Delay Estimation of DPOAEs: Is the f_1 or f_2 Sweep More Accurate?

CAPSTONE PROJECT

Presented in Partial Fulfillment of the Requirements for

the Degree Doctor of Audiology in the

Graduate School of The Ohio State University

By

Claudia A. Dome, BA

* * * * *

The Ohio State University

2007

Capstone Committee:

Stephanie Davidson, Adviser

Wayne M. King

Larry L. Feth

Peter F. Craigmile

Approved by

Adviser

Department of Speech and Hearing Sciences

ABSTRACT

ABSTRACT

Distortion product otoacoustic emission (DPOAE) delay estimation has important implications for models of emission generation and cochlear mechanics. However, the DPOAE signal environment makes the accurate estimation of DPOAE onset a difficult task. To date, two Fourier domain methods have been used exclusively in f_1 and f_2 sweep methods for DPOAE delay estimation. However, these sweep methods suffer from serious limitations. Additionlly, the two different sweep methods result in large differences in the estimated DPOAE delay, which are incompatible with the physical reality. In this study, a novel technique is employed for DPOAE latency estimation based on zero-phase filtering and penalized contrast changepoint estimation. This method has the advantage that it can be applied to a DPOAE time series evoked with a single (f_1, f_2) primary tone pair. The results are compared in the same human subjects against the traditional Fourier methods. The changepoint method provides support for the notion that the difference between the f_1 and f_2 sweeps is artifactual and that the f_2 sweep times may be closer to the true delay.

TABLE OF CONTENTS

Page

Abstract ii

Abstract iii

List of Tables vi

List of Figures vii

Chapters:

1. Introduction 1

    1.1 OAE Sources 2
        1.1.1 The DPOAE Two-source Model 5
    1.2 DPOAE Delay Estimation 7

2. DPOAE Delay Estimation-Mathematical and Algorithmic Foundations 10

    2.1 The Hilbert Space $L^2(\mathbb{R})$ 10
    2.2 Estimating $-\psi'(\omega)$ in DPOAE Data–Phase Gradients 16
        2.2.1 f_1 Sweep Method 17
        2.2.2 f_2 Sweep Method 20
    2.3 Model relating the f_1 sweep and f_2 sweep phase gradient estimates . . . 21
    2.4 Impulse Response Method 27
        2.4.1 The Impulse Response in LTI Systems 27
    2.5 Changepoint Detection 30
        2.5.1 Stationarity 31
        2.5.2 Changepoint Detection 33

3. Methods and Results 36

3.1 Methods 36
3.1.1 Human Subjects 36
3.1.2 DPOAE Acquisition 36
3.1.3 Data Analysis-Phase-gradient and Impulse Response Methods . . 39
3.1.4 Zero-phase Filtering with Changepoint Estimation 41
3.2 Results 43
3.2.1 Impulse Response Method Results 43
3.2.2 Changepoint Estimation Results 45

4. Discussion 49

References 52

LIST OF TABLES

Table **Page**

3.1 CDP ranges elicited by the f_1 and f_2 sweep paradigms for 2 and 3 kHz. By construction, the ranges overlap as much as possible while satisfying the necessary constraints. 37

LIST OF FIGURES

Figure **Page**

2.1 Top: Raw phases for subjects M1 and F1 as a function of CDP frequency for the 3 kHz f_1 sweep used in this study. Bottom: Unwrapped phases for the plots in the top panels. The delays are estimate as the slope of the least squares linear fit. 18

2.2 Top: f_1 sweep (3 kHz) phase gradient data for subjects M1 and F1. This data was shown in Figure 2.1 Bottom: The impulse response functions (voltage squared) for the 3 kHz f_1 sweep data. The impulse response functions peak at 3 msec, which is in agreement with the least squares fit to the phase gradients. 29

2.3 A DPOAE time series with f_1 and f_2 removed by phase rotation and time averaging. Note that the mean value of the data and its variance (the variability) do not change appreciably if small segments of the data are considered one at a time. 32

2.4 A DPOAE time series with f_1 and f_2 removed by phase rotation and time averaging. Note that near 30 msec and again near 180 msec, the process variance has changed. 33

2.5 A DPOAE time series with changepoints noted at 37.4 and 180.5 msec. . . 35

3.1 Top: f_1 phases recorded in the ear canal of one subject as a function of acquisition number. The initial phase is -0.90 radians. Note that the phase increments by $\pi/4$ radians. Bottom: f_2 phases recorded in the ear canal of one subject as a function of acquisition number. The initial phase is -0.94 radians. Note that the phase increments by $\pi/2$ radians. 39

3.2 Top left: Least squares linear fit to 3 kHz f_2 sweep data. Note the final unwrapped phase measurement is an outlier. Top right: Robust linear fit using the bisquare weighting function. This penalizes the large residual in the final phase measurement and thereby excludes it from the linear fit. Bottom: The impulse response function estimate (magnitude squared). The delay from the impulse response is in much better agreement with the robust fit. . 40

3.3 Impulse response delays vs phase gradient method delays for all subjects all sweeps. The correlation is quite strong except for a few outliers, which are errors in the phase gradient method. 41

3.4 Boxplots of f_1 sweep data by frequency (impulse response method). There is a significant decrease in delay as a function of frequency. 43

3.5 Top: f_1 sweep times as a function of frequency. Bottom: f_2 sweep times as a function of frequency. 44

3.6 Scatterplot of f_2 vs f_1 sweep times. The solid line represents the proposed linear relationship between the f_1 and f_2 sweep times (2 and 3 kHz) with the maximum slope, while the dashed line indicates the minimum slope. . . 45

3.7 Top left: averaged DPOAE time series. Top right: The same time series shown in the top right panel zoomed in on the 1st 50 msec. Bottom left: The filtered time series. Bottom right: The filtered series zoomed in on the 1st 50 msec with the changepoint indicated. 46

3.8 f_1 sweep delays for the changepoint method as a function of frequency. . . 47

3.9 f_1 and f_2 sweep delays for the changepoint method (2 and 3 kHz). 48

CHAPTER 1

INTRODUCTION

Otoacoustic emissions (OAEs) are retrograde pressure waves in the ear canal originating from the cochlea and generated by a physiologically vulnerable mechanism (Kemp, 1978). The most common taxonomy for evoked OAEs is based on the nature of the stimulus. Transient evoked otoacoustic emissions (TEOAEs) are elicited by clicks (80-100 μsec in duration) and tone pips (2-4 msec). Stimulus frequency emissions (SFOAEs) are evoked by a continuous tonal stimulus at a single angular frequency.

The third type of evoked OAE and the focus of this project are distortion product otoacoustic emissions (DPOAEs). DPOAEs are evoked by a superposition of sinusoidal stimuli. In the majority of experimental and clinical conditions, two tones are presented to the ear. If the two evoking tones are denoted as f_1 and f_2 (also known as the primaries) with $f_1 < f_2$, the DPOAEs will take the form $mf_1 \pm nf_2$ where $m, n \in \mathbb{Z}$. Since the exact form of the cochlear nonlinearity is unknown, it is not possible to predict which intermodulation distortion products are generated. Among those observed in humans with varying regularity are: $f_2 - f_1$, $2f_1 - f_2$, $3f_1 - 2f_2$, and $2f_2 - f_1$ (Moulin & Kemp, 1996). The so-called cubic distortion product (CDP) with frequency $2f_1 - f_2$ is by far the most robust and therefore most clinically significant.

Numerous studies have shown that the amplitude of the CDP is dependent on the overall and relative levels of the evoking stimuli as well as their frequency ratio (f_2/f_1) (Harris, Lonsbury-Martin, Stagner, Coats, & Martin, 1989; Gaskill & Brown, 1990; Kummer, Janssen, Hulin, & Arnold, 2000). If one allows L_1 to denote the sound pressure level (SPL) of f_1 and L_2 to denote the SPL of f_2, then for L_2 from 20 to 65 dB SPL, the empirically optimum L_1 is well approximated by the affine equation $L_1 = 0.4L_2 + 42$ (Kummer et al., 2000). For $L_2 = 55$ dB SPL, this implies that setting $L_1 = 64$ dB SPL will result in the maximum amplitude CDP. A large-scale study of normal and impaired listeners found that overall primary levels of $L_1 = 65$ and $L_2 = 55$ dB SPL produced the greatest separation in the CDP amplitude distributions of normal and impaired ears (Stover, Gorga, Neely, & Montoya, 1996). Reducing the overall primary tone levels resulted in a large number of normal ears resembling impaired ears, while increasing the levels caused a large number of impaired ears to more closely resemble normal ears. Further, it has been demonstrated in human ears that varying f_2/f_1 over a range of $[1.1, 1.3]$ will result in an approximately quadratic power function with the maximum amplitude CDP occurring near 1.22 (Harris et al., 1989; Gaskill & Brown, 1990; Fahey, Stagner, & Martin, 2006). Based on these findings primary tone levels of 65-55 and a primary tone ratio of 1.22 have become the clinical defaults for DPOAE testing.

1.1 OAE Sources

There is broad consensus that normally-functioning outer hair cells (OHCs) are a necessary condition for the generation of OAEs, while the mechanisms by which the emissions are produced and escape the cochlea remain poorly understood (Brown, McDowell, & Forge, 1989; Jaramillo, Markin, & Hudspeth, 1993; Camalet, Duke, Julicher, & Prost,

2000; Shera, 2004). Perhaps surprisingly, there is no consensus on how outer hair cells actually feed energy into the traveling wave to produce the emissions. The electromotility of the OHC lateral cell membrane has been demonstrated convincingly, but it is not clear that this electromotility is sufficient to produce the mechanical energy responsible for OAEs. In animals with no OHC motility, OAEs are most likely produced by nonlinearities in the OHC stereociliary bundle movements (Crawford & Fettiplace, 1985). OHC stereocilia have demonstrated a threshold displacement in the depolarizing direction. When this threshold is reached, an active process within the bundle produces a force, which causes nonlinear behavior in the displacement. There is increasing evidence that this process also contributes to mammalian OAEs (Oghalai, 2004; Ruggero, 2004).

In order for OAEs to be effective noninvasive probes of cochlear function, it is critical that their generation mechanisms be fully understood. Unfortunately, there are still considerable challenges in describing the biophysics of the cochlea as a prerequisite for understanding emissions. Models of wave motion are often highly idealized proceeding from a reduced two-chamber cochlear model with the scala vestibuli separated from the scala tympani by the basilar membrane. Clearly, this model ignores both Reissner's membrane and the organ of Corti. While the thickness and homogeneity of Reissner's membrane arguably warrant its exclusion in a simplified cochlear model, there is a complex mechanical interaction of elements in the organ of Corti, which cannot be ignored without compromising the ultimate utility of the model. However, incorporating the organ of Corti into exisiting cochlear models is a formidibable task.

Ignoring the cellular biomechanical detail, the extant theory of evoked OAEs describes two general generator mechanisms. One mechanism is the nonlinear distortion generated

when two or more traveling waves interact. This is referred to as the "wave-fixed" or "distortion" type emission (Knight & Kemp, 2001; Shera, 2004). The wave-fixed mechanism describes OAE generation as being a property of the traveling waves, specifically the maximal traveling wave interaction of f_1 and f_2 near the f_2 place. The potentially confusing part of that statement is the reference to the f_2 place. The f_2 place is important only in that it provides a spatial description of where the wave interaction takes place. The wave-fixed mechanism is an inherent part of the traveling wave interaction independent of the position along the basilar membrane where it occurs. The 2nd type of emission is thought to arise from impedance perturbations along the cochlea that produce scattering in response to the traveling wave. This is referred to as "place-fixed" or "reflection" type emissions (Zweig & Shera, 1995; Knight & Kemp, 2001; Shera, 2004). It is not unlike the scattering that occurs when a pulse traveling down a rope encounters structural inhomogeneities, e.g. changes in density. A major line of research in evoked OAEs is concerned with determining whether the stimulus-based taxonomy reflects differences in how OAEs are generated and whether these emissions break down along the lines of distortion or reflection types (Shera, 2004). Currently there is no consensus on this point. Some groups have argued that SFOAEs and TEOAEs have a common generation mechanism (reflection type), while DPOAEs are in part distinct from the other two (distortion and reflection type) (Shera & Guinan, 1999; Shera, 2004). Others have proposed that all emissions contain a mixture of emission types (sources) with the dominant contributions depending on changes in the parameters used to evoke them (Yates & Whitnell, 1999; Knight & Kemp, 2001; Goodman, Withnell, & Shera, 2003). In TEOAEs for example, the spectrum of the short duration stimulus has significant energy spread over a wide frequency range. An argument can be made that the traveling

waves generated by the stimulus participate in wave-fixed type interactions over a spatially-distributed region of the cochlea generating a number of distortion product emissions with considerable spectral extent (Yates & Whitnell, 1999).

For DPOAEs the dominant view until the early 1990s was that nonlinear distortion (wave-fixed) near the higher frequency primary was the sole source (Kemp, 1978; Kemp & Brown, 1983; Allen & Neely, 1992). The view that all distortion products are generated near the f_2 place makes specific predictions concerning DPOAE phase. The scaling symmetry of the cochlea predicts that the phase of the emission will be approximately independent of frequency for a fixed f_2/f_1 if the source is located near the peak of the traveling wave (Zweig & Shera, 1995; Shera & Guinan, 1999). This is due to the fact that the number of wavelengths a traveling wave completes before reaching its characteristic place is approximately frequency independent. Support for this scaling symmetry in the basilar membrane mechanics of healthy animals dates from the early 1970s (Rhode, 1971). Accordingly, the emission phase will be nearly identical regardless of whether it was generated with a relatively low or high frequency stimulus. For DPOAEs evoked with a fixed f_2/f_1 of ≈ 1.22 this is a good approximation. Studies have shown that fixing the ratio of the primary tones and sweeping them over a nearly three octave range results in less than a π radian shift in the phase of the emission as measured via the argument of its discrete Fourier transform (DFT) coefficient (Shera & Guinan, 1999; Knight & Kemp, 2001).

