing membrane extends down from the roof of the recess). HCs shown to exhibit cell body somatic motility are indicated by a star on their tail. White regions are fluid-filled, grey regions correspond to overlying tectorium (with grey lines indicating fibrillar structure), grey striped area represents bone and stippled areas are non-HC cellular regions (e.g. supporting cells). Distinction between scala vestibuli and scala media is omitted. Legend is as follows: AP- amphibian papilla, AR- amphibian recess, BM- basilar membrane, BP- basilar papilla, FN- fundus, LL- limbic lip, SA- sallet, SC- sallet chain, SE- sensing membrane, SM- scala media, ST- scala tympani, TC- tunnel of Corti, TM- tectorial membrane.
not exhibit somatic motility [He et al. 2003; Koppl et al., 2004]. The overlying TM is relatively quite thick, with dense radial and longitudinal fibers apparent under a light microscope. Cavities that are present in the TM over each hair cell extend back towards the homogene cells (at the neural edge), making the TM appear porous through a given cross-section [Cotanche, 1987]. For all hair cells in the papilla, the tallest row of the stereociliary bundle is tightly coupled to the TM, which is also attached to the papillar surface via fibrillar connections coupling to the micovilli of the supporting cells [Tanaka and Smith, 1975]. Unlike the mammalian ear, the avian ear is capable of regenerating itself. Upon destruction of significant portions of the papilla and hair cells, the bird ear repairs itself so that thresholds and sharpness of tuning return to values near those found in the undamaged ear [Cotanche, 1999; Smolders, 1999].

Based upon ANF recordings from previous studies, P21 chicks have a flat mean threshold of ≈ 20 dB SPL from 200-3000 Hz, increasing sharply at higher frequencies [Manley, 1990]. Psychoacoustic studies in adult chickens correlate well to these measurements [Saunders and Salvi, 1993]. $Q_{10}$ values from the single units are typically around 2-5 (though some units exhibit significantly higher $Q$ values), and increase with larger characteristic frequency [Salvi et al., 1992]. While DPOAEs have been measured in the chicken [Kettembeil et al. 1995; Ipakchi et al. 2005; Lichtenhan et al. 2005], the authors know of no published reports of SOAEs in *Gallus gallus domesticus*.

### 3.2.2 Geckos

Geckos have a short and straight BM that is 1.2-1.8 mm long in the two different species examined here⁸. The BM width and thickness vary considerably over its length and contains ≈1000-2000 hair cells [Wever, 1978]. Hair cell bundles are oriented both

---

⁷A physiologically measured $Q_{10}$ value is different from the $Q$ as typically defined by engineers (where it is defined as being proportional to the ratio of energy stored to power lost over one cycle of oscillation).

⁸The main anatomical differences between the leopard and tokay geckos is that the tokays have a slightly longer papilla that contains ≈ 40% more hair cells.
uni-directionally (all in the same direction) and bi-directionally (180° relative to one another) as shown in Fig. 3-1. There is a unique tectorial topology along one region of the papilla that consists of sallets, discretized sections of TM loosely coupled to each other via a fine strand overlying their top surface called the salllet chain. The sallets couple a single row of bi-directionally oriented hair cells together as shown in Fig. 3-1 [Wever, 1978]. There is evidence suggesting a lack of both somatic motility [Koppl et al., 2004] and traveling waves along the gecko BM [Manley et al., 1999]. A thickened tissue called the fundus (or ‘papillary bar’) runs along the length of the BM underneath the BP. Tonotopy in the lizards is believed to extend from graded changes in the hair cell bundle properties [Aranyosi and Freeman, 2004] and the associated attached TM [Authier and Manley, 1995]. Both afferent and efferent innervation are present, though the latter appears exclusive to the uni-directional segment of the papilla [Manley, 1990].

Previous studies have looked at cochlear microphonic (CM) in both species [Wever, 1978] and auditory nerve fiber (ANF) responses in the tokays [Eatock and Manley, 1981; Sams-Dodd and Capranica, 1994; Manley et al., 1999], giving an indication of their thresholds and sharpness of tuning. Based upon the CM data (which match up well to ANF measurements, in spite of the bi-directional nature of the HC orientation along a significant portion of the papilla that would presumably cause a cancellation effect and thus higher CM thresholds), the tokay ear has a threshold of ≈ 10-15 dB SPL in its most sensitive region of 500-800 Hz, increasing sharply at ≈ 12-15 dB/octave at lower and higher frequencies. Leopard geckos appear to be a further 10-15 dB more sensitive than the tokays based upon CM measures, meaning thresholds at or below 0 dB SPL (ANF studies have not been performed in leopard geckos). Derived $Q_{10}$ values from the ANF studies for the tokay were ~2-4, increasing with characteristic frequency. Spontaneous emissions have been reported.

---

9 This study showed that the tonotopic gradient along the gecko BP is opposite that of the BM/BP morphological gradients. This suggests that tuning in the lizard originates primarily at the micro-mechanical level, a manner different from mammals.

10 One needs to use caution when using cochlear microphonic measurements to ascertain ‘threshold’ due to the subjective nature of choosing a criterion value. Wever (1978) chose a 1 μV criterion which appears to correlate well to ANF-derived thresholds, at least at lower frequencies [Eatock and Manley, 1981].
in both *Eublepharis macularius* and *Gekko gecko* [Manley et al., 1996], though we know of no previous reports of evoked emissions in either species.

### 3.2.3 Frogs

Frogs have two papillae that are sensitive to sound, the amphibian papilla (AP) and the basilar papilla (BP) [Wever, 1985]. In contrast to chicks and geckos, both papillae in frogs lack a flexible BM altogether, and the hair cells sit atop relatively rigid tissue [Wever, 1973]. Unlike those of the human and chicken, the hair cells in the papilla do not show any obvious morphological distinctions, but do exhibit a degree of bidirectionality (similar to that seen in the gecko). Shaped roughly like a horseshoe and ~0.5-0.6 mm long, the AP is tonotopically organized (Lewis et al., 1982) and is sensitive to frequencies below ~1.2 kHz\(^\text{11}\). Containing \(\approx 800\) hair cells, the AP has a thick TM punctuated by many small holes (or ‘canals’) [Wever, 1973]. The hair cells couple tightly to the TM. A *tectorial curtain* (or ‘sensing membrane’) present in the AP recess extends from the bony roof of the recess down to attach the central portion of the TM. There does not appear to be a smooth gradation in either bundle or TM properties along the length of the AP [Shofner and Feng, 1983; Lewis and Leverenz, 1983]. There is still debate about what the exact mechanical mechanism for tonotopy is in the AP, but an electrical resonance (stemming from ionic flow across the membrane) may play a role at lower frequencies up to about 300 Hz [Lewis and Hudspeth, 1983; Pitchford and Ashmore, 1987]. The BP, sensitive to higher frequencies (above ~1.3 kHz) is smaller, containing only about 70 hair cells. The BP is thought to act as a singly tuned resonator (Ronken, 1991). Unlike the AP, it does not receive any efferent innervation. Similar to lizards, a great deal of diversity is seen in the inner ear anatomy across different species of frogs (the TM in particular and how it couples to the bundles).

CM measurements in species of the same family of frog indicate airborne thresholds near ~20-40 dB SPL, being smallest in the range 200-600 Hz [Wever, 1985]. ANF

\(^{11}\)The leopard frog AP length is estimated from drawings presented in Wever (1973). This estimate of 0.5-0.6 mm is similar to that of the bullfrog AP, as shown by Lewis et al. (1982).
responses in leopard frogs revealed higher mean thresholds, typically 50 dB around 500-1000 Hz and increasing at both lower and higher frequencies [Ronken, 1991]. $Q_{10}$ values range between 1-2, increasing at frequencies below 500 Hz and above 2000 Hz [Ronken, 1991]. The existence of SOAEs has been reported for *Rana pipiens pipiens* [van Dijk et al., 1996] while both DPOAEs [Meenderink and van Dijk, 2004] and SFOAEs [Meenderink and Narins, 2006] have also been reported.

### 3.2.4 Humans

Typical thresholds in a healthy human ear are relatively flat between 500 and 7000 Hz, being in the range of -5 to 15 dB SPL. Peak sensitivity occurs in the frequency range of 3-4 kHz. Since measures of tuning are typically obtained via invasive methods (such as the ANF studies described above for the non-mammals), these types of approaches are generally not applicable in humans, and non-invasive means such as psychophysical masking experiments need to be utilized. There is much current debate about which psychoacoustic method (such as using forward versus simultaneous masking paradigms) is the most valid indicator of human tuning sharpness [Glasberg and Morre, 2000; Shera et al., 2002; Ruggero and Temchin, 2005]. Thus, $Q_{10}$ values for human ears remains an open question that is addressed in Sec.3.5.5.

### 3.3 Methods

#### 3.3.1 SFOAEs

In this chapter, we focus specifically on a comparison across species of SFOAEs, emissions that arise in response to the presentation of a single tone. SFOAEs occur at the stimulus frequency, but at levels typically much lower than the stimulus, making their measurement difficult\(^{12}\). Here, we use a fixed set of stimulus conditions across all species in order to facilitate comparison (described below).

---

\(^{12}\)Various methods have been employed to measure SFOAEs. As described in Kalluri and Shera (2007), these methods all more or less produce the same results (at least when using lower level stimuli). The approach we use is a suppression paradigm.
We have three principle motivations for specifically examining SFOAEs here. First, SFOAEs ideally should be the easiest types of emissions to interpret given that they occur at the stimulus frequency (in contrast to other emission types). Second, there is a relative dearth of SFOAE measurements in the literature, both for mammals and non-mammals. As indicated above, this is likely due to the difficulty associated with their measurement. Several studies have examined SFOAEs in various non-mammalian species [birds: Manley et al. (1987), lizards: Rosowski et al. (1984b), frogs: Meenderink and Narins (2006)], but our goal here is to do so across species in a systematic way. Third, a physical theory explaining the origin of mammalian SFOAEs has been proposed [Zweig and Shera, 1995] that has shown consistent agreement with experiment and served as a foundation for current thoughts on OAE generation in mammals. The theory is based upon certain morphological features of the mammalian cochlea and describes a coherent reflection of energy from a forward traveling wave along the BM. A key feature in the model is a tuned BM that supports traveling waves, a feature absent in some of the non-mammalian species examined here. It will be of interest to examine how SFOAEs compare across species and relate emission properties back to the coherent reflection theory in light of the underlying differences in anatomy and physiology.

SFOAE data presented in this chapter comes from the following number of ears and individuals for a given species: Human - 8 ears from 7 individuals, Chicken - 12 ears from 8 individuals, Leopard gecko - 10 ears from 6 individuals, Tokay gecko - 8 ears from 5 individuals, and Frog - 8 ears from 5 individuals. Variability in a given individual is addressed in the Results section. Trend lines in the figures were computed using the loess method (a weighted polynomial regression fit), as described in Ch.2.

3.3.2 Choice of Stimulus Parameters and Methods

In order to simplify comparison across species, we measure SFOAEs in response to a limited set of stimulus parameters. There are numerous possible variations that could be made, such as using a different probe level \( L_p \), suppressor level \( L_s \) relative to \( L_p \),
suppressor frequency \((f_s)\) relative to the probe frequency \((f_p)\), and so on. However, by choosing stimulus conditions that are applied uniformly across species, we eliminate possible sources of confusion that might arise\(^\text{13}\). \(L_p = 40\) dB SPL was chosen because it is large enough to evoke a clear emission above the noise floor in all species. Ideally, we want to use as low a level stimulus as possible, since evoked emissions are known to grow fairly linearly at lower levels, but exhibit nonlinear behavior at higher intensities (typically above about 40-50 dB SPL for SFOAE)\(^\text{14}\). A suppression paradigm using \(f_s = f_p + 40\) Hz \(L_s = L_p + 15\) dB is been a common paradigm used for SFOAE detection [Brass and Kemp, 1993; Shera and Guinan, 1999]. The effects of varying these stimulus properties are explored in Ch.5 and 6.

### 3.3.3 Quantifying Phase Gradients

The change in emission phase with respect to frequency can reveal delays associated with SFOAE generation. These delays are quantified by computing the slope of the phase-frequency function, which we refer to as the phase gradient. The derivative of the phase function is computed in Matlab with the gradient function. Phase gradients are also plotted in dimensionless form by computing the \(N\)-value, defined as

\[
N_{OAE} = -\frac{1}{2\pi} \frac{\partial \phi_{OAE}}{\partial f_{OAE}} \cdot f_{OAE}
\]  

(3.1)

where \(\phi_{OAE}\) is the emission phase (in radians) and \(f_{OAE}\) is the emission frequency. \(N_{OAE}\) is the delay expressed in stimulus periods.

\(^{13}\)Emissions in the frog were masked by the acoustic noise floor above 3 kHz. So we did not measure for stimulus frequencies above 3 kHz. In an attempt to compensate however, we did extend to lower end of the range down from 500 Hz to 300 Hz. Thus, there is a slight discrepancy in frequency ranges across species.

\(^{14}\)Even lower-level stimuli can exhibit nonlinear growth (see the human figure in Fig. 6-3, where growth starts becoming nonlinear around \(L_p = 25\) dB SPL). We specifically chose \(L_p = 40\) dB SPL because that stimulus intensity was necessary to consistently evoke an SFOAE measurable above the noise floor in the chicken, the species with the smallest emissions.
3.4 SFOAE Results

Results From Individual Ears

Individual SFOAE sweeps from a single representative ear for each species are shown in Fig. 3-2. These curves were obtained using a 40 dB SPL probe tone. Save for a single ear in one of the frogs, this stimulus intensity was sufficient to evoke SFOAEs measurable above the acoustic noise floor in all species.

A number of features are readily apparent in Fig. 3-2. All species exhibited emissions whose magnitude varies with frequency. SFOAE magnitudes in the non-mammals got smaller as the emission frequency increased. Deep notches are apparent at certain frequencies in both the chicken and gecko, similar to those seen for humans. However, these notches are more numerous in the human\(^\text{15}\). Also, noting the different scales for the phase across species, all species exhibit significant phase accumulation with respect to changes in stimulus frequency. Computing the phase gradient, delays are revealed that are 1 ms or longer in all species with humans exhibiting the largest delays (Fig. 3-4 and 3-5). In all species, the delays decreased with frequency.

SFOAE properties (such as notch locations and phase gradients) were found to be highly reproducible in an individual human across sessions, whereas there was some variation in gecko. Variations in a given individual gecko were sometimes large enough to be comparable to that of the total variation seen across all individuals (as shown in Fig. 3-6). The variability in an individual gecko may have arisen due to seasonal fluctuations (which are known to effect their breeding cycles) or small changes in temperature. Further study is required to better understand the source of this variability. Session-to-session variability was not examined in chickens or frogs.

Results From All Ears

Fig. 3-3 shows the SFOAE magnitudes compiled from all the individual ears tested for all species. Each point at any given frequency for a particular species represents

\(^{15}\)These types of frequency notches are readily observed in human DPOAEs and have been called 'macro-structure'. They are typically thought to arise due to cancellation effects stemming from interference from multiple emission sources.
Figure 3-2: Comparison of SFOAE sweeps for a representative individual ear for four different species. **NOTE different scales across species.** The increased frequency resolution in the human was necessary phase unwrapping. Noise floors are shown by the dashed lines and are obtained by averaging the adjacent ±3 frequency bins. Note the different frequency scale for the frog. \([L_p = 40\ dB\ SPL, L_s = L_p + 15\ dB, f_s = f_p + 40\ Hz]\)
Figure 3-3: Comparison of SFOAE magnitude across species for $L_p = 40$ dB SPL. For any given frequency, each data point come from an individual ear (i.e. no repeats). Only points whose magnitude was at least 10 dB above the noise floor were included. Trend lines are included to guide visualization (see Methods).

Emissions are smallest in the chicken ear and typically larger in the geckos. However, emission magnitudes in the frog ear are also relatively large for $f_p$ below $\approx 1.2$ kHz, above which they fell off rapidly. Human emission magnitudes were intermediate, except at the highest frequencies where human emissions did not decrease to the noise floor as rapidly as the geckos. On average, emissions were largest at lower frequencies (around $\sim 1$ kHz) and decreased as stimulus frequency increased.

SFOAE phase curves are shown in Fig. 3-4 for all species. Note that over the frequency range tested, humans show significantly more phase accumulation relative to the other species. By computing the slope of the phase curves, emission delays are shown in the right side of Fig. 3-4. At all emission frequencies tested, delays are largest in the human, decreasing with frequency from $\sim 12$ ms at 500 Hz to 4 ms around 5 kHz. The frog shows delays of about 6 ms at the lowest frequency tested (300 Hz), decreasing to $\sim 2$ ms around 1.3 kHz, above which the delay is relatively constant with frequency. The delays in chicken and gecko are similar, ranging from
Figure 3-4: SFOAE phase curves are shown on the left and the associated delays (slope of the phase curve) on the right for $L_p = 40$ dB SPL. Both gecko species were grouped together in these plots. See caption for Fig. 3-2 for additional information.

1-2 ms and are slightly larger for the chicken. Delays in the chicken in gecko are also frequency-dependent, but do not increase rapidly with decreasing frequency as seen in the human and frog. Note that in all species, delays are greater than $\sim 1$ ms at all frequencies.

The SFOAE phase gradients are shown in dimensionless form in Fig. 3-5 as $N_{SFOAE}$ (the delay expressed in stimulus periods). Note the logarithmic scale on the y-axis. Only points where the emission magnitude was at least 10 dB above the noise floor are included (see Methods for further discussion). In all species except the frog, $N_{SFOAE}$ values systematically increase with frequency. The rate of increase with respect to frequency is larger at lower frequencies and slowed above $\sim 1$-2 kHz. The overall frequency dependence is similar across species.

In the case of the frog, there are effectively two frequency regions apparent in the $N_{SFOAE}$ trend, the transition between the two occurring at about the frequency separation of the two papillae ($\sim 1.2$ kHz). The rate of increase in $N_{SFOAE}$ with respect to frequency is different between the two regions.

Statistical results for Figs. 3-3 and 3-5 are shown in Fig. 3-6. Plots were made by pooling all the data together into octave-wide bins, from which the mean value and the 95% confidence interval were computed. Statistical values such as these are useful for comparison, but need to be viewed with caution since frequency-dependent
details are averaged out and may confound interpretation. Specifically, the notches apparent at certain frequencies shown in Figs.3-2 vary between individuals and will contribute to the mean values and their uncertainty.

Fig. 3-6 indicates that $N_{SFoAE}$ values in the chicken ear are significantly larger than gecko's in the range of 1-3 kHz. Around 1-2 kHz, $N_{SFoAE}$ values are similar in both the chicken and frog. Of the two gecko species, emission magnitudes are similar except at lower frequencies (below 2 kHz), where they are larger in the leopard gecko.

Expanding upon the emission magnitude frequency dependence, two different scales are apparent. First, there is a more global scale. To a rough approximation, the general shape of the SFOAE frequency dependence closely resembles that of the auditory threshold curve (as measured from ANF studies) for the non-mammals. For example, thresholds in the frog start to increase at $\sim 10$ dB/octave for frequencies above 1 kHz with few (if any) ANFs having a CF above 2.5 kHz [Ronken, 1991]. Similarly in terms of SFOAEs, emission magnitudes decrease towards the noise floor at
~ 20 dB/octave above 1 kHz with emissions no longer being detectable for emission frequencies above 2-2.5 kHz. Second, there are small variations with frequency in all species (i.e. the magnitude curves are not smooth). These variations are not due to uncertainty associated with the measurement, as the error bars for a given individual tended to be small with respect to changes in frequency. Additionally, deep notches are observed at certain frequencies in the human, chicken and gecko (as shown in Fig. 3-2). For a given species, the frequencies at which these notches occur varies significantly across individuals.

3.5 Discussion

We have demonstrated numerous similarities and differences in SFOAEs across various classes of vertebrates. Here, we summarize briefly the chief results found in this study. First, all species exhibit SFOAEs that are clearly measurable above the noise floor using a moderate-intensity stimulus (40 dB SPL). The magnitudes of the emissions varied across species, being largest in the gecko and smallest in the chicken. Second, all species exhibit frequency-dependent delays (as measured via the phase gradients) on the order of 1 ms or more. The mean value of these delays differs across species, being significantly larger in humans.
The first part of the discussion compares our results to those obtained in previous studies. Second, we focus on comparing emission results back to the underlying morphological differences across species. Third, in light of the framework proposed to explain the physical basis for SFOAE generation in mammals, we explore how generation mechanisms may differ across species. Lastly, we address how emission phase gradients may reveal information about time delays occurring in the inner ear and how delays may be correlated to frequency selectivity.

3.5.1 Comparison With Previous Studies

Our results are consistent with previous reports of human SFOAEs [Shera and Guinan, 1999; Schairer et al., 2006]. To the best of our knowledge, comparable measures of SFOAEs have not been reported in either the bird or lizard ear. Of the previous reports mentioned in Sec.3.3.1 examining SFOAEs in non-mammals, it is difficult to correlate the measurements made in those reports back to ours in a clear and direct fashion due to differences in measurement paradigms and the species examined.

SFOAEs have been measured in the same species of frog as the one examined here [Meenderink and Narins, 2006]; overall, our results are similar in spite of slight differences in the measurement paradigms. However, there are some differences between the previous frog research and ours that are worth pointing out. First, for similar stimulus conditions (i.e. \( L_p = 40 \) dB SPL) the emission magnitudes we measured in frog are slightly smaller (by \( \approx 5-10 \) dB). Second, we do not observe the bimodal frequency distribution reported by Meenderink and Narins, who observed a higher frequency region with a peak around 1.5-1.7 kHz that they attribute primarily to the BP. The discrepancy between the two studies may be a result of the differences in stimulus levels used (Meenderink and Narins primarily used stimulus levels in the range of 60-80 dB SPL) and growth rate differences between the AP and BP. Specifically, our stimulus intensity may have been too low to evoke a significant emission above the noise floor at the higher frequencies. Lastly, the phase gradients by Meenderink and Narins are slightly smaller than those measured here for emission frequencies less than 1 kHz by up to a factor of two. This discrepancy may be due to the difference
in stimulus intensities used in the two studies ($L_p = 40$ dB SPL for our results here while Meenderink and Narins used $L_p = 62$ dB SPL).

3.5.2 Relating Emission Properties Back to Anatomical Differences

It is curious that lower frequency emissions are largest in gecko and frog, the two species with the fewest number of total hair cells present in the inner ear. It is not clear if these larger magnitudes are the result of stronger emissions generators, some sort of cancellation effect that might arise due to a tuned BM (that supports a traveling wave), or if there is a higher degree of nonlinearity present in the inner ear of geckos and frogs (considering that SFOAEs were measured using a nonlinear suppression paradigm). As shown in Ch.4, both the gecko and frog also exhibit significantly higher intermodulation DPOAE magnitudes than either human or chicken. Another possibility is that the gecko and frog middle ears may be more efficient in terms of coupling energy from the inner ear back out to the TyM (reverse propagation).

Differences in the upper frequency limit at which emissions occur in humans and non-mammals likely stems from differences in the frequency range of hearing. ANF studies in the non-mammalian species examined here typically do not exhibit characteristic frequencies (CFs) above 5 kHz. So the machinery responsive to higher frequencies appears to be lacking in the inner ears of these non-mammals$^{16}$. The smaller upper limit for ANF CFs in non-mammals likely stems from the single ossicle of the non-mammalian middle ear, which appears less efficient at transmitting higher frequency stimuli to the inner ear [Wever, 1978].

It is also possible to compare SFOAE properties measured here to those of other mammals. Shera and Guinan (2003) examined SFOAEs in cats and guinea pigs (in addition to humans) using similar paradigms and identical stimulus conditions as those used in this study (i.e. $L_p = 40$ dB SPL). They found that phase gradients

$^{16}$Recent results in certain species of frog indicate that they appear sensitive to ultrasonic frequencies [Feng et al., 2006]. This high frequency sensitivity may be the result of a highly specific evolutionary adaptation.
Lp=40 dB SPL

Figure 3-7: Comparison of SFOAE delays in six different species. Cat and guinea pig data is from Shera and Guinan, 2003 (other data is taken from the right side of Fig. 3-4). \([L_p = 40 \text{ dB SPL}, L_s = L_p + 15 \text{ dB}]\)
in cat and guinea pig were similar to each other and significantly smaller than in humans. In fact as shown in Fig. 3-7, the delays in cat and guinea pig were similar to those observed in the non-mammals here, being no larger than a factor of ~1.5 greater than those of the chicken and smaller than the frog for emission frequencies below 1 kHz. Thus, mammalian-specific cochlear morphology (such as BM traveling waves) cannot by itself account for the differences in delay observed here between humans and non-mammals.

One of the obvious morphological differences across species is the size of the relevant auditory structures. The sensory epithelium of the human inner ear is much longer (~35 mm) than that of the non-mammals [chicken ~5 mm, gecko ~1.2-1.8 mm and frog (AP) ~0.5-0.6 mm]. Is it possible that the difference in the length of the sensory epithelium can account the differences in delays, perhaps stemming from longer travel times? Note that the SFOAE delays are close in the chicken and gecko, in spite of the differences in their lengths. Furthermore, delays at lower frequencies (below ~1 kHz) are larger in frog than chicken or gecko, in spite of the frog AP having the smallest length. Lastly, as indicated above, delays in chicken and gecko are similar to those observed in cat and guinea pig [Shera and Guinan, 2003]. These mammalian species have approximate BM lengths of 26 and 18 mm respectively [Ketten, 2000]. Thus, it does not appear that variations in size of the sensory epithelium readily explain the differences observed in the phase gradients between humans and non-mammals.

As mentioned earlier, OAEs have generally been thought of as a by-product of cochlear amplification. While active processes (i.e. those that boost the ear's response by adding in energy) are not theoretically required to produce evoked emissions, impaired ears (that lack low thresholds and sharp tuning) generally do not emit. It is quite clear that OHCs play a vital role in mammalian auditory function. However, as mentioned earlier, non-mammals lack somatic motility in their hair cells. It is thus clear that somatic motility is not required to produce relatively large SFOAE magnitudes.
3.5.3 Differences in SFOAE Generation Across Species

As described in Sec.3.3.1, one theory regarding the generation of SFOAEs in mammals uses a reflection-based argument [Zweig and Shera, 1995]. The idea is that as energy propagates along the length of the cochlea via a forward-traveling wave, there is a coherent reflection of energy back towards the stapes by randomly distributed impedance irregularities in the cochlear partition [for a review, see Shera and Guinan, 2007]. These irregularities, or mechanical roughness, can arise from a large number of different factors, such as changes in HC coupling to adjacent structures, local variations in fiber distribution in the anisotropic TM, or irregularity of HC distribution along the organ of Corti. The model predicts that significant phase gradients arise in the emissions due to the nature of the traveling wave near the peak region of the response\(^{17}\). Model predictions have shown significant agreement with measured emission properties.

However, it is not clear how or even if this reflection-based model applies to non-mammals, where the underlying anatomy is significantly different. Specifically, in the case of species that lack a tuned/flexible BM (geckos and frogs), the method of energy propagation through the inner ear is not well understood. It is clearly apparent that SFOAEs can exist even in ears where energy does not propagate along the length of the BM via a traveling wave, and that the emissions can still have significant phase gradients\(^{18}\). But are the emissions being generated in some fundamentally different way from mammals? Do SFOAE similarities across species, taken in light of anatomical differences, indicate that modification need to be made to our current explanation for how SFOAEs are generated in the mammalian ear? Or is there perhaps a mechanism in non-mammals analogous to that proposed for mammals? Specifically, might there be a mode of energy propagation in these ears that could

---

\(^{17}\)As the wave approaches the CF spot, it effectively slows down, the wavelength decreases and amplitude of response increases. For a region of place-fixed irregularities that act as scatterers, as frequency is swept across this region, the phase of the reflected component will change rapidly due to the short wavelength of the forward traveling wave.

\(^{18}\)It is worth noting that the roughness at the foundation of the reflection model would still be present for non-mammalian inner ears. Specifically, variations in various properties of the inner ear structures is not a feature specific to mammalian cochlear morphology. What is not clear though is if this roughness plays any meaningful physical role in non-mammals.
allow some sort of reflection-based mechanism to be at work? This may arise in the form of BM traveling wave-analogs, such as energy propagation by means of the TM.

Part of the difficulty in answering these questions lies with the nonlinear nature of the ear. Since a healthy ear exhibits robust nonlinear behavior in numerous regards and SFOAEs are typically measured using nonlinear paradigms, it is hard to delineate between linear and nonlinear effects underlying emission generation. Another complicating factor is the possible role of active processes and their physiological vulnerability: impaired or damaged ears do not emit (and act in a more linear fashion)\(^9\).

It is important to emphasize here that we are not directly measuring ‘reflection’. It is thus difficult to use emissions explicitly to argue for or against a reflection-based mechanism responsible for SFOAE generation. Further study in this regard is needed, particularly modeling efforts based upon observations of the motion of inner ear structures and consider flow of energy through the ear.

### 3.5.4 Emission Delays

In linear systems, when a time delay is present that is unknown, the slope of the phase-frequency function (i.e. the phase gradient, also called the *group delay*) can be used to directly quantify the delay. For background in this regard, see Appendix C. Considering that the ear behaves in a relatively linear fashion near threshold, it may be possible to interpret eOAE phase gradients in terms of time delays associated with emission generation. Caution is clearly needed in this regard. For example, for DPOAE sweeps in mammals using a fixed primary ratio \((f_2/f_1 > 1.1)\), \(2f_1 - f_2\) exhibits a frequency-independent phase for \(f_{DP} > 1.5\) kHz [Shera and Guinan, 1999]. Although the phase gradient is close to zero, this does not mean that the emission appears nearly instantaneously in the ear canal. Rather the frequency-

\(^9\)Limited observations revealed that SFOAEs (measured via the nonlinear suppression) in the chicken disappear almost immediately in the chicken upon death (intra-cardial injected overdose), similar to mammals. Geckos, in contrast, showed a significantly slower time course of emission magnitude decay upon overdose (i.p.), on the order of hours (collumella interuption immediately led to the loss of a detectable emission; see Ch.7). The difference in the time courses (upon death) between chicken in gecko may stem from physiological differences due to the warm versus cold-bloomed nature of the two species.
independent phase is thought to be a result of the scaling-symmetric nature of the
traveling wave in the mammalian ear [Shera and Guinan, 1999].

The study by Meenderink and Narins (2006) examined the connection between
phase gradients and time delays in frog SFOAEs. In spite of the difficulty associated
with clearly defining emission onset in the time-domain (particularly because it is
desirable to use ramped stimuli in order to avoid nonlinear effects associated with
a sudden onset), they were able to show a clear correlation between the onset time
delay and phase gradient measured in steady-state. Another study by Whitehead
et al. (1996) took a similar approach, comparing direct onset latencies and phase
gradients for human DPOAEs. Though there were some exceptions (likely due to
effects of scaling-symmetry), they found that phase gradients correlated to the onset
delays when \( f_2 \) was held constant. These two results thus provide confidence that
emission phase gradients (as measured in the steady-state response) can be used as
a measure of time delays associated with emission generation and the subsequent
propagation back out to the outer ear.

