Human Auditory Source Discrimination

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Stephen J. Gardeen

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Dr. John Broadhurst & Dr. Art Leuthold

May, 2011

Acknowledgements

There are many people that have earned my gratitude for their contribution to my time in graduate school.

I am very thankful to my advisors, John Broadhurst and Art Leuthold, whose guidance and support from the initial to the final level enabled me to develop an understanding of the subject.

I'd especially like to acknowledge the significant contribution of my parents Gary and Mary Gardeen. Their steady support sustained me through the lows of the graduate school experience.

I'd also like to express my appreciation for the support and encouragement of Addishiwot Woldesenbet who was the willing audience for many practice talks. I'm thankful to the many conversations with Kimia Ghanbeigi who I worked with on this research. I'm deeply thankful to all those who volunteered for this research study (volunteers with appropriate MRIs are hard to find). Lastly, I offer my regards and blessings to all of those who supported me in any respect during the completion of this project.

Steve Gardeen

Dedication

I dedicate this thesis to Isaiah, Elijah, Samuel, Anna, and Jonah.

Abstract

The aim of this study is to examine the auditory system's ability to process low energy frequency transpositions of complex sounds. The auditory processing of complex sounds such as musical instruments, voice, or environmental events is currently an active area of research. Some propose that auditory objects are represented by neurons which encode the invariant spectro-temporal acoustic properties (Griffiths & Warren,2004). These sound features tend to be heavily damped and very transient and, therefore, frequency rich. This study shows the auditorys sensitivity to detect these adjustments by detecting the pre-attentive magnetic mismatch response (MMNm) from 8 subjects passively listening to complex audio stimulations. Responses were detected from most subjects even though participants could not attentively discriminate the sounds. This result is somewhat controversial in that current views suggests that a mismatch response indicates processing that is available for higher processing i.e. should be attentively discriminable (Näätänen et al., 2010). Localization results suggest the mismatch processing is performed in the auditory associations regions (superior temporal sulcus, insula) of the auditory cortex. These results suggest that transient sounds might be essential to auditory object identification, and that the auditory system is able to distinguish at a sensory level, shifts in the heavily damped spectral structure of complex sounds even though some cannot do so attentively. This could be due to the greater analysis given by the human brain to determining the pitch center rather then the sound timbre, yet some trained musicians have the ability to distinguish these subtle differences (i.e. differences between manufacturers of the same kinds of instruments).

Contents

List of Tables

List of Figures

Chapter 1

Introduction

- Chapter 2 introduces the basic principles of vibrations and sound that are useful in understanding the nature of the work and stimulus design.
- In Chapter 3 an overview of the human auditory system is presented.
- Chapter 4 describes how brain activity can be understood by detecting magnetic fields.
- Chapter 5 details the proposed stimulus design model used in the experiment.
- Chapter 6 details data acquisition and analysis procedures with the results.
- Chapter 7 presents a final discussion of the analysis presented in the thesis.

Chapter 2

Sound

2.1 Waves and Vibrations

By understanding sound and the processes of its production, we will begin to be able to understand acoustic object identification. Acoustic properties belong to and characterize the source. Sound is the propagation of mechanical energy through vibrations of matter. Internal elastic restoring forces occur when atoms or molecules are perturbed from their equilibrium configurations. These restoring forces along with the inertia of the system produce the oscillatory behavior that generate and transmit sound waves. These waves oscillate in the direction of propagation creating a series of compressions and rarefications. Sound waves can be expressed and understood as displacements, $\mathbf{x}$, or in the case of fluids, changes in pressure, p. Solutions to the following wave equation are determined by applying appropriate boundary and initial conditions.

$$\frac{1}{v^2}\frac{\partial^2 \mathbf{x}}{\partial t^2} = \nabla^2 \mathbf{x} \qquad \frac{1}{v^2}\frac{\partial^2 p}{\partial t^2} = \nabla^2 p \tag{2.1}$$

The phase velocity, v, is determined by the properties of the medium in which the sound energy propagates.[1] The general solution for the pressure wave equation may

[1] The pressure wave equation is explained in more detail at the end of this chapter.

be expressed as,

$$p(x,t) = \phi_1(t - \frac{x}{v}) + \phi_2(t + \frac{x}{v}) \tag{2.2}$$

The function ϕ_1 represents a wave moving in the positive x direction and ϕ_2 represents a wave moving in the negative x direction. Regardless of the shape of ϕ_1 or ϕ_2, the wave pulse, propagates unchanged at a speed of sound v.

In most cases, the presence of one sound wave does not alter the medium to an extent that the properties of other waves are seriously affected. This indicates that a complicated vibration can be expressed as the superposition of individual vibrations. This is the property of linear superposition and is not true in the case of shockwaves. An undamped single vibration can be described by Newton's 2$^{\text{nd}}$ Law with a linear restoring force. Although most sound is produced by processes with multiple degrees of freedom, hearing is perceived by two one-degree of freedom oscillators... your ears. The physics of a single degree of freedom oscillator is expressed as,

$$m\frac{d^2x}{dt^2} = -kx \tag{2.3}$$

where k is the linear coefficient of the elastic force and m is the mass. This motion is also known as simple harmonic motion which can be expressed with sine and cosine functions or, with the use of Euler's formula, complex exponentials.

$$e^{i\theta} = \cos\theta + i\sin\theta \tag{2.4}$$

An undamped single vibration or tone maintains the same energy and therefore the same amplitude for all time. The sine and cosine have these properties in that their amplitudes do not change for arguments extending from -∞ to ∞. To account for damping, a resistive force proportional to velocity combines with the elastic force to accelerate a mass. Again, Newton's 2$^{\text{nd}}$ Law results in a differential equation for a

damped vibration that is,

$$m\frac{d^2x}{dt^2} = -kx - c\frac{dx}{dt} \tag{2.5}$$

where c is the damping constant. Before solving the differential equation, one property that can be expected is more transient behavior; oscillation amplitudes won't extend from -∞ to ∞. Since energy is lost to damping each cycle and based on the amount of damping, we can expect different oscillations. A very large damping constant, c, will lead to an oscillation that is so transient that the energy is dissipated before even one complete cycle. A very low damping constant would result in an oscillation that dies out slowly with many cycles occurring before that amplitude changes substantially. The differential equation can be expressed as,

$$\frac{d^2x}{dt^2} + \frac{c}{m}\frac{dx}{dt} + \frac{k}{m}x = 0 \tag{2.6}$$

If the damping and the elastic parameters are reformulated as follows the differential equation is then,

$$\tau \equiv \frac{2m}{c} \qquad \omega_0^2 \equiv \frac{k}{m} \tag{2.7}$$

$$\frac{d^2x}{dt^2} + \frac{2}{\tau}\frac{dx}{dt} + \omega_0^2 x = 0 \tag{2.8}$$

Here ω_0 is the natural frequency of the oscillation as it has units of radians/seconds and τ is the decay time constant with units of seconds. Due to the presence of damping, one can verify that simple sine and cosine functions are no longer solutions. One method of solution is to set the differential operator $\frac{d}{dt} \Rightarrow D$ and $\frac{d^2}{dt^2} \Rightarrow D^2$ which results in the equation

$$(D^2 + \frac{2}{\tau}D + \omega_0^2)x = 0 \tag{2.9}$$

The simplification arises when the operator is factored; one way is to use the quadratic equation.

$$(D + \frac{1}{\tau} + \frac{\sqrt{1-(\tau\omega_0)^2}}{\tau})(D + \frac{1}{\tau} - \frac{\sqrt{1-(\tau\omega_0)^2}}{\tau})x = 0 \tag{2.10}$$

A second order differential equation is now reduced to two first order operations. The order of operation is not important so to find the general solution, each first order operation is set to zero which results in two independent exponentials.

$$\lambda_{\pm} = -\frac{1}{\tau} \pm \frac{\sqrt{1-(\tau\omega_0)^2}}{\tau} \tag{2.11}$$

$$x(t) = A_1 e^{\lambda_+ t} + A_2 e^{\lambda_- t} \tag{2.12}$$

Here A_1 and A_2 are coefficients that can be determined by initial conditions. At this point, it is important to pay attention to the quantity under the radical, $\tau\omega_0$. If this dimensionless value is greater than 1, λ, the eigenvalue of the differential equation, would be imaginary. So to better understand the dynamics here, its important to understand the meaning and range of this parameter: $\tau\omega_0$.

From equation 2.7, since both τ and ω_0 are positive, the behavior of the system can be classified into 3 different scenarios all based on the value of the parameter $\tau\omega_0$ and its relation to unity.

Table 2.1: Damping Behavior

$\tau\omega_0$	λ	Damping
<1	real	Over-damped
$=1$	degenerate	Critically Damped
>1	imaginary	Underdamped(Oscillatory)

Based on equation 2.12, over-damped solutions with their real eigenvalues are just two linearly independent exponentials. The situation is less clear for critically damped solutions. Here the eigenvalues are identical making the two constants A_1 and A_2 no

longer independent. In order to get the final solution, the entire operation on $x(t)$ needs to be carried out before it can be set to zero. Equation 2.10 reduces to:

$$(D + \frac{1}{\tau})(D + \frac{1}{\tau})x = 0 \tag{2.13}$$

$$u \equiv (D + \frac{1}{\tau})x \qquad (D + \frac{1}{\tau})u = 0 \Rightarrow u = A_0 e^{-\frac{t}{\tau}} \tag{2.14}$$

$$A_1 e^{-\frac{t}{\tau}} = (D + \frac{1}{\tau})x \tag{2.15}$$

This first order ordinary differential equation can be solved by making use of the product rule and integrating. ($\frac{d}{dt}(xy) = \dot{x}y + x\dot{y}$. Using the notation $\frac{dz}{dt} \rightarrow \dot{z}$ to represent a time derivative.)

$$A_1 = e^{\frac{t}{\tau}}(D + \frac{1}{\tau})x = D(xe^{\frac{t}{\tau}}) \tag{2.16}$$

$$A_1 t + A_1 = xe^{\frac{t}{\tau}} \Rightarrow x(t) = e^{-\frac{t}{\tau}}(A_1 t + A_2) \tag{2.17}$$

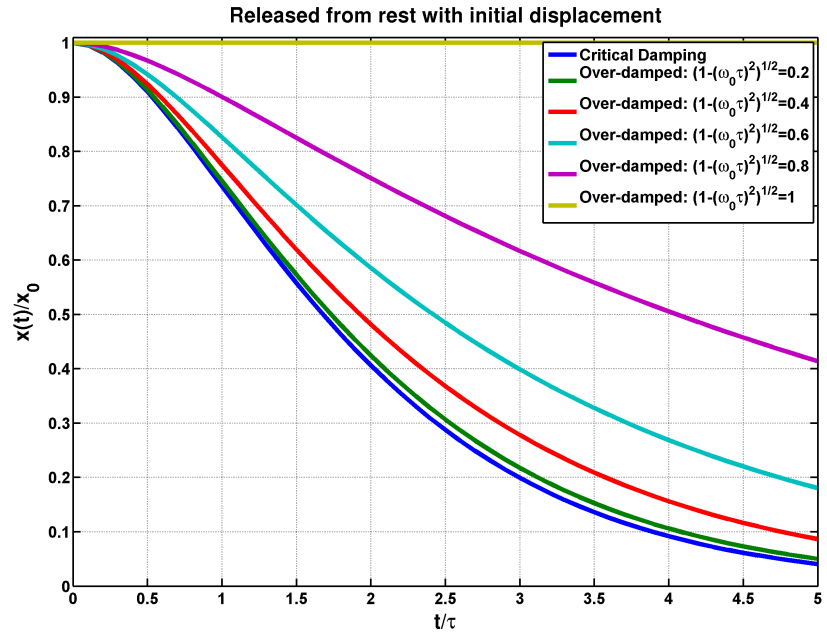
Again, A_1 and A_2 are constants determined by initial conditions. Some interesting comparisons can be made between an over-damped harmonic oscillator and a critically damped harmonic oscillator. When both systems are under the same initial conditions with identical damping constants, τ, the critically damped system returns to equilibrium faster then the over-damped system (Fig. 2.1) leading to more transient behavior. This will be useful when designing synthetic sound samples. The over-damped and critically damped position functions are, respectively,

$$x(t) = A_1 e^{-\frac{t}{\tau}(1+\sqrt{1-(\omega_o\tau)^2})} + A_2 e^{-\frac{t}{\tau}(1-\sqrt{1-(\omega_o\tau)^2})} \tag{2.18}$$

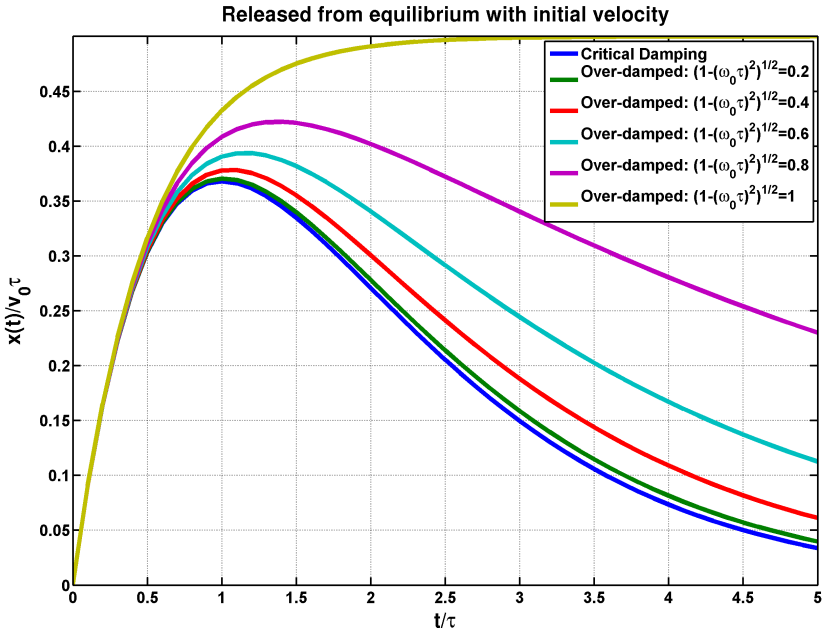
$$x(t) = A_1 e^{-\frac{t}{\tau}} + A_2 t e^{-\frac{t}{\tau}} \tag{2.19}$$

The constants A_1 and A_2 are then calculated to give similar initial conditions in both

Figure 2.1: Critically damped and over-damped displacement and velocity v. time graphs. The constants A_1 and A_2 are calculated to give similar initial conditions in both cases.



(a) initial displacement



(b) initial velocity

cases. For both situations, an initial displacement or an initial velocity, the system returns to equilibrium fastest when critically damped. To understand the behavior when the system is underdamped, the general solution (equation 2.12) with Euler's formula (equation 2.4) can be expressed as:

$$\omega_d \equiv \frac{\sqrt{(\omega_o\tau)^2 - 1}}{\tau} \tag{2.20}$$

$$x(t) = A_1 e^{-\frac{t}{\tau}} e^{i\omega_d t} + A_2 e^{-\frac{t}{\tau}} e^{-i\omega_d t} \tag{2.21}$$

$$\omega_0\tau \gg 1 \Rightarrow \omega_d \approx \omega_0 \tag{2.22}$$

It is under the conditions of underdamping that oscillations can now be expressed with respect to their natural frequency, ω_0 and their damping constant, τ; the systems are actually vibrating. When $\tau\omega_0 \gg 1$, this can be understood as many cycles occurring before the amplitude of the oscillation is significantly reduced due to the damping. This is the condition of severe underdamping and is necessary for temporally finite vibrations. Not only is the natural frequency of a vibration a relevant parameter but also the extent of its damping or its transient nature. Both of these properties are important in understanding how vibrations are perceived.

At this point, all that has been discussed are undriven harmonic oscillators. To conclude the discussion of a single degree of freedom oscillator, it is important to understand the behavior when an oscillator is driven by a periodic force, $f(t) = F_0 e^{i\omega t}$. Here Newton's 2$^{\text{nd}}$ Law can be expressed as:

$$m\ddot{x} + \frac{2m}{\tau}\dot{x} + \omega_0^2 m x = F_0 e^{i\omega t} \tag{2.23}$$

$$x(t) = Ae^{i\omega t} \Rightarrow (-\omega^2 m + i\frac{2\omega m}{\tau} + \omega_0^2 m)x = F_0 e^{i\omega t} \tag{2.24}$$

Assuming an exponential for the position function and differentiating to find the complex speed is

$$u(t) = \frac{F_0 e^{i\omega t}}{\frac{2m}{\tau} + i\frac{m}{\omega}(\omega^2 - \omega_0^2)} \tag{2.25}$$

$$\tag{2.26}$$

A useful definition of the ratio of the complex driving force, $f(t)$, to the resultant complex speed, $u(t)$, at the point of force is termed the mechanical impedance, Z.

$$\begin{aligned} Z &\equiv \frac{f}{u} && (2.27)\\ &= R_m + iX_m && (2.28)\\ &= \frac{2m}{\tau} + i\frac{m}{\omega}(\omega^2 - \omega_0^2) && (2.29) \end{aligned}$$

where $\frac{2m}{\tau}$ is the mechanical resistance, R_m, and $\frac{m}{\omega}(\omega^2 - \omega_0^2)$ is the mechanical reactance, X_m. Since $u = i\omega x$, knowledge of the Z is equivalent to solving the differential equation.

$$x(t) = \frac{f(t)}{i\omega Z} \tag{2.30}$$

where the actual displacement is given by taking the real part so that the amplitude: $A(\omega) = \frac{|f|}{\omega|Z|}$.

$$\begin{aligned} A(\omega) &= \frac{\frac{F_0}{m}}{\sqrt{(\omega_0^2 - \omega^2)^2 + (\frac{2\omega}{\tau})^2}} && (2.31)\\ \omega_r &= \frac{\sqrt{(\omega_0\tau)^2 - 2}}{\tau} && (2.32)\\ \omega_0\tau \gg 1 \Rightarrow \omega_r &\approx \omega_0 && (2.33) \end{aligned}$$

ω_r is the frequency of maximum amplitude or the resonant frequency. It is only in the severe underdamping condition that the natural frequency is roughly the same as the resonant frequency. The discussion so far has led to the point where now a definition can be made that relates the transient nature of an oscillation to its natural frequency. The sharpness of the resonance peak not only indicates the frequency of maximum amplitude but also the transient nature of resonant vibration. To understand this, $A(\omega)$, in the condition of severe underdamping can be simplified to:

$$A_{max} \equiv \frac{\frac{F_0}{m}\tau}{2\omega_0} \qquad \begin{aligned} \omega_0^2 - \omega^2 &= (\omega_0 - \omega)(\omega_0 + \omega) \approx 2\omega_0(\omega_0 - \omega) && (2.34)\\ A(\omega) &= \frac{A_{max}}{\tau\sqrt{(\omega_0 - \omega)^2 + (\frac{1}{\tau})^2}} && (2.35) \end{aligned}$$

In the case that $|\omega_0 - \omega| = \frac{1}{\tau}$ which is the same as $\omega = \omega_0 \pm \frac{1}{\tau}$. The energy is then equal to half its maximum, remembering that energy is proportional to the square of the amplitude.

$$A^2 \quad = \tfrac{1}{2} A_{max}^2 \tag{2.36}$$

Now it is clear that $\frac{1}{\tau}$ is a parameter for the width of the resonance peak when it is represented in frequency space. So a slightly damped resonance is represented by frequencies that surround the natural frequency. The width of this frequency range is

$$\Delta\omega = \frac{2}{\tau} \tag{2.37}$$

With this information relationships can now be made that relate the transient nature of on oscillation with its damping factor, τ, and the sharpness or efficiency of the resonance.

Table 2.2: Resonant Behavior

temporal properties	damping	resonant sharpness
highly transient	small τ	large $\Delta\omega$
slightly transient	large τ	small $\Delta\omega$

A more precise way to quantify the sharpness of a resonance is defined by what is called a quality factor, Q. For an undriven oscillator, the quality factor can be understood as the ratio of energy available each cycle to the energy dissipated due to damping. It is defined for resonant behavior as the ratio of the resonant frequency to the width at half the maximum energy.

$$Q \equiv \frac{\omega_0}{\Delta\omega} = \frac{1}{2}\tau\omega_0 \tag{2.38}$$

Values of Q can obviously vary but in the severe underdamping condition are much greater than unity. When representing a vibration temporally, transient properties can easily be discerned but frequencies are not as clear. However, if vibrations can

be represented spectrally or with respect to their frequency, not only are frequencies straight forward but the width of the frequency peaks give valuable timing information. The following section will show how Fourier analysis can be used to represent vibrations in frequency space as well as in time space. The complete general solution including the transient response (eqn. 2.21) with the steady-state solution (eqn. 2.30) where the real part represents the actual displacement

$$x(t) = A_1 e^{-\frac{t}{\tau}} e^{i\omega_d t} + A_2 e^{-\frac{t}{\tau}} e^{-i\omega_d t} + \frac{f(t)}{i\omega Z} \tag{2.39}$$

2.2 Fourier Analysis

Fourier analysis arises from the powerful concepts of linear superposition and orthogonality. Linear superposition, as already stated, can be understood physically in the context of waves. For sound waves, the actual wave disturbance is small enough so that the effects of a single wave going through the medium does not alter the medium for other waves. This then suggests that the total effect from many waves is just the linear addition or superposition of many waves.

The concept of orthogonality is best understood in the context of vectors. Vectors exist in N-dimensional spaces. The most familiar vector space is the 3-dimensional physical space in which we exist. Two vectors in a space are considered orthogonal if their inner product (dot product) is zero. In physical space, this is when the angle between them is 90 degrees. An orthonormal basis set is a set of vectors whose length is unity and whose orientations are all mutually orthogonal. For an N-dimensional system, there can exist at most N orthonormal vectors. And finally, every single point in the N-dimensional space can be expressed as a linear superposition (addition) of the scaled orthonormal basis vectors.

$$\mathbf{X} = a_0\mathbf{x}_0 + a_1\mathbf{x}_1 + a_2\mathbf{x}_2 \cdots + a_n\mathbf{x}_n \tag{2.40}$$

The $\mathbf{x}_n$ are the orthonormal basis vectors and a_n are the coefficients that scale the basis vectors. A wave or periodic function can be understood in this context by extending the N-dimensional space to an infinite dimensional space and defining an inner product

for an orthonormal basis set. So first represent a function $f(t) \rightarrow \mathbf{f}(t)$ as a vector where the basis is each time sample.

$$\mathbf{f}(t) = f(t_{(-\infty)})\mathbf{t}_{(-\infty)} + \cdots + f(t_0)\mathbf{t}_0 + \cdots + f(t_\infty)\mathbf{t}_\infty \tag{2.41}$$

A representation of a function in the frequency basis, $\hat{f}(\omega) \rightarrow \hat{\mathbf{f}}(\omega)$, would have a similar form where each frequency sample is then a basis vector.

$$\hat{\mathbf{f}}(\omega) = \hat{f}(\omega_{(-\infty)})\mathbf{\Omega}_{(-\infty)} + \cdots + \hat{f}(\omega_0)\mathbf{\Omega}_0 + \cdots + \hat{f}(\omega_\infty)\mathbf{\Omega}_\infty \tag{2.42}$$

Transforming a function from one representation to another is done using the orthogonality of the basis sets to express the basis vectors in the complementary basis space and the performing an inner product to calculate the coefficient of each basis vector.

$$\mathbf{\Omega}_j = e^{-i\omega_j t} = e^{-i\omega_j t_{(-\infty)}}\mathbf{t}_{(-\infty)} + \cdots + e^{-i\omega_j t_0}\mathbf{t}_0 + \cdots + e^{-i\omega_j t_\infty}\mathbf{t}_\infty \tag{2.43}$$

$$\mathbf{t}_j = e^{i\omega t_j} = e^{i\omega_{(-\infty)} t_j}\mathbf{\Omega}_{(-\infty)} + \cdots + e^{i\omega_0 t_j}\mathbf{\Omega}_0 + \cdots + e^{i\omega_\infty t_j}\mathbf{\Omega}_\infty \tag{2.44}$$

The sine and cosine functions, expressed using Euler's formula, do have this orthogonality property in that the inner product of two harmonic functions of different frequencies is zero. The inner product of identical frequencies is unity. This is what defines the Kronecker delta δ_{ij}.

$$\mathbf{\Omega}_i \cdot \mathbf{\Omega}_j = \delta_{ij} \qquad \mathbf{t}_i \cdot \mathbf{t}_j = \delta_{ij} \tag{2.45}$$

$$\hat{\mathbf{f}}(\omega) = \sum_{j=-\infty}^{\infty} (\mathbf{f}(t) \cdot \mathbf{\Omega}_j)\mathbf{\Omega}_j \qquad \mathbf{f}(t) = \sum_{j=-\infty}^{\infty} (\hat{\mathbf{f}}(\omega) \cdot \mathbf{t}_j)\mathbf{t}_j \tag{2.46}$$

The inner product for two arbitrary functions $A(t)$ and $B(t)$ can be expressed for continuous functions as integrals.

$$\mathbf{A}(t) = A(t_{(-\infty)})\mathbf{t}_{(-\infty)} + \cdots + A(t_0)\mathbf{t}_0 + \cdots + A(t_\infty)\mathbf{t}_\infty \tag{2.47}$$

$$\mathbf{A}(t) \cdot \mathbf{B}(t) = \sum_{j=-\infty}^{\infty} A(t_j)B(t_j) \Rightarrow \int A(t)B(t)dt \tag{2.48}$$

For sines, cosines, and complex exponentials in continuous space the inner product is

$$\mathbf{cos}(nt) \cdot \mathbf{cos}(mt) = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos(nt)\cos(mt)dt = \delta_{mn} \tag{2.49}$$

$$\mathbf{sin}(nt) \cdot \mathbf{sin}(mt) = \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(nt)\sin(mt)dt = \delta_{mn} \tag{2.50}$$

$$\mathbf{e}^{int} \cdot \mathbf{e}^{imt} = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{int}e^{-imt}dt = \delta_{mn} \tag{2.51}$$

The orthogonality relations can be verified by computing the appropriate integrals. The definition of the Fourier series is then

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(\frac{\pi nt}{T}) + b_n \sin(\frac{\pi nt}{T}) \tag{2.52}$$

$$a_n = \frac{1}{T}\int_{-T}^{T} f(t)\cos(\frac{\pi nt}{T})dt \tag{2.53}$$

$$b_n = \frac{1}{T}\int_{-T}^{T} f(t)\sin(\frac{\pi nt}{T})dt \tag{2.54}$$

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{i\frac{\pi nt}{T}} \tag{2.55}$$

$$c_n = \frac{1}{2T}\int_{-T}^{T} f(t)e^{-i\frac{\pi nt}{T}}dt \tag{2.56}$$

Any function $f(t)$ with period 2T can be expressed as a Fourier series. a_n and b_n are the real cosine and sine coefficients, respectively. The complex coefficient, c_n, represents the amplitude and phase of the nth harmonic term. For sound this means that any complex vibration can be expressed as a linear combination of simple harmonic (sine or cosine) terms of particular amplitudes and frequencies. These are commonly referred to as Fourier components. When the period, 2T, is extended to infinity, $T \to \infty$. Any function in time, regardless of the period can be transformed into a complex function of frequencies and corresponding phases and amplitudes by computing its Fourier Transform. The inverse transform is the operation to transform a frequency distribution into a time varying representation.

$$\hat{f}(w) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt \tag{2.57}$$

$$f(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \hat{f}(\omega)e^{i\omega t}d\omega \tag{2.58}$$

Before addressing more complicated vibrating systems, the observations of a Fourier transform of a gaussian will reinforce the main concept of this section thus far. A frequency representation of a vibrating system not only informs of spectral properties but also how those spectral properties act in time. This has been explained in the

context of single degree of freedom damped oscillatory system. With Fourier Analysis, this concept generalizes to any wave or oscillatory system. A normalized gaussian is

$$f(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\bar{t})^2}{2\sigma^2}} \tag{2.59}$$

Since the Fourier transform results in a complex function, the absolute magnitude of the result can be understood as the amplitude of each frequency component. Phase information is lost. The transform of the gaussian is then

$$|\hat{f}(\omega)| = \frac{e^{-\frac{(\sigma\omega)^2}{2}}}{\sqrt{2\pi}} \tag{2.60}$$

which is also a gaussian but with mean zero and the deviation is the inverse of the time representation deviation $\sigma \to \frac{1}{\sigma}$. A qualitative observation is that 'skinny' functions in time are 'fat' functions in frequency and 'fat' functions in time are 'skinny' functions in frequency. Therefore, a highly transient vibration will 'activate' more frequencies then a less transient vibration of similar spectral structure.

2.3 Analyzing Complex Vibrations

In order to comprehend how the brain might identify a particular sound source, Fourier analysis is used to represent a complicated vibration in terms of its frequency components or resonances. This is useful to understand the properties of the frequency structure of a particular acoustic signal which is certainly analyzed in order to identify its source. When an object is acoustically excited, whether it being a burst of air or the plucking of a string (a delta burst), waves of many frequencies radiate from the source and reverberate through the object. However, according to the geometry or boundary conditions of the object, waves of particular frequencies (called the resonant frequencies) interfere constructively and form what are called standing waves. Many vibrating systems are complex with mass distributed throughout a volume of complicated shape. Some systems are not as complex and are a good starting point for understanding the modes of vibrations for more complicated systems. A string is an example of such a system (see appendix A). The vibrations of a bar (see appendix B) represent a system that is slightly more complicated. These treatments show how an objects's physical properties relate to its frequency structure.

The spatial dependence on the selectivity of the resulting frequencies or Fourier components is known as a normal mode. The normal modes of an instrument enrich the frequency spectrum of the sound heard from the instrument giving it timbre. Timbre is defined by what it is not, according to the definition (Psychoacoustical terminology. timbre. ANSI S3.20-19731973), it is the acoustic property that characterizes a sound source that is not pitch or loudness and does not depend on duration. Timbre has also been characterized by a multidimensional nature. Possible dimensions include not only the frequency spectrum shape but also differences in the onset or offset of particular frequency components. Although some properties have been identified as characterizing timbre it is impossible to claim that any particular pattern of vibration or onset/offset shape define a source as these properties vary across sounds from the same source. The transient nature of sound due to there damping is of considerable importance when identifying sources. Remarkably, in an orchestra the human auditory system is able to keep track of the individual instruments by 'remembering' and classifying the instruments by the pattern of their transients even though it is a fraction of the total output acoustic power.

Since sounds are perceived by variations of air pressure, the conclusion of this chapter will introduce a few fundamental properties for sound waves in fluids. The states of matter that are of interest for human hearing are primarily air and liquid (inside the cochlea) both of which are fluids. The study of fluids is founded on the concept of a fluid particle. These particles are elemental volumes which are large enough to contain millions of molecules as to retain the idea of a fluid being a continuous medium yet small enough so that the macroscopic properties are nearly uniform. This leads to the concept of visualizing each particle having a displacement and a velocity. It is the vibrations of these particles that constitute the abstract concept of sound waves.

Applying the concepts of mass conservation and Newton's 2nd law, appropriately expressed for analyzing fluid elements, results in two relationships that when combined yield a partial differential equation that is known as the wave equation. Here ρ is instantaneous mass density and $\vec{V}$ is the velocity vector of the fluid particle.

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{V}) = 0 \tag{2.61}$$

The first relationship is known as the continuity equation. The first term refers to

changes of the mass inside the elemental volume and the second term refers to the flux of mass going in or out of the elemental volume. Since mass cannot be created or destroyed the terms sum to zero.

$$\rho \frac{d\vec{V}}{dt} = -\nabla p \tag{2.62}$$

The second relationship is commonly referred to as Euler‘s equation. For the most common acoustical phenomena the fluid can be assumed to be inviscid. If viscous forces were important, the right hand side of the equation would contain more terms. Newton‘s 2nd law relates the net force on a mass to its acceleration. The net force acting on a fluid particle is due to the spatial dependence of the pressure, p, that is exerted on the particle by the fluid particles surrounding it. The key assumption here is the spatial continuity of the fluid pressure.

Before proceeding it is important to state that for most of the observed acoustical phenomena the density fluctuations are small compared to the equilibrium value of the fluid. This means that the dependence of the pressure on the density can be approximated to be linear, in Taylor's Series ‘fashion’, (here p is the absolute pressure; otherwise it is the gauge pressure)

$$p = \left(\frac{\partial p}{\partial \rho}\right) \rho \tag{2.63}$$

However, the evaluation of the derivative requires knowledge of the thermodynamic process undergone by the fluid as the wave passes. Since the exchange of thermal energy between fluid particles is negligible the process is therefore adiabatic. Finally, by taking appropriate derivatives the two equations combine to form the wave equation,

$$\frac{1}{v^2}\frac{\partial^2 p}{\partial t^2} = \nabla^2 p \tag{2.64}$$

where v is the phase velocity of the sound wave and can be evaluated from the following relationship.

$$v^2 = \left(\frac{\partial p}{\partial \rho}\right)_{adiabatic:p=p_0} \tag{2.65}$$

$$= \gamma \frac{p_0}{\rho_0} \tag{2.66}$$

$$= \gamma r T_k \tag{2.67}$$

where γ is the ratio of specific heats, r is the specific gas constant, and T_k is the absolute temperature in Kelvin. This last permutation is from a manipulation of the ideal gas law. A final note is that the specific acoustic impedance in a fluid medium is $Z \equiv \frac{p}{u}$, the gauge pressure over particle velocity.

Timbre and object identification are an emerging transient property consisting of an evolving combination of acoustic properties. The focus of this study is to construct audio samples modeling that of a guitar that will examine the brain's response due to carefully adjusting temporal and spectral properties of the model. Understanding how acoustic stimuli are submitted to brain processing is the focus of the next chapter.

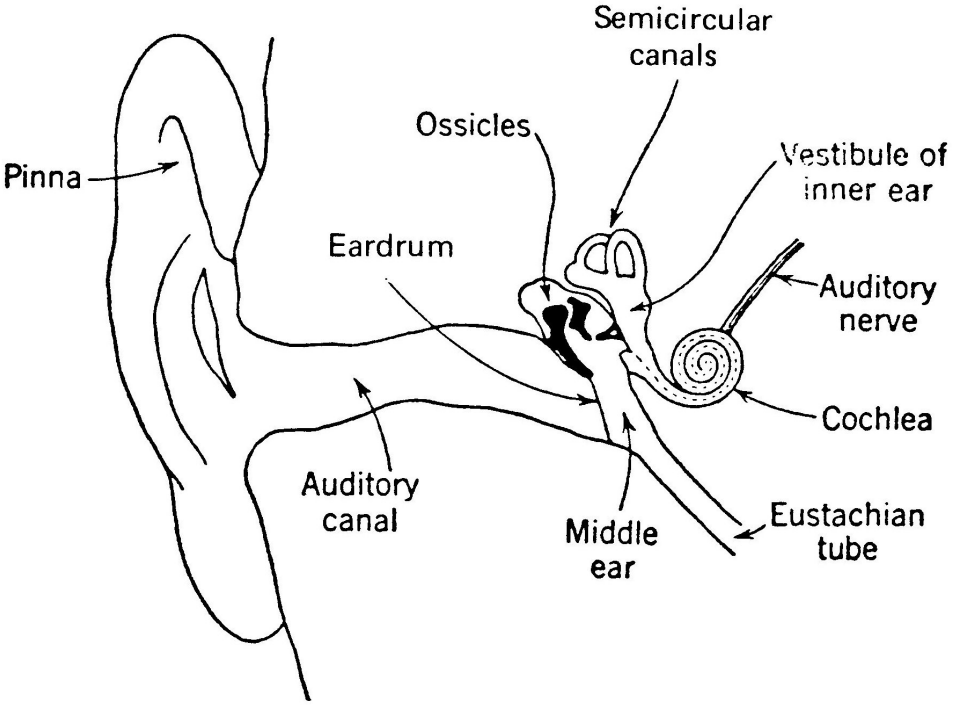
Chapter 3

The Human Auditory System

The human ear (Fig. 3.1) is one of the most sophisticated and delicate mechanical structures in the human body. The ear has an ability to detect sounds with frequencies varying from 20-20kHz, and, remarkably, even 1 kHz sounds that displace the eardrum only a tenth of the diameter of a hydrogen molecule (Kinsler, 2000). To fully appreciate the capability of the hearing system, one can imagine on the surface of a pond, two hollow tubes the diameter of the auditory canal with water bobbing up and down in the them. Now imagine disturbances at different locations in the pond that create complex wave patterns that propagate across the surface towards the tubes. The task of determining the location of where the rocks were thrown and identifying characteristics of the 'rock throwing' would be difficult, to say the least, if all that could be observed is the water level inside the hollow tubes. This is precisely what the human auditory system does. To get a better idea of the function of the various parts of the auditory system this section will explain the process starting from the ear which consists of three parts: the outer, the middle, and the inner ear.

Outer Ear. The pinna of the outer ear collects sound into the auditory canal. For humans, the gain from the pinna is small compared to that of other animals. The auditory canal is approximately a cylindrical tube of about 2.5 cm in length with diameter of about 0.7 cm with the eardrum located at the far end. The resonances of the ear canal effectively transmit sounds at the eardrum for frequencies between 2 to 6 kHz. Due to the dimensions of the canal and the limits of frequency perception, more complicated angular modes can be ignored and, therefore, approximately only plane waves

Figure 3.1: Human Ear (Kinsler, 2000).



propagate down the auditory canal.

Middle Ear. The tympanic membrane (eardrum) is considered part of the middle ear. The acoustic function of the middle ear is to transmit sounds coming from the air to the fluid filled cochlea of the inner ear while maintaining as much intensity as possible. The eardrum also serves as a barrier between the external environment and the moist warm middle ear cavity. Attached to the eardrum are a series of the 3 bones called the ossicles (malleus, incus, and stapes) that transmit vibrations to an opening in the cochlea called the oval window. Since the area of the eardrum is larger then that of the oval window of the cochlea, the total pressure acting on the oval window is increased. The ossicles serve as an impedance match for air vibrations to the fluid vibrations of the inner ear. Acoustic impedance, Z, is defined as the ratio between the acoustic pressure, p, and a particle velocity, v.

$$Z \equiv \frac{p}{v} \tag{3.1}$$

Intensity, I, is the time-averaged product of the pressure with the velocity. It is the

instantaneous rate at which work is done per unit area.

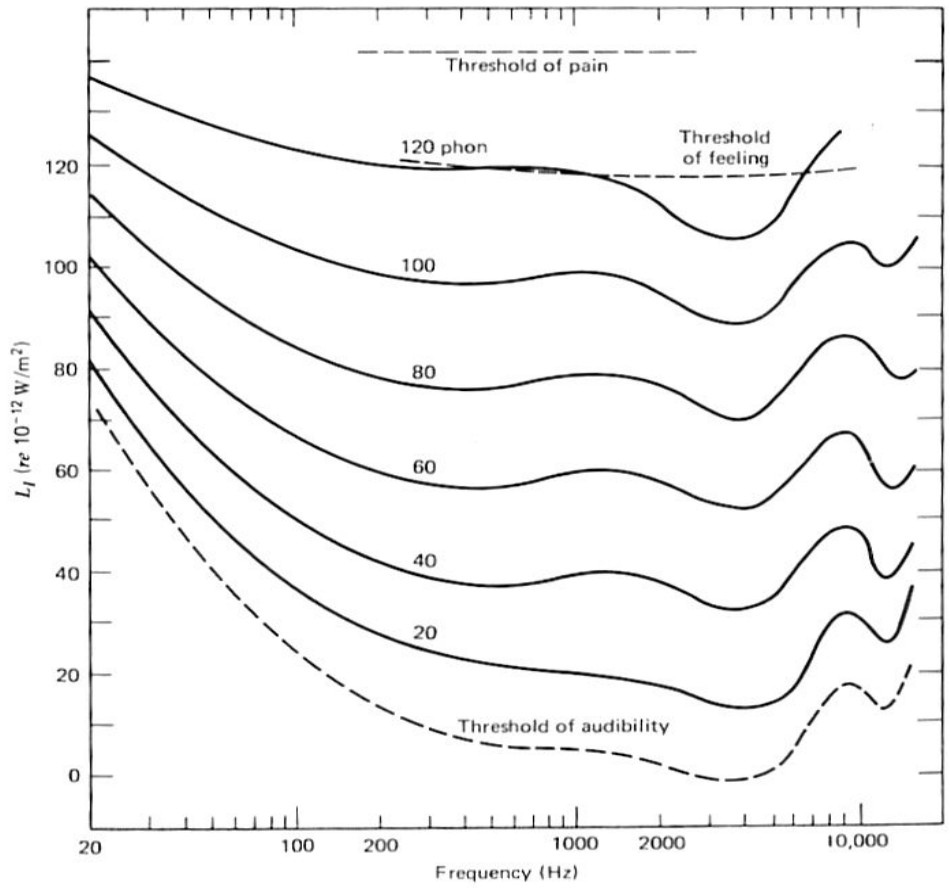
$$I = \langle pv \rangle_t \tag{3.2}$$

Therefore, intensity, as it relates to impedance is then expressed as:

$$I \sim \frac{p^2}{Z} \tag{3.3}$$

For there to be no loss in the intensity of the sound wave as it is transmitted through the middle ear, the ratio of the pressures exerted on the tympanic membrane compared to the oval window is then:

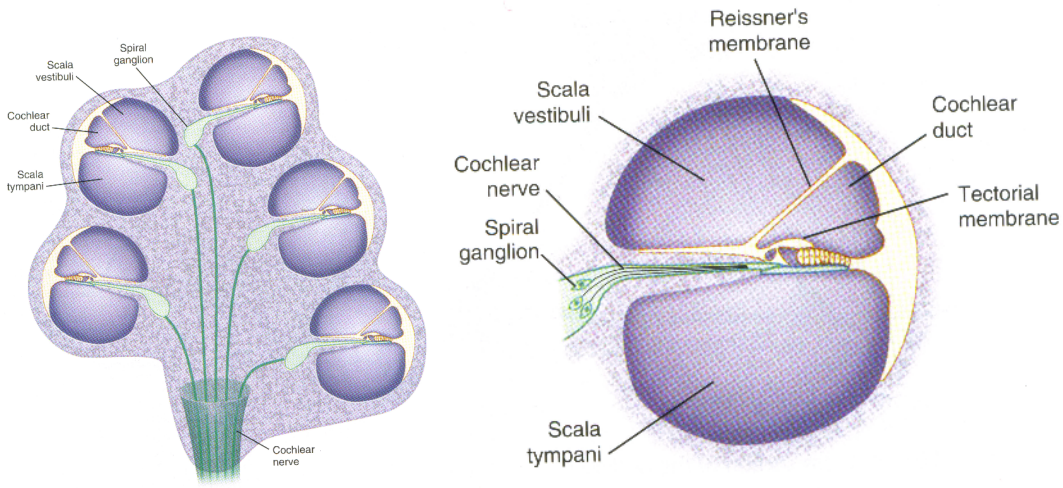
Figure 3.2: Equal loudness contours (Kinsler, 2000).



$$\frac{p_o}{p_t} \sim \sqrt{\frac{Z_{\text{fluid}}}{Z_{\text{air}}}} \sim 35\text{dB} \tag{3.4}$$

The impedance matching is not complete, but if the 3 bones where absent most of the acoustic energy would be reflected away from the inner ear because of its higher acoustic impedance, reducing the intensity transmitted to about 0.1% of the incident intensity. It is important to mention that, due to the resonances of the ear canal and the transfer function of the ossicles, not all frequencies are perceived with the same loudness. The perceived loudness follow experimentally determined equal-loudness contours (Fig. 3.2). It remains important to take this into account when comparing responses between two Fourier components such that one can know which attribute, changes in loudness or pitch, contribute to the differences in the detected response.

Figure 3.3: Cochlea (Nadeau, 2004:467).

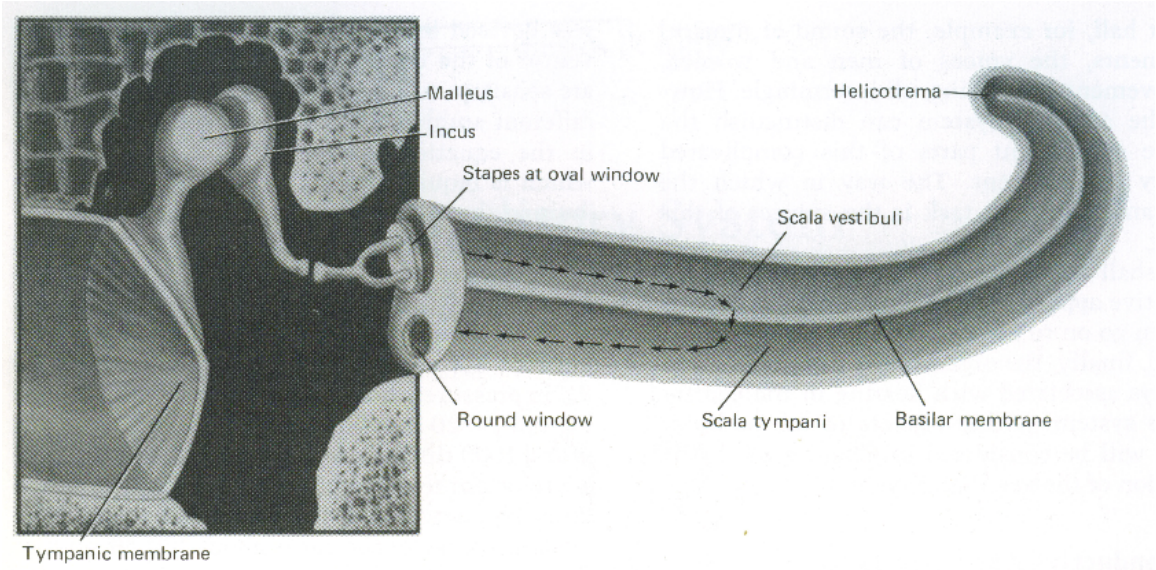


(a) cross section (b) single cross section

Inner Ear and Pathways to the Brain. The inner ear (Fig. 3.3) consists of two main parts: the vestibular apparatus and the cochlea. Only the cochlea is involved in the hearing process. The cochlea is the snail like organ that spirals for about two and a half turns and has 3 fluid filled compartments: scala tympani, scala vestibuli and scala media (cochlear duct). The vibrations of the stapes at the oval window propagate as traveling pressure waves along the scala vestibuli to the helicotrema. These pressure waves instigate movement of the basilar membrane, the scala media, and the organ of

Corti.

