MASKING OF MOTION CUES BY RANDOM MOTION:
COMPARISON OF HUMAN PERFORMANCE
WITH A SIGNAL DETECTION MODEL

by

Glenn Lewis Greig

January 1988
MASKING OF MOTION CUES BY RANDOM MOTION:
COMPARISON OF HUMAN PERFORMANCE
WITH A SIGNAL DETECTION MODEL

by

Glenn Lewis Greig

Submitted Sept. 1987

© Glenn Lewis Greig 1987

January 1988
I take great pleasure in thanking the many people who contributed to this thesis. First and foremost, I am grateful to my thesis supervisor, Professor Lloyd D. Reid, for his enthusiastic assistance and guidance. The other members of my Advisory Committee, Professors Jaap H. de Leeuw and Gordon W. Johnston, provided valuable comments during the course of the study and on the final document. Professor Neville Moray of Industrial Engineering introduced me to signal detection theory, and participated in several helpful discussions about this project.

Many of my associates among the staff and students of the Institute also contributed. I would like to mention a few who were particularly helpful. Peter Grant helped with the motion measurement instrumentation and software and ran some of the experiments. Meyer Nahon helped me with motion-drive algorithms. Steve Hitchman built some of the necessary electronic hardware, and Wolf Graf provided the technical assistance needed to keep the flight simulator running smoothly. Several members of the flight simulation group served as experimental subjects, and participated in many discussions on the nature of motion perception.

I dedicate this thesis to my wife Gail, whose patient support and encouragement were instrumental in helping me to complete this project.

This work was partially funded by the Natural Sciences and Engineering Research Council of Canada (NSERC).
Abstract

This report describes an investigation of human sensitivity to whole-body motion. Specifically, it discusses the ability of human subjects to detect a sinusoidal motion signal superimposed on a background of random motion. The purpose of the study was to determine the conditions in which motion cues are masked or hidden by concurrent random motion. The results have applications to flight simulation, and will also be of interest to other researchers working in the area of human perceptual performance.

It is proposed that for the situation under study, motion perception is a signal-in-noise detection process which can be modelled using signal detection theory. A brief review of signal detection theory is provided. Three ideal detectors adapted from the literature on auditory perception are proposed as potential models for motion perception.

Three motion perception experiments were run. In the first, a rating procedure was used to obtain receiver operating characteristic (ROC) curves for human subjects detecting sinusoidal motion in a background of low power broadband random motion. A good fit to the data was obtained using ROC curves based on Gaussian distributions of signal and noise. The second experiment used a 2-alternative forced choice task to determine the detectability of sinusoidal motion in a variety of noise conditions. The results show that detectability can be expressed as a function of signal-to-noise ratio, and that sinusoidal motion is masked primarily by noise components which are near the signal frequency. The third experiment tested the extent to which noise on one axis masks a signal on another axis. Inter-axis effects were found to be small, but significant. All three experiments provided an estimate of the slope of the psychometric curve.

Of the three ideal detectors considered, the energy detector agrees best with the experimental data. The data are compared in detail with the predictions of the energy detector. A simplified method for estimating signal detectability in arbitrary conditions is presented. Finally, the implications of the results for flight simulation are discussed.
# Table of Contents

Acknowledgements  
Abstract  
Table of Contents  
Notation  

## 1 Introduction

1.1 Motivation For This Study  
1.2 Scope and Organization  

*Figures*

## 2 A Review of the Literature

2.1 Introduction  
2.2 Reference Frames  
2.3 Vestibular Sensors  
2.3.1 Semicircular Canals  
2.3.1.1 Physiology of the Semicircular Canals  
2.3.1.2 Mathematical Models of the Semicircular Canals  
2.3.2 The Otolith Organs  
2.3.1.1 Physiology of the Otoliths  
2.3.1.2 Dynamic Model of the Otoliths  
2.3.3 Other Motion Sensors  

2.4 Motion Perception Thresholds  
2.4.1 Perception Thresholds for Angular Motion  
2.4.2 Perception Thresholds for Linear Motion  
2.4.3 Factors Influencing the Threshold  
2.4.3.1 Effects of the Visual Field
3 An Introduction to Signal Detection Theory

3.1 Introduction 3.1
3.2 Statistical Decision Theory 3.2
3.3 The ROC Curve 3.6
3.4 Ideal Detectors 3.9
  3.4.1 Discrete Representation of Signal and Noise 3.10
  3.4.2 Correlation Detector 3.12
  3.4.3 Envelope Detector 3.14
  3.4.4 Energy Detector 3.18
  3.4.5 Comparison of Detectors 3.22
Tables 3.26
Figures 3.27

4 The Experimental Approach

4.1 Introduction 4.1
4.2 Proposed Motion Perception Model 4.1
4.3 Objectives of the Experimental Program 4.4
4.4 The Experimental Facility 4.7
4.5 Generation of Signal and Noise 4.8
Tables 4.13
Figures 4.15

5 Determination of ROC Curves

5.1 Introduction 5.1
5.2 Experimental Method and Conditions 5.2
  5.2.1 Conditions 5.3
  5.2.2 The Rating Procedure 5.5
  5.2.3 Influencing the Response Criteria 5.7
5.3 Experimental Results
  5.3.1 Consistency Checks
  5.3.2 Fitting Models to the Data
  5.3.3 Psychometric Curves

Tables
Figures

6 The Effect of Noise Power and Frequency Content
  6.1 Introduction
  6.2 Experimental Method and Conditions
  6.3 Experimental Results
    6.3.1 Consistency Checks
    6.3.2 The Effect of Noise Condition
    6.3.3 The Psychometric Curve

Tables
Figures

7 Inter-Axis Effects
  7.1 Introduction
  7.2 Experimental Method and Conditions
  7.3 Experimental Results
    7.3.1 Consistency Checks
    7.3.2 Effect of Noise Condition
    7.3.3 Psychometric Curves

Tables
Figures

8 Discussion of Experimental Results
  8.1 Introduction
  8.2 Comparison Between Experiments
  8.3 Comparison With Models
  8.4 Detailed Comparison With Energy Detector
    8.4.1 Description of the Simulation
    8.4.2 Estimation of the Internal Noise Power
## Notation

A summary of the most important symbols used in this report is given below. Many other symbols appear only within a single chapter or section. These are defined when they appear and are not listed here. Unavoidably, a few of the symbols given below were assigned multiple definitions. In cases where the intended meaning is not clear from the context, the definition is repeated in the text. Also, a few of the symbols listed here take on different meanings in isolated sections or appendices. In such cases, the appropriate definition is given in the accompanying text.

A few general conventions are worth noting. Vectors are indicated by boldface type. Laplace transformed variables are indicated by an overscore. Time derivatives are indicated using the dot convention.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>inertial acceleration vector</td>
</tr>
<tr>
<td>A</td>
<td>normalized difference of means</td>
</tr>
<tr>
<td>A</td>
<td>signal amplitude (section 3.4.3)</td>
</tr>
<tr>
<td>A*</td>
<td>normalized signal amplitude (section 3.4.3)</td>
</tr>
<tr>
<td>A_cr</td>
<td>critical amplitude (amplitude for which (d'_{A} = 1.0))</td>
</tr>
<tr>
<td>b</td>
<td>slope of the ROC curve on normal probability paper</td>
</tr>
<tr>
<td>B</td>
<td>slope of the logistic function (psychometric curve)</td>
</tr>
<tr>
<td>d'</td>
<td>index of detectability (correlation detector)</td>
</tr>
<tr>
<td>(d'_{A})</td>
<td>index of detectability based on area under ROC curve</td>
</tr>
</tbody>
</table>
decibels. For amplitude, this is \(20 \log(A/A_{\text{ref}})\)
For power or energy, it is \(10 \log(P/P_{\text{ref}})\)

e
a scalar event which results from the presentation of a stimulus. Also called the evidence variable.

\(E_s\)
signal energy

\(E[x]\)
expected value of \(x\)

\(f\)
specific force vector

\(f(\cdot)\)
probability density

\(f(e \mid n)\)
conditional density of \(e\) given that the stimulus was \(n\)

\(f(e \mid s)\)
conditional density of \(e\) given that the stimulus was \(s\)

\(g\)
the gravity vector

\(k\)
a decision criterion for a decision rule based on \(e\)

\(I_0(\cdot)\)
modified Bessel function of zero order

\(L(e)\)
likelihood ratio, \(f(e \mid s)/f(e \mid n)\)

\(n\)
a stimulus comprising noise alone

\(N\)
a negative response

\(N\)
number of trials or samples

\(N_o\)
noise power density

\(p\)
roll rate in rad/sec

\(P\)
a signal or noise condition comprising roll motion

\(P_{A}\)
area under the ROC curve

\(P(\cdot)\)
probability
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(n)</td>
<td>a priori probability that the stimulus is n</td>
</tr>
<tr>
<td>P(s)</td>
<td>a priori probability that the stimulus is s</td>
</tr>
<tr>
<td>P(S</td>
<td>n)</td>
</tr>
<tr>
<td>P(S</td>
<td>s)</td>
</tr>
<tr>
<td>P_k(S</td>
<td>n)</td>
</tr>
<tr>
<td>P_k(S</td>
<td>s)</td>
</tr>
<tr>
<td>P_2(C)</td>
<td>probability of responding correctly in a two-alternative forced choice task</td>
</tr>
<tr>
<td>P_4(C)</td>
<td>probability of responding correctly in a four-alternative forced choice task</td>
</tr>
<tr>
<td>q</td>
<td>pitch rate in rad/sec</td>
</tr>
<tr>
<td>Q</td>
<td>a signal or noise condition comprising pitch motion</td>
</tr>
<tr>
<td>r</td>
<td>yaw rate in rad/sec</td>
</tr>
<tr>
<td>r(t)</td>
<td>a noise waveform</td>
</tr>
<tr>
<td>r_m</td>
<td>a sample of r(t) at t = m/(2W)</td>
</tr>
<tr>
<td>s</td>
<td>the Laplace variable</td>
</tr>
<tr>
<td>s</td>
<td>a stimulus condition comprising signal plus noise</td>
</tr>
<tr>
<td>s(t)</td>
<td>a signal waveform</td>
</tr>
<tr>
<td>s_m</td>
<td>a sample of s(t) at t = m/(2W)</td>
</tr>
<tr>
<td>S</td>
<td>a positive response</td>
</tr>
<tr>
<td>t</td>
<td>time</td>
</tr>
</tbody>
</table>
$T$ \quad \text{duration of the observation interval}

$W$ \quad \text{bandwidth of a rectangular noise spectrum}

$x$ \quad \text{location or distance along the x-axis}

$x(t)$ \quad \text{a stimulus waveform which may be either noise alone or signal plus noise}

$X$ \quad \text{a signal or noise condition comprising surge motion (motion parallel to the x-axis)}

$y$ \quad \text{location or distance along the y-axis}

$Y$ \quad \text{a signal or noise condition comprising sway motion (motion parallel to the y-axis)}

$z$ \quad \text{location or distance along the z-axis}

$z[x(t)]$ \quad \text{a scalar evidence variable based on the waveform } x(t). \text{ See "e".}

$\alpha$ \quad \text{significance level for a statistical test}

$\beta$ \quad \text{decision criterion for a decision rule based on } L(e)

$\mu$ \quad \text{mean of a probability distribution}

$\mu_n$ \quad \text{mean value of } e \text{ given that the stimulus is } n

$\mu_s$ \quad \text{mean value of } e \text{ given that the stimulus is } s

$\sigma$ \quad \text{standard deviation of a probability distribution}

$\sigma_n$ \quad \text{standard deviation of } e \text{ given that the stimulus is } n

$\sigma_s$ \quad \text{standard deviation of } e \text{ given that the stimulus is } s

$\sigma_{ab}$ \quad \text{covariance of } A \text{ and } b

\text{xi}
\( \tau \)  
a time constant

\( \Phi(\omega) \)  
noise power density at frequency \( \omega \)

\( \omega \)  
frequency in radians per second

\( \omega \)  
angular velocity vector
Chapter 1

Introduction

1.1 Motivation for This Study

This report presents the results of a study of human sensitivity to whole-body motion. The specific purpose of the study was to determine how sensitivity to low-amplitude sinusoidal motion is affected by the presence of random motion or vibration. The approach is fairly general, and the results apply to many situations in which humans are exposed to motion. However, the study was motivated by a need to improve the fidelity of moving-base flight simulators, and the conditions studied were chosen because of their relevance to flight simulation.

Ground based flight simulators were introduced shortly after the advent of powered flight, and have developed to a high degree of sophistication. Modern simulators provide a fairly realistic reproduction of conditions in the cockpit, and are widely used for pilot training. Simulators have several advantages over aircraft for training. They are cheaper to operate. More important, they can be used to train the aircrew to handle emergency situations without endangering the crew or aircraft. Also, because they can be programmed to produce difficult scenarios on command, simulators can provide more intensive training than is available in an aircraft. Commercial aircrews receive much of their training in flight simulators.

Most modern flight simulators comprise a cab mounted on a movable platform. This platform, or motion base, is used to provide the pilot with a sensation of motion similar to that felt in the aircraft. Puig et al (1978) trace the evolution of motion simulators from the early Link trainer, patented in 1931, to the multi-degree-of-freedom systems in use today. The most common motion base in use today is the 6-post synergistic base patented independently by Peterson in 1966 and Cappel in 1967. This is an arrangement of six hydraulic jacks capable of moving in all six degrees of freedom.

There is an ongoing dispute in the flight simulation community over the value of simulator motion. Its detractors claim that motion does not improve training effective-
ness and that the expense of a motion base is not justified. Its proponents claim that motion cues improve the realism of the simulation, thereby increasing pilot acceptance of simulators as trainers. They also note that a pilot’s control behaviour is different with motion than without, for several reasons. First, motion cues help the pilot anticipate the vehicle motion, and thus increase his control bandwidth. Second, high amplitude static or oscillatory acceleration can interfere with a pilot’s ability to control the vehicle. This effect would be absent in a fixed-base simulation. Finally, whole body vibration can affect performance by impairing the pilot’s visual perception.

Some researchers claim that motion cues do, in fact, aid training. Guercio and Wall (1972), Junker and Replogle (1975), and Showalter and Parris (1980) all found that flight control tasks involving complex or unstable vehicle dynamics were learned faster with motion cues than without. Some sensation of motion is induced by movement of the visual field. This is referred to as vection. However, Young (1978) found that the vestibular system must be stimulated initially in order for vection to be effective. Caro (1979) provides a good review of the arguments for and against motion.

Part of the reason for the dispute is that, because the simulator has a limited displacement, it cannot duplicate the aircraft motion. At best, the simulator motion is an approximation to the aircraft motion. At worst, the simulator motion is vastly different from that of the aircraft. In this case, the inclusion of motion may be detrimental since it teaches the pilot to respond to inappropriate cues. Therefore, if motion is to be included, care must be taken to minimize the difference in perceived motion between the aircraft and the simulator.

A typical motion drive algorithm is shown in simplified form in Figure 1.1. The aircraft motion, represented as specific forces and angular rates at the cockpit, is passed through a second or third order high pass filter to generate the simulator motion. High frequency motions are reproduced faithfully, while low frequency motions are filtered out. The coupling term allows prolonged (i.e. low frequency) translational accelerations to be approximated by tilting the simulator cab. Various filtering techniques have been employed. Reid and Nahon (1985,1986) recently developed and evaluated three different motion drive algorithms, based on classical, optimal and adaptive control theory, respectively.
Such algorithms have difficulty simulating some common aircraft maneuvers such as prolonged surge acceleration, pull-up, and coordinated turn entry. As an example, the typical simulator response (neglecting tilt coupling) to a step forward acceleration is shown in Figure 1.2. Initially, the simulator accelerates in the same direction as the aircraft, resulting in a congruent motion cue. This is short-lived, however. The simulator must stop before exceeding its displacement limit, then return to the centre of the motion envelope in order to have some travel available for the next excursion. Therefore the simulator must accelerate in a direction opposite to the aircraft acceleration, resulting in a spurious or conflicting cue. The amplitude and duration of the motion cues can be controlled to some extent by changing the parameters of the motion drive algorithm.

At the risk of grossly oversimplifying the nature of motion perception, we can at least say that we want the pilot to feel the congruent cue, but not the spurious cue. A common approach in flight simulation is to assume a motion perception threshold, then design the motion drive algorithm so that spurious cues seldom, or never, exceed the threshold.

A key question, therefore, is: how big must a congruent motion cue be to ensure that the pilot feels it? Alternatively, how small must a spurious cue be to ensure that the pilot does not feel it? How are these values affected by pilot workload, visual display conditions, and the level of background motion or vibration in the simulator? Some of these questions have been studied before. The relevant work is discussed in Chapter 2. This study addresses the last of the questions above: how is human sensitivity to motion affected by the background motion level?

1.2 Scope and Organization

As discussed above, the purpose of this study was to determine how human sensitivity to motion is affected by background motion. Early in the study, it became clear that there is a paucity of information on this aspect of motion perception. Therefore, we attempted to establish some groundwork in this area. Two specific aims were set. The first was to experimentally measure the subjective detectability of motion cues in a variety of conditions of interest in flight simulation. The second was to find a model which could be used to extrapolate the results to a broader range of conditions.
It is proposed that motion perception in a flight simulator can be represented as a signal-in-noise detection problem. The signal is the motion cue of concern, which may be either congruent or spurious. The noise is the background motion in the simulator, which might result from buffet, turbulence, or runway roughness. Given the characteristics of signal and noise, we want to know if the pilot can detect the signal.

Signal detection theory was adopted as the theoretical framework for the study. Three signal detection models were borrowed from the literature on auditory detection. These models are best suited to describing the detectability of periodic signals in random noise. Three motion detection experiments involving human subjects were run to test various aspects of these models. The experimental data were compared with the theoretical predictions. One model, an energy detection model, agreed reasonably well with experiment.

A literature review revealed no previous attempts to define a model to describe the effect of random vibration on the detectability of whole-body motion. In particular, there have been no previous attempts to apply signal detection theory, or the theory of ideal observers, to motion perception. As a result, this project became a rather basic study in human perception. Some of the implications for flight simulation are obvious. More detailed applications have been left for future researchers.

The remainder of this report is divided into eight chapters. Chapters 2 and 3 provide some background. Chapter 2 describes the vestibular sensors and discusses some of the mathematical models of these sensors that have been proposed. It also presents a brief review of the literature on motion perception thresholds. Chapter 3 gives a brief summary of signal detection theory. This includes the derivation of equations describing the performance of three ideal detectors proposed as possible motion detection models.

Chapters 4 through 7 describe the experimental program. Chapter 4 outlines the objectives of the experiments, and describes the experimental facility. The experiments are described in Chapters 5, 6 and 7, with each chapter devoted to one experiment. These chapters describe the experimental procedure, the data, and some comparisons of the data with theory.

The results of all three experiments are discussed in Chapter 8. First, the results are compared with each other and with data reported in the literature. Then the data are compared with the three signal detection models. An energy detection model provides
the best overall agreement with the data. A computer simulation of the energy detector was used to generate predictions for detailed comparison with the experimental results.

Finally, Chapter 9 provides a summary of the major findings, and suggests some topics for further study.
Figure 1.1  Simplified block diagram of a typical motion-drive algorithm (pitch and surge degrees of freedom).
Figure 1.2  Response of a typical moving-base simulator to a step acceleration of the aircraft.
Chapter 2

A Review of the Literature

2.1 Introduction

This chapter provides a brief review of the literature on human non-visual motion sensation. The review has two specific aims. The first is to show where the current study fits into the existing body of knowledge on motion sensation. The second is to introduce vestibular response models which will become components of the motion detection model developed later.

Non-visual motion and orientation cues are available from three different sources. First, changes in the pressure applied to the skin results in somatosensory cues. Second, feedback of muscular tension in the joints and limbs provides kinesthetic, or postural cues. Finally, acceleration of the head relative to an inertial frame stimulates the vestibular motion receptors in the inner ear. Of the three sources, the vestibular system is considered to be the most important.

Because of its importance in human spatial orientation, the vestibular system has been a topic of study for well over 100 years. This review summarizes the findings which are most relevant to the current study. More detailed reviews are provided by Boring (1942), Meiry (1966), Howard and Templeton (1966), Peters (1969), Young (1969), Ormsby (1974) and Guedry (1974). The most recent comprehensive review is that by Zacharias (1978). His review covers vestibular models, perception thresholds, and the application of perception models in the modelling of pilot performance.

The remainder of this chapter is divided into four sections. Section 2.2 presents the coordinate system used throughout this report. Section 2.3 discusses the physiology of the vestibular sensors, and describes mathematical models for perception of angular and translational motion. Section 2.4 summarizes the available data on motion perception thresholds, and discusses how sensitivity to motion is affected by other conditions in the cockpit. Finally, Section 2.5 introduces an alternative to the threshold model, the signal-in-noise model, which is used in the current study.
2.2 Reference Frames

In discussing motion cues and human motion sensing capabilities it is convenient to use two reference frames. The first is a body frame, \( F_H \), illustrated in Figure 2.1. It is attached to the pilot’s (or subject’s) head with its origin midway between his right and left vestibular labyrinths. The \( X \) axis points forward, the \( Y \) axis points to the right, and the \( Z \) axis points downward along the spine. The \( X-Z \) plane contains the subject’s vertical plane of symmetry.

The second frame used is an inertial frame, \( F_I \), fixed to the earth. The \( Z \) axis is aligned with the gravity vector, and the \( X \) axis is chosen to suit the problem at hand.

All whole-body motions discussed in this report are defined with respect to the body-fixed reference frame \( F_H \). The following terms are used to denote motions in the six degrees of freedom in this frame:

- surge (longitudinal): translation along the \( X \) axis.
- sway (lateral): translation along the \( Y \) axis.
- heave (vertical): translation along the \( Z \) axis.
- roll: rotation about the \( X \) axis.
- pitch: rotation about the \( Y \) axis.
- yaw: rotation about the \( Z \) axis.

2.3 Vestibular Sensors

The most important non-visual motion receptors in humans and other vertebrates are the vestibular organs, located in the inner ear. The major components of the vestibular system are shown in Figure 2.2. Each vestibular organ comprises three semicircular canals and two otolith organs, the utricle and the saccule. The semicircular canals respond to angular motion, while the otoliths respond to changes in specific force. The sensors are generally considered to be independent. However, Steer (1967) showed analytically how the semicircular canals might be stimulated by translational acceleration or static tilt. Ormsby and Young (1977) proposed a model in which otolith and canal signals are combined to give an optimal estimate of motion and spatial orientation.

The vestibular sensors play an important role in balance and orientation. Vestibular signals also help to control the oculomotor system. Stimulation of the vestibular system
results in compensatory eye movements such as nystagmus and counter-rolling, which help maintain a stable image on the retina as the head is moved.

2.3.1 Semicircular Canals

2.3.1.1 Physiology of the Semicircular Canals

The semicircular canals are sets of three roughly orthogonal thin membranous tubes. Each canal forms about two thirds of a circle; the utricle completes the circuit for each canal. The approximate orientation of the canals is shown in Figure 2.3. The utricle and canals are filled with a fluid known as endolymph. Near one end of each canal is an enlarged portion, the ampulla, which is completely sealed by the cupula. The cupula, a neutrally buoyant gelatinous wedge, is mechanically coupled to sensory hairs projecting from the crista at the base of the ampulla. A sketch of the horizontal canal showing the major components is given in Fig 2.4.

When the head undergoes angular acceleration in the plane of a canal, the endolymph, due to its inertia, lags behind the motion of the canal itself. The movement of the endolymph relative to the canal displaces the cupula. The cupular displacement is detected by the sensory hairs in the crista. Displacement of the cupula bends the sensory hairs, thereby changing the firing frequency, or discharge rate, of the afferent nerve cells. This change in frequency is interpreted by the central nervous system as motion information. For a more detailed review of the innervation of the canals, see Peters (1969).