We want to identify here what components of the emission process comprise the
total delay observed. Passage through the middle ear (both to and from) is relatively
short, on the order of 100 \( \mu s \) for mammals [Pascal et al., 1998] and may be even
shorter in non-mammals where the middle ear has a simpler mechanical structure.
Given the dimensions of the inner ear relative to the speed of sound in water, the time
due to energy propagation via compressional waves in the inner fluids would provide
a negligible contribution. So considering the lack of BM traveling waves in some of
the non-mammalian species examined here, a delay on the order of 1 ms or longer
is striking. Presumably the majority of the delay stems from travel times via slower
modes of propagation in the inner ear (such as BM traveling waves) and/or the build-
up response (i.e. tuning) at the site/region of generation. It should be noted that
delay due to wave propagation and tuning are not necessarily mutually exclusive. For
example, a traveling wave grows in amplitude as it approaches its CF spot, making
it difficult to clearly delineate between time associated with propagation to and the
build-up time at the site. At this point, we examine how SFOAE delays correlate to
empirical measures of tuning across species.

3.5.5 Tuning As a Source of Delay

It has been shown empirically in mammalian ears that SFOAE phase gradients exhibit a correlation to estimates of cochlear tuning [Shera et al. 2002]. This has motivated the notion that OAEs might be able to provide a non-invasive measure of the sharpness of tuning. The basis stems from the concept that the more highly tuned a resonant system is, the longer it requires to build up its steady-state response [see Appendix A]. In order to examine if this may be applicable across species, in Fig. 3-8 we compared $N_{SFOAE}$ for each species to $Q_{10}$ reported from various ANF studies. Although there is a lot of scatter in any given ANF study (not to mention across studies), trend lines for the ANF studies indicate similar overall frequency dependencies for $Q_{10}$ in a given species. Except for the frog, $Q_{10}$ tends to increase with frequency.

In the case of the human, psychophysically derived $Q_{ERB}$ values obtained using a forward-masking paradigm exhibit a similar frequency dependence to the $N_{SFOAE}$ values as shown in Fig. 3-8. In the case of both the chicken and gecko, $N_{SFOAE}$ values similarly show a similar frequency dependence relative to that of the $Q_{10}$ values. A vertical offset between the two measures is apparent. This offset can occur due to differences in how one measures $Q$ (for example, $Q_3$ would be shifted up vertically relative to $Q_{10}$) and at what stimulus level the SFOAEs are measured at (as shown in Ch.5, SFOAE phase gradients can vary with respect to changes in stimulus intensity, causing a vertical shift in $N_{SFOAE}$). So while the absolute values of $Q_{10}$ and $N_{SFOAE}$ might not match up (nor be expected to), the observation that they co-vary with respect to frequency suggests the two are related.

We see very different results in the frog ear. $Q_{10}$ does not increase monotonically with frequency as it does in the other species. It is not until about 1.3 kHz that we see the $Q_{10}$ values start to increase (perhaps stemming from differences between the AP and BP). This is in contrast to $N_{SFOAE}$, which increases rapidly with frequency below 1.2 kHz. Thus, $N_{SFOAE}$ and $Q_{10}$ clearly behave differently in the case of the
frog ear, suggesting that there is an additional source of mechanical delay present in the frog ear that is not related to tuning. The frog ear is also interesting in that it shows significantly longer ANF delays relative to those observed in other species [see the 'signal front delay' comparison of Ruggero and Temchin (2007)]. It is not clear what is the basis for the longer delays in frogs, though it may be related to their unique TM topology (i.e. the tectorial curtain, as described in Ch.3.2).

Having examined the empirical correlation between emission phase gradients and measures of tuning, the next step will be to address what fraction of the delay associated with the emissions can be accounted for by tuning. One potential approach would be to start with a simple case: assume that auditory filters take the form of a minimum phase system. Given the empirical data regarding how sharply tuned these filters must be (based upon ANF responses), one can directly estimate the time delay associated with the build-up time of the filter's response. Calculations of this sort can serve to help reveal what the relative contributions comprising the total delay must be and how the emission is being generated.

Furthermore, characterizing $N_{SFOAE}$ values in additional species should serve to elucidate how tuning and emissions are related. In particular, non-human primates would make ideal candidates to test predictions about how tuning may differ between humans and other species because of the similarity in their evolutionary development. As more species are examined, the data will shed light on the debate of whether humans have exceptionally sharper tuning [Shera et al. 2002; Ruggero and Temchin, 2005].

### 3.6 Summary

The key finding described in this chapter are summarized below:

- SFOAEs evoked using a moderate intensity stimulus (40 dB SPL) were detected in all species examined.

---

20 Sharper tuning in humans may have had implications for the development of speech. Furthermore along these lines, it will be insightful to see how $N_{SFOAE}$ compares in some bird species which are capable of mimicking human speech sounds with remarkable accuracy.
Figure 3-8: Comparison of $N_{SFOAE}$ (phase gradient $\times f_p$) with that of either psychophysically (human) or ANF-derived Q-values across four different species. NOTE different scales across species. Correlation is observed in all species except for the frog, which exhibits emission delays significantly longer relative to those expected from ANF measurements. Also, frog Q-values do not systematically increase with frequency as seen in the other species. $L_p = 40$ dB SPL for SFOAE measurements. Q values shown came from the following studies: Human- Oxenham and Shera (2003); Chicken- Manley et al. (1991), Salvi et al. (1991), Saunders et al. (1996); Tokay Gecko- Eatock and Manley (1981), Sams-Dodd and Capranica (1994), Manley et al. (1999); Leopard Frog- Ronken (1991).
• At lower frequencies, geckos and frogs exhibited the largest emission magnitudes, chickens the smallest. Human magnitudes were intermediate.

• In non-mammals, who have relatively poor hearing at higher frequencies (above ~5 kHz), emission magnitudes decreased with frequency. Frog emissions were typically masked by the noise floor by 2.5 kHz while the chicken and gecko persisted beyond the frequency range examined.

• Emission magnitudes varied with respect to small changes in frequency for all species, giving the magnitude curves a rough (jitter-like) appearance. Furthermore, deep notches were apparent in all species except the frog. The frequency at which these notches occurred differed among individuals of a given species.

• All species exhibited significant phase gradients in their SFOAEs. These delays were typically on the order of 1 ms or more, being significantly larger in humans by as much as a factor of 6.

• Given the lack of motility in non-mammalian hair cells, somatic motility is not required for relatively large emission amplitudes.

The comparative study of SFOAEs described here has revealed that there are significant similarities and differences across species. Clearly, SFOAEs are a general feature found in most vertebrate ears, thus the specific morphology of the mammalian cochlea is not required for their generation. Further study of SFOAE properties will serve to reveal how the underlying generation mechanism(s) may or may not be similar across species. The next chapter (Ch.4) examines how SFOAE properties compare to those of DPOAEs and what that implies about the underlying emission generators. Subsequent chapters (Ch.5 and 6) explore how emission properties depend upon stimulus parameters and how that compares across species.

We have also discussed here the suggestion that it is possible to account for emission delays by considering tuning. In the case of the chicken and gecko, there does appear to be a similar frequency dependence between the SFOAE phase gradients and ANF-derived $Q$ values. This is not the case for the frog, suggesting the presence
of an additional mechanical delay present in the frog ear that is not associated with tuning.
Chapter 4

Are Otoacoustic Emissions Generated by Multiple Mechanisms in Non-mammals?

4.1 Approaches Towards Classifying OAEs

Evoked otoacoustic emissions (eOAEs) have served to reveal much about the underlying dynamics of the ear, both in scientific and clinical contexts. However, the actual processes at work in the ear that give rise to these emissions are still not well understood. It has been proposed that different types of mammalian eOAEs (as well as spontaneous emissions, or SOAEs) arise via fundamentally different mechanisms in the ear [Shera and Guinan, 1999]. The physical framework constructed to explain this classification-based approach takes certain morphological features of the mammalian ear into account. One of the most prominent of these features is the basilar membrane (BM) and its corresponding traveling wave.

In mammals, the observed frequency independence of the distortion product emission (DPOAE) phase at 2f1 − f2 for higher frequencies when the primaries are swept at a constant ratio greater than ~1.1 is generally attributed to the scaling-symmetric nature of the BM traveling wave [Shera and Guinan, 1999]. This phase behavior is
in sharp contrast to that of $2f_2 - f_1$ and stimulus frequency emissions (SFOAEs), both of which show significant amounts of phase accumulation. These observations form the basis for the afore-mentioned classification scheme\(^1\) that describes eOAEs as arising from two fundamentally different mechanisms. For the first mechanism, it is proposed that certain emissions arise from non-linear distortion occurring in responses to BM motion, most prominently near the peak of the traveling wave. The second mechanism is described by the theory of coherent reflection [Zweig and Shera, 1995], which postulates that other emissions arise from a reflection of energy due to irregular, but densely distributed, impedance irregularities along the cochlear partition. Thus, there is a clear tie between the two hypothetical processes and differences in the rate of phase accumulation with frequency.

Our goal here is to examine phase properties in a variety of species to see if this eOAE classification applies to non-mammalian ears as well. We will focus here on four different groups: human (*Homo sapien sapiens*), chicken (*Gallus gallus domesticus*), gecko (two species: *Eublepharis macularius* and *Gekko gecko*), and frog (*Rana picipiens picipiens*).\(^2\) A key motivation here is the difference in BM properties (or lack thereof) across species, specifically how energy may propagate through the inner ear. For example, the tonotopic map along the length of the gecko was found to be reversed [Manley *et al.*, 1999], indicating that the tuning seen at the level of auditory nerve fiber (ANF) responses does not stem from mechanical tuning of the BM\(^3\). Furthermore, frogs lack a flexible BM altogether and their hair cells (HCs) sit directly atop rigid supporting structures [Wever, 1985]. It has been suggested that traveling waves may

\(^1\)In their 1999 paper, Shera and Guinan used the term *taxonomy* to refer to their classification of OAEs. Since taxonomy is a term that is generally used in classifying living organisms and our study compares OAEs across a wide array of species, we will avoid using the term in the hope of not creating any unnecessary confusion.

\(^2\)See Ch.3 for a description of anatomical and physiological differences across these species as well as motivation for using two separate gecko species.

\(^3\)As described in Ch.3, the gecko BM/BP changes its width and thickness significantly along its length. Manley *et al.* found that high CF fibers innervated the large massive end of the papilla while low CF fibers the thin, smaller end. Similar results were observed in the bobtail lizard. These observations, implying a lack of a tuned response in the BM along its length, are consistent with direct mechanical measurements in other lizard species that showed the basilar papilla has a fairly uniform motion along its length [Peake and Ling, 1980]. In lizards, tonotopy is thought to arise at the micro-mechanical level (i.e. resonance of HC bundles, interaction between the HC-TM complex, etc.).
still be present in the inner ear of the frog, but at the level of the tectorial membrane (TM) [Hillery and Narins, 1984]. By examining eOAE phase properties across an array of ears with varied anatomy and physiology, we stand to gain further insight into how emission generation mechanisms vary across species.

4.2 Definition of an eOAE Generation Mechanism

It is important from the outset to define the terminology that will be important for the interpretation of the data and subsequent discussion. The entire region of interest is schematized in Fig. 4-1. We define four key terms. First, we simply define a mechanism as a physical process. For example, a mechanism may be hair cell transduction or the propagation of energy thru cochlear fluids or along the BM. Thus, a mechanism as we define it here is not limited solely to the inner ear structure that generates the emission. Second, the system is what specifically forms the basis for an OAE. The OAE system includes all the processes involved in the transmission of sound into the inner ear, the generation of the OAE and the transmission of sound to the outer ear, and is comprised of multiple mechanisms. Contributing mechanisms to the system may include a non-linear resonator, a delay-line, and a transformer (i.e. the middle ear). In Fig. 4-1, the system of interest is bounded by the red box. It also includes the forward path (from the stimulus location to the generation site), since any relevant transfer function (such as a delay) is going to affect the ‘input’ to the generation site. Third, OAE generation occurs at a site. A site is the physical location at which the emission originates. Multiple sites can contribute to the formation of an emission. In such an extended region of generation, the OAE is the integration across all the contributing sites. When multiple sites exist, their associated emissions may or may not be produced by the same type of mechanism. Lastly, is the notion of a phase gradient (also commonly referred to as a group

---

4From Hobbie (1997): 'A system is that part of the universe that we choose to examine. The surroundings are the rest of the universe. The system may or may not be isolated from the surroundings.'

5Ideally, efferent effects would also be included in the system as we so defined it. We assume however that our stimulus levels here are low enough so not to introduce appreciable efferent effects.
delay). This is a measure of the rate of phase accumulation with increasing frequency of the emission and can reveal information about underlying time delays⁶.

Because OAEs arise from the system as a whole, they may be comprised of components stemming from both different mechanisms and different sites. Our primary approach here will be to use emission phase gradients to help us delineate among different mechanisms (with our knowledge of human phase gradient patterns serving as the reference). We make the following assumption: differences in the phase gradient properties imply differences in generation mechanisms. For example, if OAE energy were to propagate from the site(s) of generation back towards the middle ear either via a compressional wave or a backward traveling wave, we may expect two very different phase gradients due to the differences in the time scale of energy propagation (even if we make our measurements in steady-state).

4.3 eOAE Results

An overview of DPOAE results across species is given below (see Ch.3 for a similar treatment with respect to SFOAEs). We then examine how the behavior of both low-side and high-side cubic DPOAEs (measured using a fixed primary ratio) and

⁶See Appendix C for further discussion.
SFOAE phase gradients compare\textsuperscript{7}. Lastly, we present data indicating how DPOAE phase gradients depend upon the primary ratio.

4.3.1 DPOAEs

Fig. 4-2 shows representative DPOAE spectra for the different species examined here, using a uniform set of stimulus conditions to facilitate comparison ($L_1 = L_2 = 65$ dB SPL, $f_1 = 1$ kHz and $f_2/f_1 = 1.22$). Intermodulation DPOAEs were present in all species, although they tended to be smallest in humans and chickens and largest in geckos. Numerous higher order distortion products are readily apparent in the gecko ear. Higher order harmonic distortions were also present in all species, though any harmonic DPOAEs are likely to be obscured by significant (harmonic) system distortion associated with the ER-10C earphones\textsuperscript{8}. Higher order intermodulation distortions (e.g. $3f_1 - 2f_2$) were present, though our analysis here is limited to describing the cubic terms only ($2f_1 - f_2$ and $2f_2 - f_1$).

Expanding this out, Fig. 4-3 shows how DPOAE magnitudes vary across both species and individuals with respect to frequency. Each value contributing to the mean came from a unique ear (i.e. no repeated measures are included). For $2f_1 - f_2$, emissions are clearly largest in the gecko for $f_{DP} < 2$ kHz, with leopard geckos having larger magnitudes relative to the tokays. With the exception of the human and frog, $2f_1 - f_2$ decreased in magnitude with increasing frequency. In the human, $2f_1 - f_2$ increased with frequency. The frog shows a relatively flat $2f_1 - f_2$ response up to $\sim 1.5$ kHz, above which it falls off precipitously\textsuperscript{9}.

The $2f_2 - f_1$ magnitudes were largest in the geckos at all frequencies tested.

\textsuperscript{7}For DPOAEs, low-side refers to distortion products whose frequency is less than that of the lower frequency primary $f_1$. Conversely, high-side emissions have a frequency greater than that of $f_2$.

\textsuperscript{8}The ER-10C probe uses two earphones (each presenting a single tone), significantly reducing any system intermodulation distortion. However, each earphone emits a large amount of harmonic distortion, whose magnitude varies with respect to the acoustic coupling impedance between the probe and ear. This is discussed in greater detail in Ch.2.

\textsuperscript{9}At higher levels, it becomes apparent that the frog DPOAE frequency distribution takes on more of a bi-modal shape [Meenderink and van Dijk, 2004]. This shape is just starting to become apparent for $2f_1 - f_2$ in Fig. 4-3. The basis for the bi-modal shape is believed to be due to a higher threshold for evoking emissions from the higher frequency-sensitive BP. This bi-modal distribution is more clearly apparent in the phase gradients, as shown in Fig. 4-5.
Figure 4-2: DPOAE spectra from individual ears in four different species. Stimulus conditions were the same for all. \( L_1 = L_2 = 65 \text{ dB SPL}, f_1 = 1 \text{ kHz}, f_2/f_1 = 1.22 \)
Figure 4-3: DPOAE magnitude comparison (both \(2f_1 - f_2\) and \(2f_2 - f_1\)) across species for fixed set of stimulus conditions \((L_1 = L_2 = 65\ \text{dB SPL}, \ f_2/f_1 = 1.22)\). Trend lines are included to aid visualization. Both plots provide statistics where data point are averaged into octave bins. The mean and 95\% confidence intervals are plotted (see Fig. 4-6 caption for further information). \(2f_1 - f_2\) bin limits (in Hz) are as follows: frogs \([300-600, 600-1200, 1200-2400, 2400-4000]\), others \([380-800, 800-1600, 1600-3200, 3200-4000]\). \(2f_2 - f_1\) bin limits (in Hz) are as follows: frogs \([400-800, 800-1600, 1600-3200, 3200-5000]\), others \([700-1400, 1400-2800, 2800-5600, 5600-8000]\).

In humans and chickens, the magnitude of \(2f_2 - f_1\) was smaller relative to that of \(2f_1 - f_2\) by \(\sim 5-10\ \text{dB}\). At higher frequencies \((f_{DP} = 2-5\ \text{kHz})\), the \(2f_1 - f_2\) magnitude in the human actually increased relative to lower frequencies. For the geckos and frog, \(2f_1 - f_2\) was slightly smaller than \(2f_2 - f_1\) at a given distortion frequency by \(\sim 5\ \text{dB}\), though the gap widened between the two distortions for \(f_{DP} > 1\ \text{kHz}\). The bottom panels of Fig. 4-3 show statistics of the emission magnitudes when points are averaged into octave bins.

Emission phase is shown in Fig. 4-4. This figure shows all the phase curves measured for the ears tested for a given species, including repeated sessions for the same
ear. Depending upon the species, the two different distortions (as indicated by the red and green curves in Fig. 4-4) can show very different amounts of phase accumulation (this figure is further discussed in Sec.4.3.2). The amount of accumulation is quantified in Fig. 4-5 by plotting the $N$-values of the phase curves. Briefly, $N$ is the product of the phase gradient (local derivative of the phase-frequency function) and emission frequency. This is a dimension-less quantity that can be regarded as the phase delay expressed in units of stimulus cycles (see Ch.3 for further discussion). A large amount of scatter is present due to variations in the smoothness of the phase-frequency function (which are effectively amplified when taking the derivative)$^{10}$. $N_{2f_1-f_2}$ is largest at lower frequencies (below $\sim 1$ kHz) in the human and frog and decreases with increasing frequency for $f > 1-2$ kHz. The bi-modal shape is readily apparent in the frog ear, with the trough centered near the transition frequency ($\sim 1.2$ kHz) between the AP and BP [Ronken, 1991]. Though smaller in the chicken and geckos, $N_{2f_1-f_2}$ tended to increase with frequency in these species.

A different picture is apparent for $N_{2f_2-f_1}$, where humans $N$-values are significantly larger than the other species and the overall behavior is qualitatively similar to that seen for SFOAEs (see Fig. 3-5). Though the relative difference between the low and high-side distortions is not as dichotomous as the human, $N_{2f_2-f_1}$ tended to be larger than $N_{2f_1-f_2}$ in all other species (particularly at higher frequencies). Overall, phase gradients for both DPOAEs were highly similar across the two gecko species.

### 4.3.2 SFOAE vs. DPOAE Phase Gradient Comparison

Phase curves are compiled together for both SFOAEs and DPOAEs in Fig. 4-4. The same stimulus conditions were used across all species [SFOAE: $L_p = 40$ dB SPL, $L_o = 40$ dB SPL, $f_s = f_p + 40$ Hz; DPOAE: $L_1 = L_2 = 65$ dB SPL, $f_2/f_1 = 1.22$ (constant)]. In a select number of ears, repeated measures were made in different

$^{10}$ These can stem from small bends to large discontinuities/jumps in the phase. These are readily apparent in Fig. 4-4. For our present purposes here, we focus solely on the overall trends observed in the phase behavior.
Figure 4-4: Comparison of emission phase across different OAE types. **NOTE** the different scales across species. Chickens appear qualitatively similar to humans in that 2f1-f2 exhibits a strong degree of frequency independence. This is in contrast to the other three species. However, all species exhibit significant phase delay (on the order of 1 ms or greater). For each species, individual curves are plotted using varying line thickness. The total number of unique ears tested is indicated by N. Plots also include repeated measurements in some individual ears at different experimental sessions, as indicated by the bracketed number which shows the total number of curves plotted. Some phase curves were offset vertically by an integral number of cycles for clarity. [SFOAE: $L_p = 40$ dB SPL, $L_s = 55$ dB, $f_s = f_p + 40$ Hz, DPOAE: $L_1 = L_2 = 65$ dB SPL, $f_2/f_1 = 1.22$ (constant)]
Figure 4-5: N-values for DPOAEs shown in Fig. 4-3. Only points whose magnitudes were at least 10 dB above the noise floor were included.
experimental sessions and are also included in Fig. 4-4, the numbers (of both unique ears present and total number of curves) being given in the bottom-left hand corner for each species. Note the different scales across species.

Relations among the phase curves for humans are consistent with what has been previously reported [Shera and Guinan, 1999; Knight and Kemp, 2000] and form the basis for the OAE classification criteria described in the introduction. The $2f_1 - f_2$ phase is relatively independent of frequency above $\sim 1.5$ kHz while both SFOAE and $2f_2 - f_1$ show significant phase accumulation. A qualitatively similar picture as in the human is seen in the chicken phase curves. The $2f_1 - f_2$ phase is relatively flat, while the SFOAE and $2f_2 - f_1$ curves have a clear downward trend (the total amount of $2f_2 - f_1$ phase accumulation is significantly less than that for the humans).

Results in the two gecko species were highly similar to each other. In contrast to the human and chicken, all phase gradients run roughly in parallel at all frequencies. However, when making the comparison using different stimulus levels, some rare exceptions for the gecko were found as indicated in Fig. 4-6. For this specific gecko using lower intensity stimuli, the overall behavior looks qualitatively similar to the human and chicken (i.e. $2f_1 - f_2$ is significantly flatter than either the SFOAE or $2f_2 - f_1$). Thus it is not entirely clear how generalizable differences between the geckos and the humans/chickens are when varied stimulus conditions are accounted for. No ‘exceptions’ in this regard were observed in the frog. Like the majority of geckos, both the cubic DPOAE phase curves in the frog had similar amounts of phase accumulation.

Comparing $2f_2 - f_1$ phase gradients to those of SFOAEs (in Ch.3, see Fig. 3-5 and Fig. 3-6), it is apparent that $2f_2 - f_1$ gradients are smaller than those of SFOAEs by roughly a factor of 0.6 in all species. This could possibly reflect differences in the stimulus intensity between the SFOAE and DPOAE paradigms (see Ch.5).

One general observation across all species is that the phase curves exhibit some degree of frequency dependence for both DPOAEs and SFOAEs. Phase curves tend to be steeper at lower frequencies (below 1-2 kHz), with transition to a slightly more shallow slope at higher frequencies. For humans, a ‘knee-point’ is apparent around
2 kHz for both SFOAE and $2f_2 - f_1$ phase curves, above which the slope decreases roughly by a factor of two. The change for $2f_1 - f_2$ in humans is more drastic, with a knee-point occurring down around $\sim 2$ kHz and significantly steeper slopes at the lower frequencies (though $N_{2f_1-f_2}$ is still significantly less than $N_{2f_2-f_1}$ in that lower frequency region). In the case of the frog, the knee-point occurred around 1.2 kHz, the transition frequency between the two papilla.

### 4.3.3 DPOAE Dependence Upon Primary Ratio

As shown by Knight and Kemp (2000), DPOAE phase gradients for fixed primary ratios can have a significant dependence upon the ratio, particularly $2f_1 - f_2$. In the human ear, the $2f_1 - f_2$ phase is highly frequency dependent at lower ratios (typically for $f_2/f_1$ below $\sim 1.1$), in a manner similar to that of the SFOAE and $2f_2 - f_1$ phase curves. Keep in mind that as the primary ratio gets small, both $2f_1 - f_2$ and $2f_2 - f_1$ approach the SFOAE frequency. At higher ratios, there is a sharp transition
Figure 4-7: DPOAE ratio dependence in both chicken in gecko (both species pooled together) for fixed $f_2/f_1$ sweeps. For a given species, each curve represents an individual ear. For any given ear, both $2f_1 - f_2$ and $2f_2 - f_1$ were measured simultaneously. Human results were qualitatively similar to the chicken while frog results were similar to the gecko. 

$[L_1 = L_2 = 65 \text{ dB SPL}, f_1 = 500 - 5000 \text{ Hz in 100 Hz steps}]$

where $2f_1 - f_2$ becomes frequency-independent. $2f_2 - f_1$ phase behavior is relatively invariant to the primary ratio.

Figure 4-7 shows how DPOAE phase gradients depend upon primary ratio in non-mammals. Chicken results were qualitatively similar to those seen in humans. The $2f_1 - f_2$ phase shows significant phase accumulation at a small ratio while very little at higher ratios. Also similar to humans, the transition for chickens occurred around $f_2/f_1 \approx 1.09 - 1.1$. The result was quite different in the gecko ear (both species are pooled together in Fig. 4-7), where both DPOAE phase curves were fairly invariant to primary ratio. Results in the frog were similar to those of the gecko.
4.4 Discussion

4.4.1 Comparison with Previous DPOAE Studies

In this section, we discuss how our DPAOE results compare to those reported in previous studies. We know of no reports of eOAEs from the gecko ear\(^{11}\), thus we can compare only to reports examining human, chicken or frog emissions. Overall, our DPOAE measurements were fairly consistent with these previous reports.

Human DPOAEs appeared consistent in regards to both magnitude and phase gradients with previous reports [Shera and Guinan, 1999; Knight and Kemp, 2000]. For chickens, there is some variation of DPOAE properties across studies. To the best of our knowledge, phase gradients have only been reported in human and frog, thus limiting us to magnitudes only for comparisons in the chicken. Our values tended to be slightly smaller (~5 dB) than those reported by Lichtenhan et al. (2005) and closer to those reported by Ipakchi et al. (2005) and Kettembeil et al. (1995). However, all studies were typically within ±7 dB of each other\(^{12}\). Kettembeil et al. (1995) also noted that chicken DPOAEs are sensitive to the anesthesia paradigm used\(^{13}\). Our anesthesia paradigm was identical to that used Ipakchi et al. (2005). Considering that there was not very much variation across the previous studies listed (all used a different anesthesia paradigm), it does not appear that the anesthesia accounts for the smaller emission magnitudes in the chickens relative to the other species. It may be possible that chicken eOAE phase gradients exhibit some dependence upon anesthesia.

---

\(^{11}\)Both species of gecko tested here have been shown to exhibit SOAEs [Manley et al. 1996], though our observations here did not indicate them to be as highly prevalent as reported. However, it is worth noting that DPOAEs have been measured in a number of different lizard species [Rosowski et al. 1984; Manley et al. 1993].

\(^{12}\)In the study by Ipakchi et al. (2005), they note that DPOAE I/O functions are nearly identical in 0-day and 12-day post-hatch chicks (and presumably vary little subsequently with further maturation into adulthood). Thus, they propose that DPOAE magnitudes are fully mature at birth in the chick ear. This is consistent with a lack of change observed in evoked potential thresholds (measured at the cochlear nucleus) for chicks with respect to time post-hatch [Manley, 1990]. However, the Lichtenhan et al. study did show slightly higher magnitude emissions, and they used fully adult birds in their study. So there may be some slight changes with developed maturation in the chicken.

\(^{13}\)They observed that deep anesthesia, induced with either Nembutal or Halothane, led to an almost complete loss of detectable emissions. This is in contrast to our observations, where the birds were deeply anesthetized and had readily apparent emissions (though we used urethane).
With respect to the frogs, Meenderink and van Dijk (2004) examined the same species we did. While the overall trends with respect to frequency were similar, their reported magnitudes were typically 10 – 15 dB smaller than those reported here. The basis for this discrepancy is not clear. Excluding the case of a single ear, all frogs tested exhibited DPOAEs. The exceptional ear had no observable emissions at all, independent of the stimulus intensity (though the contralateral ear did emit). Phase gradients in the frog ear were similar across studies, though our measured values tended to be larger for $2f_1 - f_2$, particularly at lower frequencies.

Furthermore, it should be noted that differences in phase gradient properties are not solely limited to mammals versus non-mammals. Shera and Guinan (2003) showed that SFOAE gradients were significantly larger in human than cat and guinea pig, which were more comparable to the chicken\(^\text{14}\).

### 4.4.2 Effect of Stimulus Parameter Choice

We use the behavior of emission phase as criteria for determining if different underlying mechanisms are responsible for generation. However, we need to be careful with our choice of stimulus parameters, as emission properties can vary significantly depending upon what values are used. This stimulus dependency places a bound upon the conclusions that can be drawn from this study, where a limited set of stimulus conditions was used.

How conclusions can vary with respect to choice in stimulus conditions is most apparent with regards to the primary ratio. At small ratios in humans and chickens, both $2f_1 - f_2$ and $2f_2 - f_1$ share a similar frequency dependence and exhibit a significant amount of phase accumulation. At higher ratios, the low-side emission suddenly shifts to having a relatively frequency-independent phase (at least for emission frequencies above $\sim 1$ kHz), as shown in Fig. 4-7. This transitional ratio ($\sim 1.09$-$1.11$ for both humans and chickens) is believed to mark the point at which the dominant emission source shifts [Knight and Kemp, 2000]\(^\text{15}\). The observation that the transition

\(^{14}\)As discussed in Ch.3, emissions extend to much higher frequencies in these mammalian species than they do in non-mammals.

\(^{15}\)When two or more sources are present that contribute to the emission, the dominant one will dic-
ratio between both human and chicken is similar suggests DPOAEs are generated in a similar fashion in both species.

Emission phase gradients can also vary with respect to level [Ch.5], thereby introducing another dimension of parameter space that needs to be considered. This is apparent in Fig. 4-6, which indicates that certain gecko ears can behave qualitatively very similar to that of humans and chickens under certain stimulus intensity conditions. Both these observations, that phase gradients are sensitive to primary ratio and stimulus intensity, make clear that one needs to consider stimulus parameters when making conclusions about underlying generation mechanisms. Having said that, we continue the discussion from this point based upon the choice of stimulus levels made for this particular study. Ch.6 examines the possibility that multiple level-dependent mechanisms are present and how they may interfere to produce non-monotonic emission growth.$^6$

$^{4.4.3}$ Do Non-Mammals Show Evidence for Multiple Mechanisms?

Our results here for the human ear were consistent with the notion of multiple generation mechanisms [Shera and Guinan, 1999]: phase curves for SFOAEs and high-side cubic DPOAEs showed large phase gradients while low-side DPOAEs had a gradient close to zero (for a fixed primary ratio). As discussed in a subsequent section, this conclusion for the human ear is only valid for emission frequencies above $\sim 1$ kHz. The conclusion of the existence of multiple mechanisms rests upon our definition that significant differences in emission phase behavior are the basis for distinguishing mechanisms.

eOAE results from the chicken exhibit a very similar qualitative behavior to those seen in the human (and other mammals). Both $2f_2 - f_1$ and SFOAE manifest a significant amount of phase accumulation while $2f_1 - f_2$ is relatively frequency insensitive the actual emission properties. Un-mixing allows for these different components to be separated and their individual properties examined [Kalluri and Shera, 2001].

$^6$ Note that multiple mechanisms at a fixed level and multiple level-dependent mechanisms are not necessarily mutually exclusive.
MECHANISMS FOR eOAE GENERATION ACROSS SPECIES

[for higher frequencies ($f_{OAE}>\sim 1$ kHz) and fixed stimulus intensity]

**Figure 4-8:** Possible scheme to classify emission generation mechanisms across species. Note that this is likely to be valid only at higher frequencies (above $\sim 1$ kHz) and for a fixed stimulus intensity.