Figure 3.4: Unwound cochlea (Kandel, 2000).



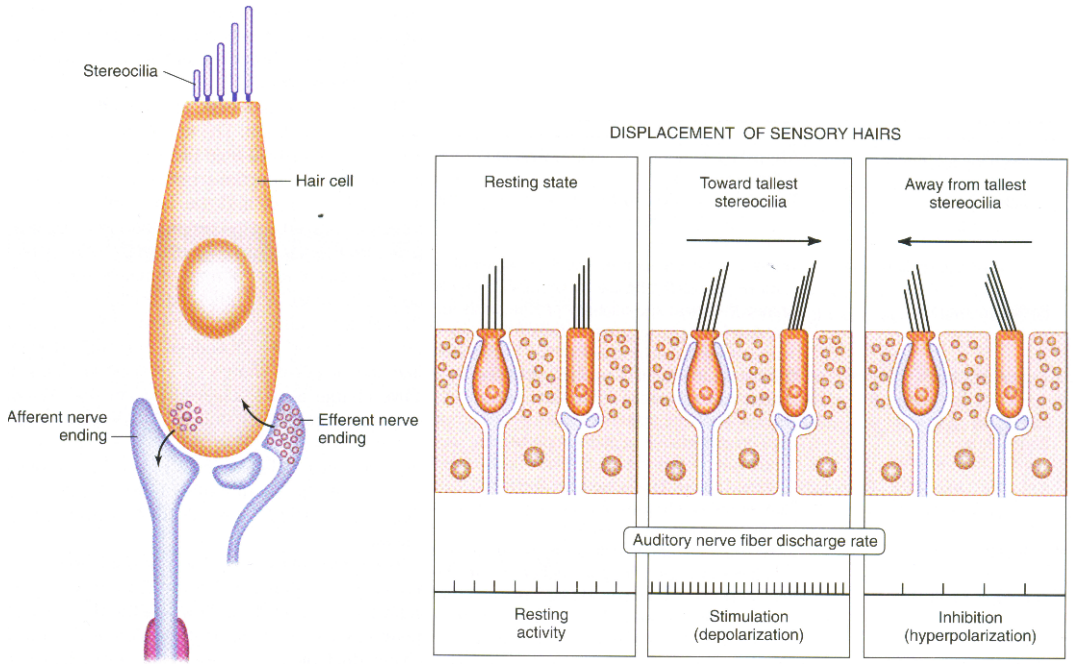
This concept is more accessible when the cochlea is 'unwound' (Fig. 3.4). The basilar membrane separates the scala tympani from the scala media and on it are the organ of Corti. The fluids in these two compartments vary in their chemical and electrical properties and are necessary for the mechano-electric transduction to the nervous system where leakages can cause significant hearing loss The organ of Corti consist of two rows of hair cells: inner hair cells and outer hair cells. At the other end of the cell are stereocilia which are connected to a relatively unstretchable membrane called the tectorial membrane. It is the relative movement between the basilar membrane and the tectorial membrane that affects the firing rates[1] of the hair cells by opening channels linked to the stereocilia. The hair cells (Fig. 3.5) maintain a tonic (low frequency) firing rate when they are not stimulated but when they are deflected, depending on the direction, the firing rate is either increased or decreased. These hair cells project their connections to the auditory nerve. Not only are there afferent connections to the brain but the hair cells can also receive efferent connections from the brain.

It should be noted that the transduction just described is a function primarily of

[1] Details of how a neuron 'fires' will be addressed in a later section.

the inner hair cells where the function of the outer hair cells is not fully understood.[2] If the transduction were to be purely mechanical, the frequency range would be much less then it is and the dynamic range would be much less then what is perceived. The outer hair cells then appear to be important in adjusting the gain of the inner hair cells through their mechanical effects on the tectorial membrane dictated by efferent signals from the brain.

Figure 3.5: Hair Cell (Nadeau, 2004:475)



(a) Hair cell.

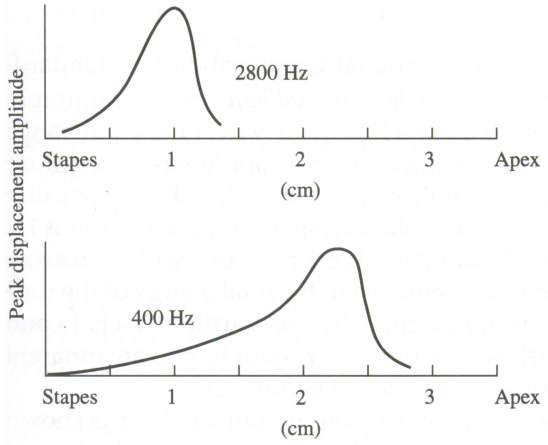
(b) How deflection of hair cell affects the firing rate.

The basilar membrane is not homogenous in that it is narrow (100 microns) and taut near the stapes, and wide (500 microns) and floppy near the helicotrema. Békésy showed, in 1960, that by direct observation of the dynamics of the basilar membrane in response to a pure tone that a traveling (Fig. 3.6) wave moving along the membrane peaks at a location that depends on the frequency of the tone. This *tonotopicity*, or a

[2] There are about 3,500 inner hair cells to about 12,000 outer hair cells. (Hearing. Its Physiology and Pathophysiology, A.R. Moller, San Diego, Academic Press, 2000.)

spatial dependence on frequency, reflects that the acoustic function of the cochlea is to perform a 'Fourier transform' on the sound stimulus; representing the sound wave by its various Fourier components. It should also be mentioned that this tonotopicity is maintained to some degree throughout the brain pathways.

Figure 3.6: Traveling waves on basilar membrane (Kinsler, 2000).



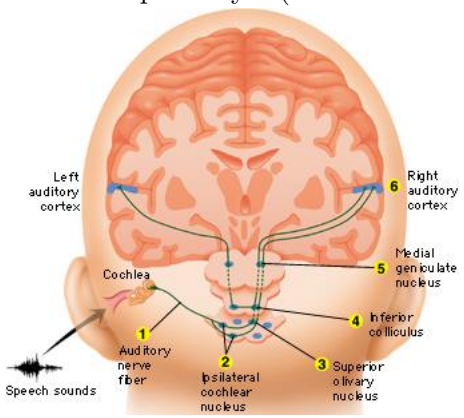
Auditory Pathways. The ascending pathways (Fig. 3.7) to the brain cross over extensively and separate information relating to latency and intensity of the signals giving the ability for stereo sound that helps to locate sources. The signals ascend in parallel to the brainstem, to the thalamus, and then to the auditory cortex where the temporal and spectral information along with the intensity are mapped.

The auditory nerve passes through the internal auditory canal and terminates in the cochlear nuclei. Then from these structures projections are made to nuclei called the superior olive, in the ventral pontine tegmentum, which are very sensitive to temporal and intensity differences between the two ears. These structures can detect the arrival difference of events as quick as 10 microseconds apart which corresponds to changes in direction of 3 degrees. Sound intensity differences are frequency dependent due in part to the diffraction of sound around the head. Higher frequencies are more attenuated then low frequencies. Therefore, neurons which are sensitive to timing differences respond better to lower frequencies where neurons that are sensitive to intensity differences respond primarily to higher frequency sounds.

From the superior olive complex, projections are also made to the inferior colliculi

which then project to the medial geniculate nuclei of the thalamus. The thalamus functions as a relay station, relaying sensory activity from the brainstem to the cortex. The thalamus also relays corticocortical connections as well as transthalamic connections. The so called 'cocktail party effect', our brain's ability to pick out and follow a conversation in a noisy environment, is a current research topic. Understanding in better detail the ability of the thalamus to relay the appropriate auditory stream is one key to understanding this remarkable ability. From the thalamus, projections are then made to the auditory cortices on both sides of the brain. Connections are predominately but not entirely contralateral. These regions lie on the temporal lobe deep with the Sylvian fissure.

Figure 3.7: Brain pathways. (brainconnection.com)

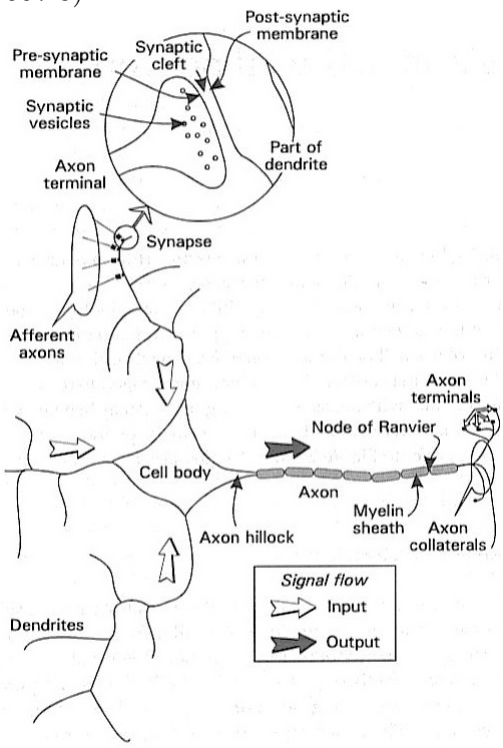


In the auditory cortex separate functional regions deconstruct sound into components to perceive location, loudness, and pitch (Kandel, 2000). The cortical surface is folded leaving much of the cortex buried in grooves. The tonotopicity is maintained throughout the auditory pathway including the cortices. Perpendicular to the tonotopic represention, neurons are organized by latency representing a variation in the temporal response for each set of neurons that respond to a particular frequency band. This is how sequential relationships between tones are represented.

The two main types of neurons found in the cortex are the pyramidal cells and stellate cells. Pyramidal cells are relatively longer and oriented parallel to each other to form layers and are perpendicular to the cortical surface. Since neurons guide the flow of current it is important to understand the basic structure and function of a neuron.

The Neuron. The neuron (Fig. 3.8) is the fundamental unit of the information processing in the central nervous system. A neuron is an electrically excitable cell that usually contains a cell body (soma), a dendritic tree, and an axon. Electrical signals are received from the dendrites and transmitted down the axon in what is called an action potential. The action potential is a fixed size and duration pulse and only occurs when the intracellular potential exceeds a certain threshold. The dendrites receive impulses from other cells through the synaptic connections. These synapses either increase or decrease the intracellular potential.

Figure 3.8: Single neuron. (Gurney, 1997:8)



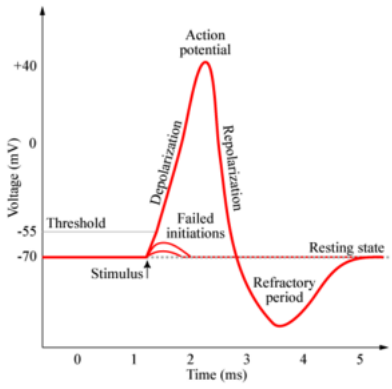
Axons can be more then a meter long or just a few millimeters. Axons are in some ways similar to electric cables. Bundles of axons constitute a nerve. The dimensions of human axons can vary with diameters ranging in sizes of about a micron to about 20 microns. Pulses travel at speeds ranging from 0.6 to 100 m/s depending on the various properties of the axon. Conduction is improved by a fatty membrane that surrounds the axon called myelin. A myelinated fiber has its sheath interrupted at intervals that are unmyelinated which are called *nodes of Ranvier*. Myelin decreases the chances of an axon making a connection with another cell. In the cortex, neurons are highly connected but the connections don't extend over much space. The brain consists mostly of what is referred to as white and gray matter where the white matter consists primarily of myelinated neurons and gray matter (common in the cortex) refers to the unmyelinated neurons.

Axons terminate at the axon terminal which contain chemical substances called neurotransmitter. When stimulated by an action potential these chemicals migrate to

the presynaptic membrane where they are then released across the synaptic cleft, about 20 nanometers wide. The transmitter then combines chemically with the receptor sites at the postsynaptic membrane which then initiates a chemical process that changes the potential across the membrane at that synapse. This change, which is based on the properties at a particular junction or synapse, can either increase or decrease the potential.

Although only the axon is responsible for transmitting signals there are many dendrites that receive signals that are then transmitted to the cell body. Dendrites often appear to have a very branched or tree-like structure. Post-synaptic potentials (PSPs) spread out from a synapse and travel along the associated dendritic structure to the soma at the axon hillock where the axon begins. Again these PSPs can be excitatory by increasing the membrane potential or inhibitory by decreasing the potential. At the same time, many other synapses are being struck at different parts of the dendritic tree influencing the intracellular potential as the PSPs migrate to the axon hillock. Each PSP exists for a certain amount of time before it decays, therefore, signals need to reach at near the same time to have a contributory effect to the potential at axon. If this potential exceeds the threshold, an action potential will be fired down the axon exciting synapses that are connected to it setting off another series of synaptic events in neighboring neurons. This series of events makes possible the logic network that makes up the central nervous system.

Figure 3.9: Action potential. (Wikipedia)



The conduction of the action potential (Fig. 3.9) is due to the changes that take place in the cell membrane. Changes in the membrane change the conductivity of ion currents through the membrane and other changes along the cell. The electric potential across the cell when it is not firing is -70 mV due to the concentrations of the ions across the membrane. There are 3 main ions that contribute to the potential difference across a cell: sodium (Na^+), potassium (K^+), and chloride (Cl^-). Not all of these

ions are equally permeable through the cell's membrane. In fact, potassium is the most permeable of the three. When ions are separated by a permeable membrane two forces are at work to establish a steady state potential. Due to diffusion, ions will tend to permeate from the side of higher concentration. Since the ions are charged, for potassium a valence of +1, any accumulation of total charge (presumably from other ions) will produce an electric field further affecting the migration of the ions. These two processes compete until a steady state is reached where the ions no longer migrate. This can be understood quantitatively by looking at the zero current condition of the Nernst-Planck equation,

$$j = D\left(-\frac{dc}{dx} + \frac{q}{k_B T}Ec\right) = 0 \tag{3.5}$$

here, c, is the concentration of an ion, E, is the electric field, k_B, is the Boltzmann constant, T, is the temperature in degrees Kelvin, and, q, is the charge of the ion. This leads to a relationship for what is called the Nernst or reversal potential.

$$\int_{c_o}^{c_i} \frac{dc}{c} = \frac{q}{k_B T}\int_{x_o}^{x_i} E dx \tag{3.6}$$

$$\ln\left(\frac{c_i}{c_o}\right) = -\Delta V \frac{q}{k_B T} \Rightarrow V \equiv V_i - V_o \tag{3.7}$$

$$\Delta V = -\frac{k_B T}{q}\ln\left(\frac{c_i}{c_o}\right) \tag{3.8}$$

The subscripts i and o correspond to concentrations inside the cell and the outside the cell, respectively, with the membrane separating the two regions. The Nernst potential is the potential where a permeant ion no longer flows through a membrane and is at equilibrium given a difference in concentration across the membrane.

Table 3.1: Ion concentrations

Ion	concentration inside	concentration outside	Nernst Potential
sodium (Na^+)	12 mM	145 mM	**+67 mV**
potassium (K^+)	140 mM	5 mM	-89 mV
chloride (Cl^-)	9 mM	125 mM	-70 mV

The table (Table 3.1) shows the reversal potentials for the 3 typical ions in a nerve cell. Although the potassium and chloride potentials are close to the resting potential

of the cell, which is approximately -70 mV, the sodium potential is much more positive[3] . One might expect the situation to be in equilibrium, but equilibrium... is death. In fact, the cell maintains this non-equilibrium concentration by pumping three sodium ions out for every two chloride ions in. This results in a net transport of one unit charge out of the cell. This accomplished by the ATPase enzyme on the cell membrane.

The differences in the permeability of an ion can be observed by how close the resting potential is to the reversal potential of a particular ion. When the potential across the membrane is different then the reversal potential, a current of those ions will flow through the membrane. The current is assumed to be proportional to the membrane potential with the constant of proportionality being the conductance. The conductance is just the inverse of the resistance. For ions that are relatively less permeable, the membrane then acts as a capacitor. We can reformulate all these ideas by drawing a circuit diagram.

Both the capacitance and the conductance are expressed as per unit area of membrane. The conductances of the three ions are in parallel and can be reduced into a single equivalent conductance. The reversal potentials can be represented as batteries on a circuit diagram driving current into the cell due to the concentration differences. The three reversals can also be reduced to a single battery representing the steady state potential, $\mathbf{V_0}$, across the membrane assuming the ion pumps have been shut down. Remember that the steady state potential tends to be near the Nernst potential of the ion species with the largest relative conductivity. Although the pumps are necessary to maintain a long enduring resting potential, many action potentials can still fire even after the pumps are shut down.[4] Going forward the assumption is that the pumps aren't crucial to the short term aspects of the action potential. Looking at a short cylindrically wrapped membrane of radius, $\mathbf{a}$, with current flowing into one end, out the other, and out through the membrane as dictated by the conductance, $\mathbf{g}$, and capacitance, $\mathbf{C}$, per unit area of the membrane. This conductivity of the interior of the cell is σ.

As with any circuit, any current going into a junction must equal the current leaving.

[3] Shown in bold in the chart to note its distinctness.

[4] The mechanism for shutting down the pumps is accomplished through medication.

For the cylinder this is,

$$I(x) = I(x+dx) + 2\pi a dx \left((V - V_0)g + C\frac{dV}{dt} \right) \tag{3.9}$$

$$-\frac{dI}{dx} = 2\pi a \left((V - V_0)g + C\frac{dV}{dt} \right) \tag{3.10}$$

The axial resistance of a segment of the cylindrical cell due to the conductivity, σ, changes the potential across a segment of cell of length dx.

$$V(x+dx) - V(x) = -I\frac{dx}{\sigma\pi a^2} \tag{3.11}$$

$$\sigma\pi a^2 \frac{d^2V}{dx^2} = -\frac{dI}{dx} \tag{3.12}$$

Combining the results of these two relationships yields the cable equation

$$\sigma\pi a^2 \frac{d^2V}{dx^2} = 2\pi a \left((V - V_0)g + C\frac{dV}{dt} \right) \tag{3.13}$$

At this point, the assumption is that the conductance, g, is a constant. In this case, the equation can be reduced by specifying some useful constants.

$$\lambda \equiv \sqrt{\frac{a\sigma}{2g}} \qquad \tau \equiv \frac{C}{g} \qquad v \equiv V - V_0 \tag{3.14}$$

$$\lambda^2 \frac{d^2v}{dx^2} - \tau\frac{dv}{dt} = v \tag{3.15}$$

This reduction to the linear cable equation realizes the observation that the relation is a diffusion equation. Of particular importance is that it does not admit traveling wave solutions. The cable equation with the assumption of constant conductance does not describe the action potential. All is not lost, however, in that the constants λ and τ do describe useful lengths and times. For a myelinated neuron, the length of the myelin sheath is approximately λ. Since the sheath greatly reduces the conductance of the membrane, the linear cable equation does represent the spread of the potential when the axon is myelinated. It also indicates the spacing between nodes of Ranvier which are just interruptions of the myelin sheath.

To comprehend the conduction of an action potential, first remember the Nernst potential of sodium is much higher then the resting potential. Reasons why this potential is nowhere near the resting potential include the observations that at the resting

potential, sodium is the least permeable ion across the membrane, and the imbalance produced by the ion pumps. The peak of the action potential (Fig. 3.9) is closest to the reversal potential of sodium. One might postulate that at some threshold voltage, the conductance of the sodium through the membrane suddenly increases. This shift now causes the potential to seek a new equilibrium nearer the sodium Nernst potential. This change is short lived thus causing the potential to return to the pre-threshold resting potential.

These ideas can be expressed quantitatively by looking at the model put forth by Hodgkin and Huxley. In 1952, they performed experiments on the giant squid axon that led them to the phenomenological formulation of their model. Instead of having a passive constant conductance, each ion conductance is represented with voltage dependent gating variables whose dynamics are set by appropriate rate equations. The model includes the three ions that have been discussed here. The conductance for the chloride ion remains as a constant 'leakage' term. The total membrane current is then modified from the simple single term expression to the multiple term expression representing each individual ion.

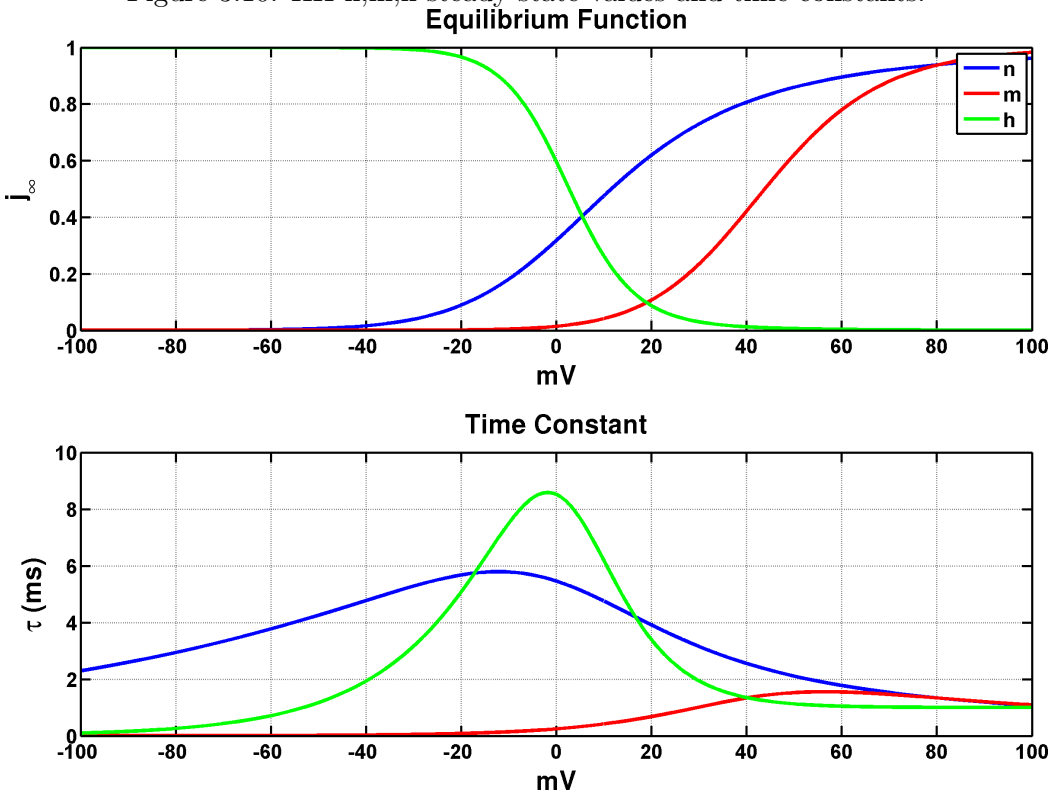
$$(V - V_0)g \Rightarrow g_0^{\mathrm{Na}^+} m^3 h(V - V_{\mathrm{Na}^+}) + g_0^{\mathrm{K}^+} n^4 (V - V_{\mathrm{K}^+}) + g_0^L (V - V_L) \tag{3.16}$$

Here g_0^i represents the maximum conductance of the i^{th} ion. m,h, and n are gating variables that vary from 0 to 1 and whose values are determined by the rate equations where j_∞ is the steady state value of the j^{th} gating variable with a time constant τ_j to reach that steady state of the form

$$\dot{m} = \frac{m_\infty(V) - m}{\tau_m(V)} \qquad \dot{h} = \frac{h_\infty(V) - h}{\tau_h(V)} \qquad \dot{n} = \frac{n_\infty(V) - n}{\tau_n(V)} \tag{3.17}$$

The gating variables (Fig. 3.10) do represent aspects of subunits of the ion channels but for the purposes here the explanation will move to simplify the Hodgkin-Huxley model in order to gain a qualitative understanding of the behavior. This will be done by considering a point neuron reducing the set of equations by eliminating any spatial dependence and making a few modifications. First all powers of the gating variable are reduced to 1, and second, the inactivation of the sodium channel h can be removed since the potassium re-polarization is the same effect. Finally, assume the sodium channel opens 'infinitely' fast ($m = m_\infty(V)$). Turning n into w leads to the Morris-Lecar model,

Figure 3.10: HH n,m,h steady state values and time constants.



(ϕ is a 'temperature' scale factor.)

$$C\frac{\mathrm{d}V}{\mathrm{d}t} = -I_{\mathrm{ion}}(V,w) + I \tag{3.18}$$

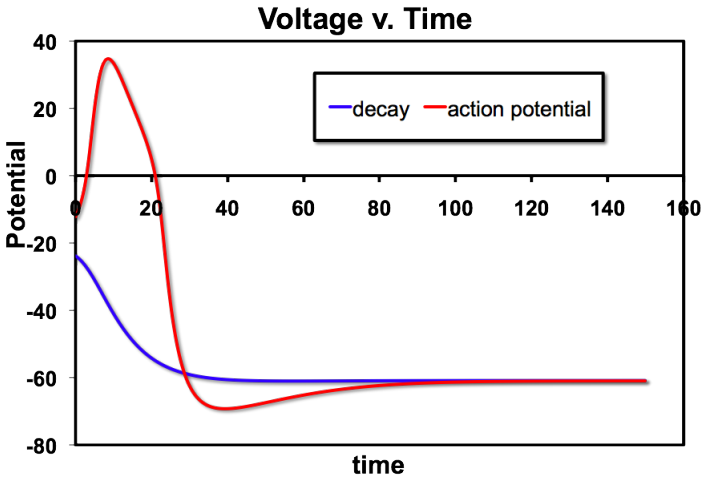
$$\frac{\mathrm{d}w}{\mathrm{d}t} = \phi\frac{w_\infty(V) - w}{\tau_w(V)} \tag{3.19}$$

where

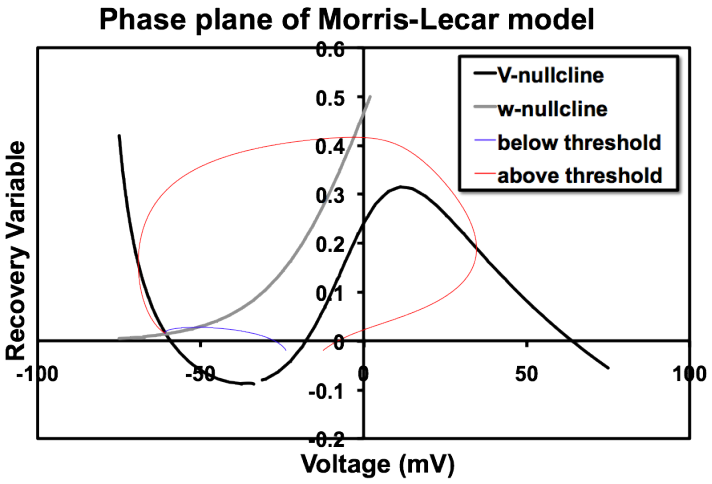
$$I_{\mathrm{ion}}(V,w) = g_0^{\mathrm{Na}^+} m_\infty(V)(V - V_{\mathrm{Na}^+}) + g_0^{\mathrm{K}^+} w(V - V_{\mathrm{K}^+}) + g_0^L(V - V_L) \tag{3.20}$$

To get an understanding of the behavior of these non-linear differential equations, it is helpful to look at the phase plots (Fig. 3.11). To end this chapter, two simulation results will show how this model produces the all or nothing threshold response of a typical action potential and also an oscillation initiated by a stimulus (current) impulse. The point of showing these simulations is to give a few introductory examples of the rich non-linear dynamics of a typical neuron, especially how a cell might obtain a continuing firing rate. There is much more to say about all of these models and the intricacies of their dynamics, but the next chapter, after two chapters of background material, will begin to focus on the chosen method of brain imaging pursued in this work.

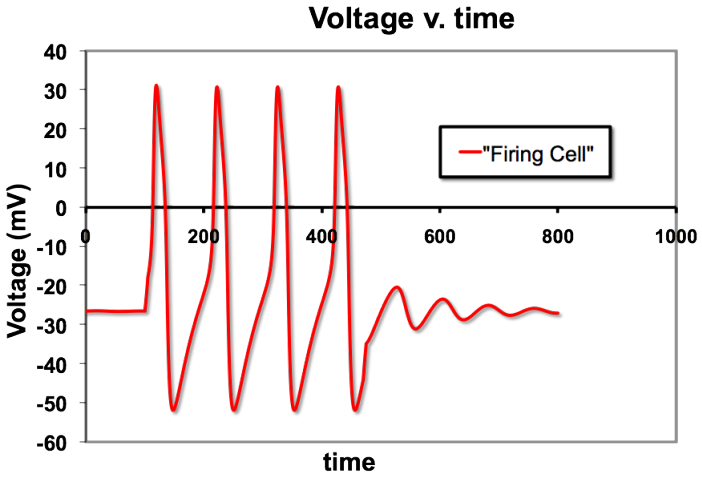
Figure 3.11: Morris-Lecar solutions run with generic inputs to show the threshold behavior of an action potential and oscillatory potentials of an electrically active cell. Showing how the non-linear dynamics of the membrane potential leads to cell firing rates.



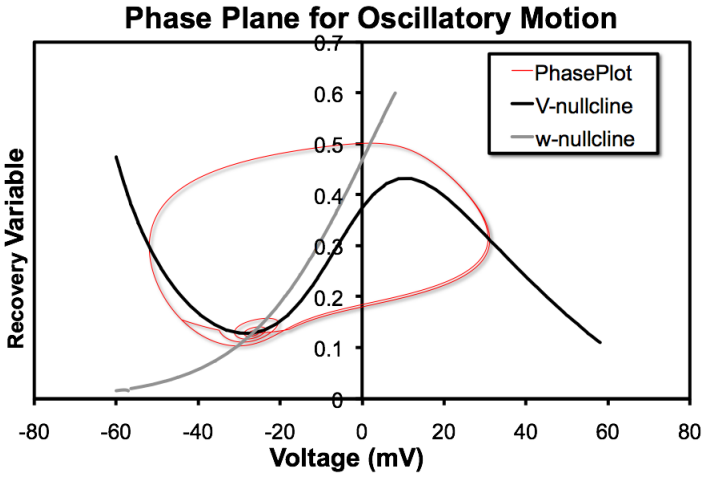
(a) Voltage v. Time



(b) Phase Plots



(c) Oscillatory Voltage v. Time



(d) Oscillatory Phase Plots

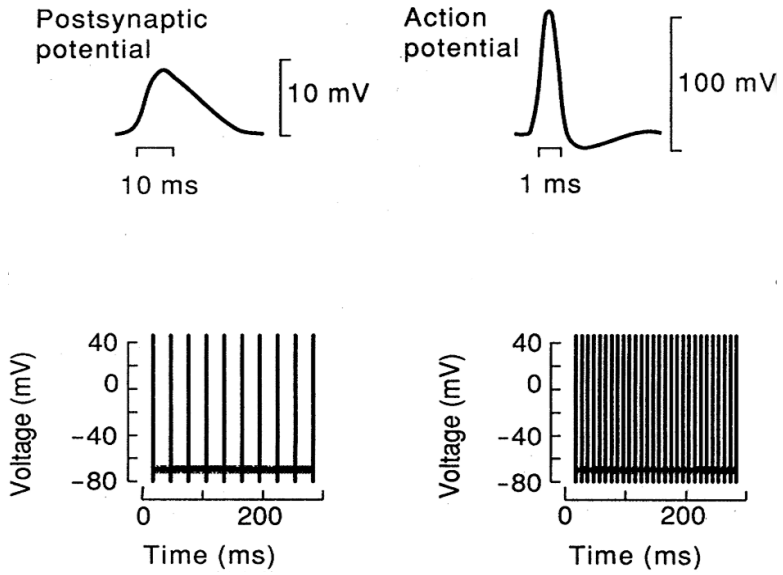
Chapter 4

Magnetoencephalography

Although much progress has been made in the ability to produce images of human brain function, the profusion of findings can be scientifically daunting. The most important reason for the complexity of the findings in brain imaging is the complexity of the brain itself. So far this work has discussed the function of the fundamental element of the human central nervous system, the neuron, but how neurons interact to form circuits that are responsible for higher more complicated function is less understood. There are many non-invasive techniques in functional brain imaging that seek to provide understanding of the electrophysiological, hemodynamic, metabolic, and neurochemical processes that underlie brain function.

Since this work images brain activity in response to rapid ($\ll$ 1 s) acoustic stimuli, a high temporal resolution is a requirement. Of all the techniques, electrophysiological techniques satisfy this need with a resolution of about 1 ms (the shortest time a single neuron takes to complete a firing cycle). This timing precision permits the exploration of basic neural function at the level of cell assembly. Although the spatial resolution of other well known hemodynamic techniques such as fMRI is better, 1-3 mm as opposed to .5 cm, the temporal resolution is at best 1s. The interpretation of fMRI data is complicated by the complex relationship between the blood oxygenation level dependent (BOLD) signal that is detected by fMRI and the underlying neural activity. The BOLD signal does not always have a one-to-one correspondence with regions of electrical activity (Baillet et al., 2001)

Figure 4.1: Neural activity. (Hämäläinen,1993)



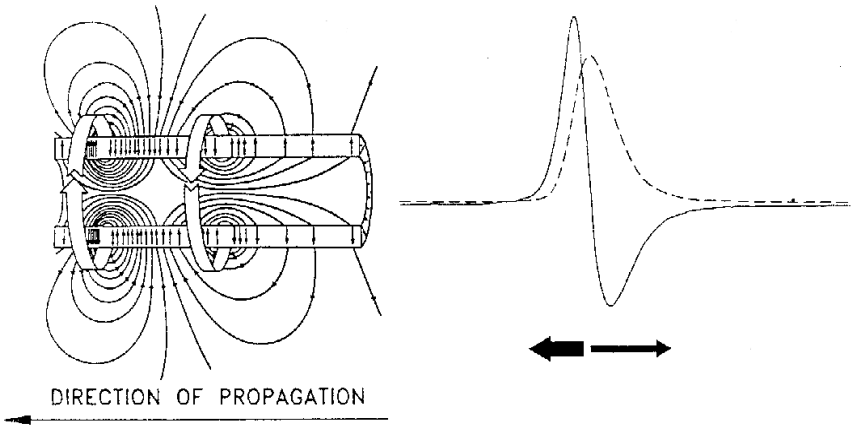
Brain Signal Sources. Magnetoencephalography (MEG) and electroencephalography (EEG) can be used to localize neural electrical activity by non-invasively detecting the electromagnetic fields that are a direct result of the neural excitations (Fig. 4.1). A neuron is excited by other neurons via a volley of action potentials (APs) that stimulate synapses. A post-synaptic potential (PSP) is then generated at its apical dendritic tree. The apical dendritic membrane becomes temporarily depolarized compared to the cell soma and the basal dendrites. This potential difference then causes a current to flow through the volume conductor from the non-excited membrane of the soma to the apical dendritic tree sustaining the PSPs. Some current travels directly through the dendritic trunk, but the rest, as required to maintain charge conservation, returns via extracellular return currents. Only those neurons with dendrites that are centered around a point that is displaced from the cell body will produce a measurable field. The signal consists of a slow depolarization followed by an even slower re-polarization ($\sim$10 ms). From a distance, the activity due to a PSP looks like a short segment, $\vec{Q}$ of current I with length $\vec{L}$.

$$\vec{Q} = I\vec{L} \tag{4.1}$$

Action potentials can be represented by two almost equal but opposing current segments

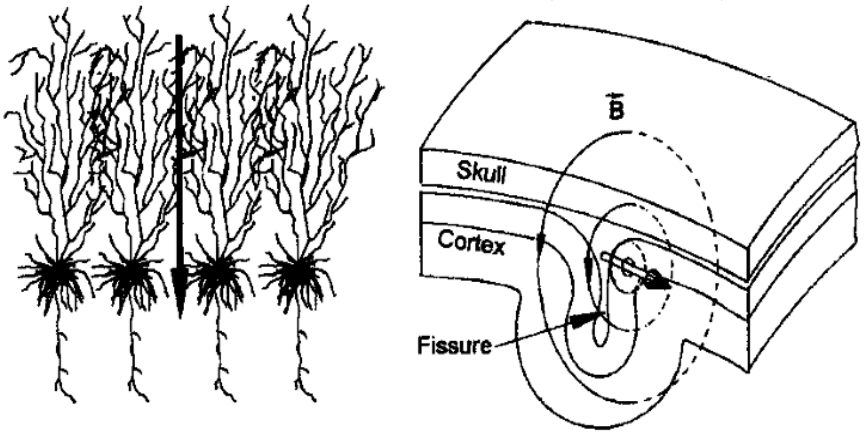
(Fig. 4.2). Due to the asymmetry combined with the differences in the temporal duration of PSPs compared to action potentials ($\sim$ 1ms), most believe that the activity of PSP's is the origin of most the magnetic fields measured outside the brain.

Figure 4.2: Action Potential. (Wikswo,1989)



Cortical Structure. The main generators of the electrophysiological signal are believed to come from the organized layered structure of the cortex (Fig. 4.3). Columns of several thousand synchronously activated large pyramidal cortical neurons are all locally arranged in the same direction perpendicular to the cortical surface. It is the coherent superposition of these currents that produce the detectable electromagnetic fields (Baillet et al., 2001;Hämäläinen,1993).

Figure 4.3: Cortical structure. (Wikswo,1989)



Bioelectromagnetic Fields. With estimates for the conductivity, σ, and noting that the permeability of biological material is very close to that of free space $\mu = \mu_0$, fields can be calculated under the quasi-static limit. Starting with Maxwell's equations:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \tag{4.2}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{4.3}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{4.4}$$

$$\nabla \times \mathbf{B} = \mu_0(\mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}) \tag{4.5}$$

To justify the quasistatic limit, the time derivative terms must be comparatively small. This can be seen by first looking at the total current density in Ampere's Law under ohmic conditions.

$$\mathbf{J} = \sigma \mathbf{E} + \epsilon \frac{\partial \mathbf{E}}{\partial t} \tag{4.6}$$

For the ohmic term to dominate, assuming a harmonic time dependence of the electric field, $\mathbf{E}(t) = \mathbf{E_0} e^{i\omega t}$, then

$$|\sigma \mathbf{E}| \gg |\epsilon \frac{\partial \mathbf{E}}{\partial t}| \tag{4.7}$$

$$\frac{\sigma}{\epsilon \omega} \gg 1 \tag{4.8}$$

With σ= 0.3 $\Omega^{-1}\mathrm{m}^{-1}$, $\epsilon = 10^5 \epsilon_0$, and $\omega = 2\pi(10^2\mathrm{Hz})$ (Hämäläinen,1993), the result is

$$\frac{\sigma}{\epsilon \omega} \approx 500 \gg 1 \tag{4.9}$$

The time derivative of the magnetic field, $\frac{\partial \mathbf{B}}{\partial t}$, must also be negligible. To justify this taking the curl of (equation 4.3) and using (equation 4.5) is,

$$\nabla \times \nabla \times \mathbf{E} = -\frac{\partial}{\partial t}(\nabla \times \mathbf{B}) \tag{4.10}$$

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\mu_0 \frac{\partial}{\partial t}(\sigma \mathbf{E} + \epsilon \frac{\partial \mathbf{E}}{\partial t}) \tag{4.11}$$

assuming that the charge density, ρ, at a macroscopic view is negligible and the previous

assumption of the harmonic time dependence of the electric field, **E**, still valid then,

$$\nabla^2\mathbf{E} = \mu_0(i\omega\sigma - \omega^2\epsilon)\mathbf{E} \tag{4.12}$$

$$\nabla^2\mathbf{E} = \frac{1}{\lambda^2}\mathbf{E} \tag{4.13}$$

$$\lambda = |\mu_0(i\omega\sigma - \omega^2\epsilon)|^{-\frac{1}{2}} \tag{4.14}$$

$$= \left(\mu_0\omega\sigma\sqrt{1+\left(\frac{\omega\epsilon}{\sigma}\right)^2}\right)^{-\frac{1}{2}} \tag{4.15}$$

Here, λ, is a characteristic length. When λ is 'large', this implies that the magnetic induction (time derivative of the magnetic field, $\frac{\partial\mathbf{B}}{\partial t}$) is small. For the values given previously, $\lambda = 65$m, much longer then the diameter of the human head. These arguments justify the quasi-static approximation. Although EEG and MEG should be seen as complementary rather then competing, the main attraction to MEG is its relatively higher sensitivity to the primary current sources located in the sulci. EEG predominantly detects radial sources typical of the gyri and deep in the sulci where as MEG is mostly blind to radial sources. Also due to the relatively low conductivity of the skull, surface potentials are spread out in contrast to the magnetic fields which are unaffected. This is thought to be advantageous for localizing sulcal magnetic sources such as those in the auditory cortex.[1] As the chapter reviews the basics of the biomagnetic forward and inverse problems these ideas will become clear.

Biomagnetic Forward Problem. Calculating the magnetic field from a current source distribution is what constitutes the forward problem in biomagnetism. When calculating the forward problem, cellular level detail is not given explicit attention. In this approximation, the whole cortex is modeled as a homogenous volume conductor with an appropriate conductivity distribution, $\sigma(\mathbf{r})$. The neural activity is modeled as current segments, $\vec{Q}$, at particular locations $\mathbf{r'}_0$.[2] The total current density due to a single current segment is then the sum of the active or primary current, $\vec{J}_0$, due to the neural activity and the second term is the ohmic volume or secondary current, $\vec{J}_v$,

[1] There is plenty of literature that question these assertions. (D. Cohen, et al., 1990; Liu,et al., 2002.)

[2] primed coordinates indicate field sources

which flow throughout the conductor in order to maintain conservation of charge.

$$\vec{J} = \vec{J}_0 + \vec{J}_v \tag{4.16}$$

$$= \vec{Q}\delta(\mathbf{r}_0' - \mathbf{r}') - \sigma(\mathbf{r}')\vec{\nabla}'V \tag{4.17}$$

Not only can a single segment be modeled this way but also many individual segments can be modeled as a larger equivalent segment. In the literature this is referred to as an equivalent current dipole (ECD) which is then an integral sum of individual current segments (the current segments are referred to as current dipoles). The magnetic field of a current segment from the Biot-Savart law is then: (the primed coordinates are the source coordinates)

$$\mathbf{B}(\mathbf{r}) = \int \frac{\vec{J}(\mathbf{r}') \times \vec{R}}{R^3} d\tau' \qquad \Rightarrow \vec{R} \equiv \mathbf{r} - \mathbf{r}' \tag{4.18}$$

$$= \frac{\mu_0}{4\pi}\frac{\vec{Q} \times \vec{R}_0}{R_0^3} - \frac{\mu_0}{4\pi}\int \frac{\sigma(\mathbf{r}')\vec{\nabla}'V \times \vec{R}}{R^3} d\tau' \qquad \Rightarrow \vec{R}_0 \equiv \mathbf{r} - \mathbf{r}_0' \tag{4.19}$$

Assuming that the conductor consists of homogeneous regions, G_i, where the conductivity, $\sigma(\mathbf{r}') = \sigma_i$ is constant the conductivity only changes across the boundaries. The biomagnetic forward solution is

$$\mathbf{B}(\mathbf{r}) = \mathbf{B}_0(\mathbf{r}) - \frac{\mu_0}{4\pi}\sum_{i=1}^{m}\sigma_i \int \frac{\vec{\nabla}'V \times \vec{R}}{R^3} d\tau' \tag{4.20}$$

$$\mathbf{B}_0(\mathbf{r}) \equiv \frac{\mu_0}{4\pi}\frac{\vec{Q} \times \vec{R}_0}{R_0^3} \qquad \text{(primary field)} \tag{4.21}$$

Using a vector identity, $\vec{\nabla}' \times (V\vec{\nabla}'(1/R)) = \vec{\nabla}'V \times \vec{\nabla}'(1/R)$, and Stokes' Theorem, $\int_G \vec{\nabla}' \times \mathbf{u} d\tau = \int_S d\mathbf{S} \times \mathbf{u}$, the volume integral can be converted to a surface integral. ($\mathbf{S}_{ij}$ is a vector normal to the area S from $i \to j$)

$$\mathbf{B}(\mathbf{r}) = \mathbf{B}_0(\mathbf{r}) + \frac{\mu_0}{4\pi}\sum_{ij}(\sigma_i - \sigma_j)\int_{S_{ij}} V(\mathbf{r}')\frac{\vec{R}}{R^3} \times d\mathbf{S}_{ij}' \tag{4.22}$$

The first term is the field due to the primary current and the second term is the contribution from the volume currents which can be calculated by integrating over the surfaces that bound regions of different conductivity, (i.e. brain-skull, skull-scalp, scalp-air).

Head Models. There are two basic approaches to head modeling that are used in the calculation of the forward problem, specifically, the volume current contribution to

the overall external magnetic field. One method is to extract anatomical information from high-resolution Magnetic Resonance (MR) images. The extracted surfaces (the skull, scalp, and brain) are then included in a boundary element method (BEM) calculation of the forward fields. This method can be computationally intense although progress has been made to increase performance (Huang et al., 1999).