Each canal is sensitive to angular acceleration in its own plane. Because the three canals are orthogonal, the vestibular system senses rotation about all three axes. It is often convenient to model the semicircular canals as a single, three degree-of-freedom rotation sensor located at the centre of the head.

2.3.1.2 Mathematical Models of the Semicircular Canals

Ernst Mach hypothesized in 1875 that motion sensation was related to inertial movement of fluid in the semicircular canals (Henn, 1984). Steinhausen (cited by Peters, 1969) proposed in 1931 that the mechanics of the semicircular canals could be represented by a heavily damped torsional pendulum. The inertia, spring constant and damping of the
pendulum represent, respectively, the inertia of the endolymph ring, the elastic restoring force of the cupula, and viscous forces arising between the endolymph and the canal walls. The equation of motion of the system is:

\[ M \ddot{\theta} + C \dot{\theta} + K \theta = M \ddot{\theta}_H \]  

(2.1)

where \( \ddot{\theta}_H \) is the angular acceleration of the head in inertial space,
\( \theta \) is the displacement of the endolymph with respect to the canal,
\( M \) is the inertia of the endolymph ring,
\( K \) is the effective spring constant,
and \( C \) is the effective damping constant.

The transfer function relating cupular displacement to head motion is obtained by taking the Laplace transform of equation 2.1. It is convenient to express the transfer function in terms of the head angular velocity, \( \omega_H \):

\[ \frac{\delta(s)}{\omega_H(s)} = \frac{H s}{s^2 + (C/M)s + K/M} \]  

(2.2)

where \( H \) is a mechanical coupling constant.

Available evidence indicates that the system is heavily overdamped, with \( C/M > K/C \). Therefore, equation 2.2 can be approximated by:

\[ \frac{\delta(s)}{\omega_H(s)} = \frac{H s}{(s + K/C)(s + C/M)} \]

\[ = \frac{H \tau_L \tau_S s}{(1 + \tau_L s)(1 + \tau_S s)} \]  

(2.3)

van Egmond et al (1949) estimated the parameters of this model based on subjective response data, and give the values \( \tau_L = 10 \) seconds and \( \tau_S = 0.1 \) seconds. The resulting transfer function suggests that the semicircular canals act as angular velocity sensors between 0.1 and 10 rad/sec. Melville Jones et al (1964) estimated the long time constant, \( \tau_L \), based on subjective response data. They report values of 6.1, 5.3 and 10.2 seconds for
rotation about the X, Y and Z body axes respectively. However, most current researchers do not distinguish between axes.

The value of $\tau_s$ noted above may be too large. van Egmond et al (1949) estimated $\tau_s$ based on hydrodynamic considerations and found a value of 0.037 seconds. Meiry (1966) estimated $\tau_s$ to be 0.04, based on observations of compensatory eye motion. More recently, Steer (1967) solved the Navier-Stokes equation for a simplified canal geometry and estimated $\tau_s$ to be about 0.005 seconds in humans.

The second order torsion pendulum model is the central feature of modern vestibular models. However, current models contain two additional components. The first is an adaptation operator of the form $\tau_A/(1 + \tau_A s)$. This was proposed by Young and Oman (1969) and by Schmid et al (1971) to account for the fact that both nystagmus and the sensation of motion during prolonged angular acceleration eventually decay to zero. Babin et al (1980) found that adaptation is unrelated to cupular displacement and suggest that it is due to neural causes.

Young and Oman (1969) set $\tau_A$ at 30 seconds for the subjective sensation of motion. The adaptation operator accounts for some of the low frequency phase lead previously associated with the long time constant $\tau_L$. Consequently, recent estimates of $\tau_L$ are higher than those noted above. Young and Oman suggest a value of 16 seconds; Schmid et al (1971) set $\tau_L$ at 18 seconds.

The second additional component is a lead operator of the form $(1 + \tau_R s)$. This implies that the canal afferent response is a function of both cupular displacement and rate of displacement. Evidence for this operator is provided by Nashner (1971) in a study of human postural control, and by Benson (1970) in a study of nystagmus records. In a neurological study, Goldberg and Fernandez (1971) found rate sensitivity of this form in the semicircular canal afferent response in the squirrel monkey. They found values of $\tau_R$ ranging from 0.013 to 0.094 seconds, with a mean of 0.049 seconds.

The final model relating change in neural firing frequency to angular velocity of the head is:

$$\frac{\Delta FR(s)}{\omega_H(s)} = \frac{HF \tau_L \tau_S s}{(1 + \tau_L s)(1 + \tau_S s)} \times \frac{\tau_A s (1 + \tau_R s)}{(1 + \tau_A s)}$$

(2.4)
where $\Delta FR$ is expressed in impulses/second, and $\omega_H$ is in rad/sec. Ormsby (1974) gives the following parameter values: $\tau_A = 30$, $\tau_L = 18$, $\tau_S = 0.005$, and $\tau_R = 0.01$. Ormsby also calculated the combined gain HF for a single neuron to be 6300. The amplitude ratio and phase characteristics of this model are shown in Figure 2.5. It is clear from Figure 2.5 that the semicircular canals act essentially as rotational velocity transducers for frequencies from 0.1 rad/sec to well over 20 rad/sec. Many systems models, such as that used by Buizza and Schmid (1982) simply model the semicircular canals as a velocity sensor following a first order high pass filter.

The total afferent firing rate equals the change in firing rate, described by equation 2.4, plus the spontaneous firing rate SFR, and the neural noise $n(t)$:

$$AFR(t) = \mathcal{L}^{-1}\{\Delta FR(s)\} + SFR + n(t) \quad (2.5)$$

where $\mathcal{L}^{-1}\{\cdot\}$ indicates the inverse Laplace transform. Goldberg and Fernandez (1971) report that the spontaneous firing rate is typically 90 impulses per second, based on their work with the squirrel monkey. This value is not critical, since it is assumed that the central nervous system processes only the change from the steady firing rate.

Only limited estimates of the neural noise characteristics are available. Ormsby has estimated the effective RMS noise based on the variance of $n(t)$ for vestibular afferents in the squirrel monkey, measured by Goldberg and Fernandez (1971). He assumes that the central nervous system first processes the vestibular signal by calculating the mean firing rate for all neurons. He sets the effective standard deviation of $n(t)$ at 0.223 impulses/second, assuming that the standard deviation for a typical regular cell is 5.1 impulses/second, and that there are 520 independent afferent noise components in each vestibular nerve. Hosman and van der Vaart (1978) derived similar results. No autocorrelations or spectral descriptions of the noise are available.

2.3.2 The Otolith Organs

2.3.2.1 Physiology of the Otoliths

As shown in Figure 2.2, there are two otolith organs in each inner ear, the utricle and the saccule. The general structure of the otolith organs is shown in Figure 2.6. The functional portion of the otoliths comprises a mass of calcium carbonate crystals, the
statoconia or otoconia, imbedded in a gelatinous membrane overlying a bed of sensory hairs known as the macula. The specific gravity of the otoconia is about 2.7. A change in the specific force in the plane of the macula results in shear displacement of the otoconia relative to the macula. This bends the sensory hairs and changes the firing frequency of the afferent neurons. The change in firing frequency is interpreted by the central nervous system as motion or tilt information.

The utricular macula is roughly parallel to a plane defined by the Y-axis and the vector lying 30° above the X axis. The utricle is sensitive primarily to lateral and longitudinal motions. The macula of the saccule, on the other hand, is roughly coplanar with the X-Z plane. It acts primarily as a vertical motion sensor. There is some disagreement in the literature about whether the saccule is sensitive to motion at all (Peters, 1969). However, studies by Melville Jones and Young (1978) and Melville Jones et al (1980) indicate that it is, and that its sensitivity is similar to that of the utricle. Therefore, for systems purposes we can consider the otolith organs as a single, three degree-of-freedom translational force sensor located at the centre of the head.

2.3.2.2 Dynamic Model of the Otoliths

The structure of the otoliths is suggestive of a mass-spring-dashpot system in which the movement of the otoconia is restrained by elastic forces from the gelatinous membrane and damped by the viscosity of the endolymph. This representation was proposed by de Vries (1950). Treating the otolith as a lumped-mass system yields the following equation of motion:

\[ M\ddot{\delta} + C\dot{\delta} + K\delta = M(\ddot{x}_H - g_H) \]  \hspace{1cm} (2.6)

where
\( \ddot{x}_H \) is the acceleration of the utricle (i.e. the head) in inertial space,<br>\( \delta \) is the displacement of the otoconia relative to the macula,<br>\( g_H \) is the component of gravity in the plane of the macula<br>\( M \) is the mass of the otoconia minus the mass of an equal volume of endolymph,<br>\( K \) is the effective spring constant,<br>and \( C \) is the effective damping constant.
It is clear from equation 2.6 that the otoliths, like all accelerometers, respond to specific force $f_H$, where:

$$f_H = a_H - g_H$$  \hspace{1cm} (2.7)

is the vector difference between the inertial acceleration and the acceleration due to gravity. Note that $f_H$ is expressed in $F_H$ components, where $F_H$ is the head-fixed reference frame.

Taking the Laplace transform of equation 2.6 yields the following transfer function for otolith displacement:

$$\frac{\delta(s)}{f(s)} = \frac{1}{s^2 + (C/M)s + K/M}$$  \hspace{1cm} (2.8)

If perceived force (or acceleration) is proportional to otolith displacement, and if $C/M \gg K/C$, then the transfer function relating perceived specific force to input specific force is:

$$\frac{\text{perceived } f(s)}{\text{actual } f(s)} = \frac{H}{(s + \omega_1)(s + \omega_2)}$$  \hspace{1cm} (2.9)

Meiry (1966) first estimated the parameters of this model by measuring the phase of subjective response to oscillatory (sinusoidal) surge acceleration. Young and Meiry (1968) later included a low frequency term in the numerator to improve the model predictions for response to sustained inputs or tilts. Their revised dynamic otolith model is:

$$\frac{\text{perceived } f(s)}{\text{actual } f(s)} = \frac{1.5(s+0.76)}{(s+0.19)(s+1.5)}$$  \hspace{1cm} (2.10)

The extra term implies that some of the sensory hairs are sensitive to rate of bending as well as bending. The amplitude ratio and phase characteristics of this model are shown in Figure 2.7. Over short periods and at frequencies greater than 1.5 rad/sec, this model acts as an integrating accelerometer, or velocity transducer.

The natural frequency of the otoliths, according to equation 2.10, is 0.53 rad/sec. However, Ormsby (1974) estimated the natural frequency of the otoliths to be between
370 and 2240 rad/sec, based on measured deflections of the otoconia. This estimate was corroborated by Fernandez and Goldberg (1976), who estimate the natural frequency of the otoliths to be on the order of 300 to 3000 rad/sec, and suggest a figure of 2500 rad/sec for man. Based on their work with the otoliths of squirrel monkeys, they found that the otolith mechanics could be modelled as a first order lag with a break frequency of 10 Hz.

Ormsby attributes the dynamics observed by Meiry to higher order processing, and proposes the following model for the otolith afferent response:

\[
\frac{\Delta FR(s)}{F(s)} = 90 \frac{s + 0.1}{s + 0.2} F(s) \tag{2.11}
\]

where \(\Delta FR\) is expressed as impulses/second, and \(F\) is specific force expressed as a fraction of gravity. The amplitude ratio and phase characteristics of this model are shown in Fig 2.8. The high-frequency behaviour of this model differs markedly from that of Meiry’s model. This model suggests that the otoliths act as acceleration transducers above 0.2 rad/sec.

Ormsby’s model has been confirmed to some extent by neurological data. Lowenstein and Saunders (1975) measured the afferent response of otolith neurons in the bullfrog, and fit the data with a transfer function identical in form to equation 2.11, but with slightly different parameters. Further supporting evidence is provided by Fernandez and Goldberg (1976), based on neurological studies of the squirrel monkey.

The total afferent firing rate is the sum of the change in firing rate, the spontaneous firing rate, and the neural noise. Fernandez and Goldberg (1976) give the spontaneous firing rate of a utricular neuron as 88 impulses/second. Following the same reasoning outlined for the semicircular canals, Ormsby estimated that the effective standard deviation of \(n(t)\) is about 0.15 impulses per second. Again, Hosman and van der Vaart (1978) derived a similar figure.

2.3.3 Other Motion Sensors

Most of the motion perception research to date has been directed toward the vestibular system, the most important non-visual motion sensing apparatus. However, motion cues in an aircraft or flight simulator are also sensed in other ways. Changes in pressure on
the skin result in somatosensory or somatic cues. Feedback of muscular tension provides kinesthetic or postural cues. Some work has been done on modelling these systems.

Meiry (1966) found that neck proprioception contributes to compensatory eye movements below 0.1 Hz. Lackner and Graybiel (1978a,1978b) discuss the effect of somatic cues on subjective orientation. Borah et al (1979) present models for tactile reception and head-neck proprioception for inclusion in an optimal control model of the human operator. Tactile receptors are modelled by a simple lead-lag network. The head-neck system is modelled as an inverted pendulum stabilized by the neck muscles.

At the same time, several studies have shown that such extra cues may be negligible in comparison with vestibular cues. Young (1967) tested the tracking performance of normal and vestibular-defective patients in a moving-base flight simulator. From his results, he concluded that "it is primarily vestibular system contributions which allow motion cues to aid pilot performance." Walsh (1961) found that translational motion perception thresholds for 4 subjects without otolith function were about 4 times those found for normal subjects. In the same study, he tested 12 subjects with spinal lesions which reduced the transmission of somatic and postural information from the trunk and limbs. Their thresholds did not differ significantly from those of normal subjects. Finally, Showalter and Parris (1980) found that whole-body motion cues improved pilot performance much more than somatic cues (provided by a g-seat), presumably because the motion cues stimulated the vestibular system. It might be concluded that, as a first approximation, human sensitivity to motion can be represented by the vestibular models alone.

2.4 Motion Perception Thresholds

For motion, as for other sensory modes, the perception threshold is an important parameter which quantifies the sensitivity and resolution of the system. There have been numerous studies of the motion perception threshold. Many of the earlier studies were carried out as part of a procedure to estimate the long time constant of the semicircular canals. This procedure, known as cupulometry, is described by Howard and Templeton (1966). Later, the threshold itself became an important result with direct applications to pilot modelling and flight simulation.
Webster's dictionary defines threshold as "the point at which a physiological or psychological effect begins to be produced." A more appropriate definition is "the signal amplitude at which a signal can be detected with a given reliability or accuracy by an observer." Traditionally, researchers have modelled the threshold as a non-linearity consistent with Webster's definition. Recently, many authors have recognized that the threshold is simply a statistical construct, which is more consistent with the second definition above. Because the threshold is a statistical construct, the values obtained are highly dependent on the experimental method and the detection accuracy required. This is discussed further in Chapter 3.

Threshold studies usually examine one of three parameters: nystagmus, the subjective sensation of motion, or the oculogyral illusion. The latter two are determined through subjective reports and are the most relevant to this study. Nystagmus refers to the compensatory eye movements which result from vestibular stimulation. The oculogyral illusion (OGI) is a visual illusion involving an apparent movement of objects in the visual field, which is caused by angular acceleration. The oculogravic illusion is a similar illusion caused by a rotation of the specific force vector relative to the observer. In both cases, the visual field is actually stationary with respect to the observer. These illusions are most prevalent for subjects fixating on an illuminated dot or cross in a dark environment. However, Huang and Young (1981) found that the OGI could also occur with a well-lit visual field.

2.4.1 Perception Threshold for Angular Motion

Most studies of motion perception thresholds report on subjective response to a step angular acceleration about a vertical axis with the head erect. Clark (1967) reviewed 25 early studies. The stimulus thresholds reported, quoted as angular accelerations, varied from 0.035 to 8.2°/s² with a median of about 1°/s². Most of the results were based on very few subjects, and Clark noted that the reports "constitute a miscellany of definitions of threshold, rotation devices, and psychophysical methods."

Following this review, Clark and Stewart carried out a series of investigations on the Man-carrying rotation device at NASA Ames (1968a, 1968b, 1969, 1970, 1972). These studies provide the most reliable source of threshold information, because a large number of subjects were tested, and because the experimental procedure was free from bias and consistent over the series of studies. Each experiment was run as a series of trials
controlled by a staircase algorithm. During each trial, the subject was exposed to a rectangular angular acceleration pulse with a duration of 10 seconds. After the end of the pulse, the subject indicated his direction of rotation. The amplitude of the pulse on subsequent trials was adjusted by a staircase procedure designed to converge on the level yielding 75% correct responses. All rotations were about an earth-vertical axis, therefore did not stimulate the otoliths.

Clark and Stewart (1968b) found that the threshold for the oculogyral illusion (OGI) is much lower than that for perception of motion in the dark. This finding was borne out by subsequent studies (1969, 1972) on over 100 subjects. The average threshold reported for the OGI is 0.11°/s²; that found for the perception of motion in the dark is about 0.4°/s². A study of 36 pilots and 56 non-pilots found no significant differences in threshold between the two groups (1972). A study of 18 subjects showed no significant differences in sensitivity to rotation about the 3 major head axes (1970). For this study, rotations were about a vertical axis with the head oriented accordingly. Finally, they noted that individual differences in sensitivity to motion are significant. Rodenburg et al (1981a) analyzed the distribution of thresholds reported by Clark and Stewart (1969). They found that, if the thresholds were expressed as dB, the distribution was normal with a standard deviation of 5.5 dB.

The values reported by Clark and Stewart are the minimum levels of acceleration perceptible after a 10 second pulse. For shorter duration signals, the threshold depends on stimulus duration. van Egmond et al (1949) measured the response latency to angular acceleration steps. Latency is the time delay between the start of the stimulus and the subject’s response. They found that the product of acceleration and latency time was about constant with a value of 1.5 to 2.0°/s. They proposed a mechanical explanation for the threshold, in which the cupular deflection must exceed some minimum value in order to produce a signal. The cupular displacement for a step acceleration input \( \alpha_H \), based on the torsion pendulum model (equation 2.3) is approximately:

\[
\delta(t) = H \alpha_H \left(1 - \exp\left(-t/\tau_L\right)\right)
\]

\[
= H \alpha_H \frac{t}{\tau_L} \quad \text{for } t \leq 0.5 \tau_L
\]

(2.12)
Thus, for a particular acceleration level $\alpha_H$, the minimum deflection is reached after a latency time $T_d$, where:

$$\alpha_H T_d = \text{constant} \quad (2.13)$$

This product is referred to in the literature as the Mulder product. Equation 2.13 implies that, because the semicircular canals act as integrating angular accelerometers, the threshold might be better expressed as a velocity threshold.

Gundry (1977) plots a summary of the results of 6 studies which shows that the Mulder product was fairly constant for latency times less than 10 seconds. He reports velocity thresholds ranging from 1.6 to 9.0°/s with a median of about 3°/s. Meiry (1966) quotes a value of 2.6°/s based on his own data and that of Clark and Stewart (1962). Huang and Young (1981) report values of 2.2 to 3.1°/s, and Rodenburg (1981b) found a value of 3.2°/s. Zacharias (1978) reinterprets the acceleration thresholds found by Clark and Stewart (discussed above), and obtains "equivalent" velocity thresholds between 2.5 and 4.0°/s. Young and Oman (1969) found that they could match Meiry's latency data by including a velocity threshold of 1.5°/s in their model following the canal dynamics.

Similar data are available for the oculogyral illusion (OGI). Doty (1969) studied the effect of stimulus duration on threshold for the OGI. For moderate durations, $\alpha_H T_d$ was about 0.4°/s. A similar value (0.37°/s) was reported by Clark et al (1980). Reinterpretation of Clark and Stewart's earlier data yields values in the range 0.6 to 1.1°/s. The difference in velocity thresholds for the OGI and for subjective sensation of motion is consistent with Clark and Stewart's findings noted above. Table 2.1 summarizes the available angular velocity threshold data.

All of the data presented above are for rotation about a vertical axis. Only four recent studies were found which investigated sensitivity to rotation about the horizontal. Kirkpatrick and Brye (1974) tested human sensitivity to pitch motion in a flight simulator. Subjects were required to determine which of two intervals contained a pitch acceleration pulse. The thresholds found, interpreted in terms of velocity, ranged from 1.6 to 3.0°/s, which is comparable to the values discussed above.

Gundry (1977) investigated sensitivity to a roll velocity pulse in a flight simulator, and reported a mean velocity threshold of 0.12°/s for 10 subjects. Hosman and van der Vaart (1978) tested the sensitivity of 3 general aviation pilots to sinusoidal pitch and roll in a
flight simulator. Subjects responded as soon as they could identify the axis of motion of a sinusoid with slowly increasing amplitude. The reported thresholds are shown in Table 2.2. For frequencies between 0.1 and 1.0 Hz, thresholds ranged from 0.30 to 0.37°/s for roll, and from 0.33 to 0.52°/s for pitch. Thresholds were slightly lower at higher frequencies.

The values reported by Gundry and by Hosman and van der Vaart are clearly much lower than the values found for rotation about a vertical axis. The most likely cause is stimulation of the otoliths. In Gundry’s study, the combined lateral specific force due to tilt and tangential acceleration had a peak of 0.004 g for threshold-level stimuli. This value is very close to the otolith threshold. In a subsequent study, Gundry (1978a) confirmed his earlier results by measuring response latency to rotational accelerations. From his data, he concludes that otolith stimulation contributes to the low thresholds found.

The results found by Hosman and van der Vaart may also be affected by otolith response. The amplitude of the sinusoidal specific force due to tilt and tangential acceleration was calculated assuming the subject’s head is 0.8 m above the axis of rotation. The results are shown in Table 2.2. The oscillatory force at threshold typically had a value of 0.003 to 0.007 g (higher at lower frequencies). These values are very near the otolith threshold.

Detection in this study may also have been aided by the oculogyral illusion (OGI). The authors do not describe the visual condition, but since a later portion of the study involved tracking a visual display, it is likely that the display was visible and may have led to the OGI. The thresholds shown in Table 2.2 are only slightly lower than those found by Doty (1969) and Clark et al (1980) for the OGI due to rotation about the vertical.

In summary, a number of authors have studied human sensitivity to rotation about a vertical axis. The data collapse best when expressed as a velocity threshold of about 2 to 4°/s. Thresholds for perception of the oculogyral illusion are lower by a factor of 4. Only four studies of thresholds for rotation about a horizontal axis were found. Two of these reported substantially lower thresholds than those found for rotation about the vertical, probably due to stimulation of the otoliths.
2.4.2 Perception Thresholds for Linear Motion

There is very little data available on thresholds of perception of translational motion or force. Much of the available data is based on tests of 1 to 3 subjects, using a wide variety of procedures. Consequently, the otolith threshold models are fairly tenuous. Peters (1969) reviewed 10 studies dating back to 1875, which reported thresholds ranging from 0.002 g to 0.027 g with a median of about 0.010 g. Gundry (1978b) updated this review and expanded it to include some studies of sensitivity to vibration between 1 and 100 Hz. Only two more recent studies were found.

Meiry (1966) measured response latency to horizontal step acceleration inputs. Three subjects were tested in two positions: first, seated with the head erect and the $X_H$ axis aligned with the direction of motion; and second, supine with the $Z_H$ axis aligned with the motion. He estimated the threshold of perception to be 0.006 g for motion along the $X_H$ axis, and 0.010 g for motion along the $Z_H$ axis. Young and Meiry (1968) found that their model could match the latency data if they assumed a threshold non-linearity following the otolith mechanics. Melville Jones and Young (1978) measured the response latency of 8 subjects for vertical acceleration, and estimated the threshold to be on the order of 0.005 g.