... dependent. Given the similarity between the two species, the proposed mammalian OAE classification scheme is also applicable in the chicken ear as well. Furthermore, the similarity in phase gradients between $2f_2 - f_1$ and SFOAE gradients in the chicken suggests that these two emission types are being generated by the same mechanism, similar to that in humans.

The behavior observed in the gecko and frog ear was very different from that observed in humans or chickens. For the gecko and frog, the emission phase gradients all showed similar phase accumulation, indicating that the classification criteria are not met. Thus, no evidence for multiple mechanisms at work generating eOAEs is present in their ears. However, as shown in Fig. 4-6, there were exceptions to this conclusion found for geckos, noticeable at higher frequencies and when using lower intensity stimuli. It is possible that the gecko ear may manifest features similar to that of mammals and birds in a stimulus-dependent way, perhaps deriving from developments to improve higher frequency hearing. It is not clear why exceptions were only found in a small number of geckos and further study is needed to better resolve this issue.
Based upon our results, Fig. 4-8 shows a possible species categorization for the classification of OAE generation mechanisms. Both mammals and birds exhibit two different mechanisms, one stemming from linear reflection and the other from nonlinear distortion. In the case of the frog, all emissions arise from a single mechanism, namely nonlinear distortion. A key difference between two groups is whether the distortion mechanism behaves in a scaling-symmetric fashion or not, described further in Sec.4.4.6. The geckos appear to be somewhere in-between the two groupings. Over the stimulus range primarily tested, the geckos would be grouped with the frogs in the single mechanism category. Placing geckos in the single mechanism category is strengthened by the observation that their emission phase gradients are invariant with respect to primary ratio (Fig. 4-7), similar to that of the frogs and in contrast to humans and chickens. However, given that a limited number of geckos exhibit a similarity in phase gradient behavior relative to humans and chickens when using lower intensity stimuli, it may also be argued that they belong in the multiple mechanism category. This notion is supported by the observation that $N_{2f_2-f_1}$ was consistently larger than $N_{2f_1-f_2}$ in the gecko ear (though this was also the case for the frog).

As indicated on Fig. 4-8, the gecko BP has two morphologically distinct parts. The bi-directional salletal region (ventral end) is where high-CF ANFs innervate, while the unidirectional continuous TM region (dorsal end) is specific to lower CF units [Manley et al., 1999]. The distinction between these two regions may form the basis that explains why (in certain instances) that geckos can exhibit phase behavior similar to humans and chickens (Fig. 4-7). Specifically, the ventral region may operate mechanically in a very different fashion relative to the dorsal region (see Sec.4.4.6)\textsuperscript{17}.

The question as to which category the gecko falls in returns to the point about choice of stimulus parameters for comparisons (Sec.4.4.2) and indicates the limitations associated with the proposed categorization. One important limitation is the actual emission frequency. As described in Sec.4.4.4, significant changes in the phase

\textsuperscript{17}Frogs also have two morphologically distinct regions that respond to lower (AP) and higher (BP) frequencies. These two regions are spatially distinct, physically separated in different portions of the inner ear. The higher frequency BP is thought to act as a very simple filter, tuned to a single frequency [Ronken, 1991]. There is no evidence it behaves in a fashion similar to that of higher frequency regions (sensitive above $\sim1.2$ kHz) of the other species examined here.
gradients occur at lower frequencies, indicating that that this categorization may not be valid in that frequency range.

4.4.4 Differences at Lower Frequencies

At lower frequencies (below ~0.5-1 kHz), we see many differences in emission properties relative to higher frequencies. Phase gradients typically increase, as a 'knee-point' can be seen in the human and frog phase curves shown in Fig. 4-4\textsuperscript{18}. One particularly notable observation is that $2f_1 - f_2$ phase gradients increase significantly at lower frequencies in both the human and chicken. Though the SFOAE and $2f_2 - f_1$ phase gradients are still significantly larger than those of $2f_1 - f_2$ at these lower frequencies, it is apparent that there is a change in the underlying $2f_1 - f_2$ generation mechanism. Two possibilities may account for this transition in the low-side DPOAE phase behavior. First, at lower frequencies the two primaries come closer together and thus $2f_1 - f_2$ may become dominated by a generation mechanism similar to that of a SFOAE. Second, the difference at lower frequencies could stem from how the basal and apical parts of the cochlea process sound differently. For example, scaling symmetry is known to break down in the apex [Shera and Guinan, 2003; van der Heijden and Joris, 2006], an effect that would most likely impact the phase response. While our measurements did not extend to very low frequencies, further study of eOAEs using stimulus frequencies below 500 Hz will provide us with a better understanding how the ear processes lower and higher frequencies differently\textsuperscript{19}.

Another consideration with regard to lower frequencies stems from the significant difference between mammals and non-mammals in the radius of curvature of the inner ear. It has recently been proposed that the coiled structure of the mammalian cochlea has an advantageous effect at lower frequencies [Chadwick \textit{et al.}, 2006]. The chicken BM curves only slightly, typically less than 90° (compared to nearly two complete

\textsuperscript{18}The knee-point is more apparent when looking at the phase gradients. Shown in Fig. 3-4 (or Fig. 3-7 for a log-log scale), SFOAE gradients vary with frequency at lower emission frequencies (3 kHz for humans, 1.2-1.3 kHz in frog) and are relatively constant at higher frequencies.

\textsuperscript{19}A major hurdle here is the technical difficulties associated with this lower frequency eOAE measurements. Chiefly among these is the difficulty in presenting sufficiently high enough intensity stimuli to evoke an OAE and its detection amid the increased acoustic noise floor.
turns in smaller mammals with a similar BM length) [Manley, 1990]. The gecko BM is fairly straight, with practically no curvature at all [Wever, 1978]. The frog AP does exhibit a high degree of curvature due to it short length and Ω-shape [Lewis et al., 1982], though it is not known what mechanical effect this may have (if any). Since non-mammals tend to hear well at lower frequencies but lack the curvature, the comparative approach can potentially shed light on whether the coiled structure of the mammalian cochlea plays an important role or not20.

4.4.5 Differences in Size

The physical dimensions of the ear may also play an important role in regard to differences in emission behavior. The sensory epithelia in the gecko (1.2-1.6 mm) and frog (≈0.5 mm for the AP) are significantly shorter than those of the human and chicken and may account for differences in low-side and high-side DPOAE phase gradients. Though the BM in chickens is much shorter than humans (or guinea pig and cat21), it is still comparable in size to smaller mammals such as mice. Though we do not know of any published reports of eOAE phase gradients in the mouse ear, since their DPOAEs are readily measurable, comparison between mouse and chicken might shed further light on how differences in size affect eOAE phase gradients and presumably tuning (though there is little overlap in frequency of their audible ranges). Further discussion on how the size of structures in the inner ear affect emission properties is provided in Ch.7.

4.4.6 Traveling Waves in Non-Mammals?

The logical step at this point is to relate our results back to what we know about the anatomy and physiology of the different species to try to better understand what role traveling waves may or may not be playing. It should be emphasized here initially that large phase delays do not necessarily equate to time delays associated with traveling

20A commonly accepted notion is that the coiling plays no functional role in cochlear mechanics, except to save space by making the ear more compact.

21BM length in guinea pig is ≈18 mm and ≈26 mm for cat [Ketten, 2000].
waves. Significant phase accumulation may be expected solely from a highly tuned filter as it builds up energy towards its steady-state response\textsuperscript{22}. Furthermore, as has been previously pointed out, it may be difficult to segregate delays into discrete components [Shera et al., 2000]. In the mammalian cochlea, the amplitude of the traveling wave grows as it moves towards it characteristic place. Thus, the delay associated with energy propagation towards the site of generation and the build-up time associated with the generator are likely to be inextricably linked.

Given the similarity between human and chicken phase gradients, coupled with our understanding of cochlear mechanics, the OAE data suggests that a scaling-symmetric traveling wave is present in the chicken ear. This observation is consistent with mechanical measurements that have been made in chicken [Bekesy, 1960] and pigeon [Gummer et al., 1987], the latter of which reported approximately scaling-symmetric traveling waves\textsuperscript{23}. This is interesting in light of the differences of the cross-sectional anatomy between human and chicken (see Fig. 3-1), difference in curvature in the inner ear, and a lack of HC somatic motility [He et al., 2003; Koppl et al., 2004]\textsuperscript{24}. Seeing that there is a lack of evidence pointing towards multiple generation mechanisms in gecko and frog, one might not find the difference relative to humans and chickens surprising given the unlikely existence of BM traveling waves in the gecko ear and a tuned BM altogether in the frog. It may however be possible that there are some analogs to traveling waves at work in these species. By analog, we mean some sort of underlying mechanism present in the ear that propagates energy throughout the inner ear such that the propagation varies in a spatial and temporal fashion. These analogs may be completely independent of the BM and may potentially arise at the level of the tectorial membrane (TM), which has been demonstrated in mammals to support shearing waves [Ghaffari et al., 2007]. Given the unique tectorial structure present in the gecko ear as well as the massive and strongly coupled TM in

\textsuperscript{22}See Appendix C for additional discussion.
\textsuperscript{23}Bekesy clearly showed a tuned BM response in the chicken. The study by Gummer et al. pointed out that this is insufficient to show the existence of a traveling wave, that one must also show a degree of phase accumulation. They found that the detection of the scaling-symmetric traveling waves was highly sensitive to the physiological state of the bird, but that a tuned BM was not.
\textsuperscript{24}Many of the gross anatomical differences between the mammalian and avian inner ears are summarized in Gummer et al. 1987.
the frog [see Ch.3], the prospect of TM-traveling wave analogs might be realizable.25

Expanding upon what might make the mammalian and avian ears unique, the frequency independent $2f_1 - f_2$ phase is thought to be a result of scaling symmetry. So not only do traveling waves appear to be an important mechanism in mammals with regards to emission generation, but so does their scaling symmetric nature. One question that arises is how common a feature scaling symmetry might be in a traveling wave given very different anatomies across these species. As mentioned in the introduction, it has been suggested that traveling waves are present in some form in the amphibian ear. This notion however, is still a current topic of debate. Converse arguments posit that observed delays in the frog (for example, either in OAEs or ANF responses, once synaptic and conduction delays are accounted for) arise purely from what might be expected from a highly tuned filter, as mentioned at the start of this section. In their study of frog DPOAE phase behavior, Meenderink et al. (2004) explicitly state that their data 'do not provide any evidence for cochlear-like traveling waves'. The basis for this argument appears to stem from their observation that frog DPOAE phase gradients differ significantly from those of mammals (similar to our observations here) and results from a simple transmission-line model [Knight and Kemp, 2001]. While true at face value, their statement may be misleading. It may be possible to have traveling waves present in the frog, but that they behave in a way very different from that in the mammalian and avian ears. In particular, frogs may exhibit traveling waves which do not behave in a scaling-symmetric fashion (that would lead to the dichotomous behavior between the upper and lower-sideband intermodulations). Perhaps more direct mechanical measurements observing TM motion applied to the amphibian papilla in the frog ear will help resolve this issue in the future [Schoffelen et al., 2007].

Further study is needed to better understand how energy propagates within the ear for different species and what physical role a scaling-symmetric response may

---

25The TM in the chicken ear is generally much more highly coupled (relative to the mammalian cochlea) to the HC bundles and the apical surface of their cell bodies. It is not currently well understood what effect this stronger coupling means mechanically, though it is thought to play an important role in the processes underlying regeneration [Cotanche, 1999]. It is also significantly more massive relative to the mammalian TM.
have. Since scaling-symmetric behavior appears limited to more basal parts of the mammalian cochlea and because non-mammals in general do not hear well at higher frequencies, it may be possible that scaled responses do not play a significant role in non-mammals. Specifically, a scaling-symmetric response may be a feature unique solely to a tuned BM and only at higher frequencies. Avian ears (as well as gecko ears under specific stimulus conditions) may represent the morphological transition between where scaling does and does not plays a significant role. A better empirical foundation of the how traveling wave behavior differs between mammals and birds will further our understanding of the precise role traveling waves play in cochlear mechanics.

4.4.7 Differences in Emission Magnitudes

One of the striking differences among the species is the differences in emission magnitudes for a similar set of stimulus conditions. At lower frequencies, DPAOEs are largest in the geckos relative to the other species while for SFOAEs, both gecko and frog show larger emission magnitudes (see Fig. 3-3). Even from the DPOAE spectra shown in Fig. 4-2, it is clear that the gecko ear emits significantly more distortion than the other species\textsuperscript{26}. Conversely, chickens showed the smallest emissions, independent of emission type or stimulus frequency\textsuperscript{27}.

The basis for relatively larger magnitudes in gecko and frog is not clear. This observation indicates that generation mechanisms must vary to some degree (independent of phase gradients). One possible explanation may be that there is significantly more mechanical nonlinearity present in the inner ear of these species. Second, the actual generation site(s) in the inner of geckos and frogs may produce emissions in a relatively more efficient way compared to the other species. Lastly, the differences

\textsuperscript{26}When clearly present, such as in the case of the gecko ear, a detailed study of higher order distortions may yield significant insight into the nonlinearity present in the ear giving rise to DPOAEs.

\textsuperscript{27}As discussed in the results section, this does not appear to be due to anesthesia effects. It had been suggested that surgery may facilitate acoustic coupling of the probe to the ear in chickens. However, this was found to cause a significant build-up of fluid on the tympanic membrane which interfered with measurement. Thus, all chicken OAEs measurements were made with a fully intact outer ear, save for a few feathers removed to allow for a clear acoustic path.

101
may have more to do with some sort of filtering of the emission subsequently after generation. This could possibly stem from differences in how energy propagates from the generation site to the middle ear (i.e. slow-wave versus fast-wave propagation). Furthermore, the larger magnitude emissions might also be explained by differences in the reverse transmission properties of the middle ear.

4.4.8 Spontaneous vs. Evoked Emissions

While this study did not closely examine SOAEs, several points about their existence relative to evoked emissions can be made. In all our experimental sessions, at least one attempt was made to measure SOAEs. Humans and geckos showed readily apparent SOAEs, while the chicken and frog showed none in all the ears tested. The proposed classification scheme predicts that both SFOAEs and SOAEs are generated by a similar mechanism. Given the presence of SFOAEs in the chicken and overall similarity of DPOAE gradients with humans, the lack of observable SOAEs in the chicken is curious, if SOAEs are even present at all.

Several possibilities may serve to account for a lack of SOAEs in ears that readily exhibit eOAEs. One explanation may be tied to the fact that SOAEs are rarely observed in most mammalian species, thus their absence in the chicken (and frog in our case) is not all that surprising. This is consistent with the observation that DPOAE phase behavior in species such as rabbit is very similar qualitatively to humans, but rabbits lack observable SOAEs. Another possibility may be due to

---

28 This may seem moot, as our stimulus paradigm measures emissions in steady-state. However, there may be fundamental differences in how emissions propagate out of the inner ear across species that could effectively filter the energy differently, independent of the time scale involved.

29 We know of no report of SOAEs in the chicken ear, though they have been reported in other avian species such as the barn owl. Gecko SOAEs were not as prevalent as those reported by Manley et al. (1996). Lastly, for the specific species of frog tested in this study, a previous report observed SOAEs in their ears in spite of our lack of measuring any spontaneous activity.

30 Within the context of laboratory mammals, the lack of observed SOAEs may be due to several explanations. First, they are generally not looked for. Second, surgery is typically performed on mammals for eOAE measurements (such as opening the bulla), which could cause damage to mechanisms in the inner ear giving rise to SOAEs.

31 It is also worthwhile to mention that rabbits appear to completely lack observable SFOAEs. [Fahey, personal communication].
anesthesia effects, as SOAEs tend to be most prevalent in humans (who are awake for measurement sessions), but not in other warm-blooded species where anesthesia is required (to prevent movement). Anesthesia is known to affect efferent feedback, possibly causing effects upon SOAEs (particularly if SOAEs are looked for after a series of evoked measurements). That geckos readily exhibit SOAEs while under anesthesia may reflect that the spontaneous emissions are being generated in a fashion very different from humans or a difference in efferent effects (or neither). Lastly, the observation that evoked emissions were smallest in the chicken relative to the other species may explain why SOAEs were not detected (i.e. they are too small in amplitude to rise above the noise floor). This explanation likely does not account for the frog ear however, where eOAE magnitudes are relatively large.

4.5 Summary

We summarize our main experimental findings here:

- DPOAEs evoked using equal level primaries at 65 dB SPL and a fixed primary ratio were found in human, chicken, gecko (two species) and frog. In all species, emission magnitudes varied with frequency.

- At frequencies below 2 kHz, emissions were largest in the gecko ear and smallest in the human and chicken, particularly for $2f_2 - f_1$.

- For emission frequencies above ~1 kHz, $2f_1 - f_2$ exhibits little (if any) phase accumulation for both human and chicken for a fixed $f_2/f_1$. The gecko and frog however showed significant $2f_1 - f_2$ phase accumulation.

- $2f_2 - f_1$ showed significant phase accumulation in all species, being largest in the humans.

- Phase gradients for $2f_2 - f_1$ were smaller than those of SFOAEs for human, chicken, and gecko emissions. However, the gradients were similar in the frog for the two emission types.
- A similar ratio dependence for \(2f_1 - f_2\) phase behavior was observed in the chicken relative to that observed for humans.

We discussed what the emissions results implied about the underlying generation mechanisms, the role of traveling waves (and scaling symmetry), and how things varied across species. This allowed us to formulated a categorization as to how generators can vary (Fig. 4-8), at least for emission frequencies above \(\sim 1\) kHz. Given that stimulus parameters can directly affect emission properties, caution is needed when making conclusions based upon phase information. Ch.5 explores directly how choice of stimulus intensity affects eOAE phase gradients.
Chapter 5

eOAE Level Dependence

5.1 Background

Evoked otoacoustic emissions (eOAEs) provide the opportunity to better understand the complex interactions among the various constituents of the inner ear. In particular, as stimulus intensity is varied, eOAE magnitudes exhibit a complex dependence on input levels that can range from linear to strongly nonlinear (e.g. non-monotonic). Additionally, stimulus intensity can have significant effects on emission phase. These observations motivate a desire to further study the level dependent properties of eOAEs in order to better characterize the underlying generation mechanisms and their associated nonlinearity. Cast in a different light, the level dependence of eOAEs can serve to reveal how the ear processes lower and higher intensity sounds differently, an important piece of the puzzle within the context of the notion of cochlear amplification. Our present purpose here is to characterize how eOAE growth varies in a number of different species, focusing on emission phase gradients.

EoAE phase gradients refer to the slope of the phase response with respect to emission frequency. These gradients indicate the amount of phase accumulation, or delay, that is occurring. These gradients have revealed significant information about how emissions are generated, providing the basis for the hypothesis that different types of emissions are generated by fundamentally different physical mechanisms [Shera and Guinan, 1999]. Since evoked emission characteristics strongly depend upon the stimuli
being used to evoke them, it stands to reason that emission phase gradients will also depend upon stimulus conditions.

The dependence of emission phase gradients upon stimulus intensity was explored within the context of stimulus frequency emissions (SFOAEs) by Schiarer et al. (2006). In the human, they found two level ranges where the phase gradients take one of two (non-zero) values, gradients being smaller in the higher stimulus level conditions. They interpreted their results in two different ways, each addressing an issue of current debate in the study OAEs. First, they addressed what phase-gradient level dependence may indicate about what the dominant emission source may be (linear reflection-based versus nonlinear distortion-based) at a particular level and how emissions might propagate back from the site(s) of generation to the middle ear\(^1\). Second, the authors addressed the level dependence in light of the proposal that eOAE gradients can provide a non-invasive measure of cochlear tuning [Shera et al., 2002].

We aim to further the study of phase-gradient level dependence here by taking a comparative approach that examines a wide range of different types of ears. We examine eOAEs in four different groups: human (*Homo sapien sapiens*), chicken (*Gallus gallus domesticus*), gecko (two species: *Eublepharis macularius* and *Gekko gecko*), and frog (*Rana pipiens pipiens*). We also examine both SFOAEs and distortion product emissions (DPOAEs) in light of the hypothesis that there are two different generation mechanisms in mammals. Overall, examination in a diverse variety of ears allows us to search for both similarities and differences in emission properties and relate those to what we know about the underlying morphology. Furthermore, the comparative approach allows us to interpret OAEs without using a rigidly bounded anatomical framework. For example, given the lack of a tuned basilar membrane (BM) in the gecko and the absence of a flexible BM in the frog, critical morphological features in OAE generation can be distinguished via a comparative approach. An ultimate goal is to be able to avoid highly specific morphological assumptions that may not have much physical relevance in OAE generation.

\(^1\)Currently, there is debate about whether the dominant mode of backward propagation occurs by energy transferred via a compressional wave in the fluids (fast) or by a basilar membrane traveling wave (quite slower).
We approach eOAE dependence on stimulus intensity in two separate ways. First, we fix the stimuli at a fixed intensity and sweep their frequency, then subsequently vary the stimulus intensity and repeat. This paradigm reveals how emission frequency dependence varies with level. We call this 'eOAE level dependence' and is the focus of this chapter. Second, we fix the frequency of the stimuli and sweep the intensity. We call this approach 'eOAE level growth' and address it in Ch.6. It is important to consider both approaches when examining how eOAEs depend on stimulus intensity, because emission frequency properties can have very different dependencies at various stimulus intensities and vice versa.

5.2 Results - eOAE Level Dependence

Fig. 5-1 shows SFOAE magnitude and phase for frequency sweeps taken at different stimulus intensities in individual human and gecko ears. First, looking at magnitude, growth at a given frequency in both species typically shows areas of nonlinearity and is sometimes even non-monotonic (at certain frequencies). Second, in the case of the human, there are features in the magnitude curves (such as peaks and valleys) which evolve in a level-dependent fashion. For example, it appears that the two peaks apparent at lower levels around 3.4 and 3.7 kHz eventually merge as level increases. Peaks and valleys can also be seen in the gecko ear, but are less numerous (geckos lack the detailed 'structure' seen in human emissions) and appear in a more haphazard fashion with respect to stimulus intensity and frequency. Third, although only tested up to 5 kHz, the higher frequency limit (before falling into the noise floor) increased as stimulus intensity increased in the gecko ear. Lastly, emission magnitudes at higher frequencies in this individual gecko (above 3-3.5 kHz) have a large degree of variability, and it is less clear what is happening above 4 kHz (potentially some system artifacts are present).

Several key observations to be made in Fig. 5-1 stem from the emission phase.

\(^2\text{Compare the 40 dB case for this gecko to that of a different individual shown in Fig. 3-2. It is apparent that the high frequency fall-off (and whether or not a clear phase gradient is still apparent) can vary significantly across individuals.}\)
Figure 5-1: Comparison of SFOAE level dependence in both an individual human and leopard gecko ear (results typical for given species). Legend to the right shows stimulus intensity used for a given frequency sweep. Phase gradients are independent of $L_p$ at lower intensities. Note that in both species at higher frequencies (above ~1-2 kHz), a level-dependent shift occurs in the phase gradients when $L_p \geq 40$ dB SPL. Some phase curves were offset vertically an integral number of cycles for clarity. $[L_a = L_p + 15 \text{ dB}, f_a = f_p + 40 \text{ Hz}]$
Both species (human and gecko) exhibit a significant amount of phase accumulation, independent of stimulus intensity. The total amount of phase accumulation is greater in the human ear. The total amount of accumulation appears largest at lower stimulus levels, and is relatively independent of stimulus intensity in that region. However, there is a shift in the phase behavior in both species around 40 dB SPL. At higher stimulus intensities, there is significantly less phase accumulation. There thus appear to be two regions of phase behavior with respect to frequency, a low-level and high-level one, with a transition between the two occurring at moderate stimulus intensities.

Results across numerous individuals for a given species are shown in Fig. 5-2 (magnitude) and Fig. 5-3 (phase gradients). We plot the phase gradients as $N$-value. $N$ is proportional to the product of the gradient and the emission frequency (Eqn. 3.1), being a physically dimensionless quantity that is the delay expressed as the number of stimulus cycles\(^3\). From Fig. 5-2, it is apparent that SFOAE magnitude growth in all species examined exhibits regions of nonlinearity in a frequency dependent fashion. Relative to the other species, growth is more uniform (across stimulus intensities) in the chicken ear, independent of frequency.

Figure 5-3 indicates that SFOAE phase gradients behave in a level dependent way in the human and gecko ears across individuals. In both species, there appear to be two distinct level ranges above $\sim 1$ kHz where gradients are larger at lower intensities (below 30-40 dB SPL) and smaller at higher intensities. The transition between the two regions appears to be at slightly lower intensities in the gecko ear. Data are more limited in the chicken and frog ear, particularly at lower levels. There is less level dependence in these two species relative to what is observed in human and gecko.

Statistics were performed on the data shown in Fig. 5-2 and Fig. 5-3 by averaging points across frequency into one octave bins. Results are shown in Fig. 5-4 and Fig. 5-5. While some features are lost, these plots serve to reveal how gross trends compare across species\(^4\). At higher frequencies in both human and gecko, the delineation

\(^3\)We use the terms phase gradient and $N$-value interchangeably.

\(^4\)Much of the detail stemming from individuals is also smeared out in Fig. 5-2 and 5-3 when the trend line is computed.
Figure 5-2: SFOAE magnitude level dependence in all four species (both gecko species are plotted together). Note that the legend differs from that used in Fig. 5-1. For any given frequency/stimulus level, each data point come from an individual ear (i.e. no repeats). Only points whose magnitude was at least 10 dB above the noise floor are included. Trend lines are included to guide visualization (see Methods). $[L_s = L_p + 15 \text{ dB}, f_s = f_p + 40 \text{ Hz}]$
Figure 5-3: $N_{SFOAE}$ level dependence in all four species. See caption for Fig. 5-2 for additional information.
Figure 5-4: SFOAE magnitude level dependence statistics stemming from Figs.5-2. For a given stimulus level, data points are collapsed into one octave bins, of which the mean value is then computed (plotted against the geometric mean frequency of all points in that given bin). Error bars indicate 95 confidence intervals (±2 times the standard error). Bin limits (in Hz) are as follows: frogs - [300-600, 600-1200, 1200-2400, 2400-3000], others - [500-1000, 1000-2000, 2000-4000, 4000-5000].

between a low and high-level region becomes more apparent. At lower frequencies (below 1 kHz) in the gecko, the phase gradients do not exhibit the level dependence apparent at higher frequencies. Behavior at lower frequencies is less clear in the human due to the large variability. One further observation is that chicken SFOAE gradients do show a slight level dependent effect in the region of 1-2 kHz.

Two additional observations not shown in the figures are worth noting. First, in one experiment involving a chicken, SFOAEs were observable using a 20 dB SPL probe (typically, the emissions were masked by the noise floor when using low stimulus intensities). In that particular case, it did appear that the phase gradient increased relative to those obtained using higher stimulus intensities. Second, in one leopard gecko, the noise floor was low enough and emissions large enough that an SFOAE
Figure 5-5: $N_{SFOAE}$ level dependence statistics stemming from Figs. 5-3. See caption for Fig. 5-4 for additional information.
evoked using a $L_p = -10$ dB SPL ($L_s = 35$ dB) was measurable. It had a phase gradient similar to that found for emissions evoked using a higher stimulus intensity (0-30 dB SPL). Thus, in geckos, very low intensity stimuli can still evoke significant emissions measurable above the acoustic noise floor.

Fig. 5-6 shows the effect of varying stimulus intensity on the phase gradients of both cubic DPOAEs in the gecko ear (using equal level primaries at a fixed ratio of $f_2/f_1 = 1.22$). DPOAE level dependence data is scarce from the other species, hence only data from geckos is shown here. Statistical analysis of the data in Fig. 5-6 reveals that there are some level dependent effects upon DPOAE phase gradients in the gecko ear, as shown in Fig. 5-7. In particular, $N$ appears larger for $2f_2 - f_1$ at the lowest stimulus intensity (45 dB SPL) when the emission frequency is in the range of 1-3 kHz. However, outside this range, there is not a significant difference in the phase gradients with respect to stimulus intensity. Also note that the $N$-values tend to be larger for $2f_2 - f_1$ than $2f_1 - f_2$ in the gecko ear, independent of stimulus intensity.\footnote{This is an extension of the observation made in Ch.4, which only considered a single stimulus intensity condition.}

The large amount of variability in the DPOAE phase makes it difficult to make any further conclusions with regard to the phase-gradient level dependence.

Also, the top row of Fig. 5-7 shows that the DPOAE rate of growth in the gecko ear depends upon the emission frequency. At lower frequencies (below ~ 1 kHz), growth is expansive at higher stimulus intensities. However at higher frequencies, the growth at higher levels is highly compressive. At lower levels, growth appears to be more linear, independent of stimulus frequency (see Ch.6 for further results on eOAE growth).

5.3 Discussion

5.3.1 Nonlinear Growth

Figure 5-4 indicates that there are varying degrees of nonlinear growth across species. We focus here on low to moderate stimulus intensities. Our stimulus intensity range
Figure 5-6: Level dependence of DPOAE phase gradients for a fixed-ratio sweep in the gecko ear
(both the lower side-band \(2f_1 - f_2\) and upper side-band \(2f_2 - f_1\) are shown). Results from both
gcko species were combined. \([L_1 = L_2, f_2 = f_1 = 1.22\) (fixed)]
Figure 5-7: Statistical analysis of gecko DPOAE level dependence stemming from Fig. 5-6. Both magnitudes (top row) and N-values (bottom row) are included. $2f_1 - f_2$ is shown in left-hand column and $2f_2 - f_1$ on the right. Note the different scales between the two distortions for the N-values. Data are averaged over frequency into one octave bins. Bin limits are as follows: $2f_1 - f_2$ - [380-800, 800-1600, 1600-3200, 3200-4000], $2f_2 - f_1$ - [700-1400, 1400-2800, 2800-5600, 5600-8000]. Only points whose magnitude was at least 10 dB above the noise floor were included in the plots. [$L_1 = L_2, f_2 = f_1 = 1.22$ (fixed)]
represents the region of interest between low and high level responses at the inner ear. A result worth emphasizing is that non-monotonicities are present in SFOAE intensity functions for all individuals in all species examined. The location of these 'dips' vary with frequency across individuals and thus tend to get averaged out. This can have the effect of smearing out important details in averaged plots such as Fig. 5-4. It is thus important to consider individual growth curves such as those shown in Fig. 6-3 when discussing eOAE growth. Furthermore, it will prove insightful to better understand the basis for differences across individuals (for a given species) in terms of eOAE growth properties. For example, what features determine whether growth is monotonic or not at a particular frequency in a given individual?.

Human, gecko, and frog SFOAE growth exhibit regions of both compressive and expansive nonlinear growth. In contrast, SFOAE growth in the chicken ear did not exhibit expansive growth and generally appeared more linear. It is not clear why growth has a more uniform rate in the chicken, though emissions are the smallest in the chicken ear. Differences aside, Fig. 5-4 clearly indicates that nonlinear emission growth is a common feature across species. The basis for this nonlinear behavior is addressed more thoroughly in Ch.6.