The second method is a simpler approximation of modeling the scalp, skull, and brain as concentric spheres. Even though the approximation is rather crude, the spherical model works rather well, as the source of the radial magnetic field is solely due to the primary currents. This and another property of the radial biomagnetic field can be understood by examining the forward solution (eqn. 4.22). If the source is radial,

$$\mathbf{Q} = Q_0\hat{\mathbf{r}} \tag{4.23}$$

$$(\mathbf{Q} \times \mathbf{R}) \cdot \hat{\mathbf{r}} = (Q_0\hat{\mathbf{r}} \times (\mathbf{r} - \mathbf{r}')) \cdot \hat{\mathbf{r}} = 0 \tag{4.24}$$

where $\mathbf{r}$ is the position vector of the field and $\mathbf{r}'$ is the position vector of the source, $\mathbf{Q}$, with orientation, $\hat{\mathbf{r}}$. Since $\mathbf{Q}$ is by definition parallel or anti-parallel to $\mathbf{r}'$, the cross product is perpendicular to both $\mathbf{r}$ and $\mathbf{r}'$ and completely tangential, $\hat{\phi}$, with no radial component. The volume current contribution to the field has vector properties that can be expressed as,

$$\mathbf{S} = S\hat{\mathbf{r}} \tag{4.25}$$

$$(\mathbf{R} \times \mathbf{S}) \cdot \hat{\mathbf{r}} = (\mathbf{R} \times S\hat{\mathbf{r}}) \cdot \hat{\mathbf{r}} = 0 \tag{4.26}$$

where $\mathbf{S}$ is the surface area vector which for a sphere is radial (Sarvas, 1987).

For both models, radial current segments do not contribute to the externally measured magnetic field. This complements EEG which is more sensitive to radial field sources. Although realistic head models are better approximations, the spherical model works remarkably well, especially for magnetic measurements which are less sensitive to the effects of volume currents then are potential measurements (Scheler et al., 2007; Tarkianinen et al., 2003), yet frontal, frontaltemporal, and deep brain neuromagnetic sources are less ideal (Hämäläinen,1989). A study done by Leahy and Mosher (Leahy et al, 1998) compared the accuracy of forward and inverse techniques for MEG and EEG current segment localization using a human skull phantom. Current segments were placed in the phantom at known locations and orientations and measurements were

taken of both the electric potential and magnetic field. Sources were localized using the locally fitted spherical model and the BEM model. With the Recursively Applied and Projected MUltiple SIgnal Classification (RAP-MUSIC) algorithm, localization errors using a spherical model are only slightly greater then those using a BEM model. The average localization error was 7-8 mm for EEG and 3 mm for MEG. Since their is little advantage in using a realistic head model, the RAP-MUSIC algorithm with a spherical model will be used in this work to localize the brain activity. Before going deeper into the topic of localization, the practical issues of detecting bio-magnetic fields will be discussed.

To understand how to detect biomagnetic fields, an approximate calculation of the field from a single neuron gives the order of magnitude of the fields that are being measured. The current segment that models the activity of a single PSP from a cell assumed to be shaped as a cylinder with a diameter, d,

$$\begin{aligned} Q &= \frac{\Delta V}{R} L && (4.27)\\ &= \Delta V \sigma A && (4.28)\\ &= \frac{\Delta V \sigma \pi d^2}{4} && (4.29) \end{aligned}$$

With a potential difference, ΔV, of 25 mV, a cell diameter of approximately 1 μm, and intracellular conductivity, σ, of 1 $\Omega^{-1}\mathrm{m}^{-1}$, (Hämäläinen,1993) the current segment is 20 fA m. Using this in the primary field formula assuming a tangential orientation,

$$\begin{aligned} \mathbf{B} &= \frac{\mu_0}{4\pi}\frac{Q R_0}{R_0^3} \qquad \Rightarrow R_0 \sim 4\mathrm{cm} && (4.30)\\ &= 2 \times 10^{-3}\mathrm{fT} && (4.31)\\ & && (4.32) \end{aligned}$$

where R_0 is the distance from the current source to field sensor. This field is not within the measurable range of current technology, however, since about 10^5 neurons are acting in a cortical area of 1 mm^2, this produces a measurable field of about 10^2 fT (Hämäläinen,1993). The only technology sensitive to these low frequency fields is the SQUID (Super Conducting Quantum Interference Devices).

SQUIDs. As the name suggests, the physical principle of which the technology is based is the quantum interference associated with some aspects of superconductivity.

Materials that are superconducting are those that exhibit no electrical resistance and expel magnetic fields (Meisner Effect). At very low temperatures, electrons interact through phonons and bind to form what are called cooper pairs. For the purposes of this discussion, a super current can be thought of as a many-particle condensate expressed by a single wave function that maintains phase coherence over macroscopic lengths.

As a consequence of the phase coherence, magnetic flux threading a superconducting loop is quantized. The quantization condition is analogous to the angular momentum quantum of the Bohr atom. The flux quantum is: (h is Plancks constant; e is electron charge)

$$\Phi_0 = \frac{h}{2e} = 2.07 \times 10^{-15} Wb \tag{4.33}$$

A super current will then adjust to maintain the flux quantization. When a small non-superconducting material interrupts a superconductor, the tunneling current through the material is dependent on the phase difference, δ, between the currents on each side of the junction. Note that in this discussion, J is the super current (not current density) and, J_0, the critical current. This is the dc Josephson relation.

$$J = J_0 \sin \delta \tag{4.34}$$

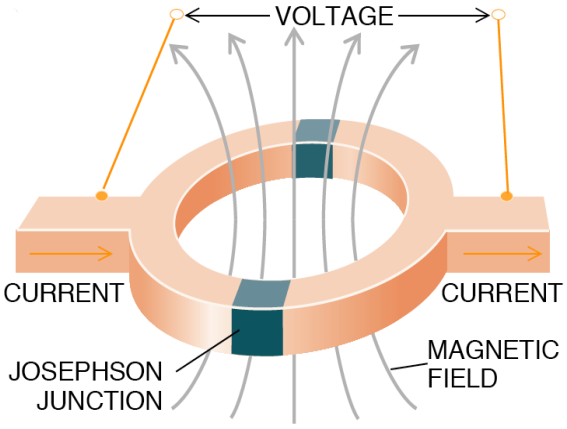
The tunneling even occurs when there is no potential difference across the junction until the super current exceeds the maximum allowed super current i.e. the critical current. Another property of Josephson junctions is the phase oscillation when a voltage, V, is applied to the junction. This is the ac Josephson relation.

$$\frac{d\delta}{dt} = \frac{2eV}{\hbar} \tag{4.35}$$

The combinations of these properties are what define Josephson junctions. There are two kinds of SQUIDS rf (radio frequency) and dc (direct current), however, most of the recent technology has been developed with dcSQUIDS which are more sensitive (for details on rfSQUIDS (Tinkham, 1996)).

A practical dcSQUID (Fig. 4.4) consists of two Josephson junctions interrupting a superconducting loop biased at the critical current. In parallel with each junction are an effective capacitor and a shunt resistor to minimize the hysteretic behavior of the junction. This is the RCSJ (resistively and capacitively shunted junction) model

Figure 4.4: dcSQUID (Clarke, 1994)



(Tinkham, 1996). Only a non-hysteretic junction shows a stable, single-valued I-V relationship. Each junction then shifts the phase according to the equations above depending on the voltage difference across them. In addition, a magnetic flux threading the loop shifts the phase such that the entire loop maintains the flux quantization condition. These three elements combine in a manner that maintains the phase at multiples of 2π. The result is that the critical current of the loop is modulated by the magnetic flux, Φ.

$$J \approx 2J_0 \left| \cos\left(\pi \frac{\Phi}{\Phi_0}\right) \right| \tag{4.36}$$

Most dcSQUID junctions are operated in an over damped condition where the capacitance is minimal to avoid complications from noise and hysteretic behavior at the critical current. The average dc voltage, V, across the entire device is then

$$V = \left(\frac{R}{2}\right) (J^2 - [2J_0 \cos\left(\pi \frac{\Phi}{\Phi_0}\right)]^2)^{1/2} \tag{4.37}$$

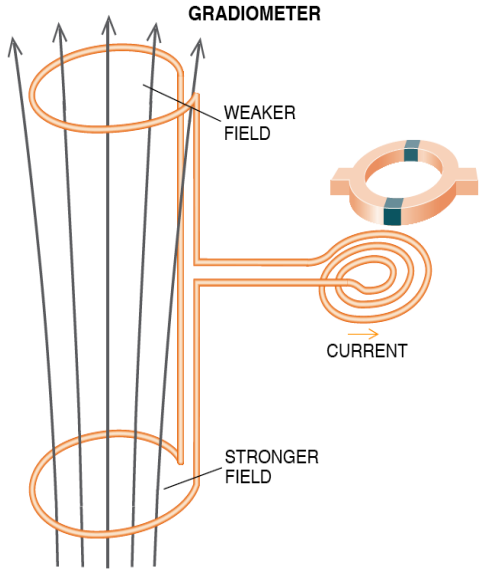
where J is the biasing current and $\frac{R}{2}$ is the resistance of the two resistively shunted junctions in parallel. This is all assuming that the two junctions are identical having the same shunted resistance R and critical current J_0. The relationship shows that the SQUID is a flux-to-voltage transducer which is periodic in Φ_0. To successfully measure biomagnetic fields it is necessary to expand the range of the SQUID by the appropriate electronics. Usually a negative feedback circuit is used to detect fields that are fractions of a flux quantum (Jenks et al., 1997). The negative feedback maintains operation of

the device at the same operating point, and the amount of feedback then becomes the output of the SQUID device. This removes the periodic voltage with applied flux, and allows the device to operate over a wider flux range. Of course, if the flux changes at a fast rate such that the feedback circuit cannot adjust, the result is what is called a flux jump.

The preceding explanation is highly idealistic but does well to explain the basic mechanism for measuring the necessary magnetic fields. The flux threading the loop also contains a term from the induced shielding current in the loop. The actual relationship flux-to-voltage relationship is determined numerically.

SQUIDs do not usually detect the original magnetic field but are connected to flux transformers that form an input circuit. The transformers include the superconducting pickup coils and together detect and transmit the magnetic fields.

Figure 4.5: Axial Gradiometer (Clarke, 1994)



SQUID applications vary in their exact implementation and detail. Most applications today use one of three implementations, the simplest being that of a magnetometer, a single loop of superconducting wire. The advantage of this design is the greater field sensitivity useful for detecting deeper biomagnetic sources. Other implementations involve different geometries, the most common is the first order gradiometer: axial and

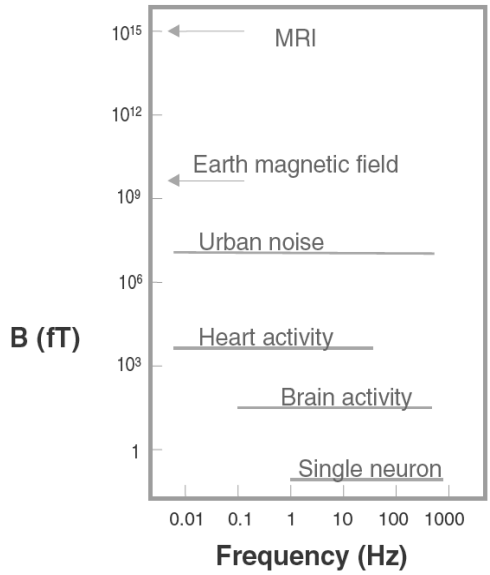
planar. These involve separating two oppositely wound magnetometers a small distance along an axis (axial) away from the specimen or along the plane (planar) of the measurement surface. One point to make is that since the pickup coils are only detecing the radial components of the magnetic field, the field is not maximum directly over the source (axial gradiometer and radially orientated magnetometer). However, for the planar gradiometer the field is maximum directly of the source. This work will use axial gradiometers (Fig. 4.5) to measure local spatial changes in the magnetic field. One can think of the closer magnetometer sensing the biomagnetic signal (signal falls off with distance) and external noise and the farther magnetometer is sensing just the external noise. The difference between the two since they are wound in different directiones gives you an external noise reduced biomagnetic signal.[3] The magnitude of the biomagnetic fields is usually much smaller than that of possible noise sources, such as the earths magnetic field ($\sim 10^{10}$fT) and magnetospheric disturbances ($\sim 10^{8}$fT). Most systems are in magnetically shielded rooms and have magnetometers (reference channels) away from the specimen of interest to measure any possible background fields (Fig. 4.6). These shields consist of layers of material with high magnetic permeability to deflect dc field lines through the shield around the enclosure. Each layer attenuates by about 3 orders of magnitude, where a typical enclosure would consist of a mu-metal layer surrounded by at least two aluminum layers. Higher frequency fields are shielded as they are attenuated through materials of high conductivity and permeability with thicknesses of at least a few skin depths (Clarke, 2004 & 2006).[4] Thermal noise sources are also minimized since the SQUID only functions at temperatures around 4K.

It suffices for the purpose at hand to instead of going into more detail to just give the basic configuration of the equipment that will be used for the experiment. The Brain Sciences Center at the Veterans Affairs hospital has the Magnes™ 3600 WH. This system consists of 248 magnetic channels (not including the reference channels) as well as 96 auxiliary channels for measuring electroencephalography (EEG, detecting surface electric potentials) signals. The sensing elements are integrated dcSQUID axial gradiometers, 1.8 cm in diameter, and are distributed over the measurement surface

[3] Other efforts to reduce noise include using a super-conducting surface to shield the SQUID sensor array (Volegove et al., 2004).

[4] skin depth: $\delta = \sqrt{\frac{2}{\omega\sigma\mu}}$(Jackson, 1999; Griffiths,1999; Hobbie, 1997 are also helpful.)

Figure 4.6: Background fields. (Pizzella et al., 2001)



with an average spacing of less then 2.2-2.3 cm. The field resolution of the sensors is better than 10 fT $Hz^{-1/2}$ at 2 kHz (typical 5 fT $Hz^{-1/2}$) (Pizzella et al., 2001).

Interpreting Bio-magnetic Fields. Interpreting the magnetic field measurements consists of extracting spatial and temporal information from neural activity. Measurements can be taken either in response to a stimulus which is referred to as an event-related field (ERF) or without stimulus i.e. spontaneously.[5] Since this work is interested in understanding auditory processing the experiment will be designed to elicit responses from designed sound stimuli. A raw magnetic measurement contains more noise then brain signal, notwithstanding the efforts to reduce the external background noise (discussed previously). Since the brain is also simultaneously performing many other tasks, the raw measurement is an aggregate of all this processing with only a fraction of it attributed to auditory processing activity.

The basic strategy to extract the desired signal from the background noise of the brain is by taking data for extended periods of time (15-20 minutes) with repeated stimulus presentation (events), and grand averaging the fields that are time-locked with each acoustic sample (1 s). The average time locked auditory activity will be scaled

[5] In the EEG literature, this is referred to as an event-related potential. For an introduction to event-related techniques (see Luck, 2005). For a more advanced treatment see (Nunez P., 2006)

by unity where the random activity will be scaled by $\frac{1}{\sqrt{N}}$ where N is the number of aggregate auditory sound epochs or events. This allows for the time locked signal to emerge from the noise. This averaging is done for each channel and can be used for localizing magnetic sources which is also known as solving the inverse problem. Even without localization the averaged responses are studied for their sub-second temporal information. These averaged ERF waveforms consist of a sequence of positive and negative field deflections which are called peaks, waves, or components. The convention for naming these peaks is to label the negative deflections with a leading 'N' and the positive deflections with a 'P' followed by the latency of the actual peak (i.e. N100 or N1 for a peak 100 ms after stimulus onset). The sequence of these peaks reflects the flow of information where the early components are more dependent on stimulus characteristics and the later components on higher processing, for example, the N200 & P300 (for reviews on methodology see Duncan et al., 2009; Kujala et al., 2007). This work will focus on the mismatch response[6] elicited during an auditory oddball detection paradigm and is believed to reflect disparity between the deviating stimulus and a sensory-memory representation of the standard stimulus (Näätänen et al., 2007; Näätänen et al., 2001; Patel et al., 2005; Winkler et al., 1993;Cowan, 1984).

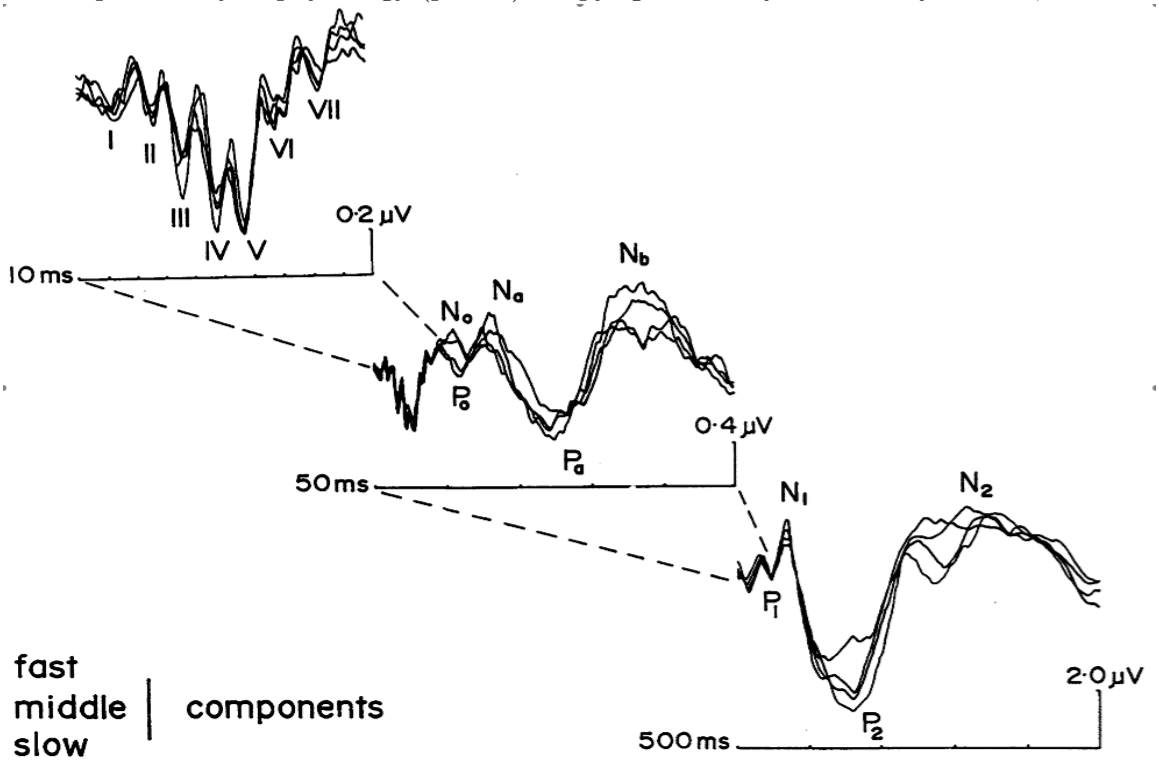
Auditory Stimulus Representation. The grand averaged, time-locked signals contain information about the different stages of auditory sensory processing in the ascending pathways to the auditory cortices. In the review (Näätänen et al., 1999), these are categorized into three consecutive stages that also overlap in time: afferent activation pattern, sensory feature traces, and a unitary sensory stimulus representation (Fig. 4.7). As has been discussed in chapter 3, air vibrations reach the ear and oscillate the ear drum. The vibrations are then transformed into fluid vibrations in the cochlea which then instigate neural activity that ascends up the auditory pathways to the cortices. The stimulus representation would be accessible to memory and correspond with perception.

During the initial stages of auditory information processing, from the cochlea to the brainstem, the afferent activation pattern rapidly changes in time carrying the specific stimulus information toward the cortex. The refractory periods of these neurons are short so that they recover almost immediately after the afferent volley passes through preparing them for the next volley of stimulation. Brain stem auditory evoked potentials

[6] Mismatch Negativity (MMN) in EEG and Mismatch Field (MMF) in MEG

(BAEPs;EEG) are a sequence of deflections (10-12 ms after onset) that emanate from structures ranging from the brain stem (auditory (8th) nerve) to the thalamus as a response to discrete acoustic stimulus. Due to such short refractory periods it is unlikely that any sort of memory processing is happening at this stage. This afferent volley reaches the cortex (Herchel's gyrus) 9-10ms from stimulus onset followed by two more prominent waves at 15 ms and 19 ms. These cortical volleys are followed by other waves at 10-40 ms after sound onset. These deflections are called the auditory middle latency response (MLR) and receive contributions from both auditory-specific and multimodal parts of the thalamocortical pathways. The neuronal circuits that make up the MLR generation do not seem to maintain auditory feature traces as their recoveries are to short after the stimulus.

Figure 4.7: Auditory brain stem (top), middle-latency (middle), and long-latency (bottom) potentials evoked by click sounds. From "The Use of Human Event-Related Potentials in Psychology," by T. W. Picton, 1980, in I. Martin & P. H. Venables (Eds.), Techniques in Psychophysiology (p. 360). Copyright 1980 by John Wiley & Sons, Ltd

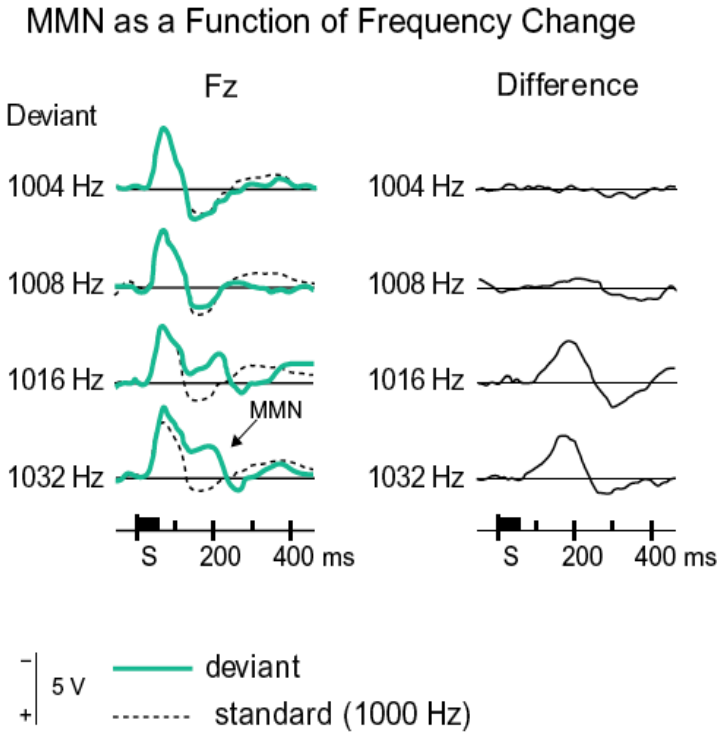


These early stages provide information to feature extraction that results in the emergence of sensory feature traces. The MLR may be partially responsible for the early part of the oscillatory gamma-band response. The gamma-band response is thought to be associated with the detection of stimulus onset and with the binding of the sensory features into a unitary percept. To find correlates of auditory feature traces, the long-latency waves (LLR), peaking at about 50 and 100ms after onset, satisfy the criteria of stability necessary for memory due to there long recovery times, up to about 10 s. The N100 (major deflection 100 ms after onset) wave does not represent a unitary stimulus evoked response but is a response from a set of simultaneously active neuronal generators. The N100 receives contributions from the frontal lobe and the lateral temporal lobe. The N100m, the magnetic counterpart of the N100, is generated from the supratemporal cortex and appears to show tonotopicity (Cansino et al., 2003). The neuronal elements of the suprtemporal N1 seem to be involved in transient detection, providing information about stimulus onsets and do not seem to be processing stimulus specific information. Their refractory behavior appears to encode the stimulus features, yet at this stage, these features are not accessible to perception. It is thought that these response patterns represent an intermediate phase on the way to the synthesis of the auditory stimulus representation that emerges when the feature integration is completed and the temporal dimension has been added. Therefore, the stimulus-specific N1 response patterns form the last stage before the representational system. According to Näätänen (Näätänen et al., 1999), the N1 does not show enough stimulus specificty necessary for an auditory stimulus representation. The mismatch negativity (MMN) (Fig. 4.8) does show more stimulus specificity and reflects the temporal aspects of stimulation.

The Mismatch Field. When a subject is exposed to infrequent (deviant) sounds in a train of otherwise similar (standard) sounds, an event related potential (ERP) component called the Mismatch Negativity (MMN or its magnetic counterpart MMNm, mismatch field MMF[7]) is evoked (Näätänen, Gaillard, & Mantysalo, 1978). The component peaks between 100-250 ms from the onset of the deviant stimulus and appears as a negative deflection with its maximum amplitude (EEG) over frontal and central scalp

[7] Studies have been done investigating the differences between MMF and MMN indicating that the responses are somewhat independent of each other (Kudo et al., 2006).

Figure 4.8: (Left) Frontal (Fz) event-related potentials (ERPs) (averaged across subjects) to randomized 1000 Hz standard (80%, black line) and to deviant (20%, green line) stimuli of different frequencies (as indicated on the left side). (Right) The difference waves obtained by subtracting the standard stimulus ERP from that of the deviant stimulus for the different deviant stimuli. Subjects were reading a book. Adapted from Sams et al. (1985a).



locations. Mismatch responses due to barely discriminable sounds can peak later at 200-300 ms (Garrido et al., 2009) and some studies show a late difference negativity (LDN) thought to be associated with higher processing of more complex sounds (Horvath et al., 2009) usually present in children processing speech. The response is more robust to speech sounds than similar non-speech sounds (Jaramillo et al., 2001) suggesting that speech sounds access long-term memory traces more easily and readily than unfamiliar sounds.

Several other components are sensitive to mismatches in task-relevant attention dependent situations, but MMN is observed even in situations where subjects are not paying attention and even coma (Kane, Butler, & Simpson, 2000). Although the response can be elicited in absence of attention it does predict behavioral accuracy (Amenedo

et al., 2000). The peak latency occurs after the N100 so the MMN amplitude should be quantified after the N100 latency window to avoid interference of the MMN by possible differences in the N100 elicited by standards and deviants, such as differences in onset/offset times.

MMN is considered to be a pre-attentive automatic process that compares incoming stimuli to a sensory memory trace of preceding stimuli. This memory trace only lasts for about 1-2 s.[8] Its biological function is to serve as a change detection mechanism in the auditory scene analysis. The frontal MMN subcomponent is an indicator that the frontal cortical mechanisms correspond to the recruitment of attention (involuntary attention switching) and have been activated in response to a change in auditory stimulation. The MMN latency predicts the behavior response latency (Tiitinen et al.,1994). A small MMN can be generated by even threshold-level differences and can, therefore, be used as an index of discrimination acuity; it provides what may be the most objective measure for this purpose.

Näätänen (Näätänen & Winkler, 1999), after examining many studies, concludes that MMN is also elicited not only when acoustic features are discriminable but also when an abstract rule or pattern is violated. Thus it is extremely important when designing MMN experiments to be mindful not only of lower level physical stimulus features but as well as higher level abstract relationships of complex auditory information. This has also led to improvement in methodology in that multiple deviants can be compared to a single standard as long as the deviants vary along different dimensions (Näätänen et al., 2004). This leads to shorter acquisition times. Music and timbre can be studied with MMN showing that despite the focus of listeners' conscious attention, relatively complex sound phenomena can be encoded by neural mechanisms that are unique to musical sounds (Tervaniemi et al., 1997;Tervaniemi, Winkler, & Näätänen, 1997;Tervaniemi et al., 2001). There are many trends in the study of the mismatch response among them is to use more complex and natural stimuli which is where this work fits in.

[8] One study has shown a 10 s duration (Bottcher-Gandor,C. Ullsperger,P., 1992)

The Inverse Problem. Detecting magnetic fields is only part of the difficulties to understanding the brains response to stimuli in MEG. The inverse problem is localizing the measured magnetic input from a number of detectors. The inverse or localization problem immediately poses difficulties since it is intrinsically non-unique due to current sources that do not produce magnetic fields i.e. radially oriented current segments. The method that will be discussed in this work, RAP-MUSIC, is one of many approaches including least-squares and beamforming approaches (Baillet et al., 2001).

A first approximation to the localization problem can be done by analyzing the radial component of the magnetic field assuming the current segment or current dipole source description. Under these assumptions, only tangential oriented current segments contribute to the externally measured radial magnetic field. For a spherical head model, the only contribution to the radial field is from the cortical current segment itself (eqn. 4.26). For a current segment at a depth of $(1-\alpha)R_0$ where R_0 is the radius of the sphere and α is a constant ranging from 0 to unity. In spherical-polar coordinates the radial magnetic field is,

$$\mathbf{Q} = Q_0\hat{\mathbf{x}} \qquad \mathbf{r}' = \alpha R_0\hat{\mathbf{z}} \qquad \mathbf{r} = R_0\hat{\mathbf{r}} \tag{4.38}$$

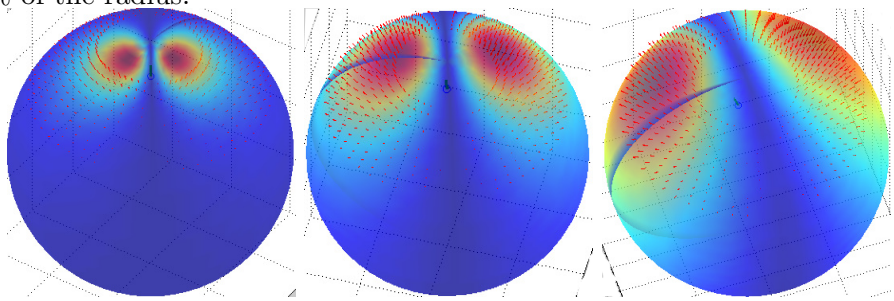
$$\hat{\mathbf{r}} = \sin\theta\cos\phi\hat{\mathbf{x}} + \sin\theta\sin\phi\hat{\mathbf{y}} + \cos\theta\hat{\mathbf{z}} \tag{4.39}$$

$$B_r = \frac{\mu_0}{4\pi}\frac{\mathbf{Q}\times\mathbf{R}}{R^3}\cdot\hat{\mathbf{r}} \tag{4.40}$$

$$= \frac{\mu_0}{4\pi}\frac{Q_0\alpha}{R_0^2}\frac{\sin\theta\sin\phi}{(1-2\alpha\cos\theta+\alpha^2)^{3/2}} \tag{4.41}$$

An estimate of the source depth can be made by estimating the polar angle, θ, of the

Figure 4.9: Radial magnetic field on spherical surface for a source depths of 20%,40%, and 60% of the radius.



radial magnetic field extremum. From figure 4.9, as the depth of the tangential source increases the angle of extremum magnetic field increases. It should also be noted that directly over the source, the radial field is zero. Maximizing the radial magnetic field with respect to θ yields a relationship between α and θ_{m}, a simple first approximation to the source depth using an external measurement.

$$\frac{\partial B_r}{\partial \theta} = \frac{n}{d} = 0 \qquad \Rightarrow \qquad nd' = d'n \tag{4.42}$$

$$\cos\theta_{\mathrm{m}}\alpha^2 + \alpha(\cos^2\theta_{\mathrm{m}} - 3) + \cos\theta_{\mathrm{m}} = 0 \tag{4.43}$$

$$\alpha = \frac{-(\cos^2\theta_{\mathrm{m}} - 3) - \sqrt{\cos^4\theta_{\mathrm{m}} - 10\cos^2\theta_{\mathrm{m}} + 9}}{2\cos\theta_{\mathrm{m}}} \tag{4.44}$$

where the negative radical is the only physical solution. The location of the tangential source below the surface of the sphere is then 1-α. This simple approximation provides a valuable check to more complicated algorithms. Taking the inverse problem further the next section will review the already mentioned RAP-MUSIC algorithm.

Recursively Applied and Projected MUltiple SIgnal Classification. The RAP-MUSIC algorithm is a variation on the MUSIC (Multiple SIgnal Classification) algorithm which was developed in the array signal processing community. These algorithms are best understood in the context of what defines a low rank signal. Before detailing the benefits of this algorithm over other approaches, this section will begin by representing the forward solution in a matrix notation. From eqn. 4.21, using the same definitions for $\vec{Q} = Q\hat{Q}$, $\vec{R_0} \rightarrow \vec{R}$, and $\hat{\mathbf{r}}$, the radial magnetic field at one particular location on the surface of the sphere, can be expressed as, (*i,j,k* sum from 1-3 when repeated and ϵ_{ijk} is the Levi-Civita symbol)

$$\mathbf{B}(\mathbf{r}) \cdot \hat{r} = \frac{\mu_0}{4\pi}\frac{\vec{Q} \times \vec{R}}{R^3} \cdot \hat{r} \tag{4.45}$$

$$B_r = \frac{\mu_0}{4\pi} Q \frac{\hat{Q} \times \vec{R}}{R^3} \cdot \hat{r} \tag{4.46}$$

$$= \frac{\mu_0}{4\pi} Q \epsilon_{ijk} \frac{(\hat{Q})_j (\vec{R})_k (\hat{r})_i}{R^3} \tag{4.47}$$

where the source orientation is independent of time and extending to an m sensor array

with T time samples and n source locations, the noiseless source field is expressed as,

$$[B_r(t)]_{mT} = \frac{\mu_0}{4\pi}\epsilon_{ijk}\frac{(\hat{Q}_n)_j(\vec{R}_{mn})_k(\hat{r}_m)_i}{R^3_{mn}}[Q(t)]_{nT} \tag{4.48}$$

(R_{mn}: vector from n$^{\text{th}}$ current source to m$^{\text{th}}$ magnetic sensor). The forward solution can then be expressed as the multiplication of an $m \times n$ or $m \times 3n$ (unknown source orientations) Lead Field matrix $\mathbf{L}$ with the $n \times T$ or $3n \times T$ source matrix, $\mathbf{s}$. The Lead Field relates a unit source at a particular grid point with its field distribution on the sensor array.

$$\mathbf{b} = \mathbf{Ls} + \mathbf{n} \tag{4.49}$$

The noise is assumed zero mean white Gaussian noise, uncorrelated between sensors, making the noise matrix $\mathbf{n}$, $m \times T$.

Taking the autocorrelation of the forward solution, (both sources and sensor noise are uncorrelated)

$$\langle \mathbf{bb}^T \rangle = \langle (\mathbf{Ls} + \mathbf{n})(\mathbf{Ls} + \mathbf{n})^T \rangle \tag{4.50}$$

$$= \mathbf{L}\langle \mathbf{ss}^T \rangle \mathbf{L}^T + \langle \mathbf{nn}^T \rangle \tag{4.51}$$

$$\mathbf{B} = \mathbf{LSL}^T + \mathbf{N} \tag{4.52}$$

where $\mathbf{B}$,$\mathbf{LSL}^T$, and $\mathbf{N}$ are $m \times m$ matrices. In order for the biomagnetic sensor array signal to be considered low-rank there must be more sensors then sources, $m > n$. Both $\mathbf{S}$ and $\mathbf{N}$ are diagonal matrices with dimensions $n \times n$ and $m \times m$, respectively, and $\mathbf{L}$ is a matrix of rank n. $\mathbf{B}$ can be partitioned into a combination of the signal subspace and noise-only subspace, $\Phi_s \lambda \Phi_s^T + \Phi_n \lambda_n \Phi_n^T$ where $\lambda = \lambda_s + \lambda_n$. $\mathbf{LSL}^T$ can be eigendecomposed as $\Phi_s \lambda_s \Phi_s^T$ where Φ_s contains the n eigenvectors corresponding to the nonzero eigenvalues that span the signal subspace.[9] The $m - n$ remaining eigenvalues of $\mathbf{B}$ are those of the noise only subspace, λ_n.

The MUSIC algorithm chooses source locations that correspond to the signal vector projecting almost entirely into the source subspace or projecting orthogonal to the noise-only subspace. This is done by performing nonlinear searches over the parameter space

[9] The noise covariance matrix can be expressed as $\Phi_n \lambda_n \Phi_n^T$.(The n here is only to indicate 'noise-only' not the number of sources.)

locating multiple peaks in a projection metric. The RAP-MUSIC algorithm improves on this by computing the source locations using a subspace projection approach that searches for global maximums of a projection metric recursively.

In the literature (Mosher et al., 1998) describes a process for computing subspace correlations using the singular value decomposition (SVD). The first source is found by searching a densely sampled source grid of the nonlinear location parameter $\vec{x}$ finding the global maximum of the subspace correlations of the lead-field, $\mathbf{l}(\vec{x})$, with the signal subspace, $\hat{\Phi}_s$, estimated from the data covariance matrix. ((Mosher et al., 1999) for how to extract the source orientation.)

$$\text{subcorr}_{max}(\mathbf{l}(\vec{x}_1), \hat{\Phi}_s) \tag{4.53}$$

The novelty of RAP-MUSIC is that instead of searching the signal subspace for multiple sources, the subspace is modified based on the computation of the first source. The search then continues by looking for the global maximum over the modified subspace. This replaces a search of n local maxima with with a search of n global maxima over their respective modified subspaces. The subspace is modified by projecting away from the first source. The orthogonal projector is defined as

$$\mathbf{\Pi}^{\perp}_{\vec{x}_1} = \mathbf{I} - \frac{\left(\mathbf{l}(\vec{x}_1)\mathbf{l}^{\dagger}(\vec{x}_1)\right)}{(\mathbf{l}^{\dagger}(\vec{x}_1)\mathbf{l}(\vec{x}_1))} \tag{4.54}$$

The projector is then applied to both arguments of the subspace correlation function. The location parameter, $\vec{x}_2$, that maximizes this modified subspace corresponds to the second source.

$$\text{subcorr}_{max}(\mathbf{\Pi}^{\perp}_{\vec{x}_1}\mathbf{l}(\vec{x}_2), \mathbf{\Pi}^{\perp}_{\vec{x}_1}\hat{\Phi}_s) \tag{4.55}$$

The projector is modified to include the second source.

$$\hat{\mathbf{L}}_2 \equiv [\mathbf{l}(\vec{x}_1), \mathbf{l}(\vec{x}_2)] \tag{4.56}$$

$$\mathbf{\Pi}^{\perp}_{\vec{x}_2} = \mathbf{I} - \frac{\hat{\mathbf{L}}_2\hat{\mathbf{L}}_2^{\dagger}}{\hat{\mathbf{L}}_{\mathbf{2}}^{\dagger}\hat{\mathbf{L}}_{\mathbf{2}}} \tag{4.57}$$

Generalizing the process for i out of n sources,

$$\hat{\mathbf{L}}_i \equiv [\mathbf{l}(\vec{x}_1), \mathbf{l}(\vec{x}_2), ..., \mathbf{l}(\vec{x}_i)] \tag{4.58}$$

$$\mathbf{\Pi}^{\perp}_{\vec{x}_i} = \mathbf{I} - \frac{\hat{\mathbf{L}}_i\hat{\mathbf{L}}_i^{\dagger}}{\hat{\mathbf{L}}_{\mathbf{i}}^{\dagger}\hat{\mathbf{L}}_{\mathbf{i}}} \tag{4.59}$$

$$\text{subcorr}_{max}(\mathbf{\Pi}^{\perp}_{\vec{x}_i}\mathbf{l}(\vec{x}_{i+1}), \mathbf{\Pi}^{\perp}_{\vec{x}_{i+1}}\hat{\Phi}_s) \tag{4.60}$$

In practice, the number of sources is estimated from the eigenvalues of the eigendecomposition of the data covariance matrix. This estimate doesn't need to be precise and an overestimate the signal subspace is satisfactory. The recursion process ceases when the subspace correlation drops below a minimum threshold. The correlations of the lead field with the remaining noise-only subspace should be rather low compared to the source correlations. To make these ideas clear, in appendix **??**, there is some sample code implementing this algorithm on a very simple source model.

Chapter 5

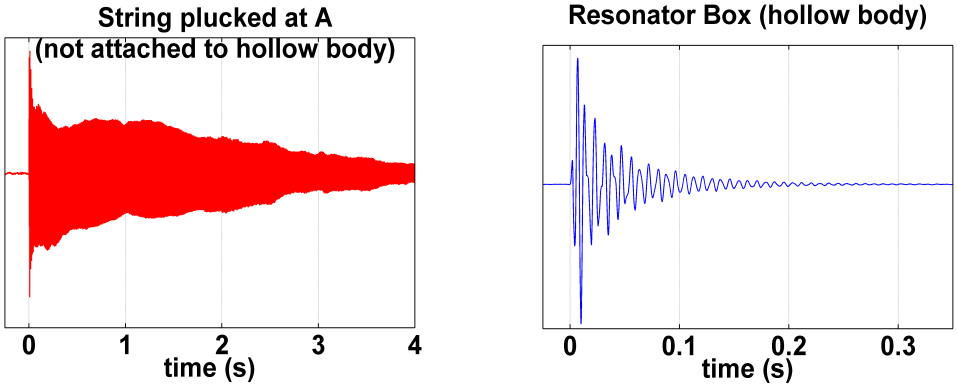
Stimulus Design

In most auditory studies, investigators routinely use synthetic sounds that have little resemblance to sounds experienced in the natural world. In some cases, this may be optimal for targeting a specific response from a neural structure, however, results from such studies may be difficult to extend to every day experience. One example of this is the ability to identify a person's voice. Studying the auditory response of pure tones or combinations of pure tones can only go so far in understanding how a person discriminates voices. A study comparing pure tones to more complex sounds with nearly the same pitch showed different anatomical locations for the more complex sounds suggesting that more processing is done for complex sounds (Ahlo, 1996). This study proposes a model for constructing real-world sounds and then examines the auditory system's ability of detecting subtle manipulations of the audio spectrum. The auditory system processes sound waves by first reconstructing them in the frequency domain (tonotopicity). Since more frequency components are required to reconstruct a heavily damped or transient vibration, this work postulates an exceptional sensitivity to subtle adjustments in the transient portions of complex sounds. It has been shown that processing in the first 300 ms of a sound contribute more to the pre-attentive processing then later processing (Grimm et al., 2006) suggesting that parts beyond 300 ms are less important for pre-attentive auditory object representation.

Audio production model. The auditory construction model is based on the fundamental principles of damping and oscillations that have been introduced in chapter 2 and influenced by the source-filter model outlined in (Handel ,1995) where the basic

idea is that the source is excited and then generates vibrational patterns that are then imposed on a filter which modifies these vibrations. Synthetic sounds resembling that of a guitar are constructed by first taking a set of measurements and then generating digital audio files that represent the measured vibrations. Measurements (Fig. 5.1 & Fig. 5.2) were taken of a plucked acoustic guitar string attached to the guitar box, the guitar box alone, and the string alone (the string acting as the source and the guitar box acting as the filter).

Figure 5.1: Recordings of Seymour Duncan Acoustic guitar made with MagMic Acoustic Pickup System. Notice the envelope differences between the string and the hollow body.



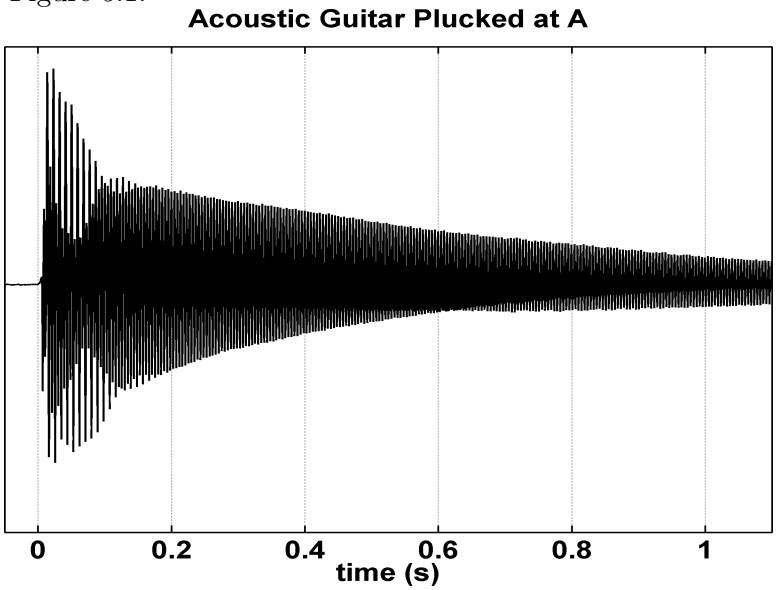
From these measurements, it is clear that the two main sources of the sound production for the guitar have different temporal envelopes. The string is a more sustained sound where the hollow body is a more transient sound. Plotting the vibrations (Fig. 5.3) in frequency space shows this property in the sharpness or quality factor (eqn. 2.38) of the various resonances. Narrower peaks indicate a more sustained resonance, the broader peaks more transient resonances. This relationship between temporal envelope and frequency component width will be exploited in the synthetic construction of the audio samples.