The response latency times for translational acceleration steps vary approximately inversely with the stimulus magnitude, $a_H$, that is:

$$a_H T_d = \text{constant}$$ \hspace{1cm} (2.14)

This indicates that a velocity threshold might be appropriate for translational motion. Melville Jones and Young (1978) analyzed their data this way and found a velocity threshold of 0.24 m/s for vertical motion. They also reanalyzed Meiry’s data and obtained a similar figure, 0.23 m/s, for horizontal motion. The only other data available for step or pulse accelerations are those obtained by Kirkpatrick and Brye (1974). Subjects were required to determine which of two intervals contained an acceleration pulse. The data show a fair bit of scatter, but estimates of the velocity threshold are 0.08 to 0.14 m/s for horizontal motion, and 0.11 to 0.44 m/s for vertical motion.

Gundry (1978b) reviewed 18 studies of human sensitivity to periodic linear motion. Many of these data are not relevant, either because the frequency range is not of interest, or because the subject oscillated relative to a stationary visual field, thus had visual cues.
to the motion. The most relevant data plotted by Gundry are those collected by Walsh (1961, 1962, 1964), Chen and Robertson (1972) and Gurnee (1934). The reported thresholds range from 0.002 to 0.020 g over the frequency range 0.067 to 0.67 Hz. The median value is about 0.006 g. The data do not show any significant differences in sensitivity between axes. Walsh's data show some frequency dependence, with thresholds at 0.33 Hz 20% lower, on average, than those at 0.11 Hz. A more recent study by Hosman and van der Vaart (1978) reports thresholds to vertical oscillation in a flight simulator ranging from 0.003 to 0.009 g for frequencies between 0.16 and 2.3 Hz. No significant frequency dependence was found.

Finally, two recent studies have reported subjective thresholds to tilt away from the vertical in a flight simulator. Hosman and van der Vaart (1978) report specific force thresholds due to tilt as 0.058 g for surge, and 0.032 g for sway. Roark and Junker (1978) report a threshold of 0.060 g for lateral force due to tilt. These values are an order of magnitude greater than the other threshold data discussed above. The reason for the discrepancy is not apparent.

In summary, very few measurements of translational motion perception thresholds have been made. However, three tenuous conclusions might be drawn. The threshold of perception for a long duration acceleration step is about 0.005 g. For shorter durations, a velocity threshold of about 0.24 m/s is more appropriate. Finally, the perception threshold for sinusoidal stimuli between 0.1 and 2.0 Hz is about 0.005 g.

2.4.3 Factors Influencing the Threshold

The threshold data discussed above were obtained for subjects in ideal laboratory conditions concentrating fully on the motion detection task. These "absolute" thresholds may have no relevance in the flight simulation environment. Gundry (1977) and Zacharias (1978) stress the need for research into "operational" thresholds more representative of pilot performance in the cockpit or simulator. Three factors which may influence sensitivity to motion are the visual field conditions, pilot workload, and background vibration levels. Each of these is discussed briefly below.
2.4.3.1 Effects of the Visual Field

There is a vast body of literature on the interaction of visual and vestibular cues. It is beyond the scope of this study to summarize this work. A good starting point for the interested reader is the review by Henn et al (1980).

The general consensus is that motion of the visual surroundings profoundly influences the perception of self-motion. The motion of the visual field with respect to the observer often tends to dominate the vestibular cues. Zacharias and Young (1981) propose a cue conflict model in which vestibular cues dominate sensation at higher frequencies, and visual cues dominate at lower frequencies. At frequencies for which visual cues dominate, movement of the visual field often decreases sensitivity to physical motion, even if the visual and motion cues are in the same direction. Sivan and Huang (1981), however, found that foveal field visual motion could enhance motion sensitivity. They related this finding to the oculogyral illusion.

In a recent study, Reid et al (1987) found that perception thresholds for sinusoidal motion were increased by a factor of up to 3 by the presence of sinusoidal visual motion on a wide angle display. The visual motion was fully congruent with the physical motion except for amplitude. The sensitivity change was largest at low frequency (0.2 Hz), and much smaller at higher frequencies (0.6 and 1.0 Hz). The visually-induced sensation of motion (vection) at 0.2 Hz was overwhelming, especially for the pitch and roll cases. It could be hypothesized that the motion detection task in this case became, perceptually if not factually, an increment detection task.

The available evidence indicates that perception threshold might be raised significantly by motion of the visual field. However, the models describing this interaction are still in flux. This remains a rich area for future study.

2.4.3.2 Effect of Pilot Workload

Several authors have studied the effect of pilot workload on motion perception thresholds. Usually this involves running a baseline study in which the subjects concentrate fully on the motion detection task, and then rerunning the experiment with the subjects performing a cognitive or control task. Kirkpatrick and Brye (1974) ran an experiment in which the subject’s detection task was to determine which of two time
intervals contained an acceleration pulse. They tested the effect of loading the subjects with an unstable steering task. The secondary task appeared to reduce sensitivity in certain degrees of freedom and certain conditions. However, an analysis of variance of all results collected did not show a significant effect.

Gundry (1977) obtained less equivocal results. He found that loading subjects with a cognitive task raised the roll threshold by 40%. The task involved mental arithmetic with aurally presented numbers. Thresholds increased monotonically with task difficulty. Gundry proposed an attenuation theory to account for the change in sensitivity. Hosman and van der Vaart (1978) tested the effect of two secondary tasks: a control task with motion cues, and an auditory binary choice task. The thresholds found increased 12 to 42% with the binary choice task alone, 25 to 85% with the tracking task alone, and 25 to 100% with both tasks.

Finally, Roark and Junker (1978) showed that the threshold of perception of lateral force due to tilt was increased by a factor of more than 2 when the pilot was engaged in a difficult roll tracking task with motion cues. They propose that a specific force "indifference" threshold of 0.1 g is appropriate for flight simulation applications. Their findings were corroborated by a brief exploratory experiment run as part of the current study.

The tasks used by Hosman and van der Vaart and by Roark and Junker both included motion cues. It is not clear what portion of the threshold shift was due to task loading, and what portion was due to masking of the motion cue to be detected by additional motion associated with the second task. Further work is needed to isolate these effects.

2.4.3.3 Effect of Background Motion and Vibration

Only four studies of the effect of background motion and vibration on motion perception thresholds were found. Gundry (1978b) cites a 1967 study by Brumaghim which assessed the perception thresholds for vertical oscillation between 4 and 10 Hz. Thresholds were typically 0.04 g. Extra vibration with RMS amplitude 0.24 to 0.43 g at 17 Hz had little effect of the thresholds. This suggests that vibration may have little masking effect. However, the frequencies examined in this study are well above those of interest for flight simulation.
More recently, a series of experiments was carried out at NASA Ames by Coleman (1979), Bury (1980) and Clark et al (1980). In all cases, the motion considered was rotation about a vertical axis. Coleman measured the choice reaction time (i.e. response latency) to constant angular acceleration in the presence of vibratory angular acceleration. He considered four different vibration conditions: no vibration, 2 Hz sinusoidal, 5 Hz sinusoidal, and 2 to 5 Hz quasirandom vibration. In each case the vibration had an RMS amplitude of $3^\circ/s^2$. Choice reaction times were measured for 10 subjects and 5 levels of constant angular acceleration between 1 and $12^\circ/s^2$. Coleman found no significant variation in performance across the 4 vibration conditions.

Bury repeated Coleman’s experiment with a few changes. The four vibration conditions considered were: no vibration, 1 Hz sinusoidal, 5 Hz sinusoidal, and 2 to 5 Hz quasirandom vibration. Bury considered three different (and higher) RMS vibration levels: 6, 8 and $10^\circ/s^2$. For the 1 Hz and quasirandom conditions, his data show an increase in reaction time (increased threshold) as the vibration amplitude increases. However, his analysis did not show this trend to be statistically significant.

Clark et al (1980) ran a similar experiment. They measured the threshold of perception for angular acceleration pulses of duration 0.5 and 1.0 seconds, in the presence of oscillatory angular acceleration. Three vibration conditions were considered: no vibration, $10^\circ/s^2$ RMS vibration at 1 Hz, and $10^\circ/s^2$ RMS vibration at 5 Hz. Their results show no significant difference across vibration conditions. The authors conclude that vibratory angular acceleration will not effectively mask constant angular acceleration in a flight simulator.

These results are somewhat at variance with those of Roark and Junker (1978), who attributed some of the sensory shift that they observed to masking. It appeared that more work was needed to resolve this issue. This provided the impetus for the current study.

2.5 The Signal-in-Noise Model

Most existing models of vestibular function represent the threshold as a non-linearity of the form:

\[
\begin{align*}
S_{\text{out}} &= 0 & \text{for } S_{\text{in}} < S_{\text{th}} \\
S_{\text{out}} &= S_{\text{in}} - S_{\text{th}} & \text{for } S_{\text{in}} > S_{\text{th}}
\end{align*}
\]
where $S_{in}$ and $S_{out}$ are the input and output signal amplitudes, respectively, and $S_{th}$ is the threshold. The origin of this model might be traced to the idea advanced by van Egmond et al (1949), that a minimum displacement of the cupula is required to produce a neural signal. This model is easy to understand, but is an inaccurate representation of the system.

The vestibular sensors are probably better modelled as a linear system with internal noise, as shown in Figure 2.9. This model is supported by neurological studies in primates by Goldberg and Fernandez (1971). They found that the relation between neural discharge rate and stimulus intensity provided no evidence for the existence of a mechanical or neural threshold. They also found considerable variability in the spontaneous (i.e. unstimulated) neural discharge rate, which could be described as noise. Some estimates of the noise levels were presented earlier in this chapter.

Several authors have noted that the signal-in-noise model is more appropriate than the threshold non-linearity model. Rodenburg et al (1981a) suggest that motion perception might be modelled as a signal detection process, but do not pursue the idea. Hosman and van der Vaart (1978) estimated the neural noise levels and reinterpreted their threshold findings as signal-to-noise ratios. The reported ratios are 0.46 to 1.0 for the semicircular canals (pitch and roll) and 2.5 to 5.0 for the otoliths (vertical).

Ormsby (1974) developed an optimal control model of the vestibular sensors. For suprathreshold perception, the model acted as a Kalman estimator which filtered the neural noise from the input signal. For near-threshold stimuli, the model compared the noisy input with an internal model of a typical movement, and made a statistical decision regarding the presence or absence of motion.

The signal-in-noise model seems particularly apt for the present study, where we are considering the effect of external noise (random motion) on subjective sensitivity to specific motion cues. Figure 2.9 shows the conceptual model for this case. It is proposed that motion perception can be analyzed using signal detection theory. This theory has proved useful in the study of auditory and electronic detection problems. Signal detection theory is discussed in Chapter 3.
<table>
<thead>
<tr>
<th>Source</th>
<th>Body Axis</th>
<th>Earth Axis</th>
<th>Visual Condition</th>
<th>Velocity Threshold (°/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>van Egmond et al (1949)</td>
<td>Z</td>
<td>Z</td>
<td>Dark</td>
<td>2.0</td>
</tr>
<tr>
<td>Meiry (1966)</td>
<td>Z</td>
<td>Z</td>
<td>Dark</td>
<td>2.6</td>
</tr>
<tr>
<td>Zacharias (1978) (note 1)</td>
<td>X,Y,Z</td>
<td>Z</td>
<td>Dark</td>
<td>2.5 - 4.0</td>
</tr>
<tr>
<td>Huang &amp; Young (1981)</td>
<td>Z</td>
<td>Z</td>
<td>Dark</td>
<td>2.2 - 3.1</td>
</tr>
<tr>
<td>Rodenburg et al (1981b)</td>
<td>Z</td>
<td>Z</td>
<td>Dark</td>
<td>3.2</td>
</tr>
<tr>
<td>Doty (1969)</td>
<td>Z</td>
<td>Z</td>
<td>OGI</td>
<td>0.4</td>
</tr>
<tr>
<td>Clark et al (1980)</td>
<td>Z</td>
<td>Z</td>
<td>OGI</td>
<td>0.37</td>
</tr>
<tr>
<td>Kirkpatrick &amp; Brye (1974)</td>
<td>Y</td>
<td>Y</td>
<td>Dim</td>
<td>1.3 - 3.0</td>
</tr>
<tr>
<td>Gundry (1977) (note 2)</td>
<td>X</td>
<td>X</td>
<td>n/a</td>
<td>0.12</td>
</tr>
<tr>
<td>Hosman &amp; van der Vaart (1978)  (note 2)</td>
<td>X</td>
<td>X</td>
<td>(note 3)</td>
<td>0.22-0.38</td>
</tr>
<tr>
<td>Hosman &amp; van der Vaart (1978)  (note 2)</td>
<td>Y</td>
<td>Y</td>
<td>(note 3)</td>
<td>0.22-0.53</td>
</tr>
</tbody>
</table>

Notes:  
2. Low values may be due to otolith stimulation.  
3. Visual condition not stated, but probably OGI.
Table 2.2  Perception thresholds for pitch and roll, based on Hosman & van der Vaart (1978). Also shown is the amplitude of the oscillatory specific force at the pilot’s head associated with the angular motion.

<table>
<thead>
<tr>
<th>Frequency (rad/s)</th>
<th>Roll</th>
<th>Pitch</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p (°/s)</td>
<td>Fy (m/s²)</td>
</tr>
<tr>
<td>0.6</td>
<td>.30</td>
<td>.088</td>
</tr>
<tr>
<td>0.8</td>
<td>.37</td>
<td>.084</td>
</tr>
<tr>
<td>1</td>
<td>.37</td>
<td>.069</td>
</tr>
<tr>
<td>2</td>
<td>.37</td>
<td>.042</td>
</tr>
<tr>
<td>4</td>
<td>.34</td>
<td>.033</td>
</tr>
<tr>
<td>6</td>
<td>.32</td>
<td>.036</td>
</tr>
<tr>
<td>8</td>
<td>.23</td>
<td>.030</td>
</tr>
<tr>
<td>10</td>
<td>.22</td>
<td>.035</td>
</tr>
<tr>
<td>12</td>
<td>.23</td>
<td>.041</td>
</tr>
<tr>
<td>14</td>
<td>.23</td>
<td>.047</td>
</tr>
</tbody>
</table>
Figure 2.1  Axes of the body-fixed reference frame, $F_H$. 

- **Y_H** Lateral Axis
- **X_H** Sagittal Axis
- **Z_H** Vertical Axis
Figure 2.2  The vestibular labyrinth (from Parker, 1980).
Figure 2.3  Top view of the head showing the approximate orientation of the semicircular canals (from Peters, 1969).
Figure 2.4a  Average dimensions of the horizontal semicircular canal. (from Igarashi, 1966)

Figure 2.4b  The structure of the ampulla and cupula (from Peters, 1969).
Figure 2.5  Response characteristics of Ormsby's semicircular canal model (equation 2.4).
Figure 2.6  The structure of the otolith organs (from Peters, 1969).
Figure 2.7  Response characteristics of Young and Meiry's revised dynamic otolith model (equation 2.10).
Figure 2.8 Response characteristics of Ormsby's otolith model (equation 2.11).
Figure 2.9  Signal-in-noise model for motion perception.
Chapter 3

An Introduction to Signal Detection Theory

3.1 Introduction

In Chapter 2 it was suggested that perception of low amplitude motion cues might be considered a signal-in-noise detection problem. This representation certainly seems valid for the current study, in which we are assessing the effect of background motion on the perception of motion cues. The theory which has evolved to tackle this type of problem is known as signal detection theory. This chapter provides a brief review of signal detection theory. For a comprehensive treatment of the theory, see the books by Green and Swets (1974) and Egan (1975).

Signal detection theory was developed by mathematicians and engineers in the early 1950’s. The earliest applications were in the analysis and design of electronic signal detection equipment such as radar receivers. The theory was adapted to model human auditory detection by Marill (1956), Tanner and Birdsall (1958), Swets et al (1961), Pfafflin and Mathews (1962), Jeffress (1964), and many others. A collection of some of the most relevant early works was compiled by Swets (1964). Signal detection theory has since been applied to other perceptual modalities, especially vision, and to a wide variety of other fields. Swets and Pickett (1982) cite applications in industrial quality control, lie detection, information science, and medical diagnostics.

To date, however, there has been no rigorous attempt to apply signal detection theory to motion perception. Kirkpatrick and Brye (1974) state that motion perception is best modelled by signal detection theory, but do not demonstrate this or attempt to define a specific detection model. Ormsby (1974) formulated a stochastic model for motion perception which has some similarities to the signal detection model. However, he does not consider the case where the observer is trying to detect motion cues in a background of random motion. Miwa et al (1984) and others model human sensitivity to whole body vibration (between 1 Hz and 64 Hz) with an energy detector, but do not concern themselves with the stochastic aspects of the detector.
The form of the theory used in this study is the same as that used to model auditory detection. It is assumed that the human observer can be modelled by a two-stage detector of the form shown in Figure 3.1. The observer's task is to detect "signals" on a sensory input channel. The input consists of an information stream, which may or may not contain a signal, contaminated by random noise. The noise may be either internal (i.e. neural noise) or external. In the present case, the signal is a specific motion cue, and the external noise is random vibration of the simulator cab.

In its simplest form, the detector consists of two components. The first stage, a signal processor, samples the input over a time interval, T, and calculates a scalar statistic based on the sample. We will refer to this statistic as the evidence variable, e. The second stage must decide, on the basis of the obtained value of e, whether or not a signal was present during the sampling interval. Because e is a random variable, this is a statistical decision.

Signal detection theory thus has its roots in two fields: stochastic signal processing, and statistical decision theory. Section 3.2 reviews the basic concepts of decision theory as it applies to signal detection. Section 3.3 describes how an observer's performance can be summarized by the receiver operating characteristic, or ROC curve. Some simple performance measures based on the ROC curve are introduced. Finally, Section 3.4 presents several alternatives for the signal processing component, and discusses the detector performance for each case. Because performance is highly dependent on the form of the signal processor, determining that form is a major aim of this study.

With the exception of one or two equations, none of the material presented in this chapter is new. It is a review, and the coverage of some topics may seem a bit brief. The treatment here is a summary of that presented by Green and Swets (1974). The reader is referred to the original work for a more comprehensive presentation.

**3.2 Statistical Decision Theory**

The discussion in this chapter is restricted to the simple scenario outlined above, in which the observer's task is to determine the presence or absence of a signal during a fixed time interval. This accurately describes the situation in most laboratory studies of detection, and is a good first approximation to more general "real world" detection tasks.
The detection problem contains the three elements common to every decision problem:

1. two possible states of the world,
2. information about the state of the world, and
3. two possible responses, or decision outcomes.

The two states of the world are the two stimulus conditions: s, the stimulus was signal plus noise; and n, the stimulus was noise alone. The two possible responses are S (or YES), accept hypothesis s; and N (or NO), accept hypothesis n.

The information which serves as the basis for the decision is the value of the evidence variable, e. Because the input consists partly of random noise, e is a random variable. Given the characteristics of signal and noise, we can calculate two conditional probability density functions. These are \( f(e \mid s) \), the probability density of e given that the input was signal plus noise, and \( f(e \mid n) \), the density of e given that the input was noise alone.

The detector must therefore make a statistical decision: given a specific value of e, was there or was there not a signal present? This decision is made by following a rule. If \( e \) falls within some given set, answer YES, there was a signal; otherwise answer NO. The simplest form of decision rule is shown in Figure 3.2. Here, the detector compares \( e \) to a criterion "k", then answers YES if \( e \geq k \), or NO if \( e < k \). The four possible combinations of stimulus and response may be summarized in the "stimulus-response matrix" shown in Figure 3.3.

The detector performance may be described statistically. Assume that the a priori probabilities of states s and n, denoted \( P(s) \) and \( P(n) \), are constant. Also assume that the characteristics of signal and noise are stationary, and that the decision rule is invariant. Then the detector performance may be summarized by the following four conditional probabilities:

- \( P(S \mid s) \), the hit rate,
- \( P(S \mid n) \), the false alarm rate,
- \( P(N \mid s) \), the miss rate, and
- \( P(N \mid n) \), the true-negative rate,

where \( P(X \mid y) \) is the probability of responding X when the stimulus is y. Only two of
these probabilities are independent, since
\[ P(N \mid s) = 1 - P(S \mid s) \]
and \[ P(N \mid n) = 1 - P(S \mid n) \] (3.1)

Clearly, hits and true negatives are correct decisions, while misses and false alarms are errors. If \( n \) is defined to be the null hypothesis, then false alarms and misses are simply Type I and Type II errors, respectively. The observer's goal is to optimize his performance by maximizing correct decisions and/or minimizing errors.

Given that the only information available is the value of \( e \), how does the observer make an optimal decision? A rational decision rule might be one based on the \textit{a posteriori} probability of each hypothesis, \( P(s \mid e) \) and \( P(n \mid e) \). One obvious decision rule is to respond \( S \) if and only if \( P(s \mid e) \geq \beta P(n \mid e) \), for some value of \( \beta \). This rule can also be expressed: respond \( S \) if and only if \( P(s \mid e) / P(n \mid e) \geq \beta \). This observation leads to the concept of likelihood ratio. According to Bayes' rule, the \textit{a posteriori} probabilities given \( e = e_k \) are:

\[ P(s \mid e_k) = \frac{P(s) f(e_k \mid s)}{f(e_k)} \] (3.2)
\[ P(n \mid e_k) = \frac{P(n) f(e_k \mid n)}{f(e_k)} \] (3.3)

The ratio of the \textit{a posteriori} probabilities is:

\[ \frac{P(s \mid e_k)}{P(n \mid e_k)} = \frac{P(s)}{P(n)} \cdot L(e_k) \] (3.4)

where \( L(e_k) = \frac{f(e_k \mid s)}{f(e_k \mid n)} \) (3.5)

The likelihood ratio, \( L(e_k) \), is therefore a convenient summary of the information provided by \( e_k \).
It may be shown that a decision rule based on likelihood ratio is optimal for a number of different decision goals. Ideally, we would like to maximize $P(S \mid s)$ and minimize $P(S \mid n)$. However, this is usually not possible. A weaker, but feasible, goal is to maximize a weighted combination:

$$J = P(S \mid s) - \beta P(S \mid n)$$  

(3.6)

The decision rule is to respond $S$ if and only if $e \in M$, where $M$ is an unknown set to be defined. Then:

$$J = \int_{M} f(e \mid s) \, de - \beta \int_{M} f(e \mid n) \, de$$

$$= \int_{M} f(e \mid n) [L(e) - \beta] \, de$$  

(3.7)

Therefore $J$ is maximized if $M$ contains those values of $e$, and only those values, for which $L(e) \geq \beta$. The optimal decision rule is simply: respond $S$ if and only if $L(e) \geq \beta$.

Consider another example of decision goal. Suppose that each correct or incorrect decision has an associated value or cost, and that the decision goal is to maximize the expected payoff. The expected value of a decision may be expressed as:

$$V^* = V_{sS} P(s) P(S \mid s) + V_{sN} P(s) P(N \mid s)$$

$$+ V_{nS} P(n) P(S \mid n) + V_{nN} P(n) P(N \mid n)$$  

(3.9)

where $V_{xy}$ is the value associated with responding $Y$ when the stimulus is $x$. Combining equations 3.1 and 3.9 yields:

$$V^* = V_{sN} P(s) + V_{nN} P(n) + (V_{sS} - V_{sN}) P(s) P(S \mid s)$$

$$- (V_{nN} - V_{nS}) P(n) P(S \mid n)$$  

(3.10)

All values in equation 3.10 are constant except $P(S \mid s)$ and $P(S \mid n)$. It is clear by inspection that $V^*$ is maximized by maximizing

$$J = P(S \mid s) - \beta P(S \mid n)$$
where \( \beta = \frac{P(n) (V_{nN} - V_{nS})}{P(s) (V_{sS} - V_{sN})} \) \hspace{1cm} (3.11)

As shown above, this may be accomplished by using a decision rule based on likelihood ratio. Neyman and Pearson (1933) showed that a decision rule based on \( L(e) \) is also optimal when the decision goal is to maximize \( P(S \mid s) \) while holding \( P(S \mid n) \) constant.