1.1.1 The DPOAE Two-source Model

While DPOAE phase behavior is nearly constant for a fixed primary tone ratio of 1.22, the same cannot be said of DPOAE magnitude when measured over a fine frequency scale, e.g. $1/32$ of an octave (He & Schmiedt, 1993). In these experiments DPOAE power

is shown to fluctuate in a quasi-periodic manner as a function of frequency with relative minima and maxima over ranges of 10-20 dB SPL. This phenomenon is commonly referred to as *DPOAE fine structure*. The intrigue of DPOAE fine structure was deepened by the discovery that these power fluctuations can be virtually eliminated by the introduction of a low-level suppressor tone near the $2f_1 - f_2$ frequency. This motivated the hypothesis that there is a 2nd DPOAE source at the characteristic place of the distortion product, which for $2f_1 - f_2$ is apical to both primary tones (Heitmann, Waldman, Schnitzler, Plinkert, & Zenner, 1998).

In the so-called *two-source model* the nonlinear interaction of f_1 and f_2 near the f_2 place generates both basally and apically propagating waves. The apically moving wave proceeds to its characteristic place where it generates a SFOAE, which then reflects basally (Talmadge, Long, Tubis, & Dhar, 1999). The CDP recorded in the ear canal is thought to be a mixture of one distortion product and one reflection emission. Phase differences in the two sources cause frequency-dependent interference patterns in the ear canal, which produce the fine structure.

As soon as the two-source model was proposed, reports emerged that questioned whether the source profile of DPOAEs was even more complicated. For example, experiments in rodents have provided support for the contribution of cochlear loci basal to the f_2 characteristic place in the generation of DPOAEs (Martin, Stagner, Jassir, Telischi, & Lonsbury-Martin, 1999). Further, there is some evidence that only the so-called *lower sideband* distortion products with frequencies lower than f_1 can be meaningful discussed in terms of wave-fixed and place-fixed contributions (Prijs, Schneider, & Schoonhoven, 2000). In any event, the ability to "unmix" the CDP has obvious theoretical and clinical implications. The direct clinical implication stems from the fact that the current hypothesis test for the

presence or absence of the CDP is used to infer cochlear status at the f_2 place (Mills, 1997; Kalluri & Shera, 2001). It may be that the conditions under which DPOAEs are acquired clinically ($f_2/f_1 \approx 1.22$, 65-55 dB SPL primaries) make the traveling wave interaction mechanism (wave-fixed) dominant (Knight & Kemp, 2000). If this is demonstrated to be the case, then the f_2 place inference is supported and there may be little clinical need for added complexity such as the introduction of a suppressor tone. In terms of the theoretical implication, if DPOAEs are to expand their clinical utility beyond the current "normal-abnormal" hypothesis test, a detailed knowledge of their generation mechanisms is critical.

1.2 DPOAE Delay Estimation

Motivated largely by the question of source localization, there has been considerable interest over the past decade in the temporal behavior of the CDP (Kimberley, Brown, & Eggermont, 1993; Stover, Neely, & Gorga, 1996; Whitehead, Stagner, Martin, & Lonsbury-Martin, 1996; Tubis, Talmadge, & Tong, 2000). Typical clinical measurements employ a Fourier transform of the ear canal pressure recordings and then retain only the estimated squared magnitudes of the Fourier coefficients. Both of these operations obscure the ability to discern the temporal evolution of the CDP. The ability to examine the temporal behavior of the CDP has important implications for source localization questions since sources from disparate regions in the cochlea would be expected to have different latencies. Techniques that measure the temporal behavior of the CDP would allow us to visualize how the contributions of these putative sources varied as acquisition parameters and cochlear status differed.

To date, two Fourier-domain methods have been used to estimate the latency of DPOAEs, the *phase gradient* (Kimberley et al., 1993; Mahoney & Kemp, 1995; Prijs et al., 2000) and

the *IFFT* method (or impulse response method) (Stover, Neely, & Gorga, 1996; Konrad-Martin et al., 2002). The phase gradient method is an approximation to the group delay, while the IFFT method approximates an impulse response for the cochlear "filter". Both methods constitute clever attempts to estimate DPOAE delay. However, they still suffer from limitations that affect their interpretation as true physical delays of the DPOAE. These limitations are exposed most clearly in the large discrepancy between delay estimates obtained with different paradigmatic implementations of the phase gradient and impulse response methods (Mahoney & Kemp, 1995; Stover, Neely, & Gorga, 1996).

The problem of estimating the temporal behavior of the DPOAE is a problem of joint time-frequency analysis. Unfortunately, DPOAE time series pose significant problems for time-frequency analysis. The evoking tones are typically at least 50 dB above the CDP and are recorded simultaneously thereby obscuring any distortion product visualization in the time domain. The phase gradient and impulse response methods are born out of a recognition of these challenges but produce results which are not completely reconcilable with the goal of estimating the DPOAE delay. It is the purpose of this study to examine these problems from a novel perspective. First, we derive an expression for phase derivatives approximated by the phase gradient method. This analysis uses a wave-fixed assumption (f_2 place generation site) and invokes the scaling symmetry and exponential frequency map of the cochlea to explain the discrepancies seen in the currently utilized delay estimation protocols. Secondly, a paradigmatic modification of DPOAE data collection is used in conjunction with time domain filtering and spectral changepoint estimation (Lavielle, 2005) to obtain DPOAE time delays. The changepoint delay estimates are then compared phase-gradient and impulse-reponse delays in normal-hearing human subjects.

The plan of this report is as follows. In the next chapter, some preliminary background on relevant time-frequency concepts and random processes is presented. Additionally, a detailed explanation of the phase-gradient and impulse response methods and the associated f_1 and f_2 sweep paradigms is given. The group delay of the CDP under the wave-fixed hypothesis and standard cochlear assumptions is derived. In chapter 3, the data acquisition methods including the pulsed primary and primary phase rotation paradigm used in human DPOAE data collection are described. Subsequently, DPOAE results from human subjects are presented. The final chapter contains the discussion of the findings and prospectus for further work.

CHAPTER 2

DPOAE DELAY ESTIMATION-MATHEMATICAL AND ALGORITHMIC FOUNDATIONS

The description of a signal jointly in time and frequency, or time and scale is a problem encountered in a wide variety of disciplines. Changing frequencies is a ubiquitous part of our sensory experience. Both our auditory and visual systems are tuned to notice change and often ignore steady state conditions. This property of the nervous system has an obvious evolutionary basis. For prey and predator alike, it is the detection of change which is critical to survival. Yet in spite of the basic role that changing frequencies play in our auditory and visual experience, time-frequency notions such as *instantaneous frequency* and *group delay* are not easy to define mathematically and often even more difficult to estimate in practice (Boashash, 1992).

2.1 The Hilbert Space $L^2(\mathbb{R})$

The most useful space for the analysis of signals is the space of square-integrable (in the Lebesque sense) complex-valued functions defined on the real numbers, $L^2(\mathbb{R})$. In the discussions to follow a bar over a function or scalar denotes the complex conjugate. In physical terms, $L^2(\mathbb{R})$ is the space of signals with finite energy. For a function, $g(t)$, to be

an element of the space we require

$$\int_{-\infty}^{\infty} |g(t)|^2 dt < \infty$$

In other words if we take the absolute value squared of the function and "sum" those values over the extended real line, a real number must be obtained. There are innumerable physical situations in which the absolute square of a signal is a key consideration making this signal space a natural one to consider. In electromagnetics, the kinetic energy imparted to a charge placed in a field is proportional to the square of the field. In sound, the energy density is proportional to the pressure squared. In circuits, the power is proportional to the voltage squared.

Let us define the real-valued scalar corresponding to the above integral as the squared *norm* of $g(t)$, which we denote as $||g(t)||^2$. Clearly when the norm ($||g(t)||$) is finite, $g(t)$ can be normalized as

$$\frac{g(t)}{||g(t)||}$$

Note in this case

$$\begin{aligned} \int_{-\infty}^{\infty} |\frac{g(t)}{||g(t)||}|^2 dt & \\ &= \frac{1}{||g(t)||^2} \int_{-\infty}^{\infty} |g(t)|^2 dt \\ &= 1 \end{aligned}$$

There are instances mathematically when it is convenient to have a unit norm. Additionally, it permits us to draw an analogy between the energy density of a signal and a probability density function, which we shall do for the calculation of moments.

It is often advantageous to represent elements of our space of signals as a superposition of basic building blocks. There are an infinite number of choices of these building

blocks. An important consideration is whether a particular representation of the signal will yield any useful insight. Certainly, one of the most useful representations for elements of $L^2(\mathbb{R})$ is the Fourier transform, which uses complex exponentials, $e^{i\omega t}$, as the "atomic" elements. It can be shown that an arbitrary element, $g(t)$, of $L^2(\mathbb{R})$ admits the following representation where equality is in the mean-square sense.

$$g(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(t') e^{-i\omega t'} dt') e^{i\omega t} d\omega$$

if we define

$$\hat{g}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(t) e^{-i\omega t} dt$$

we have

$$g(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{g}(\omega) e^{i\omega t} d\omega$$

and therefore $\hat{g}(\omega)$ gives us the contribution to $g(t)$ from the angular frequency ω. We emphasize that, in general, $\hat{g}(\omega)$ is complex-valued and therefore contains both amplitude and phase information since it can be written in the polar form $|\hat{g}(\omega)|e^{i\psi(\omega)}$ where $\psi(\omega)$ denotes the argument. Note that if $g(t)$ is a real-valued signal then

$$\begin{aligned}
\overline{\hat{g}(\omega)} &= \overline{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(t) e^{-i\omega t} dt} \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(t) e^{i\omega t} dt \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(t) e^{-i(-\omega)t} dt \\
&= \hat{g}(-\omega)
\end{aligned}$$

Therefore for real-valued $g(t)$ the Fourier transform exhibits conjugate symmetry, i.e. the Fourier transform at negative angular frequencies is the complex conjugate of the Fourier

transform at the corresponding positive angular frequencies. It also follows from the above that $|\hat{g}(\omega)|^2$ is an even function of ω.

The energy contained in a signal should not depend on the way we choose to represent it. Representations that preserve the norm of a signal are therefore quite useful. The Fourier transform has this property, namely that $||g(t)||^2 = ||\hat{g}(\omega)||^2$. In integral form this means that

$$\int_{-\infty}^{\infty} |g(t)|^2 dt = \int_{-\infty}^{\infty} |\hat{g}(\omega)|^2 d\omega$$

If we view the nonegative real-valued $|\hat{g}(\omega)|^2$ as a density, then $|\hat{g}(\omega)|^2 d\omega$ represents the fractional energy of $g(t)$ contained in a small frequency interval just as $|g(t)|^2 dt$ represents the fractional energy contained in a small time interval. It stands to reason that we can then define the average frequency in the usual way that averages are defined, i.e.

$$\mu_\omega = \int_{-\infty}^{\infty} \omega |\hat{g}(\omega)|^2 d\omega$$

Similarly, the average duration of a signal can be defined as

$$\mu_t = \int_{-\infty}^{\infty} t |gt(t)|^2 dt$$

Additionally, just as variance is defined as the 2nd central moment (about μ), the bandwidth of a signal can be defined as

$$\sigma_\omega^2 = \int_{-\infty}^{\infty} (\omega - \mu_\omega)^2 |\hat{g}(\omega)|^2 d\omega$$

and the variance of a signal in time can be similarly defined as

$$\sigma_t^2 = \int_{-\infty}^{\infty} (t - \mu_t)^2 |g(t)|^2 dt$$

From the above it seems that we need to know the Fourier transform of a function in order to compute the average frequency and bandwidth. Likewise, it seems intuitive that we cannot

compute the first moment of a signal in time solely by knowledge of its Fourier transform. Suprisingly and conveniently, this is not the case. There is a way to calculate the average frequency simply from knowledge of the time signal and its derivative. This is done via the use of a frequency *operator* (Cohen, 1995). If we define the frequency operator as

$$\mathcal{W}g(t) = \frac{1}{i}\frac{d}{dt}g(t)$$

then the average frequency can be obtained as

$$\mu_\omega = \int_{-\infty}^{\infty} \overline{g(t)}\mathcal{W}g(t)dt$$

The proof is as follows

$$\begin{aligned}
\mu_\omega &= \int_{-\infty}^{\infty} \omega|\hat{g}(\omega)|^2 d\omega \\
&= \int_\omega \omega(\frac{1}{\sqrt{2\pi}}\int_t g(t)e^{-i\omega t}dt)(\frac{1}{\sqrt{2\pi}}\int_{t'}\overline{g(t')}e^{i\omega t'}dt')d\omega \\
&= \frac{1}{2\pi}\int_\omega\int_t\int_{t'} \omega\overline{g(t')}g(t)e^{i\omega(t'-t)}dt'dtd\omega
\end{aligned}$$

note $\quad \dfrac{\partial}{\partial t'}e^{i\omega(t'-t)} = i\omega e^{i\omega(t'-t)}$

$$\begin{aligned}
&= \frac{1}{i2\pi}\int_\omega\int_t\int_{t'} \overline{g(t')}g(t)\frac{\partial}{\partial t'}e^{i\omega(t'-t)}dt'dtd\omega \\
&= \frac{1}{i}\int_t\int_{t'} \overline{g(t')}g(t)\frac{\partial}{\partial t'}(\frac{1}{\sqrt{2\pi}}\int_\omega \frac{1}{\sqrt{2\pi}}e^{-i\omega t}e^{i\omega t'}d\omega)dt'dt \\
&= \frac{1}{i}\int_{t'}\int_t \overline{g(t')}g(t)\frac{\partial}{\partial t'}\delta(t'-t)dtdt' \\
&= \frac{1}{i}\int_{t'}\overline{g(t')}\frac{\partial}{\partial t'}\int_t g(t)\delta(t'-t)dtdt' \\
&= \frac{1}{i}\int_{t'}\overline{g(t')}\frac{d}{dt'}g(t')dt'
\end{aligned}$$