Using the general solution (eqn. 2.12), the conditions of over-damping, $\tau\omega_0 < 1$, (Table 2.1) and zero initial displacement,

$$\lambda_{\pm} = -\frac{1}{\tau} \pm \frac{\sqrt{1-(\tau\omega_0)^2}}{\tau} = -\frac{1 \pm \alpha}{\tau} \tag{5.1}$$

$$\alpha \equiv \sqrt{1-(\tau\omega_0)^2} \qquad (0 < \alpha < 1) \tag{5.2}$$

Figure 5.2: Notice how the envelope of this recording has similarities to both of the recordings in Figure 5.1.



$$x(t) = A_1 e^{-\frac{t}{\tau}}(e^{\frac{\alpha}{\tau}t} - e^{-\frac{\alpha}{\tau}t}) \tag{5.3}$$

$$= 2A_1 e^{-\frac{t}{\tau}} \sinh(\frac{\alpha}{\tau}t) \tag{5.4}$$

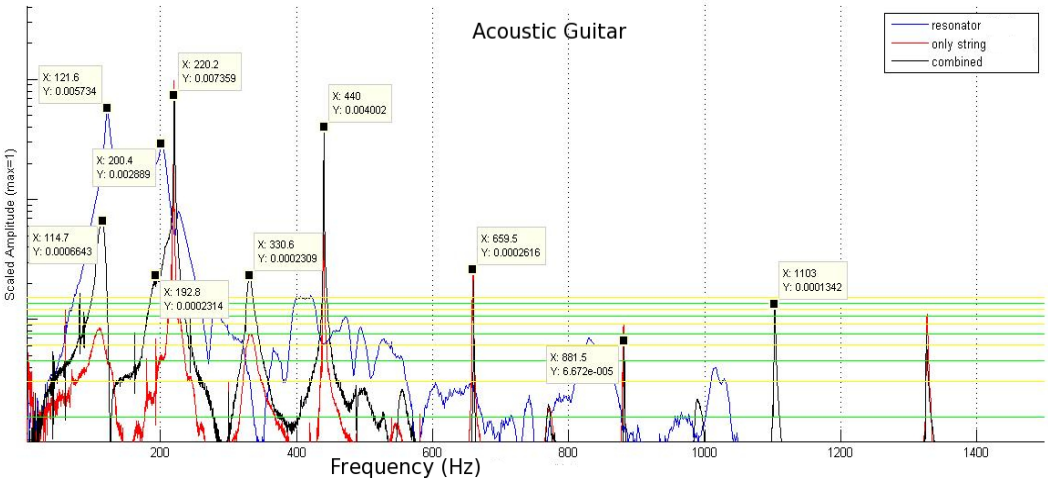
The dimensionless time courses of these functions are plotted in Figure 2.1b for different alphas α. The previous expression in dimensionless coordinates where, v_0, is the non-zero initial velocity,

$$\frac{x(t)}{v_0\tau} = \frac{1}{\alpha}e^{-\chi}\sinh(\alpha\chi) \qquad \frac{t}{\tau} \Rightarrow \chi \tag{5.5}$$

$$= \chi e^{-\chi} \qquad \alpha \to 0 \qquad \text{(critical damping)} \tag{5.6}$$

The idea is to use these solutions as temporal envelopes over various sets of Fourier components. Before moving on, it is important to note some of the characteristics of the different parameters. In the context of sound perception, the initial velocity is of little importance. The other parameters, τ and α, are important for the obvious reasons that they are the damping parameters, but they are also the parameters that determine the rise time. In many psychophysical studies (Caclin et al., 2006), the rise time (attack) is an important perceptual characteristic. These parameters provide a link between the physical characteristics of the sound production and the psychophysical properties that

Figure 5.3: The resonator is plotted in blue, the string (solo) in red, and the combination in black. The amplitude units are arbitrary as all vibrations are scaled to a max of 1. The horizontal green and yellow lines represent the digitization limits for 16 bit audio.



are used to characterize perceptual properties.

$$t_{max} = \tau \left(\frac{\tanh^{-1}(\alpha)}{\alpha} \right) \tag{5.7}$$

$$= \tau \qquad \alpha \to 0 \qquad \text{(critical damping)} \tag{5.8}$$

Referencing the frequency plot (Fig. 5.3) as a guide, two different sets of Fourier components are selected for both the sound produced by the string and the sound produced by the hollow body. The {frequency, amplitude} pairs for the string are,

$$\text{string} \Rightarrow \{\{220\text{Hz}, 1\}, \{440\text{Hz}, 0.544\}, \{660\text{Hz}, 0.0353\}, \{880\text{Hz}, 0.00898\}, ... \\ \{1100\text{Hz}, 0.0177\}, \{1327\text{Hz}, 0.00898\}\}$$

$$\text{hollow body} \Rightarrow \{\{120\text{Hz}, 1\}, \{200\text{Hz}, 0.504\}, \{410\text{Hz}, 0.0401\}, \{1103\text{Hz}, 0.1\}\}$$

Each set of Fourier components is summed on a 44.1 kHz sampled time interval of 0.75 seconds.

$$\sum_i A_i \sin(2\pi f_i t) \tag{5.9}$$

The resulting sum is then enveloped by the critically damped solution to the linear damped harmonic oscillator where C_j is a relative amplitude between enveloped set of

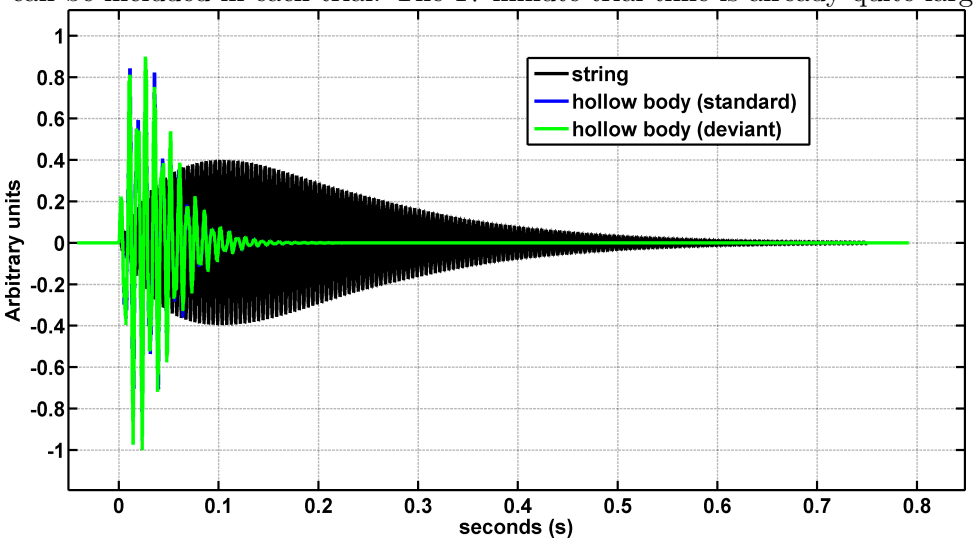
Fourier components and τ_j is the damping/attack time parameter: 20 ms for the hollow body and 100 ms for the string.

$$x(t) = C_j t e^{-\frac{t}{\tau_j}} A_i \sin(2\pi f_i t) \tag{5.10}$$

The next section describes how these parameters are set to construct the standard and deviant sounds samples of the odd-ball paradigm.

odd-ball paradigm. In order to achieve the necessary signal-to-noise ratio for detecting evoked brain responses, multiple stimulations are necessary. The audio samples above are changed slightly[1] and submitted to human subjects' right ear at normal listening volumes in a random train of stimulations lasting about 15-17 minutes. Since biaural presentation is not necessary to produce a mismatch, monoaural presentation is performed for simplicity. Subjects listen passively focusing attention on a dot as to cause minimal eye movement artifacts in the measured magnetic signal.

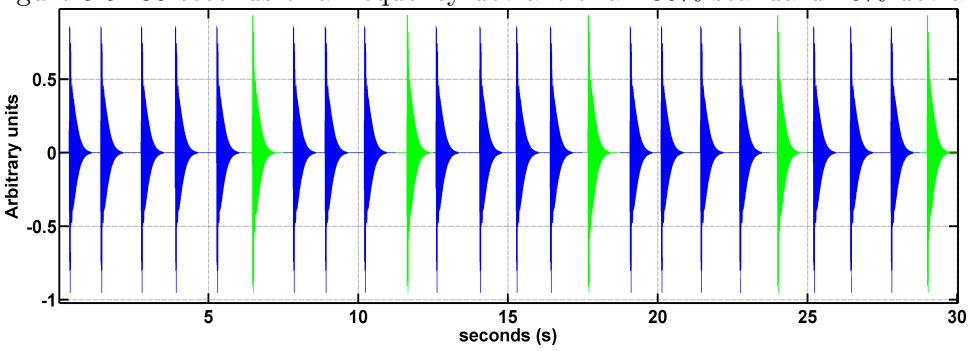
Figure 5.4: Time series representations of sound samples. Note the similarity of the envelope to the measured envelope (Fig. 5.2). Samples are not as long (.75 s) so that more can be included in each trial. The 17 minute trial time is already quite large.



Two different types of audio samples are constructed: a regularly occurring standard and a frequency deviant (Fig. 5.4). The trials (Fig. 5.5) are designed so that at least

[1] The subtle difference is what distinguishes the 'standard' from the 'deviant'.

Figure 5.5: 30 seconds of a frequency deviant trial: 80% standard 20% deviant



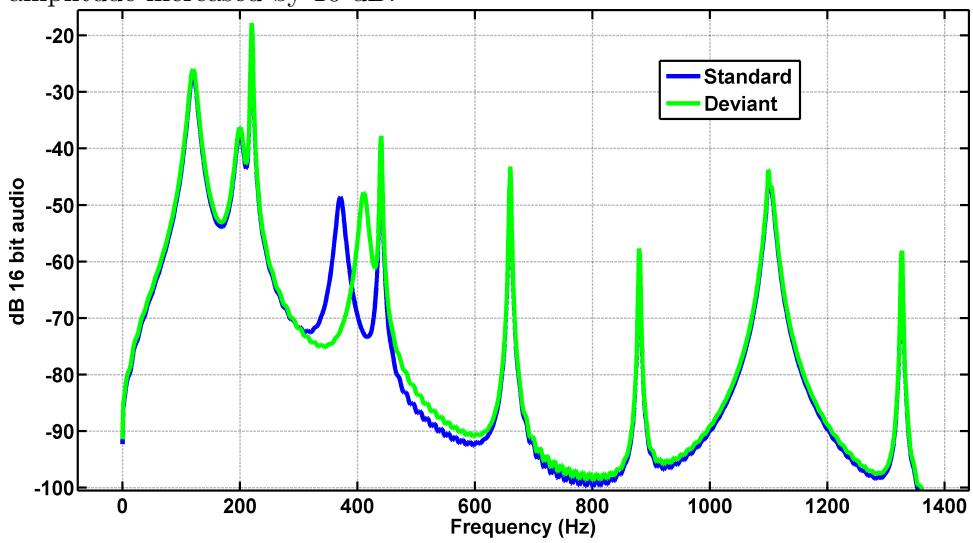
3 standards occur before a deviant, otherwise the presentation order is random. As mentioned above, the individual audio samples last 0.75 seconds but are presented with random inter-stimulus intervals (ISI) of about ∼1 second between samples. Increasing the number of deviants (for a larger signal-to-noise ratio) comes at the expense of the deviant occurence probability. This is an attempt to prevent any sort of anticipation processing from occurring. Sabri and Campbell (2001) showed that the mismatch response amplitude increases slightly as the deviant probabiliy of occurence decreases, but the results are not entirely conclusive for the interstimulus interval used in this study. Three slightly different deviant probabilities within the suggested range of probabilities (Duncan et al, 2009) are used in this study: 13%[2] ,16%[3] ,& 20%[4] resulting in total deviant counts of 107,134,&155. This was done to see the effect of an improved signal to noise while minimizing any effect to the mismatch response.

The only difference between the audio samples is a transposition of the 3rd component of the hollow body (Fig. 5.6). The string portions are identical. The frequency deviant is a shift of the 3rd component along the equal loudness contours (Fig. 3.2) by 40 Hz. The hypothesis is that the heavily damped portion of the sound creates a frequency rich pattern that can be used to identify the sound source. By changing the properties of the component with the least energy, the overall affect on the sample is minimal. The question is if the auditory system is sensitive to these types of low energy differences.

Since most conventional sound production systems have large magnetic signatures,

[2] subjectOne
[3] subjectTwo-Five
[4] subjectSix-Eight

Figure 5.6: Time series and frequency spectrum of sound samples. Frequency deviant has 3rd component shifted 40 Hz higher. Amplitude deviant has the same component with amplitude increased by 10 dB.



the sound in this study must be submitted to the subject through a plastic tube. Measurements of the response curve for the delivery system over the frequency range of interest (365-415 Hz) are required to calibrate for unintended amplitude adjustments of the frequency component shifts.

Chapter 6

Data Acquisition, Analysis, & Results

Data Acquistion. The brain responses of 8 right handed subjects (4 women) were measured using a whole-head 248 channel axial gradiometer system (Magnes 3600WH; 4D-Neuroimaging, San Diego) After the magnetic measurements, the subjects then listened to a similarly constructed single deviant 5-minute audio file and asked to actively attempt to discriminate the sounds in the sequence by counting when a deviant was encountered.

Analysis. Data were sampled at ∼1kHz and filtered, first with a 4th order high pass Butterworth filter at a .1 Hz cutoff[1] , and then with a 5th order low pass Butterworth filter with a 30 Hz cutoff. Epochs were discarded if the measurement exceeded a threshold of 8 times the standard deviation of the absolute value of a trial for a particular channel.[2] The averaging was done on all standards and deviants but also on only the standards immediately preceding the deviants. This was done to get a balanced statistical comparison of the standard with the deviant and compare the best-established standard memory trace. Animations were made of the sensor array projected onto a disc for all time samples. Sensors were selected for magnetic source localization if they contained signal that corresponded to the expected dipolar magnetic field pattern (Fig.

[1] In some cases .5 Hz

[2] SubjectFour contained dental artifacts that necessitated a more stringent approach. Eventually the affected channels were discarded from the analysis.

4.9). This meant paying attention to both the expected spatial dipolar pattern and temporal correlations. This was done visually be examining the sensor array animations. Anywhere from 10-65 sensors were used for the localization per subject.

Localizations of both the standard and deviant responses were done using the RAP-MUSIC algorithm implemented in Brainstorm (http://neuroimage.usc.edu/brainstorm) for a single sphere head model approximated from T1-weighted anatomical MRI data. A novel approach to determine a mismatch was accomplished by observing maximums in the appropriate temporal window of the sensor array animation of the difference between the standard field and deviant field at every channel in time. Then the animation for each standard and deviant was examined for the expected temporal evolution of a stationary[3] dipolar pattern and sources were localized. Mismatches are confirmed by performing a Welch's t-test on each average at every time sample and by significant (greater then .5 cm) differences in localization. Due to possible morphological differences in cell arrangement in the brains of different subjects, the comparison of standard to deviant mismatch was made between signals applied to the same subject and no global fitting was considered.

Determining a mismatch. The process that was used to determine a mismatch consisted of a number of steps. The first step involved analyzing 3 different animations. The first animation consisted of the difference between the standard and deviant for each channel in time (for example, for sujbectOne Fig. 6.3). The other animations were that of the standard (Fig. 6.4) and deviant (Fig. 6.5). The difference animation was used to locate channels that had large differences at the appropriate time ranges for a mismatch ∼150-300 ms. Once a large difference was found, both the standard and deviant animations were checked to see if this difference corresponded to a possible source. This was to make sure the mismatch corresponded to activity that could be localized near the auditory cortex and was done by not only checking for the expected dipolar pattern of Fig. 4.9, but also verifying how the dipolar pattern changed in time. If the positive radial field did not appear or disappear at the same rate as the negative radial field, then these fields did not represent the fields from the expected stationary

[3] Localizing moving dipoles for this particular method involves using shorter windows and therefore fewer samples. The interpretation of of what constitutes a moving dipole could be somewhat vague and open ended i.e. what exactly, given the complex contours of the cortex, should a moving source look like.

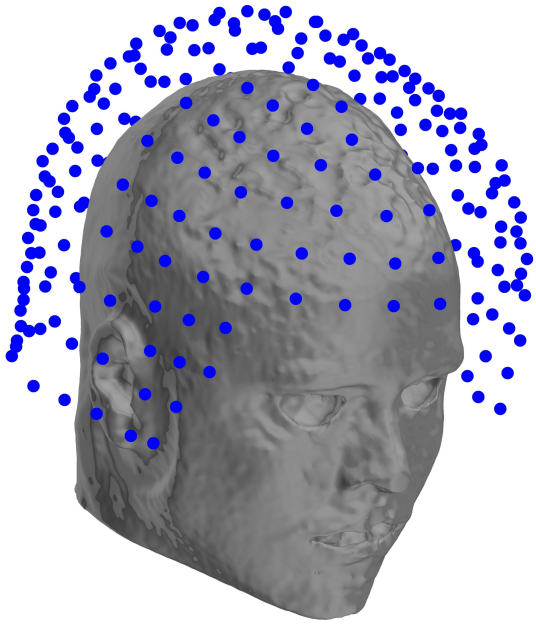


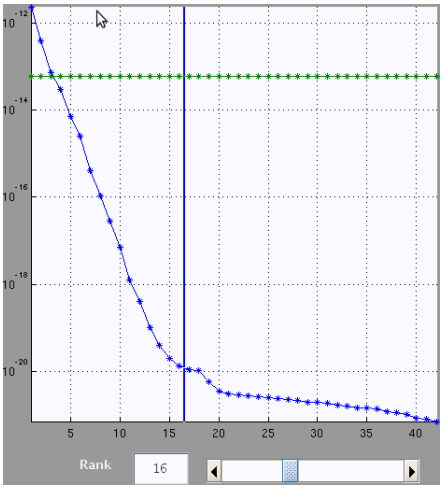
Figure 6.1: The brain responses were measured using a whole-head 248 channel axial gradiometer system (Magnes 3600WH; 4D-Neuroimaging, San Diego)

current segment source. [4]

If the observed difference from the difference animation corresponds to an expected source then only the channels that detected the dipolar pattern are used as inputs in the localization algorithm. In the figures below, the channel selection is shown by the color of the small crosses next to the channel labels. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard.As mentioned in chapter 4, the RAP-MUSIC algorithm works by an eigendecomposition that separates the data covariance matrix into a signal subspace and a noise-only subspace. In practice, this is done by estimating the rank (number of sources) from the eigenvalues of the eigendecomposition of the data covariance matrix, as seen in Figure 6. The estimate doesn't need to be precise but an underestimate could

[4] An interesting functional observation was made in that sometimes the positive and negative radial field patterns would oscillate in size. One getting smaller as the other larger and then reversing, but the position of zero field never changing. A possible explanation for this could be that the postsynaptic activation represented by a current segment (assumed to be straight, eqn. 4.1 and Fig. 4.3) can actually be curved.

Figure 6.2: The eigendecomposition done in brainstorm. The eigenvalues corresponding to the noise subspace should all be the same value, therefore was this occurs the rank should be selected.



yield undesired results. For most of the subjects the estimate was rather large ∼10, considering that only one or two sources are actually being considered. The algorithm ceases to look for sources once the correlation drops below the minumum threshold, which for the results shown here was ∼98%. Only the first or second source met this criteria.

subjectOne. SubjectOne was a right handed, caucasian male in his mid 40s. This subject was the only one to have his head immoblized with a fitted foam helmet to reduce the head movement noise. Figures 6.3 (standard-deviant field difference map) and 6.6 (channel 97) show evidence of a mismatch over the entire range of 150-300 ms. Other channels did as well but this channel had the most distinct mismatch. The difference seen on figure 6.3 does correspond to the expected dipolar patterns of current segment sources as can be seen in figures 6.4 and 6.5 (field maps). The sources were localized (both localizations were done on the same channels) to the left auditory cortex (Fig. 6.7). The standard source more lateral then the deviant source. [5] These results suggest that indeed a mismatch does occur for this subject. Unfortunately, the attentive

[5] Many other localizations were peformed on this subject that yielded locations outside the vicinity of the auditory cortex. These were at different times, i.e. 50ms and 450ms. The clarity of the data as visualized in the animations allowed for this. Helmet fits were no longer used due to their extreme discomfort.

discrimination task was not done for this subject.

subjectTwo. SubjectTwo was a 26 year old, right handed, asian female. Figures 6.8 (standard-deviant difference map) and 6.11 (channel 131) show evidence of a mismatch over the range of 150-250 ms. When analyzing the field map difference animation, it was immediately apparent that there existed unique sources at both 150ms and 300ms. The difference seen at 150 ms (Fig. 6.3) corresponds to the unique deviant source seen in figure 6.5. The unique standard source at 300ms can be seen in figure 6.4. The sources were localized near the left auditory cortex (Fig. 6.12). Although there was significant difference in brain activation indicating a mismatch response, unique sources for each sound sample, the subject could only count one sound out of 28 as being deviant when being exposed to a total of 240 sounds over five minutes in the attended task suggesting the sounds were not attentively discriminated.

subjectThree. SubjectThree was a 29 year old, right handed, caucasian male. Figures 6.13 (standard-deviant difference map) and 6.16 (channel 145) show evidence of a mismatch over the range of 140-200 ms. The standard (Fig. 6.14) and deviant (Fig. 6.15) field maps do show the typical dipolar patterns near the same location on the right side. The sources were localized near the right auditory cortex (Fig. 6.17), however, the standard source was more lateral then the deviant source. This subject was able to count 15 sounds out of a possible 28 as being deviant suggesting a small possiblity that the sounds were attentively discriminated. Although this subject was right handed there was more activation coming from the right side.

subjectFour. SubjectFour was a 22 year old, right handed, caucasion female. Figures 6.18 (standard-deviant difference map) and 6.21 (channel 156) show evidence of a mismatch over the range of 240-310 ms. For this subject, only the field maps for the time slice corresponding to the unique standard source (Fig. 6.19) are included; no deviant source was localized that contained this time slice (Fig. 6.15). Due to the subject's dental apparatus having a magnetic signature some channels (those closest to the mouth) were discarded from the analysis. This made the field maps and the source localizations less robust over longer time periods, therefore the localizations only consisted of a few time samples. Inspite of this difficulty, the locations of the sources are near the left auditory cortex (Fig. 6.22) both deviant and standard had source at the same location for an earlier time range but the standard then has a unique source later,

yielding the mismatch. In the attended discrimination task, only 8 out of a possible 28 deviants were counted in a presentation of 240 sounds suggesting the sounds were unlikely to be attentively discriminated.

subjectFive. SubjectFive was a 22 year old, right handed, caucasian male. Figures 6.23 (standard-deviant difference map) and 6.26 (channel 174) show evidence of a mismatch over the range of 130-175 ms. The field maps for the standard (Fig. 6.24) and the deviant (Fig. 6.24) display there respective sources. The source localizations (Fig. 6.27) for this subject were on the right side of the head and over a short time period 130-140 ms. The field map animations for this subject did not display many sources as others had, yet there is some consistency in that the deviant source is less lateral then the standard. In the attended discrimination task, only 7 out of a possible 28 deviants were counted in a presentation of 240 sounds suggesting the sounds were unlikely to be attentively discriminated.

subjectSix. SubjectSix was a 22 year old, right handed, hispanic female. Figures 6.28 (standard-deviant difference map) and 6.31 (channel 155) show evidence of a mismatch over the range of 230-270 ms. When analyzing the field map difference animation, a unique source at 250 ms is observed corresponding to the deviant (Fig. 6.30). Both standard (Fig. 6.29) and deviant showed a source pattern. For this particular subject, the localizations (Fig. 6.32) do not cover the same time range but seems to suggest that, initially, the standard and deviant are processed in separate locations. Although not contained in a figure here, this can be seen on the field map difference for that appropriate time slice. The larger difference is due to the unique deviant source. In the attended discrimination task, the subject did not identify any of the possible 28 deviants in a presentation of 240 sounds suggesting the sounds were not attentively discriminated.

subjectSeven. SubjectSeven was an 18 year old, right handed, caucasian male, the only subject with musical training. Figures 6.33 (standard-deviant difference map) and 6.36 (channel 208) show evidence of a mismatch over the range of 210-270 ms. However, the standard (Fig. 6.34) and deviant (Fig. 6.35) field maps along with the localizations (Fig. 6.37) show the sources coming from about the same location yet the deviant is slightly less lateral. Although this subject performed the best in the attended discrimination task, the animated field maps and localizations were the most

difficult to locate near the auditory cortex. Many of the attempted localizations, for the appropriate time range, resolved the sources on the lower fissures of the temporal lobe and were thus were discarded. Regardless, the channel plot (Fig. 6.36) suggests a mismatch. In the attended discrimination task, 20 out of a possible 28 deviants were counted in a presentation of 240 sounds.

subjectEight. SubjectEight was a 28 year old, right handed, asian female. Figures 6.38 (standard-deviant difference map) and 6.41 (channel 194) show evidence of a mismatch over the range of 150-250 ms. When analyzing the field map difference animation, a unique source at 160 ms is observed corresponding to the standard (Fig. 6.39). Although a field map of the deviant (Fig. 6.40) is included, the time (117ms) is not appropriate for the mismatch processor. The map is included because no other sources were found over those channels for the appropriate time ranges. The localizations do correspond to activity near the auditory cortex (Fig. 6.42) but in different locations. In the attended discrimination task, only 5 out of a possible 28 deviants were counted in a presentation of 240 sounds suggesting the sounds were not attentively discriminated.

Table 6.1: Summary of Results

subject	**count (Total 28)**	**source localization summary**
6.1 subjectOne	NA	Both standard and deviant produce sources at about the same time, but the standard location is more lateral than the deviant.
6.2 subjectTwo	1	Standard has unique source at 270-340 ms. Deviant has unique source at 135-165 ms
6.3 subjectThree	15	Both standard and deviant produce sources at about the same time, but the standard location is more lateral than the deviant.
6.4 subjectFour	8	Standard has unique source at 264-271 ms.
6.5 subjectFive	7	Both standard and deviant produce sources at about the same time, but the standard location is more lateral than the deviant.
6.6 subjectSix	0	Deviant has unique source at 230-260 ms.
6.7 subjectSeven	20	Difficult to localize the mismatch yet standard location is slightly more lateral than the deviant. (only subject with musical training.)
6.8 subjectEight	5	Standard has unique source at 140-165 ms.

Summary of Results. Most subjects did not seem to be able to attentively discriminate the subtle difference between the standard sound and the deviant sound. However, for nearly all subjects (subjectSeven being the exception; localizations are not distinct) there was strong evidence of a pre-attentive discrimination due to the elicitation of the mismatch response. For a mismatch to occur, a relatively large difference needed to be observed on the field difference map as well as a statistically significant difference on one of the channels either used in the localization or near to the dipolar source pattern. Localizations were performed over time ranges that corresponded to the expected dipolar patterns to see if the locations were also different. For most subjects[6] all these criteria were met.

Another unexpected interesting result was the consistency and distinctness of the functional differences for each gender. For all the males, the deviant locations and the standard locations were at about the same time with the deviant location less lateral. For all the females, unique sources either due to the standard, deviant, or both were observed. This was not observed for the males. The final chapter (Chap. 7) will discuss

[6] Mismatch responses were observed for at least one subject across all deviant probabilities

these results in the context of the current literature.

6.1 subjectOne results

Caucasian male age 45. Head immobilized with foam helmet.

Figure 6.3: subjectOne: **(standard-deviant)** at 0.185 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

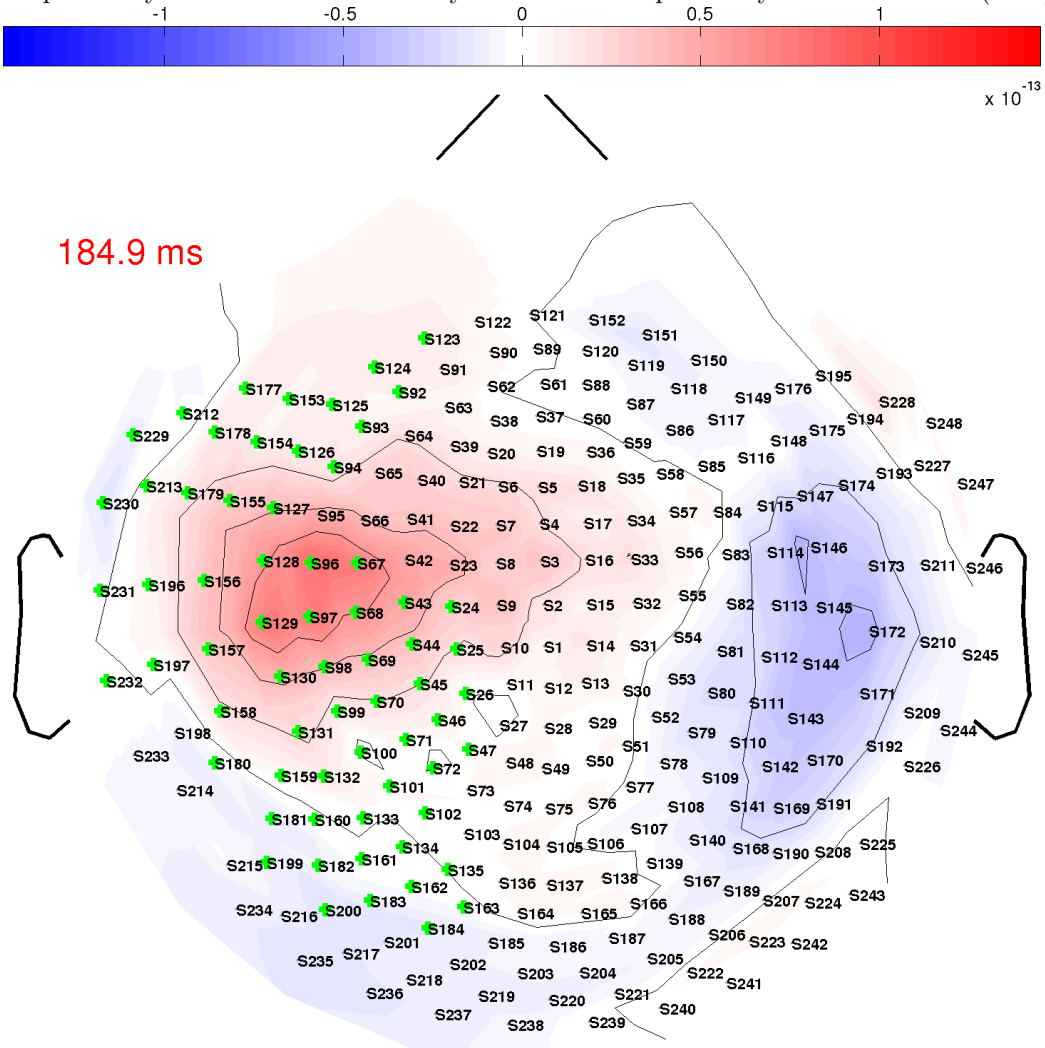


Figure 6.4: subjectOne: **standard** at 0.185 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

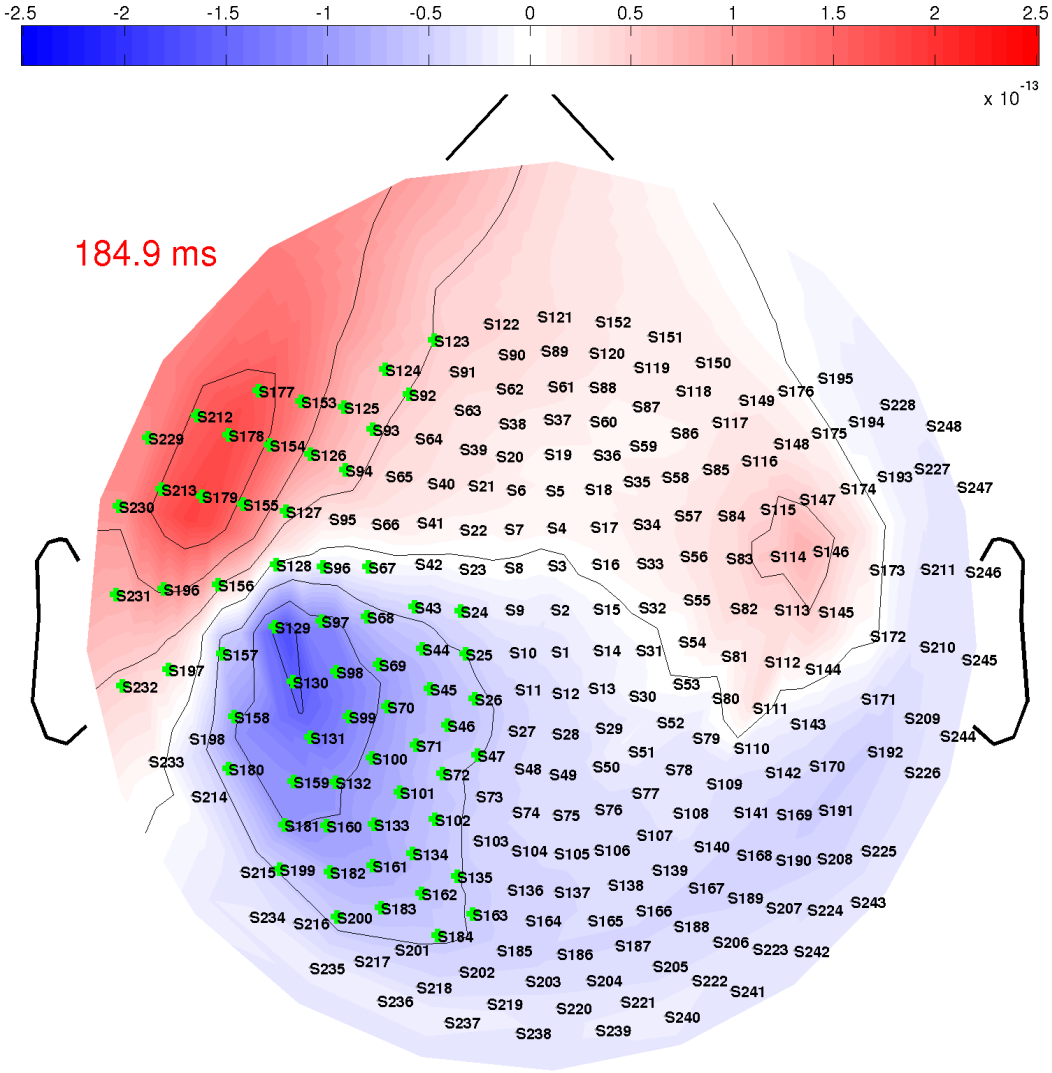


Figure 6.5: subjectOne: **deviant** at 0.185 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

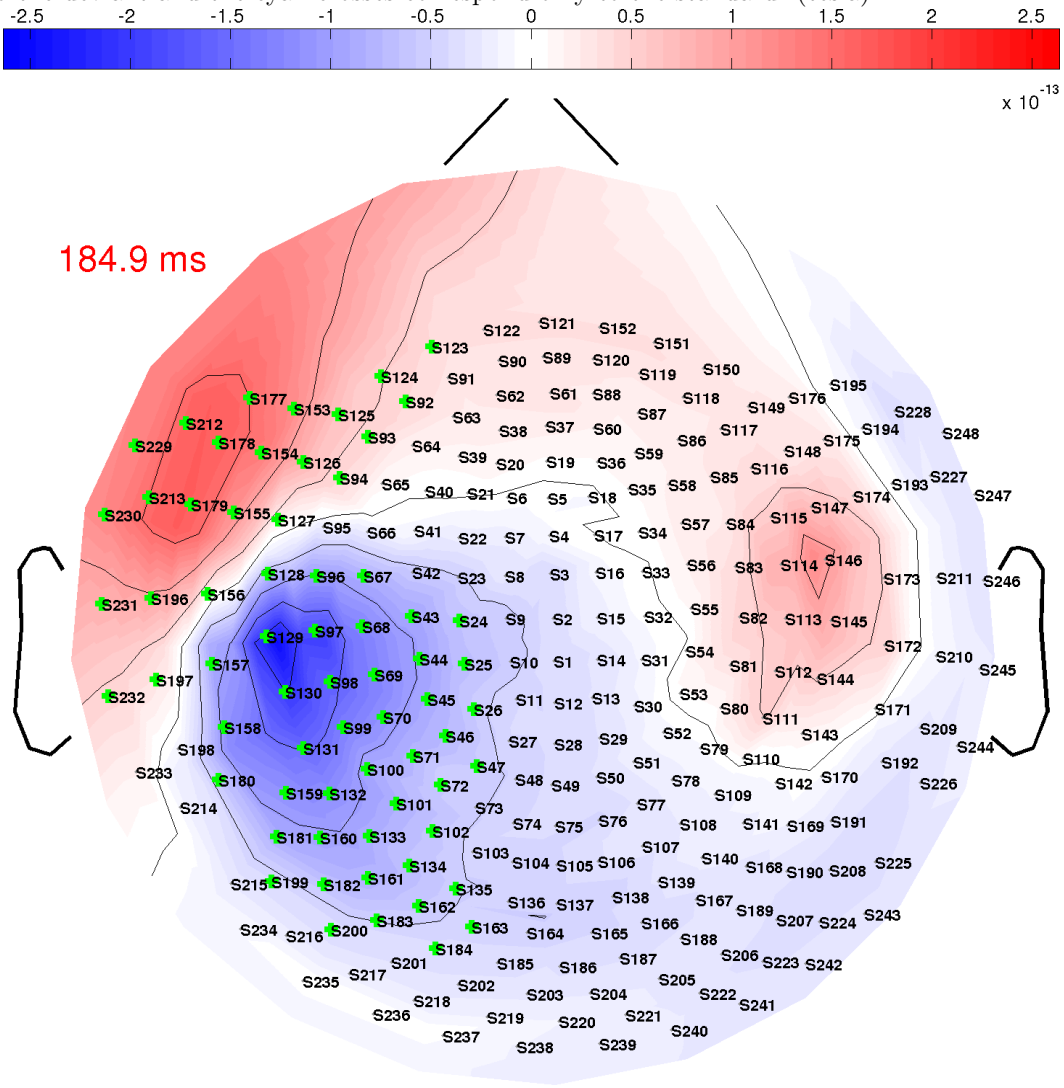


Figure 6.6: subjectOne: channel 97

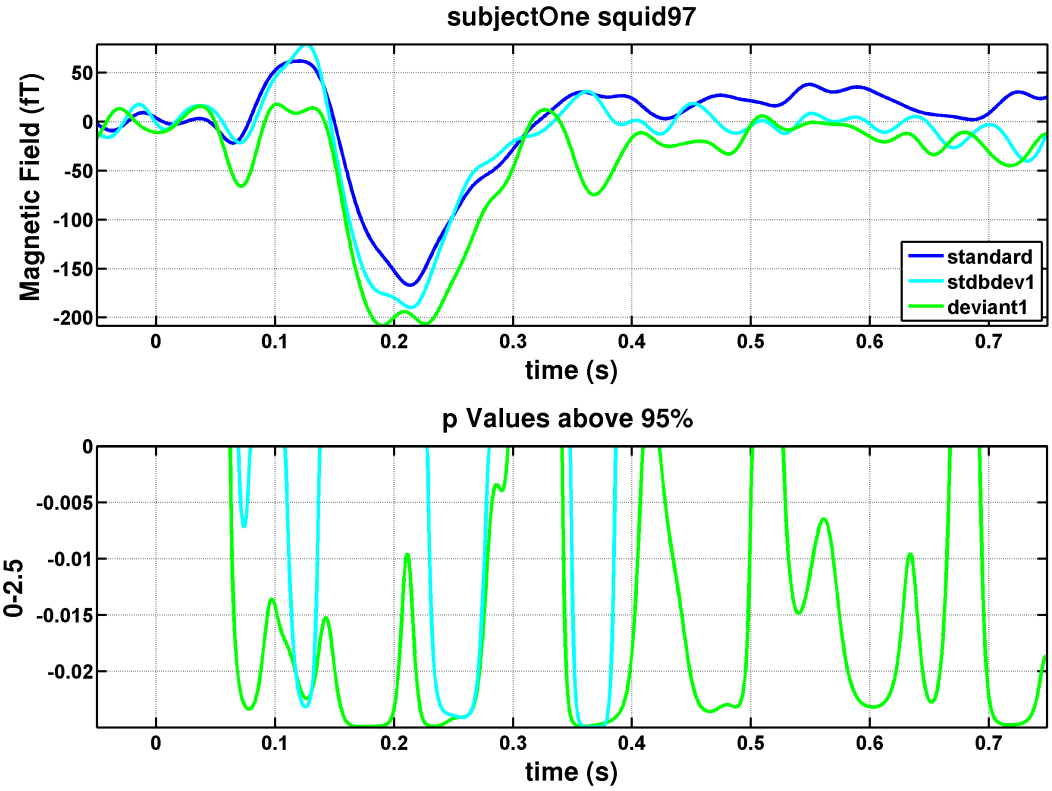
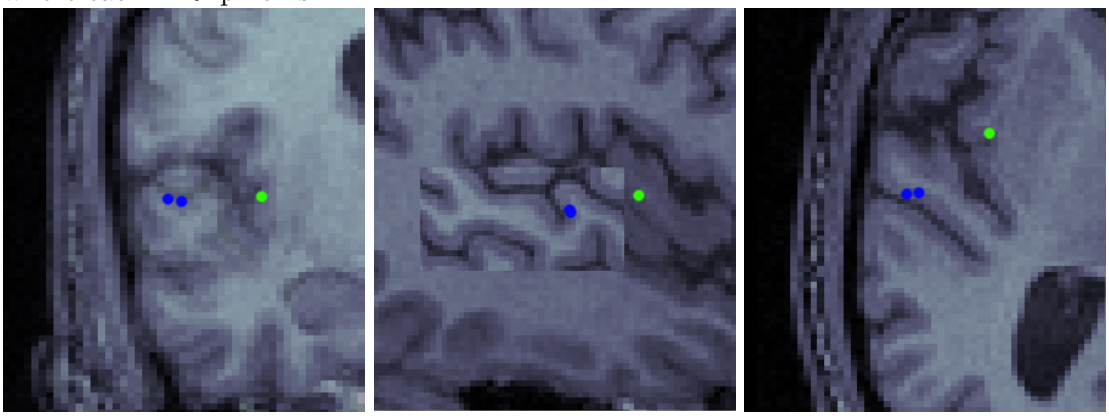


Figure 6.7: subjectOne: The two blue localizations correspond to the standard responses at 170-210 ms (right) and 200-240 ms (left) with signal subspace correlations of 99.2% and 99.1%, respectively. The deviant response is the green localization at a time of 170-240 ms with a signal subspace correlation of 98.9%. Precision to within $\sim$ 5 mm where each MRI pixel is $\sim$ 1 mm.



6.2 subjectTwo results

Asian female age 25.