In summary, the observer's task is to detect signals on a noisy input channel. This involves a statistical decision based on a scalar evidence variable, \( e \). For most common decision goals, the optimal decision rule is one based on the likelihood ratio, \( L(e) \). It should be noted that if \( L(e) \) is monotonic with \( e \), then the rule "respond \( S \) if and only if \( e \geq k \)" is equivalent to the rule "respond \( S \) if and only if \( L(e) \geq \beta \), where \( \beta = L(k) \)." Therefore an optimal decision rule can be formulated based on any function of \( e \) which is monotonic with \( L(e) \). This observation has important consequences in the discussion of ideal detectors.

### 3.3 The ROC Curve

In Section 3.2 we showed that the response characteristics of a detector can be summarized by two conditional probabilities, \( P(S \mid s) \) and \( P(S \mid n) \). This can be shown graphically by plotting \( P(S \mid s) \) versus \( P(S \mid n) \). This graph, known as a receiver operating characteristic (ROC) graph, provides a vehicle for summarizing the overall detector performance.

For example, assume that \( e \) is distributed according to the conditional distributions shown in Figure 3.4a. Also assume that the observer's decision rule is to respond \( S \) if and only if \( e \geq k \), for some criterion \( k \). Note that, in general, \( e \) may or may not be monotonic with \( L(e) \). If it is, then the decision rule is optimal. Then:

\[
P_k(S \mid s) = \int_k^\infty f(e \mid s) \, de \hspace{1cm} (3.12)
\]

\[
P_k(S \mid n) = \int_k^\infty f(e \mid n) \, de \hspace{1cm} (3.13)
\]
When $k = k_1$, these probabilities define point $k_1$ on the ROC graph in Figure 3.4b. Now, suppose that the observer is induced to change his decision criterion, perhaps by changing the payoffs and penalties for correct and incorrect responses. Assume that the new criterion is $k_2$. Both $P(S|s)$ and $P(S|n)$ change, defining a new point, $k_2$, on the ROC graph. Similarly, the use of criterion $k_3$ yields a third point on the ROC graph. Varying the criterion $k$ continuously over the range $(-\infty, \infty)$ defines a smooth curve, known as the ROC curve, which summarizes the detector’s performance.

Note that the ROC curve describes the detector performance for one particular set of stimulus conditions. Only the observer’s criterion, or bias, changes from one point on the curve to the next. Green and Swets (1974) note that human observers can be induced to adjust their response criterion to meet verbal specifications or to optimize performance in terms of some decision goal. In other words, human observers do have the ability to operate at different points along the ROC curve.

Several characteristics of the ROC curve are worth noting. First, it is clear from equations 3.12 and 3.13 that $P(S|s)$ increases monotonically with $P(S|n)$. Second, it may be seen by differentiating equations 3.12 and 3.13 with respect to $k$ that $dP_k(S|s)/dP_k(S|n) = f(k|s)/f(k|n)$. In other words, the slope of the ROC curve at the point defined by $k$ is simply the likelihood ratio, $L(k)$. It follows that if $L(e)$ is monotonic with $e$, then the slope of the ROC curve decreases monotonically with $P(S|s)$ and $P(S|n)$. That is, the ROC curve is concave downward.

In studying detection, we need a measure of the signal detectability which is independent of the observer’s bias. The hit rate, $P(S|s)$, has often been used in the past as a measure of detectability. However, this is an inadequate measure, since it may vary from 0 to 1, depending on the observer’s bias. To make any sense of $P(S|s)$, we must also know the corresponding value of $P(S|n)$. The area under the ROC curve, $P_A$, summarizes the relative values of $P(S|s)$ and $P(S|n)$ over all possible values of the decision criterion. Therefore, it is a suitable measure of signal detectability. $P_A$ may be estimated directly by some experimental procedures. As shown in Chapter 6, $P_A$ is equal to the fraction of correct responses in a two-alternative forced choice detection task.

Hypothetical ROC curves for two different signal strengths are shown in Figure 3.5. The lower curve corresponds to the weaker signal. As the signal strength increases, the distance between the distributions $f(e|s)$ and $f(e|n)$ increases, and the ROC curve moves...
upward. If the signal is completely indistinguishable from the noise, then $P(S | s) = P(S | n)$, and the ROC follows the diagonal. If the signal is very strong, the detector can achieve 100% correct detections with no false alarms. In this case, the ROC follows the upper boundary of the unit square. Thus the area under the ROC curve, $P_A$, gives an indication of overall performance. It ranges from 0.5 for chance detection to 1.0 for perfect discrimination.

A related measure of detectability can be defined directly from the conditional distributions, $f(e | s)$ and $f(e | n)$. Assume that $f(e | s)$ and $f(e | n)$ are Gaussian distributions with means $\mu_s$ and $\mu_n$, respectively, and equal variances $\sigma_s^2 = \sigma_n^2$. The index of detectability, $d'$, is defined:

\[ d' = \frac{\mu_s - \mu_n}{\sigma_n} \]  

and increases monotonically with $P_A$. For $d' = 0$, $P_A = 0.5$. As $d'$ becomes very large, $P_A$ approaches 1.0. A commonly used benchmark, $d' = 1.0$, corresponds to $P_A = 0.76$. A similar but more general index may be defined for the case when $\sigma_s \neq \sigma_n$:

\[ d'_A = \frac{\mu_s - \mu_n}{\sqrt{\sigma_s^2 + \sigma_n^2}/2} \]  

Assuming that $f(e | s)$ and $f(e | n)$ are Gaussian, the value of $P_A$ associated with a particular value of $d'_A$ is independent of the ratio $\sigma_n/\sigma_s$.

While the area under the ROC curve gives an indication of the relative distance between $f(e | s)$ and $f(e | n)$, the shape of the ROC yields some information about the shape of the underlying distributions. The curves shown in Figure 3.5 are based on Gaussian distributions with equal variance. These ROC curves are symmetrical about the negative diagonal. If the variances are unequal, then the ROC curve is skewed, as shown in Figure 3.6.

If $f(e | s)$ and $f(e | n)$ are both Gaussian, the resulting ROC curve is a straight line when plotted on normal probability axes. The ROC curve shown in Figure 3.6 is plotted on probability axes in Figure 3.7. The slope of the line, $b$, summarizes the asymmetry of the ROC curve. In this report, $b$ is referred to as the ROC slope parameter. Assuming that
both distributions are Gaussian, \( b \) is simply the ratio of the standard deviations:

\[
b = \frac{\sigma_n}{\sigma_s}
\]  

(3.16)

Further examples showing how the shapes of the underlying distributions influence the ROC curve are given by Egan (1975).

### 3.4 Ideal Detectors

The previous sections discussed the elements of decision theory and showed how a detector’s performance can be summarized by the ROC curve. Clearly, the ROC curve is simply a representation of the conditional distributions \( f(e \mid s) \) and \( f(e \mid n) \). These, in turn, depend on the characteristics of signal and noise and the form of the signal processor. One of the chief aims of this study is to determine how the signals are processed, because that governs the overall performance of the detector.

In this section, we consider three different formulations of the processor and show the implications for \( f(e \mid s) \) and \( f(e \mid n) \), and hence for the ROC curve. In each case, we assume that the observer performs optimally given the limits of his knowledge of the signal waveform. As shown in Section 3.2, a decision rule based on likelihood ratio yields the optimum performance. Therefore, we assume that the observer is able to calculate the likelihood ratio based on an observation, or at least is able to evaluate observations on a scale which is monotonic with likelihood ratio.

Let us summarize the detection problem. First, the detector observes a waveform, \( x(t) \), over a finite time interval, 0 to \( T \). It must then decide which of two hypotheses to accept. The waveform is either \( s \), signal plus noise, or \( n \), noise alone. The decision is based on a statistic \( z[x(t)] \) calculated by the signal processor. Note that \( z \) is simply an evidence variable which is monotonic with likelihood ratio. Evaluating the performance of an ideal detector involves three steps:

1. Determine the conditional densities \( f[x(t) \mid s] \) and \( f[x(t) \mid n] \).

2. Determine the likelihood function \( L[x(t)] \) and a function \( z[x(t)] \) which is monotonic with \( L[x(t)] \).

3. Determine the ROC curve by integrating \( f(z \mid s) \) and \( f(z \mid n) \).
Steps 1 and 2 show how the waveform should be processed. Step 3 relates the detector performance to the physical characteristics of signal and noise.

This section is divided into five parts. The first outlines a discrete representation of signal and noise which serves as the basis of the analysis. The next three present the derivation of the ideal detector for three different situations. Three different detectors are derived: a correlation detector, an envelope detector, and an energy detector. All three have been applied to auditory detection and masking: the correlation detector, by Tanner and Birdsay (1958); the envelope detector, by Jeffress (1964); and the energy detector, by Pfafflin and Mathews (1962). The final subsection compares the performance of the three detectors.

The presentation of the theory in Sections 3.4.1 through 3.4.4 is somewhat tedious; some readers may prefer to proceed directly to the discussion in Section 3.4.5. The main results are given in equation 3.36 for the correlation detector, equations 3.52 and 3.53 for the envelope detector, and equations 3.66 to 3.68 for the energy detector. The main conclusion is that performance is degraded as the detector's uncertainty about the signal waveform increases.

### 3.4.1 Discrete Representation of Signal and Noise

The theory of ideal detectors can be developed for continuous signals, but the analysis is very complicated. The derivation of the theory is simplified considerably by using a discrete representation of the signal and noise waveforms. The general concepts are demonstrated more clearly by this approximate analysis. The use of a discrete-time representation of the stimuli does not imply a sampling process in the human observer.

Assume that the noise is band-limited Gaussian noise, with power density $N_0$ and bandwidth $W$ Hz. The assumed rectangular noise spectrum is shown in Figure 3.8. Also assume that the observation time, $T$, is chosen such that the product $WT$ is an integer. The analysis presented here is for the case where the lower limiting frequency of the noise, $W_L$, is zero. The theory may be extended to the case where $W_L$ is an integer multiple of $W$, simply by changing the limits on the summation in equation 3.17. Extension to more arbitrary bands is more complex. For the case where $W_L = 0$, the
noise waveform on \((0,T)\) can be approximated by a finite Fourier series:

\[
    r(t) = \sum_{k=0}^{\text{WT}} \left[ a_k \cos \frac{2\pi kt}{T} + b_k \sin \frac{2\pi kt}{T} \right]
\]  

(3.17)

where \(a_k\) and \(b_k\) are independent Gaussian variables with zero mean and variance \(\sigma_k^2 = N_o/T\), for \(k = 1\) to \(\text{WT} - 1\). The coefficients \(a_0, b_0, a_{\text{WT}},\) and \(b_{\text{WT}}\) (which multiply the sinusoidal components at the limiting frequencies of the noise band) are independent Gaussian variables with zero mean and variance \(N_o/(2T)\). The expected value of the noise power is:

\[
    \mathbb{E} \left[ \frac{1}{T} \int_0^T [r(t)]^2 \, dt \right] = N_oW
\]  

(3.18)

We may also represent \(r(t)\) as a vector of discrete-time samples. According to the Nyquist sampling theorem, all of the information in \(r(t)\) can be represented by a vector of \(2\times WT\) discrete samples of \(r(t)\), where the time increment between samples is \(1/(2W)\). Let \(\mathbf{r}\) denote this vector. Then

\[
    \mathbf{r} = \{r_m\}
\]  

(3.19)

where \(r_m = r(m/(2W))\). According to equation 3.17, \(r_m\) is a linear function of the Gaussian variables \(a_k\) and \(b_k\). Therefore \(r_m\) is Gaussian. It is readily demonstrated using the Fourier series representation of \(r(t)\) that:

\[
    \mathbb{E}[r_m] = 0 \quad \text{for all } m
\]  

(3.20)

\[
    \mathbb{E}[r_m^2] = N_oW \quad \text{for all } m
\]  

(3.21)

\[
    \mathbb{E}[r_mr_n] = 0 \quad \text{for } m \neq n
\]  

(3.22)

Thus the \(2\times WT\) samples of the noise waveform, \(r_m = r(m/(2W))\), are independent Gaussian variables with zero mean and variance \(\sigma_r^2 = N_oW\).
We further assume that the signal can be represented in the same way as the noise. The two representations are:

\[ s(t) = \sum_{k=0}^{WT} \left[ p_k \cos \frac{2\pi k t}{T} + q_k \sin \frac{2\pi k t}{T} \right] \quad (3.23) \]

\[ s = \{s_m\} \quad (3.24) \]

where \( s_m = s(m/(2W)) \). The signal energy is:

\[ E_s = \int_0^T [s(t)]^2 \, dt = \frac{1}{2} \frac{2WT}{W} \sum_{m=1}^{2WT} s_m^2 \quad (3.25) \]

### 3.4.2 Correlation Detector

The first situation we will consider is one in which the signal is defined exactly. For this analysis, we use the discrete-time representation of signal and noise. We assume that the detector knows in advance the exact signal waveform, \( s(t) \), or the value of \( s_m = s(m/(2W)) \) at each of the \( 2WT \) sample points. The detector uses the information in \( x(t) \) and its advance knowledge of \( s(t) \) as the basis for its decision. We assume that the ideal detector bases its decision on a scalar function \( z[x(t)] \), which is monotonic with likelihood ratio.

Let \( x = \{x_m\} \) be the temporal representation of the input waveform, \( x(t) \). For each sample, \( x_m \), the conditional distributions are both Gaussian. If the input is noise alone, \( x_m = r_m \), and:

\[ f(x_m | n) = \frac{1}{\sqrt{2\pi N_o W}} \exp \left( -\frac{x_m^2}{2N_o W} \right) \quad (3.26) \]

If the input is signal plus noise, \( x_m = s_m + r_m \), so:

\[ f(x_m | s) = \frac{1}{\sqrt{2\pi N_o W}} \exp \left( -\frac{(x_m-s_m)^2}{2N_o W} \right) \quad (3.27) \]
3.13

Because $r_m$ and $r_n$ are independent for $m \neq n$, the conditional densities for the entire waveform may be written as:

$$f[x(t) \mid n] = f(x \mid n) = \prod_{m=1}^{2WT} f(x_m \mid n)$$

$$f[x(t) \mid s] = f(x \mid s) = \prod_{m=1}^{2WT} f(x_m \mid s)$$

The likelihood ratio can be found by combining equations 3.26 through 3.29:

$$L[x(t)] = \frac{f[x(t) \mid s]}{f[x(t) \mid n]} = \exp \left( \frac{\sum 2x_m s_m - \sum s_m^2}{2 N_o W} \right)$$

Because $\sum s_m^2$ is constant, the likelihood ratio is monotonic with an "evidence variable" $z[x(t)]$, where:

$$z[x(t)] = \sum_{m=1}^{2WT} x_m s_m$$

Note that $z[x(t)]$ is essentially the correlation of $x(t)$ with $s(t)$. Because $z$ is monotonic with the likelihood ratio, a decision rule of the form "respond S if and only if $z \geq k$" is an optimal decision rule. Therefore, if the signal waveform is known exactly in advance, the optimal detector uses the correlation of the observed waveform with the expected signal waveform as the basis for its decision.

To evaluate the detector performance, we must determine the conditional densities $f(z \mid n)$ and $f(z \mid s)$. Since $z$ is a linear function of the Gaussian variables $x_m$, $z$ is Gaussian. If the input is noise alone,

$$\mu_n = E[\sum_{m=1}^{2WT} r_m s_m] = 0$$

$$\sigma_n^2 = E[ (\sum_{m=1}^{2WT} r_m s_m)^2 ] = N_o W \sum_{m=1}^{2WT} s_m^2 = 2 N_o W^2 E_s$$
If the input is signal plus noise,

\[
\mu_s = E\left[ \sum_{m=1}^{2WT} s_m (r_m + s_m) \right] = \sum_{m=1}^{2WT} s_m^2 = 2WE_s
\]  

(3.34)

\[
\sigma_s^2 = E\left[ \left( \sum_{m=1}^{2WT} s_m (r_m + s_m) - \sum_{m=1}^{2WT} s_m^2 \right)^2 \right]
\]

\[
= E\left[ \sum_{m=1}^{2WT} r_m s_m \right]^2 = \sigma_n^2
\]  

(3.35)

Therefore we have derived the conditional densities \( f(z|s) \) and \( f(z|n) \) in terms of the physical parameters of the signal and noise. The distributions are both Gaussian and have equal variance. Therefore the ROC curves associated with this detector are symmetric, like those shown in Figure 3.5. The index of detectability, \( d' \), for this case is:

\[
d' = \frac{\mu_s - \mu_n}{\sigma_n} = \sqrt{\frac{2E_s}{N_0}}
\]  

(3.36)

This shows that detectability is a function of the ratio of signal energy to noise power density, and is independent of the noise bandwidth. For a sinusoidal signal, \( d' \) is proportional to the signal amplitude.

3.4.3 Envelope Detector

The second case we will consider is one in which the signal is defined statistically. Specifically, assume that the signal is a sinusoid of known frequency but unknown phase, \( \phi \):

\[
s(t,\phi) = A \cos \left( 2\pi mt - \phi \right)
\]

\[
= A \cos \phi \cos \frac{2\pi mt}{T} + A \sin \phi \sin \frac{2\pi mt}{T}
\]  

(3.37)
Assume that \( \phi \) is a uniformly distributed random variable with probability density:

\[
p(\phi) = \frac{1}{2\pi} \quad \text{for} \ 0 \leq \phi \leq 2\pi \quad (3.38)
\]

For the analysis of this case, we use the Fourier series representation for the observed waveform:

\[
x(t) = \sum_{k=0}^{\text{WT}} \left[ g_k \cos \frac{2\pi kt}{T} + h_k \sin \frac{2\pi kt}{T} \right] \quad (3.39)
\]

If the observed waveform is noise alone, then \( x(t) = r(t) \), and \( g_k = a_k \) and \( h_k = b_k \), where \( a_k \) and \( b_k \) are independent Gaussian variables with zero mean and variance \( \sigma_k^2 = N_o/T \) (or \( N_o/(2T) \) for \( k = 0 \) or \( k = \text{WT} \)). Therefore,

\[
f[x(t) \mid n] = \prod_{k=0}^{\text{WT}} f(g_k \mid n) f(h_k \mid n) \quad (3.40)
\]

where

\[
f(g_k \mid n) = \frac{1}{\sqrt{2\pi} \sigma_k} \exp \left( -\frac{g_k^2}{2\sigma_k^2} \right) \quad (3.41a)
\]

and

\[
f(h_k \mid n) = \frac{1}{\sqrt{2\pi} \sigma_k} \exp \left( -\frac{h_k^2}{2\sigma_k^2} \right) \quad (3.41b)
\]

If the input is signal plus noise, then \( x(t) = r(t) + s(t, \phi) \). Therefore, \( x(t) - s(t, \phi) \) must have the same distribution as \( r(t) \). Combining equations 3.37 and 3.39 gives the following:

\[
x(t) - s(t) = \sum_{k=0}^{\text{WT}} \left[ g_k \cos \frac{2\pi kt}{T} + h_k \sin \frac{2\pi kt}{T} \right] \\
+ (g_m - A \cos \phi) \cos \frac{2\pi mt}{T} + (h_m - A \sin \phi) \sin \frac{2\pi mt}{T} \quad (3.42)
\]

Since this is simply \( r(t) \), \( g_k \) and \( h_k \) (for \( k \neq m \)) and \( (g_m - A \cos \phi) \) and \( (h_m - A \sin \phi) \) are independent Gaussian variables with zero mean and variance \( \sigma_k^2 \). It follows that
\[ f(g_k \mid s, \phi) = f(g_k \mid n) \text{ and } f(h_k \mid s, \phi) = f(h_k \mid n) \text{ for } k \neq m. \] Also, \( f(g_m \mid s, \phi) = f(g_m - \text{Acos} \phi \mid n) \) and \( f(h_m \mid s, \phi) = f(h_m - \text{Asin} \phi \mid n). \) Therefore, \( f[x(t) \mid s, \phi] \) may be expressed:

\[
\begin{align*}
    f[x(t) \mid s, \phi] &= \prod_{k=0}^{\infty} f(g_k \mid s, \phi) f(h_k \mid s, \phi) \\
    &= \left( \prod_{k=0}^{\infty} f(g_k \mid n) f(h_k \mid n) \right) f(g_m - \text{Acos} \phi \mid n) f(h_m - \text{Asin} \phi \mid n)
\end{align*}
\]

(3.43)

Integrating over all values of \( \phi \), we obtain:

\[
\begin{align*}
    f[x(t) \mid s] &= \prod_{k=0}^{\infty} f(g_k \mid n) f(h_k \mid n) \int_0^{2\pi} f(g_m - \text{Acos} \phi \mid n) f(h_m - \text{Asin} \phi \mid n) \frac{d\phi}{2\pi}
\end{align*}
\]

(3.44)

The likelihood ratio is obtained by dividing \( f[x(t) \mid s] \) by \( f[x(t) \mid n] \). Combining equations 3.40 and 3.44 and simplifying:

\[
\begin{align*}
    L[x(t)] &= \frac{1}{2\pi} \int_0^{2\pi} \exp \left( -\frac{T}{2N_o} \left( (g_m - \text{Acos} \phi)^2 + (h_m - \text{Asin} \phi)^2 - g_m^2 - h_m^2 \right) \right) d\phi \\
    &= \exp \left( -\frac{A^2 T}{2N_o} \right) \frac{1}{2\pi} \int_0^{2\pi} \exp \left( -\frac{c_m^2 A T}{N_o} \cos(\phi - \gamma) \right) d\phi 
\end{align*}
\]

(3.45)

where \( c_m \) and \( \gamma \) are the amplitude and phase of the observed sinusoid at the \( m \)th frequency, that is:

\[
\begin{align*}
    g_m &= c_m \cos \gamma \\
    h_m &= c_m \sin \gamma
\end{align*}
\]

(3.46)

Equation 3.45 can be simplified by noting that the modified Bessel function of order zero is defined:

\[
I_0(x) = \frac{1}{2\pi} \int_0^{2\pi} \exp(x \cos \phi) \, d\phi
\]

(3.47)
Therefore the likelihood ratio of \( x(t) \) is:

\[
L[x(t)] = I_o \left( \frac{c_m AT}{N_o} \right) \exp \left( -\frac{A^2 T}{2N_o} \right)
\]

(3.48)

\( I_o(x) \) increases monotonically with \( x \), therefore the likelihood ratio is monotonic with \( c_m \). A decision rule of the form "respond S if and only if \( c_m \geq k \)" is an optimal decision rule. Thus if the signal frequency is known but the phase is not, the optimal detector bases its decision on \( c_m \), the amplitude of the observed waveform at the signal frequency. Davenport and Root (1958) showed that \( c_m \) is equivalent to the envelope of the output of a narrow band filter centred at the signal frequency.