To illustrate the utility of this result consider the following amplitude and frequency modulated signal

$$g(t) = (\frac{\alpha}{\pi})^{1/4} e^{-\alpha t^2/2 + i(\omega_0 t + \beta t^2/2)}$$

with $\alpha > 0$. The amplitude is modulated in an exponentially decaying manner by $e^{-\alpha t^2/2}$ while the phase is quadratically modulated. The average frequency is seen to be ω_0. However, to prove this result by using the Fourier transform is not a trivial task. Fortunately, the frequency operator provides a much easier way to find μ_ω. To show that we can compute this without knowing the Fourier transform recognize that if we set $\alpha = \frac{1}{2\sigma^2}$ then

$$\sqrt{\frac{\alpha}{\pi}} e^{-\alpha t^2}$$

is the form of a Gaussian probability density function with mean zero. Next we note that

$$\mathcal{W} g(t) = \frac{1}{i}\frac{d}{dt} g(t) = \sqrt{\frac{\alpha}{\pi}} (i\alpha t + \beta t + \omega_0) e^{-\alpha t^2/2 + i(\omega_0 t + \beta t^2/2)}$$

Therefore

$$\begin{aligned} &\int_{-\infty}^{\infty} \overline{g(t)} \frac{1}{i}\frac{d}{dt} g(t) \\ &= \int_{-\infty}^{\infty} \sqrt{\frac{\alpha}{\pi}} (i\alpha t + \beta t + \omega_0) e^{-\alpha t^2} dt \\ &= \int_{-\infty}^{\infty} \omega_0 \sqrt{\frac{\alpha}{\pi}} e^{-\alpha t^2} dt = \omega_0 \end{aligned}$$

since the the integral does not depend on ω_0, it may be pulled outside the integral sign. We then recognize the integral as a Gaussian probability density function which integrates to one.

The first moment of the signal in time can be computed solely by knowledge of its Fourier transform in an analogous manner through the use of the time operator

$$\mathcal{T} \hat{g}(\omega) = -\frac{1}{i}\frac{d}{d\omega} \hat{g}(\omega)$$

with

$$\mu_t = \int_{-\infty}^{\infty} \overline{\hat{g}(\omega)} \mathcal{T} \hat{g}(\omega) d\omega$$

The proof is similar to the one for the frequency operator. Note that if we write the Fourier transform of a signal in polar form as

$$\hat{g}(\omega) = |\hat{g}(\omega)| e^{i\psi(\omega)}$$

where $\psi(\omega)$ is the spectral phase then

$$\begin{aligned}
\mu_t &= \int_{\omega} \overline{\hat{g}(\omega)} - \frac{1}{i} \frac{d}{d\omega} \hat{g}(\omega) d\omega \\
&= \int_{\omega} |\hat{g}(\omega)| e^{i\psi(\omega)} - \frac{1}{i} \{ |\hat{g}(\omega)|' e^{i\psi(\omega)} + i\psi'(\omega) |\hat{g}(\omega)| e^{i\psi(\omega)} \} d\omega \\
&= \int_{\omega} -\psi'(\omega) |\hat{g}(\omega)|^2 d\omega
\end{aligned}$$

From the above we see that the first moment of a signal in time can be obtained by integrating the negative of the first derivative of the spectral phase against the energy density in frequency. In this way $-\psi'(\omega)$ can be interpreted as the instant of appearance of a specific frequency ω. This is termed the *group delay*. If we take ω to be the angular frequency of a DPOAE then $-\psi'(\omega)$ is the instant that the DPOAE appears in the time signal. The key question then becomes how do we estimate $-\psi'(\omega)$. In DPOAE data this is done by the use of phase gradients.

2.2 Estimating $-\psi'(\omega)$ in DPOAE Data–Phase Gradients

A number of investigators have attempted to use the phase (argument of the Fourier coefficient) as an estimate of DPOAE delay (Kimberley et al., 1993; Whitehead et al., 1996; Bowman, Eggermont, Brown, & Kimberley, 1998). One of the most common ways

to estimate the delay (latency) of DPOAEs is a paradigm in which one of the primary frequencies is fixed and the other is swept in frequency (Kimberley et al., 1993; Mahoney & Kemp, 1995; Prijs et al., 2000). These are referred to as *phase gradient* methods and have been widely used both to estimate cochlear traveling wave delay (Kimberley et al., 1993; Mahoney & Kemp, 1995; Prijs et al., 2000) and to provide theoretical models for DPOAE generation (Tubis, Talmadge, & Tong, 2000; Tubis, Talmadge, Tong, & Dhar, 2000). In an f_1 sweep paradigm, f_2 is fixed while f_1 is swept in frequency. As in all swept frequency paradigms, the (f_1, f_2) frequency pairs are constrained to vary over a range of ratios from $\approx [1.1, 1.3]$. Outside of this range, it is difficult to obtain distortion products even in normal-hearing listeners with large CDP amplitudes under typical testing conditions. The f_1 sweep method is implemented as follows.

2.2.1 f_1 Sweep Method

Fix f_2 at an arbitrary frequency (e.g. 2 kHz) while f_1 is swept in some small frequency step (e.g. 25 Hz) from ≈ 1540 to ≈ 1820. This would result in 12 (f_1, f_2) pairs of the form:

$$(1540, 2000) \quad (1565, 2000) \quad (1590, 2000) \ldots (1665, 2000) \ldots (1815, 2000)$$

Obviously with each of these pairs a different CDP is elicited. Just as the pairs are given in order of increasing f_1 frequency, the CDP will increase in frequency from 1080 to 1630. Note that with a fixed f_2 if f_1 is moved in 25-Hz increments, the CDP will move in $2f_1$ or 50 Hz increments. The algorithm is implemented as follows.

1. Data is acquired at each (f_1, f_2) pair. This will require a requisite amount of time domain averaging in order to elicit a detectable CDP if at all.

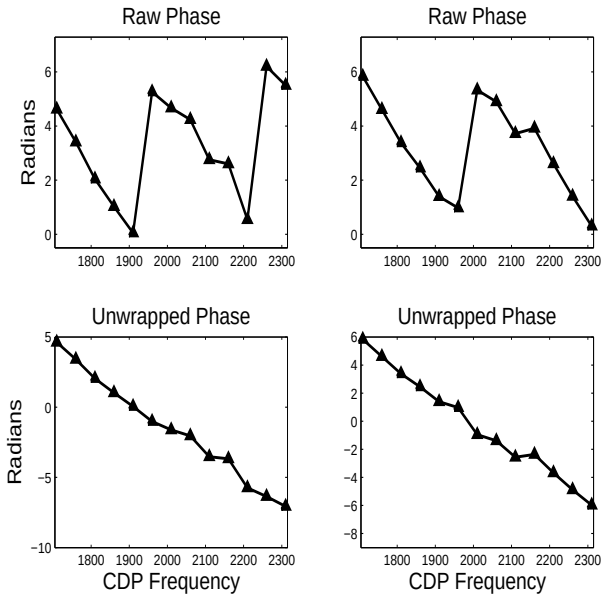

Figure 2.1: Top: Raw phases for subjects M1 and F1 as a function of CDP frequency for the 3 kHz f_1 sweep used in this study. Bottom: Unwrapped phases for the plots in the top panels. The delays are estimate as the slope of the least squares linear fit.

2. The data is Fourier transformed and the Fourier coefficient (complex-valued) for each significant CDP is retained in a vector.

3. The phase is unwrapped, i.e. phase changes exceeding π radians are replaced by their 2π complement and a simple first order linear regression model of the form

$$\phi_{CDP} = \beta_0 + \beta_1 f_{CDP} + \epsilon_{CDP}$$

is fit where ϕ_{CDP} represents the unwrapped DP phase, f_{CDP} the CDP frequency (in Hz) and ϵ_{CDP} represents the error term. The negative of the slope of the least squares fit (scaled by $1/2\pi$ if Hz are used) is used as the estimate of the group delay. In Figure 2.1, f_1 sweep data is shown for a fixed f_2 of 3 kHz for two subjects.

Recall that using the time operator, it was shown that the negative of the derivative of the spectral phase has the interpretation as the instant when a frequency appears.

There are three chief problems with the phase-gradient method. First, the sweep paradigm acquires a number of phase measurements at different frequencies, all the (f_1, f_2) pairs, and estimates the derivative as the best linear fit to the unwrapped phases. As a result, data acquisition must always involve a number of (f_1, f_2) pairs. This requires both excess time and renders the overall delay estimate sensitive to the behavior of just one or two (f_1, f_2) pairs. Next, the derivative is assumed to be constant requiring the group delay to be the same at each of the CDP frequencies. Note that there is nothing in the definition of group delay that requires the derivative to be constant. In principle, this is a severe constraint. Phase gradient methods are only able to arrive one delay estimate. In other words, if the two-source model described in the introduction is accurate, the phase gradient method would have no way to unmix the sources. Both wave-fixed and place-fixed emissions would be measured as a single phase contribution at the CDP frequencies via the argument of the Fourier coefficient. In fact, the phase gradient method is implicitly based on the distortion model of DPOAEs. For example, in an odd order nonlinearity, a superposition of two tones,(f_1, f_2), results in a cubic distortion product (among others) with a phase that depends on the evoking tone phase via the relationship $2\theta_1 - \theta_2$ with θ_1 denoting the phase of f_1 and θ_2 denoting the phase of f_2. Under the distortion model of OAE emission, the CDP phase in a fixed f_2-swept f_1 paradigm reflects the phase of f_1 at the f_2 place. Recall that the scaling symmetry of the cochlea results in a constant phase for a fixed f_2/f_1. However, the f_1 sweep paradigm changes the relationship between the primary tone traveling wave patterns by systematically moving f_1 and thereby changing the phase of the resulting CDP. Finally, the linear regression model assumes equal variance of the error terms, ϵ_{CDP}. Given

our knowledge of human DPOAEs evoked with different primary tone ratios, this is not a valid assumption.

2.2.2 f_2 Sweep Method

In the f_2 sweep method, the f_1 primary is fixed and f_2 is swept in a small frequency step. Again, the (f_1, f_2) ratios are constrained to fall in the $[1.1, 1.3]$ range in order to elicit CDPs with a reasonable probability in normal-hearing persons. In the case of the f_2 sweep, the CDP frequency will move the same increment as the f_2 step because of its $2f_1 - f_2$ dependence. The measurement paradigm proceeds exactly as described for the f_1 sweep method with the reversed roles of the primaries.

In may seem incompatible with the f_2 place hypothesis to attempt to measure DPOAE delay with a f_2 sweep paradigm. The key is to understand that in a f_2 sweep paradigm, the higher frequency primary is only being moved a few hundred Hz. Even if it could be recorded with perfect accuracy, the difference in delay for a CDP emerging from say the 4300 Hz place on the cochlea as opposed to the 4000 Hz place would be infinitesimally small.

In spite of biologic arguments for why the f_1 sweep and f_2 sweep paradigms should exhibit negligible differences, striking contrasts are routinely observed between the two. DPOAE delays recorded with the f_2 sweep are approximately twice as long as those recorded with the f_1 sweep method (Kimberley et al., 1993; Mahoney & Kemp, 1995; Moulin & Kemp, 1996). It is important to consider how these differences arise. Based on purely physical considerations, they must be artifacts of the measurement process. This question has been addressed in detail in a previous paper and an approximate relationship between the f_1 and f_2 sweep methods derived (Tubis, Talmadge, Tong, & Dhar, 2000). In

the next section, a complementary analysis is presented that reaches substantively the same conclusions reached by the authors of the earlier paper.