Figure 6.8: subjectTwo: **(standard-deviant)** at 0.15 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

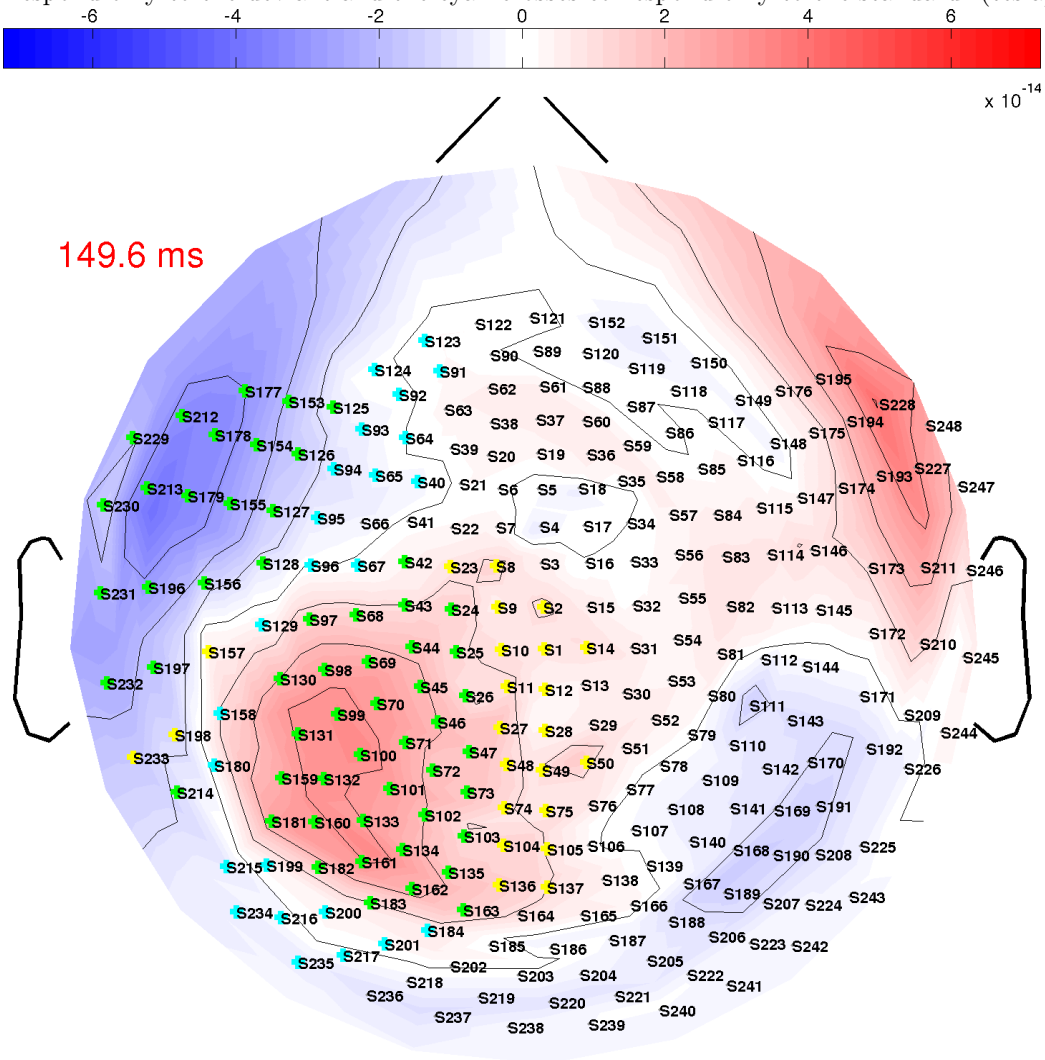


Figure 6.9: subjectTwo: **standard** at 0.3 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

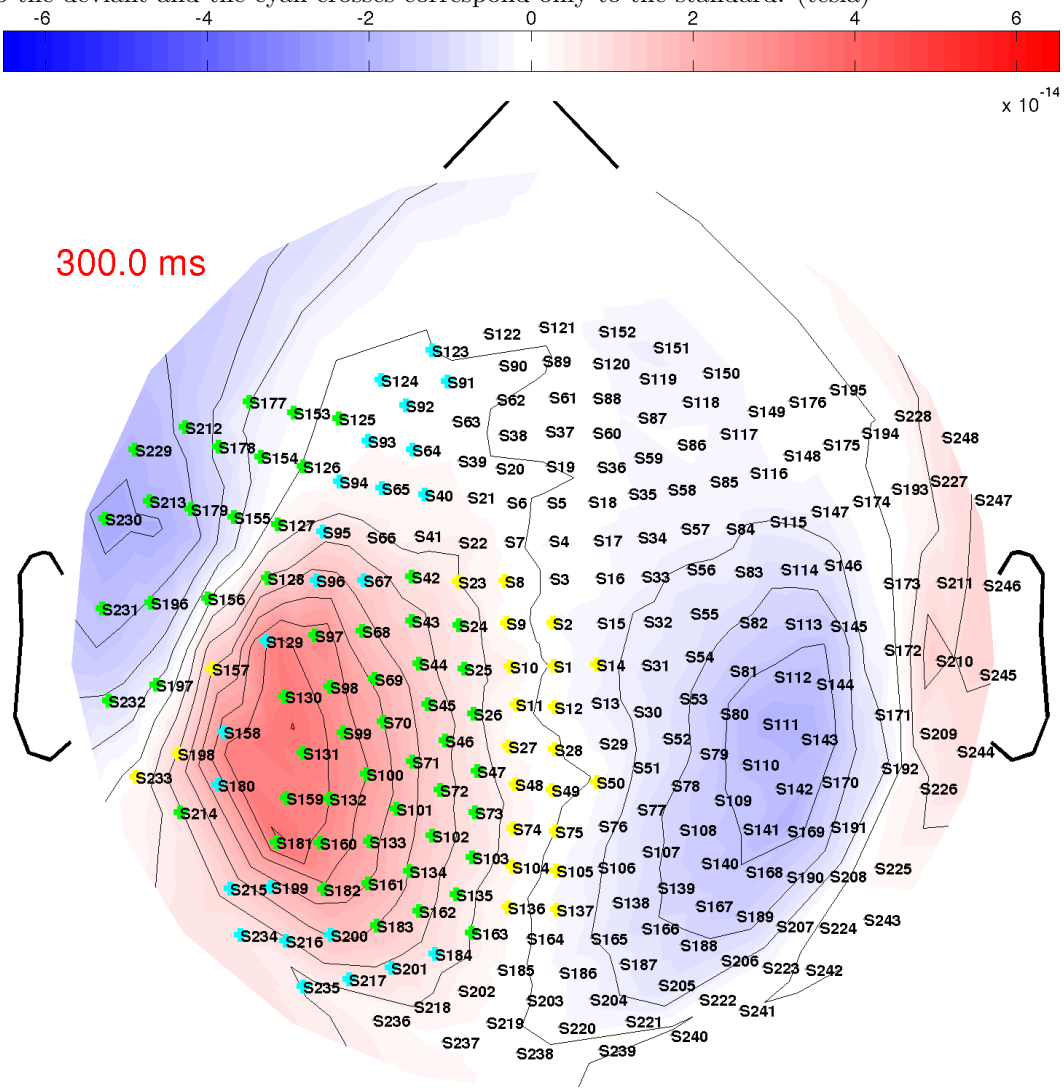


Figure 6.10: subjectTwo: **deviant** at 0.15 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

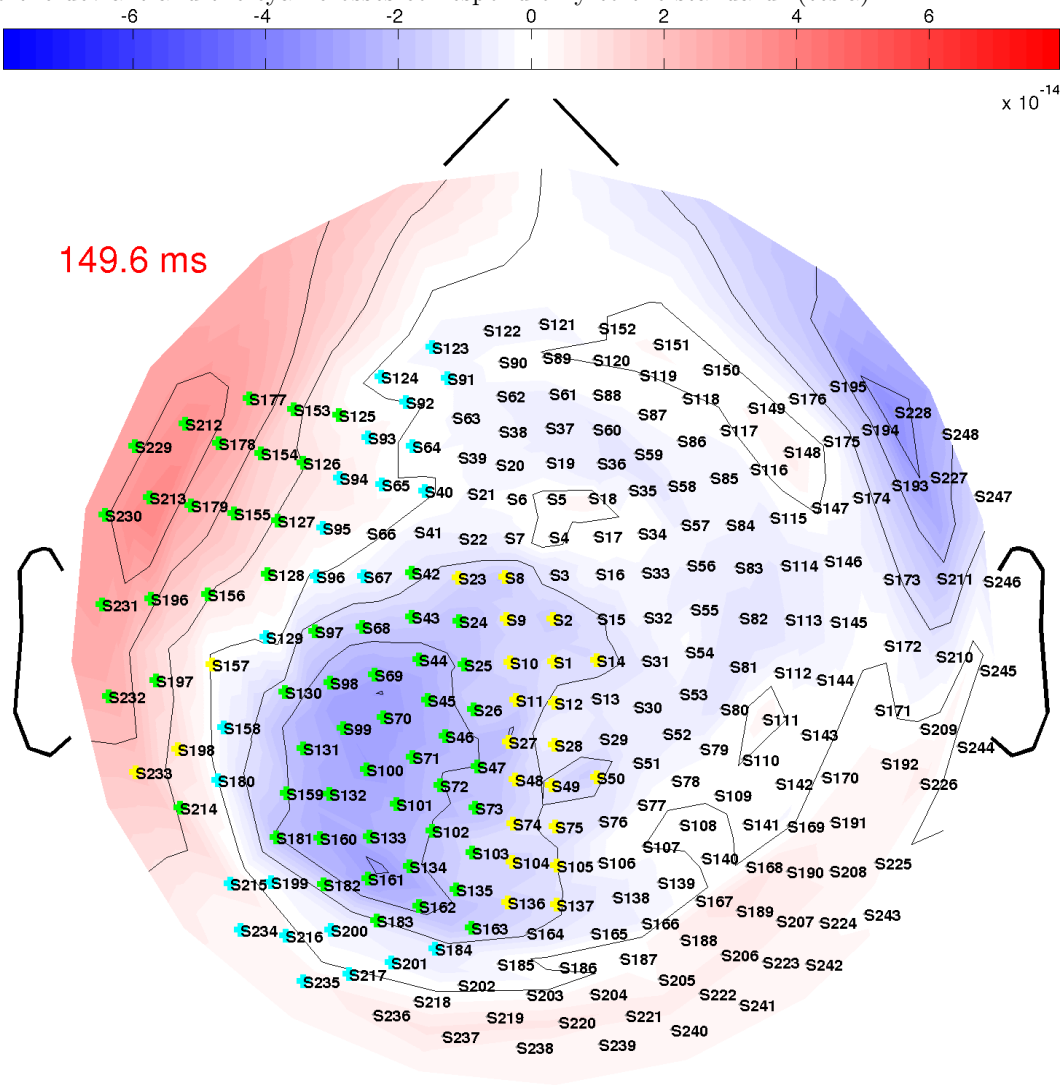


Figure 6.11: subjectTwo: channel 131

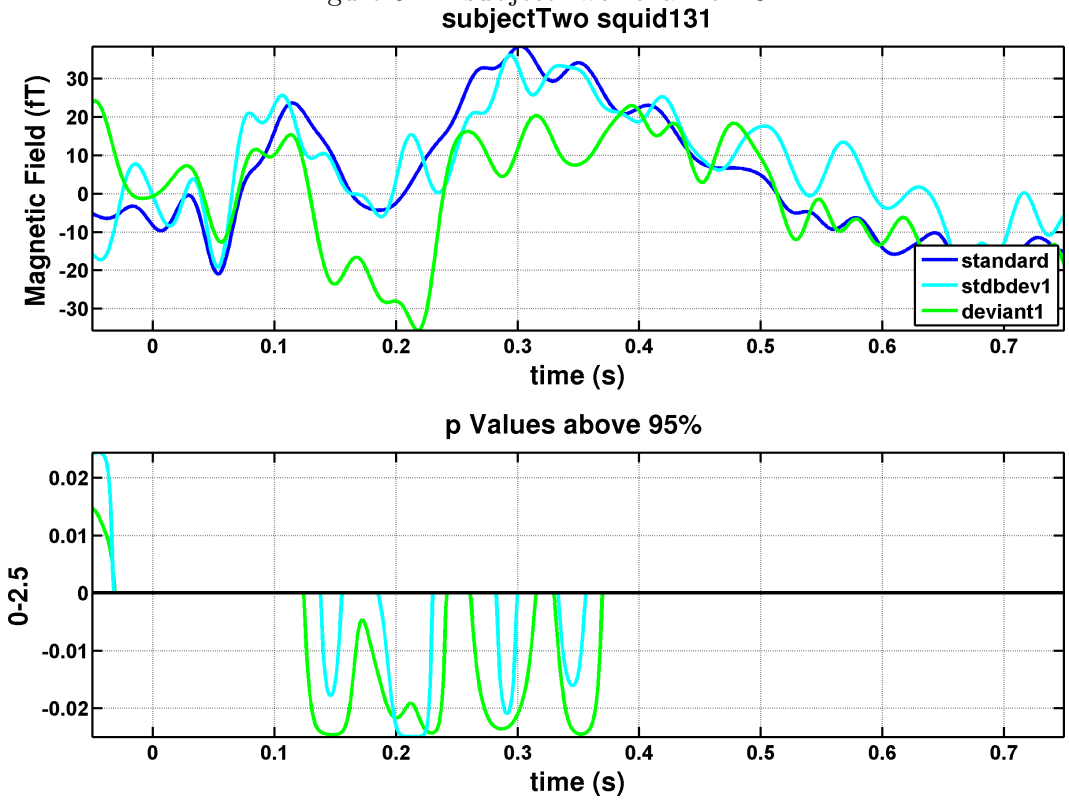
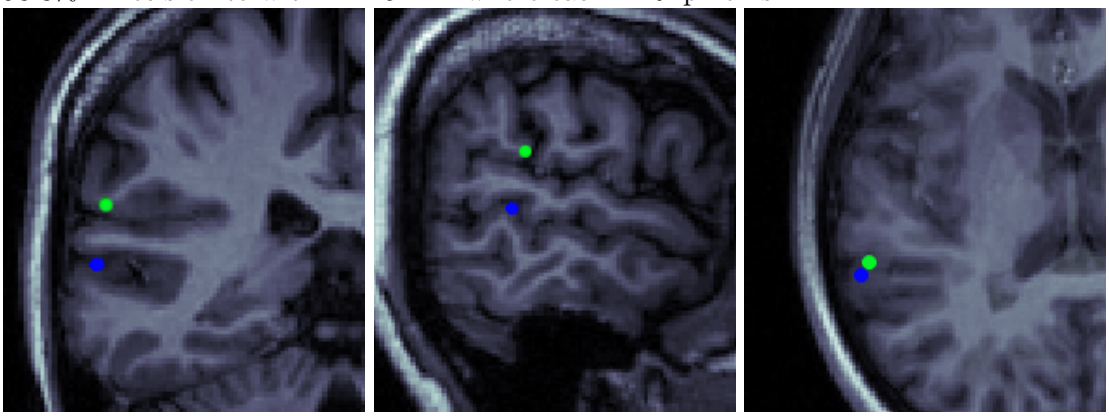


Figure 6.12: subjectTwo: The blue localization corresponds to the standard responses at 270-340 ms with a signal subspace correlation of 99.1%. The green localization corresponds to the deviant response at 135-165 ms with a signal subspace correlation of 98.3%. Precision to within $\sim$ 5 mm where each MRI pixel is $\sim$ 1 mm.



6.3 subjectThree results

Caucasian male age 29. Right handed but left footed.

Figure 6.13: subjectThree: **(standard-deviant)** at 0.163 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

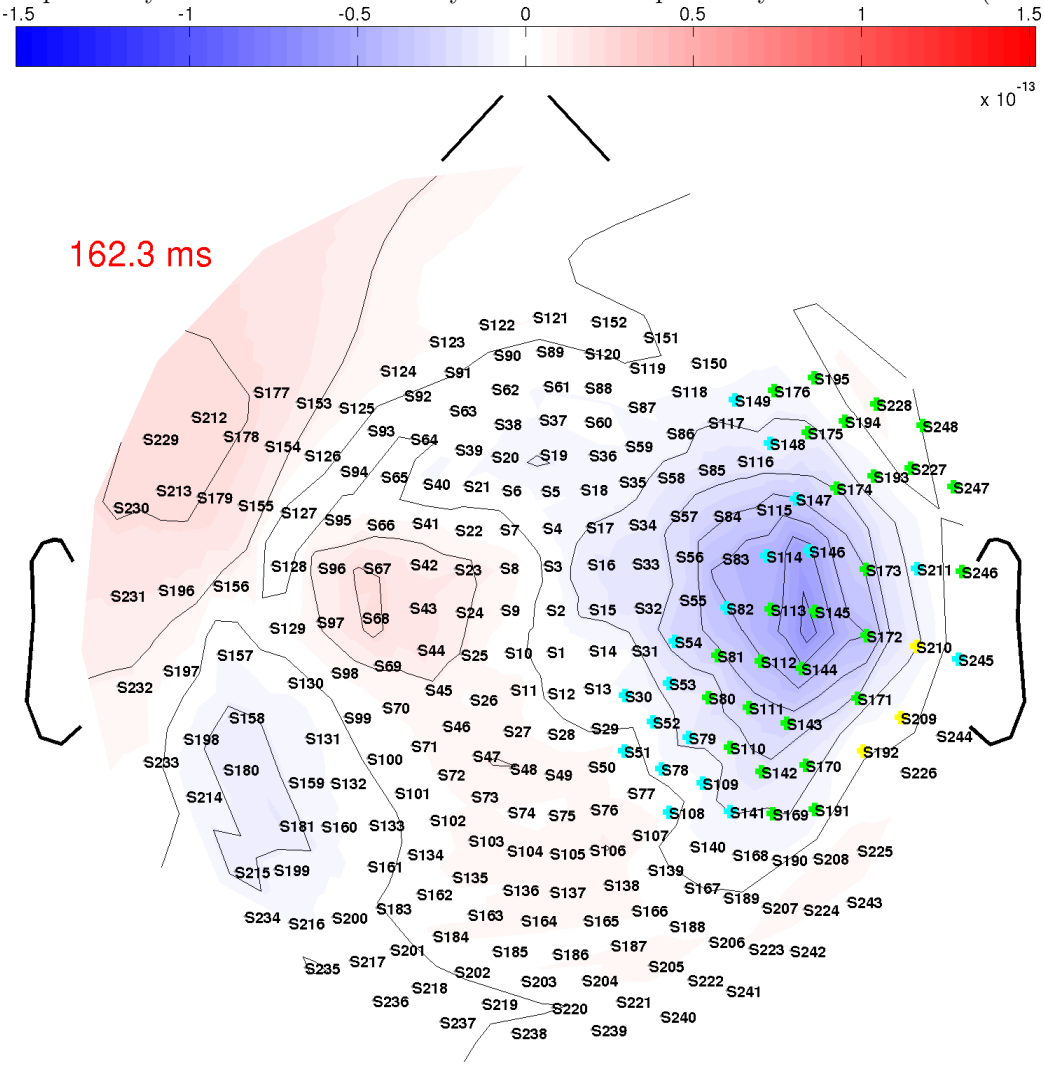


Figure 6.14: subjectThree: **standard** at 0.163 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

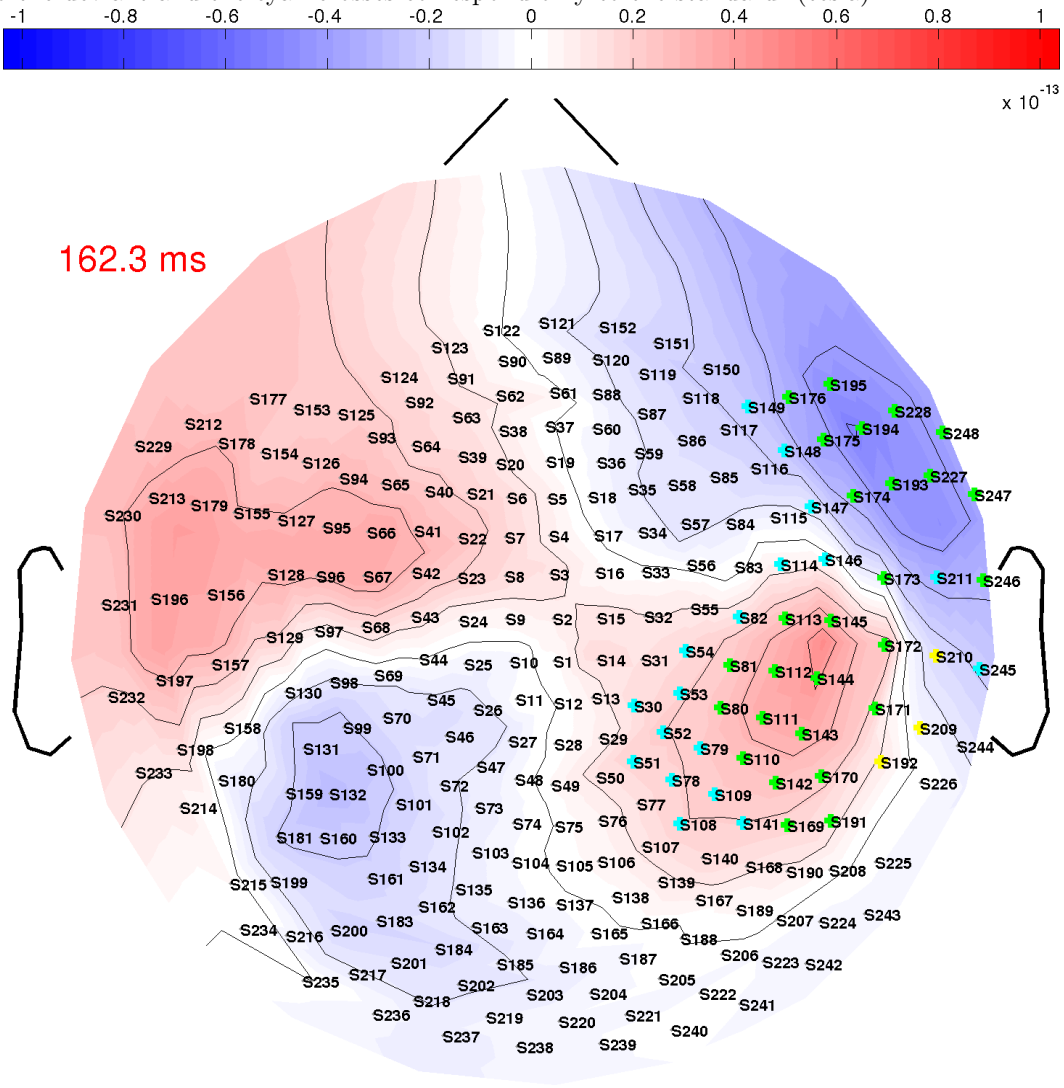


Figure 6.15: subjectThree: **deviant** at 0.163 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

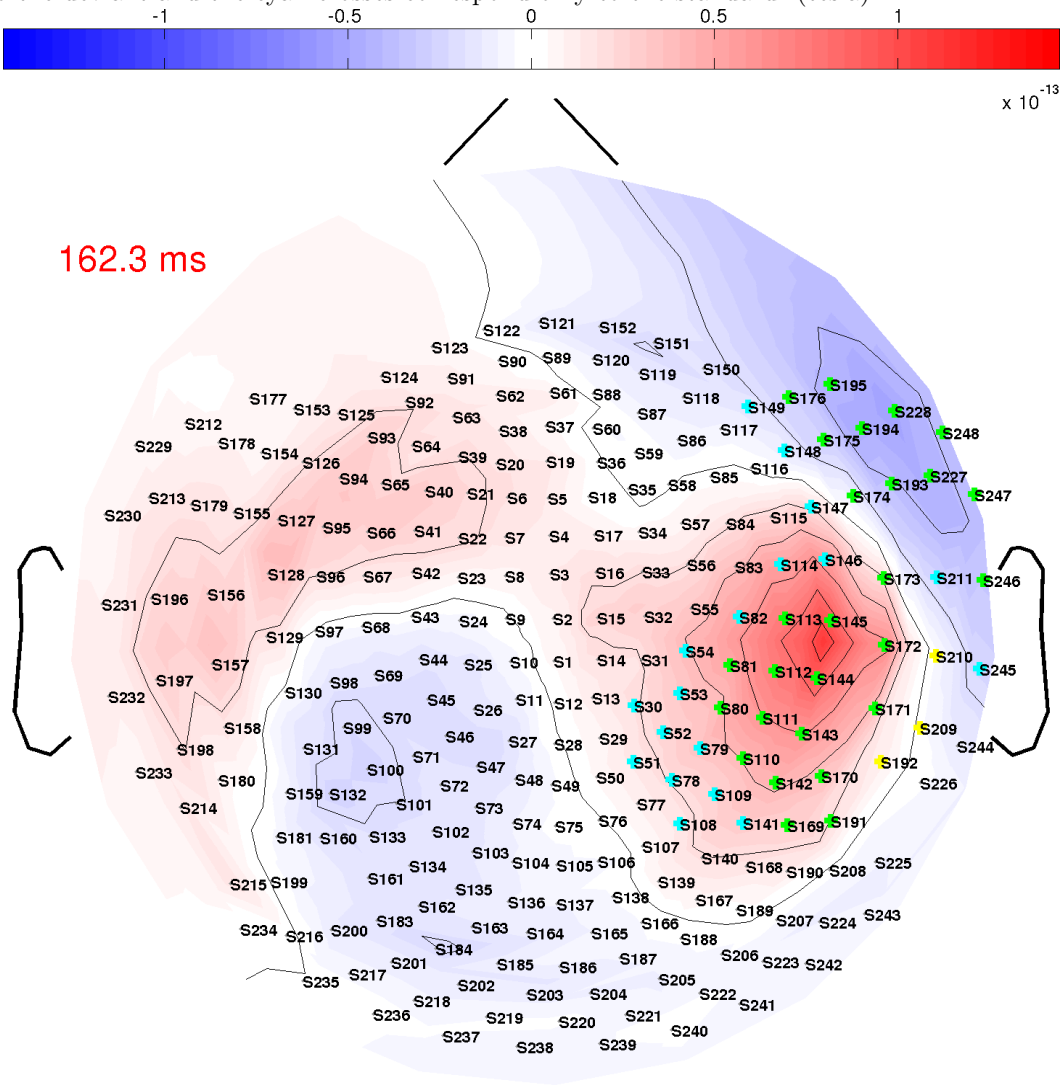


Figure 6.16: subjectThree: channel 145

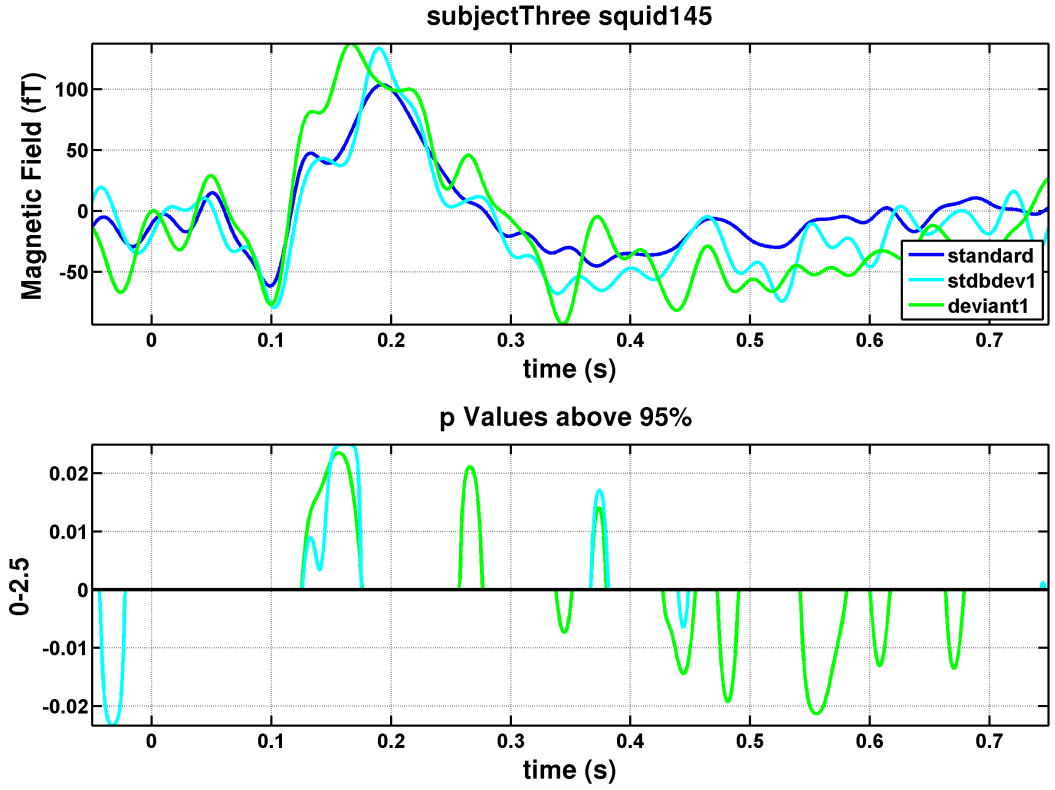
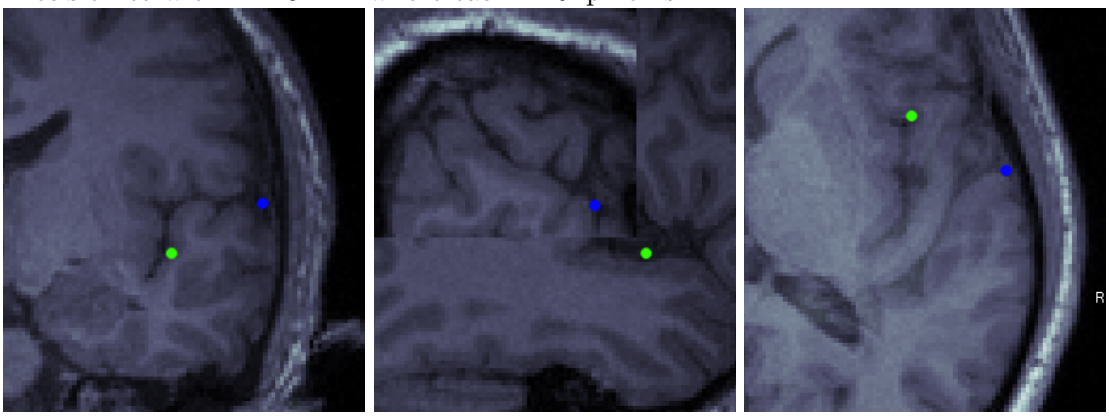


Figure 6.17: subjectThree: The blue localization corresponds to the standard response at 143-168 ms with a signal subspace correlation of 99.3%. The deviant response is the green localization at a time of 143-168 ms with a signal subspace correlation of 99.0%. Precision to within $\sim$ 5 mm where each MRI pixel is $\sim$ 1 mm.



6.4 subjectFour results

Caucasian female age 22. Dental artifact in channels nearest jaw.

Figure 6.18: subjectFour: **(standard-deviant)** at 0.263 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

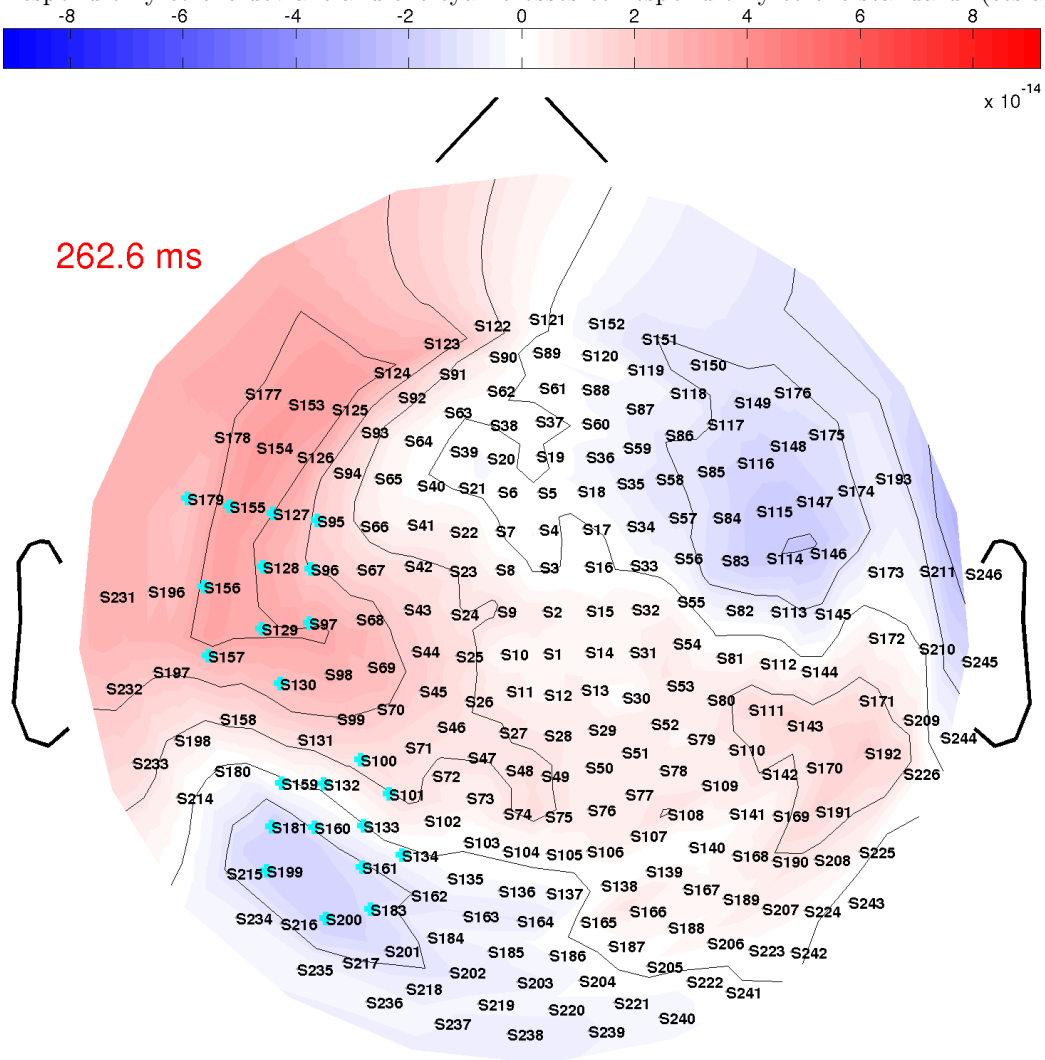


Figure 6.19: subjectFour: **standard** at 0.263 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

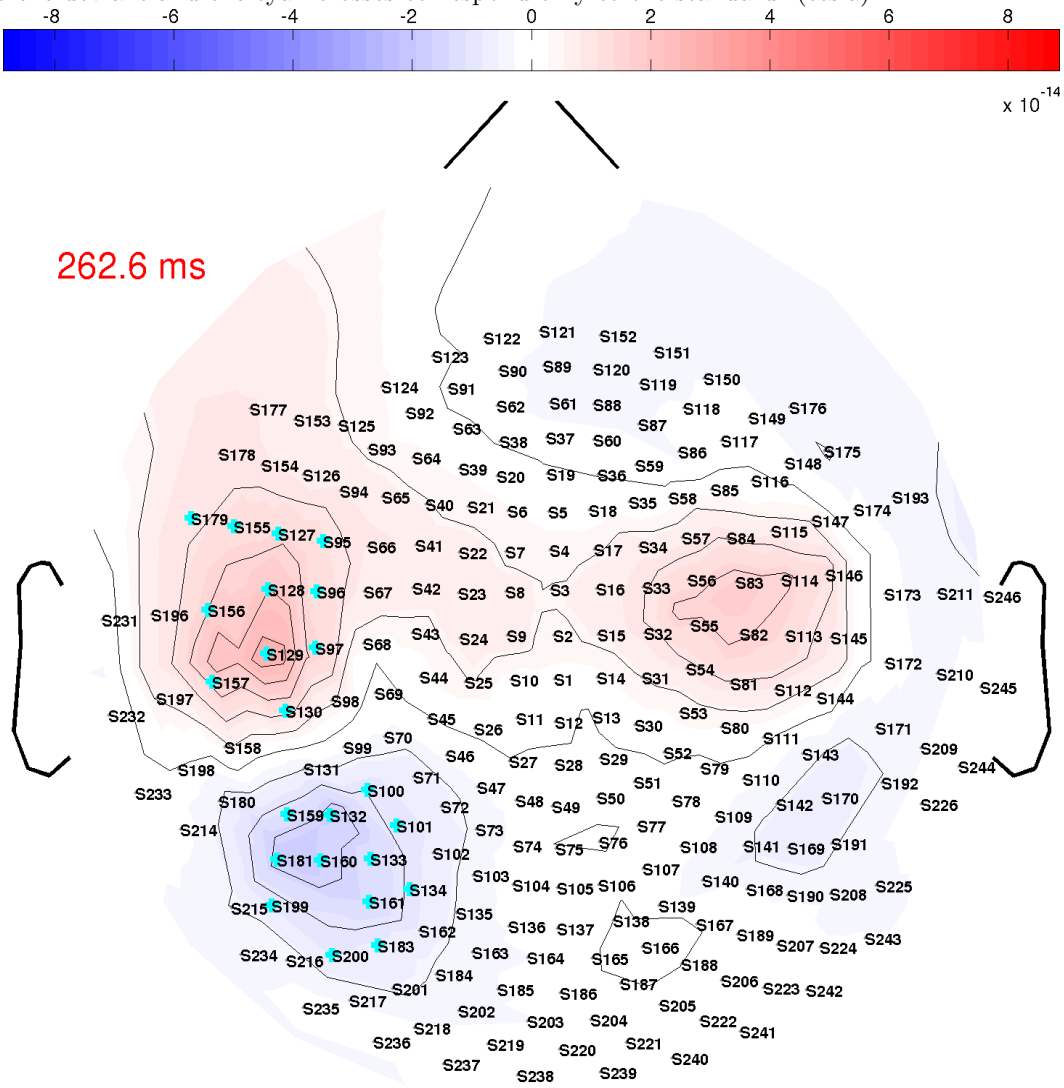


Figure 6.20: subjectFour: **deviant** at 0.263 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

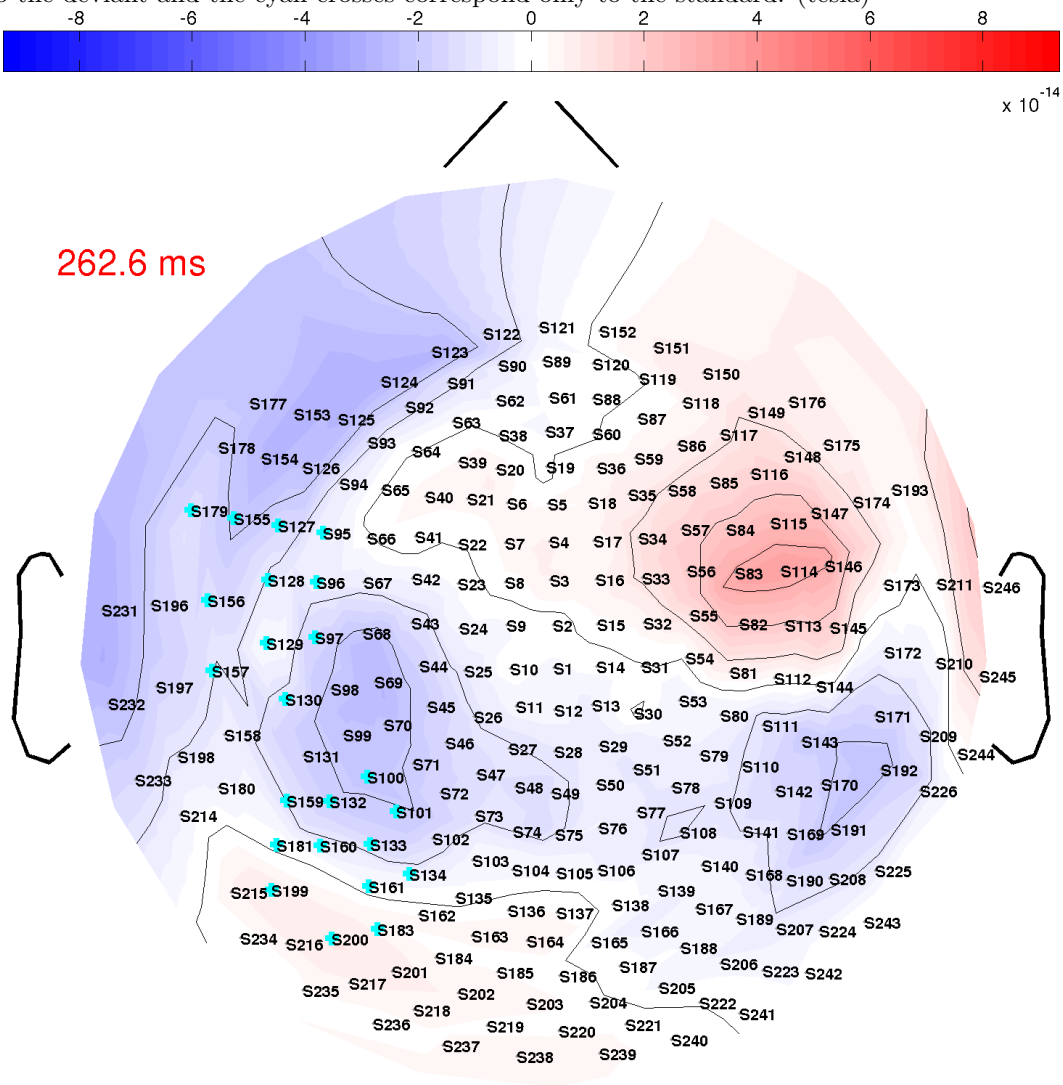


Figure 6.21: subjectFour: channel 156

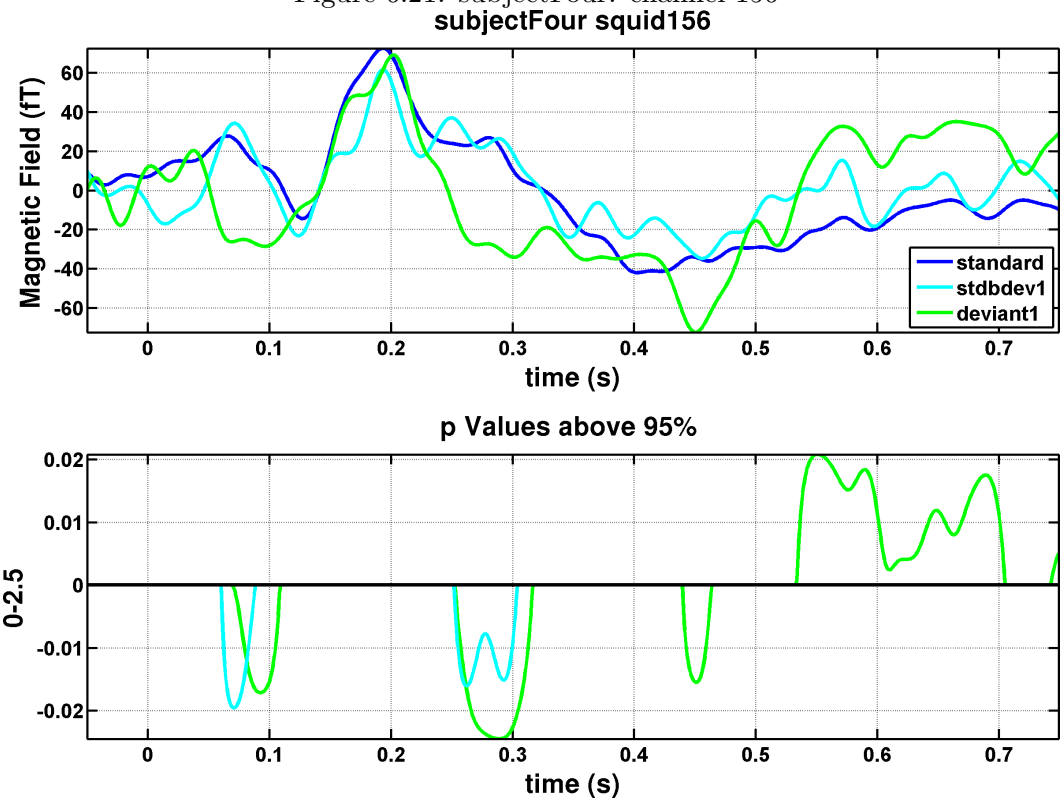
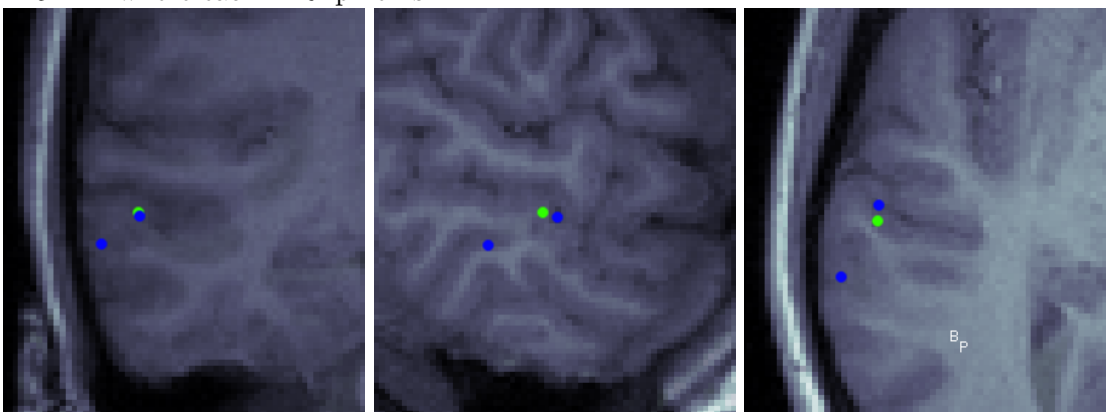


Figure 6.22: subjectFour: The two blue localizations correspond to the standard responses at 190-200 ms (superior) and 264-271 ms (inferior) with signal subspace correlations of 99.6% and 99.7%, respectively. The deviant response is the green localization at a time of 190-200 ms with a signal subspace correlation of 99.8%. Precision to within $\sim$ 5 mm where each MRI pixel is $\sim$ 1 mm.



6.5 subjectFive results

Caucasian male age 22.