It is convenient at this point to define the following nondimensional amplitudes:

\[
c_* = \frac{c_m}{\sigma_k}
\]

\[
A_* = \frac{A}{\sigma_k}
\]

(3.49)

where \( \sigma_k^2 = N_o/T \). The likelihood ratio, expressed in terms of these variables, is:

\[
L[x(t)] = I_o(c_* A_*) \exp(-A_*^2/2)
\]

(3.50)

To determine the detector performance, we must find the distribution of \( c_* \) under the two hypotheses, \( s \) and \( n \). If the input is noise alone, then \( g_m \) and \( h_m \) are independent Gaussian variables with variance \( \sigma_k^2 \). Therefore:

\[
c_*^2 = (g_m/\sigma_k)^2 + (h_m/\sigma_k)^2
\]

(3.51)

is a \( \chi^2 \) variable with two degrees of freedom. It follows that:

\[
f(c_* | n) = f(c_*^2 | n) \frac{d}{dc_*}(c_*^2)
\]

\[
= c_* \exp(-c_*^2/2)
\]

(3.52)

This distribution, known as the Rayleigh distribution, is shown by the solid curve in Figure 3.9.
Derivation of \( f(c_\star \mid s) \) is straightforward. We know that \( f(c_\star \mid s) = L(c_\star) f(c_\star \mid n) \). Also, because \( L[x(t)] \) depends only on \( c_\star \), \( L(c_\star) = L[x(t)] \). Therefore:

\[
f(c_\star \mid s) = c_\star I_0(c_\star A_\star) \exp\left[-\frac{(c_\star^2 + A_\star^2)}{2}\right]
\]  

(3.53)

Some examples of this distribution are shown in Figure 3.9, for various values of the normalized signal amplitude, \( A_\star \). As expected, \( f(c_\star \mid s) = f(c_\star \mid n) \) when \( A_\star = 0 \). Theoretical ROC curves based on these distributions are shown in Figure 3.10. As shown in Figure 3.9, the variance of \( c_\star \) is greater when the stimulus is signal plus noise than when the stimulus is noise alone. As a result, the ROC curves for this detector are not symmetrical about the negative diagonal. When plotted on probability axes, the ROC curves predicted by this detector are nearly straight lines, with a slope of 0.80 to 0.90. The slope decreases as the signal detectability increases.

Detectability is a function of the normalized signal amplitude, \( A_\star \). It may also be expressed in terms of the signal to noise ratio, \( E_s/N_o = A_\star^2/2 \). The variation of \( P_A \) with signal energy is discussed in Section 3.4.5.

### 3.4.4 Energy Detector

The final situation considered is one in which the detector is very uncertain about the form of the signal. For this analysis we use the discrete-time representation of signal and noise: the input waveform is summarized by the vector \( x = \{x_m\} \). First, consider the case where the signal is simply a sample of noise with bandwidth \( W \). Let \( N_o \) be the power density of the noise, and \( S_o \) be the power density of the signal. Then the conditional densities for each sample \( x_m \) are:

\[
f(x_m \mid n) = \frac{1}{\sqrt{2\pi N_o W}} \exp\left(-\frac{x_m^2}{2N_o W}\right)
\]  

(3.54)

\[
f(x_m \mid s) = \frac{1}{\sqrt{2\pi (N_o+S_o) W}} \exp\left(-\frac{x_m^2}{2(N_o+S_o) W}\right)
\]  

(3.55)
The likelihood function is:

$$L[x(t)] = \frac{2^{WT}}{\prod_{m=1}^{2^{WT}} f[x_m | s]} \prod_{m=1}^{2^{WT}} f[x_m | n] = \left( \frac{N_o}{N_o+S_o} \right)^{WT} \exp \left( \frac{S_o}{2N_oW(N_o+S_o)} \sum_{m=1}^{2^{WT}} x_m^2 \right)$$

(3.56)

$N_o$, $S_o$ and $WT$ are constant, so the likelihood ratio is monotonic with $z$, where:

$$z = \sum_{m=1}^{2^{WT}} x_m^2$$

(3.57)

Clearly, $z$ is proportional to the energy of the observed waveform. Therefore, if the signal is simply a sample of noise, the optimal detector uses a decision rule based on the energy of the observed waveform.

It can be argued that the energy detector is also close to optimal for the case in which the signal is sinusoidal, but the detector is uncertain of the signal phase and frequency. Let us determine the detector performance for the case in which the signal waveform (represented by $s = \{s_m\}$) is always the same, but its characteristics are unknown. A sinusoidal signal of unknown frequency and phase is a specific example of this case.

If the waveform is noise alone, then the waveform energy is:

$$z_n = \sum_{m=1}^{2^{WT}} r_m^2 \Delta t$$

(3.58)

where $r_m$ is Gaussian with zero mean and variance $\sigma_r^2 = N_oW$, and $\Delta t = 1/(2W)$. Recall that $r_m$ is independent of $r_n$ for $m \neq n$. By definition, the sum of the squares of $N$ independent, normalized Gaussian variables is a $\chi^2$ variable with $N$ degrees of freedom. Therefore:

$$z_n^* = \frac{2z_n}{N_o} = \sum_{m=1}^{2^{WT}} \left( \frac{r_m}{\sigma r} \right)^2$$

(3.59)
is a $\chi^2$ variable with $2WT$ degrees of freedom. The expected value of a $\chi^2$ variable is the number of degrees of freedom. The variance is twice the number of degrees of freedom. Therefore:

$$\mu_n = E[z_n^*] = 2WT$$ (3.60)

$$\sigma_n^2 = E[(z_n^* - \mu_n)^2] = 4WT$$ (3.61)

If the waveform is signal plus noise, then the normalized waveform energy is:

$$z_s^* = \frac{2z_s}{N_0} = \frac{2^{WT}}{\sum_{m=1}^{2WT} \left( \frac{r_m + s_m}{\sigma_r} \right)^2}$$ (3.62)

$z_s^*$ is the sum of the squares of $2WT$ independent Gaussian variables. Each variable has a mean of $s_m/\sigma_r$ and unity variance. Therefore, $z_s^*$ is a noncentral $\chi^2$ variable with $2WT$ degrees of freedom. The noncentrality parameter is:

$$\lambda = \sum_{m=1}^{2WT} (s_m/\sigma_r)^2 = \frac{2E_s}{N_0}$$ (3.63)

The mean of a noncentral $\chi^2$ distribution is the number of degrees of freedom plus the noncentral parameter. The variance is twice the number of degrees of freedom plus four times the noncentral parameter. Therefore,

$$\mu_s = E[z_s^*] = 2WT + \frac{2E_s}{N_0}$$ (3.64)

$$\sigma_s^2 = E[(z_s^* - \mu_s)^2] = 4WT + \frac{8E_s}{N_0}$$ (3.65)

This completes the derivation of $f(z \mid s)$ and $f(z \mid n)$ for the energy detector. Calculation of the detector performance directly from the $\chi^2$ distributions is difficult. However, if the number of degrees of freedom is large ($2WT \geq 10$), then both the $\chi^2$ and non-central $\chi^2$ distributions are approximately Gaussian. Green and Swets use this fact to derive simple approximate expressions for the energy detector performance. From equation 3.15, the
The index of detectability is:

\[ d'_{A} = \frac{\mu - \mu_n}{\sqrt{\sigma_n^2 + \sigma_s^2}/2} = \frac{E_f/N_0}{\sqrt{WT + E_f/N_0}} \]  

(3.66)

From equation 3.16, the slope of the ROC curve on probability axes is:

\[ b = \frac{\sigma_n}{\sigma_s} = \sqrt{\frac{WT}{WT + 2E_f/N_0}} \]  

(3.67)

The above approximations (equations 3.66 and 3.67) were checked computationally for a representative range of values. For each pair of values (N degrees of freedom and \( \lambda \) non-centrality), eight evenly spaced points on the ROC curves were calculated from the exact \( \chi^2 \) distributions. To obtain accurate estimates of the index of detectability, \( d'_{A} \), and the ROC slope parameter, \( b \), an ROC curve based on Gaussian distributions of \( f(e|s) \) and \( f(e|n) \) was fitted to the computed points using the maximum likelihood procedure described in Appendix F. The resulting values were compared with the approximate values calculated using equations 3.66 and 3.67. This comparison is shown in Table 3.1.

The conditional distributions for a typical case are shown in Figure 3.11. The exact ROC points and the fitted Gaussian ROC curve based on these distributions are shown in Figure 3.12. The true \( \chi^2 \) ROC curve is very close to a straight line, so a Gaussian model provides a good approximation. However, the approximate ROC curve obtained using equations 3.66 and 3.67 is not correct. Specifically, the approximate expression for the ROC slope, \( b \), is seriously in error. The skewness of the \( \chi^2 \) distributions tends to reduce the "effective" difference in variance. The approximate expression for \( d'_{A} \) is reasonably accurate, although it can underestimate the true value by 5 to 10% when \( \sigma_n/\sigma_s < 0.5 \).

The following expression was found to give a much better approximation for the ROC slope parameter:

\[ b = \left( \frac{\sigma_n}{\sigma_s} \right)^{1/4} = \left( \frac{WT}{WT + 2E_f/N_0} \right)^{1/8} \]  

(3.68)
As shown in Table 3.1, this expression is accurate within 2% over the range of parameters considered. As shown in Figure 3.12, the ROC curve based on the new expression agrees quite well with the data for the example case.

If $E/N_o << WT$, $d_A'$ is proportional to the signal energy (or the square of the amplitude) and the ROC is almost symmetrical. As $E/N_o$ increases relative to WT, the ROC curve becomes more skewed ($b < 1$).

### 3.4.5 Comparison of Detectors

The three previous sections showed how the ideal detector should process the observed waveform in three different situations. If the signal shape is known exactly, then the detector should base its decision on the correlation between the observed waveform and the expected signal. As the uncertainty about the signal increases, the form of the optimal detector also changes. If the detector is uncertain about both the phase and frequency of the signal, the optimal decision rule is one based on the energy of the observed waveform.

The derivation of these ideal detectors was based on some rather sweeping assumptions. We assumed that the noise was Gaussian with a rectangular power spectrum, and that both signal and noise could be represented adequately by discrete samples or a finite Fourier series. However, the analysis is instructive because it shows how the form of the ideal detector depends on the uncertainty about the signal. The three detectors discussed above span the range of likely situations.

Some results are available for nonrectangular spectra. Pfafflin and Mathews (1962) developed an energy detection model in which the input noise was passed through a tuned second order bandpass filter. They later showed that, for a given bandwidth and noise power, the variance of the noise energy, $\sigma_n^2$, is half as large with the second order filter as with a rectangular filter (Mathews and Pfafflin, 1965). Swets et al (1962) reported similar findings. For more complex cases, the analysis becomes very tedious. An alternative approach is to assume the form of the detector, then evaluate its performance by Monte Carlo simulation.

Let us compare the performance of the three detectors. All three show that signal detectability is a function of $E/N_o$, the ratio of signal energy to noise power density.
Random noise therefore masks the signal. Increasing the noise power decreases the signal detectability. For the correlation and envelope detectors, performance is independent of the noise bandwidth, \( W \). This implies that narrow band noise near the signal frequency should mask a signal as effectively as broadband noise. The performance of the energy detector, on the other hand, decreases as the noise bandwidth increases.

The shape of the ROC curve is different for the three detectors. For the correlation detector, \( f(e|s) \) and \( f(e|n) \) are Gaussian with equal variance. The resulting ROC, plotted on linear axes, is symmetrical about the negative diagonal. Plotted against normal probability axes, the ROC is a straight line with slope \( b = 1.0 \). For the energy detector, \( f(e|s) \) and \( f(e|n) \) are approximately Gaussian, but have unequal variance. This yields a skewed ROC when plotted on linear axes, as shown in Figure 3.12. Plotted on probability axes, the ROC is a straight line with slope \( b < 1.0 \). The slope decreases as the ratio of signal power to noise power increases. Finally, the ROC for the envelope detector is also slightly skewed. The underlying functions are not Gaussian, so when plotted on probability axes, the ROC is slightly curved rather than a straight line. An example envelope detector ROC is shown in Figure 3.10.

The overall detector performance is substantially different for the three detectors. \( P_A \), the area under the ROC curve, is shown as a function of \( E_s/N_0 \) for all three detectors in Figure 3.13. These curves are known as psychometric curves. The correlation detector clearly performs best. The envelope detector and energy detector need a higher signal to noise ratio to achieve the same level of detectability. This is expected, because these detectors are less certain than the correlation detector of the signal shape. Also note that the psychometric curve for the correlation detector is relatively shallow, rising from chance to near-perfect performance over about 25 dB. The curve for the energy detector is much steeper, rising from chance to perfect performance over about 10 dB.

As a final note, it is interesting to compare these detectors with the predictions of various threshold theories. We will consider the high-threshold theory proposed by Blackwell (1953), and the low-threshold theory proposed by Luce (1963). A review of these and other threshold theories, and a discussion of their implications, is given by Green and Swets (1974, Chapter 5).

The classical interpretation of the threshold is a minimum level of sensory stimulation which must be exceeded for the observer to detect the signal. If the stimulus does not
exceed the threshold, it has absolutely no effect on the observer. The idea of an invariant threshold is in conflict with most perceptual data, so many researchers have adopted the view that either the threshold or the sensory effect of a given signal varies randomly with time. However, the threshold is still considered to be a physiological nonlinearity of the observer; the stimulus must exceed the threshold in order to have an effect. This is in contrast with the variable decision criterion suggested by signal detection theory.

High threshold theory assumes that the threshold is never exceeded by noise alone, and is exceeded by signal plus noise with a probability of $P(X>T | s)$. If the threshold is exceeded, the observer always responds $S$. If it is not exceeded, the observer guesses, responding $S$ with a probability $P(S | X<T)$. The conditional probabilities of responding $S$ are therefore:

$$P(S | s) = P(X>T | s) + P(S | X<T) [1 - P(X>T | s)]$$  \hspace{1cm} (3.69)

$$P(S | n) = P(S | X<T)$$  \hspace{1cm} (3.70)

Combining these equations yields:

$$P(S | s) = P(X>T | s) + P(S | n) [1 - P(X>T | s)]$$  \hspace{1cm} (3.71)

$P(S | s)$ is a linear function of $P(S | n)$. The ROC curve predicted by high-threshold theory is therefore a straight line. Some examples are shown in Figure 3.14.

Luce (1963) proposed a two-state low-threshold theory in which the threshold is exceeded by noise alone with a probability of $P(X>T | n)$. In this case, the conditional probabilities of responding $S$ are:

$$P(S | s) = P(S | X>T) P(X>T | s) + P(S | X<T) [1 - P(X>T | s)]$$  \hspace{1cm} (3.72)

$$P(S | n) = P(S | X>T) P(X>T | n) + P(S | X<T) [1 - P(X>T | n)]$$  \hspace{1cm} (3.73)

If the observer always responds $S$ when the threshold is exceeded and $N$ when it is not, then $P(S | s) = P(X>T | s)$ and $P(S | n) = P(X>T | n)$. However, the observer may change his operating point on the ROC by guessing. If the observer always responds $S$ when the threshold is exceeded, and guesses with probability $P(S | X<T)$ when it is not, then both $P(S | s)$ and $P(S | n)$ are linear functions of $P(S | X<T)$. Changing the guessing probability $P(S | X<T)$ traces out a linear ROC segment from the point $(P(X>T | n), P(X>T | s))$ to the
point (1,1). Similarly, if the observer never responds S when the threshold is not exceeded, and guesses with probability \( P(S \mid X>T) \) when it is, then both \( P(S \mid s) \) and \( P(S \mid n) \) are linear functions of \( P(S \mid X>T) \). Changing the guessing probability traces out a linear ROC segment from the point \((P(X>T \mid n), P(X>T \mid s))\) to \((0,0)\). Thus the low-threshold theory predicts piecewise linear ROC's such as those shown in Figure 3.15.

The ROC curves predicted by both of these theories differ markedly from the curvilinear ROC's predicted by signal detection theory. To test whether or not signal detection theory is applicable to motion perception, a good first step is to determine the shape of the ROC curve.
Table 3.1  Detectability and ROC slope parameters based on the $\chi^2$ distributions.

<table>
<thead>
<tr>
<th>DOF</th>
<th>$\lambda$</th>
<th>approximate (eqn. 3.66)</th>
<th>MLE</th>
<th>approximate (3.67)</th>
<th>(3.68)</th>
<th>MLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1.5</td>
<td>0.500</td>
<td>0.496</td>
<td>0.707</td>
<td>0.917</td>
<td>0.907</td>
</tr>
<tr>
<td></td>
<td>3.0</td>
<td>0.866</td>
<td>0.900</td>
<td>0.577</td>
<td>0.872</td>
<td>0.870</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>1.250</td>
<td>1.357</td>
<td>0.480</td>
<td>0.833</td>
<td>0.846</td>
</tr>
<tr>
<td>5</td>
<td>2.0</td>
<td>0.535</td>
<td>0.533</td>
<td>0.745</td>
<td>0.929</td>
<td>0.921</td>
</tr>
<tr>
<td></td>
<td>4.0</td>
<td>0.943</td>
<td>0.972</td>
<td>0.620</td>
<td>0.887</td>
<td>0.888</td>
</tr>
<tr>
<td></td>
<td>7.0</td>
<td>1.429</td>
<td>1.532</td>
<td>0.513</td>
<td>0.846</td>
<td>0.864</td>
</tr>
<tr>
<td>10</td>
<td>3.0</td>
<td>0.588</td>
<td>0.588</td>
<td>0.791</td>
<td>0.943</td>
<td>0.937</td>
</tr>
<tr>
<td></td>
<td>6.0</td>
<td>1.061</td>
<td>1.084</td>
<td>0.674</td>
<td>0.906</td>
<td>0.906</td>
</tr>
<tr>
<td></td>
<td>10.0</td>
<td>1.581</td>
<td>1.658</td>
<td>0.577</td>
<td>0.872</td>
<td>0.884</td>
</tr>
<tr>
<td>30</td>
<td>5.0</td>
<td>0.598</td>
<td>0.597</td>
<td>0.866</td>
<td>0.965</td>
<td>0.959</td>
</tr>
<tr>
<td></td>
<td>10.0</td>
<td>1.118</td>
<td>1.128</td>
<td>0.775</td>
<td>0.938</td>
<td>0.934</td>
</tr>
<tr>
<td></td>
<td>16.0</td>
<td>1.668</td>
<td>1.703</td>
<td>0.696</td>
<td>0.913</td>
<td>0.914</td>
</tr>
<tr>
<td>60</td>
<td>6.0</td>
<td>0.522</td>
<td>0.522</td>
<td>0.913</td>
<td>0.977</td>
<td>0.972</td>
</tr>
<tr>
<td></td>
<td>12.0</td>
<td>1.000</td>
<td>1.003</td>
<td>0.845</td>
<td>0.959</td>
<td>0.953</td>
</tr>
<tr>
<td></td>
<td>20.0</td>
<td>1.581</td>
<td>1.597</td>
<td>0.775</td>
<td>0.938</td>
<td>0.934</td>
</tr>
<tr>
<td>120</td>
<td>16.0</td>
<td>0.970</td>
<td>0.972</td>
<td>0.889</td>
<td>0.971</td>
<td>0.965</td>
</tr>
</tbody>
</table>

Notes:  
1. equations 3.66 and 3.67 are approximate expressions proposed by Green and Swets.  
2. equation 3.68 is a new approximate expression for b.  
3. MLE is the maximum likelihood estimate.
Figure 3.1  The two main components of a signal detector.
Typical conditional distributions of the evidence variable, $e$. A simple decision rule is to respond $S$ if and only if $e \geq k$.

**Figure 3.3** The Stimulus Response Matrix.
Figure 3.4 Derivation of the ROC curve by varying the response criterion.
Figure 3.5  Effect of signal strength on the ROC curve.
a) conditional distributions with unequal variance

b) the corresponding ROC curve.

Figure 3.6 Skewed ROC curve based on Gaussian distributions with unequal variance. The points on the curve are for comparison with Figure 3.7.
Figure 3.7 An ROC curve based on Gaussian distributions, plotted on normal probability axes. The points on the curve are the same as those shown in Figure 3.6.
Figure 3.8  Assumed rectangular noise spectrum.
Figure 3.9  Conditional distributions $f(e \mid s)$ and $f(e \mid n)$ based on the envelope detector.
Figure 3.10  ROC curves based on the envelope detector.
Figure 3.11  Chi-square conditional distributions based on the energy detector, with $\lambda = 10$, and $N = 10$. 
Figure 3.12  ROC points based on the distributions of Figure 3.11, fit with Gaussian ROC curves.
Figure 3.13  Psychometric curves showing the detection performance of the three models as a function of signal to noise ratio.
Figure 3.14  Typical ROC curves based on high-threshold theory.
Figure 3.15  Typical ROC curves based on Luce's two-state low threshold theory. For the curves shown, the probability of the threshold being exceeded by noise alone is 0.10.
Chapter 4

The Experimental Approach

4.1 Introduction

Recall from Chapter 1 that the objective of this study is to find a model to describe human sensitivity to periodic motion cues in the presence of random motion. This chapter describes a motion perception model based on the vestibular models discussed in Chapter 2, and the ideal detectors introduced in Chapter 3, and outlines the experimental program which was set up to test the model.

This chapter is divided into several sections. Section 4.2 presents a complete perception model which includes receptor dynamics, an internal filter, and an ideal detector. Section 4.3 outlines the study program, which comprises three experiments. In each experiment, we tested the ability of human observers to detect a sinusoidal motion "signal" superimposed on a background of random motion, or "noise." The three experiments were complementary. Each tested a different aspect of the signal detection models. All of the experiments were carried out in the UTIAS Flight Research Simulator, which is described in Section 4.4. Finally, Section 4.5 discusses some aspects of the signal and noise motion which were common to all experiments, and presents graphs of typical motion conditions used in the experiments.

4.2 Proposed Motion Perception Model

Chapter 2 provided a brief description of the human motion sensors and presented mathematical models which have been developed to describe their response characteristics. At the end of Chapter 2, it was suggested that perception of low-amplitude motion is a signal-in-noise detection problem which can be modelled using signal detection theory. Chapter 3 summarized the basics of signal detection theory, and introduced three "ideal detectors" which might be applicable to motion perception. The purpose of this section is to pull all of this information together, in order to formulate a complete motion perception model.
The proposed motion perception model is shown in Figure 4.1. The model consists of four major components: the receptor dynamics, an internal noise source, an internal filter, and an ideal detector. The inputs to the model are the signal, \( s(t) \), and the external noise, \( r(t) \). The model monitors the input waveform over some sampling interval \((0,T)\), then outputs a response which is either S (a signal was present), or N (a signal was not present). It should be emphasized that the purpose of the model is to provide an engineering tool for predicting human motion perception performance. The individual components of the model do not necessarily have physiological parallels in the human observer.

The first component of the model represents the receptor dynamics. Unity transfer functions were used to model the receptors. The output of the translational motion sensors (otoliths) is assumed to be proportional to the input specific force. The output of the angular motion sensors (semi-circular canals) is assumed to be proportional to angular velocity. Both transfer functions are assumed to be constant over the frequency range of interest. These are simplified versions of the vestibular models proposed by Ormsby (1974) and shown in Figures 2.5 and 2.8. Over the frequency range of interest (0.1 Hz to about 4 Hz), the gain of Ormsby's models is constant.

Motion cues in an aircraft or simulator are sensed by somatic and kinesthetic sensors as well as the vestibular system, and a comprehensive model should include the contributions of all three. However, as discussed in Chapter 2, motion sensation is dominated by the vestibular sensors, so the contribution of other receptors is neglected in the approximate model used here.

The assumed dynamic models for the receptors should also account for the dynamics of the seat and body in transmitting motion to the head. No measurements of head acceleration or rotation were made during the course of this study. Therefore, estimates of head motion are based on data from the literature. Barnes and Rance (1974) tested the transmission of yaw oscillations to the head of seated subjects. They define the term transmissibility as the gain of the describing function relating head motion to chair motion. For subjects restrained at the shoulders, transmissibility was unity below 1 Hz, dropped off to 0.8 at 4 Hz, and then fell off proportional to \( 1/\omega \) above 4 Hz. For unrestrained subjects, transmissibility was 1.0 at 0.6 Hz, rose to 1.5 at 3 Hz, then fell off proportional to \( 1/\omega^2 \) above 3 Hz. Woods (1967) measured the transmission of lateral (Y) acceleration to the head of unrestrained seated subjects. He obtained values of 0.8 at
1 Hz, rising to 1.0 at 1.5 Hz, and falling off to 0.4 at 4 Hz. Clearly, the data for unrestrained subjects show some resonant amplification between 1 and 3 Hz, but fall off fairly sharply at higher frequencies.

In the current study, subjects were restrained by a five point aviation harness which had spring-loaded inertial reels on the shoulder straps. Therefore, the subjects were partially restrained at the shoulders. The data provided by Barnes and Rance for restrained subjects were judged to be the most applicable. For the purposes of all analysis in this report, it was assumed that transmissibility was 1.0 below 4 Hz, and attenuated above 4 Hz.