SFOAE level growth rates also varied with stimulus frequency in different ways across species. In the human at lower frequencies, growth is compressive only at lower intensities and more linear at higher ones. However, at higher frequencies, there are two distinct areas of compressive growth, consistent with the individual human curve shown in Fig. 6-3 (which exhibits an extended region of compressed growth at lower levels and a saturated response at higher levels). Growth rates in the chicken ear appear relatively independent of frequency. Gecko SFOAE growth rates at lower intensities are typically independent of frequency, except at the highest frequencies tested where growth is highly compressed. Growth at higher levels though is more rapid in the gecko ear at lower frequencies relative to higher ones (below 2 kHz). Growth in the frog is overall compressive, except around ~1-2 kHz when

---

6Compressive means a slope of less than one (on a log-log scale) for the I/O function and expansive means a slope greater than one.
using higher intensity stimuli where growth becomes expansive\textsuperscript{7}. In all species except the frog, emission magnitudes are generally largest at the lowest frequencies for a given stimulus intensity, this trend being most readily apparent in the gecko. To summarize, the nonlinear growth of SFOAEs clearly exhibits a frequency dependence in all species, in spite of significant morphological differences.

DPOAE growth for the gecko ear is shown in the top row of Fig. 5-7. These growth curves are also highly nonlinear, appearing expansive at lower frequencies and compressive at higher frequencies (particularly for higher intensity stimuli). In general, SFOAE and DPOAE growth curves were qualitatively similar in the gecko ear. Both SFOAEs and DPOAEs have different growth properties at higher stimulus intensities depending upon whether the emission frequency is above or below 1-2 kHz. This observation suggests that different parts of the papilla are generating emissions in different ways.

It should be noted that in mammalian DPOAE magnitudes, there is a complex dependence upon the relative difference between $L_1$ and $L_2$ (as well as the primary ratio) [Johnson et al. 2006], which may also have an effect on the phase gradients. While this was not systematically examined in this study, preliminary observations (not shown) indicate that gecko and frog do not exhibit a similar dependence on the relative primary levels and DPOAEs tended to be largest when $L_1 = L_2$. Better characterizing this difference between mammals and non-mammals will help elucidate how generation mechanisms differ between species (and what effects phenomena such as two-tone suppression are having within an anatomy-specific context).

A result worth emphasizing here is the prevalence of SFOAEs in the gecko using stimulus intensities as low as 10 dB SPL (and even as low as −10 dB SPL in one individual). This observation is significant in the sense that it implies that the gecko ear is able to generate and/or emit eOAEs in a highly \textit{efficient} manner (relative to

\textsuperscript{7}Our SFOAE magnitude results were consistent with frog SFOAE growth functions reported by Meenderink and Narins (2006). They tested towards higher stimulus intensities and found that expansive growth did not start until $L_p$ got above 60-70 dB SPL, independent of frequency. While not apparent in Fig. 5-4, our frog measurements extended towards lower levels and indicated growth was fairly linear for stimulus intensities below 40 dB SPL (see Fig. 6-3). There were however differences between their study and our in terms of the phase gradients (as discussed in Sec.5.3.2).
the stimulus), even at very low sound intensities where the incident acoustic energy is likely to be comparable to the thermal energy present in the ear [Rosowski, 1991]. It should be pointed out that geckos have significantly fewer hair cells (~1000-2000) relative to humans (~14000). It is likely to be important to consider the effective hair cell density in a given frequency region for this comparison to be of value. While the gecko basilar papilla (BP) does exhibit an exponential tonotopic mapping [Manley et al. 1999], the physical area of generation region may be larger (and thus have more contributing hair cells) relative to the generation region in the human cochlea.

### 5.3.2 Level Dependence of eOAE Phase Gradients

In both the human and gecko, there are two separate level regions where the SFOAE phase gradients behave in different ways. In the low-level region (typically below ~ 40 dB SPL), phase gradients tend to be the largest and relatively independent of stimulus intensity. At moderate stimulus intensities (~ 40 – 50 dB SPL) and frequencies above ~ 1.5 kHz, there appears to be a shift, or transition, to a high-level region, where phase gradients are smaller (but still significant) and are not necessarily independent of level.

A *transition* implies an abrupt change. As shown in Fig. 5-1 for an individual human and gecko, the shift occurs within a 10 dB change in stimulus intensity. Furthermore, the notion that the shift between the two regions occurs relatively abruptly is consistent with the observations of Schrairer et al. (2006). The transition may be analogous to the observation in human DPOAEs that there is an abrupt shift in the frequency dependence of the $2f_1 - f_2$ phase for a fixed primary ratio paradigm around $f_2/f_1 \sim 1.09-1.11$ [Knight and Kemp, 2000]. This observation in DPOAEs has been attributed to a change in the dominant source generating the emission (*i.e.* a transition between a place-fixed component being the largest to that of a wave-fixed component).

---

8Seeing that lizards are cold-blooded, the noise due to thermal energy of the inner ear fluids would be lower than that of mammals. This trade-off however would most likely come at a cost, since any metabolic activity in the *cooler* ear that might be present to boost detection would effectively be reduced.
It is important to point out that there are two major factors that affect the delineation between the low and high level regions. First, there can be considerable differences across individuals for a given species. For example, while the shift from one region to the other is typically clear in a given individual (as in Fig. 5-1), the actual transition intensity can vary across subjects (e.g. occurring between 30-40 dB SPL for one and 40-50 dB for another). This has the effect of giving the appearance of an intermediate transition region when averaging that may potentially be misleading. Second, the transition point clearly varies with respect to frequency (another factor that can vary across individuals). In both human and gecko, the delineation between the two regions shifts towards lower levels as frequency increases (this can be seen in Fig. 5-5).

The level dependence of eOAE phase gradients has been examined in previous studies. The results by Schiarer et al. (2006) are similar to those reported here for human SFOAEs. Some differences were apparent, as their measured phase gradients tended to be larger at lower frequencies than those measured here. Further aspects relative to their particular study are discussed below. Whitehead et al. (1996) also examined delays using both phase gradients and onset times in the time domain for human DPOAEs. Though not the primary focus of their report, they found that delays systematically decreased with respect to stimulus intensity. Long and Talmadge (2006) looked at level dependence of DPOAE phase gradients in the human ear using a (large) fixed primary ratio for the frequency sweeps. They found that there were significant differences across individuals, where some individuals had gradients insensitive to stimulus intensity and others clearly exhibited larger gradients at lower levels. In that study, they focused solely on $2f_1 - f_2$ (which has a gradient close to

---

9Whitehead et al. (1996) used a paradigm where one primary was at a fixed frequency and the other swept, in contrast to a fixed $f_2/f_1$ paradigm used here. The delays they saw between the two methods (phase gradient versus onset time) were similar in spite of the difficulties associated with determining the emission onset in the time domain and differences stemming from which primary tone was swept (phase gradients were smaller in the case of fixed $f_2$). However, there were some discrepancies which they discuss in their paper as likely stemming from interference effects from multiple sources. A similar overall paradigm was also employed by Stover et al. (1996), who also performed an unmixing paradigm to examine DPOAE components with different latencies.

10They used a range of $L_2 = 25 - 75$ dB SPL where $L_1 = 39 + 0.4L_2$ dB for $L_2 < 65$ and equal level primaries for $L_2$ at or above 65 dB SPL.
zero at higher frequencies for a fixed $f_2/f_1$), which has been shown in the human ear to be comprised of two different components, each with different latencies [Kalluri and Shera, 2001]. The results of Long and Talmadge imply that the relative contribution of these two DPOAE components can vary in a level-dependent manner, with large variations across individuals.

A study by Meenderink and Narins (2006) examined SFOAE phase gradients in the frog and concluded that the gradients were relatively insensitive to stimulus intensity, at least in the range of 60-80 dB SPL. This conclusion may be somewhat questionable in two regards. First, they did not statistically show there is not a level-dependent difference, as visual inspection of their data might suggest. Second, there was some discrepancy between our phase gradients and those measured in their study in that our gradientss tended to be larger for frequencies below 1 kHz. An obvious difference between the two studies was that our highest stimulus intensity tested was 60 dB SPL, while that value was the lowest used by Meenderink and Narins. However, even at $L_p = 60$ dB SPL, the phase gradients we measured were larger, by up to a factor of 1.5 for frequencies between 0.6-1 kHz. At frequencies below 0.6 kHz, the gap widens even further as they reported a decrease in the gradients with frequency, a feature we did not observe\textsuperscript{11}. The discrepancy between the two studies, given the differences in the stimulus intensity range employed, may indicate that there is indeed a level-dependent transition in the frog SFOAE phase gradients, but that our stimulus range did not extend high enough to detect it.

It is curious that this level-dependent transition in the phase gradients does not appear to be present in the chicken or frog over the stimulus range tested\textsuperscript{12}. Given other similarities in emission properties between human and chicken, which suggest

\begin{footnote}{\textsuperscript{11}At low frequencies, Meenderink and Narins (2006) found that the phase gradients decreased for SFOAEs, but found the converse for DPOAEs, where DPOAE phase gradients (particularly for $2f_2-f_1$) increased for frequencies below 1 kHz [Meenderink and van Dijk, 2004]. The difference between the two emission types is interesting, though it may partly stem from the fact that Meenderink and van Dijk averaged their phase slopes across ratios, potentially smearing out frequency-dependent details.}

\begin{footnote}{\textsuperscript{12}For the chicken results shown in Fig. 5-5, it does appear that there is a slight delineation between a low and high-level region in the 1-2 kHz range. However, this is much smaller than observed for the human and gecko and does not further extend towards higher frequencies as observed in the other species).}

121
similar underlying generation mechanisms in the two species [see Ch.4], we might expect to see a similar qualitative behavior in the phase gradients of both species. One possible explanation for the discrepancy might be tied to the fact that emission magnitudes are smallest in the chicken. It is possible that since emissions evoked using lower intensities were undetectable, the acoustic noise floor is effectively masking the low-level region where the emissions with larger gradients would be found. This is consistent with a single observation as described in Sec.5.2. Conversely, our measurements may have been entirely in the low-level region and the stimulus intensities did not extend high enough to see the transition. If the level dependence is indeed absent in the frog, it may be less surprising given the gross morphological differences (such as a lack of a BM) and notion that each of the two papilla contribute to the overall emission in a relatively complex fashion.\textsuperscript{13}

With consideration to the comparative approach taken here, we come back to the specific points regarding human SFOAEs made by the Schiärer et al. (2006) study raised in the introduction. Relative to the human, geckos also exhibited a qualitatively similar level dependence in their SFOAEs. The similarity is striking in light a lack of BM traveling waves in the gecko. This observation suggests complex BM interactions are not necessary to account for eOAE level dependent properties. Similarity between the two species is consistent with the point raised by Schiärer et al. that the phase gradient level transition is not due to the dominant source shifting from one based upon reflection to nonlinear distortion. If this were the case for the human ear, we would expect the high-level region to exhibit gradients much closer to zero due to the scaling-symmetric nature of the traveling wave in the cochlea. This is furthered by an overall lack of evidence for a scaling-symmetric effect in eOAE generation in the gecko ear [see Ch.4].

While the shift between the two level regions may be due to a change in the mechanism for backward propagation, the similarity between human and gecko argues that it is not due to a change in the mode of motion of the BM given the lack of BM

\textsuperscript{13}The notion that the two papilla in the frog contribute to the emissions in a complex fashion was described by Meenderink and Narins on a 2007 poster at ARO. Unfortunately due to scheduling issues, their abstract never made it into the final version of the conference proceedings.
traveling waves in the gecko ear. However, differences need to be considered. The shift in the phase gradients between the two regions in the human ear was close to a factor of two, while it tended to be larger in the gecko ear (and not necessarily constant at higher and higher levels). It may thus be difficult here to reach any definitive conclusion with regards to the means of backward propagation from the generation site. Changes in tuning are addressed in the next section.

In light of the similarity in SFOAE phase gradient level dependence between humans and geckos, the question arises whether the mechanism responsible is similar between the two species given the gross morphological differences. Additionally, it is desirable to further explore whether or not chickens and frogs truly lack the phase gradient level dependence by using a wider range of stimulus conditions. As indicated above, there is evidence hinting that the stimulus paradigms used in the present study were insufficient to clearly reveal the transition in chicken and frog. Furthermore, the choice of stimulus parameters such as the suppressor level or frequency (relative to the probe) may be crucial factors. Given the lack of similarity in level dependence across all species for our particular choice of stimulus parameters, it will be important to ascertain whether or not phase gradient level dependence is a feature unique only to certain species and not general to all ears. If a level dependence is absent in other species, the similarity between human and gecko may be misleading in the sense that the underlying cause for the level dependence might be very different between the two. Several possible explanations for the level-dependent transition in the phase gradients are discussed in the next section.

5.3.3 Level-Dependent Generation Mechanisms

As indicated by Schiarer et al. (2006), upon transition to the high level region where the phase gradient is smaller, the emissions still exhibit a significant amount of phase accumulation. This observation is in contrast to what one might expect if the emission were to become dominated by a nonlinear distortion source that behaved in a scaling-symmetric fashion. The mechanisms underlying $2f_1 - f_2$ and SFOAE generation (at any level for the latter) therefore must be different in the human ear. By mechanism,
we simply mean a physical process that is part of the emission generation (as defined in Ch.4.2). The case is not so clear for the gecko ear, which generally appears to lack the frequency-independent $2f_1 - f_2$ phase seen in humans. Though gecko DPOAEs lack a clear level dependent transition in the phase gradients (Fig. 5-6) like that observed for SFOAEs, this discrepancy may result from inherent differences in the measurement paradigms. SFOAEs are measured at the stimulus frequency while DPOAEs are measured at frequencies different than those of the stimuli and are typically of much smaller relative intensity. These inherent differences in stimulus paradigms indicate that caution is warranted when directly comparing SFOAEs to DPOAEs, regardless of similarities between the phase gradients of the two emission types.

Another observation worth noting is that by comparing Fig. 5-4 and Fig. 5-5, both human and gecko magnitude growth tends to be highly compressive in the intensity region where the transition in the phase gradients occurs. The correlation between magnitude and phase here suggests that there may be some interference (destructively) between two level-dependent generation mechanisms when the stimulus intensity is near the transition point. It should also be emphasized that the SFOAE phase gradient level dependence is only observed at higher emission frequencies in both human and gecko, above $\sim$1.5 kHz. This observation indicates that the processes underlying the emission generation are likely different between lower and higher frequencies.

It has been proposed that SFOAE phase gradients provide a non-invasive means to estimate tuning of the inner ear [Shera et al., 2002]14. Our data clearly indicate though that in certain species, these gradients are level dependent and thus caution is needed when making comparisons between OAEs and other measures of tuning such as $Q$-values. Tuning is typically measured near threshold, where it is sharpest. Regardless of the measurement method (ANF, BM motion, psychophysical, etc.), it is known that sharpness of tuning broadens as stimulus intensity increases15. It thus stands to reason that the phase gradients that would be comparable to direct

---

14 This is discussed more thoroughly in Ch.3.
15 A reference is needed here....

124
measures of tuning near threshold should be evoked using lower intensity stimuli.

As shown in Fig. 3-8 for the gecko ear, choice of stimulus level clearly has an effect. That figure shows $N$-values plotted for both $L_p = 20$ dB SPL (solid red line) and $40$ dB (dashed red line) together with ANF-derived $Q$-values. Two differences are obvious. First, there is a vertical offset depending upon stimulus intensity, consistent with Fig. 5-3. Second, there is a frequency dependent effect in the lower intensity condition, such that the the frequency dependence of $N$ differs significantly from that of $Q$ (or the higher intensity $N$-values) below $\sim1$ kHz.

So a new question is raised: Does the level-dependent shift in eOAE phase gradients correlate to the change in the sharpness of tuning as intensity is increased? Furthermore, regardless of the answer to the previous question, what is the mechanism responsible for a decrease in the sharpness of tuning for higher intensity stimuli? For a linear system (which the ear clearly is not), $Q$ would be independent of stimulus amplitude [French, 1971]. So what nonlinear features in the ear give rise to this broadening? While this study did not directly examine the level dependence of $Q$-values, the similarity between the decrease in SFOAE delays at higher levels and the broadening of tuning at higher levels (broader tuning means shorter rise times in a second order resonator) suggest the two are correlated. Extending this one step further, the similarity between human and gecko in terms of their SFOAE properties suggests that decrease in the sharpness of tuning does not depend strongly upon highly specific morphological arguments.

Several possible physical mechanisms that could explain the level dependence of eOAE phase gradients and potentially the broadening of tuning are discussed here. First, if there is some amplification mechanism present that effectively boosts the response to low intensity stimuli, losing its effect at moderate to higher levels could account for the observed effects. For example, the active mechanism could take the form of a negative damping. If the active contribution was effectively swamped out at higher levels (such that the negative damping no longer provided a significant contribution to the response), the (nonlinear) shift in the effective damping term could account for the decreased $Q$ and thus the shorter phase delays. Second, it is
possible that efferent effects could have some influence. Efferent effects in this regard however are unlikely because the transition between the low and high levels regions occurs at sound intensities lower than where efferent responses likely start to play a significant role. Another point arguing against efferent effects is that only the low frequency region of the gecko BP receives any efferent innervation [Manley, 1990], yet we clearly see level dependent effects at higher frequencies in their emissions. Lastly, there is the possibility of reciprocal neural networks that act locally solely within the inner ear [Thiers et al., 2006]. These neural paths, being neither afferent or efferent, may exist to quickly affect the mechanical responses in the inner ear, similar to a role the efferents might play. Since their existence has only been shown anatomically in mammals, it is not clear if similar networks exists in other species and what specific role(s) these connections may be playing. In summary, it may not necessarily be one or the other of these possible mechanisms described above that are playing a significant role, but several in conjunction with one another.

Further study comparing eOAE phase gradients and measures of tuning sharpness will help resolve some of the questions so far raised. In particular, careful comparison between the intensity dependence of SFOAE phase gradients and sharpness of tuning (as measured either via ANF responses or psychophysically) will serve to illuminate the connection between the mechanisms giving rise to OAE generation and sharp frequency selectivity in the ear.

5.4 Summary

Summarizing the main experimental findings presented in this paper:

- Humans, chickens, geckos, and frogs all exhibit varying degrees of nonlinear growth in SFOAE magnitude, with regions of both compressive and expansive growth in SFOAE magnitude, with regions of both compressive and expansive

---

16Efferent effects have been demonstrated to have a reduced effect in highly anesthetized mammals [Liberman (1989); Harel et al., 1997]. Seeing that humans were awake while geckos were anesthetized, yet both exhibited similar effects, argues against an efferent effect.

17Careful consideration needs to be put into the psychophysical paradigms used to estimate tuning. For example, forward versus simultaneous masking can yield two very different estimates [Oxenham and Shera, 2003].

126
growth.

- Growth behavior varied with respect to frequency in a species-dependent fashion. For example, SFOAE growth at higher intensities in the gecko ear was expansive at lower frequencies (below 2 kHz) and compressive at higher frequencies while growth rate was relatively independent of frequency in the chicken.

- SFOAE phase gradients show a stimulus intensity-dependent transition from long to short delays with increasing level in both human and gecko ears at frequencies above ~1 kHz.

- No SFOAE phase gradient level dependence was observed in chicken or frog ears.

- Emission growth tends to be most compressive in the region where the level-dependent transition of the phase gradients occurred (for human and gecko).

- DPOAE phase gradients in the gecko ear did not exhibit the degree of level dependence observed for SFOAEs.

Evoked emissions depend upon the stimulus intensity used to evoke them in a complex fashion, exhibiting both linear and nonlinear behavior. One particular result emphasized in this report was that SFOAE phase gradients exhibit a level-dependent transition, indicating some sort of nonlinear shift in the underlying generation mechanisms. In a subsequent chapter [Ch.6], we explore the notion of whether this nonlinear effect derives from a single nonlinear mechanism or a level-dependent transition between two mechanisms. Better characterizing and understanding the basis for these level-dependencies will serve to further our understanding of how the ear processes low and high intensity sounds differently. Additionally, a comparative approach that systematically explores OAEs in a number of different types of ears holds the promise to reveal general principles underlying the level dependence of emission generation.
Chapter 6

Evidence For Multiple
Level-Dependent eOAE Sources?

6.1 Nonlinear and Non-Monotonic eOAE Growth

Level growth functions for evoked otoacoustic emissions (eOAEs) indicate changes in emission magnitude and phase as a function of the intensity of the evoking stimulus. Level functions have been utilized in a wide array of contexts, for example to examine temporary threshold shifts in humans or to look at post-acoustic trauma recovery in bird ears where the inner ear regenerates itself. eOAE level functions typically show a high degree of nonlinearity as well as non-monotonicities for some stimulus parameters. The goal of this chapter is two-fold. First, we characterize similarities and differences in eOAE growth in an array of species exhibiting different morphologies. Second, we examine whether e.i.e.OAE growth is consistent with a proposed model that explains non-monotonicities as arising from a single nonlinear point-source. By exploring these two paths, we aim to better assess how the ear might process low and high intensity sounds differently.

We begin by reviewing previous efforts that examined eOAE level functions\[1\].

---

\[1\] A partial listing of these studies include (categorized by the species relevant to the present study): Human - Brown and Gaskill 1990 (DPOAE), Abdala 2000 (DPOAE), Schraier et al. 2003
Numerous studies have looked at emission growth in an array of species (both mammalian and non-mammalian), all of which indicated nonlinear behavior present to some extent. A feature observed in many of these studies was the presence of non-monotonic growth\(^2\). In these cases, notches were present that indicated there were regions where the emission magnitude would actually decrease with increases in the stimulus intensity. At yet higher level, emission magnitude started to increase again and eventually saturate. In studies where emission phase was reported, notches in the magnitude also exhibited a corresponding jump in the phase. Based upon previous studies alone, it is apparent that non-monotonic eOAE growth is not unique to a particular species or anatomy.

The basis for nonlinear growth in eOAEs is not currently well understood. It has been argued that the presence of non-monotonicities suggests there must be at least two level-dependent sources that are out of phase with respect to one another and interfere destructively [Brown, 1987]. The dominant source at lower stimulus intensities is more physiologically vulnerable and it has been suggesting that this low-level source derives from an active process present in the inner ear [Liberman et al., 2004]. Emissions evoked using higher level stimuli (above \(\sim 60-70\) dB SPL), being less physiologically vulnerable, are thought to arise from passive nonlinear sources in the ear. Unmixing studies support the presence of multiple sources comprising a DPOAE [Stover et al., 1996; Kalluri and Shera, 2001], but it is not clear what exactly these different sources are nor how they may interact in a level-dependent manner. An alternative point of view was raised by Weiss and Leong (1985), who suggested that a single nonlinear point-source is sufficient to explain the presence of non-monotonic behavior in the ear. The notion of a single-source model was further extended and proposed to explain DPOAE growth functions [Lukashkin and Russell, (SFOAE); Bird - Kettembeil et al. 1995 (DPOAE), Ipakchi et al. 2005 (DPOAE), Lichtenhan et al. 2005 (DPOAE); Lizard - Rosowski et al. 1984 (DPOAE), Manley and Koppl 1993 (DPOAE); Frog - Meenderink 2005 (DPOAE), Vassilakis et al. 2004 (DPOAE), Meenderink and Narins 2006 (SFOAE). Furthermore, studies have examined eOAE growth in non-vertebrate such as moths and grasshoppers [Kossl and Boyan, 1998; Coro and Kossl, 2001].

\(^2\)Lack of non-monotonic behavior in reports that only presented data averaged across individuals is not surprising, as the frequencies at which notches occur vary across individuals and averaging can thus smear out features such as notches.
An attractive feature of a single-source approach is its simplicity: it does not rely upon complex mechanical or physiological assumptions.

This chapter systematically compares both SFOAE and DPOAE level functions across species. Four advantages of this approach are as follows. First, comparison across an array of species allows us to clearly delineate features of eOAE growth that are specific to a given species and correlate them back to what we know about the anatomy and physiology. Certain species lack features such as a flexible basilar membrane (BM) and/or hair cell somatic motility (two features generally considered important in mammalian eOAE generation), providing a simpler anatomy to facilitate interpretation. Second, we use the same measurement system and stimulus paradigms in all species, thereby reducing the possibility of methodological differences that may confound interpretation. Third, our system allows us to measure emission phase and the associated gradients (slope of the phase-frequency curve), providing another quantification of emission characteristics. Lastly, it has been proposed that low-level SFOAEs and the "generator" component of DPOAEs arise from two fundamentally different mechanisms in the mammalian ear [Shera and Guinan, 1999]. By characterizing the growth properties for both types of emissions across species, we stand to better understand how these two mechanisms differ.

Ultimately, better understanding of emission growth holds the promise of providing us with a better understanding of whether the ear processes lower and higher intensity sounds differently, elucidating the mechanisms which provide the ear with a large dynamic range. We start by first fleshing out the foundation of the single-source model [Lukashkin et al., 2002]. By understanding the model's underlying assumptions, predictions, and limitations, we will have a basis for addressing whether or not the notion of a single point-source is sufficient to describe eOAE growth.
6.2 Single-Source Model

6.2.1 Overview

The essence of the single-source model was first described by Weiss and Leong (1985), who were modeling the effects of transduction currents upon the hair cell membrane in the alligator lizard ear. Having observed non-monotonic behavior in some of their measurements at intermodulation distortion frequencies, they state “Thus the observation of such minima in level functions and the accompanying change in phase angle is not sufficient evidence for inferring the existence of two mechanisms that produce signals that are in phase opposition.” They point out that a point-source nonlinearity is capable of predicting magnitude notches in level growth functions as well as phase jumps of 1/2 cycle. Using a sigmoidal nonlinearity (derived from the saturating properties of the mechano-electro transduction apparatus of the hair cell), the model is highly generalizable and can be applied in a wide variety of contexts. Relative to other possible modeling approaches, the strength of the model lies in its simplicity: a single nonlinear source is capable of describing some aspects of the behavior of a complex physiological system.

We attempt here to briefly provide some intuition for how the single-source model gives rise to notches and phase jumps. There are a number of different approaches that one could take, such as a graphical description. Our specific approach taken here stems from a series expansion of the nonlinear function describing the source (such as a Taylor, Volterra, or Bernstein expansion of the function into an infinite series). The expansion provides a means of expressing a function by a series of higher order (i.e. nonlinear) polynomials, each polynomial constituting a different term. The contribution of each n’th order term is tempered by the local derivative of the nonlinear function as well as a reciprocal factorial of order n. Basically, nonmonotonicities occur when higher order terms of the expansion interfere with one another.

Upon series expansion for an input of two sinusoids (or primaries, with frequencies
$f_1$ and $f_2$, amplitudes $A_1$ and $A_2$ and operating point $x_o$)\(^3\) to a (arbitrary) nonlinear function, harmonic and intermodulation distortions become apparent via the presence of the higher order terms\(^4\). Collecting all the terms for a particular distortion product (such as $2f_1 - f_2$), each of the higher order terms will grow nonlinearly with respect to the amplitudes of the sinusoids [Engebretson and Eldredge, 1968]. As $A_1$ and $A_2$ increase, some (higher order) terms of the expansion grow faster relative to the others. Depending upon the nature of the nonlinearity (i.e. the local derivatives about the operating point), each higher order term will have either a $+$ or $-$ sign (or equivalently in terms of a sinusoids, a phase difference of $\pi$ radians or 180°). When one term grows faster than the dominant term and becomes comparable in magnitude, a difference in sign between the two terms can cause a cancellation of the two. The result is a notch. As $A_1$ and $A_2$ increase further, the dominant term shifts to the previously subordinate one. The phase of the distortion product is determined by the dominant term, so a notch will be associated with a phase jump. The given distortion will grow with respect to the dominant term’s dependence upon $A_1$ and $A_2$ (away from regions of interaction). For example, when the cubic term is dominant well below a notch, the $2f_1 - f_2$ distortion will grow in amplitude as $A_1^2A_2$.

The specific nature of the nonlinearity is also important due to the presence of the function’s derivatives in the series expansion. When the function is sigmoidal, growth for larger and larger $A_1$ and $A_2$ values eventually saturates, regardless of the order of the dominant term. There is still some debate about the necessity of asymmetry in the nonlinearity about the operating point in order to get cancellation in the single-source model [Fahey et al., 2000; Kossi and Coro, 2006].

The single-source model explains the non-monotonic DPOAE growth in terms of a single nonlinear point-source [Lukashkin et al., 2002] and is schematized in Fig. 6-1. Within the context of the model, the eOAE point-source is considered to be a single

---

\(^3\)The operating point is basically the zero set point about which the sinusoids oscillate.

\(^4\)By ‘input of two sinusoids to the the nonlinear function’, we refer to the mapping of the input (two sinusoids) via the nonlinear function (the nonlinearity) to an output (the distorted sinusoids). For example, suppose our nonlinear function was $y = f(x) = x^3$ and the input is a single sinusoid $x(t) = \sin \omega t$, then the output would be $y(t) = \sin 3\omega t$. In this case, the nonlinearity produces cubic harmonic distortions, with energy at both the frequencies $\omega$ and $3\omega$. 

133
hair cell. The general form of the nonlinearity stems from the hair cell transduction current - bundle displacement relation and has the general form [Crawford et al., 1989]

\[
y(x) = \frac{A}{[1 + e^{b(x+c)}] \cdot [1 + e^{d(x+e)}]}
\]  

(6.1)

where \( A, b, c, d, \) and \( e \) are all constants. Lukashkin et al. add in a positive feedback loop in order to account for the effect of the cochlear amplifier. Aside from the observation of notches/phase jumps in DPOAE growth functions, the authors validate the single-source model with observations of horizontal shifts in notch location after injection of furosemide. This drug is known to lower the endocochlear potential [Sewell, 1984], which the authors model as a reduction in the gain of the cochlear amplifier. They further assume that if two sources were present and interfering to give rise to the notch, the lower level source would stem from an active mechanism while the higher level source would be passive. Thus, the effect of the furosemide would be to shift the notch to the left since the active source would be contributing to a lesser degree. But they report observing the opposite, that the shift in the notch is towards the right. This observation is in agreement with their model predictions based upon attenuation of the gain introduced by a positive feedback loop. They argue that the agreement between their model and the DPOAE data with respect to furosemide vulnerability validates the single-source model and disproves the multiple-source model.

Although the single-source model has only been proposed to explain non-monotonic DPOAE growth, it can be extended towards SFOAEs, which also empirically been observed to exhibit notches in their growth functions [Schairer et al., 2003]. While growth of the primaries is explicitly linear in the single-source model (for the input of two sinusoids), suppression paradigms that isolate the deviations from linearity associated with SFOAEs [see Ch.2] demonstrate complex growth that can be non-monotonic.

What are the key assumptions associated with the single-source model for DPOAE growth? First, whereas Lukashkin et al. (2002) assume the presence of a gain mech-
Figure 6-1: Overview of the single-source model. The underlying nonlinearity of the model derives from the transduction current (ionic flow into hair cell through bundle channels), which has a sigmoidal dependence upon bundle displacement and can be described either with a hyperbolic tangent or exponential-based function [Crawford et al., 1989]. The mapping of two sinusoids through this function generates distortion. As the amplitudes of the sinusoids are increased, higher order terms start to have an appreciable effect and can cause destructive interference. The result is a notch and phase jump. Higher order distortions can exhibit multiple notches and phase jumps.

anism, no such mechanism is required to produce non-monotonic growth. A purely passive, single nonlinear source can give rise to the notches and phase jumps [Weiss and Leong, 1985]. So amplification is not required for the single-source model to work. Second, the single-source model makes no specific anatomical or physiological assumptions beyond that of a sigmoidal-type nonlinearity. Therefore, the model could apply to a wide range of ears, particularly those with sterociliary based sensory cells where mechanoelectro transduction is known to saturate for both positive and negative bundle deflections. Lastly, the model is frequency independent (due to its instantaneous nature). It predicts that non-monotonic growth occurs at all frequencies. Furthermore, delay is a feature not captured by the model.