Figure 6.23: subjectFive: **(standard-deviant)** at 0.136 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

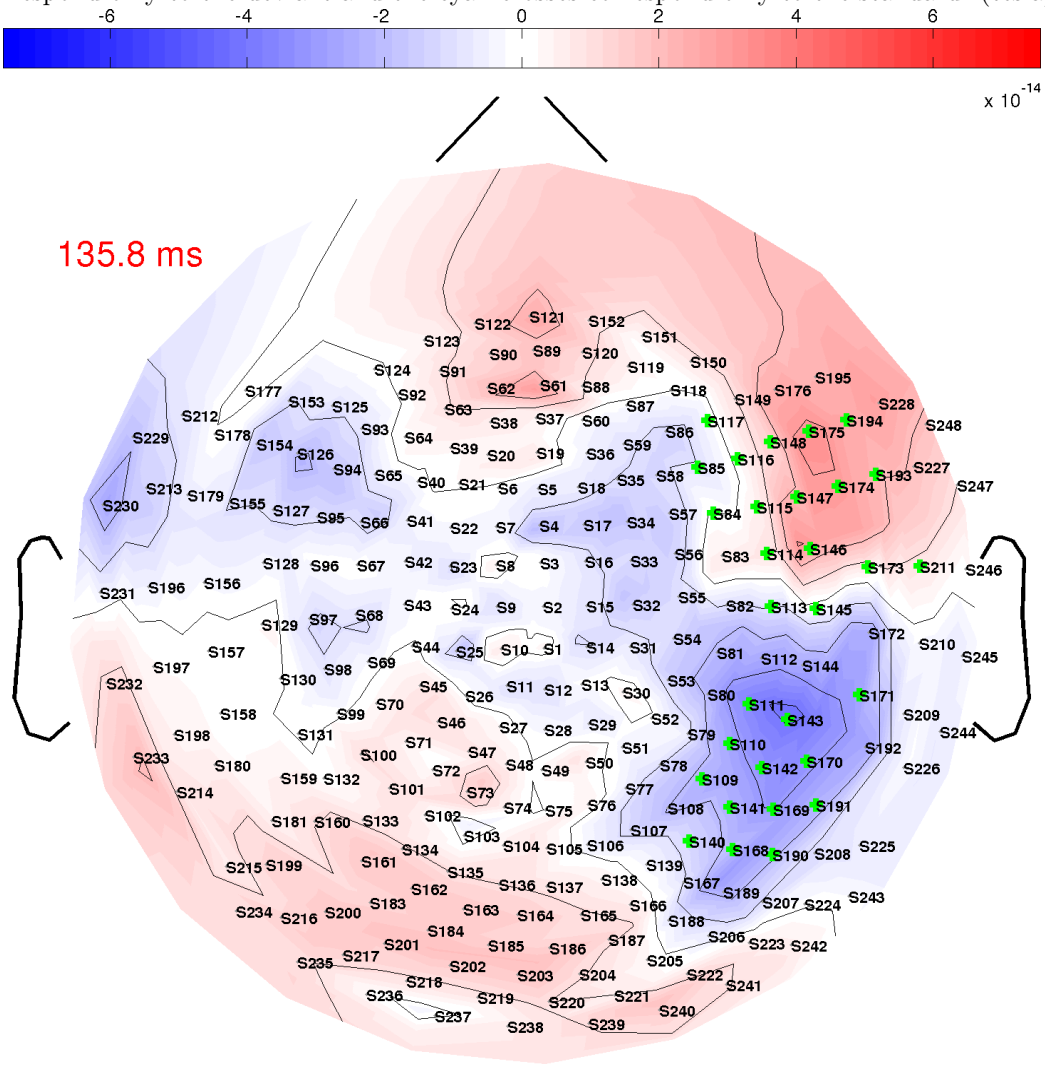


Figure 6.24: subjectFive: **standard** at 0.136 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

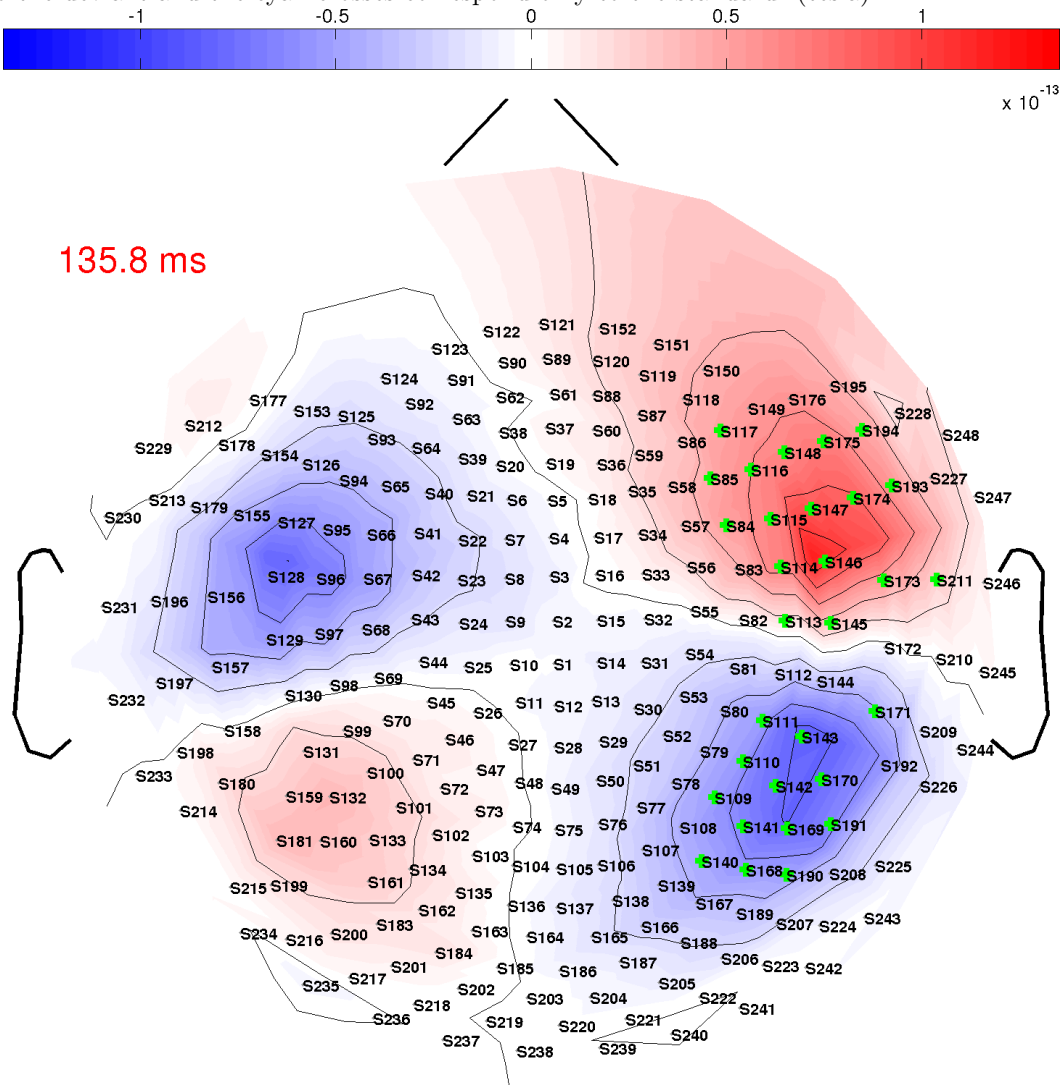


Figure 6.25: subjectFive: **deviant** at 0.136 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

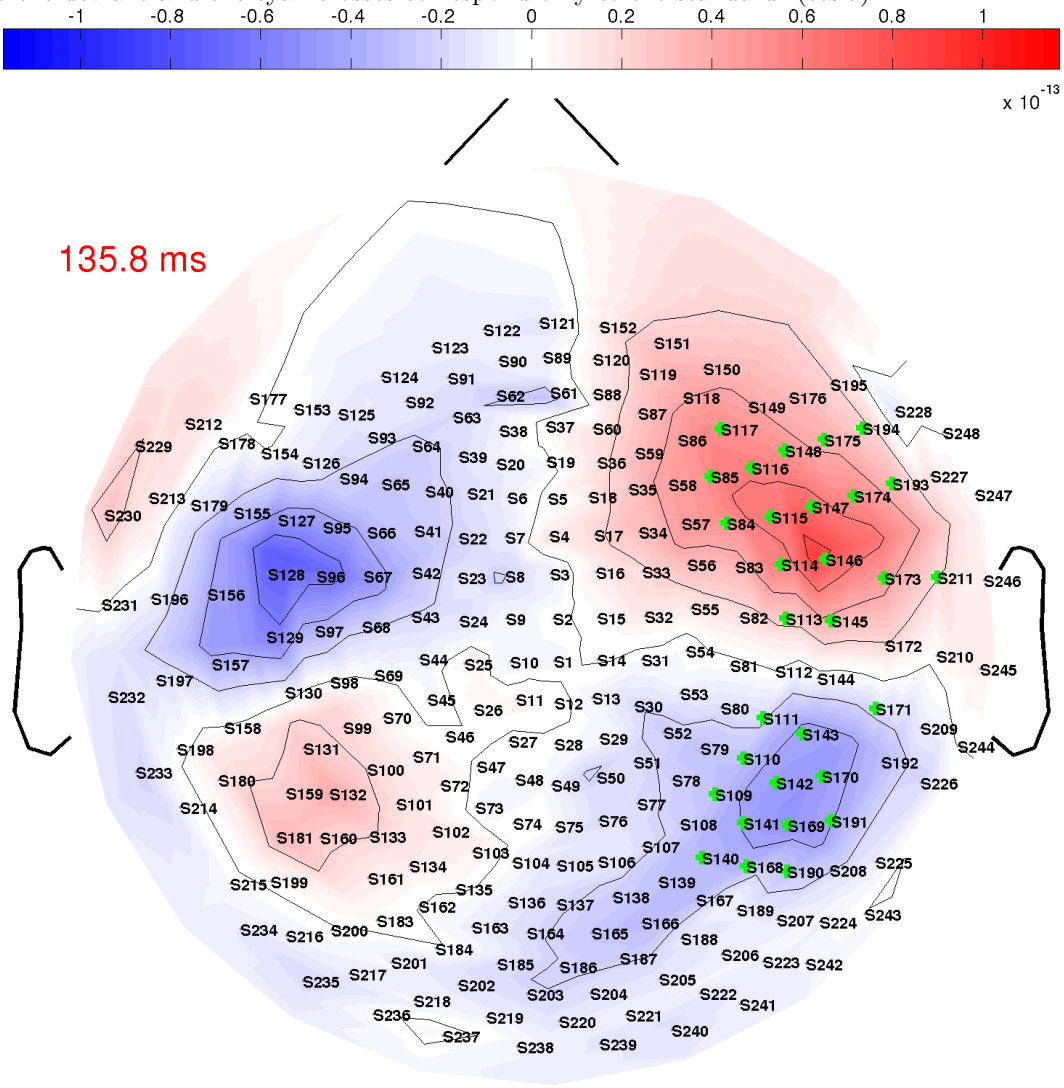


Figure 6.26: subjectFive: channel 174

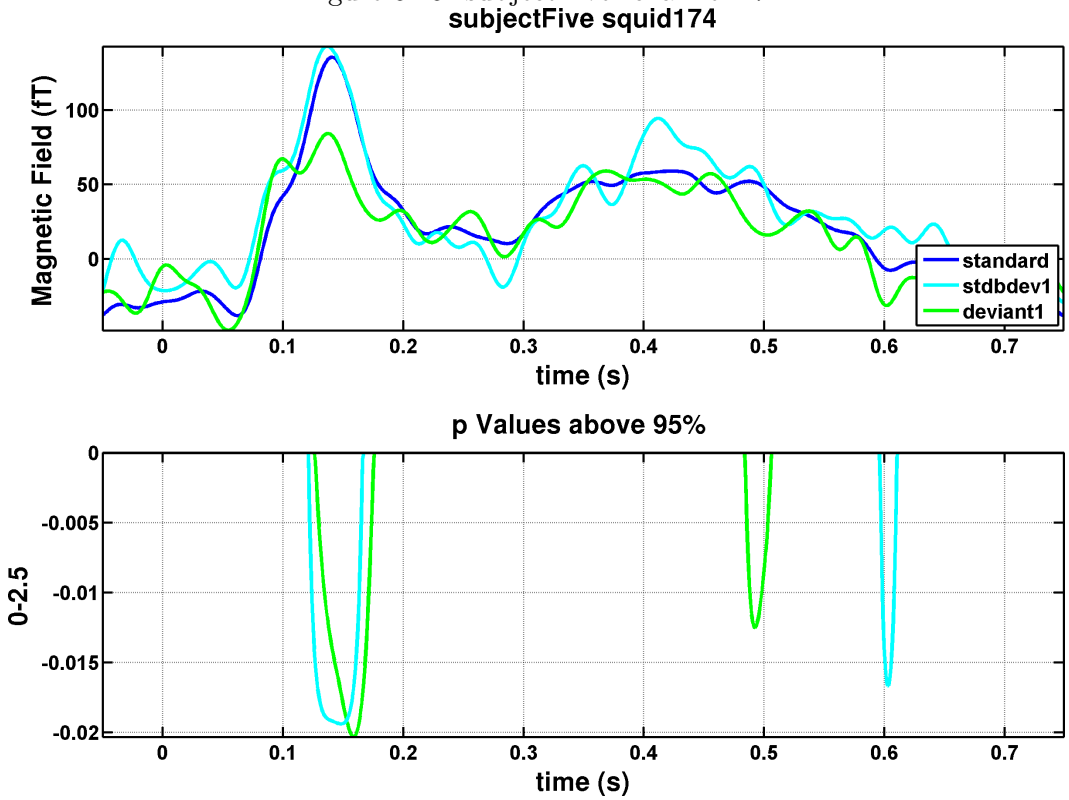
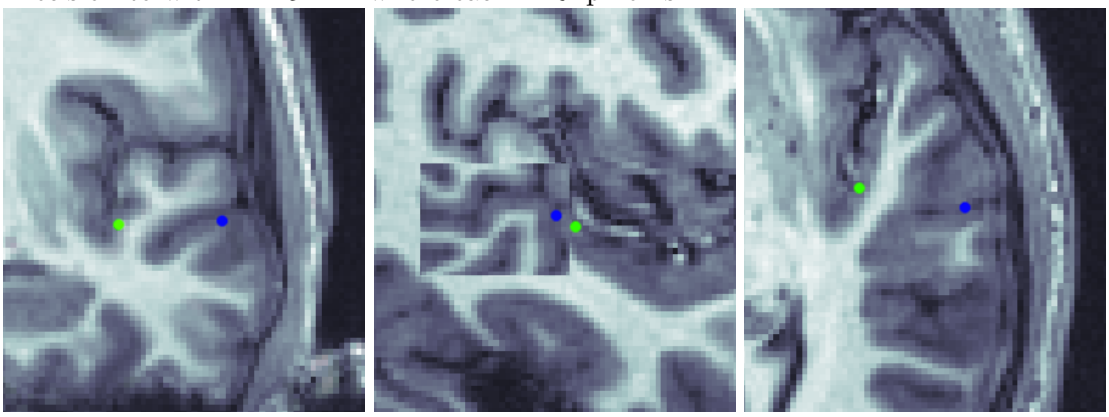


Figure 6.27: subjectFive: The blue localization corresponds to the standard response at 130-140 ms with a signal subspace correlation of 99.6%. The deviant response is the green localization at a time of 130-140 ms with a signal subspace correlation of 97.5%. Precision to within $\sim$ 5 mm where each MRI pixel is $\sim$ 1 mm.



6.6 subjectSix results

Hispanic female age 22.

Figure 6.28: subjectSix: **(standard-deviant)** at 0.249 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

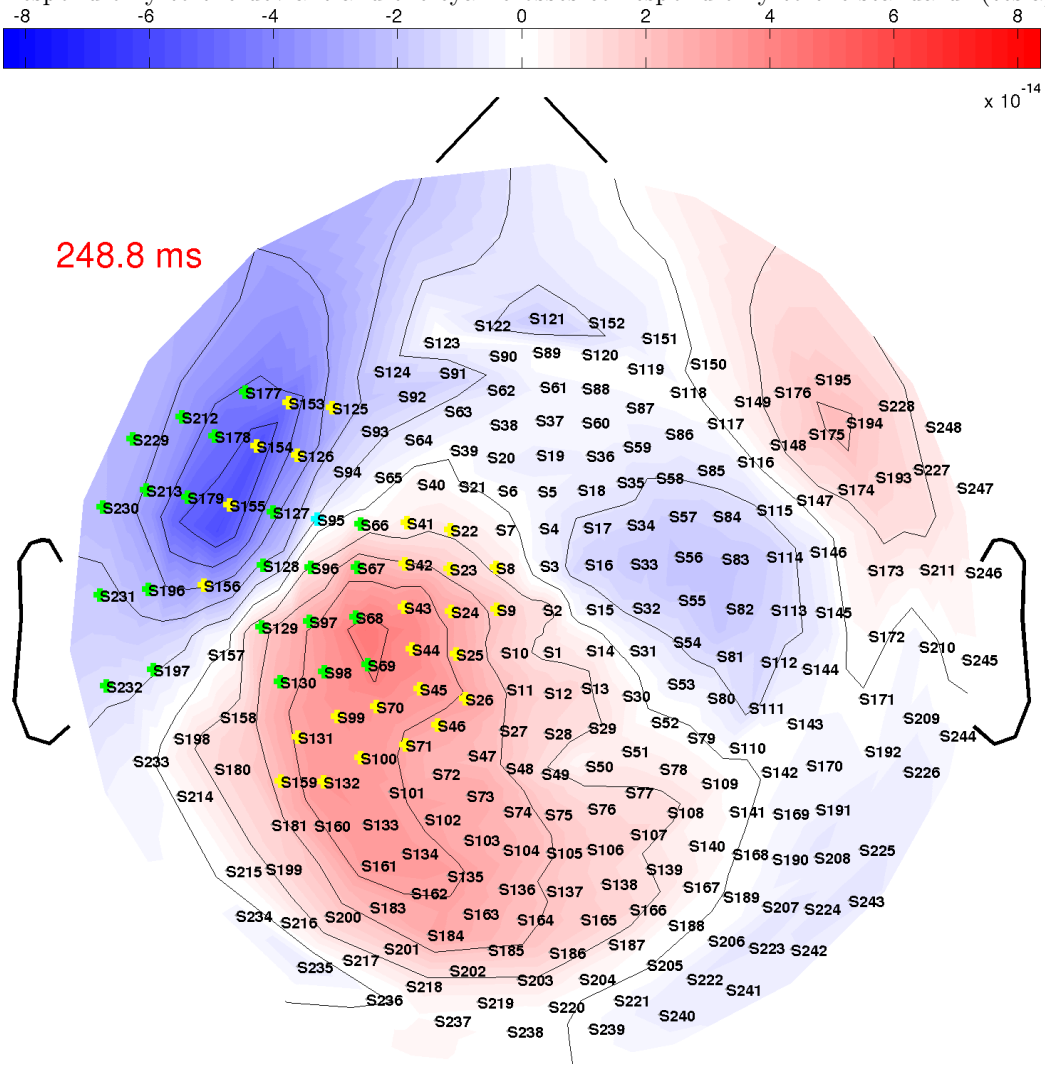


Figure 6.29: subjectSix: **standard** at 0.225 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

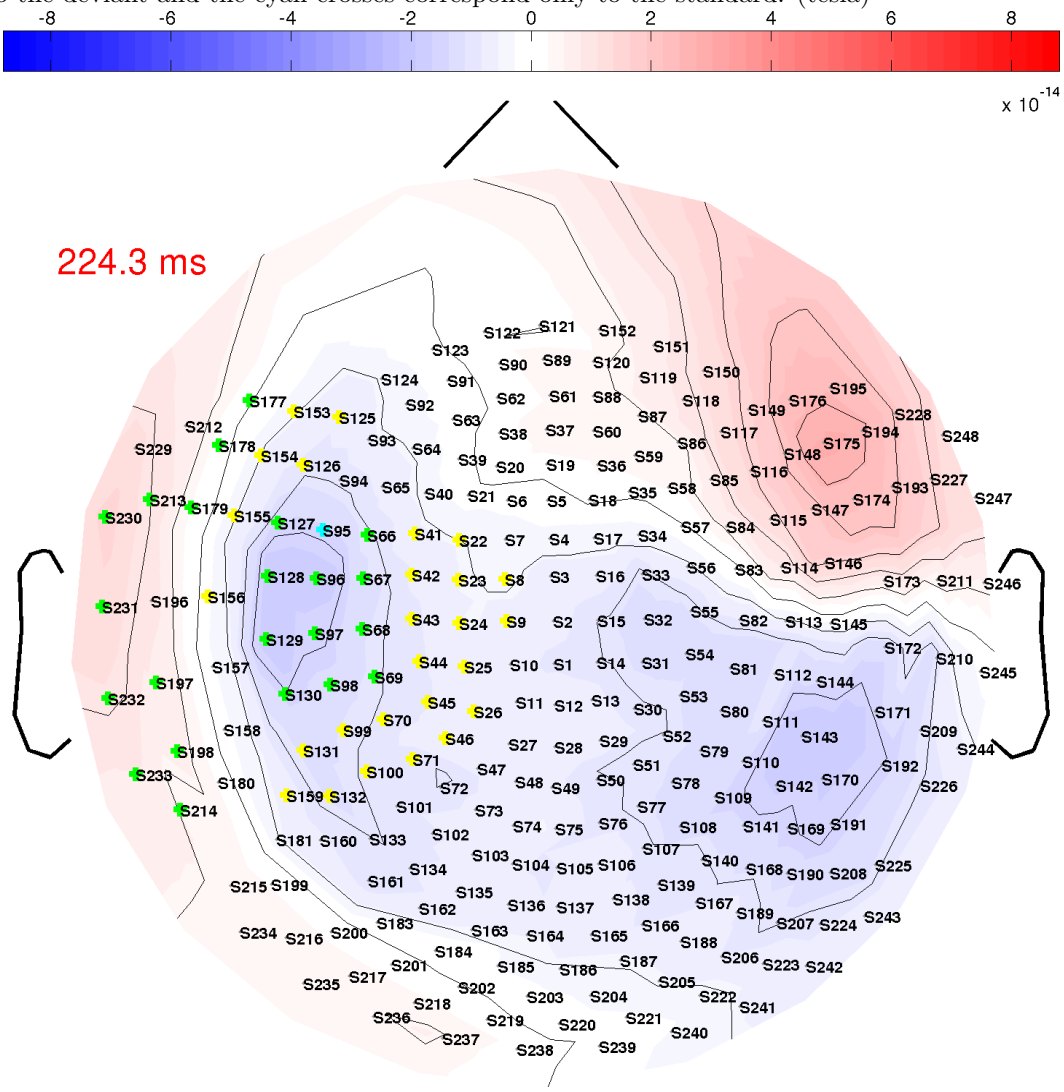


Figure 6.30: subjectSix: **deviant** at 0.249 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

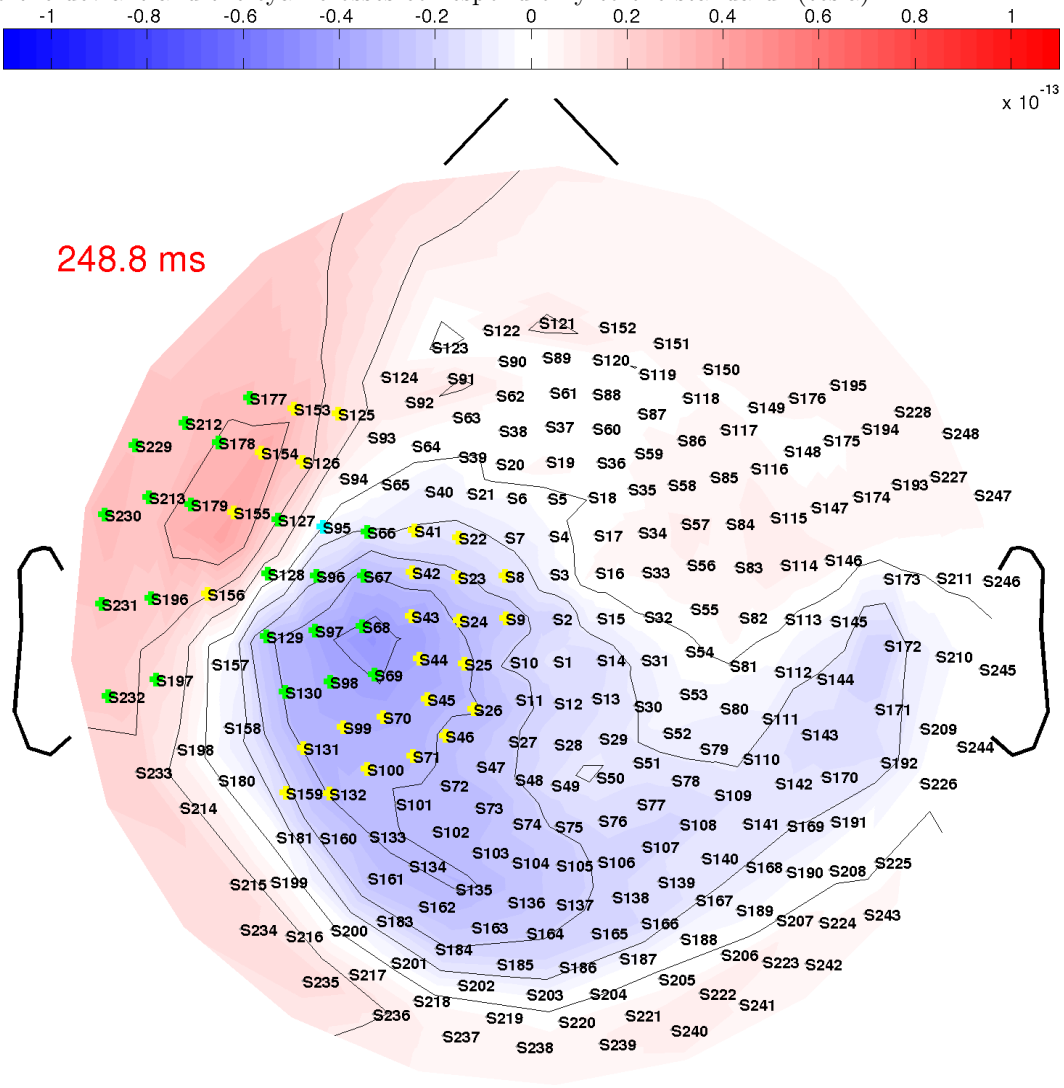


Figure 6.31: subjectSix: channel 155

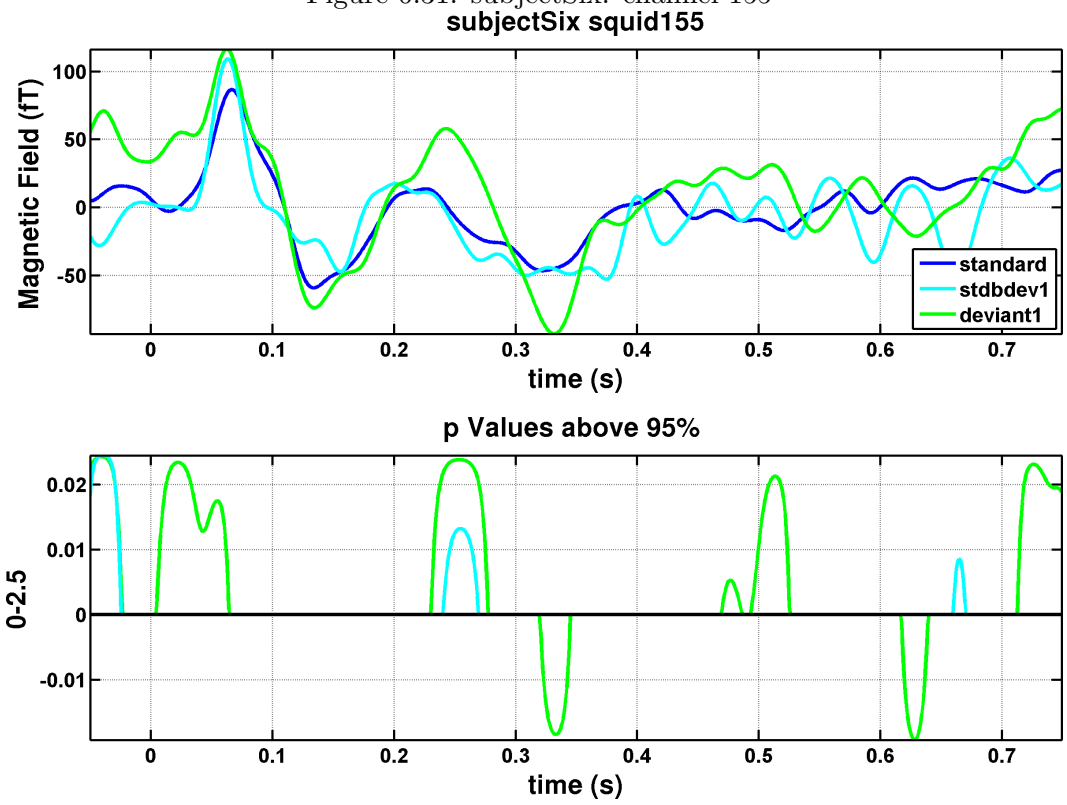
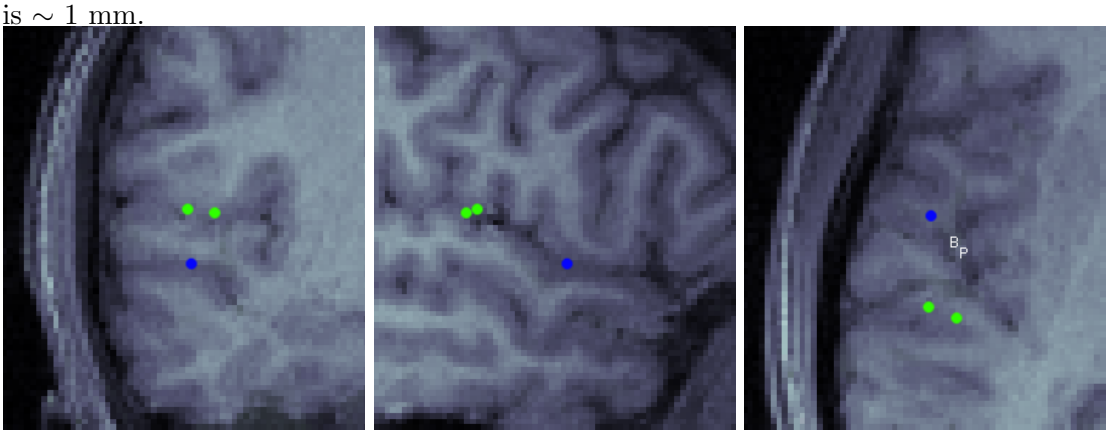


Figure 6.32: subjectSix: The blue localization corresponds to the standard response at 215-230 ms with a signal subspace correlation of 99.2%. The two green localizations correspond to deviant responses at 195-215 ms (left) and 230-260 ms (right) with each signal subspace correlation of 99.2%. Precision to within ~ 5 mm where each MRI pixel is ~ 1 mm.



6.7 subjectSeven results

Caucasian male age 18. Reports musical training.

Figure 6.33: subjectSeven: **(standard-deviant)** at 0.242 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

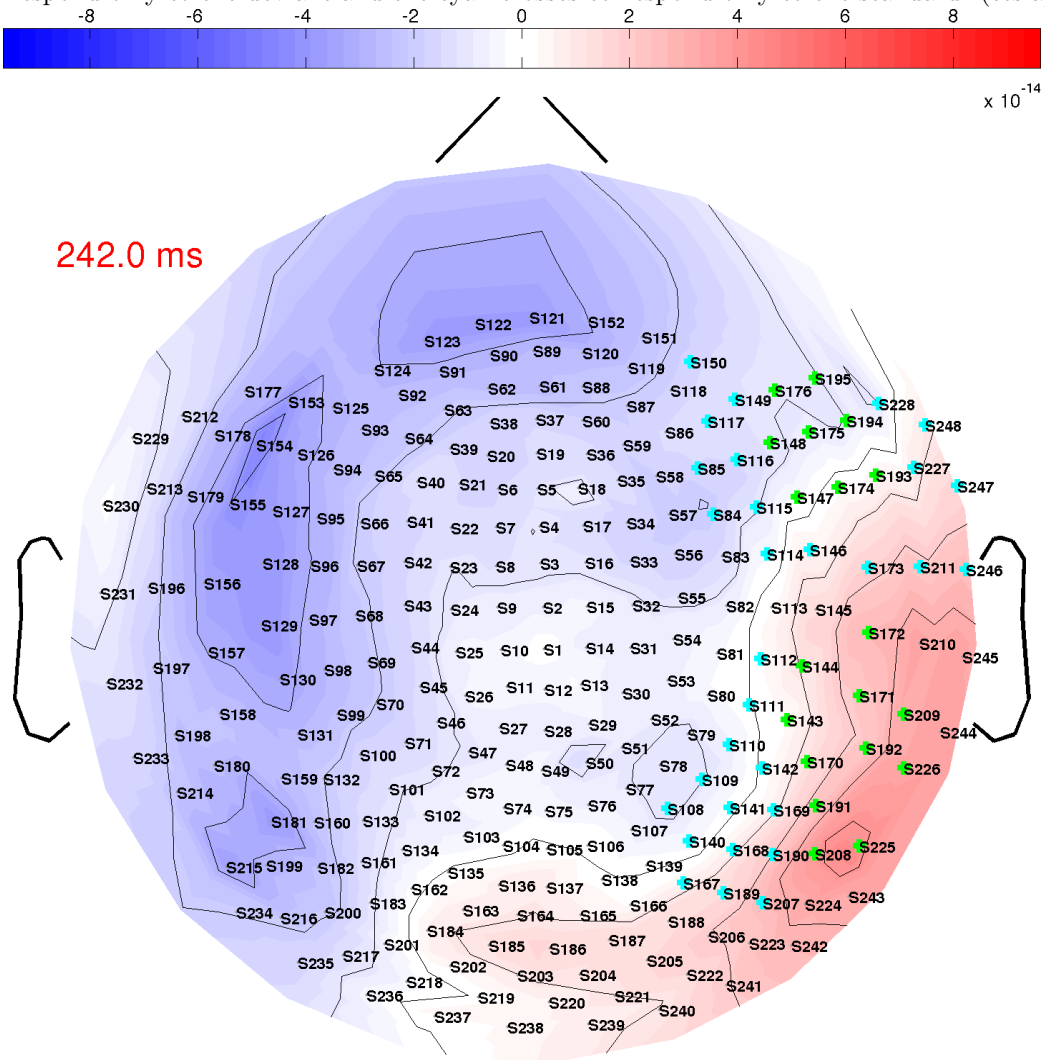


Figure 6.34: subjectSeven: **standard** at 0.242 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

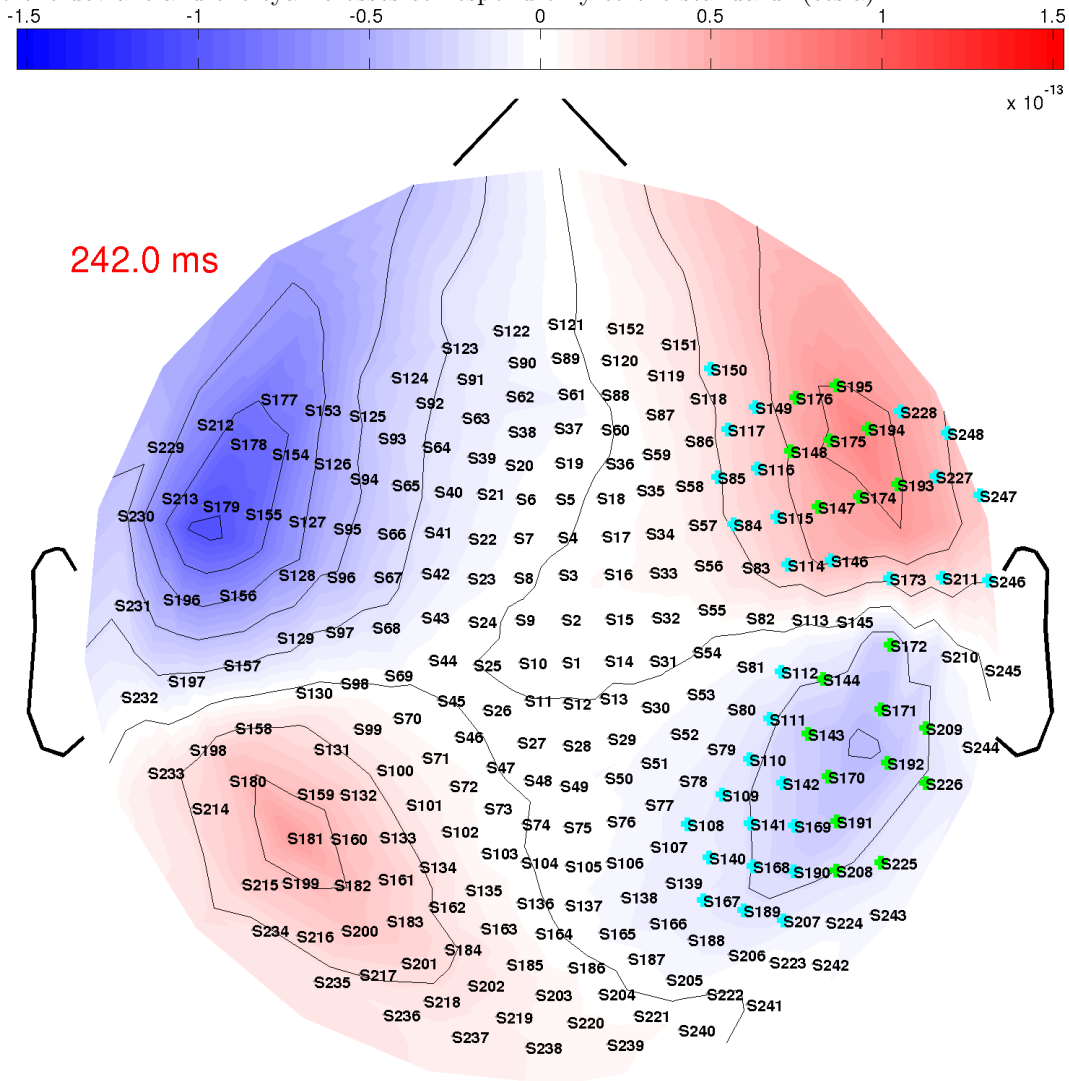


Figure 6.35: subjectSeven: **deviant** at 0.242 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

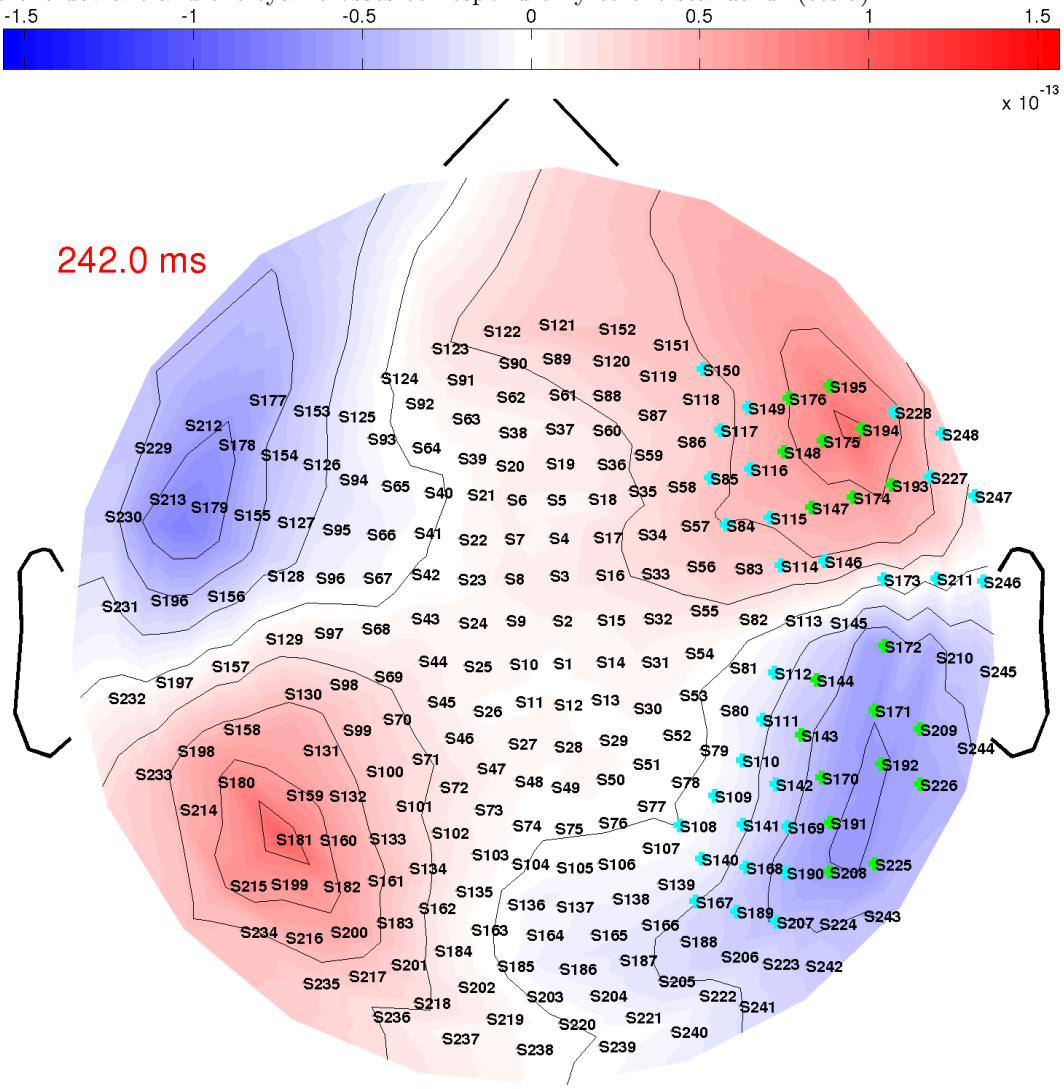


Figure 6.36: subjectSeven: channel 208

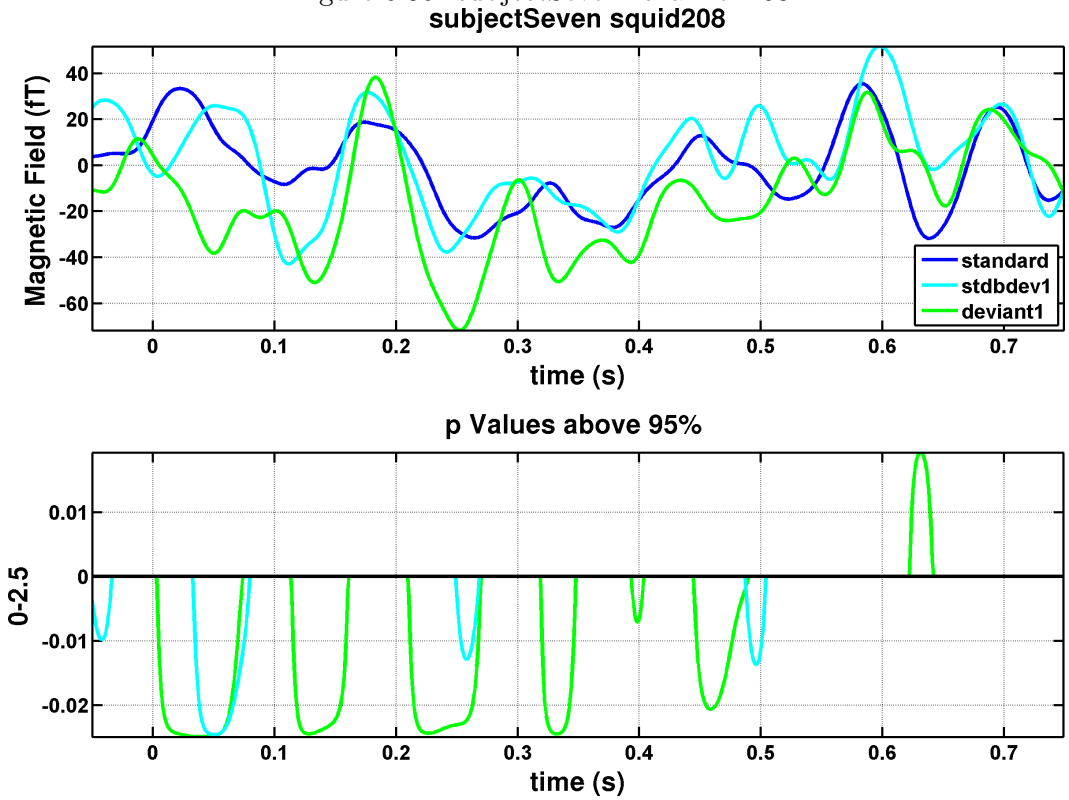
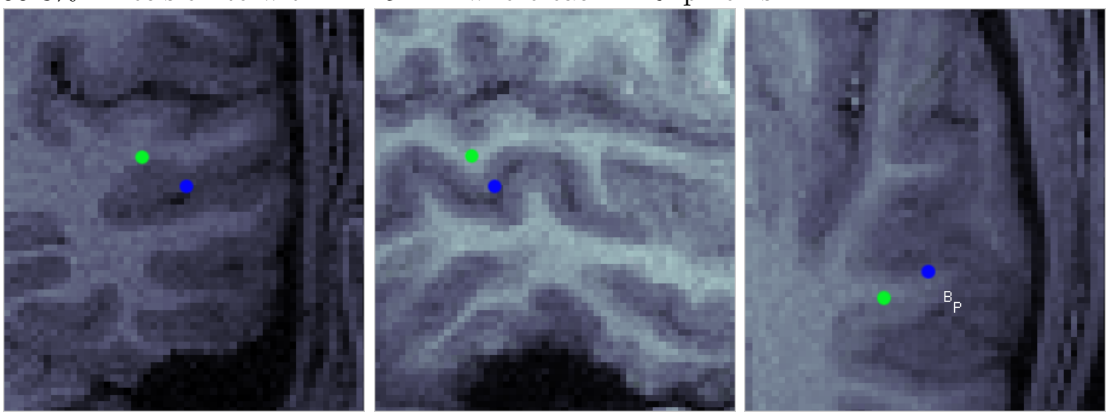


Figure 6.37: subjectSeven: The blue localization corresponds to the standard response at 240-265 ms with a signal subspace correlation of 99.6%. The green localization corresponds to the deviant response at 240-260 ms with signal subspace correlations of 99.8%. Precision to within $\sim$ 5 mm where each MRI pixel is $\sim$ 1 mm.



6.8 subjectEight results

Asian female age 28.