2.3 Model relating the f_1 sweep and f_2 sweep phase gradient estimates

In this model, the DPOAE spectral phase, $\psi(\omega_{dp})$, is expressed as the sum of two components, an anterograde component, $\overrightarrow{\phi}(x_g, \omega_{dp})$, which depends on the phases of f_1 and f_2 at the generator region and a retrograde component $\overleftarrow{\phi}(x_g, \omega_{dp})$ so that

$$\psi(\omega_{dp}) = \overrightarrow{\phi}(x_g, \omega_{dp}) + \overleftarrow{\phi}(x_g, \omega_{dp})$$

In order to derive an expression for the derivative of the spectral phase in terms of the anterograde and retrograde components, the total differentials are examined term by term. The total differential of the anterograde component, ($\overrightarrow{\psi}(\omega_{dp})$), is

$$d\overrightarrow{\psi}(x_g, \omega_{dp}) = \frac{\partial \overrightarrow{\phi}(x_g, \omega_{dp})}{\partial \omega_{dp}} d\omega_{dp} + \frac{\partial \overrightarrow{\phi}(x_g, \omega_{dp})}{\partial x_g} dx_g$$

Similarly, the total differential of the retrograde component, ($\overrightarrow{\psi}(\omega_{dp})$), is

$$d\overleftarrow{\psi}(x_g, \omega_{dp}) = \frac{\partial \overleftarrow{\phi}(x_g, \omega_{dp})}{\partial \omega_{dp}} d\omega_{dp} + \frac{\partial \overleftarrow{\phi}(x_g, \omega_{dp})}{\partial x_g} dx_g$$

Next the explicit dependence of ω_{dp} on ω_1 and ω_2 is worked into the model. Equating the differentials the following relations are obtained

$$d\omega_{dp} = 2d\omega_1$$

$$d\omega_{dp} = -d\omega_2$$

The total differential for the anterograde component is given by

$$2\frac{\partial \overrightarrow{\phi}(x_g, \omega_1)}{\partial \omega_1} d\omega_1 - \frac{\partial \overrightarrow{\phi}(x_g, \omega_2)}{\partial \omega_2} d\omega_2 + \frac{\partial \overrightarrow{\phi}(x_g, \omega_{dp})}{\partial x_g} dx_g$$

where the dependence of the phase-place derivative on ω_{dp} has been momentarily maintained on purpose. Note that ω_{dp} in the retrograde component is the emission phase and therefore cannot be explained via a straightforward dependence on the primary tone phases. The DPOAE group delay becomes

$$-\frac{d\psi(\omega_{dp})}{d\omega_{dp}} = -2\frac{\partial\overrightarrow{\phi}(x_g,\omega_1)}{\partial\omega_1}\frac{d\omega_1}{d\omega_{dp}} + \frac{\partial\overrightarrow{\phi}(x_g,\omega_2)}{\partial\omega_2}\frac{d\omega_2}{d\omega_{dp}} - \frac{\partial\overrightarrow{\phi}(x_g,\omega_{dp})}{\partial x_g}\frac{dx_g}{d\omega_{dp}}\cdots$$
$$-(\frac{\partial\overleftarrow{\phi}(x_g,\omega_{dp})}{\partial\omega_{dp}} + \frac{\partial\overleftarrow{\phi}(x_g,\omega_{dp})}{\partial x_g}\frac{dx_g}{d\omega_{dp}})$$

Now we return to the anterograde phase-place derivative given by

$$\frac{\partial\overrightarrow{\phi}(x_g,\omega_{dp})}{\partial x_g}\frac{dx_g}{d\omega_{dp}}$$

again utilizing the relationship $\omega_{dp} = 2\omega_1 - \omega_2$ and noting that $d\omega_{dp}/d\omega_1 = 2$ and $d\omega_{dp}/d\omega_2 = -1$, the above may be rewritten as

$$\frac{\partial\overrightarrow{\phi}(x_g,\omega_{dp})}{\partial x_g}\frac{dx_g}{d\omega_{dp}} = 2\frac{\partial\overrightarrow{\phi}(x_g,\omega_1)}{\partial x_g}\frac{dx_g}{d\omega_{dp}} - \frac{\partial\overrightarrow{\phi}(x_g,\omega_2)}{\partial x_g}\frac{dx_g}{d\omega_{dp}}$$

Using the above, the group delay may be rewritten as

$$-\frac{d\psi(\omega_{dp})}{d\omega_{dp}} = -2(\frac{\partial\overrightarrow{\phi}(x_g,\omega_1)}{\partial\omega_1}\frac{d\omega_1}{d\omega_{dp}} + \frac{\partial\overrightarrow{\phi}(x_g,\omega_1)}{\partial x_g}\frac{dx_g}{d\omega_{dp}}) + (\frac{\partial\overrightarrow{\phi}(x_g,\omega_2)}{\partial\omega_2}\frac{d\omega_2}{d\omega_{dp}} + \frac{\partial\overrightarrow{\phi}(x_g,\omega_2)}{\partial x_g}\frac{dx_g}{d\omega_{dp}})$$
$$-(\frac{\partial\overleftarrow{\phi}(x_g,\omega_{dp})}{\partial\omega_{dp}} + \frac{\partial\overleftarrow{\phi}(x_g,\omega_{dp})}{\partial x_g}\frac{dx_g}{d\omega_{dp}})$$

Note that for an f_1 sweep $d\omega_2 = 0$, while in an f_2 sweep $d\omega_1 = 0$. Further, the generator region is assumed to be fixed at the f_2 place and therefore can be written solely as a function of f_2, i.e. $x_g(\omega_2)$. Accordingly, in an f_1 sweep paradigm

$$\frac{dx_g}{d\omega_{dp}} = 0$$

while in an f_2 sweep paradigm this derivative is nonzero.

In simplifying the expression for group delay, we consider f_1 sweep first. In order to subsequently distinguish the group delays obtained with different sweep paradigms, let D_1 denote $-d\psi/d\omega_{dp}$ obtained with the f_1 sweep and D_2 denote the group delay observed with the f_2 sweep. It follows that D_1 can be written as

$$D_1 = -\frac{\partial \overrightarrow{\phi}(x_g, \omega_1)}{\partial \omega_1} - \frac{\partial \overleftarrow{\phi}(x_g, \omega_{dp})}{\partial \omega_{dp}}$$

where we have used

$$2\frac{d\omega_1}{d\omega_{dp}} = 1$$

In order to analyze D_2, the group delay for the f_2 sweep method, it is instructive to invoke the scale-invariance of the cochlea, which results in a constant anterograde phase for $\overrightarrow{\phi}(x_i, \omega_i)$ where x_i is the characteristic place for the frequency ω_i. Mathematically, this is expressed as

$$\overrightarrow{\phi}(x_i, \omega_i) = K$$

where K denotes a constant and therefore

$$\frac{d\overrightarrow{\phi}}{d\omega_i} = \frac{\partial \phi(x_i, \omega_i)}{\partial \omega_i} + \frac{\partial \phi(x_i, \omega_i)}{\partial x_i}\frac{dx_i}{d\omega_i} = 0$$

It is well known that the rate of change of frequency with respect to place in the cochlea is proportional to frequency with the characteristic frequencies monotonically decreasing from the base of the cochlea (near the oval window) to the apex. This relationship can be written as a differential equation where the independent variable is x, the spatial position along the cochlea, and the dependent variable is $\omega(x)$.

$$\frac{d\omega}{dx} = -\lambda\omega \qquad \lambda > 0$$

The above is an ordinary first order differential equation, which can be solved as follows

$$\frac{d\omega}{dx} = -\lambda\omega$$
$$\frac{d\omega}{dx} + \lambda\omega = 0$$
$$\text{set} \quad \mu(x) = e^{\int \lambda dx} = e^{\lambda x}$$
$$\frac{d\omega}{dx}e^{\lambda x} + \lambda e^{\lambda x}\omega = 0$$
$$\frac{d}{dx}(\omega(x)e^{\lambda x}) = 0$$
$$\int \frac{d}{dx}(\omega(x)e^{\lambda x})dx = C$$
$$\omega(x) = Ce^{-\lambda x}$$

where C is a constant of integration. Returning to the expression for the derivative of the anterograde phase of an arbitrary frequency at its characteristic place, we obtain

$$\begin{aligned}\frac{\partial \overrightarrow{\phi}(x_i, \omega_i)}{\partial \omega_i} &= -\frac{\partial \overrightarrow{\phi}(x_i, \omega_i)}{\partial x_i}\frac{dx_i}{d\omega_i} \\ &= \frac{\partial \overrightarrow{\phi}(x_i, \omega_i)}{\partial x_i}\frac{1}{\lambda\omega_i}\end{aligned}$$

and therefore

$$\frac{\partial \overrightarrow{\phi}(x_i, \omega_i)}{\partial \omega_i}\lambda\omega_i = \frac{\partial \overrightarrow{\phi}(x_i, \omega_i)}{\partial x_i}$$

In order to derive an expression for D_2, recall the expression for group delay derived earlier

$$-\frac{d\psi(\omega_{dp})}{d\omega_{dp}} = -2(\frac{\partial \overrightarrow{\phi}(x_g, \omega_1)}{\partial \omega_1}\frac{d\omega_1}{d\omega_{dp}} + \frac{\partial \overrightarrow{\phi}(x_g, \omega_1)}{\partial x_g}\frac{dx_g}{d\omega_{dp}}) + (\frac{\partial \overrightarrow{\phi}(x_g, \omega_2)}{\partial \omega_2}\frac{d\omega_2}{d\omega_{dp}} + \frac{\partial \overrightarrow{\phi}(x_g, \omega_2)}{\partial x_g}\frac{dx_g}{d\omega_{dp}})$$
$$-(\frac{\partial \overleftarrow{\phi}(x_g, \omega_{dp})}{\partial \omega_{dp}} + \frac{\partial \overleftarrow{\phi}(x_g, \omega_{dp})}{\partial x_g}\frac{dx_g}{d\omega_{dp}})$$

In the wave-fixed model, the generation region of the CDP is held to vary with ω_2. Accordingly, we have

$$\frac{dx_g}{d\omega_{dp}} = \frac{dx_g}{d\omega_2}\frac{d\omega_2}{d\omega_{dp}}$$

substituting this into the expression for group delay, we obtain

$$-\frac{d\psi(\omega_{dp})}{d\omega_{dp}} = -2(\frac{\partial\overrightarrow{\phi}(x_g,\omega_1)}{\partial\omega_1}\frac{d\omega_1}{d\omega_{dp}} + \frac{\partial\overrightarrow{\phi}(x_g,\omega_1)}{\partial x_g}\frac{dx_g}{d\omega_2}\frac{d\omega_2}{d\omega_{dp}}) + \dots$$
$$\frac{d\omega_2}{d\omega_{dp}}(\frac{\partial\overrightarrow{\phi}(x_g,\omega_2)}{\partial\omega_2} + \frac{\partial\overrightarrow{\phi}(x_g,\omega_2)}{\partial x_g}\frac{dx_g}{d\omega_2}) + -(\frac{\partial\overleftarrow{\phi}(x_g,\omega_{dp})}{\partial\omega_{dp}} + \dots$$
$$+\frac{\partial\overleftarrow{\phi}(x_g,\omega_{dp})}{\partial x_g}\frac{dx_g}{d\omega_2}\frac{d\omega_2}{d\omega_{dp}})$$

In order to simplfy the expression for D_2, the anterograde and retrograde components are examined separately. First, the anterograde component, $\overrightarrow{D}_2$, can be written as

$$\overrightarrow{D}_2 = -2(\frac{\partial\overrightarrow{\phi}(x_g,\omega_1)}{\partial\omega_1}\frac{d\omega_1}{d\omega_{dp}} + \frac{\partial\overrightarrow{\phi}(x_g,\omega_1)}{\partial x_g}\frac{dx_g}{d\omega_2}\frac{d\omega_2}{d\omega_{dp}}) + \dots$$
$$\frac{d\omega_2}{d\omega_{dp}}(\frac{\partial\overrightarrow{\phi}(x_g,\omega_2)}{\partial\omega_2} + \frac{\partial\overrightarrow{\phi}(x_g,\omega_2)}{\partial x_g}\frac{dx_g}{d\omega_2})$$

The first term is seen to be zero because $d\omega_1 = 0$ in an f_2 sweep. Further, due to the scale invariance of the cochlea, the term

$$\frac{\partial\overrightarrow{\phi}(x_g,\omega_2)}{\partial\omega_2} + \frac{\partial\overrightarrow{\phi}(x_g,\omega_2)}{\partial x_g}\frac{dx_g}{d\omega_2}$$

is recognized as the perfect differential of $\overrightarrow{\phi}(x_i,\omega_i) = K$ and therefore is equal to zero. Therefore, the anterograde component of D_2 simplifies to

$$\overrightarrow{D}_2 = -2\frac{\partial\overrightarrow{\phi}(x_g,\omega_1)}{\partial x_g}\frac{dx_g}{d\omega_2}\frac{d\omega_2}{d\omega_{dp}}$$

Converting the phase-place derivative above into a phase-frequency derivative, we obtain

$$\overrightarrow{D}_2 = -2\frac{\partial\overrightarrow{\phi}(x_g,\omega_1)}{\partial\omega_1}\lambda\omega_1\frac{dx_g}{d\omega_2}\frac{d\omega_2}{d\omega_{dp}}$$

Given the differential equation describing the cochlear place-frequency map and the relation $\omega_2 = 2\omega_1 - \omega_{dp}$, the above can be rewritten as

$$\overrightarrow{D}_2 = -2\frac{\partial \overrightarrow{\phi}(x_g, \omega_1)}{\partial \omega_1}\lambda\omega_1 \frac{dx_g}{d\omega_2}\frac{d\omega_2}{d\omega_{dp}}$$
$$\overrightarrow{D}_2 = -2\frac{\partial \overrightarrow{\phi}(x_g, \omega_1)}{\partial \omega_1}\lambda\omega_1 - \frac{1}{\lambda\omega_2}(-1)$$
$$\overrightarrow{D}_2 = 2\frac{\omega_1}{\omega_2}(-\frac{\partial \overrightarrow{\phi}(x_g, \omega_1)}{\partial \omega_1})$$

Next we write the retrograde component of D_2 ($\overleftarrow{D}_2$) as

$$\overleftarrow{D}_2 = -\frac{\partial \overleftarrow{\phi}(x_g, \omega_{dp})}{\partial \omega_{dp}} - \frac{\partial \overleftarrow{\phi}(x_g, \omega_{dp})}{\partial x_g}\frac{dx_g}{d\omega_2}\frac{d\omega_2}{d\omega_{dp}}$$
$$= -\frac{\partial \overleftarrow{\phi}(x_g, \omega_{dp})}{\partial \omega_{dp}} - \frac{\partial \overleftarrow{\phi}(x_g, \omega_{dp})}{\partial \omega_{dp}}(\lambda\omega_{dp})(-\frac{1}{\lambda\omega_2})(-1)$$
$$= 2\frac{\omega_1}{\omega_2}(-\frac{\partial \overleftarrow{\phi}(x_g, \omega_{dp})}{\partial \omega_{dp}})$$