The second component of the model is an internal noise source, indicated in Figure 4.1 as \( r_1(t) \). One obvious source of internal noise is the random variation in firing frequency of the vestibular neurons, which was discussed in Chapter 2. However, keeping in mind that we are trying to model overall motion detection performance, internal noise must not be limited to vestibular sources. Some other possible contributors of internal noise are postural sources such as muscular tremor and spontaneous body sway. Except for the rough estimates of RMS neural noise presented in Chapter 2, no information is available about the characteristics of the internal noise.

The third component of the model is an internal filter. We assume that the observer has some imperfect knowledge of the signal frequency. As a result, he is able to ignore any random motion which has a frequency much different from the signal frequency. Only noise components near the signal frequency affect the signal detectability. In the model, this is represented by a bandpass filter. There is, as yet, no basis for estimating the internal filter bandwidth. Note that the internal filter affects performance only if the detection component of the model is an energy detector. In the development of the theory for the correlation and envelope detectors, it was assumed that the observer knows the signal frequency exactly.

The concept of an internal filter has been widely used in auditory detection models. For auditory detection, the internal filter or "critical band" has been explained in terms of the physiology of the primary auditory receptor, the cochlea (Fletcher, 1940). Because of the simple structure of the vestibular receptors, it seems unlikely that there is a simple physiological explanation for an internal filter for motion perception. Frequency
discrimination is more likely a function of higher order processing, perhaps even conscious filtering.

The final component of the model is an ideal detector which monitors the incoming waveform over a time interval \((0,T)\), and decides whether or not a signal was present. Chapter 3 introduced three possible formulations for this detector. All of them are based on the assumption that the observer's task is to detect a sinusoidal signal imbedded in white noise or band-limited noise. The correlation detector is based on the assumption that the signal frequency and phase are known exactly. The envelope detector assumes that the observer is uncertain of the signal phase, and the energy detector assumes that he is uncertain of both phase and frequency. Many other formulations are possible, but these three are distinct enough to cover the likely range of performance.

We would like to keep an open mind about which, if any, of the models is valid. However, a review of the assumptions suggests that the energy detector might be the most appropriate. Even with repeated sinusoidal signals, an observer is likely to be somewhat uncertain of frequency and phase. In a "real-life" motion detection task in an aircraft or simulator, the pilot is normally quite uncertain of the signal characteristics. There, most of the motion cues are transient rather than sinusoidal, and the onset time of each cue is not known in advance.

4.3 Objectives of the Experimental Program

The initial aim of this study was to answer a practical question. How big must a motion cue be for a pilot to feel it, and how does this level depend on the overall background motion level in the simulator? To answer this question, we have suggested that motion perception is a signal-in-noise detection problem, and can be modelled by signal detection theory.

The original question is now replaced by a multitude of others. First, does signal detection theory apply to motion perception, or is human performance described better by a threshold theory? Second, if signal detection theory does apply, which of the above models is the best? Finally, we would like to estimate a number of model parameters: the statistics of the internal noise, the bandwidth of the internal filter, and the length of the sampling interval, \(T\).
To test the models, we must rely on measurements of subjective perception, and compare these data with the model predictions. The only two behavioural data available are the hit rate, $P(S|s)$, and the false alarm rate, $P(S|n)$, which together define a point on the ROC curve. By choosing a suitable experimental procedure, we can either trace out the entire ROC curve for a given subject and motion condition, or estimate the signal detectability, as expressed by $P_A$ or $d_A'$. The shape of the ROC curve or the variation of detectability with signal amplitude or noise condition can then be compared with model predictions.

Three experiments were proposed. The purpose of the first was to determine the shape of the ROC curve for a few representative cases. The resulting data provided a basis for assessing the applicability of signal detection theory to motion perception. As shown in Chapter 3, the threshold model predicts a straight-line ROC, while the signal detection models predict a smooth curvilinear ROC. The results also helped to evaluate the suitability of the three models. The correlation detector predicts a symmetrical ROC curve ($b = 1$), while the envelope and energy detectors predict skewed ROC curves for which $b < 1.0$. ROC curves were obtained for both pitch and surge signals at 0.6 Hz, at two different signal amplitudes. A low-power broadband random motion was used as the noise condition in all cases. Data were obtained using a rating procedure which provided four points on the ROC curve. This experiment is described in Chapter 5.

The purpose of the second experiment was to assess the effect of noise power density and frequency content on signal detectability. We determined the signal amplitude which had a detectability index of $d_A' = 1.0$ for seven different noise conditions. The data were obtained using the forced choice procedure. Both pitch and surge were tested, and the noise was always on the same axis as the signal. Three of the noise conditions considered were broadband spectra with different power densities. For all three ideal detectors, detectability is a function of the signal-to-noise ratio, $E_s/N_0$. Therefore, the signal energy required to obtain $d_A' = 1.0$ should be proportional to the noise power density. A finding that the signal-to-noise ratio is constant for all conditions would support signal detection theory; a finding that it is not would force us to reject it.

The remaining four noise conditions were narrow band spectra with the same power density but different centre frequencies. One reason for including these conditions was to distinguish between the energy detector and the other two models. For the correlation and envelope detectors, signal detectability is independent of the noise bandwidth. For
these models, a narrow band noise spectrum containing the signal frequency should mask a signal as effectively as a broadband spectrum with the same power density. Performance of the energy detector, on the other hand, does depend on the noise bandwidth. In this case, the narrow band spectrum should not mask the signal as effectively as the broadband spectrum.

Results for the narrow band noise cases also helped to test the concept of an internal filter. If the observer is able to filter out noise which is far from the signal frequency, then the signal amplitude (at $d' = 1.0$) should be highest when the noise band is near the signal frequency. Analysis of the data provided estimates of the internal filter bandwidth. The second experiment is described in Chapter 6.

The first two experiments concentrated on testing the validity of the signal detection model and estimating some parameters. The third experiment, on the other hand, was designed to answer a more practical question: to what extent does random motion on one axis affect the detectability of signals on other axes? As in the second experiment, we estimated the signal amplitude which has a detectability of 1.0 in several different noise conditions. In this experiment, however, we also considered cases where the signal and noise were on different axes. The results provide an indication of the amount of interaction between axes. This experiment is discussed in Chapter 7.

All three experiments provided data showing how detectability varies with signal amplitude with all other conditions fixed. This provides some basis for choosing between models. As seen in Figure 3.13, the predicted psychometric curve grows steeper and moves to the right as we progress from the correlation detector to the energy detector. The second and third experiments tested two or more signal frequencies, to provide some rough information about the receptor dynamics. All three experiments considered both translational and rotational motions. The only parameter which was not specifically tested was the sampling time of the detector, $T$.

The three experiments were designed to be complementary, each testing a different aspect of the models. However, one baseline condition was common to all three. Also, the results from all three experiments have some bearing on the same model parameters. A comparison of the data from the three experiments, and a discussion of the overall implications for the signal detection models, are presented in Chapter 8.
4.4 The Experimental Facility

All experiments were carried out in the UTIAS Flight Research Simulator. The simulator is a mock-up of a DC-8 cab, mounted on a Series 300 hydraulic motion base manufactured by CAE Electronics Ltd. of Montreal. The motion base consists of six hydraulic actuators in a synergistic configuration, and is capable of motion in all six degrees of freedom: surge, sway, heave, pitch, roll and yaw. The simulator is equipped with an instrument package containing three accelerometers and three rate gyros which allow accurate measurement of its motion. Figure 4.2 is a photograph of the facility.

The motion base is driven by digital signals from the controlling computer, which in this case is a Perkin-Elmer 3250 minicomputer. The drive signals comprise a position and acceleration command for each actuator. An analog control circuit uses these command signals, in combination with feedback from position and force transducers on each actuator, to drive the servo valves for each actuator. In all six degrees of freedom, the response of the motion base is almost linear, and has a near-unity transfer function from 0 to 10 Hz. The dynamic response characteristics of the motion base have been investigated in detail and are reported by Grant (1986).

The maximum acceleration, velocity and displacement in each degree of freedom are summarized in Table 4.1. The amplitude of sinusoidal motion on each degree of freedom can be varied almost continuously from zero to the maximum. The bit resolution on the actuator commands is 2.3×10⁻⁴ m for position, and 0.005 m/s² for acceleration. The resolution of the simulator motion depends on the motion base geometry and is different for each axis. However, it is of the same order as that for the actuator motion. At 0.6 Hz, the resolution is approximately 2×10⁻³ m/s² (0.2 milli-g) for translational accelerations, and 2×10⁻³ rad/s² (0.11°/s²) for angular accelerations. Expressed as velocities, these values are 5×10⁻⁴ m/s for translations and 0.03°/s for rotation. The minimum commandable motion is equal to the resolution. However, at very low amplitudes the response of the simulator is quite noisy (signal-to-noise ratio less than 2:1), so the practical minima are about 8×10⁻³ m/s² for translations, and 8×10⁻³ rad/s² for rotations.

All of the experiments were run in the rear workstation of the simulator, which is shown in Figure 4.3. The workstation consists of a seat mounted in front of an automobile dashboard. The seat is a typical aviation pilot’s seat, equipped with armrests and a five point harness. The subject’s head was not restrained, but subjects were told to keep their
head erect and facing forward, and not to slouch. During each experimental run, the subject sat strapped into the seat with his arms resting on the armrests and his feet flat on the floor. All lights inside the simulator were off except for the dashboard lights, which were illuminated at a low level representative of night flying. No "out-the-window" forward display was used. Subjects were told to keep their eyes open and to look straight ahead.

Sound measurements inside the simulator showed that the background sound level during periods of no motion was about 30 dB(A). With large amplitude motions, the hiss of the motion base hydraulics is clearly audible inside the cab, with peak levels of about 45 dB(A) when the actuators are at their peak velocity. For most of the motion conditions used in this study, the hydraulic noise is inaudible. However, to ensure that auditory cues did not affect the results, the subjects wore an audio headset during the experiments. The headset shell provided 15 to 20 dB attenuation over the frequency range 125 - 8000 Hz. Also, broadband noise was played over the headset to mask any aural cues about the motion. The headset also provided a line of communication between the subject, the computer which ran the experiment, and the experimenter.

Each experiment was fully automated. The controlling program for each experiment was written in Fortran and ran in real time on the Perkin Elmer 3250 computer. The controlling program had several functions. It calculated the desired simulator motion in real time and sent the appropriate command signals to the motion base. It controlled a sound generator which issued prompts, warnings and feedback to the subject. It adjusted the experimental conditions or made a random selection of conditions when necessary. Finally, it recorded the response data entered by the subject. Automation of the experiment streamlined the data collection process and allowed us to eliminate all visual and auditory contact between the subject and experimenter. This helped to ensure that the data are unaffected by any interaction between the subject and experimenter.

4.5 Generation of Signal and Noise

Each experimental run was a series of trials. The start and end of each trial was indicated by auditory tones. During each trial, the subject was exposed to a motion stimulus which was either noise alone or signal plus noise. The required simulator motion was calculated on-line in real time.
The random background motion or noise was continuous throughout each run, that is, it did not start and stop on each trial. Also, the power and frequency content of the noise were constant over the duration of a run. The noise was generated by passing digital white noise through a fifth order digital bandpass filter, as described in Appendix C. The gain, break frequencies and damping of the filter were selected to obtain the desired noise spectrum. The white noise was simply a sequence of Gaussian random numbers. A different seed value was used to initialize the random number generator at the start of each run. As a result, the actual time history of noise motion was unique for each run.

The motion stimulus on some trials was signal plus noise. On these trials, a gated sinusoidal signal was superimposed on the noise motion. The gated signal consisted of three sections: a fade-in ramp of duration $T_r$, a period of constant amplitude of duration $T_s$, and a fade-out ramp of duration $T_r$. In order to avoid any discontinuities in commanded displacement, velocity or acceleration, the following consistent ramp was used:

$$x(t) = \frac{-A(t)}{\omega} \cos \omega t + \frac{B(t)}{\omega^2} \sin \omega t$$  \hspace{1cm} (4.1)

$$\dot{x}(t) = A(t) \sin \omega t$$  \hspace{1cm} (4.2)

$$\ddot{x}(t) = A(t) \omega \cos \omega t + B(t) \sin \omega t$$  \hspace{1cm} (4.3)

where $x(t)$ is the translational or angular displacement at time $t$, and $A(t)$ is an amplitude function defined as:

$$A(t) = \begin{cases} 
A_o(t/T_r) & \text{for } 0 \leq t < T_r \\
A_o & \text{for } T_r \leq t \leq T_s + T_r \\
A_o(T_s+2T_r-t)/T_r & \text{for } T_s + T_r < t \leq T_s + 2T_r
\end{cases}$$

and $B(t) = \frac{d}{dt} [A(t)]$

It may be shown that if $T_r = m\pi/\omega$ and $T_s = n\pi/\omega$, where $m$ and $n$ are integers, then $x(t)$ as defined by equation 4.1 is continuous and twice differentiable, for $0 \leq t \leq T_s + 2T_r$. The
first and second derivatives are given by equations 4.2 and 4.3, respectively. The third derivative of \( x(t) \) is discontinuous.

For all cases in the current study, \( T_r \) was set equal to \( 2\pi/\omega \), the signal period. \( T_s \) was 10 seconds, which satisfied the condition \( T_s = \pi \omega/\omega \) for each signal frequency considered. Figure 4.4 shows the waveform of a typical gated signal at 0.6 Hz. Note that the nominal signal amplitude expressed as displacement, velocity or acceleration, is \( A_s/\omega \), \( A_o \) or \( \omega A_o \), respectively.

In order to reduce the predictability of the signal, a random delay was included between the start of the trial and the start of the gated signal. For each trial, the delay was selected randomly from a uniform distribution between 1 and 2 seconds. Also, the gated signal was multiplied by a random sign function, which took the value +1 or -1 with equal probability on each trial.

The required motion commands were computed on-line as the experiment was in progress. The motion calculations were updated every 40 milliseconds, that is, at 25 Hz. At each time step, the desired simulator motion was simply the sum of the noise motion and the signal motion. The command motion is completely specified by two vectors: \( \omega = [p \ q \ r]^T \), the angular velocity, and \( \Delta f_H \), the change in specific force in the body frame.

Command signals for the motion base were generated by a motion drive algorithm developed specifically for this study. This algorithm was an adaptation of the classical washout filter algorithm described by Reid and Nahon (1985). On each time step, the input specific force and angular velocity vectors were transformed and integrated to obtain the simulator position and attitude in an earth-fixed inertial reference frame. The motion command vector, which specifies the length and acceleration of each actuator, was then calculated based on the simulator geometry and the desired position and attitude. A brief description of the motion drive algorithm is given in Appendix A.

The most notable feature of this algorithm is its ability to produce pitch and roll motions (both signal and noise) which are "pure" in the sense that they do not result in any change in specific force at the observer's head. Rotating about a horizontal axis which is fixed in inertial space causes a change in specific force at the centre of rotation in the body frame, due to the changing orientation of the gravity vector with respect to the body frame. The motion drive algorithm used here accelerated the simulator translationally in such a way
as to eliminate the change in specific force at the origin of the body frame. The resulting motion was much like that of a swing, where the centre of rotation was some distance above the observer’s head. The point of this was to avoid stimulating the otoliths (which sense specific force) during a rotation. This is discussed in more detail in Appendix A.

Prior to the start of the experiments, the motion quality was tested by recording the actual simulator motion during typical experimental conditions. Measurements were made using the onboard instrumentation package equipped with three accelerometers and three rate gyros. The measured data were then processed to obtain time histories and spectra of the specific force and angular rates at the observer’s head. A brief discussion of the signal processing procedure is given in Appendix B.

The motion measurements showed some coupling, or crosstalk. This is parasitic motion on one axis caused by commanded motion on another. The coupling can be attributed to small asymmetries in the motion base geometry and mass distribution, and in the analog control circuitry. The crosstalk is small enough that it is not of concern for ordinary flight simulation applications. However, for the present study we felt that it should be minimized. Two methods of reducing crosstalk were considered: one open loop, the other closed loop. The open loop method was chosen as the most practical. This solution involved multiplying the body-frame motion vector by a $6 \times 6$ filter matrix prior to passing it to the motion drive algorithm. This measure did not completely eliminate the crosstalk, but the remaining levels are small, as summarized in Table 4.2. The crosstalk reduction method is discussed briefly in Appendix A.

The combined transfer function of the motion drive algorithm and motion base was determined for a range of frequencies between 0.2 and 2.0 Hz. The results are shown in Table 4.2. The transfer function was rechecked every few days throughout the experimental phase of the study. The results did not vary substantially from those shown in Table 4.2. All signal amplitudes quoted in this report have been corrected for the motion base transfer function, and therefore indicate actual acceleration or angular rate.

Some typical samples of measured motions are shown in Figures 4.5 to 4.9. Figure 4.5a shows a time trace of a 7.0 milli-g surge signal, with no added noise. Figure 4.5b shows a typical sample of broadband noise, again on the surge channel. Figure 4.5c shows the case where both signal and noise are present. Power spectra for the cases of signal alone and signal plus noise are shown in Figure 4.6. Each spectrum is based on a 10-second
sampling interval during which the signal amplitude was constant. The theoretical noise spectrum is also shown in Figure 4.6b.

Figures 4.7 through 4.9 show similar measurements of pitch motion. Figure 4.7 shows typical measured time histories of pitch rate. Figure 4.8 shows the corresponding time histories of pitch acceleration. Finally, power spectra for the signal alone and the signal plus noise cases are shown in Figure 4.9. It is interesting to note that for both cases shown (surge and pitch), the index of detectability is about 1.7.
Table 4.1  Maximum displacement, velocity and acceleration of the UTIAS Flight Research Simulator.

<table>
<thead>
<tr>
<th>Axis</th>
<th>Maximum Displacement</th>
<th>Maximum Velocity</th>
<th>Maximum Acceleration</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>0.61 m</td>
<td>1.0 m/s</td>
<td>15 m/s²</td>
</tr>
<tr>
<td>Y</td>
<td>0.59 m</td>
<td>1.0 m/s</td>
<td>15 m/s²</td>
</tr>
<tr>
<td>Z</td>
<td>0.49 m</td>
<td>1.0 m/s</td>
<td>15 m/s²</td>
</tr>
<tr>
<td>P</td>
<td>21 °</td>
<td>30 °/s</td>
<td>400 °/s²</td>
</tr>
<tr>
<td>Q</td>
<td>20 °</td>
<td>30 °/s</td>
<td>400 °/s²</td>
</tr>
<tr>
<td>R</td>
<td>24 °</td>
<td>30 °/s</td>
<td>400 °/s²</td>
</tr>
</tbody>
</table>
Table 4.2  Combined transfer function of the motion drive algorithm and the motion base, at selected frequencies. Values given are $|H_{ij}(j\omega)|$.

Units of $H_{ij}(j\omega)$:

1. translation/translation - dimensionless
2. translation/rotation - $(m/s^2)/(rad/s^2)$
3. rotation/translation - $(rad/s^2)/(m/s^2)$
4. rotation/rotation - dimensionless

<table>
<thead>
<tr>
<th>Driven Axis</th>
<th>Freq. (Hz)</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>P</th>
<th>Q</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>0.2</td>
<td>0.842</td>
<td>0.131</td>
<td>0.008</td>
<td>0.014</td>
<td>0.002</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>1.030</td>
<td>0.086</td>
<td>0.008</td>
<td>0.028</td>
<td>0.010</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>1.073</td>
<td>0.060</td>
<td>0.016</td>
<td>0.032</td>
<td>0.019</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>1.044</td>
<td>0.061</td>
<td>0.019</td>
<td>0.036</td>
<td>0.022</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>1.010</td>
<td>0.045</td>
<td>0.021</td>
<td>0.025</td>
<td>0.028</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>0.894</td>
<td>0.039</td>
<td>0.044</td>
<td>0.027</td>
<td>0.042</td>
<td>0.020</td>
</tr>
<tr>
<td>Y</td>
<td>0.2</td>
<td>0.049</td>
<td>0.854</td>
<td>0.006</td>
<td>0.002</td>
<td>0.017</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.008</td>
<td>1.095</td>
<td>0.017</td>
<td>0.010</td>
<td>0.038</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.039</td>
<td>1.188</td>
<td>0.025</td>
<td>0.024</td>
<td>0.045</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.020</td>
<td>1.190</td>
<td>0.042</td>
<td>0.045</td>
<td>0.049</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.042</td>
<td>1.143</td>
<td>0.038</td>
<td>0.039</td>
<td>0.053</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>0.035</td>
<td>0.967</td>
<td>0.075</td>
<td>0.078</td>
<td>0.046</td>
<td>0.023</td>
</tr>
<tr>
<td>P</td>
<td>0.2</td>
<td>0.134</td>
<td>0.109</td>
<td>0.062</td>
<td>0.826</td>
<td>0.090</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.046</td>
<td>0.108</td>
<td>0.016</td>
<td>0.950</td>
<td>0.091</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.021</td>
<td>0.064</td>
<td>0.026</td>
<td>0.979</td>
<td>0.094</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.145</td>
<td>0.021</td>
<td>0.032</td>
<td>1.028</td>
<td>0.080</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.072</td>
<td>0.039</td>
<td>0.041</td>
<td>0.983</td>
<td>0.091</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>0.083</td>
<td>0.058</td>
<td>0.045</td>
<td>0.988</td>
<td>0.082</td>
<td>0.019</td>
</tr>
<tr>
<td>Q</td>
<td>0.2</td>
<td>0.131</td>
<td>0.304</td>
<td>0.053</td>
<td>0.096</td>
<td>0.823</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.122</td>
<td>0.107</td>
<td>0.056</td>
<td>0.054</td>
<td>0.976</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.085</td>
<td>0.049</td>
<td>0.047</td>
<td>0.018</td>
<td>1.047</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.055</td>
<td>0.019</td>
<td>0.032</td>
<td>0.021</td>
<td>1.046</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.061</td>
<td>0.058</td>
<td>0.041</td>
<td>0.029</td>
<td>1.057</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>0.106</td>
<td>0.032</td>
<td>0.016</td>
<td>0.037</td>
<td>1.004</td>
<td>0.016</td>
</tr>
</tbody>
</table>
Figure 4.1  Block diagram of the proposed motion perception model.
Figure 4.2   The UTIAS Flight Research Simulator.
Figure 4.3  The experimental workstation.
Figure 4.4  Time histories of acceleration, velocity and displacement for a gated sinusoidal signal with frequency 0.6 Hz. For the signal shown, \( T_s = 10 \) seconds, and \( T_r = 1.67 \) seconds.
Figure 4.5  Typical measured time histories of surge specific force.
Figure 4.6  Typical measured power spectra of surge specific force (same conditions as Figure 4.5).
a) signal alone (amplitude 1°/sec)

b) noise alone (condition B — see Chapter 6)

c) signal plus noise

Figure 4.7 Typical measured time histories of pitch rate.
Figure 4.8  Time histories of pitch acceleration, based on the pitch rate time histories of Figure 4.7.
Figure 4.9  Typical measured power spectra of pitch acceleration (same conditions as Figures 4.7 and 4.8).
Chapter 5

Determination of ROC Curves

5.1 Introduction

This chapter describes the first of three experiments run to test some aspects of the signal detection model for motion perception. The objective of this experiment was to determine the shape of the motion detection receiver operating characteristic (ROC) curve, for comparison with model predictions. Recall from Chapter 3 that the high threshold model predicts a straight-line ROC, while the signal detection models predict a curved ROC. These curved ROC's may be either symmetrical, as predicted by the correlation detector, or skewed, as predicted by the envelope and energy detectors.