Figure 6.38: subjectEight: **(standard-deviant)** at 0.165 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

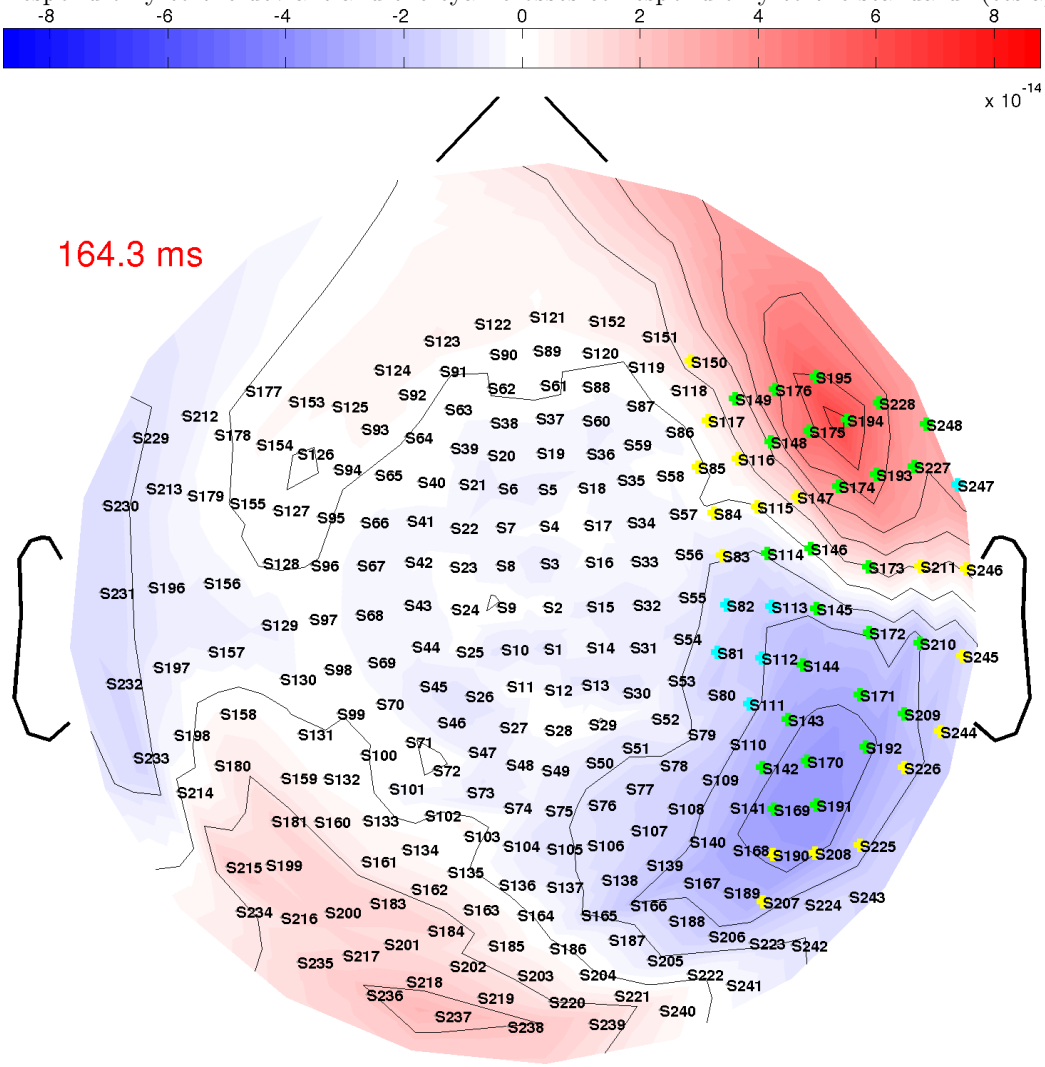


Figure 6.39: subjectEight: **standard** at 0.165 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

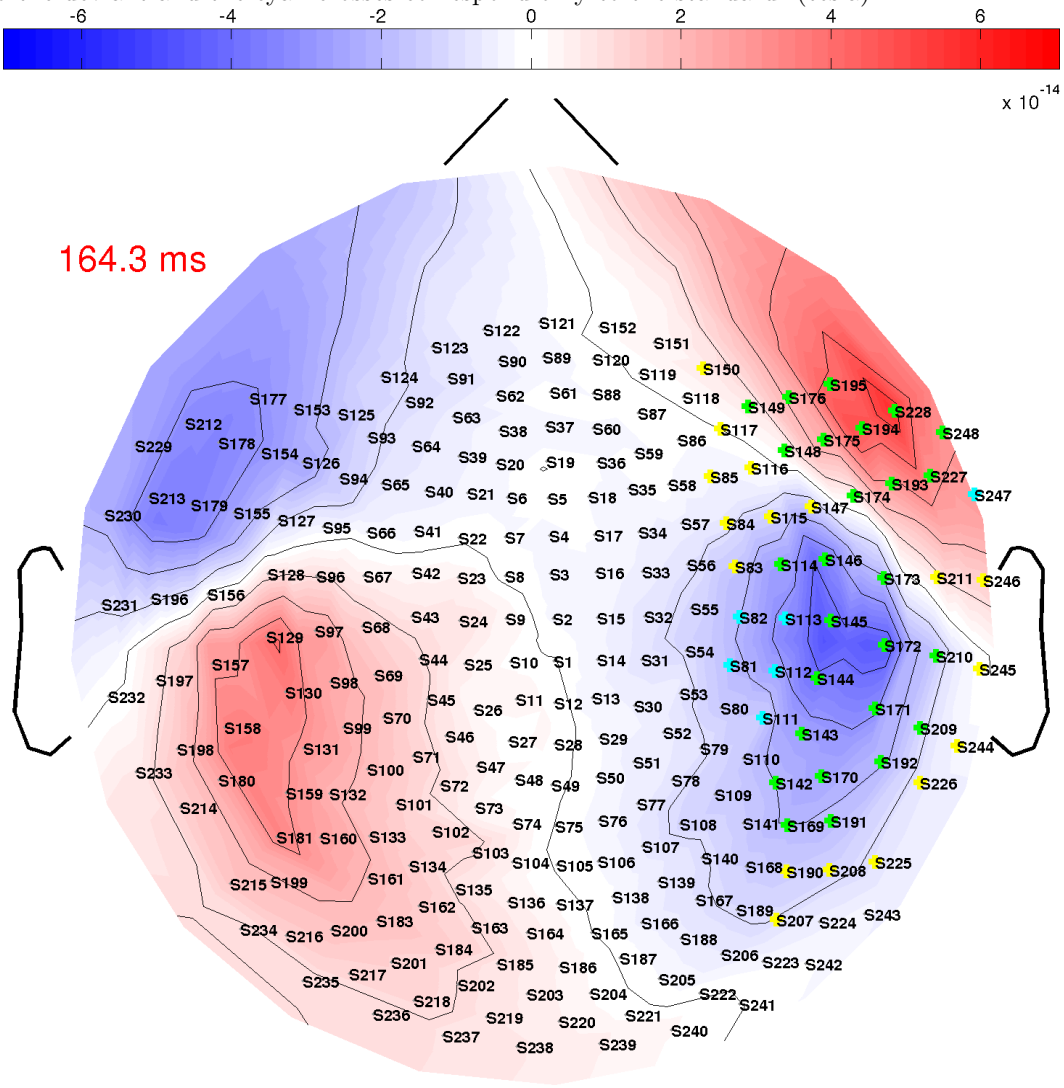


Figure 6.40: subjectEight: **deviant** at 0.117 s. The green crosses indicate channels that were selected for the localizations of both standard and deviant. The yellow crosses correspond only to the deviant and the cyan crosses correspond only to the standard. (tesla)

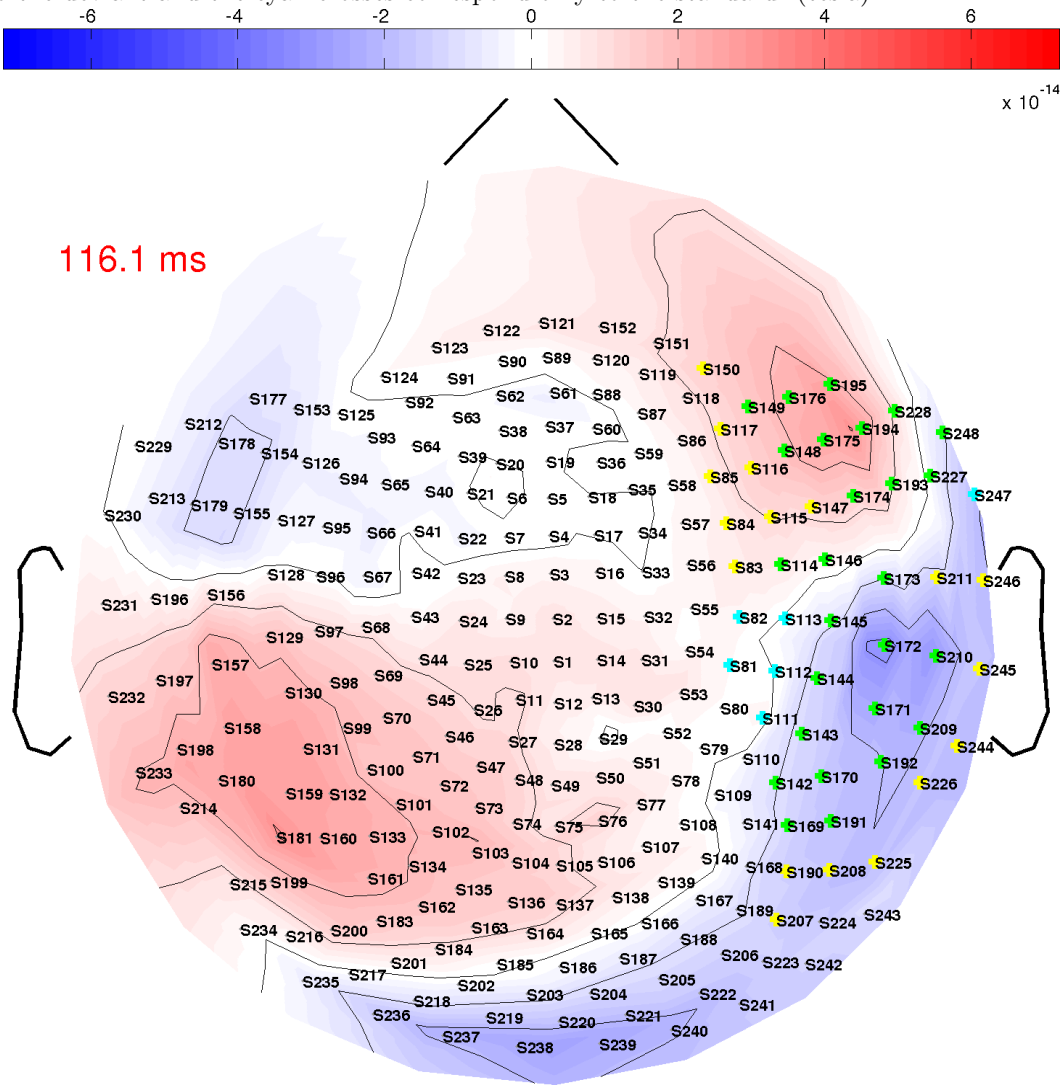


Figure 6.41: subjectEight: channel 194

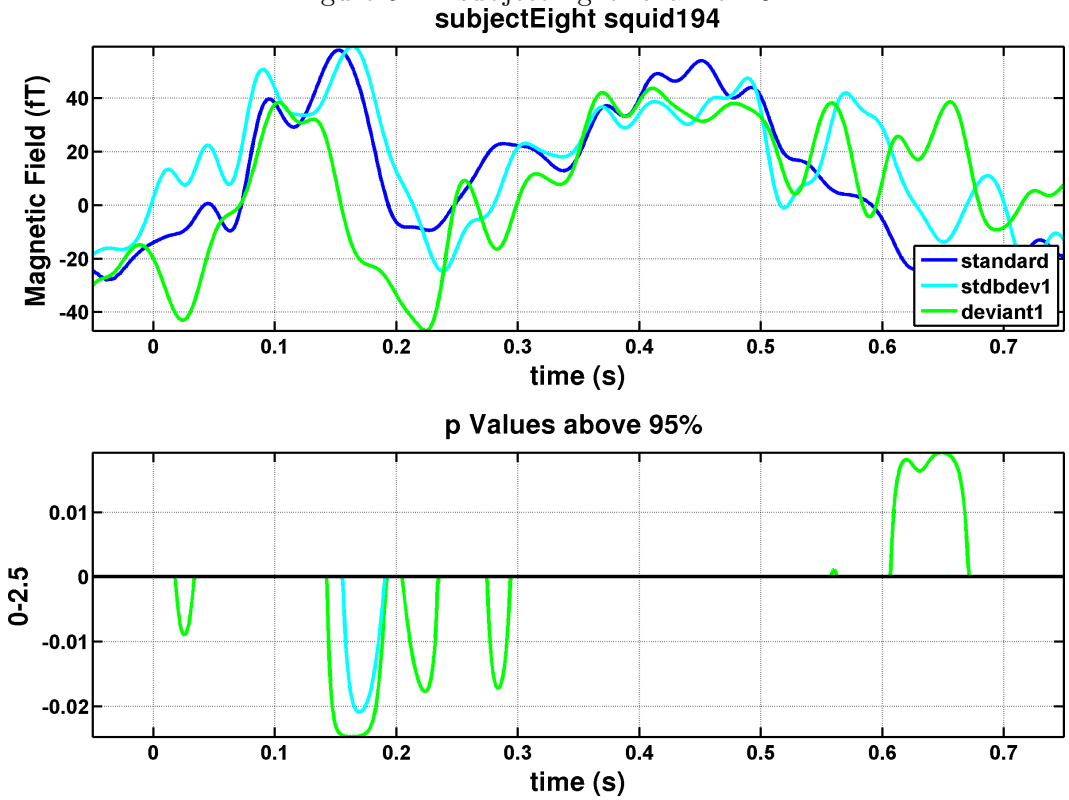
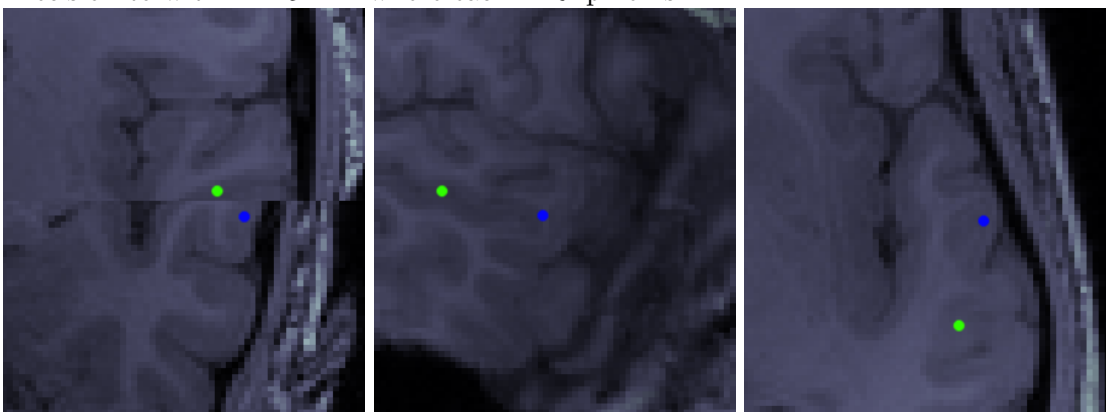


Figure 6.42: subjectEight: The blue localization corresponds to the standard response at 140-165 ms with a signal subspace correlation of 99.7%. The deviant response is the green localization at a time of 105-130 ms with a signal subspace correlation of 99.1%. Precision to within $\sim$ 5 mm where each MRI pixel is $\sim$ 1 mm.



Chapter 7

Conclusion

The motivation for this work is to understand the process by which the human auditory system can identify a sound source solely by analyzing the incoming sound energy. This is a remarkable human skill and even occurs when the incoming sound waves are degraded, as in a telephone call from a recognizable voice. In fact, when identifying speakers from phone conversations, human listeners[1] are better then typical algorithms and quite competitive with the best algorithms (Schmidt-Nielsen and Crystal, 2000). Human listeners are also more robust to degradation then are algorithms. Instead of asking individuals to actively discriminate different speakers, the mismatch response can be used as an objective measure since although being pre-attentive it reflects behavior performance (Titova & Näätänen, 2001). The ability that humans have to discriminate voices is still not matched by modern technology. One could imagine the many applications of automated voice identification technology if it could perform at the same level as humans. So how is this ability to discriminate voices or to identify sounds studied? Where in the auditory system does this happen? What are the limits of this ability?

Most often, the auditory environment consists of several sources each producing sound at the same time. The human auditory system has to segregate this mixture into separate streams and attribute them to their original sources[2] . In a review by

[1] The comparisons were made using combined group data as individual listener ability varies significantly.

[2] This is described by Bregman (Bregman, 1990) as auditory scene analysis

Griffiths and Warren, this is referred to as auditory object analysis (Griffths and Warren, 2004). They go on to develop a framework for defining and studying auditory objects, stating that object analysis involves the processing of information that corresponds to things in the sensory world where the sensory inputs about an object can be separated from other inputs and the background. Also, this segregated information can be abstracted and generalized to multiple experiences (i.e. identifying a familiar tune played on different instruments) and might be accessed across sensory modalities. In the paper, different models are reviewed pertaining to different aspects of auditory object analysis: psychophysics (Caclin et al., 2006; McAdams et al. 2004), auditory image formation in the cochlea present in the firing pattern of the auditory nerve fibers (see Griffiths and Warren, 2004 for references), and abstraction of object characteristics usually referencing activity in the primary auditory cortex or beyond (Griffiths and Warren, 2002; Näätänen et al., 1999; see Griffiths and Warren, 2004 for references)[3] . The intent of the review is to present the models as a hierarchy of operational stages, but it does give a working, yet general definition that can be incorporated into current models. Auditory objects are described as an acoustic experience that produces a two-dimensional image with time and frequency dimensions. Computations are performed on these spectrotemporal patterns appropriately separating and matching them to learned representations integral to auditory functions such as language processing, sound recognition/identification, auditory spatial analysis, streaming, and perception of melody. They propose that these computations are performed in the planum temporale, which is just posterior to the primary regions of the auditory cortex. Two basic outputs are produced from these computations: sound object information corresponding to regions anterior to primary auditory cortex and sound position information corresponding to regions posterior-dorsal to the primary auditory cortex (Griffths and Warren, 2002). The auditory cortex can be subdivided into core (primary), belt (secondary), and parabelt regions. Adjacent regions have a larger number of connections and information generally proceeds from core, to belt, to parabelt, and to more distal areas as processing becomes more complex (Kaas J H , Hackett T A, 2000; Davis and Johnsrude, 2003).

[3] McLachlan and Wilson propose a sound identification model stating that the cortical identification mechanism drives modulations of the spectrotemporal response fields of the inferior colliculus and monitors the encoding of the echoic trace in the thalamus (McLachlan and Wilson,2010)

Psychophysical timbre processing models (Caclin et al., 2006, McAdams et al. 2004) are also relevant to the study of auditory objects. These studies model timbre by rating usually 3 different perceptual dimensions: temporal attributes (attack time), spectral envelopes (spectral centroid), and spectral flux (spectral fine structure). Most studies use similarity/dissimilarity ratings and are not anatomically specific, but the Caclin study observed the mismatch negativity component to determine that the timbre dimensions partially correspond to separate anatomical locations. In her study she designed synthetic sounds that varied along the various timbre dimensions. Another study by McAdams (McAdams et al. 2004), uses a physical model of bar to design the sound files which are then rated for timbre dimension similarity to compare changes in physical properties to changes in timbre. In studies involving complex sounds, the sounds used are either natural, actual vibrating structures or the unaltered recordings of those events, or synthetic by superimposing sine waves.

The advantage of using real sounds is the authenticity of the usual acoustic experience but the disadvantage is not knowing what exactly about the sound the listener is perceiving/processing. The advantage of synthetic sounds is that they are simple and totally controllable, but the disadvantage is that the sounds don't truly reflect real-life vibrations. The compromise is to use physical models as in the McAdams study. Chapter 2 of this work demonstrates a variety of objects that have mathematical descriptions. Mathematically modeling real world vibrations is complex and can be computationally intense, but the model proposed in this work, chapter 5, is an attempt at simplification. The model is based on the fundamentals of oscillatory damping but contains inputs from actual measurements. The model could possibly be expanded to other sound sources besides those of a guitar by superimposing different sets of components with different temporal envelopes; these inputs coming from a variety of measurements. The benefit of this sound production model is that the properties of the two-dimensional image, described in the Griffiths and Warren review, can be manipulated since those are the fundamental building blocks of the object identification model. These inputs could test for fundamental features such as damping, amplitude adjustments, or frequency transposition of heavily damped components as is done here.

The uniqueness of this work is that all the tones have the same pitch and therefore should not generate any sort of higher melodic processing, therefore the only differences

being the pre-attentive discrimination of the lowest energy component transposition of the heavily damped portion of the sound. The results (Table 6.1) of this work show only the locations of maximum mismatch that correspond to stationary dipolar current sources; other sources at different times ($\sim$50-100 ms & $\gtrsim$350 ms) are not reported. Most of these mismatches are localized[4] to the superior temporal sulcus (STS) both ventral-posterior and ventral-anterior to the primary auditory cortex (which also had a few localizations). All but one of the males localized the deviant to the insular cortex. Neuroimaging evidence suggests that the insula is involved in allocating auditory attention, sound detection, entry of the sound into awareness, and in processing novel versus familiar auditory stimuli (Bamiou, Musiek, and Luxon, 2003). The results reported here are also somewhat consistent with a meta-analysis done by Arnott et al. which reviewed evidence from positron emission tomography (PET) studies and functional magnetic resonance imaging (fMRI) studies to verify the validity of the 'where' and 'what' model which states separate processing pathways exist for spatial analysis and object identification. The object pathways were predominately inferior frontal gyrus, anterior temporal lobe, and posterior temporal lobe. The spatial analysis pathways were superior frontal sulcus and inferior parietal lobule. (Arnott et al, 2004;Maeder et al., 2001;Binder et al., 2004).

The channel graphs, localization information, and counting tasks in this work suggest that although the sounds are (for most subjects) not consciously discriminable, the brain shows the pre-attentive ability to discriminate the sounds. A current model of auditory attention (Näätänen et al.,2010) explicitly illustrates the stages or aspects of central auditory processing that can be consciously experienced which includes the mismatch response. Although not many studies show mismatch responses that are not consciously perceived, there is evidence of the mismatch response elicited prior to, and therefore, predicting conscious perception (Tremblay, 1998). Since the sounds are modeled after a guitar, subjects seem to be conditioned to cue on the pitch representation which is the same for both sounds. Since context doesn't require listeners (except for some professional musicians) to cue on the intricacies of the sound source (of the same instrument), most aren't able to do so without some form of training.[5]

[4] Since the precision of the localizations are not less then .5 cm, indicates anatomical locations are more of an estimation.

[5] When I defended this work, I played a 30s audio sample with a visual aid for the audience. After

The most striking observation gained from observing the sensor array animations is the difference between genders. Many studies have shown sex differences in cognition where men excel in most visuospatial tasks and women in certain verbal tasks. It has been hypothesized that these sex differences in cognition result from a more bilateral pattern of language representation in women than in men. When adjusted for volume, women have more neurons in the auditory cortex. Sex-related differences in volume fractions of the total brain have been found in the planum temporale and Broca's area, but not in Heschl's gyrus. By studying postmortem brains, it has been concluded that language-related auditory association cortex, but not the primary hearing area, are proportionally larger in the female brain (Rademacher, J et al., 2001; Im Kiho et al., 2006).

Koelsch et al shows that an electrophysiological indicator of music-syntactic processing (ERAN, or music-syntactic MMN) is generated bilaterally in females, and with right hemispheric predominance in males. Their findings indicate that gender differences for the analysis of auditory information are not restricted to processes in the linguistic domain such as syntax, semantics, and phonology (Koelsch et al., 2003). However, a meta-analysis of mostly fMRI and PET done by Sommer et al. where 14 studies providing data on 377 men and 442 women were selected based on statistical criteria to show that no significant difference in gender language lateralization, signifying that differences are absent at the population level or that differences are only present in, as of yet undefined, listening tasks (Sommer et al., 2004).

The sample size for this work is rather small at only four, yet the differences are quite distinct in that, for females, the mismatch response is due to extra sources where for males the mismatch responses all come from nearly similar sources at relatively the same time. The observation here is not stating anything about lateralization but somewhat consistent differences in the response pattern. Gender difference reporting is also conflicting in the electrophysiological literature as well. Matsubayashi et al. report gender differences when recording the MMNm but Kasai et al. report no significant gender differences. Matsubayashi gives the possibility for this being the innate differences between MEG and EEG (Matsubayashi, J et al., 2008; Kasai, K. et al., 2002). However,

---

months of 'training', I could clearly distinguish the difference, however I wasn't 100% convinced if what was distinguishable for me would also be distinguishable for those who attended my defense. It wasn't.

Ikezawa et al.(Ikezawa et al., 2008)do report lateralization differences between genders for phonetic MMN, attributing this to how attention is deployed. Since the MMN required attention to be focused away from the phonetic stimuli this could be the reason for the observed gender differences.[6]

Female head sizes are smaller and although processing might be similar, due to volume constraints, the locations of the various populations that perform the mismatch processing might be different, explaining some of the anatomical differences: orientations due to cortical folding could be different. This is important when detecting magnetic fields since radial oriented current sources do not contribute to the measured magnetic field. In this task the subjects attention was to remain fixed on a dot. This attention devoted to this task, as mentioned in Ikezawa et al., could be the reason for the observed differences. This brings to mind the common observation that women seem to be better at multitasking where as men are more focused. Referring to the hunter-gather ideas, multitasking is useful for child rearing where focusing one's attention on a single task might be useful for hunting. It seems that the attention necessary to remain visually focused, without blinking, on dot could be the reason for the observed differences although the task did not involve language.

Language consists of syllables where consonants are partitioned by vowel sounds. Vowels sounds tend to be more sustained and consonants less so and more damped. One explanation for there being fewer vowels and more consonants is due to the frequency component richness of the heavily damped consonant sounds. This allows for more categories (i.e. variations in the 2-D spectrotemporal templates that ascend to the auditory association areas) to discriminate different heavily damped sounds. The vowels serve to encapsulate and chunk the consonant sounds. Most studies don't address the temporal information that can be discerned from the spectral information. This work suggests that the human brain is sensitive to shifts in these heavily damped frequency components. Further work should then focus on the distinction between what is lightly damped and heavily damped. This can be done by changing the single damping parameter, τ_j, (eqn. 5.10) using the sound production model discussed in the stimulus design chapter 5.

To the knowledge of the author, few (if any) studies use a sensor array animation

[6] The study references Voyer and Flight (2001) to support their claim.

to guide the selection of time samples and channel selection. Since a subject's position inside the sensor helmet varies by subject, different sensors have different positions relative to particular anatomical locations on the subjects head. This is not the case in EEG (electroencephalography) where all sensors are positioned in roughly the same scalp locations for all subjects. Without seeing the complete sensor array activation, individual channels are difficult to interpret. It is routine to view field maps of the measurement surface but not as an animation.

This work shows the ability of the human brain to discriminate subtle differences in complex sounds at a pre-attentive level. The idea that sound source identification is accomplished by cueing on heavily damped sounds is supported by the results here.

Chapter 8

References

Alho, K., et al. "Processing of Complex Sounds in the Human Auditory Cortex as Revealed by Magnetic Brain Responses." Psychophysiology 33.4 (1996): 369-75. Print.

Amenedo, E., and C. Escera. "The Accuracy of Sound Duration Representation in the Human Brain Determines the Accuracy of Behavioural Perception." The European journal of neuroscience 12.7 (2000): 2570-4. Print.

Arnott, S. R., et al. "Assessing the Auditory Dual-Pathway Model in Humans." NeuroImage 22.1 (2004): 401-8. Print.

Baillet, S., J. C. Mosher, and R. M. Leahy. "Electromagnetic Brain Mapping." Signal Processing Magazine, IEEE 18.6 (2001): 14-30. Print.

Bamiou, D. E., F. E. Musiek, and L. M. Luxon. "The Insula (Island of Reil) and its Role in Auditory Processing. Literature Review." Brain research.Brain research reviews 42.2 (2003): 143-54. Print.

Binder, Jeffrey R., et al. "Neural Correlates of Sensory and Decision Processes in Auditory Object Identification." Nature neuroscience 7.3 (2004): 295-301. Print.

Bottcher-Gandor, C., and P. Ullsperger. "Mismatch Negativity in Event-Related Potentials to Auditory Stimuli as a Function of Varying Interstimulus Interval." Psychophysiology 29.5 (1992): 546-50. Print.

Bregman, Albert S. Auditory Scene Analysis: The Perceptual Organization of Sound. Cambridge, Mass. : MIT Press, 1994. Print.

Caclin, A., et al. "Separate Neural Processing of Timbre Dimensions in Auditory Sensory Memory." Journal of cognitive neuroscience 18.12 (2006): 1959-72. Print.

Cansino, S., A. Ducorps, and R. Ragot. "Tonotopic Cortical Representation of Periodic Complex Sounds." Human brain mapping 20.2 (2003): 71-81. Print.

Clarke, John, and Alex I. Braginski. The SQUID Handbook : Volume 1: Fundamentals and Technology of SQUIDs and SQUID Systems. Wiley-VCH, 2004. Print.

—. The SQUID Handbook, Volume 2: Applications of SQUIDs and SQUID Systems. Wiley-VCH, 2006. Print.

Clarke, John. "Squids." Scientific American 271.2 (1994): 46. Print.

Cohen, D., et al. "MEG Versus EEG Localization Test using Implanted Sources in the Human Brain." Annals of Neurology 28.6 (1990): 811-7. Web.

Cowan, N. "On Short and Long Auditory Stores." Psychological bulletin 96.2 (1984): 341-70. Print.

Davis, Matthew H., and Ingrid S. Johnsrude. "Hierarchical Processing in Spoken Language Comprehension." Journal of Neuroscience 23.8 (2003): 3423-31. Print.

Duncan, C. C., et al. "Event-Related Potentials in Clinical Research: Guidelines for Eliciting, Recording, and Quantifying Mismatch Negativity, P300, and N400." Clinical

neurophysiology : official journal of the International Federation of Clinical Neurophysiology 120.11 (2009): 1883-908. Print.

Garrido, M. I., et al. "The Mismatch Negativity: A Review of Underlying Mechanisms." Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology 120.3 (2009): 453-63. Print.

Griffiths, David B. Introduction to Electrodynamics. Upper Saddle River, N.J. : Prentice Hall, c1999., 1999. Print.

Griffiths, T. D., and J. D. Warren. "What is an Auditory Object?" Nature reviews.Neuroscience 5.11 (2004): 887-92. Print.

Griffiths, Timothy D., and Jason D. Warren. "The Planum Temporale as a Computational Hub." Trends in neurosciences 25.7 (2002): 348-53. Print.

Grimm, S., et al. "Mechanisms for Detecting Auditory Temporal and Spectral Deviations Operate Over Similar Time Windows but are Divided Differently between the Two Hemispheres." NeuroImage 32.1 (2006): 275-82. Print.

Gurney, K. An Introduction to Neural Networks. UCL Press, 1 Gunpowder Square, London EC4A 3DE, UK., 1997. Print.

Hämäläinen, M. S., and J. Sarvas. "Realistic Conductivity Geometry Model of the Human Head for Interpretation of Neuromagnetic Data." IEEE transactions on biomedical engineering 36.2 (1989): 165-71. Print.

Hämäläinen. "Magnetoencephalographytheory, Instrumentation, and Applications to Noninvasive Studies of the Working Human Brain." Reviews of Modern Physics 65.2 (1993): 413. Print.

Handel, Stephen. "Timbre Perception and Object Identification." Hearing: Handbook

of Perception and Cognition. Ed. B. C. J. Moore. 2nd Edition ed. San Diego: Academic Press, 1995. 425. Print.

Hobbie, Russell K. Intermediate Physics for Medicine and Biology. 3rd ed. New York: Springer, 1997. Print.

Horvth, Jnos, Urte Roeber, and Erich Schrger. "The Utility of Brief, Spectrally Rich, Dynamic Sounds in the Passive Oddball Paradigm." Neuroscience letters 461.3 (2009): 262-5. Print.

Huang, M. X., J. C. Mosher, and R. M. Leahy. "A Sensor-Weighted Overlapping-Sphere Head Model and Exhaustive Head Model Comparison for MEG." Physics in Medicine and Biology 44.2 (1999): 423-40. Print.

Ikezawa, S., et al. "Gender Differences in Lateralization of Mismatch Negativity in Dichotic Listening Tasks." International journal of psychophysiology : official journal of the International Organization of Psychophysiology 68.1 (2008): 41-50. Print.

Im, Kiho, et al. "Gender Difference Analysis of Cortical Thickness in Healthy Young Adults with Surface-Based Methods." NeuroImage 31.1 (2006): 31-8. Print.

Jackson, John. Classical Electrodynamics. New York : Wiley, c1999., 1999. Print.

Jaramillo, M., et al. "Are Different Kinds of Acoustic Features Processed Differently for Speech and Non-Speech Sounds?" Brain research.Cognitive brain research 12.3 (2001): 459-66. Print.

Jenks, William G., Ian M. Thomas, and John P. Wikswo Jr. "SQUIDS." Encyclopedia of Applied Physics.VCH Publishers Inc., 1997. 457. Web.

Kaas, Jon H., and Troy A. Hackett. "Subdivisions of Auditory Cortex and Processing Streams in Primates." Proceedings of the National Academy of Sciences of the United

States of America 97.22 (2000): 11793-9. Print.

Kandel, Eric R. Principles of Neural Science. Eds. Eric R. Kandel, James H. Schwartz, and Thomas M. Jessell. New York: McGraw-Hill, Health Professions Division, 2000. Print.

Kane, N. M., S. R. Butler, and T. Simpson. "Coma Outcome Prediction using Event-Related Potentials: P(3) and Mismatch Negativity." Audiology & Neuro-Otology 5.3-4 (2000): 186-91. Print.

Kasai, K., et al. "No Effect of Gender on Tonal and Phonetic Mismatch Negativity in Normal Adults Assessed by a High-Resolution EEG Recording." Brain research.Cognitive brain research 13.3 (2002): 305-12. Print.

Kinsler, Lawrence E. Fundamentals of Acoustics. 4th ed. ed. New York: John Wiley & Sons, Inc., 2000. Print.

Koelsch, S., and A. D. Friederici. "Toward the Neural Basis of Processing Structure in Music. Comparative Results of Different Neurophysiological Investigation Methods." Annals of the New York Academy of Sciences 999.1 (2003): 15-28. Print.

Kudo, N., et al. "Comparison between Mismatch Negativity Amplitude and Magnetic Mismatch Field Strength in Normal Adults." Biological psychology 71.1 (2006): 54-62. Print.

Kujala, T., M. Tervaniemi, and E. Schroger. "The Mismatch Negativity in Cognitive and Clinical Neuroscience: Theoretical and Methodological Considerations." Biological psychology 74.1 (2007): 1-19. Print.

Leahy, R. M., et al. "A Study of Dipole Localization Accuracy for MEG and EEG using a Human Skull Phantom." Electroencephalography and clinical neurophysiology 107.2 (1998): 159-73. Print.

Liu, A. K., A. M. Dale, and J. W. Belliveau. "Monte Carlo Simulation Studies of EEG and MEG Localization Accuracy." Human brain mapping 16.1 (2002): 47-62. Web. 9/21/2010.

Luck, Steven J., and Stephen J. Luck. An Introduction to the Event-Related Potential Technique. Cambridge, MA : MIT Press, c2005., 2005. Print.

Maeder, Philippe P., et al. "Distinct Pathways Involved in Sound Recognition and Localization: A Human fMRI Study." NeuroImage 14.4 (2001): 802-16. Print.

Matsubayashi, Junko, et al. "The Influence of Gender and Personality Traits on Individual Difference in Auditory Mismatch: A Magnetoencephalographic (MMNm) Study." Brain research 1236 (2008): 159-65. Print.

McAdams, Stephen, Antoine Chaigne, and Vincent Roussarie. "The Psychomechanics of Simulated Sound Sources: Material Properties of Impacted Bars." The Journal of the Acoustical Society of America 115.3 (2004): 1306-20. Print.

McLachlan, N., and S. Wilson. "The Central Role of Recognition in Auditory Perception: A Neurobiological Model." Psychological review 117.1 (2010): 175-96. Print.

Mosher, J. C., and R. M. Leahy. "Recursive MUSIC: A Framework for EEG and MEG Source Localization." Biomedical Engineering, IEEE Transactions on 45.11 (1998): 1342-54. Print.

—. "Source Localization using Recursively Applied and Projected (RAP) MUSIC." Signal Processing, IEEE Transactions on 47.2 (1999): 332-40. Print.

Näätänen, R., T. Kujala, and I. Winkler. "Auditory Processing that Leads to Conscious Perception: A Unique Window to Central Auditory Processing Opened by the Mismatch Negativity and Related Responses." Psychophysiology (2010).

Näätänen, R., A. W. Gaillard, and S. Mantysalo. "Early Selective-Attention Effect on Evoked Potential Reinterpreted." Acta Psychologica 42.4 (1978): 313-29. Print.

Näätänen, R., et al. "The Mismatch Negativity (MMN) in Basic Research of Central Auditory Processing: A Review." Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology 118.12 (2007): 2544-90. Print.

Näätänen, R., et al. "The Mismatch Negativity (MMN): Towards the Optimal Paradigm." Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology 115.1 (2004): 140-4. Print.

Näätänen, R., and I. Winkler. "The Concept of Auditory Stimulus Representation in Cognitive Neuroscience." Psychological bulletin 125.6 (1999): 826-59. Print.

Näätänen, R. "The Perception of Speech Sounds by the Human Brain as Reflected by the Mismatch Negativity (MMN) and its Magnetic Equivalent (MMNm)." Psychophysiology 38.1 (2001): 1. Print.

Nadeau, Stephen, and Patrick J. Antonelli. Medical Neuroscience. Philadelphia : Saunders, c2004., 2004. Print.

Nunez, P. L., and R. B. Silberstein. "On the Relationship of Synaptic Activity to Macroscopic Measurements: Does Co-Registration of EEG with fMRI make Sense?" Brain topography 13.2 (2000): 79-96. Print.

Nunez, P. L., and R. Srinivasan. Electric Fields of the Brain: The Neurophysics of EEG. New York: Oxford University Press, 2006. Print.

Patel, S. H., and P. N. Azzam. "Characterization of N200 and P300: Selected Studies of the Event-Related Potential." International journal of medical sciences 2.4 (2005): 147-54. Print.

Pizzella, Vittorio, et al. "SQUID Systems for Biomagnetic Imaging." Superconductor Science and Technology 14.7 (2001): R79-R114. Print.

Psychoacoustical Terminology. Timbre. ANSI S3.20-1973. New York:, 1973. Print.

Rademacher, J., et al. "Human Primary Auditory Cortex in Women and Men." Neuroreport 12.8 (2001): 1561-5. Print.

Sarvas, J. "Basic Mathematical and Electromagnetic Concepts of the Biomagnetic Inverse Problem." Physics in Medicine and Biology 32.1 (1987): 11-22. Print.

Sabri, Merav, and Kenneth B. Campbell. "Effects of Sequential and Temporal Probability of Deviant Occurrence on Mismatch Negativity." Cognitive Brain Research 12.1 (2001): 171-80. Web.

Scheler, G., et al. "Spatial Relationship of Source Localizations in Patients with Focal Epilepsy: Comparison of MEG and EEG with a Three Spherical Shells and a Boundary Element Volume Conductor Model." Human brain mapping 28.4 (2007): 315-22. Print.

Schmidt-Nielsen, Astrid, and Thomas H. Crystal. "Speaker Verification by Human Listeners: Experiments Comparing Human and Machine Performance using the NIST 1998 Speaker Evaluation Data." Digital Signal Processing 10.1-3 (2000): 249-66. Print.

Sommer, I. E., et al. "Do Women really have More Bilateral Language Representation than Men? A Meta-Analysis of Functional Imaging Studies." Brain : a journal of neurology 127.Pt 8 (2004): 1845-52. Print.

Tarkiainen, A., et al. "The 3D Topography of MEG Source Localization Accuracy: Effects of Conductor Model and Noise." Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology 114.10 (2003): 1977-92. Print.

Tervaniemi, M., et al. "The Musical Brain: Brain Waves Reveal the Neurophysiological Basis of Musicality in Human Subjects." Neuroscience letters 226.1 (1997): 1-4. Print.

Tervaniemi, M., I. Winkler, and R. Näätänen. "Pre-Attentive Categorization of Sounds by Timbre as Revealed by Event-Related Potentials." Neuroreport 8.11 (1997): 2571-4. Print.

Tervaniemi, M. "Musical Sound Processing in the Human Brain. Evidence from Electric and Magnetic Recordings." Annals of the New York Academy of Sciences 930 (2001): 259-72. Print.

Tiitinen, H., et al. "Attentive Novelty Detection in Humans is Governed by Pre-Attentive Sensory Memory." Nature 372.6501 (1994): 90-2. Print.

Tinkham, Michael. Introduction to Superconductivity : Second Edition (Dover Books on Physics). Dover Publications, 2004. Print.

Titova, N., and R. Näätänen. "Preattentive Voice Discrimination by the Human Brain as Indexed by the Mismatch Negativity." Neuroscience letters 308.1 (2001): 63-5. Print.

Tremblay, K., N. Kraus, and T. McGee. "The Time Course of Auditory Perceptual Learning: Neurophysiological Changes during Speech-Sound Training." Neuroreport 9.16 (1998): 3557-60.

Volegov, P., et al. "Noise-Free Magnetoencephalography Recordings of Brain Function." Physics in Medicine and Biology 49.10 (2004): 2117-28. Print.

Voyer, D., and J. Flight. "Gender Differences in Laterality on a Dichotic Task: The Influence of Report Strategies." Cortex; a journal devoted to the study of the nervous system and behavior 37.3 (2001): 345-62. Print.

Wikswo, J. P., Jr. "Biomagnetic Sources and their Models." Advances in Biomagnetism. Ed. S. J. Williamson. New York: Plenum Press, 1989. 1. Print.

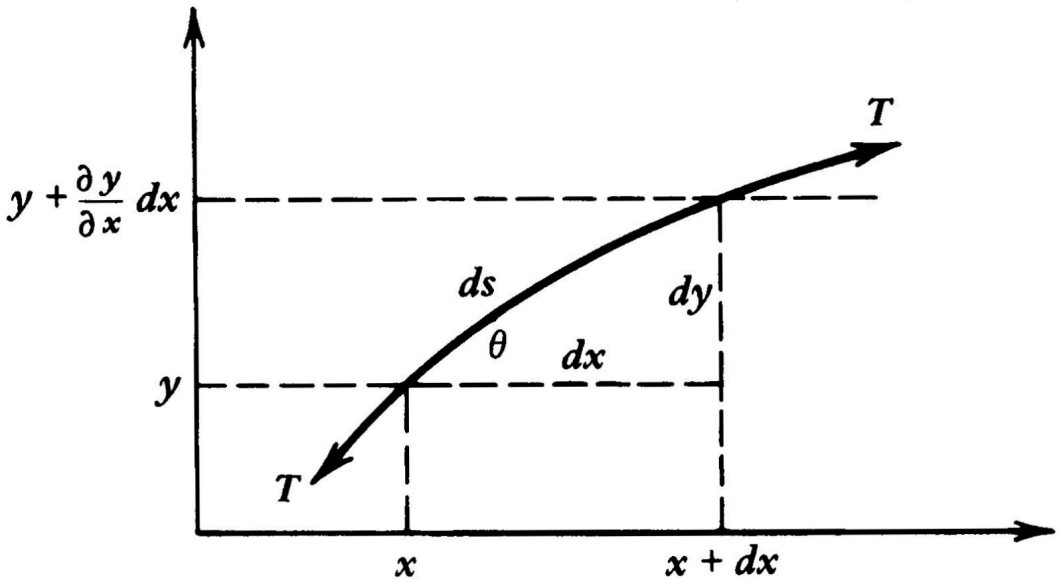
Winkler, I., K. Reinikainen, and R. Näätänen. "Event-Related Brain Potentials Reflect Traces of Echoic Memory in Humans." Perception & psychophysics 53.4 (1993): 443-9. Print.

Appendix A

Vibrating Strings

When analyzing a string (Fig. A.1), the assumptions are that the stiffness is minimal and the density is uniform over the length of the string. The string is stretched so that it is under tension with the ends fixed and the displacements are not to large. For a string, the string segment displacements (y direction) are perpendicular to the direction of propagation (x direction). This defines a transverse wave. The net force in the vertical direction on a segment of string, ds, under tension, T, is

Figure A.1: Force diagram of string segment. (Kinsler, 2000)



$$dF_{net_y} = T\sin(\theta)_{x+dx} - T\sin(\theta)_x \tag{A.1}$$

and using Taylor's expansion

$$f(x+dx) = f(x) + \frac{\partial f(x)}{\partial x}dx \tag{A.2}$$

for small y displacements $\tan(\theta) \approx \sin(\theta)$

$$dF_{net_y} = T\sin(\theta) + \frac{\partial(T\sin(\theta))}{\partial x}dx - T\sin(\theta) = T\frac{\partial(\tan(\theta))}{\partial x}dx \tag{A.3}$$

$$dF_{net_y} = T\frac{\partial(\frac{\partial y}{\partial x})}{\partial x}dx \quad \Rightarrow \quad \tan(\theta) \approx \frac{\partial y}{\partial x} \tag{A.4}$$

$$\rho dx\frac{\partial^2 y}{\partial t^2} = T\frac{\partial^2 y}{\partial^2 x}dx \quad \Rightarrow \quad dF = dm\frac{\partial^2 y}{\partial t^2} \tag{A.5}$$

where ρ is the density (mass/length).

$$\frac{1}{v^2}\frac{\partial^2 y}{\partial t^2} = \frac{\partial^2 y}{\partial^2 x} \Rightarrow v = \sqrt{\frac{T}{\rho}} \tag{A.6}$$

which is the wave equation from section 2.1 but now a defined phase velocity, v. In order to solve this equation, the problem, with a string of length, L, must be constrained by some set of boundary conditions (i.e what happens to $y(0,t)$;$y(L,t)$ and its derivative) and its initial conditions (i.e. what happens to $y(x,0)$ and its derivative).

One way to solve this equation is to write the function as the product of functions each of a different independent variable.

$$y(x,t) = f(x)g(t) \tag{A.7}$$

this assumes the function y is separable. Rewriting the wave equation so that the independent variables are on their own sides of the equation is

$$\frac{1}{v^2}\frac{\ddot{g}}{g} = \frac{f''}{f} \Rightarrow -k^2 \tag{A.8}$$

here the prime, $'$, represents a derivative with respect to the spatial coordinate. Since the spatial terms, right side of the equation, have no time dependence, their ratio must be constant for all time. Since that constant is equal to the ratio of the time

dependent terms, their ratio also must be constant in time. The minus sign is necessary for oscillatory motion. If the sign of a function and its concavity (i.e., second derivative) are the same the function will grow without bound which is not wave motion. If they are different then the function after changing sign will begin changing slope to revert back to the equilibrium position; natural oscillatory behavior can be described by harmonic functions with complex coefficients A_i.

$$f'' = -k^2 f \Rightarrow f(x) = A_i e^{\pm ikx} \tag{A.9}$$

Applying the boundary conditions $f(0) = 0$ and $f(L) = 0$ for a general function of the form

$$\begin{aligned} f(x) &= A_0 e^{ikx} + A_1 e^{-ikx} && \text{(A.10)} \\ f(0) &= A_0 + A_1 = 0 \Rightarrow A_0 = -A_1 && \text{(A.11)} \\ f(L) &= A_0(e^{ikL} - e^{-ikL}) = A_0(2i)\sin(kL) = 0 && \text{(A.12)} \\ & && \text{(A.13)} \end{aligned}$$

Since A_0 is complex and not zero, the argument of the sine function must be for an integer, n.

$$kL = n\pi \tag{A.14}$$

At this point, k, is just a constant that facilitated the separation of variables. Looking at the time dependent part of the wave equation

$$\ddot{g} = -(kv)^2 g \tag{A.15}$$

and using the definition of ω_0 from equation 2.7[1] to rewrite equation 2.3 is $\ddot{x} = -\omega_0^2 x$ which has the same form leading to an expression for k, the wave number, which is related to the wavelength λ.

$$k = \frac{\omega}{v} \tag{A.16}$$

$$k = \frac{2\pi}{\lambda} \tag{A.17}$$

[1] the k in equations 2.3 and 2.7 is the elasticity constant, not the k constant from the wave equation ... check their units.