Combining the previous result for $\overrightarrow{D}_2$, we obtain

$$D_2 = 2\frac{\omega_1}{\omega_2}(-\frac{\partial \overrightarrow{\phi}(x_g, \omega_1)}{\partial \omega_1} - \frac{\partial \overleftarrow{\phi}(x_g, \omega_{dp})}{\partial \omega_{dp}})$$

Recalling that

$$D_1 = -\frac{\partial \overrightarrow{\phi}(x_g, \omega_1)}{\partial \omega_1} - \frac{\partial \overleftarrow{\phi}(x_g, \omega_{dp})}{\partial \omega_{dp}}$$

it follows that

$$\frac{D_2}{D_1} = 2\frac{\omega_1}{\omega_2}$$

In the above it was shown that under a wave-fixed hypothesis and by invoking the scale invariance of the cochlea, the f_2 sweep delay is expected to exceed the f_1 sweep delay by a factor of $2\omega_1/\omega_2$. Given ω_2/ω_1 ratios in the interval $[1.1, 1.3]$ it follows that the f_2 sweep

delay should be ≈ 1.54 to 1.82 times the f_1 sweep delay. While theoretically the D_2/D_1 ratio is a function of both ω_1 and ω_2, the phase gradient method produces only one group delay estimate. An important question that remains to be addressed for the phase gradient methods is how the reductionist approach necessitated by the phase gradient method affects the estimates of D_1 and D_2 and accordingly the proposed relationship

$$D_2 = 2\frac{\omega_1}{\omega_2}D_1$$

In the above, the group delay obtained with the f_2 sweep paradigm is predicted to be approximately two times the group delay obtained with the f_1 sweep. Note that for the range of primary frequency ratios used in the sweep methods, this factor will vary from 2(1/1.1) to 2(1/1.3). This prediction will subsequently be tested against data collected in human subjects.

2.4 Impulse Response Method

In addition to the phase-gradient methods, the impulse response method has been extensively used to estimate DPOAE delay (Stover, Neely, & Gorga, 1996; Konrad-Martin, Neely, Keefe, Dorn, & Gorga, 2001). In the previous section, it was noted that the phase gradient methods are only capable of producing one group delay estimate. This estimation method is incompatible with the theory that there are at least 2 sources of the ear-canal CDP. The impulse response method of DPOAE delay estimation was developed to overcome this limitation(Stover, Neely, & Gorga, 1996).

2.4.1 The Impulse Response in LTI Systems

The use of the impulse response is indispensible in the theory of linear translation-invariant systems. For an LTI system, the response to an arbitrary input is completely

characterized by the convolution of the input with the system's impulse response function. Equivalently, we may take the Fourier transform of the input and multiply it by the Fourier transform of the system's impulse response, known as the system *transfer function*. This follows from the well-known result that convolution in one domain is equivalent to multiplication in the canonically conjugate domain. In the conventional sense, filtering an input signal invovles convolving it with a given impulse response, $h(t)$, or equivalently multiplying the Fourier transform of the input by $\hat{h}(\omega)$.

Recall that in the f_1 and f_2 sweep paradigms, CDPs are elicited at a number of different frequencies. For each of these CDP frequencies, a complex number (the DFT coefficient) is obtained. If these DFT coefficients are placed in a vector with some appropriate spacing indicative of their frequency separation, this vector can be viewed as the transfer function of a cochlear *filter*. It then follows that the inverse DFT of the transfer function is the impulse response. The peak of the cochlear impulse response is taken as the CDP delay.

If the DFT coefficients are only placed at positive frequencies, the impulse response will be complex-valued. A real-valued signal must display conjugate symmetry in the Fourier domain. Previous work with the impulse response method has created real-valued impulse response functions by satisfying the conjugate symmetry property, i.e. by placing the complex conjugate of the DFT coefficients in the appropriate DFT bins. In this work, the analytic signal corresponding to the impulse response is preferred. The envelope of the analytic signal provides a better estimate of the peak of the impulse response and therefore the putative CDP delay. The following algorithm is used to produce the impulse response estimate.

1. The DFT of the averaged DPOAE time data is computed.

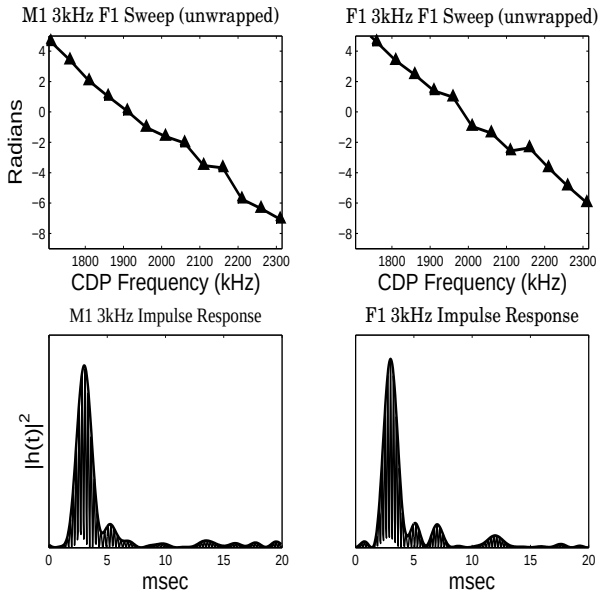

Figure 2.2: Top: f_1 sweep (3 kHz) phase gradient data for subjects M1 and F1. This data was shown in Figure 2.1 Bottom: The impulse response functions (voltage squared) for the 3 kHz f_1 sweep data. The impulse response functions peak at 3 msec, which is in agreement with the least squares fit to the phase gradients.

2. The DFT coefficient for the CDP is retained and placed in the appropriate place in a vector of zeros (positive frequencies only).

3. The DFT vector is multipled by a factor of 2 to create the Fourier transform of the analytic signal.

4. The inverse DFT is computed and the resulting vector is the analytic signal corresponding the impulse response. The modulus is used to estimate the CDP delay. In Figure 2.2, the estimated impulse response functions are shown for two subjects along with the corresponding phase gradient estimates.

2.5 Changepoint Detection

In this study, a novel application of changepoint detection for DPOAE time series is introduced. This section serves as a brief introduction to changepoint detection in general and to the specific technique for dependent observations used in this study.

In order to properly define the problem of changepoint estimation, it is necessary to define the notion of a random, or stochastic process. A real-valued stochastic, or random, process is a mapping from the Cartesian product of some index set and an outcome space into the real numbers. Recall that a random variable (RV) is a measurable function mapping some outcome space, Ω, into the real numbers such that the inverse image of the mapping is a contained within the σ-algebra of subsets of Ω. If we denote the real number assigned to an arbitrary $\omega \in \Omega$ as $X(\omega)$, then a real-valued random process assigns a real number to each pairing of elements from the index set and the outcome space. If we denote the index set as T, then a stochastic process may be denoted as $X(t, \omega)$ with $t \in T$. The index set may be countable or uncountable. The index set is often time, but need not be. The probability measure defined on the underlying outcome space, $T\mathrm{x}\Omega$, induces a probability measure on $X(t, \omega)$. Therefore, a random process can be simply thought of a sequence of random variables. Examples of such processes in audiology are evoked potential and OAE waveforms.

Note that there is nothing in the definition of a random process which requires that the probability measure is the same for each t. As an illustration, a random proccess can be defined on a countable index set, such that for even t, the random variables follow a $N(1, 1)$ probability law, while for odd t the random variables are exponentially distributed with parameter 1.

2.5.1 Stationarity

Since a random process is a sequence of random variables, it stands to reason that we will be interested in the relationships among the RVs that make it up. Specifically, we will be interested in the joint distribution of the constituent RVs, i.e. $P(X_1 \leq x_1 \cap X_2 \leq x_2 \cap X_3 \leq x_3 \ldots)$. Clearly, for all but trivial processes, the joint distribution will be too large to specify. This is especially true in practice when the process is being estimated from a finite number of observations. In many cases, it helps to make simplifying assumptions about the process. If these assumptions constitute a reasonable approximation to the process, the estimation problem may be tractable. One of the most common assumptions in the study of random processes is that of *stationarity*. The most common form of stationarity is actually *second order stationarity* A random process, X_t, is said to be second order stationary if

1. $\mathbb{E}(x_t) = \mu$ where μ is a constant independent of time.

2. $\mathbb{E}(|x_t|^2) < \infty$.

3. $\gamma(r, s) = \mathbb{E}\{(x_r - \mu_r)(x_s - \mu_s)\} = \mathbb{E}\{(x_{r+t} - \mu_r)(x_{s+t} - \mu_s)\} = \gamma(r + t, s + t)$.

where $\mathbb{E}$ denotes the expectation operator and $\gamma(r, s)$ is the autocovariance between points r and s. If the process does not satisfy the above it is referred to as *nonstationary*. In summary, the above requirements require that the process has a mean value and variance that do not depend on what segment of the process that we are looking at. In Figure 2.3, we show an example of a DPOAE time series where the contribution of the primary tones has been removed by a method to be described in a later section. This process constitutes one realization of what would be considered a stationary process.

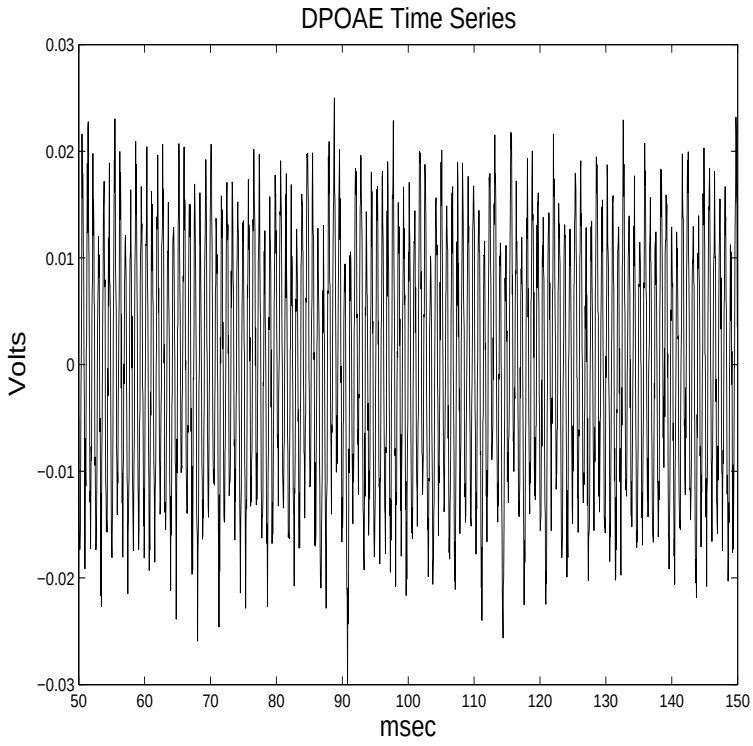

Figure 2.3: A DPOAE time series with f_1 and f_2 removed by phase rotation and time averaging. Note that the mean value of the data and its variance (the variability) do not change appreciably if small segments of the data are considered one at a time.

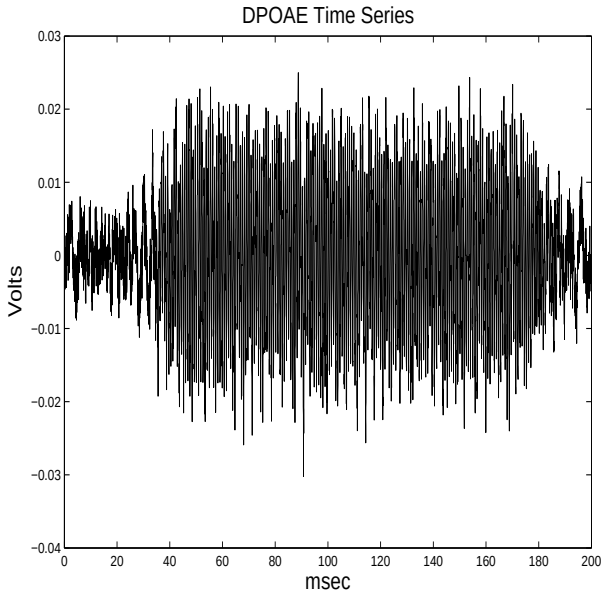

Figure 2.4: A DPOAE time series with f_1 and f_2 removed by phase rotation and time averaging. Note that near 30 msec and again near 180 msec, the process variance has changed.

In actuality, Figure 2.3 was produced by zooming in on 100 msec of the DPOAE time series. In Figure 2.4, the entire 200 msec of data is shown. It is clear while the mean of the data is independent of the index (time), the variance of the process is not. There is a marked change in the variance somewhere around 30 msec and again near 180 msec. The data may be regarded as piecewise stationary, but not globally stationary.

2.5.2 Changepoint Detection

In many applications, it is critical to determine when some aspect of a random process has changed (Basseville & Nikiforov, 1993). In general, we assume that the statistics of the process are determined by some parameter $\theta \in \Theta$ which remains constant between changepoints. As a simple gross characterization, the time series shown in Figure 2.4

contains two changepoints, one near 30 msec and another near 180 msec. The process is parameterized by some parameter, θ from 0 to 30 msec (and perhaps again from 180 to 200 msec), and by $\theta' \neq \theta$ from 30 to 180 msec.