Each experiment was run as a series of trials. During each trial, the subject was exposed to either noise alone, or signal plus noise. After each trial, the subject rated the subjective likelihood that a signal was present on a scale from 1 to 5. The data were analyzed to give four points on the ROC curve.

A total of twenty ROC curves were obtained, each based on about 430 trials. Five subjects were tested in four representative signal conditions: pitch and surge, at 0.6 Hz, at two different amplitudes. In all cases, the noise was a multi-axis, low-amplitude, broadband random motion.

The experimental ROC data were compared with various models using a maximum likelihood curve-fitting procedure. The ROC's were very definitely curvilinear, indicating that the high threshold model is inappropriate for the conditions studied. The curves were slightly skewed, rather than symmetrical, implying that the energy detector or envelope detector is probably more applicable than the correlation detector.

The experimental conditions and the rating procedure used to obtain the ROC data are described in more detail in Section 5.2. The results are presented and discussed in Section 5.3.
5.2 Experimental Method and Conditions

The objective of this experiment was to determine the shape of the ROC curve of the human observer as a detector of whole-body motion. The ROC curves obtained describe an observer's ability to detect a sinusoidal motion signal superimposed on a background of random motion (noise). Five subjects were tested in four different motion conditions, yielding a total of twenty ROC curves.

Each subject participated in a total of 20 production runs, five for each motion condition (except for two cases where, due to scheduling difficulties, only four were run). There were 80 to 90 trials per run. The ordering of the conditions for each subject was randomized to minimize any effects due to order of presentation. Prior to the start of the production runs, each subject had two or three training runs for each condition. The purpose of the training runs was to familiarize the subjects with the rating task and the signal and noise conditions, to ensure that the data obtained were from practiced, consistent observers. Each subject took part in five or six runs per week over a period of six weeks. The total testing time for each subject, including both training and production runs, was about 20 hours.

The data were collected using the rating procedure. Each experimental run consisted of a series of trials. On each trial, the subject was exposed to a motion stimulus which was either noise alone, or signal plus noise. The subject's task was to estimate the likelihood that a signal was present, and to rate that likelihood on a scale from 1 to 5, according to the following criteria:

1. There almost certainly was NOT a signal present.
2. There probably was NOT a signal present.
3. It is equally likely that there was or was not a signal.
4. There probably WAS a signal present.
5. There almost certainly WAS a signal present.

The result from each run was a table of ten numbers indicating the number of times the subject selected each of the five responses when the stimulus was noise alone, and when the stimulus was signal plus noise. These data can be interpreted to give four points on the ROC.
Each trial proceeded as shown in Figure 5.1. The random noise motion was continuous over the duration of the run. A warning tone indicated the start of the trial. The next 16 seconds was the "signal interval". On about half of the trials, a gated sinusoidal signal with a duration of 10 seconds was superimposed on the noise during this interval (see Section 4.5). At the end of the signal interval, a buzzer sounded to prompt the subject for a response. The subject entered his response by pressing a button on a lap-held keyboard. A tone acknowledged the subject's response and provided feedback. A high tone indicated that there had, in fact, been a signal present; a low tone indicated that there had not. Following this feedback tone, there was a delay of about 5 seconds before the next trial began.

The signal and noise conditions were kept constant throughout each experimental run, and the subject was told at the beginning of the run what conditions to expect. The probability of a signal occurring on any particular trial was 0.50, and was independent of all other trials. The run was continued either until completion of 40 trials with a signal and 40 without, or until completion of 90 trials, whichever came first. On average, each run comprised about 86 trials and took 35 minutes to complete.

Details about the subjects and the experimental conditions are provided in Section 5.2.1. Section 5.2.2 describes the rating procedure in more detail, and explains how the rating data are analyzed to give points on the ROC curve. Finally, Section 5.2.3 discusses how the observers' response criteria can be influenced in order to obtain points which are well distributed along the ROC curve.

5.2.1 Conditions

Five male research workers in the flight simulation laboratory served as experimental subjects. This subject pool included both the experimenter and his assistant. All subjects were between the ages of 24 and 31, were in good health, and had no history of vestibular disorder or other inner ear trouble. All subjects had previous experience in motion perception studies, amounting to about 20 hours of testing time for each. The rating task was new to all subjects. A summary of the relevant data for each subject is given in Table 5.1. Subjects were paid $5.00 per hour for participating in the experiment.

Each subject was tested under four different experimental conditions, as summarized in Table 5.2. Two different signal axes were considered: pitch and surge. This allowed us
to collect data for both a rotation (pitch), where sensation is dominated by the response of
the semicircular canals, and a translation, where sensation is dominated by the otolths.

For each axis, two different signal amplitudes were used. These amplitudes were
selected individually for each subject based on his performance during the training runs.
One amplitude was set at a fairly low level which would yield a $d'$ value between 0.7 and
1.0. The other was set a bit higher, in an attempt to obtain a $d'$ value of about 1.5. The
signal detectability was kept in this range in order to obtain as much information as
possible about the shape of the ROC. For very weak or very strong signals, all of the
detection models predict ROC curves which approach the same theoretical limits.
Differences between the models are most pronounced at moderate detectability levels.
This implies that, for a given amount of data, we can draw stronger conclusions about the
form of the underlying model by using a signal amplitude which results in a moderate $d'$
value, say between 0.5 and 2.0.

The signal frequency in all cases was 0.6 Hertz. This value was selected rather arbitrarily
as being roughly in the middle of the range of interest for flight simulator motion. The
signal duration, excluding the fade-in and fade-out ramps, was 10 seconds. Again, this
was chosen in the range of interest for flight simulation. If a pilot has not responded to a
stimulus after ten seconds, then we might conclude that, for all intents and purposes, he
could not detect it.

In all four cases, the background motion or noise was a multi-axis, low amplitude, broad-
band random motion. The noise was present on all three translational degrees of freedom
and on all three rotational degrees of freedom. The motion on each degree of freedom
was uncorrelated with that on the other five. This noise condition is referred to in later
chapters as spectrum A.

Figure 5.2 illustrates the power spectrum of the noise on each axis. Expressed in terms of
acceleration, the power spectrum was flat from 0.25 Hz to 4.0 Hz, with a power spectral
density of 0.0002 $(m/s^2)^2/Hz$ for translations, and 0.0002 $(rad/s^2)^2/Hz$ for rotations. The
spectrum rolled off at 40 dB per decade above 4.0 Hertz, and at 60 dB per decade below
0.25 Hertz. The resulting measured RMS motion levels were approximately 0.003 g
along each translational axis, and 1.7°/s^2 about each rotational axis. The motion was
representative of flight through light turbulence. The rationale for choosing this noise
case is discussed in Chapter 6.
During each experimental run, the general physical, visual, and auditory conditions were as described in Chapter 4.

5.2.2 The Rating Procedure

The rating procedure used in this experiment yields several points on the ROC curve in a single experimental run. This makes it much more efficient than the simpler binary choice procedure.

The simplest psychophysical procedure is the binary choice, or YES/NO procedure. A YES/NO experiment is run as a series of trials. On each trial, the subject is exposed to a stimulus which is either noise alone, or signal plus noise. After each trial, the subject makes a binary decision: yes, there was a signal present, or no, there was not. The subject's performance over a series of trials may be summarized by the stimulus response matrix discussed in Chapter 3. Plotting the hit rate, $P(S \mid s)$, versus the false alarm rate, $P(S \mid n)$, gives one point on the ROC curve. To obtain another point on the ROC, we must induce the subject to change his response criterion, then rerun the experiment.

The rating procedure is identical to the YES/NO procedure in every respect except the subject's task. For this procedure, he must select a response from a field of, say, 5 categories. The first category represents near certainty that a signal was not present (the stimulus was noise alone), and the last represents near certainty that a signal was present (the stimulus was signal plus noise). The intermediate categories represent smaller degrees of certainty about the occurrence or non-occurrence of the signal. Thus, the subject's response reflects his estimate of the likelihood that a signal was present. The subject's performance over a series of trials may be summarized by the stimulus-response matrix shown in Figure 5.3.

To interpret the rating data, we consider each decision boundary between categories as a different response criterion. We can then analyze the data with respect to that criterion. First, consider the criterion represented by the boundary between categories 4 and 5. We assume that, given a binary decision task, the observer adopting this criterion would say YES only on the trials where he selects category 5, "there almost certainly was a signal
present". In this case:

$$P_d(S \mid s) = P(5 \mid s)$$
and
$$P_d(S \mid n) = P(5 \mid n)$$  \hspace{1cm} (5.1)

These probabilities define a single point on the ROC curve. Next, consider the boundary between categories 3 and 4. We now interpret both responses 4 and 5 as YES, so that for this criterion:

$$P_3(S \mid s) = P(4 \mid s) + P(5 \mid s)$$
and
$$P_3(S \mid n) = P(4 \mid n) + P(5 \mid n)$$  \hspace{1cm} (5.2)

These values define a second point on the ROC curve.

The remaining points on the ROC curve may be found by analyzing the data with respect to the remaining criteria. In the general case, for an m-category rating procedure, the hit rate and false alarm rate defined by the boundary between category j and category j+1 are:

$$P_j(S \mid s) = \sum_{k=j+1}^{m} P(k \mid s)$$  \hspace{1cm} (5.3a)

$$P_j(S \mid n) = \sum_{k=j+1}^{m} P(k \mid n)$$  \hspace{1cm} (5.3b)

Note that the number of points obtained, excluding the endpoints, is one less than the number of categories. Also note that the endpoints of the ROC are defined consistently by equation 5.3 by setting j equal to 0 or m:

$$P_0(S \mid s) = P_0(S \mid n) = 1$$
$$P_m(S \mid s) = P_m(S \mid n) = 0$$  \hspace{1cm} (5.4)

The primary advantage of the rating procedure over the binary choice procedure is its superior efficiency. Obtaining four points on an ROC curve requires four runs using the YES/NO procedure, and only one using the rating procedure. The trade-off is that the observer must perform a more difficult task. Our implicit assumption is that the subject
can maintain several distinct response criteria throughout the experiment, and that his response is consistent in terms of those criteria. Green and Swets (1974) discuss this issue and conclude that the two procedures give comparable results.

### 5.2.3 Influencing the Response Criteria

The five-category rating procedure used for this experiment provides four points on the ROC curve. In order to obtain as much information as possible about the shape of the ROC, we would like to ensure that these four points are well distributed along the curve. Four points clustered close together provide little more information than a single point. Similarly, points which are very near to the endpoints of the ROC provide little information, since all the models predict ROC curves which pass through (0,0) and (1,1).

To achieve this goal, we encouraged the subjects to adopt response criteria which would yield well-distributed points. As noted in Chapter 3, observers can be induced to adjust their criteria to meet verbal specifications or to maximize the expected value of a utility or payoff function. Both of these methods were used here.

Subjects were instructed to respond according to the following scale:

1. There almost certainly was NOT a signal present.
2. There probably was NOT a signal present.
3. It is equally likely that there was or was not a signal.
4. There probably WAS a signal present.
5. There almost certainly WAS a signal present.

To help the subjects interpret these categories more specifically, we set up a payoff matrix and scored subjects on their performance on each run. Subjects could maximize their scores by maintaining a set of distinct, well-distributed response criteria.

The concept of a payoff matrix deserves some discussion. It was shown in Chapter 3 that the expected value of the payoff in a binary YES/NO decision is maximized by selecting a criterion $\beta$ such that:

$$
\beta = \frac{P(n) (V_{nN} - V_{nS})}{P(s) (V_{sS} - V_{sN})}
$$

(5.5)
where $V_{xY}$ is the value associated with responding Y when the stimulus is x. For each observation, the optimal strategy is to respond yes (S) if $L(e) \geq \beta$, or no (N) if $L(e) < \beta$.

This may be extended to the rating procedure. The optimal response criterion for decision boundary j is:

$$\beta_j = \frac{P(n) (V_{nj} - V_{nj+1})}{P(s) (V_{sj+1} - V_{sj})} \quad (5.6)$$

For each observation, the optimal strategy is to select response category $j+1$ (or higher) if the likelihood ratio is greater than $\beta_j$, and category j (or lower) if it is less than $\beta_j$. For consistency with the verbal definitions of the categories given above, the optimal criteria defined by equation 5.6 must be strictly increasing with j. A proof of equation 5.6 is given in Appendix E.

For this experiment, we used the payoff matrix shown in Figure 5.4a. The resulting optimal decision criteria range from 0.35 to 2.86, and are strictly increasing from left to right. Figure 5.4b shows a hypothetical ROC corresponding to a signal with $d' = 1.0$. The points shown would yield the maximum score.

The scoring method and the payoff matrix were explained carefully to the subjects at the beginning of the experiment. Also, subjects were instructed that they could maximize their score by adopting the following response policy, where the response criteria are defined explicitly in terms of likelihood ratio.

<table>
<thead>
<tr>
<th>Estimated likelihood</th>
<th>Optimal Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(s \mid e):P(n \mid e)$</td>
<td></td>
</tr>
<tr>
<td>less than 1:3</td>
<td>1</td>
</tr>
<tr>
<td>1:3 to 2:3</td>
<td>2</td>
</tr>
<tr>
<td>about 1:1</td>
<td>3</td>
</tr>
<tr>
<td>3:2 to 3:1</td>
<td>4</td>
</tr>
<tr>
<td>more than 3:1</td>
<td>5</td>
</tr>
</tbody>
</table>

Subjects were coached during the training runs to develop several distinct response criteria. After each training run, the subject was given a numerical score based on the payoff matrix of Figure 5.4, and shown how to improve that score by adjusting his
criteria nearer to the optimal values. During production runs, subjects were encouraged to apply their criteria consistently over the course of each run and between runs.

The experimental results show that this incentive scheme was fairly successful. A comparison of the observers' actual criterion values with the optimal values is shown in Figure 5.5. The criteria shown are criteria for a decision rule based on the likelihood ratio. The optimal values are based on the payoff matrix, as discussed above. To determine the actual values, the experimental data were fitted with a theoretical curve based on Gaussian distributions of signal and noise. The slope of the curve at each data point was taken to be the observer's criterion for that decision boundary.

Figure 5.5 shows that the subjects adopted response criteria which were reasonably close to the optimal values. The two lines in Figure 5.5 show the average criteria used. One line is the average over all subjects and conditions where the signal was weak (dA' less than one); the other is the average over all subjects and conditions for signals with dA' greater than one. It is interesting to note that subjects adopted slightly different criteria for weak signals than for strong signals. For weak signals, the average slope of the curve is less than one, indicating that subjects tend to use less extreme values of β than optimal. For strong signals, they tended to use more extreme criteria than optimal. With a weak signal, the subject is seldom certain about the occurrence or non-occurrence of the signal, so he should seldom respond 1 or 5. Conversely, with a strong signal, the subject should usually respond 1 or 5. The tendency to distribute responses evenly over the five categories leads to the observed results. This turns out to be beneficial, since it does lead to a more uniform distribution of points along the ROC curve.

5.3 Experimental Results

The experiment described above comprised a total of 98 production runs. These runs covered 20 different combinations of independent variables: five subjects and four experimental conditions. In all cases but two, five production runs were conducted for each subject and condition. In the other two cases, only four were run. The experimental result from each run was a stimulus-response matrix like that shown in Figure 5.3, which summarizes the subject's response over a series of 80 to 90 trials.

The analysis of the data included three main stages. First, the data were reviewed to see if the subjects' performance was consistent, both from run to run, and over the duration
of each run. After the data were found to be consistent, a summary stimulus-response matrix was calculated for each subject and condition by summing the stimulus response matrices from individual runs. Experimental ROC points were calculated from the summary matrices.

The second stage of the analysis considered the general shape of the ROC curves. Several different theoretical models were fit to the experimental points. A \( \chi^2 \) test was used to determine which models differed significantly from the data. Curved, slightly skewed ROC's such as those predicted by the envelope detector and energy detector provided the best fit. The symmetrical curves predicted by the correlation detector did not fit the data as well. The straight-line ROC did not fit the data at all.

The final step of the analysis was to determine the slope of the psychometric curve. This was found by plotting the area under the ROC curve versus signal amplitude for each subject and axis, and fitting these points with a generic psychometric function known as the logistic function. The data suggest a psychometric curve which is steeper than that predicted by the correlation detector, but not quite as steep as those predicted by the envelope and energy detectors.

Section 5.3.1 outlines the consistency checks; Section 5.3.2 presents the comparison of theoretical ROC curves to the data; Section 5.3.3 discusses the slope of the psychometric curve.

5.3.1 Consistency Checks

Each subject participated in about 20 production runs: five runs in each of the four experimental conditions. The object of conducting five runs for each condition was to obtain enough data to provide relatively low variance estimates of the signal detectability and ROC curve shape. Our intent was to sum the stimulus-response matrices over the five runs for each subject and condition, and analyze the pooled data.

However, before pooling the data, we wanted to ensure that the data for each subject and condition were reasonably consistent over the duration of the experiment. The data for each subject were collected over a period of about five weeks. It seemed quite plausible that there might be some learning effect or change of response strategy over this time. Also, each run lasted about 35 minutes. This suggested that there might be some degra-
5.11

dation of performance toward the end of a run due to fatigue, boredom or adaptation. Therefore the data were checked to see if the subjects’ performance was reasonably stationary.

Stationarity was tested by calculating the run-to-run variance of several parameters. The null hypothesis was that the subjects’ performance was stationary from run to run. If the actual variance obtained was significantly larger than that expected for a true stationary process, the null hypothesis was rejected, and we concluded that performance did vary from run to run.

Three different parameters were considered: the index of detectability, \(d'\); the ROC slope parameter, \(b\); and the actual probabilities associated with each ROC point, \(P_k(S|s)\) and \(P_k(S|n)\). The variance of \(d'\) and \(b\) was not significantly larger than expected, indicating that the overall performance was reasonably stationary. The average variance of \(P_k(S|s)\) and \(P_k(S|n)\) was about 40% higher than that expected for a stationary observer. An F-test showed this value to be statistically significant (\(\alpha < 0.05\)). This suggests that subjects did not hold their response criteria exactly constant. The effect of this variability was assessed by comparing the average values of \(b\) and \(d'\) found for each subject and condition with the values determined from the pooled data. No significant effect was observed for either \(b\) or \(d'\). While the observers did not hold their criteria perfectly constant, they did not vary enough to preclude pooling the data.

The combined data were analyzed to determine if performance varied over the duration of the run. Each data set was divided into three subsets. The first subset included the first 30 trials of each run; the second included trials 31 to 60 of each run; the third included trial 61 through the end of each run. The three subsets were compared to each other following the same procedure outlined above for the comparison between runs. For all of the parameters tested, the variability between subsets was, on average, about equal to the expected variability. On average, subjects performed equally well in each third of each run.

5.3.2 Fitting Models to the Data

This section describes the analysis of the data in terms of the general shape of the ROC curves. The purpose of this analysis was to determine which, if any, of the models predict an ROC shape which is consistent with the data.
Following the consistency checks outlined above, a summary stimulus-response matrix was calculated for each subject and condition by summing the stimulus-response matrices from the individual runs. Each summary data set was processed to produce four points on an ROC curve, using equation 5.3. The summary stimulus-response matrices and ROC graphs for all 20 cases are given in Appendix G.

Theoretical ROC curves based on threshold and detection models were fit to the data. Three different models were considered: the high threshold model, the Gaussian model, and the envelope detection model. The Gaussian model is based on Gaussian distributions of $f(e \mid n)$ and $f(e \mid s)$, and includes the correlation detector and the energy detector as special cases.

Curve fitting for the high threshold model was done by the method of least squares. Curve fitting for the Gaussian model and the envelope detection model was done using the maximum likelihood procedure described in Appendix F. This fitting procedure was developed by Dorfman and Alf (1969) to fit Gaussian ROC curves with unequal variance to rating-procedure data. As part of the current study, the procedure was extended to two other cases: Gaussian ROC curves with a fixed slope, $b$; and ROC curves based on the envelope detector. In each case, the model parameters were selected to maximize the likelihood of obtaining the experimental results actually recorded.

The "goodness of fit" for each model was evaluated using a $\chi^2$ test. A fit variable was calculated for each data set. This was defined as:

$$\gamma_i = N_n \sum_{j=1}^{m} \left( \frac{(R_{nj} - P_{nj})^2}{P_{nj}} \right) + N_s \sum_{j=1}^{m} \left( \frac{(R_{sj} - P_{sj})^2}{P_{sj}} \right)$$  (5.7)

where

- $N_n$ is the number of trials with noise alone,
- $N_s$ is the number of trials with signal plus noise,
- $R_{nj}$ is the observed probability of selecting response $j$ when the input is noise alone,
- $P_{nj}$ is the calculated probability of selecting response $j$ when the input is noise alone,
- $R_{sj}$ is the observed probability of selecting response $j$ when the input is signal plus noise, and
P_{sj} is the calculated probability of selecting response j when the input is signal plus noise.

Note that R_{nj} and R_{sj} are calculated directly from the response data, while P_{nj} and P_{sj} are calculated from the fitted model, based on maximum likelihood estimates of the response boundaries. If the model fits the data, then γ_i is a χ^2 variable with m−k−1 degrees of freedom, where m is the number of response categories, and k is the number of model parameters which can be varied.

This may be verified as follows. If A_1 is a χ^2 variable with n_1 degrees of freedom and A_2 is χ^2 with n_2 degrees of freedom, then A_1 + A_2 is χ^2 with n_1 + n_2 degrees of freedom (Hogg and Craig, 1970). Each summation in equation 5.7 is χ^2 with m−1 degrees of freedom. Therefore the right hand side of equation 5.7 has, initially, 2(m−1) degrees of freedom. From this we must subtract the number of parameters which can be varied to improve the fit: m−1 decision boundaries and k model parameters. Therefore, γ_i is χ^2 with m−k−1 degrees of freedom. Further discussion of the multivariate maximum likelihood estimator is provided by Kendall and Stuart (1961).

The evaluation of the fit is a standard χ^2 test. We hypothesize that the model fits the data and that, as a result, γ_i is χ^2 with m−k−1 degrees of freedom. We then examine the distribution of the γ_i obtained from the 20 curve fits. If this distribution is approximately χ^2 with m−k−1 degrees of freedom, then we accept the hypothesis. If the γ_i distribution lies significantly to the right of the χ^2 distribution, then we must reject the hypothesis, and conclude that the model does not fit the data. A good summary test may be done, based on a global fit variable:

$$\gamma_g = \sum_{i=1}^{20} \gamma_i$$

If the model fits the data, then by the above rule this variable is χ^2 with 20(m−k−1) degrees of freedom. The probability of exceeding γ_g, given that the hypothesis is true, is:

$$\alpha = P(\chi^2 > \gamma_g)$$

If α is sufficiently small, then we must conclude that the model does not fit the data.
The first model tested was the straight-line ROC based on the high threshold model. Figure 5.6 shows a comparison between this model and a typical data set. The solid line shown is the straight-line ROC which fits the data best in the sense of least squared deviation. The fit is very poor. Clearly, the data show a curvilinear trend which cannot be fit by a straight line. The value of $\gamma_i$ for this fit is 51. With this model, there is only one parameter which can be varied (the value of $P(S|I)$ when $P(S|N) = 0$. Thus from the discussion above, $\gamma_i$ is $\chi^2$ with three degrees of freedom if the underlying ROC curve is actually a straight line. The probability that a $\chi^2$ variable with three degrees of freedom will exceed 51 is much less than 0.0001. Because the fit is this poor for all 20 data sets, and because the deviation from the model is systematic, we can reject the high threshold model with almost perfect confidence.

The second model considered was based on Gaussian distributions of signal and noise. The assumption that the underlying distributions are Gaussian leads to a family of ROC curves with two parameters:

$$A = \frac{\mu_s - \mu_n}{\sigma_s} \quad (5.10)$$

$$b = \frac{\sigma_n}{\sigma_s} \quad (5.11)$$

A is a measure of the distance between $f(e|n)$ and $f(e|s)$, thus gives an indication of overall detectability. b is the ROC slope parameter, the slope of the ROC when plotted on normal probability axes. This is a very general model which includes the correlation detector ($b = 1$) and the energy detector ($b < 1$) as special cases.