6.2.2 Predictions

We provide several specific predictions of the Lukashkin et al. model. First, growth of the distortion at lower levels will go as the order of the distortion. For example, $2f_1 - f_2$ and $2f_2 - f_1$ will grow as the cube $A_1^2 A_2$ (assuming both $A_1$ and $A_2$ are increased together) at levels below the notch and saturation regions. Second, non-monotonic growth occurs at all frequencies. Third, for a sigmoidal nonlinearity, the model predicts the distortion product to take one of two values for its phase, separated
by 180°. This occurs because the dominant term (of the expansion) is either positive or negative. Thus, the model predicts phase jumps of 1/2 cycle and a constant phase between transitions. Fourth, the model predicts growth patterns resembling that shown in Fig. 6-1, where compressive growth would occur only in the region immediately preceding a notch and upon moving into the saturation region. Thus, there are only a single limited region below saturation that shows a compressed growth rate (i.e. 0 to 1 dB/dB). Fifth, the model predicts non-monotonic SFOAE growth when measured using a 2TS paradigm (see Ch.2), but that the growth should be quadratic (with respect to probe intensity) at lower levels. Sixth, the model can not produce level-dependent delays.

To summarize, the single-source model predicts notches and phase jumps that occur in eOAE level functions as deriving from a nonlinear point source. This is in contrast to the notion of two (or more) spatially or mechanistically distinct level-dependent sources that are out of phase relative to each other. The distinction here is important because each provides very different interpretations of how the eOAEs are being generated.

6.3 Methods

6.3.1 OAE Measurements

Both DPOAE and SFOAE growth is examined in four different groups: human (Homo sapien sapiens), chicken (Gallus gallus domesticus), gecko (two species: Eublepharis macularius and Gekko gecko), and frog (Rana pipiens pipiens). See Ch.3 for a description of anatomical and physiological differences across these species as well as motivation for using two separate gecko species.

Measurement techniques were similar to those employed in previous chapters. See Ch.2 for a more complete description. A suppression paradigm (two tone suppression, or 2TS) is employed for measuring SFOAEs. This method looks at the complex difference (at the probe frequency, \( f_p \), with level \( L_p \)) across two conditions:
the probe presented alone (where presumably the emission is present) and the probe
and suppressor (at frequency $f_s$ and level $L_s$) presented together (where the emission
is significantly reduced).

Stimulus intensities ($L_p$ and $L_s$ for SFOAEs, $L_1$ and $L_2$ for DPOAEs) were always
varied together. One important difference from previous chapters was that the stimu-
lus intensities were presented in randomized order to prevent possible adaptive effects
(though all averages for a given intensity were computed at once, typically lasting $\sim 17$
$s$). Error bars indicate standard error of the mean over the averages taken at a given
stimulus intensity (35 averages/level). Measurements at certain stimulus intensities
were repeated over the course of a run (typically 5-10 minutes apart) to examine
variability in emissions over the time scale of minutes and possible adaptation effects.

6.3.2 Analysis

Caution is needed here in one important regard: averaging. Emission properties
can vary significantly from ear to ear. So averaging many different curves together
can result in the loss of important details. For example, given a fixed set of stimulus
conditions across ears, some ears may exhibit curves that are highly linear while others
have deep notches. Taking the average of all these curves will effectively smear out
these details. For this reason, the results presented here mostly stem from individual
ears that are representative for a given species.

Figure 6-2 shows the method used to determine the size of the phase jump when
a notch was present. A hyperbolic tangent function of the form

$$y(x) = A \cdot \tanh[b \cdot (x + c)] + d$$

was used to fit to the phase curves (which were unwrapped in order to remove any
phase discontinuities associated with phase curves that varied smoothly with respect
to stimulus level). There are four free parameters: $A$, $b$, $c$, and $d$ ($y$ represents the
emission phase and $x$ is the stimulus level). An exhaustive search method was used
over a specified grid to find the best fit parameters by minimizing $\chi^2$ (indicated by the
Figure 6-2: Example of SFOAE phase dependence with respect to probe level (this is from a leopard gecko ear). Green curve represents the best fit of a hyperbolic tangent function used to estimate the size of the phase jump. $[L_s = L_p + 15 \text{ dB}, f_p = 1000 \text{ Hz and } f_s = f_p + 40 \text{ Hz}]$

green curve in Fig. 6-2) [Bevington and Robinson, 1992]. The uncertainty associated with each phase value (which appears in the denominator of each term appearing in the sum for $\chi^2$) were excluded in order to obtain better fits to the actual phase jump (the small uncertainty for high-level phase values typically had an adverse effect on the fit parameters). When all points were equally weighted (i.e. points far away from the jump), the fits tended to underestimate the size of the jump due to the phase varying smoothly with level at levels above and below the jump. For this reason, the fit procedure was reiterated to recompute $A$, $b$, and $d$ using a Gaussian weighting function centered about $c$ (now fixed) in order to get a better fit to the phase values close to the jump. The width of the weighting function had a standard deviation of 10 dB. To summarize, the first fit determines $c$, the stimulus level at which the jump occurs. The second fit determines $A$, the size of the jump then being $|2A|$. 

Gradients were also computed to determine how emission magnitude and phase varied with respect to stimulus level. The gradients were computed using Matlab's \textit{gradient.m} function (central differences).
6.4 eOAE Growth

6.4.1 SFOAEs

Figure 6-3 shows SFOAE level growth functions for four different species. These are single curves from individual ears, but serve to give an overall indication of the phenomena. The frequencies chosen for each species shown in Fig. 6-3 were close to the most sensitive frequency for the given species. Measurable emissions were highly stable over short periods of time (~17 s, the length of time required for averaging at a particular intensity) and showed little variability. Typically, only measurements made very close to a notch showed any significant degree of variability. In all the species except the frog, emissions appeared stable over longer periods of time, as indicated by multiple measurements made at a given stimulus intensity. For the frog, there was some shift apparent in emission magnitude that was not due to changes in the probe coupling or earphone calibration. Unlike the chicken and gecko, frog body temperatures were not regulated and there might have been a slight shift in temperature.

Several features are worth emphasizing in Fig. 6-3. First, all species exhibit some degree of non-monotonicity in SFOAE growth with varying degrees of phase transitions at the notches. While not shown explicitly for the chicken here (where growth was fairly linear at all intensities), other measures have indicated non-monotonic SFOAE growth in the chicken⁵. The notch in this human growth curve is much broader than that observed in the gecko and frog. Second, SFOAEs exhibited varying rates of growth at low stimulus levels. In all species except the chicken, SFOAEs grew expansively as they emerged from the noise floor. In the human and frog, growth was only slightly faster than linear, though in the gecko it was closer to cubic. Third, the properties of the phase transition can vary. In both the gecko and frog, the phase jump occurs fairly rapidly (with respect to increasing stimulus intensity) and is close to 1/2 cycle. However for the human case, the phase transition is much more gradual

⁵Direct measures of SFOAE level curves in the chicken are limited. However, the level dependence of frequency sweeps, such as that shown for the human and gecko in Fig. 5-1, indicate that there are multiple frequencies in an individual chicken ear where non-monotonic growth occurs.
(and also close to 1/2 cycle). As shown in Fig. 6-4 for the gecko, there are also cases where the SFOAE phase shift is clearly less than 1/2 cycle. Fourth, there is some degree of smooth variation in emission phase with respect to level at higher stimulus intensities. This is observed in all four species, being most readily apparent in the gecko.

Additional curves for the gecko (at both a low and high frequency) are shown in Fig. 6-4. Each of the two curves was obtained in a different individual. In the 600 Hz case, it appears that growth is close to quadratic at low levels (below 30 dB SPL), becoming compressive up to about 40-45 dB and then close to linear above 45-50 dB. There is not a clear notch in the magnitude, though a transition in the phase of ~1/4 cycle is present. The phase transition in this case is much more gradual relative to the sharp transition seen in Fig. 6-3. In the 2000 Hz curve, growth is close to linear up to ~25 dB SPL, above which there is an extended region of compressed growth up to 50 dB. A sharp notch occurs around 55 dB, above which growth is extremely rapid. At the notch intensity, two separate measurements were made and differ significantly. The discrepancy between these two measurements (10-15 dB difference in magnitude and different degrees of uncertainty), likely taken 5-10 min. apart, indicate that variability in the emission can occur over long periods of time (at least greater than 17 s) when the stimulus intensity is close to a notch and the notch is close to the noise floor. In the compressed region before the notch, the phase varies slowly by ~1/2 cycle, with a jump of 1/4 cycle occurs upon traversing the notch. Note that the phase also exhibits greater variability at the notch intensity. The results in Fig. 6-4 indicate that non-monotonic SFOAE growth can occur at both relatively low and high frequencies in the gecko.

Additional SFOAE level curves in the gecko ear are compiled into a single plot as shown in Fig. 6-5 (both gecko species are included together). Stimulus frequency is effectively included by varying the line thickness: the thicker the line, the higher the probe frequency the level function was obtained at (the range was $f_p = 500$-5000 Hz). For a given level curve, loess trends were fit to the data points to facilitate visualization [see Ch.2]. The noise floor curve shown (computed by averaging the
Figure 6-3: SFOAE level growth curves from four different species. Both emission magnitude and phase are plotted. Gecko species shown here was a leopard gecko. Note the difference in probe frequency across species. The particular curves shown here are representative for a given species, though there can be significant variation in the exact shape across individuals (for a fixed set of parameters) and stimulus frequencies (in a given individual). Presentation of levels was randomized and some repeat measurements at certain levels were made. Dashed lines indicate the noise floor. Some phase values were shifted vertically by one cycle and points dominated by noise were excluded for clarity. $[L_s = L_p + 15 \text{ dB}, f_s = f_p + 40 \text{ Hz}]$
Figure 6-4: Additional plots of non-monotonic SFOAE level growth curves from individual leopard geckos at two different frequencies. A different animal was used for each frequency. \([L_s = L_p + 15 \text{ dB}, f_s = f_p + 40 \text{ Hz}]\)

±3 frequency bins adjacent to that of the probe frequency) is from an individual ear and is typical of what was observed in the level growth functions. The noise floor was relatively independent of the emission frequency, varying no more than by a few dB. The increase in the noise floor at higher stimulus intensities likely stems from increased measurement system nonlinearity.

While some features of the level growth have been smoothed out by computing the trend (obscuring details such as a notch), it is observed that growth is not non-monotonic at all frequencies. In many cases, SFOAE growth was very close to linear over the entire intensity range tested. Except for the highest frequencies tested, the intensity that emissions emerged from the noise floor (≈10-20 dB SPL) was relatively independent of frequency.

6.4.2 DPOAEs

Figures 6-6 and 6-7 show cubic DPOAE growth in four different species. For a given species, a single individual was used in both figures. A primary pair \((f_1 \text{ and } f_2)\) was varied over the intensity range (with \(L_1 = L_2\)) and both cubic DPOAEs extracted, (low-side) \(2f_1 - f_2\) shown in Fig. 6-6 and (high-side) \(2f_2 - f_1\) in Fig. 6-7. These two figures indicate how both cubic distortions grow independently with respect to
Figure 6-5: SFOAE level growth curves from all gecko ears (both species). Line thickness correlates to probe frequency, with thicker lines meaning higher frequencies (range: $f_p = 500 - 5000$ Hz). Color corresponds to different experimental runs. For clarity, curves represent loess fits to actual data points and thus some features have been smoothed out (such as notches). Dashed line indicates a typical noise floor observed during a given run. [$L_s = L_p + 15$ dB, $f_s = f_p + 40$ Hz]
Figure 6-6: DPOAE (2\(f_1 - f_2\)) level growth curves from an individual ear in four different species. Note that primary frequencies vary across species. See caption for Fig. 6-3 for additional details. \([L_1 = L_2]\)
changes in the (fixed-frequency) primary pair intensity. Note that the stimulus frequencies (and the ratio between them) vary across species. In a particular species, the frequencies chosen tended to be those where emissions were largest and thus reveal growth properties extending down towards lower levels.

Focusing first on Fig. 6-6 ($2f_1 - f_2$), growth in all species is close to linear at lower levels. With the exception of the gecko, emissions appear gradually out of the noise floor as stimulus level is increased. While the magnitude shows little uncertainty once above the noise floor, there is a large degree of variability in the phase up to moderate stimulus intensities (especially in the frog). In the case of the gecko, the emission appears to effectively jump out of the noise floor, at which point growth is close to linear. In all species, growth becomes compressive at moderate stimulus intensities. Growth remains monotonic in the human and frog and the phase is relatively constant in those two species, with no clear sharp transitions. The human phase does vary by a significant fraction of a cycle over the range of 50-65 dB SPL. The chicken and gecko both have a small degree of non-monotonic growth (or a 'dip'), which was also apparent in the phase as there is a relatively sudden transition at the point where the magnitude dips. In the gecko, the transition was sharp and ~1/4 cycle, the phase varying smoothly with stimulus level above the jump. In the chicken, the phase transition was more gradual and about 1/2 cycle. The gecko also exhibits a region of compressed growth at higher levels (60-70 dB SPL), with a second non-monotonic dip present that is shallow and relatively broad. At higher levels in the frog (above 70 dB SPL), the rate growth increases suddenly and becomes close to cubic. Note that a well defined notch is not observed in any of the four species. In all species, repeated measurements at a given intensity showed little change (in either magnitude or phase), indicating that the DPOAE at $2f_1 - f_2$ was relatively stable over time (5-10 minutes).

If we now examine Fig. 6-7 ($2f_2 - f_1$), the magnitudes were relatively smaller than those of $2f_1 - f_2$ for the human and to a lesser extent, the chicken. In both those species, low-level growth appears faster than linear (but not quite quadratic). For the humans, the emission magnitude starts to decrease above 60 dB SPL, bottoming
Figure 6-7: DPOAE ($2f_2 - f_1$) level growth curves from an individual ear in four different species. The curves shown here come from the same data set as that shown in Fig. 6-6.
out in a notch around 70 dB that also exhibits a phase transition of ~1/2 cycle. For levels above the notch, growth becomes cubic. Though not exhibiting a notch as seen in the human, the chicken has a shallow non-monotonic dip with a corresponding phase jump of 1/4 cycle. The dip occurs at approximately the same stimulus intensity as that seen for $2f_1 - f_2$, ~80 dB SPL. At moderate intensities for the chicken, the emission magnitude shows little variability while there is a large degree of uncertainty in the phase (similar to $2f_1 - f_2$). It is not until the dip that the phase starts to become clearly defined. The gecko exhibits two extended regions of highly compressed growth, from ~35-50 dB SPL and 55-65 dB, giving the appearance of a staircase. Growth is close to cubic in-between. The phase in the gecko is relatively constant until the end of the first ‘step’, at which point it starts to vary smoothly with level. Note that the $2f_1 - f_2$ dip for the gecko occurs at a slightly higher intensity by 2-4 dB than that of the first step for $2f_2 - f_1$ (where the phase starts varying with level). There is no sudden phase transition in the gecko. The frog exhibits a wide range of compressive growth from 50-70 dB SPL, where the emission is neither well defined nor stable (in magnitude or phase). However, at 70 dB SPL, there is a transition similar to that exhibited by $2f_1 - f_2$ where the rate of growth suddenly increases, becoming cubic, and the emission takes on a clearly defined magnitude and phase.

In some cases for the gecko ear, multiple notches for cubic DPOAEs were observed. Between notches, there would be a phase transition of ~0.5 cycles. In ears where these multiple notches were apparent, they occurred only at smaller primary ratios and using primary frequencies around 1-1.2 kHz.

Comparing the growth of $2f_1 - f_2$ to $2f_2 - f_1$ indicates how the two emissions may or may not be generated in a similar fashion. In the human, it is apparent that non-monotonic growth can occur for one DPOAE but not the other. This indicates that the behavior of $2f_2 - f_1$ does not mirror that of $2f_1 - f_2$ in the human. In contrast, both cubic DPOAEs in the non-mammalian species behaved in a similar fashion relative to each other (particularly for the chicken and frog). Quantitative differences between the low and high-side emissions are apparent in the gecko (such as the degree of compressive growth, the transition point between compressive and
cubic growth, etc.), but qualitatively both exhibited multiple regions of compressed growth (stair-case).

Not shown directly here but alluded to earlier when discussing averaging, DPOAE level functions can vary widely from ear to ear. Using the same set of stimulus parameters in different individuals of a given species can yield different results. Whereas one ear might clearly exhibit non-monotonic growth, another might be linear over the entire intensity range tested. Thus there is a large degree of variability in the eOAE level functions from ear to ear (similarly with SFOAEs).

6.4.3 Frequency Dependence

Up to this point, we have considered emission growth for a fixed set of stimulus frequencies (with the exception being Fig. 6-5). Fig. 6-8 shows human SFOAEs for frequency sweeps taken using varied stimulus intensities. The connection between Fig. 6-8 and the level-growth curves (such as that in Fig. 6-3) would be to draw a vertical line at a particular frequency and examine how the emission growth varies between the intersecting curves (which were made using 10 dB steps in the stimulus intensity). Linear growth would be indicated by a 10 dB spacing between curves while compressive growth would exhibit spacings smaller than 10 dB.

Figure 6-8 shows that for frequencies below 2925 Hz and above 3050 Hz, growth is nonlinear but monotonic over the intensity range tested. The vertical spacing between the curves is less than 10 dB, indicating that the growth is compressive. The degree of compression varies both with respect to frequency and intensity. For example, at 3150 Hz, an increase in $L_p$ from 30 to 40 dB SPL produces almost no change in the emission magnitude. Conversely at 2800 Hz, the same change produces a difference of ~5 dB (still less than linear). So the degree of compression at a given intensity range can vary significantly with respect to frequency, even over a few hundred Hz. The growth actually becomes non-monotonic for probe frequencies between 2925 and 3050 Hz, with the deepest notch occurring at frequencies just below 3 kHz. At the nadir, a 20 dB SPL stimulus produces an emission 10 dB larger than that of a stimulus at 40 dB SPL. Moving away in frequency from the notch, the difference in the emission
magnitude between those two conditions becomes smaller and eventually the 40 dB probe elicits a larger emission.

Another level-dependent feature stemming from Fig. 6-8 is the effect of stimulus intensity upon the emission phase. At low intensities (30 dB SPL or below), the slope of the phase-frequency function (phase gradient) is independent of intensity. When $L_p = 40$ dB SPL, the intensity where the notch is present, a transition in the phase gradient occurs such that the slope becomes smaller. At higher stimulus levels (50 dB and above), the slope has decreased roughly by a factor of two and indicates a decrease in the emission phase gradient. The level-dependent shift in the phase gradients observed here indicates that the SFOAE frequency dependence varies with respect to changes in stimulus intensity. Similar results, in that both notches and phase gradients were level/frequency-dependent, were found in the gecko (see Fig. 5-1). For both chickens and frogs, while no change was detected in the phase gradients with respect to level, (Fig. 5-3), the existence of notches in the SFOAE level functions depended upon the particular stimulus frequency. So while Fig. 6-8 only shows results for the human, it is observed that in all species examined that at least some subset of notches that occur in frequency sweeps and intensity sweeps are correlated.

6.4.4 Non-Monotonicities with Respect to $L_s$

Another factor to be considered when examining SFOAE level functions is the relative intensity difference between probe and suppressor when using a 2TS paradigm. Fig. 6-9 shows how SFOAEs are affected by the relative difference between $L_p$ and $L_s$ in a gecko ear. In the majority of cases where similar measures were made in the gecko ear using different stimulus frequencies and fixed $L_p$ values, notches of this type were not observed and the response magnitude was typically saturated by the point where $L_s = L_p$. While notches of these type appear to be rare, they indicate that the relative difference between $L_p$ and $L_s$ is an important consideration when examining SFOAE growth measured using a 2TS paradigm. This is of particular concern for our study, where a fixed suppressor level of $L_s = L_p + 15$ dB is used; the notch in the emission phase gradient relates to a delay.

---

6 See Ch. 3 and App. C for background on how the emission phase gradient relates to a delay.
Figure 6-8: SFOAE frequency dependence measured at different stimulus intensities in a single human subject. A notch, both with respect to frequency and level, is apparent around $f_{p} \approx 3$ kHz and $L_{p} \approx 40$ dB SPL. Furthermore, there is a transition in the phase gradients, depending upon whether the stimulus level is below (larger gradient) or above (smaller gradient) 40 dB SPL. [$L_{s} = L_{p} + 15$ dB, $f_{s} = f_{p} + 40$ Hz]
Figure 6-9: SFOAE dependency upon suppressor level ($L_s$) for a fixed probe level ($L_p$) in an individual gecko. Suppression saturates once the suppressor is $\approx 10$ dB above the probe level, but subsequently becomes non-monotonic when $L_s = L_p + 16$ dB. Above the notch level, the emission magnitude increases further. Dashed line indicates the noise floor. [$L_p = 40$ dB SPL, $f_p = 500$ Hz Hz, $f_s = 540$ Hz]
Fig. 6-10: Size of SFOAE phase jump from all species and stimulus frequencies. The jump size was computed by fitting a sigmoidal curve to the phase data when a notch in the magnitude or jump in the phase was visually apparent. The peak of the distribution indicates that the phase jumps tend to be smaller than 1/2 cycle (red dashed line), typically 0.3-0.4 cycles.

Fig. 6-9 occurred when the suppressor level was ~15 dB above that of the probe.

In the specific case shown in Fig. 6-9, $L_p$ is fixed at 40 dB SPL and $L_s$ is varied (the frequency of both the probe and suppressor are fixed). When $L_s$ smaller than $L_p$-10 dB, the measured response is dominated by noise. Once $L_s$ gets above 30 dB SPL, the emission starts to grow quadratically with respect to $L_s$ (except for a small dip that occurs when $L_s = L_p$) and the phase is constant. Once $L_s$ is ~10 dB above $L_p$, the response starts to saturate. However, further increase in $L_s$ cause the response to decrease, resulting in a notch and phase jump of 1/2 cycle. Above the notch, the SFOAE response grows linearly at first (the phase varying smoothly with $L_s$) and then appears to start saturating out.
Figure 6-11: Histogram showing how SFOAE magnitude and phase vary with respect to stimulus intensity at points at least 5 dB below a notch. All species and stimulus frequencies are included. Gradients at the endpoints are excluded. Magnitude growth tends to be slightly compressive (linear growth indicated by red line). Low-level SFOAE phase is typically constant with respect to stimulus intensity (distribution peak centered about zero), but the width of the distribution indicates...

6.4.5 SFOAE Phase Jumps

The size of the SFOAE phase jumps (when a notch was present) were quantified from 40 level growth curves (all species) and is shown in Fig. 6-10. See Fig. 6-2 for an example of how the fits were computed. The jump size distribution is centered about 0.3-0.4 cycles and is clearly less than 1/2 cycle (as indicated by the dashed line). The phase unwrapping algorithm used does not allow jumps greater than 0.5 cycles, so the phase changes greater than 0.5 cycles are not due to unwrapping artifacts.

6.4.6 Sub-Notch Intensity Gradients

Figure 6-11 indicates how SFOAE magnitude and phase vary with respect to stimulus intensity at lower stimulus levels, at least 5 dB below a notch (if one is present). Only points where the magnitude is at least 10 dB above the noise floor are included. Each contributing value is the slope with respect to intensity at a given probe level (of the 40 curves, each contributed ~15-25 values). Low-level SFOAE growth tends to be slightly compressive (the red line on the left-hand figure indicates linear growth), typically ~0.8 dB/dB. The phase is, on average, constant below a notch with respect to changes in stimulus intensity (this can be seen in the phase-level curve shown in Fig. 6-2). The width of the phase distribution arises due to two factors. First, the
phase is not always constant and varies slightly with level as shown for the chicken in Fig. 6-5. Second, the presence of broad notches (where the phase jump is gradual and drawn out over a region extends beyond 5 dB below a notch/phase jump) yielded contributing values (see the human in Fig. 6-5).

6.5 Discussion

6.5.1 Nonlinear Emission Growth

Comparison Across Species

It is clear that there are overall similarities in eOAE level growth properties across species. First, all species exhibit some degree of non-monotonic SFOAE and DPOAE growth. Furthermore, non-monotonic growth was not always marked by deep notches, but shallow dips were also apparent that would be accompanied by smaller phase transitions, less than 1/2 cycle. Second, all species exhibit non-monotonic growth in a frequency dependent fashion. Notches were not always present, indicating that growth could also be monotonic. But there appears to always be some choice of stimulus frequencies for a given individual where growth would behave non-monotonically. This frequency dependence was apparent in species with a tuned BM (human and chicken) as well as those that lack a tuned BM (gecko) or a flexible BM altogether (frog).

Two common features that all the species examined in this study have are hair cells and overlying tectorial structure(s) that couples hair cells together. It is however important to also consider non-vertebrates, which lack even these two features. The growth of eOAEs in insect ears show many similarities with eOAE in the species described in this report, such as non-monotonic growth and level-dependent phase transitions that are physiologically vulnerable [Kossl and Boyan, 1998; Coro and Kossl, 2001; Kossl and Coro, 2006].

Our results make it clear that mammalian specific features, such as hair cell somatic motility and a tuned BM, are not required to account for an array of eOAE
level growth features. Whatever the underlying mechanism(s) is that gives rise to non-monotonic behavior, it appears to exist in a wide range of ears that exhibit greatly differing morphologies. Furthermore, if SFOAEs and DPOAEs are generated by two different mechanisms as has been proposed for mammals [Shera and Guinan, 1999], both these mechanisms are capable of exhibiting a non-monotonic dependence on stimulus intensity.

Comparisons with Previous Studies

Overall, our human results were consistent with previous studies [Schraier et al. 2003 (SFOAE); Abdala 2000 (DPOAE)]. For the sake of brevity, we only discuss studies that examined eOAE growth in non-mammals here. For the chicken ear, our results were consistent with previous reports. Kettembeil et al. (1995) and Lichtenhan et al. (2005) observed instances of non-monotonic growth for $2f_1 - f_2$ in the chicken, but reported that growth generally appeared monotonic for the majority of stimulus conditions. Ipakchi et al. (2005) also reported growth functions, but averaging was done both across individuals and frequency and it is difficult to distinguish features that would be present in individual curves. Their reported values indicated growth was only nonlinear for stimulus intensities above $\sim 80$ dB SPL, where it became expansive at all but the highest frequencies they tested (2-2.5 kHz). Consistent with our data, growth at lower stimulus levels (below $\sim 75$ dB SPL) in all studies was linear. Note that these studies used different anesthesia paradigms (we used the same one as Ipakchi et al., urethane injection i.p.), but this difference appeared to have little effect on the overall growth functions.

While we know of no reports of evoked emissions in the gecko ear, DPOAE growth functions have been reported in other lizard species. Rosowski et al. (1984) found nonlinear growth at $2f_1 - f_2$ in the alligator lizard, a species that lacks a TM over a significant portion of its papilla. They found that growth was cubic from the low-

---

7As described in a subsequent section describing unmixing experiments, the measured emission is most likely comprised of multiple generation sources, each of which is operating under a different mechanism. So it may be possible that the non-monotonic growth represent a shift in the dominant emission mechanism in the mammalian ear.
est stimulus intensity (50 dB SPL; the emission was dominated by noise at lower intensities) up to about 60 dB, at which point growth slowed and even became non-monotonic. At higher levels between 65-75 dB SPL, growth again became cubic and eventually started to saturate at higher levels. This result is qualitatively similar to what we observed in the gecko, though we were able to measure at lower stimulus intensities and see compressive growth below 50 dB SPL. Though growth appeared mostly monotonic in their report, instances of non-monotonic growth were found in certain cases for the alligator lizard. When notches were apparent, growth was either compressed or linear at levels below the notch and the was also a corresponding ~0.4-0.5 cycle phase jump. They also noted that emission phase could vary smoothly with level, particularly at higher stimulus intensities. Having performed numerous physiological manipulations (such as acoustic overexposure, pharmacological manipulations and BM destruction), Rosowski et al. concluded that emissions in the lower level region (50-60 dB SPL) were highly vulnerable while the emissions evoked at higher intensities were not. They suggested that this observation indicates the presence of at least two different generations sources. In another lizard species, the bobtail lizard, Manley et al. (1993) found growth to be close to linear at lower levels and cubic at higher levels, sometimes with a deep notch in between.

In the case of the frog, Meenderink and Narins (2006) examined SFOAE level functions. They found that growth functions in the frog varied significantly in a frequency dependent way, depending upon whether \( f_p \) was above or below ~1-1.5 kHz. At lower \( f_p \) values, Meenderink and Narins observed growth to be compressive at moderate levels (40-60 dB SPL) and quadratic at higher levels. For larger \( f_p \) values, growth at lower levels was quadratic and starts to saturate out for stimulus levels above 70-80 dB SPL. Our results are consistent with their report. The frequency dependence of growth behavior was attributed to the notion that the two different auditory papilla in the frog act as the primary generation source in a given frequency range and do so in different fashions. We were able to measure SFOAEs in the frog

---

8The amphibian papilla (AP) is sensitive to lower frequencies while the basilar papilla (BP) sensitive to higher frequencies [see Ch.3 for further description of the two]. A frequency transition between the two (in terms of ANF innervation patterns) is thought to be ~1.2 kHz in this particular
at lower stimulus levels and found that below 40 dB SPL, emission growth at lower frequencies (below 0.8 kHz) is linear. At moderate levels (40-60 dB SPL), growth was highly compressive, becoming quadratic at higher levels (consistent with Meenderink and Narins). Over the entire intensity range tested at these lower frequencies, the SFOAE phase tended to be constant, except in specific cases when sharp notches were present at the particular frequency tested. As frequency increases towards 1 kHz, linear growth extends toward higher levels before bending over and starting to become compressive. At frequencies above 1.5 kHz, we observed that growth remained quadratic at lower levels and was typically in the noise floor below 40 dB SPL. While frog SFOAE growth did exhibit non-monotonicities, growth was generally observed to be monotonic. Both studies indicate that there is an extended intensity region (~20 dB) in the frog ear where SFOAE growth is highly compressed and that the compression is strongest at lower frequencies (below 800 Hz).

As for DPOAE level growth functions in the frog, our results were consistent with those reported by Meenderink and van Dijk (2004) for the same species. Both cubic DPOAEs in their study showed a transition around 70-75 dB SPL where emissions started to grow cubically, similar to our observations. Meenderink and van Dijk also reported that DPOAE phase can vary smoothly with level in the frog ear at higher stimulus levels (above 75-80 dB SPL).

6.5.2 Comparison of eOAEs and Single-Source Model

Comparing eOAEs to Model Predictions

In Sec.6.2.2, we made predictions using the single-source model. We compare our eOAE results to those predictions and discuss how the model does and does not account for emission growth. First, the slopes of the DPOAE level functions do not match at lower levels. For distortions such as $2f_1 - f_2$ and $2f_2 - f_1$, the model predicts growth to go as the cube (when $L_1 = L_2$) in the region well below the notch. However, the emission growth appears to be much closer to linear. Also, SFOAE growth (using species [Ronken, 1991]).
a 2TS) paradigm would be expected to grow as the square of the primaries, but again, growth appears closer to linear (Fig. 6-11). So in the case of both SFOAEs and DPOAEs, low-level emission growth is slower than predicted. Lukashkin et al. (2002) noted this discrepancy and suggested the presence of an additional nonlinearity that compresses the cubic growth. It is not clear what this additional nonlinearity might be and how it would compare across species.