Changepoint techniques may be used to detect changes in the mean, variance, empirical distribution function, and the empirical spectral distribution of a stationary process in a specified band of frequencies (Picard, 1985; Basseville & Nikiforov, 1993; Lavielle & Ludena, 2000; Lavielle, 2005). In this study we use a method, which simultaneously detects all the changepoints in the empirical spectral distribution function of a 2nd order stationary random process by minimizing a penalized contrast function (Lavielle & Ludena, 2000; Lavielle, 2005). If we denote the periodogram of the process x_t over the j-th segment as

$$I_j(\omega) = \frac{1}{2\pi n_k} | \sum_{t=t_{j-n_k+1}}^{t_j} x_t e^{-i\omega t} |^2$$

Then

$$F_j(\lambda) = \int_0^\lambda I_j(\omega) d\omega$$

denotes the energy in the band $[0, \lambda]$. Lavielle & Cardenas (2000) suggest the following contrast estimator

$$-\frac{n_k}{N} \sum_j F_j(\lambda)^2$$

This contrast estimator is then weighted by a term which penalizes over estimation of the number of changepoints. In Figure 2.5, the penalized contrast estimator is applied to the DPOAE time series shown in Figure 2.4. The estimator returns two changepoints, one at 37.4 msec and one at 180.5 msec.

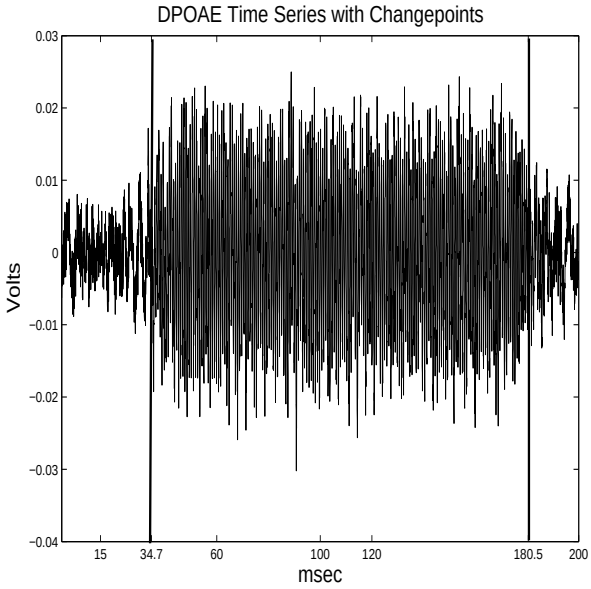

Figure 2.5: A DPOAE time series with changepoints noted at 37.4 and 180.5 msec.

CHAPTER 3

METHODS AND RESULTS

3.1 Methods

3.1.1 Human Subjects

Nine subjects, 3 male and 6 female, participated in the experiment. All subjects were paid for their participation. The subject selection criteria for this experiment were: 1.) normal hearing as defined by air conduction thresholds ≤ 15 dB HL at the standard audiometric frequencies and the absence of any significant air-bone gap, 2.) clear canals and normal tympanograms on the day of testing, and 3.) normal click-evoked TEOAEs on the day of testing.

3.1.2 DPOAE Acquisition

DPOAE data were acquired in f_1 and f_2 sweep paradigms "centered" on f_2 frequencies of 2, 3, and 4 kHz. In the f_1 sweep paradigm, the f_2 frequencies were fixed at 2, 3, or 4 kHz respectively. For each acquisition, f_1 was initially set at the nearest Fourier frequency resulting in an approximate f_2/f_1 ratio of 1.3. Subsequently, f_1 was increased in 25 Hz steps until an approximate f_2/f_1 ratio of 1.1 was reached. This resulted in 11 (f_1, f_2) pairs for the 2-kHz f_1 sweep and 13 (f_1, f_2) pairs for the 3 and 4-kHz f_1 sweeps.

Frequency	f_1 Sweep	f_2 Sweep
2 kHz	$[1080, 1580]$	$[1130, 1480]$
3 kHz	$[1710, 2310]$	$[1760, 2110]$

Table 3.1: CDP ranges elicited by the f_1 and f_2 sweep paradigms for 2 and 3 kHz. By construction, the ranges overlap as much as possible while satisfying the necessary constraints.

An error in the setup of the 4 kHz f_2 sweep made that data unsuitable for analysis. Subsequently, in discussing the f_2 sweeps, reference will only be made to 2 and 3 kHz. For the f_2 sweeps, the lower frequency primary (f_1) was fixed at a Fourier frequency yielding an approximate primary tone ratio of 1.22 when the higher frequency primary was 2 or 3 kHz. For example, in the 3-kHz f_2 sweep, the lower frequency primary was fixed at 2455 Hz yielding $3000/2455 = 1.22$. The f_1 frequency for the 2-kHz f_2 sweep was fixed at 1640 Hz. In each of the f_2 sweeps, the higher frequency primary was moved in 50-Hz steps from approximately 150-300 Hz above to 200-300 Hz below its "centerpoint" (2 or 3 kHz). The f_1 and f_2 sweeps were designed so that the essentially the same range of cubic distortion products were elicited in both paradigms subject to the constraint that the $1.1 \leq f_2/f_1 \leq 1.3$ ratios. The ranges of CDP frequencies elicited in 2 and 3 kHz paradigms are given in Table 3.1.

For each (f_1, f_2) frequency pair, eighty acquisitions of 200 msec in duration were obtained. Each of these acquisitions consisted of 4000 points sampled at 20 kHz. The resulting 4000x80 data matrix was stored for offline analysis. The lower frequency primary was presented at 65 dB SPL for the entire 200 msec. The higher frequency primary was pulsed on for 150 msec beginning at 25 msec and terminating at 175 msec. At the beginning and end of the pulsed primary, 2.5 msec cosine-squared windows were used to reduce spectral

spread. The higher frequency primary was presented at 55 dB SPL. The pulsed paradigm was utilized to provide a clear marker for the onset of the CDP. Evoking the distortion product requires stimulation of the cochlea by both primary tones. The delayed f_2 onset facilitates the detection of the CDP because each time series contains a number of sample points prior to the onset (in this case a minimum of 500).

In addition to the pulsed f_2 primary, the phases of both primaries were rotated in a systematic way designed to greatly reduce the energy contributions of primary tones in the data. The phase of the f_1 primary was rotated by $\pi/4$ radians in each acquisition, while the phase of the f_2 primary was rotated by $\pi/2$. Due to the phase rotation, averaging any 8 (or integer multiple of 8) consecutive acquisitions largely eliminated ear canal contributions from both f_1 and f_2 in the recording. Note that the anterograde component of the CDP, $\overrightarrow{\phi}(x_g, \omega_{dp})$, is not eliminated since the actual starting phases of f_1 and f_2 could not be controlled precisely enough in the ear canal. This is illustrated in Figure 3.1 on the next page. In Figure 3.1 the intial phases of f_1 and f_2 were -0.90 and -0.94 radians respectively, it is clear that the $2\theta_1 - \theta_2$ contribution of the primary phases is nonzero. However, the phase rotation of the primaries is maintained at integer multiples of $\pi/4$ for f_1 and $\pi/2$ for f_2, which maintains the phase cancellation.

In theory the contribution of f_1 and f_2 is completely eliminated with this technique. However, these recordings were made in the ear canals of human subjects and it is impossible to eliminate small movements, e.g. swallowing, head movement, etc., that introduce phase perturbations in the recording. As a result, in practice this phase cancellation of the primary tones is rarely perfect, but in compliant subjects it does result in substantial cancellation of f_1 and f_2. The DPOAE data acquisition paradigm is summarized as follows:

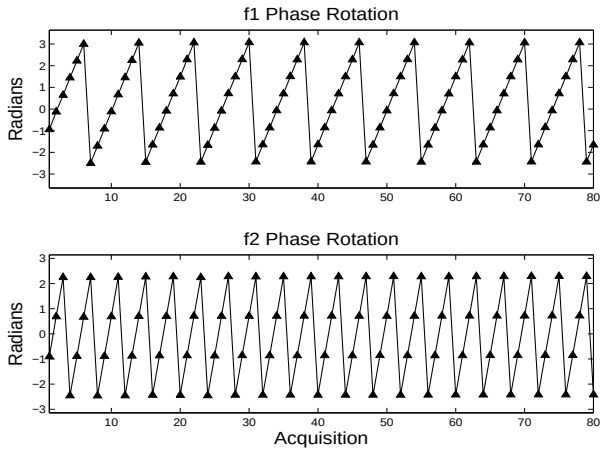

Figure 3.1: Top: f_1 phases recorded in the ear canal of one subject as a function of acquisition number. The initial phase is -0.90 radians. Note that the phase increments by $\pi/4$ radians. Bottom: f_2 phases recorded in the ear canal of one subject as a function of acquisition number. The initial phase is -0.94 radians. Note that the phase increments by $\pi/2$ radians.

- Each acquisition was 200 msec in length. The sampling rate was fixed at 20 kHz. For each acquisition 4000 data points were acquired. This resulted in a frequency resolution of 5 Hz. Eighty 200 msec acquisitions were obtained for each (f_1, f_2) pair.

- The phase of f_1 was rotated $\pi/4$ radians and the phase of f_2 was rotated $\pi/2$ radians for each acquisition.

- f_1 was presented for the entire 200 msec acquisition, while f_2 was on from 25-150 msec.

3.1.3 Data Analysis-Phase-gradient and Impulse Response Methods

The resulting f_1 and f_2 sweeps were subjected to analysis with the appropriate phase-gradient and impulse response methods. All the (f_1, f_2) pairs which elicited a significant

CDP were used in the analysis. In fitting a linear regression model to the unwrapped phases, a bisquare weighting function was used to avoid the sensitivity of the standard least squares technique to outliers with large residuals. The sensitivity of the standard least squares technique is illustrated in Figure 3.2 The estimated delays obtained with the phase gradient

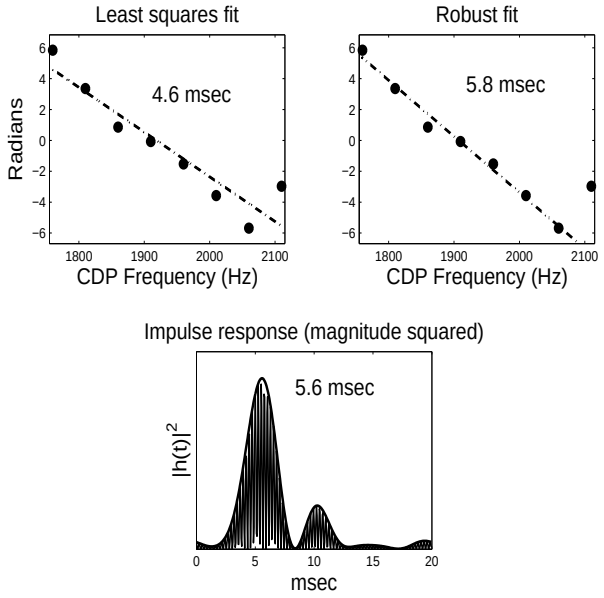

Figure 3.2: Top left: Least squares linear fit to 3 kHz f_2 sweep data. Note the final unwrapped phase measurement is an outlier. Top right: Robust linear fit using the bisquare weighting function. This penalizes the large residual in the final phase measurement and thereby excludes it from the linear fit. Bottom: The impulse response function estimate (magnitude squared). The delay from the impulse response is in much better agreement with the robust fit.

and impulse response methods were in good agreement. This is illustrated in Figure 3.3 The linear fit is seen to be quite strong with the exception of a few outliers. Upon inspection, it was clear that the outliers were the result of the phase gradient estimates. Even with robust regression methods, there were cases in which a regression model could not be fit

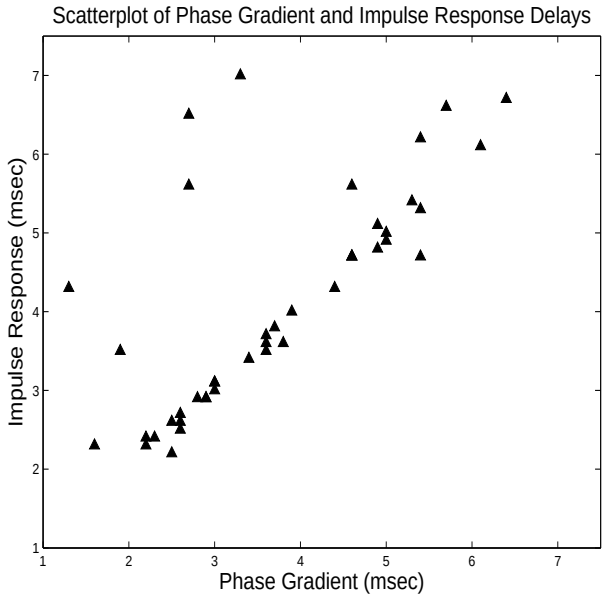

Figure 3.3: Impulse response delays vs phase gradient method delays for all subjects all sweeps. The correlation is quite strong except for a few outliers, which are errors in the phase gradient method.

to the data rendering the phase gradient method unusable. In subsequent data analysis, the impulse response method delays will be used exclusively.