Equation A.16 is the relationship that defines phase speed, v, which for this wave equation (eqn. A.6) is constant. The solution to the time dependent equation is of the form of harmonic functions, with complex coefficients B_i, and is satisfied by the initial conditions which in this case is not restrictive, so the final wave solution is

$$g(t) = B_0 e^{i\omega t} + B_1 e^{-i\omega t} \tag{A.18}$$

$$y(x,t) = A_0(e^{ikx} - e^{-ikx})(B_0 e^{i\omega t} + B_1 e^{-i\omega t}) \tag{A.19}$$

$$y(x,t) = \sin(kx)(c_0 e^{i\omega t} + c_1 e^{-i\omega t}) \tag{A.20}$$

where the real part of the $y(x,t)$ is the physical oscillation and c_i are also complex constants. The constraint of equation A.14 with the definition of k, equation A.17, constrains the allowed frequencies to those that satisfy

$$k_n = \frac{\pi n}{L} \qquad n = 1, 2, 3, \ldots \tag{A.21}$$

$$\lambda_n = \frac{2L}{n} \tag{A.22}$$

$$\omega_n = \frac{\pi n v}{L} \tag{A.23}$$

Notice that $y(x,t) = 0$ whenever $\sin(\frac{2\pi}{\lambda_n}x) = 0$ regardless of t. When $n = 1$ this is the boundary condition, but when $n > 1$, $y(x,t) = 0$ occurs at other points between the boundaries in the middle of the string. These positions are called nodes which are a segments where the string does not oscillate. A quarter of a wavelength away are positions called antinodes which are positions where the oscillation amplitude is maximum. The larger n is the more nodes and antinodes.

The complete general solution for a vibrating string fixed at both ends is

$$y(x,t) = \sum_{n=1}^{\infty} \sin(k_n x)(c_{0n} e^{i\omega_n t} + c_{1n} e^{-i\omega_n t}) \tag{A.24}$$

Since each mode, n, is independent of any other mode, the solution is a linear superposition of all modes. The coefficients c_{in} are determined by the initial conditions in how the string is excited.

To conclude this section on vibrating strings, we present two solutions of the general equation corresponding to two different ways to excite a stretched string. The first is a string plucked at some point $x_p : 0 \leftrightarrow L$ a distance d off the x-axis and then released

from rest at $t = 0$. The second will be a string struck at some point $x_p : 0 \leftrightarrow L$ transferring momentum when it is not displaced. Both of these represent methods for exciting stringed instruments such as the guitar and the piano.

For the first solution, equation A.24 with the velocity initial condition that $\dot{y}(x,0) = 0$ and therefore $c_{0n} = c_{1n} = c_n$ reduces to

$$y(x,t) = \sum_{n=1}^{\infty} \sin(k_n x)(2c_n)\cos(\omega_n t) \tag{A.25}$$

when $\alpha \equiv \frac{d}{L}$ and $\gamma \equiv \frac{x_p}{L}$ and defining the displacement initial condition the Fourier coefficient (equation 2.54 $t \to x$ and $T \to L$) is

$$\begin{aligned} y(x,0) &= \frac{\alpha}{\gamma}x \qquad 0 < x < \gamma & \text{(A.26)} \\ &= \frac{\alpha}{(1-\gamma)}(1-x) \qquad \gamma < x < 1 & \text{(A.27)} \\ 2c_n &= \frac{2}{L}\int_0^L y(x,0)\sin(k_n x)dx & \text{(A.28)} \\ &= -\frac{2\alpha\sin(n\pi\gamma)}{n^2\pi^2(\gamma-1)\gamma} & \text{(A.29)} \end{aligned}$$

The final expression for the vibration of a plucked string is then

$$y(x,t) = \sum_{n=1}^{\infty} \sin(k_n x)(-\frac{2\alpha\sin(n\pi\gamma)}{n^2\pi^2(\gamma-1)\gamma})\cos(\omega_n t)D(n,t) \tag{A.30}$$

I've included $D(n,t)$ to represent a damping function in the form of $e^{-\frac{t}{\tau_n}}$ which is a reference to equation 2.21. This form of the equation won't satisfy the wave equation but then the wave equation derivation did not account for damping. One could include a damping term in the wave equation proportional to $\frac{\partial y}{\partial t}$. The simulated results [2] depending on the actual form of τ_n are reasonable.

For the second solution, equation A.24 with the displacement initial condition that $y(x,0) = 0$ and therefore $c_{0n} = -c_{1n} = c_n$ reduces to

$$y(x,t) = \sum_{n=1}^{\infty} \sin(k_n x)(\frac{2ic_n}{\omega_n})\sin(\omega_n t) \tag{A.31}$$

[2] Mathematica (.nba) animations available on request.

Defining the velocity initial condition with a delta function, where v_0 is the vertical velocity of the string segment at x_p, the Fourier coefficient (equation 2.54 $t \to x$ and $T \to L$) is

$$\dot{y}(x,0) = \delta(x-\gamma L)v_0\frac{L}{2} \tag{A.32}$$

$$2ic_n = \frac{2}{L}\int_0^L \dot{y}(x,0)\sin(k_n x)dx \tag{A.33}$$

$$\tag{A.34}$$

Using the delta function definition for a Fourier series, the velocity coefficients are

$$\delta(x-\gamma L) = \frac{2}{L}\sum_{m=1}^{\infty}\sin(k_m x)\sin(m\pi\gamma) \tag{A.35}$$

$$2ic_n = \frac{2v_0}{L}\sum_{m=1}^{\infty}\sin(m\pi\gamma)\int_0^L \sin(k_m x)\sin(k_n x)dx \tag{A.36}$$

$$= \frac{2v_0}{L}\sum_{m=1}^{\infty}\sin(m\pi\gamma)\delta_{mn}\frac{L}{2} \tag{A.37}$$

$$= v_0\sin(n\pi\gamma) \tag{A.38}$$

$$\tag{A.39}$$

The final expression for the vibration of a struck string is then

$$y(x,t) = \sum_{n=1}^{\infty}\sin(k_n x)\left(\frac{v_0\sin(n\pi\gamma)}{\omega_n}\right)\sin(\omega_n t)D(n,t) \tag{A.40}$$

again with the damping function $D(n,t)$. Both of these solutions have a term $\sin(n\pi\gamma)$ which is shows that the position of where the string is struck or plucked will determine which Fourier components are excited. If the damping term is unity (i.e no damping), the waveform will not change its shape each cycle. For a more realistic solution, the damping constant, τ_n, in the function $D(n,t)$ should become smaller as $n \to \infty$. This matches most observations in that the energy is sustained in the lower frequency components longer then in the higher frequency components. These observations are all assuming that the damping is light (i.e. underdamping).

Some more realistic boundaries conditions can be understood by representing the boundary in terms of its mechanical impedance. For instance suppose at $x = L$, the support to maintain that condition gives a little to the motion of the string. This will modify the wave behavior. If the mechanical impedance of the support is known, Z_{m0} then the velocity at $x = 0$ is

$$y(0,t) = \frac{1}{i\omega Z_{m0}} T \left(\frac{\partial y}{\partial x}\right)_{x=0} \Rightarrow \text{(eqn.A.3)} \tag{A.41}$$

$$u(0,t) = \frac{1}{Z_{m0}} T \left(\frac{\partial y}{\partial x}\right)_{x=0} \tag{A.42}$$

If the mechanical impedance at the support is very large, then the boundary conditions are $y(0,t) \approx 0$ and $u(0,t) \approx 0$, the condition for a fixed end. If the support does not apply any force to the string then the impedance is zero and the boundary condition must be $\left(\frac{\partial y}{\partial x}\right)_{x=0} \approx 0$ to maintain a physical value for y and u. This is the condition for a free end. Impedance is critical when solving many types of boundary conditions as later sections will attest.

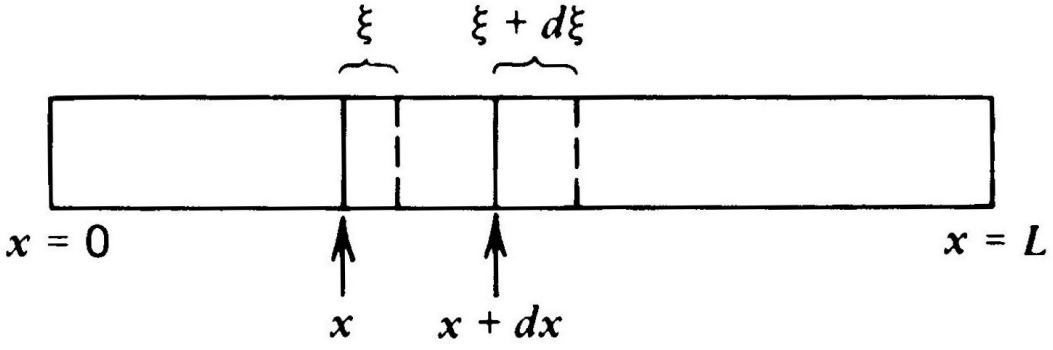
These are two solutions for vibrations common to sound production in stringed instruments. A violin is most often bowed; how does that fit into the description so far? When a bow passes over a string, friction between the bow is enough to move the string a certain distance. When the distance is large enough so that the restoring elasticity overwhelms the friction it 'snaps' back. Friction is still being applied but this time reduces the recoil so that the string can 'catch' the bow again and resume its steady displacement from equilibrium. The pattern is that of a sawtooth. The sawtooth vibration maintains itself to some degree even when there is no bowing. Friction is applied to varying degrees and direction throughout the cycle to set up this vibration. This complex vibration is the superposition of excited Fourier components. Using a method similar to the previous two methods, one could calculate the Fourier components for a sawtooth waveform. Are there any other simple objects that shed insight on to how more complex objects vibrate?

Appendix B

Vibrating Bars

The previous section discusses the vibration of strings under the conditions that stiffness is minimal. From experience, metal strings and nylon strings sound different even when playing the same note or fundamental tone (i.e. $n = 1$ Fourier component). One can imagine a setup where the tension is calibrated with the density so that almost identical vibrations occur for both types of strings. The components through out the vibrations should be identical for both strings. So what accounts for the perceived difference in sound? This section will look at a vibration of a bar where the only restoring force is that due to tensile, compressive, and shear stresses.

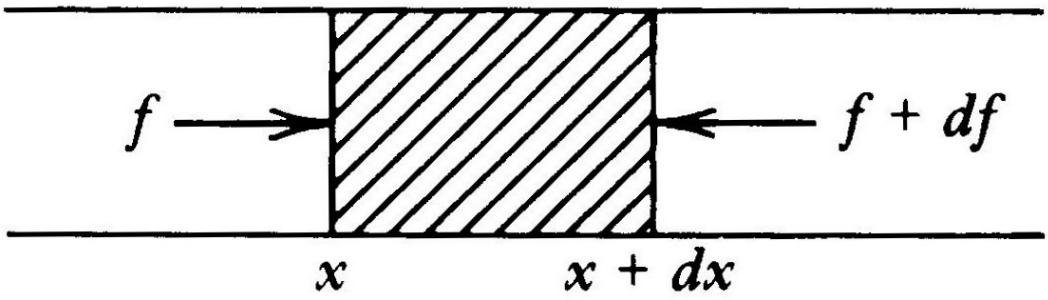
Figure B.1: Bar of length L defining strain. (Kinsler, 2000)



When a long thin bar (Fig. B.1) of uniform cross-sectional area, A, is subjected to a longitudinal force, the bar will deform by either stretching or compressing. Strain is a measure of the degree of deformation. Stress is the external force acting on a mass

per unit area. Just like the vertical displacements of a string, for sufficiently small stresses, strain is proportional to stress where the proportionality constant is referred to as the elastic modulus. Since the stress is applied in the same direction as the resulting displacement, any wavelike stresses will produce longitudinal waves, waves that displace in the direction of propagation. This is in contrast to the waves on a string which are transverse.

Figure B.2: Bar segment under compressive force. (Kinsler, 2000)



For a long thin bar, the coordinate of deformation, $\xi(x,t)$, is assumed to only depend on the longitudinal direction and time. Consider the deformation on an element of length dx with its cross-sectional area A at x (Fig. B.2). The application of a force in a particular direction when balanced by on equal opposing force at the other end of the solid causes a displacement of the plane at x a distance ξ. Correspondingly, the force also causes the displacement of the plane at $x + dx$ a distance $\xi + d\xi$. Here the assumption is that only small deformations that do not permanently deform the solid occur and, thus, can be approximated as linear using a Taylor series expansion. The deformation displacement, ξ, at $x + dx$ is then

$$\xi(x + dx) = \xi(x) + (\frac{\partial \xi}{\partial x})dx \tag{B.1}$$

The net deformation of the element is then

$$d\xi = (\frac{\partial \xi}{\partial x})dx \tag{B.2}$$

The quantitative description for the strain, ϵ, is defined as the ratio of the change of the length to the original length. When the element is stretched the strain is positive and negative when the element is compressed.

$$\epsilon \equiv \frac{(\partial \xi/\partial x)}{dx}dx = \frac{\partial \xi}{\partial x} \tag{B.3}$$

Stress, σ, is defined as the ratio of the internal force, F_x, to the cross-sectional area. This internal force can be understood as the restoring force exerted by the solid to maintain its shape.

$$\sigma \equiv \frac{F_x}{A} \tag{B.4}$$

The proportionality relationship between stress and strain is then

$$\sigma = -E\epsilon \tag{B.5}$$

A close resemblance to equation 2.3. Here the elastics modulus, E, is named Young's Modulus which is specific to the longitudinal nature of the stress to the strain. The minus sign is included because E is always positive and the elastic stresses are always restoring to equilibrium. The sign convention for stresses is then a positive stress causes a compression (negative strain) and a negative stress causes a tension (positive strain). For wave motion to occur, F_x varies in time and space and to first approximation the assumption is that the force variations are small and therefore linear. The force acting at point x and at point $x + dx$ is

$$F_x = -EA\frac{\partial \xi}{\partial x} \tag{B.6}$$

$$F_x(x + dx) = -F_x - (\frac{\partial F_x}{\partial x})dx \tag{B.7}$$

The force at $x + dx$ on dx is acting to the left or in the negative x direction. The net force to the right is then

$$dF_x = -(\frac{\partial F_x}{\partial x})dx \tag{B.8}$$

$$= EA\frac{\partial^2 \xi}{\partial x^2}dx \tag{B.9}$$

The volume of the element dx equal to Adx and the mass equal to ρAdx, where ρ is the volume density (mass/volume) of the bar. Applying Newton's 2$^{\text{nd}}$ Law to replace dF_x with $dm\ddot{\xi}$

$$\rho Adx\frac{\partial^2 \xi}{\partial t^2} = EA\frac{\partial^2 \xi}{\partial x^2}dx \tag{B.10}$$

$$\frac{1}{v^2}\frac{\partial^2 \xi}{\partial t^2} = \frac{\partial^2 \xi}{\partial x^2} \Rightarrow v^2 = \frac{E}{\rho} \tag{B.11}$$

yields the familiar wave equation with phase speed, v, for a linear elastic solid.

Just like for the string, the wave equation can be solved for a specific set of boundary and initial conditions. For a bar of length L, three basic sets of boundary conditions are first, when both ends of the bar a fixed, second, both ends of the bar are free, and, finally, one end is fixed on end is free. The approach at this point is the same as that of the solution of the string. Expressing the wave equation as separable in terms of time in space as in equations A.7 and A.8 leads to the general solution for the spacial coordinate similar to equation A.10.

$$\begin{aligned} \xi(x,t) &= f(x)g(t) \Rightarrow g(t) = B_0 e^{i\omega t} + B_1 e^{-i\omega t} && \text{(B.12)} \\ f(x) &= A_0 e^{ikx} + A_1 e^{-ikx} && \text{(B.13)} \end{aligned}$$

The first set of boundary conditions are the same as that of the string, fixed at both ends. The boundary condition that $f(0) = 0$ leads to $A_0 = -A_1 = A$.

$$f(x) = A(e^{ikx} - e^{-ikx}) = A(2i)\sin(kx) \qquad \text{(B.14)}$$

and the boundary condition that $f(L) = 0$ leads to

$$\begin{aligned} f(L) &= A(2i)\sin(kL) = 0 && \text{(B.15)} \\ kL &= n\pi \qquad n = 1, 2, 3, \ldots && \text{(B.16)} \end{aligned}$$

which is the same condition as that for the string. Leading to $\omega_n = \frac{\pi n v}{L}$. The second set of boundary conditions where the bar is free at both ends is understood as there being no force at positions $x = 0$ and $x = L$. Since force is force is proportional to $\frac{\partial \xi}{\partial x}$, this leads to $f'(0) = 0$ and $f'(L) = 0$ where the $\frac{\partial f}{\partial x} \Rightarrow f'(x)$.

$$\begin{aligned} f'(x) &= ik(A_0 e^{ikx} - A_1 e^{-ikx}) && \text{(B.17)} \\ f'(0) &= ik(A_0 - A_1) = 0 \Rightarrow A_0 = A_1 = A && \text{(B.18)} \\ f'(L) &= ikA(2i)\sin(kL) = 0 && \text{(B.19)} \\ kL &= n\pi \qquad n = 1, 2, 3, \ldots && \text{(B.20)} \end{aligned}$$

This is the same condition as that for a bar fixed at both ends meaning that the frequencies for both boundary conditions are the same. However what is not the same is the wave pattern as can be seen by looking at the positions of zero displacement or

nodal points.

$$f(x) = C_n \sin(k_n x) \quad \Rightarrow \quad \textbf{Fixed} - \textbf{Fixed} \tag{B.21}$$

$$f(x) = C_n \cos(k_n x) \quad \Rightarrow \quad \textbf{Free} - \textbf{Free} \tag{B.22}$$

The nodal points and anti-nodal points are shifted by $\frac{\lambda}{4}$, respectively, even though the frequencies of vibrations are the same. And finally for the Fixed-Free condition, the boundary conditions are $f(0) = 0$ and $f'(L) = 0$.

$$f(0) = A_0 + A_1 = 0 \Rightarrow A_0 = -A_1 = A \tag{B.23}$$

$$f'(0) = (2i)kA\cos(kL) = 0 \Rightarrow kL = \frac{\pi}{2}(2n-1) \qquad n = 1, 2, 3, \ldots \tag{B.24}$$

This leads to frequencies that are not the same as those for the previous two conditions.

$$\omega_n = n\frac{\pi v}{L} \qquad \textbf{Free} - \textbf{Free}/\textbf{Fixed} - \textbf{Fixed} \tag{B.25}$$

$$\omega_n = (n - \frac{1}{2})\frac{\pi v}{L} \qquad \textbf{Fixed} - \textbf{Free} \tag{B.26}$$

$$n = 1, 2, 3, \ldots \tag{B.27}$$

The spacing between each set of frequency components is the same but the first frequency ($n = 1$) or the fundamental for the Fixed-Free bar is half that of a Free-Free/Fixed-Fixed bar. Finally a general solution would be of the form,

$$\xi(x,t) = \sum_{n=1}^{\infty} f(x, C_n, k_n)(B_0 e^{i\omega_n t} + B_1 e^{-i\omega_n t}) \tag{B.28}$$

where B_i is determined by initial conditions.

At this point, one can see how different boundary conditions on the same solid can have different effects. Four simple sound production systems have been analyzed. All four have had similar spectral properties in that the frequency spacing between each successive component is the same. When the ratio between a set of frequencies and fundamental ($n = 1$) is an integer then the components are specially designated as harmonics. In this context, the Fixed-Fixed conditions contains all its harmonics where as the Fixed-Free only contains the odd harmonics. So this might beg the question, are all sounds harmonically related? As the remaining sections will show, the quick answer is, no, but the reader should now begin to have an intuition for how simple

sound production systems will have different frequency distributions that are based on the shape (i.e boundary conditions) and the physical properties of the system (i.e. densities, elasticities, damping properties, etc.). Going back to the bar, this next section will show how a slight modification in the setup leads to not so 'elegantly' [1] related frequency components.

Consider a bar of length L with mass m concentrated at its end at $x = L$ and the other end $x = 0$ free. This leads to the condition that $\frac{\partial \xi}{\partial x} = 0$ at $x = 0$ reducing the spatial term to a cosine function. The time dependent term, $g(t)$, is that of equation B.12 with the properties of equation A.15.

$$\xi(x,t) = C\cos(kx)g(t) \tag{B.29}$$

The boundary condition at $x = L$ is understood in the context of Newton's 2nd Law. For all time, a positive internal force F_x compresses the bar but at the right end this results in an acceleration of the mass to the right.

$$F_x(L,t) = m\left(\frac{\partial^2 \xi}{\partial t^2}\right)_{x=L} \tag{B.30}$$

Keeping in mind the Young's modulus is $E = \rho v^2$ (eqn. B.11) where ρ is the volume density and $\omega = kv$ the boundary condition reduces to

$$F_x(L,t) = -EA\frac{\partial \xi}{\partial x} = (EAk)C\sin(kL)g(t) \tag{B.31}$$

$$m\left(\frac{\partial^2 \xi}{\partial t^2}\right)_{x=L} = -C\cos(kL)(kv)^2 g(t) \tag{B.32}$$

$$\frac{\tan(kL)}{k} = -\frac{mv^2}{EA} \tag{B.33}$$

$$\frac{\tan(kL)}{kL} = -\frac{m}{m_b} \Rightarrow m_b = \rho AL \tag{B.34}$$

m_b is the total mass of the bar. Although the analysis isn't complete, it wouldn't hurt to stop and analyze the relationship to this point. Look at the added mass and see what becomes of the condition on k. As $m \to \infty$ [2] should restrict the movement at $x = L$ so that it approaches that of the Free-Fixed boundary condition. Under this condition $\tan(kL) \to \infty$ when $\cos(kL) \to 0$ which is the condition for the Free-Fixed case (eqn.

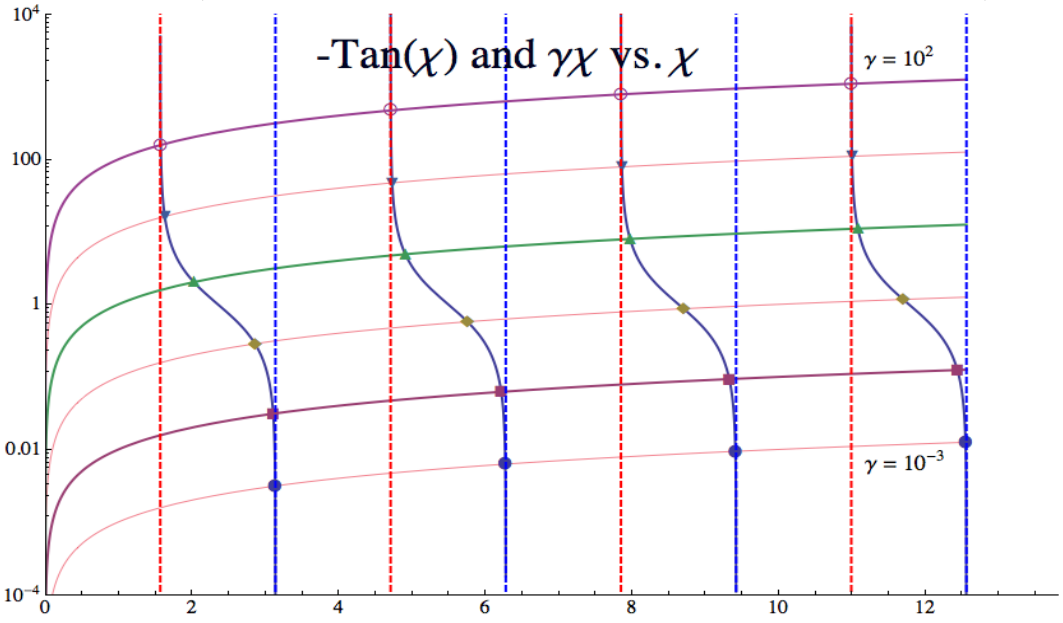
[1] depends on your perspective and/or maybe your profession. Harmony is good for selling records...
[2] masses can't be infinite but the idea is a large mass relative to the bar

B.24). As $m \to 0$ as in the $\tan(kL) \to 0$ as $\sin(kL) \to 0$, the condition for the Free-Free bar (eqn. B.19). When solving for k, the equation is a transcendental equation and can only be solved numerically. First set $kL \to \chi$ and $\frac{m}{m_b} \to \gamma$ so the equation is now dimensionless.

$$-\tan(\chi) = \gamma\chi \tag{B.35}$$

Plotting both sides of the equation with respect to χ on the same set of axes and calculating the intersections will yield the χ_0 that satisfy the equation. Figure B.3 shows the solutions, χ_0, that satisfy the transcendental relationship.

Figure B.3: This is a log plot in $\chi \to kL$ where $\gamma \to \frac{m}{m_b}$. The horizontal sloping solid lines represent $\gamma\chi$ for $\gamma = \{10^{-3} \ldots 10^{2}\}$. The vertically sloping solid lines represent $-\tan(\chi)$. These lines drift horizontally from $\chi = m\frac{\pi}{2} \to (m+1)\frac{\pi}{2}$ for $m = 1, 3, 5 \ldots$ which are represented by the red and blue dashed vertical lines, respectively. The intersections of the two solid curves are the solutions to the transcendental equation. As can be seen on the graph, the limits of γ correspond to the boundary conditions in those situations (i.e. $\gamma \to \infty$ Free-Fixed Condition; $\gamma \to 0$ Free-Free Condition).



$$\chi_0(\gamma, n) = \begin{pmatrix} 1.57714 & 4.71451 & 7.85525 & 10.9965 \\ 1.63199 & 4.73351 & 7.86669 & 11.0047 \\ 2.02876 & 4.91318 & 7.97867 & 11.0855 \\ 2.86277 & 5.76056 & 8.70831 & 11.7027 \\ 3.1105 & 6.22105 & 9.33173 & 12.4426 \\ 3.13845 & 6.27691 & 9.41536 & 12.5538 \end{pmatrix} \tag{B.36}$$

The rows correspond to the different values of γ which are the same as figure B.3. The first row is $\gamma = 10^2$ where the other rows are orders of magnitude smaller. The columns correspond to the n^{th} intersection or the n^{th} mode. So for this situation the wave number and the frequencies are

$$k_n = \frac{\chi_0(\gamma, n)}{L} \tag{B.37}$$

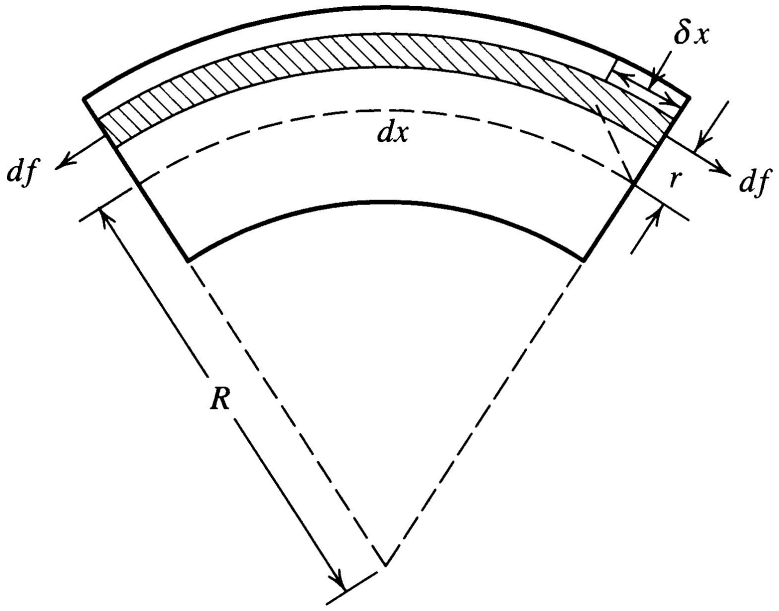
$$\omega_n = v\frac{\chi_0(\gamma, n)}{L} \tag{B.38}$$

v is the wave phase speed and $n = 1, 2, 3 \ldots$ The successive frequencies are not integer multiples of each and therefore are not harmonically related. When $\gamma = 1$ (the bar and the point mass are equal) now the nodal point, the points of no vibration, for $n = 1$ is shifted towards the end away from the middle as it approaches the Free-Fixed condition from the Free-Free condition. A minor adjustment to the system results in some differences in the frequency structure or distribution of the vibrating solid.

Longitudinal motion is not the only motion that a bar can experience. In fact, when a bar is struck imprecisely down its longitudinal axis most of the vibration is transverse. Analyzing the transverse vibrations of a bar is challenging but will illuminate some properties of sound production that have not yet been encountered.

The figure (Fig. B.1) shows a long bar of length, L, with a uniform cross sectional area of A. The x coordinate measures positions along the bar and the y coordinate the transverse displacements from its equilibrium configuration. Figure B.4 shows a segment of a bent bar where the upper portion is stretched while the lower portion is compressed. At some point in the middle, called the neutral axis, the bar is neither stretched nor compressed. Consider the segment of the bar, dx, where the bending of the bar is observed with reference to radius of curvature, R, of the neutral axis. If $\delta x = (\partial\xi/\partial x)dx$ is the increment, due to the stretching, of length of a filament located

Figure B.4: Bent bar segment. (Kinsler, 2000)



r above the neutral axis then the longitudinal force df is equal to

$$df = -EdA\left(\frac{\partial\xi}{dx}\right) = -EdA\left(\frac{\delta x}{dx}\right) \tag{B.39}$$

$$= -EdA\left(\frac{r}{R}\right) \tag{B.40}$$

dA is the cross-sectional area of the filament. The angle of this arc in radians is $\theta = dx/R = (dx + \delta x)/(R + r)$ which yields the relationship $\delta x/dx = r/R$. The total longitudinal forces on the bar is zero since tensile (negative) forces above the neutral axis are canceled by compressive (positive) forces below it. There is a bending moment M in the bar,

$$M = \int r df = -\frac{E}{R}\int r^2 dA \tag{B.41}$$

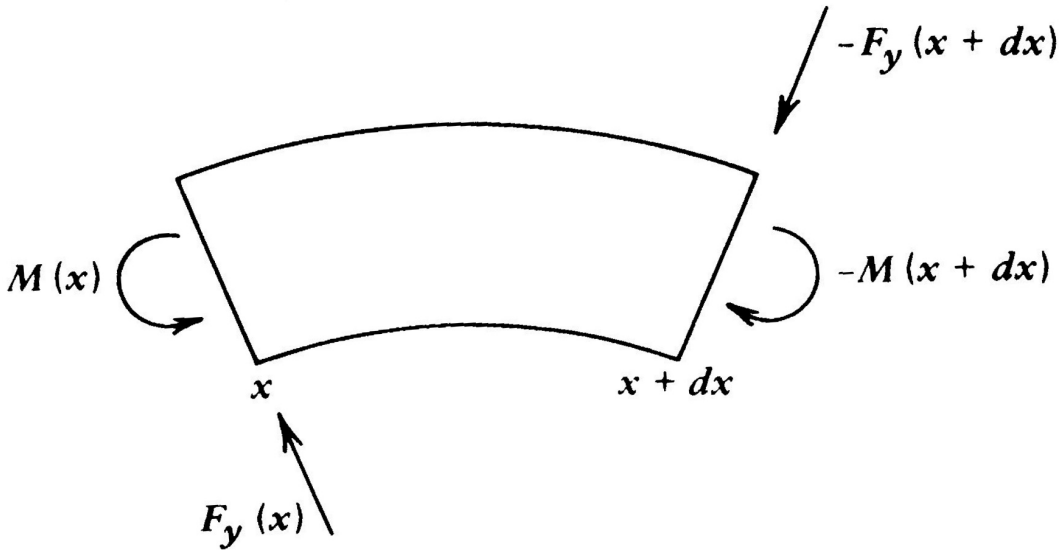
This moment (Fig. B.5) can be thought of as a net torque acting about the neutral axis which isn't a point but a line so not a pure analogy. Defining the constant κ by

$$\kappa^2 = \frac{1}{A}\int r^2 dA \tag{B.42}$$

then the moment M is

$$M = -EA\frac{\kappa^2}{R} \tag{B.43}$$

Figure B.5: Moment of bent bar segment. (Kinsler, 2000)



κ is the second moment of area, a property of shape. In this case its the moment of the cross sectional area. Radius of curvature can be shown to equal

$$R = \frac{[1 + \left(\frac{\partial y}{\partial x}\right)^2]^{3/2}}{\frac{\partial^2 y}{\partial x^2}} \approx \frac{1}{\frac{\partial^2 y}{\partial x^2}} \tag{B.44}$$

with the assumption that displacements are small so that $\partial y/\partial x \ll 1$. Then the moment M is

$$M = -EA\kappa^2 \left(\frac{\partial^2 y}{\partial x^2}\right) \tag{B.45}$$

The sign convention for the moment M or torque is that counterclockwise bending is caused by a positive moment and clockwise bending is caused by a negative moment. Distorting the bar not only produces bending moments but also shear forces, F_y, on the edges of each segment. The sign convention for the shear forces on the edge of the segments is positive for a shear acting upward and negative for a shear acting downward. For the condition of static equilibrium, the net turning moment should be zero. Taking the moments with respect to the left end of the segment is

$$M(x) - M(x + dx) = F_y(x + dx)dx \tag{B.46}$$

with the expansion of a Taylor's series and only keeping linear terms

$$F_y = -\left(\frac{\partial M}{\partial x}\right) = EA\kappa^2\left(\frac{\partial^3 y}{\partial x^3}\right) \tag{B.47}$$

When a bar undergoes transverse vibrations it is not in static equilibrium, but since the displacements are small the condition of static equilibrium is valid since the contributions to the net torque (eqn. B.46) on the segment from changing angular momentum are negligible: $\sum \tau = \frac{dL}{dt} \approx 0$. The net upward force acting on the segment is then

$$dF_y = F_y(x) - F_y(x+dx) = -\left(\frac{\partial F_y}{\partial x}\right)dx = -EA\kappa^2\left(\frac{\partial^4 y}{\partial x^4}\right) \tag{B.48}$$

and to get the wave equation using Newton's 2$^{\text{nd}}$ Law for the acceleration of the segment with mass $dm = \rho A dx$ where ρ is the density (mass/volume) of the bar.

$$\frac{\partial^2 y}{\partial t^2} \quad = \quad -(v\kappa)^2\left(\frac{\partial^4 y}{\partial x^4}\right) \tag{B.49}$$

where $v^2 = \frac{E}{\rho}$ from equation B.11. By assuming $y(x,t)$ is separable as in equation A.7 the wave equation can be expressed as

$$-\frac{1}{(\kappa v)^2}\frac{\ddot{g}}{g} = \frac{f^{(4)}}{f} \Rightarrow k^4 \tag{B.50}$$

Remembering that by separating the space coordinates from the time coordinates each ratio must equal a constant. Since the second time derivative needs to have the opposite sign as the original function so that it is bounded in time the k^4 has a positive sign. This leads to a relationship between k and ω

$$\omega = (\kappa v)k^2 \tag{B.51}$$

The solution to the fourth order spatial differential equation in f leads to four equations of the form

$$\frac{f^{(4)}}{f} = k^4 \Rightarrow f(x) = e^{\pm kx}; e^{\pm ikx} \tag{B.52}$$

$$f(x) = A_1 e^{kx} + A_2 e^{-kx} + A_3 e^{ikx} + A_4 e^{-ikx} \tag{B.53}$$

where the coefficients A_i can also be complex. The time coordinate equation has the familiar general solution

$$g(t) = B_1 e^{i\omega t} + B_2 e^{-i\omega t} \tag{B.54}$$

for simplicity only the positive ω will be kept. This leads to a complete general solution for a single frequency of the form

$$y(x,t) = (A_1 e^{kx} + A_2 e^{-kx}) e^{i\omega t} + A_3 e^{i(\omega t + kx)} + A_4 e^{i(\omega t - kx)} \tag{B.55}$$

where the actual displacement is the real part. The behavior has both oscillatory and decay behavior characterized by the value of k. So k acts as a wave number and an spatial attenuation coefficient. k's dependance on ω is different then that of both the longitudinal bar and the string where the dependence was linear.

$$k = \sqrt{\frac{\omega}{\kappa v}} \tag{B.56}$$

For both the string and the longitudinal bar the phase speed for the waves was the same for each frequency, which came out of the form of the wave equation. The phase speed, v_p, for transverse vibrations is

$$v_p = \kappa v k = \sqrt{\kappa v \omega} \tag{B.57}$$

This shows that each frequency component will propagate at its own unique phase speed which is known as dispersion. Higher frequencies will travel faster then lower frequencies which destroys the initial wave shape. To get a better understanding of these properties, it is helpful to work at solutions to some specific boundary and initial conditions. First thing to note is that since it is a fourth order differential equation, there will be 4 boundary conditions. For a clamped end the boundary conditions are that both the displacement and the slope are zero for all time. $y = 0; \frac{\partial y}{\partial x} = 0$. For an end to be free the boundary conditions require that the shear force, F_y, (eqn. B.47) and the moment M (eqn. B.45) be zero, which leads to: $\frac{\partial^2 y}{\partial x^2} = 0; \frac{\partial^3 y}{\partial x^3} = 0$. Another condition could be that of where the end is constrained to a position ($y = 0$) but no moment or torque on the end ($\frac{\partial^2 y}{\partial x^2} = 0$). For a bar of length L, using the definitions of $\cosh(\theta)$ and $\sinh(\theta)$ equation B.55 can be expressed as

$$y(x,t) = (B_1 \cos(kx) + B_2 \sin(kx) + B_3 \cosh(kx) + B_4 \sinh(kx)) \cos(\omega t + \phi) \tag{B.58}$$

where all the B_i's are real. For the conditions of a bar clamped at both ends, $B_1 = -B_3$, for $y(0,t) = 0$ and$B_2 = -B_4$ for $\frac{\partial y}{\partial x}|_{x=0} = 0$. The second set of boundary condition as

$x = kL$ is

$$B_1(\cos(kL) - \cosh(kL)) + B_2(\sin(kL) - \sinh(kL)) = 0 \tag{B.59}$$

$$-B_1(\sin(kL) + \sinh(kL)) + B_2(\cos(kL) - \cosh(kL)) = 0 \tag{B.60}$$

Maneuvering the equations so that the coefficients are on opposite sides of the equals and then cross multiplying after dividing the equations yields the transcendental equation

$$\cos(kL)\cosh(kL) = 1 \tag{B.61}$$

Simplifications come from the identities $\cos^2(\theta)+\sin^2(\theta) = 1$ and $\cosh^2(\theta)-\sinh^2(\theta) = 1$. Since $\cosh(x) \to e^x$ as $x \to \infty$, after a few terms the solutions can be approximated as the zeros of $\cos(kL)$. With the modes

$$k = (3.01124, 4.99951, 7.00002, 9...)\frac{\pi}{2L} \tag{B.62}$$

$$\omega = (3.01124^2, 4.99951^2, 7.00002^2, 9^2...)\frac{\kappa\upsilon\pi^2}{4L^2} \tag{B.63}$$

Clearly the frequency components are not harmonically related to the lowest component. On top of that, each component will have its unique phase speed and attenuation constant. The higher frequency components are damped more quickly, leaving only the lowest component vibrating. This is perceived initially as a metallic sound the blends into a longer duration pure tone at the fundamental. From this example, one can see that solids have complicated time varying frequency component structure and it is the fast heavily damped components that give the sound its identification which in this case is metallic.

Appendix C

Sample Code

In most cases, algorithms are best understood by attempting to replicate the procedure. The following code represents this process implemented in Mathematica. The assumptions are rather restrictive but illustrate the key points of the algorithm.

Defining physical constants and sensor grid. (m sensors: $\theta\left\{0 \to \frac{\pi}{2}\right\}$ and$\phi\{0 \to \pi\}$

```
μ0 = 4π; R0 = Q^.5; nθ = 5; nϕ = 5;
θ = Table[πu, {u, -nθ 1/2/(nθ+1), nθ 1/2/(nθ+1), 1/2/(nθ+1)}]; Length[θ]; θ
ϕ = Table[πv, {v, 1/(nϕ+1), nϕ 1/(nϕ+1), 1/(nϕ+1)}]; ϕ
θϕ = Flatten[Table[{θ[[i]], ϕ[[j]]}, {i, Length[θ]}, {j, Length[ϕ]}], 1];
θϕ = DeleteDuplicates[θϕ];
Length[θϕ]
```

Defining source grid. (q source location pts.)

```
α = Table[w, {w, .05, .95, .1}]; α
```

Lead field Matrix. Maps source locations to field sensors. (mxq)

```
L = Table[μ0/(4π) Qα[[j]]/R0^2 Sin[θϕ[[i,1]]]Sin[θϕ[[i,2]]]/(1-2Cos[θϕ[[i,1]]]α[[j]]+α[[j]]^2)^(3/2), {i, Length[θϕ]}, {j, Length[α]}];
MatrixRank[L]
Dimensions[L]
{ul, wl, vl} = SingularValueDecomposition[L];
wl//MatrixForm
```

Generate sample data. (m sensors x nt time samples)

```
nt = 30; dt = .01; nQ = 2; aQ = 10;
T = ntdt;
f = Table[w.5, {w, Length[α]}]

iQ = {};
s0 = Table[0, {w, Length[α]}];
While[Length[iQ] < nQ,
iQ = RandomInteger[{1, Length[α]}, nQ];
iQ = DeleteDuplicates[iQ];]
Q0 = Table[aQ(RandomReal[] - .5), {i, nQ}];
For[i = 1, i ≤ nQ, i++, s0[[iQ[[i]]]] = Q0[[i]]]
S = Table[s0[[w]]Sin[2πf[[w]]t], {w, Length[α]}, {t, dt, T, dt}];
s0
Dimensions[S]
Dimensions[L.S]
σ = 0.05√(Mean[Mean[(L.S)^2]])
n = Table[RandomReal[{-σ, σ}], {i, Length[θϕ]}, {j, nt}]; Dimensions[n]
b = L.S + n;
ListPlot[L.S]
ListPlot[n]
ListPlot[b]
```

SVD of data covariance matrix of the sensor measurements.

$R = \text{Table}[\text{Mean}[b[[i]]b[[j]]], \{i, \text{Length}[\theta\phi]\}, \{j, \text{Length}[\theta\phi]\}]; \text{Dimensions}[R]$
$\{u, w, v\} = \text{SingularValueDecomposition}[R]; \text{Dimensions}[u]$
$\text{MatrixRank}[w]$
$w//\text{MatrixForm}$

SVD of Lead field matrix for loop of all locations for subspace correlations.

s0
α
$\text{sourceIndex} = \{\};$
$\text{correlations} = \{\};$
$\text{ListLinePlot}[\text{SingularValueList}[R]]$
$\text{numS} = \text{nQ} + 1;$
$\text{aProj} = \text{IdentityMatrix}[\text{Length}[\theta\phi]];$
$\Phi\text{s} = u[[\text{All}, 1;;\text{numS}]];$
$\text{Ahat} = \text{Table}[\{\}, \{i, \text{Length}[\theta\phi]\}];$
$\text{AhatT} = \{\};$
$n = 0;$

The loop below repeats for the number of estimated dipoles. In practice this is estimated from the singular values of the data covariance matrix. Here this is set near ' nQ' times. Also in practice the algorithm should exit once the user specified correlation threshold is no longer exceeded.

$\text{While}[n \leq \text{numS} - 1,$
$l = \text{Transpose}\left[\text{Table}\left[\frac{\text{aProj}.L[[\text{All},i]]}{\text{Norm}[\text{aProj}.L[[\text{All},i]]]}, \{i, \text{Length}[\alpha]\}\right]\right];$
$\text{Cc} = \text{Table}[\text{Abs}[\text{Transpose}[\text{aProj}.\Phi\text{s}].l[[\text{All}, i]]], \{i, \text{Length}[\alpha]\}];$
$\text{If}[n \neq 0, \text{Cc}[[\text{Flatten}[\text{sourceIndex}], n + 1]] = 0];$

$n = n + 1;$
$\text{iS} = \text{Flatten}[\text{Position}[\text{Cc}[[\text{All}, n]], \text{Max}[\text{Cc}[[\text{All}, n]]]]];$

```
a = l[[All, iS]];
at = Flatten[Transpose[a]];
AppendTo[correlations, Max[Abs[Cc[[All, n]]]]];
AppendTo[sourceIndex, iS];

Table[AppendTo[Ahat[[i]], a[[i, 1]]], {i, Length[θϕ]}];
AppendTo[AhatT, at];
Ahat//MatrixForm;
AhatT//MatrixForm;
aProj = IdentityMatrix[Length[θϕ]] − Ahat.Inverse[AhatT.Ahat].AhatT; ];
```

Finding the time series estimates.

```
correlations
Qi = Flatten[sourceIndex]
At = Table[L[[All, Qi[[i]]]], {i, Length[Qi]}];
A = Transpose[At];
St = Inverse[At.A].At.b;
```

Plot time series.

```
p1 = ListPlot[St];
p2 = ListLinePlot[S[[iQ]]]; iQ
Show[p1, p2]
```