Second, we only observe notches at a subset of stimulus frequencies for either SFOAEs and DPOAEs. This observation is in contrast to the frequency-independent nature of the single-source model, which would predict notches at all frequencies. Based upon the results shown in Fig. 6-8, emission growth behavior clearly depends upon stimulus frequency, indicating that there is a correlation between at least some subset of notches that are seen in constant-intensity frequency sweeps and constant-frequency intensity sweeps. Seeing that eOAE level functions are both stimulus intensity and frequency dependent, this correlation implies that the nonlinear single source must also have a means of frequency selectivity if it is to accurately describe eOAE growth. While not built into the single-source model, a hair cell (i.e. the point-source) exhibits a degree of tuning due to the mechanical properties of the stereociliary bundle and the nearby fluids, which can lead to a resonant behavior⁹. It is unlikely that the addition of a frequency dependence to the single-source model will be sufficient to account for the observed eOAE growth properties. A point-source model described by Meenderink and van Dijk (2004) was proposed to account DPOAE frequency dependence. Their model was a single nonlinear resonator, and while level growth was not specifically examined, the model accounted for a number of qualitative features in the frequency dependence of the DPOAE magnitude. However, it was not able to account for the phase behavior (their model is a second order resonator and showed a 1/2 cycle shift in the phase about the oscillator's resonant frequency), and some sort of additional phase shift is needed. Even with a resonance built in to the single-source model, it is unlikely that it would be able to account for the level-dependent shift in

⁹This type of micro-mechanical response is believed a primary mechanism for tuning in lizard species that lack a TM and where the BM exhibits no tonotopic behavior [Aranyosi and Freeman, 2005].
the delays (as shown in the bottom part of Fig. 6-8).

Third, DPOAE phase discontinuities (i.e. jumps) differ significantly from the predicted transition of 180°. The model constrains the phase to one of two values (separated by 1/2 cycle) and the data clearly indicate transitions that range between 0 to 1/2 cycle (as shown for SFOAEs in Fig. 6-10). The size of the phase transition roughly correlates to the depth of the notch: the deeper (and wider) the notch is, the closer to 1/2 cycle (and more gradual) the phase jump is. The phase transition is not necessarily sharp and can appear more gradual with the phase taking on intermediary values in between the lower and upper limiting values (this is apparent for the human SFOAE phase curve shown in Fig. 6-3 or both cubic DPOAEs for the chicken in Fig. 6-4 and 6-5). The model itself only predicts a sharp transition. The implication is that there must be some means in the model (for it to work) to account for a slower phase transition when crossing a notch that does not constrain the phase to be one of two values separated by a fixed amount. Furthermore, at higher levels, the phase can vary smoothly with level (like in the gecko SFOAE curve in Fig. 6-3 for example), a feature that is difficult to explain within the context of the single-source model.

Fourth, eOAE growth functions tend not to exhibit the generic shape of the growth curves in the single-source model (as shown in Fig. 6-1). Aside from the discrepancies listed above, we observe extended regions of highly compressed emission growth. For example, the gecko in Fig. 6-7) shows two separate regions where over the course of a 10-15 dB change in stimulus intensity, the emission magnitude is practically unchanged. In general, a wide range of unique features are present in the growth functions (such as sudden dips, shallow/deep wide/sharp notches, extended regions of compressive growth, etc.), which vary significantly across both individual ears and the choice of stimulus parameters. The model does not capture any of these features and its generic shape often differs significantly from that of an individual growth function\(^\text{10}\).

\(^\text{10}\)Personal observations of DPOAE growth in the guinea pig ear has shown those curves to take a relatively generic shape that is similar to that exhibited by the model (though discrepancies such as the rate of growth are still present). It is not clear why guinea pigs exhibit a more uniform growth across individuals and further study is needed. This observation though urges that caution is needed when trying to make conclusions about a very general class of model to a wide range of different
What Do Shortcomings of the Model Imply?

It is clear that there are numerous inconsistencies between the single-source model predictions and eOAE growth. Given the simplifying assumptions present for the model and its inherent limitations, the single-source model is obviously not going to provide an all encompassing description of eOAE growth. However, the numerous discrepancies between the model and the data indicate that the single-source model is too simple.

In addition to the differences between model predictions and the data described in the previous section, the degree of physiological vulnerability that has been observed in eOAE growth works against the notion of a single-source model. It has been observed that physiological changes induced in an animal (such as by acoustic overexposure, hypoxia, genetic variations, and even death) tend to strongly affect level growth only at low stimulus levels and have little effect at higher stimulus levels [Rosowski et al., 1984; Coro and Koss, 2001; Liberman et al., 2004]. Though Lukashkin et al. (2002) use a physiological variation to support the single-source model (via injection of furosemide into the cochlear fluids, which they assumed would lower cochlear amplification and thus only affect the lower stimulus level growth if two sources were present), a number of factors could account for the shift toward higher levels in the notch location they observed (such as some sort of dc biasing induced by the decreased EP)\textsuperscript{11}. In general, changes in the underlying nonlinearity in the single-source model (as might be affected by physiological changes) have the effect of translating the level growth curve (vertically and/or horizontally), but it does not appear possible to affect growth of one portion of the curve (i.e. below the notch) and not the other (above the notch). Clearly, physiological vulnerability of different portions of the growth curve need to be accounted for. The effects of induced physiological variations are discussed further in Sec.6.5.4.

\textsuperscript{11}Lukashkin et al. (2002) argue that if two sources are present and furosemide affects only the (active) low level source, than a notch should shift towards lower intensities because the active source has a smaller response and thus will interfere to a lesser degree with the passive, high level source. In the data they report, the notch shifts towards higher levels. Furthermore, the role of furosemide upon non-monotonic ANF rate functions is described in Sec.6.5.4.
That is not to say that the single-source model is not without its merits. The comparative approach has validated one strength of the model, its anatomical-independence. We have observed that many eOAE growth features are similar across the species examined here in spite of gross morphological differences. This similarity supports the notion that the generalizable nature of the model is a necessary feature.

Additionally, the single-source model may be applicable in morphologically simpler ears, such as those of invertebrates. DPOAE level functions have been measured in species such as grasshoppers and moths that have exhibited non-monotonicity with many properties similar to DPOAE growth in vertebrates, including physiological vulnerability [Koss, Boyan, 1998; Coro and Koss, 2001]. However, even in these species, there are inconsistencies between model predictions and the DPOAE data that would need to be accounted for. For example, Coro and Koss (2001) show phase jumps less than 1/2 cycle in the moth when making a transition across a DPOAE notch and growth is quadratic at lower levels (not cubic as the model would predict).

How Might Models Explaining eOAE Growth Be Improved?

Modifications to the single-source model have been proposed that enable it to better account for observed eOAE properties. Chief among these is the addition of a dynamic operating point [Lukashkin and Russell, 2005]. Studies have argued that this operating point can be shifted about using a low frequency biasing tone, effectively mapping out the cochlear transduction function [Bian et al. 2002]. Additions such as a dynamic operating point can greatly complicate the dynamics of the single-source model, leading to increasing complex behavior. It is not clear at this point if these modified versions of the single-source model are able to effectively capture more of the features observed in the data.

We take a different approach here and argue that a greater degree of versatility is necessary for models of nonlinear eOAE growth, as the primary goal is to gain insight into the underlying physiology. There are three aspects that will need to be considered in future modeling efforts. First, the ear is made of many parts (or sources), which are all coupled together mechanically via fluids and tectorial structures. In
this regard, eOAEs should not be considered to derive from a single source, but as a spatial summation of distributed sources. In other words, eOAEs are not generated by a point-source, but from an extended region comprising a number of different generators. Second, it is important to consider the frequency dependence of emission growth. One approach may be to consider each source as a nonlinear resonator that is coupled to other nearby sources. Furthermore in terms of frequency dependence, the model will need to be able to capture the level-dependent transition in the emission phase gradients. It may prove that the shorter delays at higher levels are due to differences in reverse propagation from the generation region (back to the external ear). It is however difficult to directly distinguish between emission generation and reverse propagation and any model that is to be of value needs to encompass both these aspects. Lastly, the observations of physiological vulnerability upon eOAE growth will need to be considered. Understanding how low-level growth can be strongly affected while that of higher levels is not will likely serve to further our understanding of the possible role of active elements (i.e. amplification) present inside the ear.

6.5.3 Further Evidence for Multiple Sources

We have demonstrated that eOAE growth functions exhibit numerous inconsistencies with predictions of the single-source model, suggesting that non-monotonic growth functions stem from the interaction of multiple sources\(^{12}\). The question then becomes, what might these different sources be?

The studies described below have employed methods that effectively allow for an unmixing of the emission, deconstructing the measured emission into different components. All of these studies indicate that emissions (at least in mammals) arise from different sources, each contributing to the emission in varying extents and each with their own unique characteristics. Stover et al. (1996) separated DPOAEs in the

\(^{12}\)Even though they emphasize the 'single' nature of their model, Lukashkin et al. (2002) appeared aware of the need for a more distributed approach. In their paper, they state that difference between model and data "... might be anticipated because DPOAEs are generated by a distributed nonlinearity from an extended region of the cochlear partition with some phase gradients along the region. The vector summation of the DPOAEs from different parts of this prolonged region might lead to a partial cancellation of the notch and corresponding phase change."
time domain using a signal analysis technique (exploiting the fact that the components from different sources have different latencies). They observed multiple components, the dominant one varying with respect to changes in stimulus intensity (thus providing an explanation for the observed level-dependent shift in the emission phase gradients). Individual components exhibited nonlinear growth (with respect to changes in $L_2$), larger latency (i.e. longer delay) components showing compressive growth that was non-monotonic in some cases. Fahey et al. (2000) employed a suppression paradigm by presenting a third tone, whose frequency was varied to effectively preclude different sources. In their report, they argued that suppression maps can be explained by the presence of multiple emission components that interact in a nonlinear fashion. Kalluri and Shera (2001) extended both methods (time-domain windowing and third tone suppression), indicating that not only are there different generation sources present, but that each is generated by a different underlying mechanism [see Ch.1 for further discussion]. Goodman et al. (2003) examined SFOAE unmixing in the guinea pig ear and also found multiple sources, each of which exhibited significantly different latencies (similar to DPOAEs). In that study they also examined level dependence and found results similar to those of this study: non-monotonic SFOAE growth that varied with respect to frequency and also showed level-dependent shifts in the phase gradients.

For a constant stimulus intensity, comparing the frequency dependence of the two unmixed components (for either SFOAEs or DPOAEs), it is observed that the two interact with one another. For example, the long latency component present in DPOAEs ($2f_1 - f_2$) is typically smaller in amplitude than the shorter latency component for emission frequencies above $\sim$1 kHz. The effect of the long latency component however produces a periodic structure in the emission magnitude (called fine structure) with peaks and valleys spaced approximately 1 cycle apart (relative to the phase delay of the long latency component). The more comparable the two components are in magnitude, the more strongly they will interact. The dominant

---

13In their paper, Stover et al. use the term latency when referring to emission phase gradients, whereas we have used the term delay up to this point. For the purposes of this report, the two terms are equivalent.
stimulus component determines the phase of the emission (and thus the phase gradient). To summarize, the results of all these unmixing studies indicate that the emission measured at the ear canal is comprised from multiple sources, which can interact with one another to affect the total measured response.

It will most likely prove worthwhile to perform unmixing on non-mammalian emissions to determine if they are also comprised of multiple components as observed for mammalian DPOAEs. While emission latencies are shorter in non-mammals relative to mammals (and humans in particular), the delays are long enough (~1 ms) that the time-domain windowing methods are still possible. If multiple sources are present, analysis of their relative magnitudes and phases (with respect to both stimulus intensity and frequency) could reveal how these sources interact to give rise to non-monotonic growth functions.

While it is not clear how the different sources interact, they all must exhibit some degree of non-linearity (otherwise, no distortion would be present). While physiologically it is unreasonable to have a purely expansive nonlinearity, it need not necessarily be purely sigmoidal in nature. Specifically, it appears reasonable that hair cell mechano-electro transduction is not the only source of nonlinearity present in the inner ear that could be playing an appreciable role. Focusing just at the level of the hair cell, there are multiple sites where nonlinearity can manifest: sigmoidal transduction at the bundle tip, capacitance across the hair cell membrane or the relationship between OHC length changes and the transmembrane potential [Patuzzi, 1996]. Furthermore, other macromechanical nonlinearities can be present that effect the driving stimulus to the hair cells, such as the passive relationship between the sound pressure induced force on the BM and the actual displacement (i.e.i.e.i.e. a nonlinear stiffness).

It is also worth briefly discussing different modes of motion and how they contribute to level functions. By mode, we refer to an extra dimension in which motion or energy propagation can occur\textsuperscript{14}. For example, Aranyosi and Freeman (2005) de-

\textsuperscript{14}For example, a wheel traveling along a slippery surface can be considered to have two modes of motion, one rotational and one translational.
scribe two modes of motion in the alligator lizard ear, one translational and one rotational. These two modes could possibly be dependent upon each other in some complex level-dependent fashion. In the mammalian cochlea, there may be other modes of motions such as relative motion between the TM and BM or fluid movement along the tunnel of Corti [Karavitaki and Mountain, 2007; Guinan et al., 2005]. Lukashkin et al. (2002) originally argued against the notion of multiple modes affecting DPOAE growth functions based upon the observation that $2f_1 - f_2$ and $2f_2 - f_2$ produce notches at different levels for a given primary pair. Their argument assumes that both cubic DPOAEs are produced at the same location, an unlikely prospect given the very different behavior of each emission’s phase gradient [Knight and Kemp, 2000].

6.5.4 Non-Monotonic Behavior In ANF Responses

As discussed earlier, the model described by Weiss and Leong (1985) predicts the possibility of non-monotonicities arising at the mechano-electro stage of transduction. This feature has been observed to be preserved to a certain extent in the response of auditory nerve fiber (ANF) rate functions. Below threshold, a fiber discharges at its spontaneous rate. When driven with stimulus intensities above threshold, the firing rate increases nearly linearly, reaching saturation (i.e. maximum firing rate) 20-40 dB above threshold. With further increases to stimulus intensity, ANF responses remain constant until much higher levels (typically 40-50 dB above saturation), at which point a non-monotonic notch can occur [Kiang et al. 1986]. Looking at the phase response of the fiber (assuming it is a lower-CF fiber, where phase locking to the stimulus occurs), phase is constant up until the notch, at which point a phase jump of $\sim 1/3-2/3$ cycle occurs.

In the case of mammals, manipulations of cochlear variables (i.e. acoustic trauma that affects IHCs, selective OHC destruction, efferent stimulation and reduction of the endocochlear potential) have been made to assess their effect upon cochlear tuning and ANF response at low and high stimulus levels. Many of these studies were summarized by Kiang et al. (1986), who discuss whether ANF tuning occurs through one or more
excitation mechanisms. What they found was that different manipulations produced different effects in the tail and tip of tuning curves as well as the low and high level regions of the ANF level functions. For example, efferent stimulation affected only the tip of the tuning curve, with little or no effect on the tail nor on the shape of the ANF level function. Conversely, OHC loss (leaving IHC bundles intact) abolished the tip altogether, making the tail hypersensitive. Additionally, loss of OHCs caused the ANF level function notch to disappear.

Kiang et al. (1986) suggest that for higher-CF fibers, the existence of two different regions in the ANF tuning curves (taking into account the effects of various manipulations) indicates the presence of at least two different physiological mechanisms giving rise to tuning in the mammalian cochlea. The first, relatively insensitive and broadly tuned (accounting for the low-frequency tail of the curve), likely stems from the passive BM response and IHCs. The second, highly sensitive and sharply tuned (describing the tip region), appears directly related to the state of the OHCs. It is also pointed out that the tip thresholds are more susceptible to the various manipulations, indicating "that the sharply tuned tip requires a generating mechanism that is more easily damaged than the mechanism that generates the tail" [Kiang et al. 1986].

Given that these two mechanisms work together to provide a large dynamic range and sharply tuned responses, interaction between them could lead to destructive effects (i.e. a notch/phase jump). Within the context of the comparative approach, non-mammals also show existence of tail and tip regions in the ANF tuning curves [Sams-Dodd and Capranica, 1994], though the dichotomy is not as great as that seen in mammals.

It appears unlikely that if the non-monotonicity arose from a single-source, that

---

15 An alternate view is given by Patuzzi (1996), who provides a simple argument that the tail-tip shape of the ANF tuning curve simply stems from the nature of the traveling wave pattern of excitation along the BM.

16 This may be due to the fact that non-mammals have a lower limit on the upper range of their hearing, meaning that CF for a most units is usually well under 10 kHz. As in mammals, units with CFs at or below 2-3 kHz tend to see more v-shaped with a less pronounced tail. It is worth noting that for some non-mammals, asymmetric tuning curves are not always present. In some species such as birds, ANF tuning curves have a highly symmetric v-shape, independent of CF [Koppl, 1997; Manley et al. 1985]. Frog tuning curves also appear v-shaped, though there is evidence for tails in certain populations of low frequency neurons [Frishkopf and Goldstein, 1963; Feng et al. 1975].
manipulation of the cochlear environment would selectively affect one level region and not the other. With respect to the Lukashkin et al. (2002) study, the furosemide effects on DPOAEs they reported chiefly caused a horizontal shift in the notch location and steepening up of the slope below the notch. This is in contrast to the effect upon ANF notches seen by Kiang et al. (1986), where furosemide had little effect upon horizontal notch location but a big effect upon the lower level region of the rate function. In contrast, we might expect in the single-source model that the entire level function would maintain its overall shape (depending upon the specific type of manipulation) and be shifted along one or more dimensions.

It has not been demonstrated however that the basis for non-monotonicities in ANF responses is the same as that for eOAEs. Experimentally, a possible correlation between the two could be probed by simultaneous ANF and eOAE measurement. Along these lines, it is worth pointing out that non-monotonicities have also been observed in summed potentials (such as the cochlear microphonic and compound action potential) [personal observations in alligator lizard CM; Henry, 1995] and to a more limited degree, BM responses [Patuzzi et al. 1984; Robles et al. 1997; Cooper and Rhode 1997; Rhode, 2007].

6.6 Summary

We summarize the main empirical findings here:

- Nonlinear SFOAE growth at moderate to high stimulus intensities was apparent in all species examined: human, chicken, gecko and frog.

- At certain frequencies, SFOAE growth was observed to be non-monotonic in all species. For a given species, the frequencies at which non-monotonic growth occurred varied across individuals.

- SFOAE phase exhibited a transition when non-monotonic notches were present. The rate at which transitions occurred and their size (ranging from 0 to 1/2 cycle) varied with respect to notch depth.
• Similar behavior was observed in DPOAEs for all species: non-monotonic growth and phase jumps at certain stimulus frequencies.

• For a given primary pair, the two cubic DPOAEs could exhibit different level dependencies.

• SFOAE phase gradients exhibited a level-dependent transition in both human and gecko\(^\text{17}\).

• For a fixed primary level, SFOAE magnitude in the gecko can exhibit a non-monotonic dependence on suppressor level.

Similarities in emission growth across species suggest that the basis for nonlinear behavior is not highly morphological-dependent. Given its underlying assumptions and inherent limitations, the single-source model is not able to effectively capture the wide range of eOAE growth features observed. Supported by other lines of evidence, such as DPOAE un-mixing experiments, our results suggest that there are multiple level-dependent sources out of phase relative to each other that can interfere, thereby causing destructive interference that results in the observed notches and phase jumps. We suggest that eOAEs need to be considered as arising from a spatial summation of distributed sources. Furthermore, correlating eOAEs to non-monotonicities observed in other measures of inner ear behavior (such as ANF notches) will serve to reveal how interactions among multiple sources take place. As we learn more about how level-dependent sources are similar and different across species, we will further elucidate the possibility of active mechanisms present in the inner ear to improve perception.

\(^1\text{7}\)Also see Ch.5.
Chapter 7

Potpourri

7.1 Other OAE Properties

We briefly discuss here some other OAE properties that we explored over the course of this thesis and were not discussed in Ch.3-6.

7.1.1 Effect of Death

A striking feature of OAEs is their physiological vulnerability. This property of emissions has repeatedly formed the basis for many arguments about how evoked OAEs directly infer the existence of a cochlear amplifier. Similar to results seen in mammals, Fig. 7-1 shows the effects immediately following death upon both SFOAEs and DPOAEs in the chicken ear. Upon death, emissions disappear within minutes.

7.1.2 Ossicular Interruption

Figure 7-2 shows the effect of ossicular interruption upon leopard gecko DPOAEs post-mortem. While the time course of OAE decay post-mortem is very different in the gecko relative to the chicken, it is clear that the measured OAEs arise from the inner ear. Fig. 7-2 shows that upon removal of the columella, DPOAE magnitude
Figure 7-1: Effect of death upon SFOAEs $2f_1 - f_2$ in the chicken ear. Death was administered via an intra-cardial overdose, upon which the bird immediately stopped breathing and was placed back in the sound chamber for measurement. [SFOAEs: $L_p = 50$ dB SPL DPOAEs: $L_1 = L_2 = 65$ dB SPL, $f_2/f_1 = 1.07$]

falls to the noise floor and the characteristic downward phase slope is vanishes\(^1\).

### 7.2 OAEs a By-product of Cochlear Amplification?

As alluded to throughout the thesis, OAEs have generally been thought of as a by-product of cochlear amplification. This is particularly true with regard to mammalian DPOAEs, where these emissions are commonly used as a direct assay of OHC function\(^2\). One of the most likely reasons for this is the striking observation of the physiological vulnerability of OAEs. In mammals and birds, low-level emissions have been shown to be extremely sensitive to the state of the animal (under such manipulations as hypoxia) and disappear almost immediately upon death\(^3\). The presence of spontaneous emissions, whose properties statistically have been shown to deviate from those expected from filtered noise and more closely resemble a self-sustained oscillator, has

\(^1\)The high-side DPOAE was chosen here because $2f_1 - f_2$ exhibited some significant system distortion around 3-3.5 kHz due to 1-way electric cross-talk between the earphones.

\(^2\)Are OHCs really sufficient to constitute the basis for the cochlear amplifier? It is important for one to keep in mind that OHCs are just one part of a complex system that is highly sensitive to even slight manipulations. For example, in the absence of the large potential difference across their apical surface (EP), OHCs are no longer able to allow the ear to achieve its remarkable sensitivity. Thus in this regard, might one not call the stria vascularis the basis of the cochlear amplifier?

\(^3\)This has also been observed to a certain extent in non-mammals. Manipulations such as varying body temperature or introduction of pharmacological substances can have a significant impact upon certain emission features, particularly upon SOAEs.
also been argued as evidence for an underlying amplifier. Furthermore, more indirect arguments stemming from cochlear modeling efforts indicate the need for an active contribution inside the ear⁴.

Caution is however needed here. While numerous observations suggest the underlying presence of active mechanisms present in the ear, it is a process that is currently poorly understood⁵. Furthermore, one could possibly conceive a purely passive model which could explain evoked emission behavior, taking the physiological vulnerability into account⁶.

It is quite clear that OHCs play a vital role in mammalian auditory function.

---

⁴This argument stems as far back as Gold (1948), who suggested an active force would be necessary to overcome viscous forces present in the inner ear fluids. Neely and Kim (1983) proposed that a cochlear model must provide some sort of active contribution (they used negative damping) in order to better match the sensitivity and sharp tuning that had been observed in more recent experiments (which indicated that the physiological state of the animal was a major factor in what was observed). And Zweig (1991) argued by using an inverse approach that the BM shows a region about the peak of the traveling wave where the real part of the impedance must be negative. Further examples could be listed ad infinitum here.

⁵It is important to define what is meant by an active process. An active process is one which is capable of adding energy to the system, presumably via some sort of metabolic activity. A passive system has a response that is purely reactive and/or dissipative, not adding any energy to the response.

⁶Though obviously, SOAEs and their observed properties are much more difficult to account for in regards to a purely passive model.
However, as mentioned earlier, non-mammals lack somatic motility in their hair cells. Some researchers have suggested that it is not the cell body so much that plays an active role, but the stereociliary bundle [Manley et al. 2001]. Hair cells from non-mammalian species have been observed to be able to generate appreciable forces at the bundle that may be sufficient to couple back the rest of the auditory organ and have some appreciable effect. Recent observations in the mammalian cochlea have indicated a similar process may be at work there as well. Seeing that we still have much to learn about what the underlying physiological mechanism(s) may be that gives rise to the ear's sensitive responses, it is important at this point to reserve absolute judgment about cochlear amplification. It is in this regard that a comparative study may be so beneficial, by better understanding how the processes may be similar and/or different across species.

7.3 Effects of Physical Dimensions

The data clearly indicate that the $2f_2 - f_1$ and SFOAE phase gradients are significantly larger in humans than the other non-mammalian species. The discussion in Ch.3 provides an argument that the basis for this arises from inherent differences in the sharpness of tuning across species. We extend that discussion here in regards to the inherent differences across species in terms of the dimensions of inner ear structures. Differences in tuning might be expected due to the varying lengths of the sensory epithelium.

It is first insightful to identify whether there is a spatial dimension that is roughly constant across species. The general shape and size of hair cells varies somewhat across species, but overall appears to be fairly uniform in vertebrate hearing organs. Thus we might expect that the distance between adjacent hair cells (from center to center) represents a dimension that is relatively constant across species, on the order

---

7And a clear delineation between cell types as is apparent in mammals. However, in spite of a lack of somatic motility, chicken hair cells do have a HC dichotomy (short and tall HCs) similar to that in mammals despite the lack of membrane-based motility.
of $\approx 10 - 20\mu m^8$. If this is a critical dimension where other important dimensions vary significantly across species, it may be expected that mechanical properties such as the spatial wavelength near the peak of the traveling wave will also vary$^9$.

Seeing that the human has a significantly larger hearing organ than any of the other species tested here$^{10}$, perhaps one might expect that the phase gradients are largest in human. It will be insightful to see how gradients compare in other species with dimensions similar to or larger than humans (such as primates, cows or elephants). It is worthwhile to point out that the phase delays were similar in both chicken in gecko despite the significant difference in the length of the auditory papilla. It may be possible that the comparison is not valid here due to fundamental differences in the underlying mechanics (such as stemming from the lack of BM traveling waves in the gecko ear).

### 7.4 Future Work

Some suggestions for future work are described below:

- Examine eOAEs in a wider range of species, particularly species with larger ears (e.g. cow, elephant) to see how emission gradients might correlate to size. Are humans truly unique in having larger SFOAE gradients?

- Test physiological vulnerability of eOAEs in non-mammalian species using different paradigms (such as knocking out EP with furosemide, hypoxia, variation in body temperature, etc.). By effectively knocking out different components in a selective way, can we move ahead in characterizing the OAE generation process?

---

$^8$This varies by roughly a factor of two in a given species. For example, along the radial direction in the chicken ear where the apical IHC surface area is larger in the SHCs relative to the THCs. Also, OHC radial spacing varies longitudinally along the cochlear length in mammals. We neglect the radial distance between IHC and OHC in the above value.

$^9$This notion is consistent with the observation made by Bekesy (1960) that larger cochleas, such as that of an elephant, appeared to be much more highly tuned.

$^{10}$Dimensions are covered in more detail in Ch.3. Briefly, the length of the sensory epithelium for each species is approximately: human $\sim 30-35$ mm, chicken $\sim 5-6$ mm, gecko $\sim 1.2-1.8$ mm and frog (AP) $\sim 0.6-1.2$ mm.
• Try un-mixing approaches in non-mammals. Are different emission components are present in their ear and if so, how do they vary with respect to changes in stimulus parameters (e.g. stimulus intensity)?

• Can non-mammalian emissions in species such as the gecko (where emission magnitudes are relatively large) help reveal how differences manifest across individuals? Insights in this regard could potentially could be of great benefit in furthering the potential use of OAEs for clinical diagnoses.

• Simultaneously measure emissions and other auditory measures. For example, how do nonlinearities in eOAE correlate to those observed in the cochlear microphonic? Or how do emission delays compare to those seen in the auditory nerve fiber (once synaptic and conduction delays have been accounted for)?

• How do non-mammalian spontaneous emissions correlate to eOAEs? Are these similarities in SOAE and SFOAE properties as has been demonstrated for mammals [Shera, 2003]?

• Based upon ANF tuning curves (which tell us how tuned the ear must be), can we estimate the delay associated with that response if we make some simplifying assumptions as to what the filter must be (e.g. minimum phase)? How do these delays correlate to eOAE delays?

• Can modeling efforts in the gecko ear (where the anatomy is relatively simpler compared to the mammalian cochlea) reveal insight into energy propagation though the inner ear and what role (if any) traveling waves play?

• Are there ways to exploit the bi-directional hair cell orientation patterns in the gecko and frog to learn something about hearing (for example, along the lines of Manley et al. 2001)?

• Do non-mammalian click-evoked OAEs show a degree of correlation to SFOAEs as has been observed in humans [Kalluri and Shera, 2007]?
7.5 Recapitulation

This thesis has taken a comparative look at evoked otoacoustic emissions in an array of vertebrates (human, bird, lizard and amphibian) that differ considerably in their auditory anatomy and physiology. We summarize our key findings as follows:

1. SFOAEs and DPOAEs are present in all species examined. Lower frequency emissions (below 3-4 kHz) are largest in species that lack both hair cell somatic motility and a tuned basilar membrane.

2. Emissions exhibit significant delays greater than 1 ms in all species. While these delays were largest in the human, the difference was not due to non-mammalian morphology (cats and guinea pigs exhibit delays similar to the chicken and are actually smaller than those of the frog). By examining the correlation between SFOAE delays and ANF frequency selectivity, it appears than in all species except the frog that the mechanisms associated with eOAE generation are the same as those giving rise to tuning in the inner ear.

3. Based upon observation of eOAE phase gradients, evidence suggests the presence of two different generation mechanisms in the human and chicken (and maybe the gecko). The case is not so clear in the gecko or frog, as the delineation between mechanisms depends upon the presence of a scaling-symmetric response, a feature that appears unlikely to be present in these two species.

4. A single-source model is insufficient to describe non-monotonic emission growth, even in ears with relatively simple anatomies. This observation motivates the need to consider OAEs as arising from a spatial summation from a distributed region.

The work presented here has raised a number of new questions and provides further motivation for future work that takes a comparative approach, integrating across the spectrum of knowledge (i.e. anatomy, mechanics, physiology, emissions, evoked potentials, etc.). Given the value of clinical applications utilizing OAEs and
current efforts towards realizing a new range of potential applications (such as a non-invasive monitor of intra-cranial pressure for head trauma victims or for diagnosing learning disabilities in children), I hope that future researchers will see the value of asking the question: *Why does the ear emit sound?*
Appendix A

Harmonic Oscillator

Our purpose here is to develop the equations of motion describing the harmonic oscillator [French, 1971]. This apparently simple and intuitive problem takes on a large degree of elegance within the context of how solid a foundation it sets for quantitatively describing a number of problems in acoustics and cochlear mechanics. Much of the terminology used in cochlear models (such as complex impedance\(^1\)) can be described in a straight-forward manner based upon first principles applied to the harmonic oscillator. Much of this quantitative machinery is applied to a simple cochlear model in the subsequent appendix section.

Fig. A-1 shows the physical basis for the harmonic oscillator\(^2\) A block of mass \(m\) sits on a rough surface (thereby introducing friction as a dissipative force). It is connected to a spring of stiffness \(k\) and can be driven by an external force \(F_{\text{ext}}\). For simplicity, we are only going to consider a linear system here\(^3\).

---

\(^1\)On the surface, many find it surprisingly unintuitive as to why values describing a physical problem are complex. This is merely an issue of convenience. Complex notation allows for us more compactly express two real and physical quantities of interest, magnitude and phase, with a single complex variable.

\(^2\)One could come up with countless variations of this simple arrangement, or venture from the realm of mechanics to electronics using an RLC circuit (we return to this analogy between electronics and mechanics a bit further on).