3.1.4 Zero-phase Filtering with Changepoint Estimation

Prior to changepoint estimation all time series were convolved with an equiripple FIR filter to improve the SNR. The time series were convolved with the filter, the output reversed, and run through the filter again to remove phase effects. The filters were constructed so that the CDP frequency was centered in a 600-Hz passband. The stopbands were an additonal 400 Hz below and above the respective passband edges. The transfer function of the filter was purposely designed to be very conservative in order to avoid artifactual ringing that could be detected as a spurious changepoint in the spectral distribution. After filtering

each time series was checked for the presence of spectral artifacts at the CDP frequency by taking the DFT of the 1st 500 points of the series. Since f_2 was presented beginning at 25 msec, only artifact could account for the presence of the CDP in the 1st 500 points. In those cases where f_1 was not perfectly cancelled and was within 300 Hz of the CDP, an IIR notch filter was first used to remove f_1 to prevent artifactual results in the changepoint analysis. The removal of f_1 when necessary was also accomplished with zero-phase filtering.

Changepoint estimation was performed on empirical spectral distribution function as outlined in (Lavielle, 1999). A conservative α-level test of 10^{-5} was used. This is the level recommended in Lavielle(2005) because the method is trying to simultaneously estimate all the changepoints in the process. In order to obtain the relative delay of the CDP, an f_2 time series was obtained by averaging every 4-th acquisition of the same data matrix. This series was used to obtained an absolute f_2 delay (recall that f_2 was turned on at 25 msec and reached its maximum amplitude at 27.25 msec). No filtering was necessary in these time series due the magnitude of f_2. The absolute f_2 delay was then subtracted from the absolute CDP delay to obtain the relative CDP delay.

Unlike the phase-gradient and impulse response methods, the changepoint technique can be applied to every primary tone ratio yielding in principle N different delays where N is the number of (f_1, f_2) pairs. A larger SNR will yield a more accurate estimate of the changepoint. Accordingly, a weighted average of the changepoint delay was computed by multiplying each delay by the ratio of its local SNR to the total SNR and summing the result.

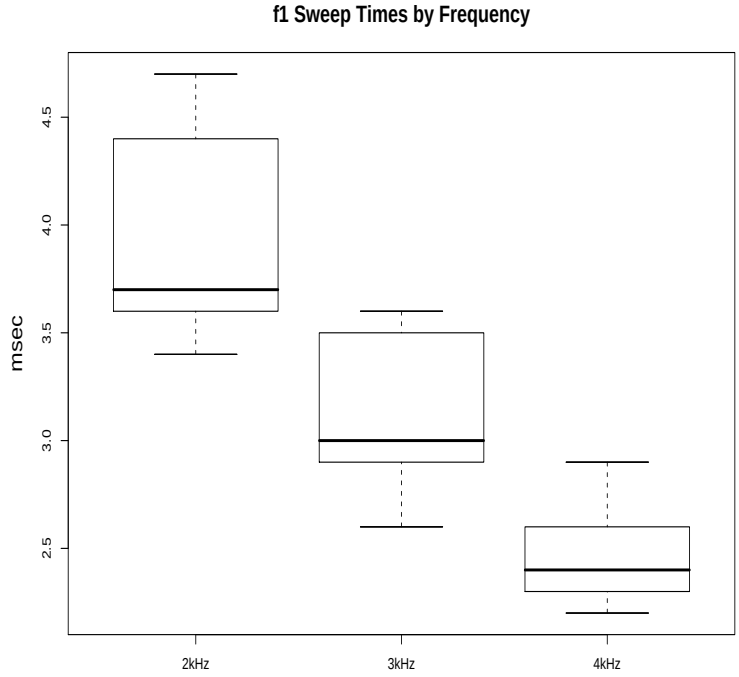

Figure 3.4: Boxplots of f_1 sweep data by frequency (impulse response method). There is a significant decrease in delay as a function of frequency.

3.2 Results

3.2.1 Impulse Response Method Results

The delay times obtained in the f_1 and f_2 sweep paradigms in this study were in good agreement with those reported previously (Kimberley et al., 1993; Mahoney & Kemp, 1995; Stover, Neely, & Gorga, 1996). In particular, a repeated measures ANOVA showed a significant decrease in both the f_1 and f_2 sweep times as a function of frequency. This is shown for the f_1 sweep data with 2,3, and 4 kHz in Figure 3.4.

Both f_2 sweeps resulted in significantly larger estimated group delays than their f_1 sweep counterparts. Figure 3.5 shows the delay times by frequency and sweep method.

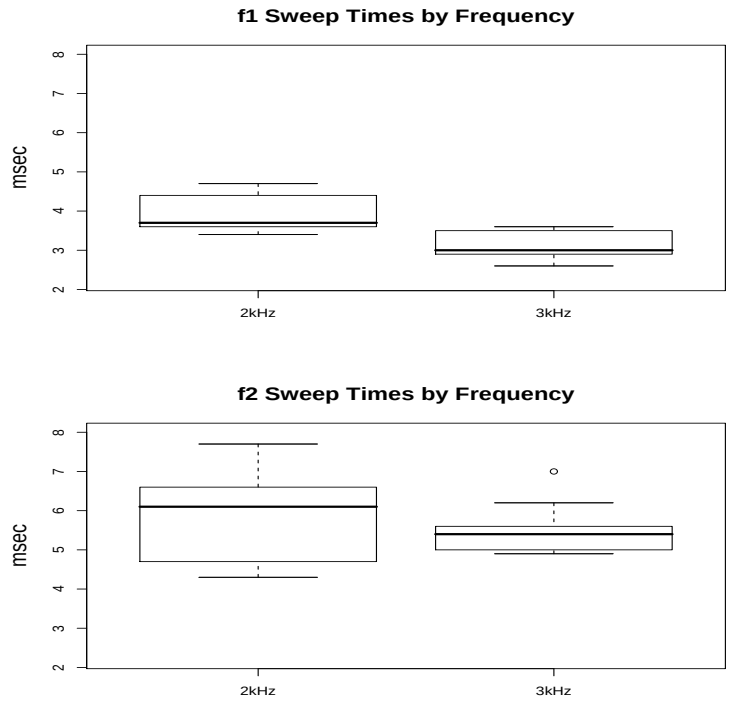

Figure 3.5: Top: f_1 sweep times as a function of frequency. Bottom: f_2 sweep times as a function of frequency.

Recall in chapter 2, a relationship was derived between the f_1 and f_2 sweep delays based on the scale invariance property of the cochlea and a wave-fixed assumption. The proposed relationship was given by

$$D_2 = 2\frac{\omega_1}{\omega_2}D_1$$

For the ratios of ω_1/ω_2 used in this study, the minimum value of this factor is 1/1.3 and the maximum is given by 1/1.1. In Figure 3.6, a scatterplot of f_1 vs. f_2 sweep times for the 2 and 3 kHz f_1 and f_2 sweeps is presented along with two lines indicating the minimum and maximum slopes of the proposed relationship given above.

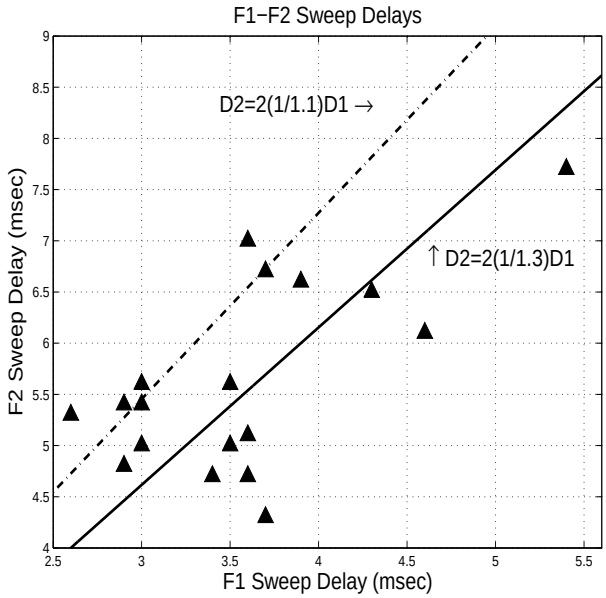

Figure 3.6: Scatterplot of f_2 vs f_1 sweep times. The solid line represents the proposed linear relationship between the f_1 and f_2 sweep times (2 and 3 kHz) with the maximum slope, while the dashed line indicates the minimum slope.

3.2.2 Changepoint Estimation Results

Empirically a local SNR of at least 28 dB was required for the detection of the CDP delay as a changepoint. This resulted in a failure to produce changepoint delay estimates in a few subjects, particularly in the f_2 sweeps (recall the conservative $\alpha = 10^{-5}$) in one or more of the sweep methods. For another subset of subjects, the local SNRs were high enough that significant changepoints were recorded at virtually every (f_1, f_2) pair in several of the sweeps. In Figure 3.7, we present an averaged DPOAE time series and its filtered counterpart with the significant changepoint (in absolute time) indicated

Similar to the impulse response (phase gradient) method, there was a significant decrease in delay as a function of frequency for the changepoint method. These results are

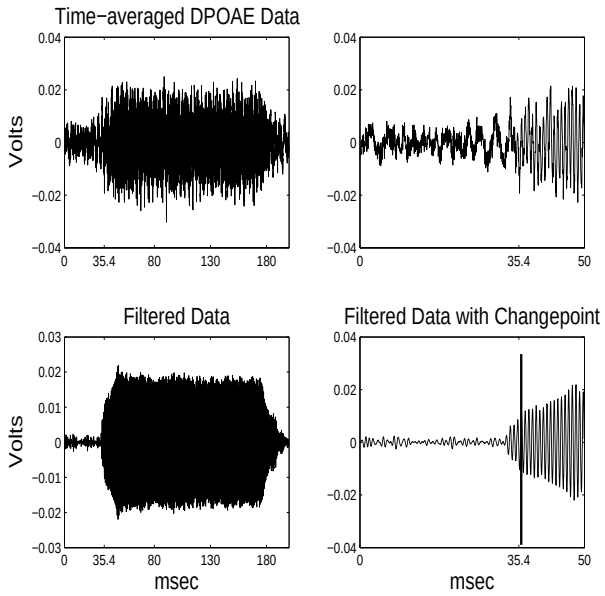

Figure 3.7: Top left: averaged DPOAE time series. Top right: The same time series shown in the top right panel zoomed in on the 1st 50 msec. Bottom left: The filtered time series. Bottom right: The filtered series zoomed in on the 1st 50 msec with the changepoint indicated.

presented in Figure 3.8. The reader is invited to compare these results with those presented in Figures 3.4 and 3.5.

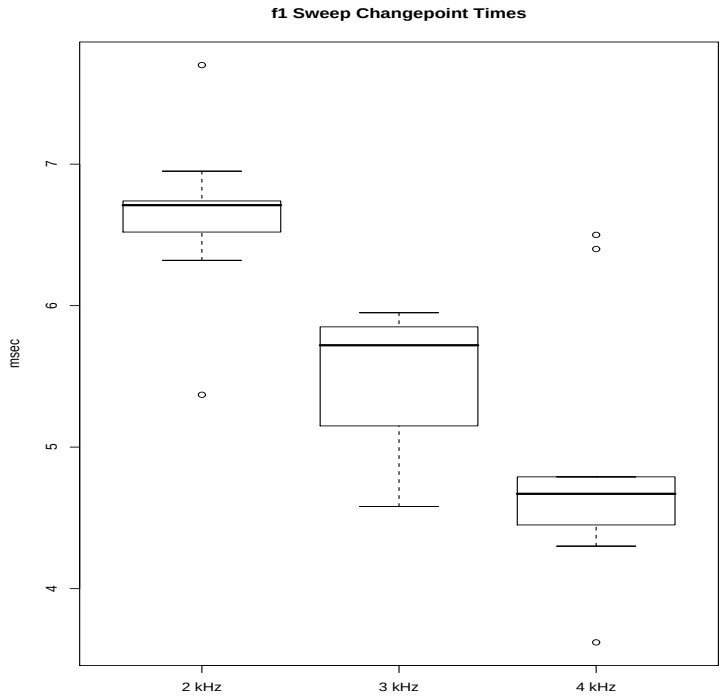

Figure 3.8: f_1 sweep delays for the changepoint method as a function of frequency.

Finally, the f_1 and f_2 sweeps were compared with the changepoint method. While there was a significant difference between the f_1 and f_2 sweep times for the impulse response method, the different sweep paradigms failed to show any significant difference with the changepoint method as illustrated in Figure 3.9

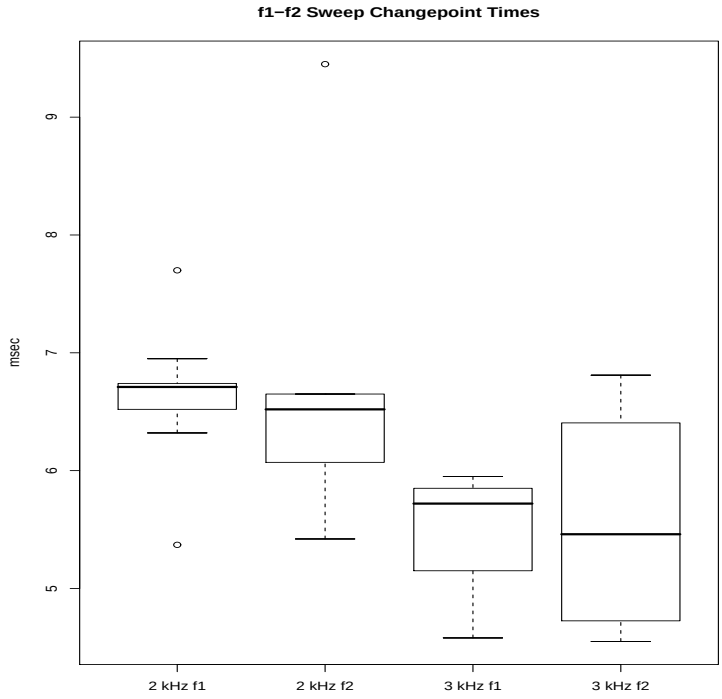

Figure 3.9: f_1 and f_2 sweep delays for the changepoint method (2 and 3 kHz).