\(^3\)One can imagine adding nonlinearity into the system by a plethora of different ways, such as the spring stiffness depending upon the square of the displacements.
A.0.1 Undamped, Undriven Case - Simple Harmonic Oscillator

Let’s start with the simplest case, assuming no driving force or friction. At equilibrium, the block simply sits at rest. However, if there is some force displacing the block from rest, it will start to move. Combining Hooke’s law \( F = -kx \), where \( x \) is the position of the block relative to its resting position) with Newton’s second law \( F = ma \) where \( F \) is the force the block experiences and \( a \) is the acceleration of the block, or the second derivative of \( x \) with respect to time), we have

\[
F = ma = m\ddot{x} = -kx
\]

We use the dot notation to represent a derivative with respect to time (i.e. \( \ddot{x} \equiv \frac{d^2x}{dt^2} \)). We can rewrite the above equation as

\[
\ddot{x} + \frac{k}{m} x = 0
\]

This equation has the simple solution

\[
x(t) = A \cos (\omega_o t + \phi)
\]

where \( A \) is the amplitude of response and \( \phi \) is the phase at time \( t \). Plugging this solution in, we find that \( \omega_o = \sqrt{k/m} \). This value is called the natural frequency (or resonant frequency for reasons we will subsequently flesh out) and describes the rate
at which the block will oscillate back and forth. This value depends solely upon \( k \) and \( m \) and can be thought of as the preferred frequency the system oscillates at.

We can also describe the problem in terms of energy considerations. Suppose we displace the block a distance \( x \) from rest. The total work done \((U)\) is then given by

\[
U = \int_0^x F \cdot dx = -\int_0^x kx \cdot dx = -\frac{1}{2}kx^2
\]  

(A.4)

This represents work done to the spring (hence the negative sign), specifically the potential energy stored in the spring due to its displacement from rest. When let go, the block will move back towards its initial position and when \( x = 0 \), all this energy will have transferred from the spring to the block (appearing as kinetic energy given by \( K = \frac{1}{2}m\dot{x}^2 \)). For this simple case, the total energy of the system is given by \( E = K + U \) = constant. This energy term has two terms (one corresponding to the mass of the block and the other to the stiffness of the spring), each effectively storing energy that is being transferred back and forth between the two as time progresses. This notion of energy being stored and transferred between the various components of the system is important for our purposes here and we will return to it within the context of better understanding what \( \omega_o \) really represents.

### A.0.2 Driven Harmonic Oscillator

Now let's examine the problem when the system is being driven by an external sinusoidal force. We can express this as

\[
\ddot{x} + \frac{k}{m}x = F_o \cos \omega t
\]  

(A.5)

Where \( \omega \) is the frequency of the driving force. Remember that the system prefers to oscillate at \( \omega_o \), but we are forcing it to move at \( \omega \). Let us consider only steady-state solution (i.e. ignore any transient effects associated with start-up and that the block is oscillating solely at frequency \( \omega \)). Again, we can assume a

---

4Work has energy per unit time. We are only interested in the total work done, hence our expression has units of energy.
Figure A-2: Schematic showing response of the driven harmonic oscillator in its steady-state response to a sinusoidal driving force of frequency $\omega$. Both magnitude ($\kappa$) and phase ($\alpha$) are shown.

solution of the form $x(t) = B \cos(\omega t + \alpha)$. Plugging this in we obtain

$$-mw^2 B \cos \omega t + kB \cos \omega t = F_0 \cos \omega t$$  \hfill (A.6)

Upon solving for $B$, we can express our solution as

$$x(t) = \frac{F_0/m}{\omega^2 - \omega_o^2} \cos(\omega t + \alpha)$$  \hfill (A.7)

Letting $\kappa = |B|$, we plot the response magnitude and phase (of the block, relative to the resting position and driving force phase respectively) in Fig. A-2. This figure shows two very important concepts.

Resonance

First is the notion of resonance. Certain mechanical systems such as the harmonic oscillator\(^5\) oscillate at their own characteristic frequency when left on their own (i.e. the undriven case). However, when driven with an external sinusoidal force, energy is being put into the system with every cycle. Depending upon whether the driving frequency ($\omega$) is close to or far away from the characteristic frequency ($\omega_o$), the steady-state response can either be very large or very small. As indicated in Fig. A-2, when $\omega \approx \omega_o$, the response shoots off towards $\infty$, indicating that there effectively is no steady-state response when $\omega = \omega_o$. This happens because the system stores all the energy being put into it on a cycle-by-cycle basis, thereby continually increasing the amplitude of the response. This is exactly what resonance is: the ability of a

\(^5\)Here, we are referring to systems that have multiple modes of energy storage. Both the spring and mass can store energy (as potential and kinetic energies respectively).
mechanical system to store energy when being driven by an external sinusoidal force close to the system's characteristic frequency\(^6\). This is why the system's characteristic frequency is often referred to as the *resonant frequency*.

**Phase Shift**

As shown in Fig. A-2, when you change the driving frequency and pass through the resonant frequency, there is a shift in the phase of the response (of the block relative to the driving force) of 1/2 cycle. This can be reasoned out by looking at the extreme cases. When \(\omega\) is very small (well below \(\omega_0\)), the block will move in step with the driving force with amplitude \(\approx F/k\). The amplitude is large and the accelerations the block experiences are small (due to low frequency of oscillation). In this limit, we say the response is *stiffness dominated*. In the other limit, when \(\omega\) is very large, the amplitude becomes very small and the accelerations large. Here we would expect the motion to be 1/2 cycle out of phase from the driving force because the acceleration and displacement are 180° out of phase (*i.e.* accelerations are most positive and displacements are the most negative). In this region, we say the response is *mass dominated*.

We can also think about this by going back to Eq.A.6. The first term on the left-hand side (corresponding to the mass) will dominate when \(\omega\) is large. The second term (corresponding to the stiffness) will be large when displacements are big. These two are fundamentally 180° out of phase with each other (as indicated above). The resonant frequency represents the transition between the two.

### A.0.3 Damped Driven Oscillator

Let us move on to the more realistic case where some fricative force is present that removes energy from the system. We can write the equation of motion as

---

\(^6\)The term resonance is also used in other contexts which appear similar, but are somewhat fundamentally different. For example, resonance is used within the context of standing waves in a tube, which result from interference between forward and backward traveling waves.
\[ m\ddot{x} + kx + b\dot{x} = F_0 \cos\omega t \]  

(A.8)

If we let \( \gamma \equiv b/m \), this becomes

\[ \ddot{x} + \gamma \dot{x} + \omega_0^2 x = \frac{F_0}{m} \cos\omega t \]  

(A.9)

Now for convenience, we are going to assume the solution in the form of a complex exponential\(^7\). The solution will have the form

\[ z(t) = Ae^{-i(\omega t + \delta)} \]  

(A.10)

where \( x = \Re(z) \). Plugging in this solution, we obtain

\[ A(\omega) = \frac{F_0/m}{[(\omega_0^2 - \omega^2)^2 + (\gamma m)^2]^{1/2}} \]  

(A.11)

\[ \delta(\omega) = \frac{\gamma\omega}{\omega_0^2 - \omega^2} \]  

(A.12)

The resonant frequency of the system is shifted to a slightly smaller value due to the resistive term and is given by

\[ \omega_r = \sqrt{\frac{k}{m} - \frac{b^2}{4m^2}} \]  

(A.13)

Fig. A-3 shows a schematic of how the system's response with the effect of damping. The resistive term has the effect of keeping the amplitude finite at resonance and makes the phase shift more gradual. The damped harmonic oscillator is a versatile problem, with applications spanning a wide array of physical phenomena. The next few sections outline some useful quantities stemming from the oscillator which will be useful in the context of describing the physical processes occurring in the ear.

\(^7\)Euler's identity is the connection between sinusoids and complex exponentials. This notation allows for convenience in that exponentials are much easier to manipulate algebraically as well as vector visualization in the complex plane allows for more straight-forward inspection.
A.0.4 Quality Factor

It is apparent from Fig. A-3 that there is a peaked response in the magnitude of the system response about the resonant frequency. Describing the shape in this region is a useful quantity for understanding properties of the underlying physical system. When this peak is narrow, we say that the system is *highly tuned*. That is, only driving frequencies close to that of system's resonant frequency will produce a large response. Frequencies far away will do little to excite the system (meaning that very little energy is stored up). When the peak is broad, a large range of frequencies elicit a large response from the system and we say that it is *broadly tuned*.

This degree of tuning is typically quantified by a value called the *quality factor*, or \( Q \). This is a dimension-less quantity that is typically defined in a number of different ways. The most general definition is the ratio of \( 2\pi \) times the energy stored in the oscillator to the energy lost during a single complete oscillation. In the undriven, damped harmonic oscillator, \( Q \) is simply given as

\[
Q = \frac{\omega_0}{2\gamma} \tag{A.14}
\]

Without going into too much detail\(^8\), the higher \( Q \) is, the less of an effect the dissipative force has on the system and the more sharply tuned it is. It is clear from Eq.A.14 that \( Q \) depends upon the mass, stiffness and resistive terms.

Other definitions of \( Q \) are typically used. A particularly common one looks directly at the width of the measured response (such as in ANF tuning curves). Some set value

\(^8\)See French, 1971 for further discussion
below the maximum amplitude (such as 10 dB down) is set and the frequencies both above and below the resonant frequency are determined. The spacing between these two frequencies is then divided by the resonant frequency to obtain $Q$.

### A.0.5 Build-up Time

Our discussions up to this point have focused solely upon steady-state solutions. However, there is going to be some amount of time associated with energy going into the system as it builds up towards the steady-state response. Far away from resonance, the maximum response is small and the oscillator will quickly reach its steady-state. Near resonance, the response is much larger and it is going to take significantly longer for the driving force to put the energy into the system (one a cycle-by-cycle basis) which it is capable of storing. The more highly tuned an oscillator is, the longer it will take for the system build up its response. It is in this regard that highly resonant systems can be described as *slow*.

### A.0.6 Impedance

A very useful quantity for describing an oscillator's behavior is the *impedance*\(^9\). The complex impedance $Z$ (it will soon become apparent as to why it is complex) is defined as the ratio of the driving force over the resulting velocity of oscillation of the system\(^{10}\). Using the complex solution $z$ to the driven harmonic oscillator (Eq. A.10), we can rewrite our equation of motion as

$$F_{\text{ext}} = m\ddot{z} + b\dot{z} + kz$$  \hspace{1cm} (A.15)

We can express the impedance than as

$$Z = \frac{F_{\text{ext}}}{\dot{z}} = \frac{z \cdot (-m\omega^2 + k + i\omega b)}{z \cdot i\omega}$$  \hspace{1cm} (A.16)

---

\(^9\)The impedance is a typical quantity seen early on in an engineering curriculum, particularly for electrical engineers when describing electronic circuits.

\(^{10}\)In electronics, it is defined as the ratio of the driving voltage to the current. In acoustics, it is the ratio of the sound pressure to the resulting volume velocity.
Rewriting this, the complex impedance becomes

\[ Z = b + i \left[ m\omega - \frac{k}{\omega} \right] \] (A.17)

The real part of the impedance describes the *resistance*, specifically how the system gains or loses energy\(^{11}\). The imaginary part of the impedance describes the *reactance*, or how the system stores energy. Resonance occurs when the imaginary part of the impedance goes to zero. More specifically, the mass and stiffness terms cancel each other out, leading to a *maximally efficient transfer of energy* between the two. We can also express the impedance in terms of its magnitude and phase.

\[ |Z| = \sqrt{b^2 + \left( m\omega - \frac{k}{m} \right)^2} \] (A.18)

\[ \angle Z = \arctan \left( \frac{\omega m - k/m}{b} \right) \] (A.19)

When expressed in this form, \(|Z|\) and \(\angle Z\) reveal the magnitude and phase of oscillation relative to the driving force respectively. It is important to keep in mind that these are all *second-order systems*\(^{12}\). A second order system will have a single resonant peak for the steady-state solution, but higher order systems (e.g. ones with additional masses or springs in the case of the mechanical oscillator) can have more complicated dynamics.

### A.0.7 Analogy to Acoustics and Electronics

As mentioned earlier, the harmonic oscillator has a direct analogy to electronics and acoustics. The physical systems that would correlate would be an *RLC* circuit and a Helmholtz resonator respectively. In terms of energy storage elements, we have the mass and spring in the mechanical resonator. For the circuit, the analogs would be

---

\(^{11}\)The external driving force to putting energy into the system. However, the system itself may have its own means to put additional energy into itself, thereby increasing its responses. One straightforward way to do this would be some kind of feedback loop.

\(^{12}\)That is, there are two elements that energy can be *stored* in.
the inductor and capacitance. In the acoustic case, we have an acoustic mass (such as a lumped volume of air) and an acoustic compliance (a volume of air that can be compressed). Also, each case has resistive elements that can take away energy from the system.

With each of these cases, the overall all qualitative aspects appear very different, but the underlying quantitative descriptions are for the most part identical. It is important to keep in mind that these are all second-order systems$^{13}$. A second order system will have a single resonant peak for the steady-state solution, but higher order systems (e.g. ones with additional masses or springs in the case of the mechanical oscillator) can have more complicated dynamics.

A.0.8 Applying the Harmonic Oscillator to Problems in Hearing

The harmonic oscillator has provided us with a set of straight-forward quantitative tools that we can now apply towards problems in hearing (next section). It is insightful to take a step back though at this point and consider the purpose for taking this direction. The ear can be considered a collection a coupled highly tuned resonators. While any given element in the ear will be a higher-order system than a simple mass on a spring (and always consist of smaller sub-elements coupled together), the harmonic oscillator provides a very good starting point to think about the underlying mechanics of the ear when the appropriate assumptions and simplifications are made.

$^{13}$That is, there are two elements that energy can be stored in.
Appendix B

Traveling Waves in the Mammalian Cochlea - A Simple Model

We outline here a simple model for mechanics of the mammalian cochlea. The main purpose is to show how simple considerations with regard to the anatomy and physics lead to the presence of traveling waves, which can propagate energy throughout the cochlea in a manner analogous to that of an electric transmission line.

B.0.9 Transmission Line

We briefly spend time here to discuss properties of an electric transmission line (TL), which will be apparent when used for analogy later. As shown in Fig. B-1, a coaxial transmission line can be formed by a long, straight wire surrounded by a conducting sheath. An infinitesimal element of length Δz can be modeled by the circuit in Fig. B-2.

The space between the conductors will have a conductance $G$ depending upon what the medium is. Both wires effectively act as parallel-plate capacitors and will have the resulting capacitance $C$ between them. The resistance and inductance of the wires (per unit length) will be given by $R$ and $L$ respectively. In the loss-less condition, both $G$ and $R$ would identically be zero. Applying Kirchhoff’s laws, we
have

\[ v(z, t) - R \cdot \Delta z \cdot i(z, t) - L \cdot \Delta z \cdot \frac{\partial i(z, t)}{\partial t} - v(z + \Delta z, t) = 0 \]  \hspace{1cm} (B.1)

and

\[ i(z, t) - G \cdot \Delta z \cdot v(z + \Delta z, t) - C \cdot \Delta z \cdot \frac{\partial v(z + \Delta z, t)}{\partial t} - i(z + \Delta z, t) = 0 \]  \hspace{1cm} (B.2)

Upon rearranging and taking the limit as \( \Delta z \to 0 \), we obtain the general TL equations

\[ -\frac{\partial v}{\partial z} = R \cdot i(z, t) + L \frac{\partial i}{\partial t} \]  \hspace{1cm} (B.3)

\[ -\frac{\partial i}{\partial z} = G \cdot v(z, t) + C \frac{\partial v}{\partial t} \]  \hspace{1cm} (B.4)

Solutions to these equations will be of the form \( v(z, t) = R \Re[V(z)e^{j\omega t}] \) and \( i(z, t) = R \Re[I(z)e^{j\omega t}] \). Note that we use \( j \) to represent \( \sqrt{-1} \) to avoid confusion. This leads to

\[ \frac{\partial^2 V}{\partial z^2} = \gamma^2 V(z) \]  \hspace{1cm} (B.5)

and

\[ \frac{\partial^2 I}{\partial z^2} = \gamma^2 I(z) \]  \hspace{1cm} (B.6)

where \( \gamma = \sqrt{(R + j\omega L)(G + j\omega C)} \), the propagation constant. The \((R + j\omega L)\) term is generally referred to as the series impedance \( Z \) while the \((G + j\omega C)\) term is the shunt admittance \( Y \). Generally \( R \) is neglected and we have

\[ \gamma \approx j\omega \sqrt{LC \left[ 1 + \frac{G}{j\omega C} \right]} \]  \hspace{1cm} (B.7)

A general solution to the equations will be the sum of a forward and backward trav-
eling wave. For the voltage \( v(z, t) \), this will be given by

\[
v(z, t) = e^{j\omega t}[V^+(z) + V^-(z)] = e^{j\omega t}[V_o^+ e^{-\gamma z} + V_o^- e^{\gamma z}]
\]

which in turn solves the wave equation

\[
\frac{\partial^2 v}{\partial z^2} - \frac{1}{\mu^2} \frac{\partial^2 v}{\partial t^2} = 0
\]

where \( \mu = \omega/\gamma \) and describes the velocity of the traveling wave. It is insightful to stop here a moment and examine the result. Our analysis has shown that as the voltage changes across an infinitesimal cross-section of the coaxial line, this change will propagate outward with speed and attenuation defined by \( \gamma \) (it is the real part of \( \gamma \) which dictates how the wave amplitude changes). This derivation will provide insight by analogy as we develop the first stage of our cochlear model.

One other useful quantity to describe here is the characteristic impedance of the line, \( Z_o \). This is defined as the ratio between the voltage and the current. For an infinitely long line with only a forward-traveling wave, we have

\[
Z_o = \frac{V_o^+}{I_o^+} \approx \sqrt{\frac{j\omega L}{G + j\omega C}}
\]
Cochlear Geometry

One of the first assumptions we make is to unravel the snail-shaped cochlea into a straight tube. Previous calculations have shown that cochlear curvature can safely be neglected when considering the macromechanics because the radius of curvature is large compared to the spatial wavelength of the BM traveling waves [Steele and Zais, 1984]. Fig. B-3 shows a cross-section through the basal turn of a guinea pig. Separating the scala vestibuli (ScV) and scala media (DC) is a structure called Reissner's membrane. This avascular barrier serves to separate the ionic constituents of the fluids on each side, but does not add any impedance and can thus be neglected.

One important observation to note is the location of the basilar membrane (BM). As the bony projections juts out from the left (which encloses the nerves of the spiral ganglion), the BM stretches only a short distance across (from beneath the IHCs over to the right side wall). This is schematized in Fig. B-4, which shows the specific structures considered in our model. The wavelengths are large compared to the scala height and thus pressure can be considered uniform across the cross-section (long wavelength approximation). It is important to note that the organ of corti sits on top of the basilar membrane and surely plays an active role in the dynamics. We however here are going to focus solely on the macroscopic behavior and not the specifics of what is happening at the microscopic level.
Figure B-3: Micrograph showing anatomy of a cochlear turn. [ScV - scala vestibuli, ScT - scala tympani, MT - tectorial membrane, IHC - inner hair cell, OHC - outer hair cell, DC - scala media]

Figure B-4: Coordinate system for cross-section.
We will spend a moment here to describe the geometry of our problem such that we can make as many simplifying assumptions as possible for clarity and still retain the dynamical behavior as best we can. The distance \( x \) represents the distance along the length of the straightened cochlea from the stapes. At each point \( x \), we consider the cross-section as shown in the right of Fig. B-4. We have chosen a rectangular shape for the cochlear aqueduct and neglected the bony projection, assuming that the BM simply stretches all the way across. This will be an adequate assumption since we are chiefly concerned with how BM motion effects the change of each scalae area in the cross-section. The cross-section has width \( b(x) \) and a height \( h(x) \) (with a reference of 0 at the bony top of the scala vestibuli). The quantity \( d \) describes the position of the BM. There are two points to be mentioned in regard to the figure. First, the magnitude of \( d \) has been greatly exaggerated for clarity. The height of a cochlear aqueduct in the guinea pig varies from 1.8 mm in the basal end to around 0.6 mm in the apical end. However, BM amplitude ranges only from 0.05 nm at the lowest levels up to 10-20 nm at the highest levels. We have also shown the lateral sides of the BM fixed. This will introduce a constant factor in how the area changes which can be safely neglected from our calculations below. One important initial assumption which we will make is that \( b \) and \( h \) are independent of \( x \) and thus the cross-sectional area of a volume element is constant (denoted by \( A_v \) or \( A_t \)). For example, this simplifies the expressions \( \frac{\partial}{\partial t}[A_v(x) \cdot u_v(x, t)] \) to \( A_v \cdot \partial u_v(x, t)/\partial t \) where \( u_v(x, t) \) is the fluid velocity across the scala vestibuli. Since \( A_v(x) = b(x) \cdot d(x, t) \), this is equivalent to assuming

\[
\left| \frac{\partial u_v}{\partial x} \right| \gg \left| \frac{\partial d}{\partial x} \cdot \frac{\partial b}{\partial x} \right|
\]  

(B.11)

We have already seen that \( \partial d/\partial x \) will be quite small since the magnitude of the BM motion is relatively small.

Model

We assume that the fluid is incompressible. It is also assumed that cochlear fluids are not viscous, such that all energy entering the cochlea is dissipated by movement.
of the cochlear partition. The fluid dynamics are also assumed to be linear, since the wavelengths are relatively large and the velocities small so we only need to worry about inertial forces, [deBoer, 1996]. As the pressure difference across the element \( p(x + \Delta x, t) - p(x, t) \) is incremented, there will be a corresponding change in fluid flow through the element given by

\[
\Delta t \cdot A_v \cdot [p_v(x, t) - p_v(x + \Delta x, t)] = \rho \cdot \Delta x \cdot A_v \cdot [u_v(x, t + \Delta t) - u_v(x, t)] \tag{B.12}
\]

If we divide and take the limit, we obtain

\[
\frac{\partial p_v}{\partial x} = -\rho \frac{\partial u_v}{\partial t} \tag{B.13}
\]

The negative term arises because as \( \partial p/\partial x \) increases, the pressure will grow larger on the right compared to the left and thus fluid is less likely to flow through. Because of our chosen coordinate system, we similarly find for the scala tympani

\[
\frac{\partial p_t}{\partial x} = -\rho \frac{\partial u_t}{\partial t} \tag{B.14}
\]

Because the fluid is incompressible and the basilar membrane is flexible, we can derive an expression for fluid velocity to membrane displacement \( [d(x, t)] \). Again considering a cross-sectional slice of thickness \( \Delta x \), the change of mass inside the element will equal the product between the density and the change in volume

\[
\Delta t \cdot A_v \cdot \rho \cdot [u(x, t) - u(x + \Delta x, t)] = \Delta x \cdot \rho \cdot b \cdot [d(x + \Delta x, t) - d(x, t)] \tag{B.15}
\]

Which upon dividing and taking the limit, we obtain

\[
A_v \frac{\partial u_v}{\partial x} = -b \frac{\partial d}{\partial t} \tag{B.16}
\]

A consequence of our chosen coordinate system means that increasing \( d \) means a
smaller scala tympani cross-sectional area, so we have

\[ A_t \frac{\partial u_t}{\partial x} = b \frac{\partial d}{\partial t} \]  \hspace{1cm} (B.17)

We now turn towards the forces acting upon the cochlear partition. There are three forces in addition to the inertial force which will have an effect:

\[ F = F_{\text{stiffness}} + F_{\text{drag}} + F_{\text{pressure}} = \mu \cdot \Delta x \cdot \frac{\delta^2 d}{\delta t^2} \]  \hspace{1cm} (B.18)

where \( \mu \Delta x \) is the effective mass of section. The pressure force will be given by area \cdot pressure difference = \( \Delta x \cdot b \cdot [p_v(x, t) - p_t(x, t)] \). This is positive since \( p_v > p_t \) means that \( d \) will increase. The drag force is the product of the velocity and the effective damping coefficient, \( \alpha \cdot \Delta x \). This force will be negative since it works opposite the velocity when \( \partial d/\partial t > 0 \). Finally, the force due to stiffness will be the product of the displacement \( d \) and the effective stiffness \( \kappa \cdot \Delta x \). Note that we assume that \( \mu, \alpha \) and \( \kappa \) are all independent of \( x \). This yields the following equation of motion

\[ b \cdot [p_v(x, t) - p_t(x, t)] = \mu \cdot \frac{\delta^2 d(x, t)}{\delta t^2} + \alpha \cdot \frac{\delta d(x, t)}{\delta t} + \kappa \cdot d(x, t) \]  \hspace{1cm} (B.19)

We now wish to show that \( A_{cs} \cdot (u_v(x, t) + u_t(x, t)) \) is constant with respect to position and equal to zero (where \( A_{cs} = A_v + A_t \)). Consider the derivative

\[ \frac{\partial}{\partial x} A_{cs} \cdot [u_v(x, t) + u_t(x, t)] \]  \hspace{1cm} (B.20)

Which from our previous result, we can express this as

\[ A_{cs} \cdot \left[ -b \frac{\partial d}{\partial t} + b \frac{\partial d}{\partial t} \right] = 0 \]  \hspace{1cm} (B.21)

Thus \( u_v(x, t) + u_t(x, t) \) is constant at any place along the cochlea. At the base, any fluid change at the oval window (i.e. at the stapes footplate) must be equal and opposite in magnitude of the change at the round window. So we have \( u_v(0, t) = -u_t(0, t) \) for
all \( t \). This then leads to

\[
u_v(x, t) = -u_t(x, t) \quad (B.22)
\]

This result can now be used to consider \( \partial / \partial x [p_v(x, t) + p_t(x, t)] \), which will identically be 0. So we obtain

\[
p_v(x, t) + p_t(x, t) = \alpha \quad (B.23)
\]

where \( \alpha \) is a constant.

We can now make some simplifications by introducing the variables \( p \equiv p_v - p_t \) and \( u \equiv A_{cs}(u_v - u_t)/2 \). We assume that the time dependence of both \( p \) and \( u \) is sinusoidal and each can be expressed as \( p(x, t) = \Re[P(x, \omega)e^{j\omega t}] \) and \( u(x, t) = \Re[U(x, \omega)e^{j\omega t}] \) (\( \omega \) is \( 2\pi \) times the frequency of stimulation). We can now write

\[
\frac{\partial p}{\partial x} = \frac{\partial p_v}{\partial x} - \frac{\partial p_t}{\partial x} = -\rho \left( \frac{\partial u_v}{\partial t} - \frac{\partial u_t}{\partial t} \right) \quad (B.24)
\]

which leads to

\[
\frac{\partial p}{\partial x} = -\frac{2\rho}{A_{cs}} \frac{\partial u}{\partial t} \quad (B.25)
\]

But since \( \partial p/\partial x = e^{j\omega t}\partial P/\partial x \) and \( \partial u/\partial t = i\omega e^{j\omega t}U \), we have

\[
\Rightarrow \frac{\partial P}{\partial x} = -\frac{2\rho}{A_{cs}} i\omega U = -ZU \quad (B.26)
\]

We can carry out a similar analysis for \( u \) to obtain an expression for \( \partial U/\partial x \):

\[
\frac{\partial p}{\partial x} = \frac{A_{cs}}{2} \left[ -\frac{b}{A_{cs}} \left( \frac{\partial d}{\partial t} + \frac{\partial d}{\partial t} \right) \right] = -b \frac{\partial d}{\partial t} \quad (B.27)
\]

Rearranging terms,

\[
\frac{\partial d}{\partial t} = -b \frac{\partial u}{\partial x} = -\frac{e^{j\omega t}}{b} \frac{\partial U}{\partial x} \quad (B.28)
\]

We can now integrate this to get an expression for \( d(t) \).

\[
d(t) = \int -\frac{e^{j\omega t}}{b} \frac{\partial U}{\partial x} dt = -\frac{e^{j\omega t}}{j\omega b} \frac{\partial U}{\partial x} \quad (B.29)
\]
This result can then be plugged back into the equation of motion (eqn. B.19) as follows

\[ b p(x, t) = \frac{e^{j\omega t}}{j\omega b} \frac{\partial}{\partial x} \left[ -\mu^2 + \alpha i\omega + \kappa \right] = b \cdot P e^{j\omega t} \]  

(B.30)

Upon rearranging terms, we can write this as

\[ \frac{\partial U}{\partial x} = -\frac{P}{j\omega b \frac{\mu}{\kappa} + \frac{\alpha}{\kappa} + \frac{1}{j\omega b^2}} = -YP \]  

(B.31)

We can now use these results to develop a single expression for each \( P \) and \( U \). We first assume here that \( Z \) and \( Y \) have no \( x \) dependance. Plugging B.26 into B.31, we obtain

\[ \frac{\partial^2 P}{\partial x^2} + \frac{1}{\ell^2} P = 0 \]  

(B.32)

where \( \ell = j\sqrt{1/ZY} \). This is just an expression for a traveling wave (the time dependance has already been effectively accounted for and this is the basic wave equation). So our result so far does allow for traveling waves (as we had hoped for). We get an analogous equation for \( U \). Since we assumed \( \ell = \ell(\omega) \) (and not \( \ell(x, \omega) \)), we can easily solve this to get \( p(x, t) \). Allowing for both forward and backward traveling waves, we have

\[ p(x, t) = P_1 e^{i(\lambda x/\ell + \omega t)} + P_2 e^{i(-\lambda x/\ell + \omega t)} \]  

(B.33)

since \( P(x, \omega) = P_1 e^{ix/\ell} + P_2 e^{-ix/\ell} \). The wavelength of the traveling wave will be given by \( \lambda = 2\pi\ell \). When \( \lambda \) is purely real, this wave will travel decrement-free. However, imaginary parts of \( \lambda \) will cause a decay in the amplitude of the wave as it travels along (due to \( Z \) and \( Y \), whose real and imaginary parts will always be positive for our system).

**Comparison to Transmission Line Result**

So we have developed a model based upon simple physical descriptions of our system which allows for traveling waves to propagate along the BM. Time was spent initially developing the dynamical equations for a transmission line and we see here that
our two results are analogous. Consider equations (B.5) and (B.6) in comparison to equations (B.26) and (B.31). Comparison of a simple case shows that these two pairs of equations are equivalent to each other. In the loss-less TL case, both $R$ and $G$ are identically 0. Thus $\gamma$ is purely imaginary such that $\gamma^2$ is real and negative. Now consider the loss-less cochlea, where there is no damping or inertial terms (i.e. $\mu, \alpha = 0$). So we must have

$$\gamma^2 = -\frac{1}{\ell^2} \tag{B.34}$$

In comparing these two, we get

$$C \cdot L = \frac{v^2}{\kappa} \cdot \frac{2\rho}{A_{cs}} \tag{B.35}$$

So the capacitive term in the transmission line is analogous to the restorative spring term in the cochlea while the electrical inductance term corresponds to a mass term. It is also easy to see that the resistive term $R$ carries over to the damping term $\alpha$.

It is interesting that our method for determining a model of the cochlea produced something which so closely quantitatively parallels the treatment of an electrical transmission line. A benefit gained is that results and insights into transmission line dynamics (which have been quite thoroughly developed) can potentially be used to provide insight back to the cochlear problem. For example, suppose we wanted to add a single pressure source at some point $x_0$ along the cochlea (the reason will become more apparent later on when we discuss amplification). We can think of this being analogous to adding a single voltage source along the TL and being able to use a similar set of quantitative tools to describe the resulting behavior.