In summary, both the impulse response (phase gradient) and changepoint methods show significant decreases in delay as a function of frequency. The f_2 sweep delays were significantly longer than the f_1 sweep delays with the impulse response method, while there was no significant difference between the sweep methods with the changepoint technique.

CHAPTER 4

DISCUSSION

In this study, a novel technique for the delay estimation of DPOAEs was presented. The method relied on a pulsed-primary paradigm with phase rotation to reduce the significant contribution of the nuisance frequencies, f_1 and f_2. A penalized contrast method was then used to estimate the onset of the DPOAE as a change in the spectral distribution in a local frequency neighborhood. Additionally, a complementary approach to Tubis et al.(2000) was presented which predicted under a wave-fixed hypothesis that the f_1 delay estimated with the phase gradient method should be approximately twice the f_2 sweep delay. This has been reported in previous studies and was shown to be in good agreement with the delays observed in these data. Further, significant decreases in estimated delay were found in the f_1 and f_2 sweeps as the frequency increased. This also is in good agreement with previous work. The average changepoint delays also decreased significantly as the frequency increased from 2 to 4 kHz for the f_1 sweep and 2 to 3 kHz for the f_2 sweep. Unlike with phase gradient and impulse response methods, there was no significant difference between the f_1 and f_2 sweep delays estimated with the changepoint method.

Purely based on physical reasoning, there is no way that the infintesimal movement in the generator region could account for a doubling in the delay. For this reason, others have cautioned against interpreting the phase gradient results (and hence impulse response

methods results as well) as true delays (Tubis, Talmadge, Tong, & Dhar, 2000). The question arises if any of the current methods yield results which are close to the true delay. In other words, should more faith be placed in the f_1 sweep or f_2 sweep delays. The delays measured with the changepoint method more closely resembled f_2 sweep delays. However, it is unclear how the conservative α-level used in the changepoint method may have affected the results. Recall that since the penalized contrast method attempts to estimate all the changepoints in the data, Lavielle (2005) recommends a very small rejection region in order to eliminate false positives. It could be argued that in this application, there is substantial *a priori* information about the changepoints, which may permit a less stringent α level without the risk of false alarms. Relaxing the α level, would presumably not affect the hypothesis that the differences between the f_1 and f_2 sweeps is an artifact of the phase gradient and impulse response methods, but it could very well affect which delay appeared to be more accurate. Finally, it should be noted that visual inspection of high-SNR waveforms show the DPOAE beginning a few cycles before the estimated changepoint, but still in much better agreement with the f_2 sweep delays.

There are several cautionary notes. First and foremost, the sample size and frequencies used in this study were extremely limited. Frequencies lower than 2 kHz were not included because of the high noise environment and highly nonwhite background. Additionally, the CDP is further away in frequency space as the frequency increases, which permits better phase cancellation (and or filtering) of f_1. Still, the data acquisition error can be rectified and the 4 kHz f_2 sweep included in future work. It is also plausible to include 5 kHz and perhaps 6 kHz in future work. While these data provide limited support for the hypotheses that the delay differences between the f_1 and f_2 sweeps is artifactual and the delay time of the DPOAE is underestimated by the f_1 sweep method, further study

is needed. Further study is planned to build on these results. The filtering method used in this study may not be optimal and there is certainly additional work to be done in that regard. The measurement paradigm also should be given more consideration. Currently, there is more data collected in the f_1 sweep paradigm than in the f_2 sweep. This was done to replicate previous work, but in the changepoint method it was detrimental to the calculation of an average delay. With fewer measurements, it was less probable that enough CDPs would be elicited to produce an average delay estimate with reasonable variance. Finally, the flexibility of the changepoint method was not exploited to its fullest extent. For example, questions surrounding the validity of the second source model were not addressed in this study. Putative interference within the time series or an asymmetry between DPOAE onset and offset are questions which may be amenable to this type of changepoint analysis. If the time between the end of f_2 and the offset of the CDP is found to be significantly longer, than the time between the onset of f_2 and the CDP onset, that may provide evidence for a longer latency source.

In spite of its limitations, the study has resulted in some intriguing findings. It has provided several testable hypotheses to guide future work. Further, it has provided some theoretical basis for the observed differences in the f_1 and f_2 phase gradients. Finally, it has provided a methodological framework for DPOAE delay estimates which are more readily interpretable as true delays than the current methods.

References

Allen, J., & Neely, S. (1992). Micromechanical models of the cochlea. *Physics Today, 45*, 40-47.

Basseville, M., & Nikiforov, I. (1993). *Detection of abrupt changes-theory and applications.* Englewood Cliffs, N.J.: Prentice-Hall.

Boashash, B. (1992). Extimating and interpreting the instantaneous frequency of a signal - part 1: Fundamentals. *Proceedings of the IEEE, 80*, 520-538.

Bowman, D., Eggermont, J., Brown, D., & Kimberley, B. (1998). Estimating cochlear filter response properties from distortion product otoacoustic emission (dpoae) phase delay measurements in normal hearing adults. *Hearing Research, 119*, 14-26.

Brown, A., McDowell, B., & Forge, A. (1989). Acoustic distortion products can be used to monitor the effects of chronic gentamicin treatment. *Hearing Research, 42*, 143-156.

Camalet, S., Duke, T., Julicher, F., & Prost, J. (2000). Auditory sensitivity provided by self-tuned critical oscillations of hair cells. *PNAS, 97*, 3183-3188.

Cohen, L. (1995). *Time frequency analysis*. Englewood Cliffs: Prentice Hall.

Crawford, A., & Fettiplace, R. (1985). The mechanical properties of ciliary bundles in the turtle cochlear hair cells. *Journal of Physiology, 364*, 359-379.

Fahey, P., Stagner, B., & Martin, G. (2006). Mechanism for bandpass frequency characteristic in distortion product otoacoustic emission generation. *Journal of the Acoustical Society of America, 119*, 991-996.

Gaskill, S., & Brown, A. (1990). The behavior of the acoustic distortion product $2f_1 - f_2$, from the human ear and its relation to auditory sensitivity. *Journal of the Acoustical Society of America, 99*, 821-839.

Goodman, S., Withnell, R., & Shera, C. (2003). The origin of sfoae microstructure in the guinea pig. *Hearing Research, 183*, 7-17.

Harris, F., Lonsbury-Martin, B., Stagner, B., Coats, A., & Martin, K. (1989). Acoustic distortion products in humans: Systematic changes in amplitude as a function of $\frac{f_2}{f_1}$ ratio. *Journal of the Acoustical Society of America, 85*, 220-228.

He, N., & Schmiedt, R. (1993). Fine structure of the $2f_1 - f_2$ acoustic distortion product: Changes with primary level. *Journal of the Acoustical Society of America, 94*, 2659-2669.

Heitmann, J., Waldman, B., Schnitzler, H., Plinkert, P., & Zenner, H. (1998). Suppression of distortion product otoacoustic emissions (dpoaes) near the 2f1 - f2 removes dp-gram fine structure: Evidence for a secondary generator. *Journal of the Acoustical Society of America, 103*, 1527-1531.

Jaramillo, F., Markin, V., & Hudspeth, A. (1993). Auditory illusions and the single hair cell. *Nature*, *364*, 527-529.

Kalluri, R., & Shera, C. (2001). Distortion-product source unmixing: A test of the two-mechanism model for dpoae generation. *Journal of the Acoustical Society of America*, *109*, 623-637.

Kemp, D. (1978). Stimulated acoustic emissions from within the human auditory system. *Journal of the Acoustical Society of America*.

Kemp, D., & Brown, A. (1983). An integrated view of cochlear mechanical nonlinearities observable from the ear canal. In E. de Boer & M. Viergever (Eds.), *Mechanics of hearing* (p. 75-82). The Hague: Martinus Nijhoff.

Kimberley, B., Brown, D., & Eggermont, J. (1993). Measuring human cochlear traveling wave delay using distortion product emission phase responses. *Journal of the Acoustical Society of America, 94*, 1343-1350.

Knight, R., & Kemp, D. (2000). Indications of different distortion product otoacoustic emission mechanisms from a detailed f_1, f_2 area study. *Journal of the Acoustical Society of America, 107*, 457-473.

Knight, R., & Kemp, D. (2001). Wave and place fixed dpoae maps of the human ear. *Journal of the Acoustical Society of America*, *109*, 1513-1525.

Konrad-Martin, D., Neely, S., Keefe, D., Dorn, P., Cyr, E., & Gorga, M. (2002). Sources of dpoaes revealed by suppression experiments, inverse fast fourier transforms, and sfoaes in normal ears. *Journal of the Acoustical Society of America*, *111*, 1800-1809.

Konrad-Martin, D., Neely, S., Keefe, D., Dorn, P., & Gorga, M. (2001). Sources of dpoaes revealed by suppression experiments and inverse fast fourier transforms in normal ears. *Journal of the Acoustical Society of America*, *109*, 2862-2879.

Kummer, P., Janssen, T., Hulin, P., & Arnold, W. (2000). Optimal $l_1 - l_2$ primary tone level separation remains independent of test frequency in humans. *Hearing Research*, *146*, 47-56.

Lavielle, M. (1999). Detection of multiple changes in a sequence of of dependent variables. *Stochastic Processes and their Appliciations, 83*, 79-102.

Lavielle, M. (2005). Using penalized contrasts for the change-point problem. *Signal Processing, 85*, 1501-1510.

Lavielle, M., & Ludena, C. (2000). The multiple change-points problem for the spectral distribution. *Bernoulli, 6*, 845-869.

Mahoney, C., & Kemp, D. (1995). Distortion product otoacoustic emission delay measurement in human ears. *Journal of the Acoustical Society of America, 97*, 3721-3735.

Martin, G., Stagner, B., Jassir, D., Telischi, F., & Lonsbury-Martin, B. (1999). Suppression and enhancement of distortion-product otoacoustic emissions by interference tones above f_2. i. basic findings in rabbits. *Hearing Research, 136*, 105-123.

Mills, D. (1997). Interpretation of distortion product otoacoustic emission measurements. *Journal of the Acoustical Society of America*, *102*, 413-429.

Moulin, A., & Kemp, D. (1996). Multicomponent acoustic distortion product otoacoustic emission phase in humans. i. general characteristics. *Journal of the Acoustical Society of America, 100*, 1617-1639.

Oghalai, J. (2004). The cochlear amplifier: Augmentation of the traveling wave within the inner ear. *Curr Opin Otolaryngol Head Neck Surg*, *12*, 431-438.

Picard, D. (1985). Testing and estimating changepoints in time series. *Advances in Applied Probability*, *17*, 841-867.

Prijs, B., Schneider, S., & Schoonhoven, R. (2000). Group delays of distortion product otoacoustic emissions: Relating delays measured with f_1- and f_2- sweep paradigms. *Journal of the Acoustical Society of America*, *107*, 3298-3307.

Rhode, W. (1971). Observations of the vibration of the basilar membrane in squirrel monkeys using the mössbauer technique. *Journal of the Acoustical Society of America*, *49*, 1218-1231.

Ruggero, M. (2004). Comparison of group delays of 2 $f_1 - f_2$ distortioin product otoacoustic emissions and cochlear travel times. *ARLO*, *5*, 143-147.

Shera, C. (2004). Mechanisms of mammalian otoacoustic emission and their implications for the clinical utility of otoacoustic emissions. *Ear and Hearing*, *25*, 86-97.

Shera, C., & Guinan, J. (1999). Evoked otoacoustic emissions arise by two fundamentally different mechanisms:a taxonomy for mammalian oaes. *Journal of the Acoustical Society of America*, *105*, 782-798.

Stover, L., Gorga, M., Neely, S., & Montoya, D. (1996). Toward optimizing the clinical utility of distortion product otoacoustic emission measurements. *Journal of the Acoustical Society of America*, *100*, 956-967.

Stover, L., Neely, S., & Gorga, M. (1996). Latency and multiple sources of distortion product emissions. *Journal of the Acoustical Society of America*, *99*, 1016-1024.

Talmadge, C., Long, G., Tubis, A., & Dhar, S. (1999). Experimental confirmation of the two-source interference model for the fine structure of distortion product otoacoustic emissions. *Journal of the Acoustical Society of America*, *105*, 275-292.

Tubis, A., Talmadge, C., & Tong, C. (2000). Modeling the temporal behavior of distortion product otoacoustic emissions. *Journal of the Acoustical Society of America*, *107*, 2112-2127.

Tubis, A., Talmadge, C., Tong, C., & Dhar, S. (2000). On the relationships between the fixed- f_1, fixed- f_2, and fixed-ratio phase derivatives of the 2 $f_1 - f_2$ distortion product otoacoustic emission. *Journal of the Acoustical Society of America*, *108*, 1772-1785.

Whitehead, M., Stagner, B., Martin, G., & Lonsbury-Martin, B. (1996). Visualization of the onset of distortion-product otoacoustic emissions, and measurement of their latency. *Journal of the Acoustical Society of America*, *100*, 1663-1670.

Yates, G., & Whitnell, R. (1999). The role of intermodulation distortion in transient-evoked otoacoustic emissions. *Hearing Research*, *136*, 49-64.

Zweig, G., & Shera, C. (1995). The origin of periodicity in the spectrum of evoked otoacoustic emissions. *Journal of the Acoustical Society of America*, *98*, 2018-2047.