**WKB Approximation**

Unfortunately, some of the made assumptions were not too good for what happens in a real cochlea. For example, we know that BM stiffness and width changes along the length of the cochlea. So both $\alpha$ and $b$ are functions of $x$ and thus $\lambda$ is as well. Since the wavelength changes along the cochlea, we will need to take a modified approach to obtain an expression for $p(x, t)$ [150].
In 1926, three authors (Wentzel, Kramers and Brillouin) independently published papers which approximated a wave function as an oscillatory wave depending upon a phase integral. This allowed for obtaining approximated solutions to the time-independent Schrödinger equation in one dimension [43]. If a particle with energy $E$ is moving in a field of constant potential $V$ where $E > V$, the wave function is given by

$$\psi(x) = Ae^{\pm jkx}$$

(B.36)

where $k \equiv \sqrt{2m(E - V)/\hbar}$. This is just a wave with constant amplitude $A$ and wavelength $\lambda = 2\pi/k$. Now what if $V$ was not constant, but varied slowly in comparison to $\lambda$ (so that over a region containing many wavelengths, $V$ is approximately constant)? The WKB approximation states that $\psi$ will remain practically sinusoidal, but that the wavelength and amplitude will vary slowly with $x$. If we write $k$ as $k(x) \equiv \sqrt{2m[E - V(x)]/\hbar}$, we get an approximate solution of the form [43]

$$\psi(x) \cong \frac{C}{\sqrt{k(x)}} e^{\frac{j}{k(x)} \int k(x)dx}$$

(B.37)

What is so useful about this result is that it can be applied to many other types of problems, such as describing light in a medium where the index of refraction is changing slowly or in our case, wave traveling along the cochlea.

If we assume that BM stiffness and scalae area changes rather gradually, then the cochlea could act as a uniform transmission line locally. Considering only forward traveling waves, our original solution had $P(x, \omega)$ in the form $P_0 e^{\pm jx/\lambda}$. So we can try a trial solution of the form

$$P(x, \omega) = A(x)e^{-j \int_0^x dx' / \ell(x', \omega)}$$

(B.38)

which is analogous to going from (B.36) to (B.37) in the quantum problem. We can substitute this back into the wave equation (B.32) to check its validity and determine the function $A(x)$. Since we assume that things vary gradually, we can assume that
\( \delta^2 A / \partial x^2 \) is negligible and we have

\[
\left[- \frac{2j}{\ell(x)} \frac{dA}{dx} - \frac{1}{\ell(x)^2} + \frac{j}{\ell(x)^2} \frac{d\ell}{dx}\right] + \frac{1}{\ell(x)^2} = 0 \quad (B.39)
\]

This yields

\[
\frac{1}{\ell(x)} \frac{d\ell}{dx} = 2 \frac{1}{A(x)} \frac{dA}{dx} \quad (B.40)
\]

This is easily solved and shows that the amplitude will be proportional to the square root of the wavelength. Also note that propagation velocity is now going to vary as well. We can now derive a specific expression for \( \ell(x, \omega) \), expanding out \( Z \) and \( Y \). This yields

\[
\ell(x, \omega) = j \left[ \frac{j\omega L + R + 1/j\omega C}{j\omega M} \right]^{1/2} \quad (B.41)
\]

Now we introduce some new variables. Let \( \omega_r(x) \equiv 1/\sqrt{LC} \), \( \beta(x, \omega) \equiv \omega / \omega_r(x) \), \( \delta \equiv \omega_r(x)RC \) and \( N \equiv (l/4)\sqrt{M/L} \). Here, \( \omega_r(x) \) is just the local resonant frequency of a given spot along the length of the BM. In a human ear, this ranges from about 20 kHz at \( x = 0 \) to 20 Hz at \( x = 35 \) mm in a logarithmic fashion. The variable \( l \) is a constant which will be further described in the following section. With these variables, we can rewrite the wavelength as

\[
\ell(x, \omega) = \frac{l}{4N} \frac{(1 - \beta^2 + j\delta\beta)^{1/2}}{\beta} \quad (B.42)
\]

So we are thus able to now write the wavelength as a function of \( \ell(\omega/\omega_r(x)) \). The pressure variation at a given spot \( x_o \) depends upon the relationship between the stimulation frequency \( \omega \) and the local resonant frequency \( \omega_r(x_o) \). We can now use this expression for the wavelength to derive the BM transfer function.

**Transfer Function**

Making measurements of BM motion is difficult due to the cochlea being encased in a bony structure. Removing the bone to expose the cochlear aqueduct runs the risk of greatly degrading the viability of a living cochlea. However, careful measurements
have been made and take the form of comparing BM motion at different points along
the cochlear length to that of the stapes. Thus we will develop an expression for the
complex transfer function, given by

\[ T(x, \omega) = \frac{\text{BM velocity}}{\text{stapes velocity}} \]  \hspace{1cm} (B.43)

We can now use the WKB approximation (B.38) and our expression for the wavelength
(B.42) to find \( T(x, \omega) \). BM velocity is just given by \( \partial d/\partial t \), which we can use equations
(B.29), (B.31), (B.38) and (B.42) to obtain. Stapes velocity will be the scala vestibuli
volume velocity at \( x = 0 \) divided by the area of the stapes (which we assume to just
be \( A_v \)). Expanding this out, we have

\[ \text{BM velocity} = \frac{e^{i \omega t}}{b} \ell^2(x, \omega) Y e^{-i \int_0^x dx'/\ell(x', \omega)} \]  \hspace{1cm} (B.44)

We will get something similar for the stapes velocity, except that this will be evaluated
solely at \( x = 0 \) and thus no integral will appear in the denominator. Thus it becomes
apparent that we are going to need to solve for a specific expression of equation (B.38)
using (B.42). For convenience, we will assume that \( N \) and \( l \) are constant with respect
to \( x \). We need to expand out our definition for \( \omega_r \) since this is a function of position.
Remember that this represents the local resonant frequency at a given point along the
BM. Experimental evidence has shown that the tonotopic organization of the cochlea
is logarithmic with higher frequencies at the basal end. Thus we can characterize this
by

\[ \omega_r(x) = \omega_{max} e^{-x/l} \]  \hspace{1cm} (B.45)

where \( \omega_{max} \) is the highest frequency which can be heard (at \( x = 0 \)) and \( l \) is the distance
along the BM which this characteristic resonant frequency changes by a factor of \( e \n(and is independent of \( x \)). Since we have \( \beta = \omega/\omega_r \), we have

\[ dx = \ell \frac{d\beta}{\beta} \]  \hspace{1cm} (B.46)
Using this substitution, the integral to be evaluated can be written as

\[ 4N \int_0^x \frac{d\beta}{[1 - \beta^2 + j\delta\beta]^{1/2}} \] (B.47)

This can be simplified even further by the substitution \( u(x, \omega) = \beta(x, \omega) - j\delta/2 \). If we assume that \( \delta^2 \) term is negligible (the validity of which is physiologically reasonable), the expression reduces to

\[ \int_0^x \frac{du}{\sqrt{1 - u^2}} = \sin^{-1} u(x) - \sin^{-1} u(0) \] (B.48)

Putting everything together, we can write a closed expression for the transfer function, given by [151]

\[ T(x, \omega) \approx T_0 j\beta(x, \omega) \left[ \frac{\omega_{\text{max}}}{\omega_T(x)} \right] e^{i4N\left\{ \sin^{-1}[\beta(x, \omega) - j\delta/2] - \sin^{-1}[\beta(0, \omega) - j\delta/2] \right\}} \frac{1 - \beta^2(x, \omega) + j\delta\beta(x, \omega)^{3/4}}{[1 - \beta^2(x, \omega) + j\delta\beta(x, \omega)]^{3/4}} \] (B.49)

Here, \( T_0 \) is a constant real coefficient. This equation now gives us a means to directly test the model, as we will show in the following section.

### B.0.11 Model Comparison to Experimental Data

In the 1930’s, Georg von Bekesy made important measurements in human cadaver temporal bones that showed the existence of traveling waves along the BM (for his extensive work in hearing, he was subsequently awarded the Nobel prize for Physiology or Medicine in 1961). It was however not until 1971 when measurements made by Bill Rhode in a living squirrel monkey using the Mossbauer technique showed strongly non-linear behavior such as compressive growth [Rhode, 1971]. This indicates some source of nonlinearity inside living ears (haven taken into account the nonlinear nature of the Mossbauer technique). This data from Rhode was used as the basis for an inverse model proposed by the physicist George Zweig that argues that the cochlear impedance exhibits a region where the real part becomes negative (i.e. amplification) [Zweig, 1991]. It is these results from Rhode which we will use to compare the model
Figure B-5: Data obtained by Rhode from a single squirrel monkey. The transfer function $T$ here is shown as BM motion relative to the stapes. Imaging was done at one specific place while frequency was swept. [Rhode, 1971]

to. Data from a single monkey is shown in Fig. B-5.

The transfer function of the model has the following free parameters which we will have to adjust to get the best possible fit (as well as constrain to stay within physiological bounds if the model is to work!): $T_0$, $N$ and $\delta$. It appears that the resonant frequency $\omega_r$ (or center frequency, CF) at the place the measurements were made is approximately 7.8 kHz. Using the mapping given by equation (B.45) where $\omega_{\text{max}}/2\pi \approx 50$ kHz and $l \approx 5$ mm for the squirrel monkey, we have $x \approx 9$ mm. Reasonable physiological values for $N$ and $T_0$ have an order of magnitude of $10^0$ and $\delta$ is around $10^{-1}$. Optimizing these values within that range, we obtain a fit to the data as shown in Fig. B-6. The fit is quite reasonable considering the number of simplifying assumption which were made. However, higher frequencies are not well covered as the amplitude falls off much more quickly. Additionally, there is a dip in the phase not seen in the actual data. We can also make a plot of the wavelength $\lambda$ (B.42). Figure B-7 has two parts: one at $x = 9.25$ mm and frequency varied (top) and the other with frequency fixed at $\omega = 7.8$ Hz. The real part of $\lambda$ gives the physical wavelength while the imaginary part give the inverse of the damping constant. Below the CF, $\lambda$ is purely real and the result is a traveling wave propagating unperturbed.
Figure B-6: Fit to Rhode's data using the transfer function given in equation (B.49). The model fit is the bold solid line. The fitting parameters used are $\delta = 1/40, T_0 = 2, N = 5$ and $x = 9.25$ mm.

Figure B-7: A. The real and imaginary parts of the wavelength $\lambda(x, \omega)$. Here, $x$ is held constant at 9.25 mm. B Similar to above, but $\omega$ is held constant at the CF $2\pi \cdot 7.8$ kHz.
However, as you come closer to the CF, the imaginary part starts to increase from zero, causing the wave to damp out. Above the CF, $\lambda$ is purely imaginary and there is only damping and no wave propagation.

Presumably, cochlear amplification is most prevalent at low sound intensities. At higher levels, this amplifier is effectively turned down (or washed out), thereby causing compressive (i.e. nonlinear) growth observed in a number of different measurements. Rhode's data presented in Fig. B-5 were likely taken at higher intensities (70-90 dB SPL). The data has subsequently been extrapolated to lower levels near threshold [Zweig, 1991], which we can further use to test our model. This fit to the new data is shown in Fig. B-8.

The fit is decent, though we had to push $\delta$ beyond the physiological range. Since this term relates to damping combined with the model's peak being much smaller at CF, this implies that there is some active mechanism which is adding energy into the system. It should also be noted that the peak in the model is much narrower than the experimental data. This may stem from the fact that when using the WKB approximation, we assume that the wavelength and amplitude vary slowly compared to $x$. This is probably not true in the region around the CF where amplification is likely to occur.
Conclusions

While we have seen that this simple model works reasonably well in comparison to experimental data, the model is insufficient to capture all the features observed in the ear. Most notably (albeit not shown here), the model is entirely linear, which we know the ear not to be. And while the model can account for responses in a dead ear, it does not exhibit the features seen in a living, healthy ear (such as sharp tuning and large phase delays).

A logical next step in the model would be to include features which would account for active amplification mechanisms that have been proposed to exist in the ear. These could take the form of pressure sources distributed along the length of the cochlea. This effectively could describe feedback mechanisms which could arise from possible generators such as OHC motility or stereociliary bundle/transduction channel dynamics. A good starting point would be to model these with delta functions and work out the transfer function with these additional terms to see if the model predictions better match the data. A further step would be to try to model these active elements more as physiological entities rather than lumped elements. This would allow for the model to make better predictions about what the possible amplification mechanisms are and the significance of their roles.

Differences between data and model aside, this simple view has given us considerable insight into the function of the inner ear. In responses to sound pressure (propagated through the external and middle ears to the cochlea), traveling waves are set up that propagate energy thru the cochlea to its characteristic place (analogous to an electric transmission line). This conclusion stems from a few simple physical laws as well as basic knowledge about mammalian cochlear anatomy (i.e. a flexible membrane with graded mass/stiffness, etc.). It is interesting to ask whether this simple physical description is valid in the non-mammalian ear, where inner ear anatomy differs greatly.

One final point to emphasize here: the model described here, while fairly simple, borrowed from a number of different flavors of science such as electro-dynamics,
quantum mechanics and physiology. It is in this regard that the study of hearing (or most of biology for that matter) is truly a multi-disciplinary science, making it so engaging and challenging.
Appendix C

Using Phase Gradients to Estimate Time Delays

C.1 Gradients in the Steady-State Condition

Biological activity takes time, creating a *time delay* between the stimulation and the subsequent response. As is apparent throughout this thesis, it is often quite informative to determine the length of these delays in order to better understand the underlying system. The basic problem is schematized in Fig. C-1. Part of the problem is that measuring transient responses (i.e. specific time of onset) can be difficult. Thus, the time delay is difficult to ascertain looking solely for the onset of the response. One method to circumnavigate this difficulty is to consider the system once it has achieved steady-state and make use of the response phase sensitivity to input frequency to determine the delay.

We present a sinusoidal stimulus to the system which always has zero phase at time $t = 0$. This is the *input* to the system. The *output* is the system's response. This is delayed by time $\tau_1$ and may include some transient onset effects. Once in steady-state at a time time $\tau_2$ (this value is arbitrary), we measure the input and

---

1See Whitehead et al. (1996) or Meenderink and Narins (2006) for further discussion in this regard. In spite of it, these authors were able to overcome it and show for human DPOAEs and frog SFOAEs a correlation between time delays measured directly at onset and using the steady-state phase gradient method outlined here.
output phase ($\theta_2$ and $\phi_2$ respectively). $\theta_2$ serves as the reference from which we can now compute the phase delay as shown in Fig. C-2.

As the frequency ($f$) of the input stimulus is swept, we have the relationship

$$\tau_1 = \frac{\partial(\theta_2 - \phi_2)}{\partial f}$$  \hspace{1cm} (C.1)

Thus the time delay is effectively revealed by the phase gradient. This technique for estimating time delays is commonly used in a wide range of applications. However, caution is needed within the context of application towards OAEs. For example, the phase gradient for $2f_1 - f_2$ in mammals when the primaries are swept using a constant ratio ($f_2/f_1$) greater than $\sim 1.08$ is close to zero. A naive interpretation would be that the emissions appear almost instantaneously relative to the evoking stimulus. However, the physical geometry needs to be considered as this frequency independence of the phase likely arises due to the scaling symmetric nature of traveling waves in the cochlea [Shera and Guinan, 1999].
Figure C-2: Schematic showing how frequency dependence of response phase can reveal time delay. The slope of the phase versus frequency function (phase gradient) is the quantity of interest.

Figure C-3: Schematic showing build-up of response of a second order system (e.g. a harmonic oscillator) when driven by an external sinusoidal stimulus with frequency $\omega$. After some time, the system reaches its steady-state response. The more sharply tuned the resonator is (and the closer the driving frequency is to the system's natural frequency), the longer the build-up time will be. The steady-state response amplitude $[A(\infty)]$ also varies with respect to $\omega$. 
C.1.1 Delays Associated With Tuning

One goal of measuring OAE phase gradients is to try to understand something about tuning (this is described in Ch.3). To try to understand the connection, we examine a simple case: a second order system (e.g. a harmonic oscillator). A second order system driven by an external (sinusoidal) force can be described by

\[ m\ddot{x} + b\dot{x} + kx = A\cos\omega t \]  

(C.2)

and has the natural frequency \( \omega_n \approx \sqrt{k/m}. \) Assuming the driving force starts at time \( t = 0, \) the response of the system will exhibit a build-up that goes as

\[ x(t) = A(\infty) \left[ 1 - e^{-t/\tau} \right] \]  

(C.3)

Here, \( A(\infty) \) is the steady-state response amplitude and depends upon the different system parameters. This build-up time is schematized in Fig. C-3. The system stores energy until the response eventually saturates, being governed by the time constant \( \tau. \) This time constant is simply given as \( \tau = m/b, \) being reciprocally related to the amount of damping present in the system. We can express this time constant in another form, as

\[ \tau = Q/\omega_n \]  

(C.4)

where \( Q \) is the quality factor that describes how sharply ‘tuned’ the system is (it is defined as the ratio between the amount of energy lost to the energy stored over one cycle of oscillation while in steady-state). The larger \( Q \) is, the longer the build-up time will be.

As shown in Fig. C-4, high-\( Q \) systems exhibit very large (and narrowly-tuned) responses when driven close to the natural frequency. Low-\( Q \) systems (i.e. greater damping) store less energy and exhibit broader frequency responses. For a fixed \( Q, \) a bandwidth can be defined for the system as the frequency region bounded both about the peak by a fixed decrease (say 10 dB) in the response amplitude from the maximum. The larger the bandwidth, the smaller \( Q \) is (Fig. C-4). Looking at the phase response
of the system (relative to the driving force), the system exhibits a 1/2 cycle phase change as the driving frequency ($\omega$) varies through the system's natural frequency.

Now depending on the size of $Q$, the transition in the phase can be gradual (small $Q$) or sudden (large $Q$). The speed at which the transition occurs can be quantified by the dimensionless parameter $N$, the product of the slope of the phase curve at $\omega_0$ (i.e. the phase gradient, or group delay) and the system's resonant frequency. It can be shown [French, 1971] that

$$Q \propto N$$  \hspace{1cm} (C.5)

So there is a direct connection between the phase gradient and the degree of tuning, at least for a second order system.
Appendix D

Correcting for I/O Delays in Phase Measurements of eOAEs

D.1 Overview

The purpose of this appendix is to provide an overview of how certain corrections needed to be made in the OAE phase measurement\(^1\). This was necessary in order to account for the delays associated with both digital to analog (D/A) and analog to digital (A/D) conversions. Both these steps introduce significant time delays and need to be accounted for and subtracted out. The delays stem from the actual hardware on the sound card that does the conversions. They become slower as well as bit-depth increases. Thus for 24-bit sample conversion (as was used for OAE recordings), the delays can become significantly longer than those stemming for the OAE generation mechanisms.

While the total amount of delay associated with both these steps was easily characterized, as described below (typically on the order of ~105 samples), the main issue stemmed from the lack of knowledge of the relative contributions between the D/A and A/D delays. This was of significant consequence for DPOAE measurements, where knowledge of the primary phases at the ear canal is essential and required a reference phase that was generated using a diode circuit. SFOAE phase correction

\(^1\)The quantitative basis for these corrections derives from notes of C. Shera.
was relatively much more simple and merely required knowledge of the total delay.

A schematic showing an overview of the system and the key variables is shown in Fig. D-1. The quantity $\phi$ represents the phase of the particular signal. The subscript indicates the particular frequency (e.g. $\phi_1$ is the phase of the primary $f_1$). The superscript indicates the particular stages occurring between the delays. Quantities with $A$ represent the stimuli as defined in the software that are to be sent out from the computer to drive the probe. Those with $B$ represent the state at the ear canal, as delivered via the earphones. Note that an additional term appears here, one associated with the emission (for simplicity, we specifically show the term $\phi_{dp}$ associated with the distortion product $2f_1 - f_2$) and the other ($\phi_x$) associated with the actual generation mechanism. Lastly, terms with a superscript $C$ represent those actually measured back at the computer (after the A/D conversion) via the mic response.

The subsequent sections here outline the specific computations performed to correct the OAE phase values. We attempt to be as complete as possible here in order to allow for the method to be reproducible for those who may need to take this approach if their relative measurement system delays are not well defined.\footnote{Some analog-digital conversion cards have clearly specified delays for the various stages of I/O. In that case, the methods described here will be of use, but employing a reference phase via a diode is likely unnecessary.}
D.2 Phase Correction

D.2.1 Total System Delay

The total delay associated with D/A and A/D conversion ($\tau_{sys}$) could be found by simply hooking the output and input of the soundcard together, totally bypassing the OAE measurement system (we call this loop-back mode). An underlying assumption made throughout the entirety of this appendix is that the soundcard is capable of simultaneous input and output (i.e. synchronized I/O). The total delay can be quantified in one of two ways. First, the most straightforward means, is to examine the delay in the time domain using a simple signal such as a step or sinusoid. Second, the delay can be quantified in the frequency domain using a sinusoid and varying frequency (this method is described in some detail in Appendix C). Using this latter method, the delay is expressed (in polar format) as a phase value ($\phi_{sys}(f)$) that is a function of frequency. The connection between the two is

$$\phi_{sys}(f) = \tau_{sys} \cdot f$$  \hfill (D.1)

such that $\phi_{sys}(f)$ is the phase of a sinusoid of frequency $f$ delayed by time $\tau_{sys}$ (and has a phase of zero at $t = 0$; see Fig. D-2). For the most part, all delays below are expressed in polar format, with the particular frequency of interest appearing the subscript.

The second method described above is method we employ here as it derives directly from the method used to calibrate the earphones (EPs) when the probe is coupled in situ to the ear. Calibration is done using a flat spectrum, random phase gaussian noise signal. We call this signal $N^A(f)$, which is the complex Fourier amplitude of the signal (which conversely can just as easily be expressed in the time domain)$^3$. Upon presentation, the signal $N^C(f)$ is subsequently measured. The ratio is then taken, such that

$$C(f) = N(f)^C / N(f)^A$$  \hfill (D.2)

$^3$The complex notation can be expressed in one of two ways, either as a real and imaginary number or as a magnitude and phase. We use the latter.
where $C(f)$ is the calibration spectrum containing the information regarding the total system delay$^4$. Acoustic delays at the ear canal are also included here, and while relatively negligible compared to the soundcard delays, still need to be accounted for as described below.

**D.2.2 SFOAEs**

For a description of the suppression paradigm employed to measure SFOAEs, see Ch.2. Briefly, two tones are used ($f_p$ and $f_s$), with the emission occurring at the probe (i.e. stimulus) frequency $f_p$. With the total delay known from the calibration as described above, the corresponding phase correction for SFOAEs can easily be computed. First, we account for the delay between the start of the acquisition window (where the phase is defined to be zero) and the actual time window extracted for the segment of interest, called $\tau_{\text{win}}$. For a any given frequency, this is given by $\phi_{\text{win}}(f)$. Thus, we effectively set $\phi^A_p$ to zero, where the subscript $p$ is indicative of the probe frequency. This is schematized in Fig. D-2. This delay can be accounted for either before (when defining the stimuli to be output to the EPs) or after (once the mic response has been digitized) the actual measurement.

Next, the corrected probe phase is then computed as

$$\phi_p^{\text{corr}} = \phi_p^C - C_p$$  \hspace{1cm} (D.3)

where $C_p$ is just the in situ calibration as described above. Once the phase correction is done for both the probe-alone and probe+suppressor segments, the complex subtraction is performed to obtain the emission phase $\phi_{\text{SFOAE}}$ [Ch.2, also described in detail in Shera and Guinan (1999)].

Accounting just for the total delay is sufficient for SFOAEs. This is because the (relevant) delay occurs at a single frequency ($f_p$) and can simply be subtracted off by accounting for the total roundtrip. More specifically, the phase of the primary tone

$^4$When taking the ratio of complex values expressed in polar form, the amplitudes divide and the phases subtract.
Figure D-2: Schematic showing delay between start of acquisition window and extracted time waveform that needs to be accounted for. The delay $\tau_{\text{win}}$ is purposely specified so to let the system come to a steady-state response. The total acquisition window is presented repeatedly presented to allow for averaging (in order to minimize noise effects). For both SFOAE and DPOAE measurements, the total acquisition buffer is typically on the order of ~0.25-0.5 s long and multiple segments (shaded region) are extracted, each with their own unique $\tau_{\text{win}}$ value (see Ch.2).

at the ear canal (which will vary with frequency) is unimportant here. As described in the following section, this is not the case for DPOAEs where the relative phase of the primaries at the ear canal needs to be known to ascertain the phase at the distortion product frequency.

D.2.3 DPOAEs

As alluded to above, the phase correction for the delays in DPOAEs is a bit more complicated and requires an additional reference phase. This is done by running the DPOAE measurement paradigm while the system is connected to a diode, which presumably acts as an instantaneous nonlinearity. Running across a diode produces a rich spectrum of distortion products which can be saved to file and whose phase (for a particular distortion product, given the identical choice of frequency parameters used in the DPOAE measurements) can then be used as described below. The ear canal as shown in Fig. D-1 is simply replaced with a single diode and two sinusoids
are output simultaneously on a single D/A channel\textsuperscript{5}. ‘Calibration’ for the diodes is done using single tones, as a noise signal introduces complications that affect relevant values and complicates the process. The only delay accounted for when creating the diode reference spectrum is that associated with the time between the start of the acquisition window and the actual time waveform extracted, $\phi_{\text{win}}$ (similar to that described above for the SFOAE).

In the description that follows, we specifically focus upon $2f_1 - f_2$ for clarity. We attempt to be as general as possible though and make it apparent where the equations will vary for different distortions.

When measuring the actual DPOAEs, the complex spectra are saved directly to file with no computation performed on the phase (save for the correction associated with the delay $\phi_{\text{win}}$). That is all done post-processing in Matlab. Using Fig. D-1 as a guide, the DPOAE phase for a given spectrum (with uniquely defined primary frequencies) is

$$\phi_{dp}^C = \phi_{dp}^B - \phi_{dp}^i = 2\phi_1^B - \phi_2^B - \phi_{dp}^i + \phi_x = (2\phi_1^A - \phi_2^A) - (2\phi_1^o - \phi_2^o) - \phi_{dp}^i + \phi_x \quad \text{(D.4)}$$

where $\phi_{dp}^C$ is the actual phase value read in from the stored file. The quantity $\phi_x$ is the value we ultimately wish to determine. The subscript always indicates the frequency of interest (except for the case of $x$ for the sake of clarity). Superscript $o$ label delays associated with the DAC pathway and $i$ for the ADC path. Thus then,

$$\phi^o = \phi^B - \phi^A \quad \text{(D.5)}$$

and

$$\phi^i = \phi^C - \phi^B \quad \text{(D.6)}$$

Basically, in order to get $\phi_x$ from $\phi_{dp}^C$, we just need to subtract all the other phase components present in Eqn.D.4. As described above, we can simply make $\phi_1^A = \phi_2^A = 0$, which will be done from this point on. However, one complicating

\textsuperscript{5}This avoids the necessity of creating a voltage adder if both D/A channels are used, each with a single sinusoid.
factor stems from the fact that when the probe is coupled to the ear, the delays $\phi_f$ and $\phi_e$ are not pure time delays, but include phase offsets due to the acoustic coupling stemming from the airspace between the probe and eardrum.

Out of necessity to be more specific, we introduce two further superscripts to distinguish between diode-produced and ear-produced distortions. The first, $e$, indicates purely electrical I/O delays. The second, $ea$, includes both electro and acoustical effects. Using this scheme, the phase shift due to the acoustic component when coupled to the ear is given by

$$\phi^{ac} = \phi^{eaC} - \phi^{eC}$$  \hfill (D.7)

The inclusion of the superscript $C$ indicates the phase as is measured post-digitization of the microphone signal (we will maintain this notation below, hopefully for the sake of clarity).

From the corresponding diode distortion spectrum (which has identical primary frequencies as that of the specific DPOAE spectrum being examined), the diode DP reference phase ($\phi_f^{\text{diode}} \equiv \phi_f^{eC}$) is extracted. Re-expressing relative to Eqn.D.4), we have

$$\phi^{eC} = -(2\phi_1^{eo} - \phi_2^{eo} + \phi_{dp}^{ei}) + \phi_x^{\text{diode}}$$  \hfill (D.8)

Presumably, $\phi_x^{\text{diode}} = 0$ as the diode generates the distortion instantaneously. Similarly, for the DPOAE measured in the ear canal, we have

$$\phi^{eaC} = -(2\phi_1^{eao} - \phi_2^{eao} + \phi_{dp}^{eai}) + \phi_x^{eaC}$$  \hfill (D.9)

Note that when the system is in loop-back mode, $\phi_f^{sy} = \phi_f^{eo} + \phi_f^{ei}$. Also, our target value now has a superscript, $\phi_x^{eaC}$.

We can rewrite the acoustic delay as

$$\phi^a = \phi^{eo} + \phi^{ai}$$  \hfill (D.10)

---

6We combine multiple superscripts below. Each component still retains its particular definition.
where
\[ \phi^{ao} = \phi^{eao} - \phi^{eo} \] (D.11)

and
\[ \phi^{ai} = \phi^{eai} - \phi^{ei} = 0 \] (D.12)

The last term of Eqn.D.10 is zero because we assume that the microphone response is instantaneous, thus there is no acoustic delay associated with that path. Thus, we have
\[ \phi^{ao} = \phi^{a} - \phi^{ai} \approx \phi^{a} \] (D.13)

Now comparing the DP data (ear versus diode),
\[ \phi^{eaC} - \phi^{eC} = -[2 \phi^{ao} - \phi^{ao} + \phi^{ai}] + [\phi^{eaC}_x - \phi^{eC}_x] \] (D.14)

which we can solve for \( \phi^{eaC}_x \). Keep in mind that we presume that \( \phi^{eC}_x = \phi^{ai}_x = 0 \).

Finally, we arrive at the expression for the corrected DPOAE phase term:
\[
\phi^{eaC}_x = \phi^{eaC}_x - \phi^{eC}_x - 2(\phi^{eaC}_1 - \phi^{eC}_1) + 2(\phi^{eaC}_2 - \phi^{eC}_2) \] (D.15)

D.2.4 Summary

Summarizing everything up to this point, it is important to account for time delays associated with one's measurement system when quantifying eOAE phase values. When determining the DPOAE phase, we have used a diode reference phase to effectively account for the unknown relative difference between various stages of system delay. Thus, the phase associated with the actual DPOAE generation mechanism can be revealed even when the relative I/O delays are unknown. As a reminder, the example explicitly described above was specific to \(2f_1 - f_2\). It is trivial then to adapt this method to any particular distortion product, whether it is harmonic or an inter-modulation. For example, \(3f_2 - 2f_1\) would be corrected as
\[
\phi^{eaC}_x = \phi^{eaC}_x - \phi^{eC}_x - 3(\phi^{eaC}_2 - \phi^{eC}_2) - 2(\phi^{eaC}_1 - \phi^{eC}_1) \] (D.16)
In the case of SFOAEs, it is even simpler and only the total system delay needs to be accounted for.
Bibliography


223
emission generation mechanisms.” Association for Research in Otolaryngology, Mid-Winter-Meeting Denver, CO), Abstract #811.


[57] Johansson J. EMC of Telecommunication Lines, Master’s Thesis, Lulea University of Technology (http://www.sm.luth.se/ urban/master/)


229


231


Financial Support

The primary source of support for the work done in this thesis comes from the Harvard-MIT Health Sciences and Technology Speech and Hearing Biosciences and Technology Program training grant (NIH T32 DC00038). Further support came from NIH grants T3200038, R01 DC0023821 (DMF) and R01 DC03687 (CAS).