Figure 5.7 shows the example data fit by an ROC curve based on the Gaussian model. A maximum likelihood curve fitting routine was used, in which both $A$ and $b$ were varied to produce the best fit. The curve shown in Figure 5.7 is a good fit to the data. The curved ROC shape predicted by this model agrees much better with the data than the straight-line ROC considered earlier. This model has two parameters. Therefore, if the model fits the data, $\gamma_i$ is $\chi^2$ with 2 degrees of freedom, and $\gamma_g$ is $\chi^2$ with 40 degrees of freedom. For the example data, $\gamma_i = 0.17$, which indicates almost a perfect fit. The global fit variable $\gamma_g$ has a value of 46.4 ($\alpha = 0.23$, 40 d.o.f), which indicates that the data do not differ significantly from the model.
5.15

ROC curves for all 20 experimental data sets, fit with this Gaussian model, are provided in Appendix G. A summary of the estimated model parameters and their associated variances, for all 20 data sets, is given in Table 5.3.

It is worthwhile to study the statistics of the ROC slope parameter b, because it reflects the shape of the ROC curves. Table 5.3 provides a summary of the estimated ROC slopes. These estimates range from 0.8 to 1.0, although one subject did have two higher values, 1.06 and 1.29. The standard deviation of the estimate of 1.29 is high (0.21), since the four ROC data points are clustered fairly close together. An analysis of variance showed that the slope does not vary significantly between subjects, and is not a function of signal axis or amplitude. The geometric mean is 0.90. This implies that $\sigma_s > \sigma_n$, and indicates that the ROC curves are slightly skewed.

The variance of the slope estimates is about 0.012. This compares closely with the average variance of each estimate of b, which is also 0.012. We conclude that the ROC slope parameter is approximately the same for all subjects and all conditions, and that the observed variation reflects primarily the uncertainty in estimating b from a limited number of trials. As noted above, the mean value is 0.90. The standard deviation of the mean is 0.025. A 95% confidence interval for the mean is the sample mean plus or minus two standard deviations. Therefore:

$$P(0.85 < \bar{b} < 0.95) = 0.95$$

This gives the likely range of the ROC slope parameter, b.

The likely range of b can also be estimated by a $\chi^2$ test. First, assume that b is the same for all data sets. Then, for each value of b, refit the model to the data allowing A to vary but holding b fixed. In this case, there is only one model parameter varied in each fit. If the model fits the data, the global fit parameter is $\chi^2$ with 60 degrees of freedom.

Figure 5.8 shows the variation of $\gamma_g$ with b. The right hand axis shows the probability of exceeding $\gamma_g$ assuming that the model fits. The curve is roughly parabola with a minimum near 0.90. For $b = 0.9$, $\gamma_g = 60$ and $\alpha = 0.49$. Therefore, the model does fit all the data well if we assume $b = 0.9$. For values of b greater or less than 0.9, $\gamma_g$ increases, indicating a poorer fit. It is quite unlikely that b is much less than 0.8 or much greater than 1.0.
A particular case that is worth noting is the case where $b = 1$. This gives the symmetrical ROC predicted by the correlation detector. The value of $\gamma_g$ found for this case is 83. The probability of exceeding this value with a $\chi^2$ distribution with 60 degrees of freedom is less than 0.025. Therefore, we can reject the correlation detection model with a significance level of $\alpha = 0.025$. Figure 5.9 shows the example data compared with the equal variance ROC. Note the systematic deviation of the data from the curve.

Finally, the experimental data were compared with theoretical ROC curves based on the envelope detector. Figure 5.10 compares the example data with the best fit envelope detector ROC. This model was expected to fit the data well, since the above analysis shows that the experimental ROC curves are slightly skewed. As with the Gaussian model, a maximum likelihood procedure was used to fit the data. With this model, there is only one parameter to vary: the distance between the distributions, expressed nondimensionally as $A/\sigma_r$. Therefore the fit variable $\gamma_i$ is $\chi^2$ with 3 degrees of freedom, and $\gamma_g$ is $\chi^2$ with 60 degrees of freedom. The value of $\gamma_g$ is 63 ($\alpha = 0.37$, 60 d.o.f.). Therefore, the data do not vary significantly from the model. A summary of the parameter estimates for this model is provided in Table 5.4.

The foregoing analysis has shown that the envelope detector and the Gaussian model with unequal variance both provide a good fit to the data. For the Gaussian model, the ROC slope parameter, $b$, does not vary substantially between subjects or between conditions. Using a value of 0.9 provides an acceptable fit to all data sets. A 95% confidence interval for $b$ is (0.85, 0.95). This value of $b$ (less than 1) is consistent with the energy model. The high threshold model does not fit the data and can be rejected with nearly perfect confidence. The correlation detector, which predicts a symmetrical, curved ROC, does not yield a good fit. It was rejected with a significance of 0.025.

5.3.3 Psychometric Curves

The last stage of the analysis was to examine the absolute detection levels found. This fulfilled two purposes. The first was to determine the slope of the psychometric curve, for comparison with model predictions. Recall that the psychometric curve shows how detectability (expressed as area under the ROC) varies with signal amplitude. The second purpose was to determine the "critical amplitude," the amplitude for which $d_{A'} = 1.0$ and the area under the ROC curve is 0.76.
The area under each ROC curve, $P_A$, was estimated using the $A$ and $b$ parameters from the Gaussian model discussed in Section 5.3.2. It may be shown that the area under the ROC is:

$$P_A = P(e_n < e_s)$$

(5.12)

where $e_n$ is a sample from $f(e|n)$, the noise distribution, and $e_s$ is a sample from $f(e|s)$, the distribution of signal plus noise. A proof of this is given in Chapter 6. It follows directly that:

$$P_A = P\left(\frac{e_n - e_s - (\mu_n - \mu_s)}{\sqrt{\sigma_n^2 + \sigma_s^2}} < \frac{\mu_s - \mu_n}{\sqrt{\sigma_n^2 + \sigma_s^2}}\right)$$

$$= P\left(Z < \frac{\mu_s - \mu_n}{\sqrt{\sigma_n^2 + \sigma_s^2}}\right)$$

(5.13)

Since $e_n$ and $e_s$ are independent and Gaussian with means $\mu_n$ and $\mu_s$ and variances $\sigma_n^2$ and $\sigma_s^2$, then $e_n - e_s$ is Gaussian with mean $\mu_n - \mu_s$ and variance $\sigma_n^2 + \sigma_s^2$. Therefore, $Z$ is a standard normal variable and the area is:

$$P_A = N_z\left(\frac{\mu_s - \mu_n}{\sqrt{\sigma_n^2 + \sigma_s^2}}\right)$$

$$= N_z\left(A/\sqrt{1+b^2}\right)$$

(5.14)

The variance of an estimate of $P_A$ is given by:

$$\sigma_p^2 = \frac{\exp[-A^2/(1+b^2)]}{2\pi} \left(\frac{\sigma_A^2}{1+b^2} + \frac{A^2b^2\sigma_b^2}{(1+b^2)^3} - \frac{2Ab\sigma_{Ab}}{(1+b^2)^2}\right)$$

(5.15)

Table 5.5 summarizes the area and standard deviation of area found for each ROC curve.

The data relating $P_A$ to signal amplitude for each subject and axis were fit with the logistic function, a generic function commonly used to approximate the psychometric
curve (Taylor and Creelman, 1967). The form of the function is:

\[ P_A = 0.5 + \frac{0.5}{1 + \exp[-B(X-M)]} \]  \hspace{1cm} (5.16)

where \( X \) is the signal amplitude in dB, \\
\( M \) is the amplitude corresponding to the middle of the psychometric function, \\
and \( B \) is the slope.

This is a useful function, because it allows us to quantify the slope of the psychometric curve, for comparison with model predictions.

For regression analysis it is convenient to define the following transformation:

\[ Y = \ln \left( \frac{2P_A - 1}{2 - 2P_A} \right) \]  \hspace{1cm} (5.17)

Equation 5.16 can now be rearranged into a standard linear expression:

\[ Y = \bar{Y} + B(X - \bar{X}) \]  \hspace{1cm} (5.18)

where \( \bar{Y} = B(\bar{X} - M) \) \hspace{1cm} (5.19)

and \( \bar{X} \) is the mean of \( X \). The variance of \( Y \) is given by:

\[ \sigma_Y^2 = \frac{\sigma_p^2}{(2P-1)^2(1-P)^2} \]  \hspace{1cm} (5.20)

Maximum likelihood estimates for \( M \) and \( B \) and their associated variances were then found following standard linear regression procedures (Hogg and Craig, 1970, pp. 335-338). The results are tabulated in Table 5.5.

The parameter \( B \) indicates the slope of the psychometric curve. Observed values range from 0.12 to 0.40, with a mean of 0.28. There is no significant variation between subjects or axes. The standard deviation is 0.094, which compares closely to the standard deviation of each estimate, which is, on average, 0.089. We therefore conclude that the
slope is not significantly different for the different cases. The standard deviation of the mean slope is 0.031.

Figure 5.11 shows a comparison between the data and the logistic function with $B = 0.28$. The abscissa of each data point is the relative amplitude, $X - M$.

It is interesting to compare the data with psychometric curves predicted by the detection models. The correlation detector is closely approximated by a logistic function with a slope of 0.20; for the envelope detector, the slope is 0.33. For the energy detector, the slope is in the range 0.35 to 0.39, depending on its bandwidth and integration time. Therefore, the psychometric curve obtained from the data is steeper than that predicted by the correlation detector, but not quite as steep as those predicted by the envelope and energy detectors.

The critical amplitude for each subject and axis was calculated from equation 5.16 by setting $P_A = 0.76$, the area under the ROC when $d_A' = 1.0$. The values obtained are summarized in Table 5.6. The average value found for pitch was $0.34^\circ/s$, and that for surge was $0.026 \text{ m/s}^2 (0.0026 \text{ g})$. All subjects were about equally sensitive, except subject PR, who was significantly less sensitive to pitch rotation ($0.52^\circ/s$). These values are compared to results from the other two experiments and from the literature in Chapter 8.
### Table 5.1 Summary of subject data.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Age</th>
<th>Height (cm)</th>
<th>Weight (kg)</th>
<th>Participated in experiments*</th>
</tr>
</thead>
<tbody>
<tr>
<td>DG</td>
<td>25</td>
<td>183</td>
<td>84</td>
<td>*</td>
</tr>
<tr>
<td>PG</td>
<td>24</td>
<td>183</td>
<td>73</td>
<td>*</td>
</tr>
<tr>
<td>GG</td>
<td>31</td>
<td>178</td>
<td>73</td>
<td>*</td>
</tr>
<tr>
<td>MN</td>
<td>29</td>
<td>168</td>
<td>64</td>
<td>*</td>
</tr>
<tr>
<td>PR</td>
<td>25</td>
<td>178</td>
<td>68</td>
<td>*</td>
</tr>
<tr>
<td>DS</td>
<td>23</td>
<td>185</td>
<td>75</td>
<td>*</td>
</tr>
<tr>
<td>BW</td>
<td>33</td>
<td>190</td>
<td>110</td>
<td>*</td>
</tr>
</tbody>
</table>

* experiments were run in the order 3, 1, 2.
Table 5.2  Experimental conditions for each subject. Values shown are the signal amplitudes for each case.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Surge, 0.6 Hz (m/s²)</th>
<th>Pitch, 0.6 Hz (°/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Case 1</td>
<td>Case 2</td>
</tr>
<tr>
<td>DG</td>
<td>.022</td>
<td>.033</td>
</tr>
<tr>
<td>PG</td>
<td>.022</td>
<td>.033</td>
</tr>
<tr>
<td>GG</td>
<td>.027</td>
<td>.038</td>
</tr>
<tr>
<td>MN</td>
<td>.022</td>
<td>.033</td>
</tr>
<tr>
<td>PR</td>
<td>.027</td>
<td>.038</td>
</tr>
</tbody>
</table>
Table 5.3  Maximum likelihood estimates of parameters $A$ and $b$ for the Gaussian model, and the associated standard deviations.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Axis</th>
<th>Ampl.</th>
<th>$A$</th>
<th>$b$</th>
<th>$\sigma_A$</th>
<th>$\sigma_b$</th>
<th>$\sqrt{\sigma_{Ab}}$</th>
<th>$\gamma_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DG</td>
<td>X</td>
<td>.022</td>
<td>.892</td>
<td>.860</td>
<td>.107</td>
<td>.089</td>
<td>.059</td>
<td>5.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.033</td>
<td>2.100</td>
<td>1.292</td>
<td>.238</td>
<td>.208</td>
<td>.194</td>
<td>.98</td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.30</td>
<td>.901</td>
<td>1.057</td>
<td>.114</td>
<td>.100</td>
<td>.062</td>
<td>.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.42</td>
<td>1.151</td>
<td>.897</td>
<td>.118</td>
<td>.101</td>
<td>.076</td>
<td>1.85</td>
</tr>
<tr>
<td>PG</td>
<td>X</td>
<td>.022</td>
<td>.709</td>
<td>.796</td>
<td>.109</td>
<td>.081</td>
<td>.049</td>
<td>.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.033</td>
<td>1.258</td>
<td>.849</td>
<td>.122</td>
<td>.095</td>
<td>.079</td>
<td>5.84</td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.30</td>
<td>.852</td>
<td>.984</td>
<td>.112</td>
<td>.104</td>
<td>.064</td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.42</td>
<td>1.346</td>
<td>.848</td>
<td>.124</td>
<td>.089</td>
<td>.076</td>
<td>.70</td>
</tr>
<tr>
<td>GG</td>
<td>X</td>
<td>.027</td>
<td>.906</td>
<td>.851</td>
<td>.107</td>
<td>.081</td>
<td>.057</td>
<td>2.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.038</td>
<td>1.458</td>
<td>.926</td>
<td>.134</td>
<td>.105</td>
<td>.090</td>
<td>.82</td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.30</td>
<td>.776</td>
<td>.883</td>
<td>.107</td>
<td>.085</td>
<td>.056</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.42</td>
<td>1.348</td>
<td>.884</td>
<td>.124</td>
<td>.098</td>
<td>.082</td>
<td>3.01</td>
</tr>
<tr>
<td>MN</td>
<td>X</td>
<td>.022</td>
<td>.853</td>
<td>.826</td>
<td>.104</td>
<td>.082</td>
<td>.054</td>
<td>.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.033</td>
<td>1.183</td>
<td>1.020</td>
<td>.128</td>
<td>.127</td>
<td>.088</td>
<td>2.20</td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.24</td>
<td>.932</td>
<td>.886</td>
<td>.111</td>
<td>.091</td>
<td>.063</td>
<td>7.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.36</td>
<td>1.676</td>
<td>.796</td>
<td>.149</td>
<td>.110</td>
<td>.105</td>
<td>.63</td>
</tr>
<tr>
<td>PR</td>
<td>X</td>
<td>.027</td>
<td>.742</td>
<td>.866</td>
<td>.107</td>
<td>.087</td>
<td>.058</td>
<td>.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.038</td>
<td>1.149</td>
<td>.864</td>
<td>.134</td>
<td>.137</td>
<td>.105</td>
<td>4.99</td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.30</td>
<td>.538</td>
<td>.890</td>
<td>.103</td>
<td>.085</td>
<td>.051</td>
<td>5.44</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.42</td>
<td>.764</td>
<td>.883</td>
<td>.111</td>
<td>.091</td>
<td>.066</td>
<td>.30</td>
</tr>
</tbody>
</table>

Geometric mean .902  $\sum \gamma_i = 46.39$
Std. deviation .110
Table 5.4  Maximum likelihood estimates of parameter $A_*$ for the envelope detection model, and the associated standard deviation.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Axis</th>
<th>Amplitude</th>
<th>$A_*$</th>
<th>$\sigma_A$</th>
<th>$\gamma_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DG</td>
<td>X</td>
<td>.022</td>
<td>1.669</td>
<td>.115</td>
<td>4.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.033</td>
<td>2.530</td>
<td>.135</td>
<td>7.77</td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.30</td>
<td>1.562</td>
<td>.115</td>
<td>6.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.42</td>
<td>1.930</td>
<td>.113</td>
<td>2.34</td>
</tr>
<tr>
<td>PG</td>
<td>X</td>
<td>.022</td>
<td>1.494</td>
<td>.130</td>
<td>.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.033</td>
<td>2.082</td>
<td>.114</td>
<td>5.50</td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.30</td>
<td>1.543</td>
<td>.117</td>
<td>3.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.42</td>
<td>2.180</td>
<td>.115</td>
<td>.94</td>
</tr>
<tr>
<td>GG</td>
<td>X</td>
<td>.027</td>
<td>1.697</td>
<td>.116</td>
<td>1.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.038</td>
<td>2.231</td>
<td>.115</td>
<td>2.15</td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.30</td>
<td>1.520</td>
<td>.118</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.42</td>
<td>2.155</td>
<td>.115</td>
<td>2.34</td>
</tr>
<tr>
<td>MN</td>
<td>X</td>
<td>.022</td>
<td>1.638</td>
<td>.116</td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.033</td>
<td>1.894</td>
<td>.115</td>
<td>4.49</td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.24</td>
<td>1.693</td>
<td>.116</td>
<td>7.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.36</td>
<td>2.574</td>
<td>.122</td>
<td>.44</td>
</tr>
<tr>
<td>PR</td>
<td>X</td>
<td>.027</td>
<td>1.490</td>
<td>.119</td>
<td>.59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.038</td>
<td>1.942</td>
<td>.119</td>
<td>5.02</td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.30</td>
<td>1.212</td>
<td>.127</td>
<td>5.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.42</td>
<td>1.501</td>
<td>.118</td>
<td>.19</td>
</tr>
</tbody>
</table>

$\sum \gamma_i = 63.33$
Table 5.5  Values of $P_A$ and $\sigma_p$ based on the Gaussian model, and estimates of parameters M and B for the psychometric function.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Axis</th>
<th>Amplitude</th>
<th>$P_A$</th>
<th>$\sigma_p$</th>
<th>M(dB)*</th>
<th>B</th>
<th>$\sigma_B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DG</td>
<td>X</td>
<td>.022</td>
<td>.751</td>
<td>.024</td>
<td>6.76</td>
<td>.395</td>
<td>.081</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.033</td>
<td>.901</td>
<td>.017</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.30</td>
<td>.732</td>
<td>.024</td>
<td>10.26</td>
<td>.200</td>
<td>.091</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.42</td>
<td>.804</td>
<td>.021</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PG</td>
<td>X</td>
<td>.022</td>
<td>.710</td>
<td>.028</td>
<td>7.90</td>
<td>.282</td>
<td>.083</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.033</td>
<td>.831</td>
<td>.020</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.30</td>
<td>.728</td>
<td>.025</td>
<td>10.05</td>
<td>.342</td>
<td>.092</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.42</td>
<td>.848</td>
<td>.019</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>X</td>
<td>.027</td>
<td>.755</td>
<td>.024</td>
<td>8.58</td>
<td>.302</td>
<td>.089</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.038</td>
<td>.858</td>
<td>.018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.30</td>
<td>.720</td>
<td>.025</td>
<td>10.23</td>
<td>.353</td>
<td>.093</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.42</td>
<td>.844</td>
<td>.019</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MN</td>
<td>X</td>
<td>.022</td>
<td>.744</td>
<td>.024</td>
<td>7.14</td>
<td>.119</td>
<td>.076</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.033</td>
<td>.796</td>
<td>.022</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.24</td>
<td>.757</td>
<td>.024</td>
<td>7.46</td>
<td>.395</td>
<td>.077</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.36</td>
<td>.905</td>
<td>.015</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PR</td>
<td>X</td>
<td>.027</td>
<td>.712</td>
<td>.026</td>
<td>9.85</td>
<td>.264</td>
<td>.097</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.038</td>
<td>.808</td>
<td>.022</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.30</td>
<td>.656</td>
<td>.027</td>
<td>13.94</td>
<td>.180</td>
<td>.112</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.42</td>
<td>.717</td>
<td>.026</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>.283</th>
<th>.094</th>
</tr>
</thead>
</table>

* The 0 dB reference levels are: 0.01 m/s$^2$ for surge, and 0.1°/s for pitch.
Table 5.6 Estimated critical amplitudes for surge and pitch.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Surge, 0.6 Hz (m/s²)</th>
<th>Pitch, 0.6 Hz (°/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DG</td>
<td>0.022</td>
<td>0.34</td>
</tr>
<tr>
<td>PG</td>
<td>0.026</td>
<td>0.33</td>
</tr>
<tr>
<td>GG</td>
<td>0.028</td>
<td>0.33</td>
</tr>
<tr>
<td>MN</td>
<td>0.025</td>
<td>0.24</td>
</tr>
<tr>
<td>PR</td>
<td>0.032</td>
<td>0.52</td>
</tr>
<tr>
<td>Geometric mean</td>
<td>0.026</td>
<td>0.34</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.0037</td>
<td>0.10</td>
</tr>
</tbody>
</table>
Figure 5.1  Main events during a rating-procedure trial.
Figure 5.2  
Power spectrum for noise condition A.
For translations, $N_{\text{ref}} = 0.005 \text{ (m/s}^2\text{)}^2/\text{Hz}$.
For rotations, $N_{\text{ref}} = 0.005 \text{ (rad/s}^2\text{)}^2/\text{Hz}$. 
Figure 5.3  The stimulus-response matrix for a 5 category rating task.
Stimulus Noise alone Signal plus noise

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise alone</td>
<td>14</td>
<td>7</td>
<td>0</td>
<td>-10</td>
<td>-30</td>
</tr>
<tr>
<td>Signal plus noise</td>
<td>-30</td>
<td>-10</td>
<td>0</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>Optimal value of decision criterion, $\beta$</td>
<td>0.35</td>
<td>0.70</td>
<td>1.42</td>
<td>2.85</td>
<td></td>
</tr>
</tbody>
</table>

a) The payoff/penalty matrix.

![Figure 5.4](image)

b) Optimal distribution of ROC points, for $d'=1$.

Figure 5.4 Influencing the response criteria with a payoff matrix.
Figure 5.5  A comparison of the observed response criteria with the optimal criteria.
Figure 5.6  Example ROC data fit with the high-threshold model. Surge, 0.022 m/s², subject PG. Error bars indicate ± 1 standard deviation.
Figure 5.7  Example ROC data fit with the Gaussian model. Surge, 0.022 m/s², subject PG. Error bars indicate ± 1 standard deviation.
Figure 5.8  Variation of the global fit variable, $\gamma_g$, as a function of the ROC slope parameter, b.
Figure 5.9  Example ROC data fit with the Gaussian model, with the constraint that $b = 1$. Surge, 0.022 m/s$^2$, subject PG. Error bars indicate $\pm 1$ standard deviation.
Figure 5.10  Example ROC data fit with the envelope detector model. Surge, 0.022 m/s², subject PG. Error bars indicate ± 1 standard deviation.
Figure 5.11  Variation in signal detectability as a function of signal amplitude. The data are fit with a logistic function with $B = 0.28$. Error bars indicate $\pm 1$ standard deviation.
Chapter 6

The Effect of Noise Power and Frequency Content

6.1 Introduction

This chapter describes the second of the three motion detection experiments run. In the first experiment, only a single noise condition was considered. In this experiment, on the other hand, several different noise spectra were considered. The purpose of this experiment was to learn how the detectability of sinusoidal motion is influenced by the power and frequency content of the noise, or background motion. Recall that signal detection theory predicts that detectability is a function of signal-to-noise ratio.

A total of 28 cases were considered: two axes of motion (pitch and surge), at two frequencies (0.2 Hz and 0.6 Hz), with seven different background motion (noise) conditions. Three of the noise conditions were broadband spectra with different power densities. The other four were narrow band (one octave) spectra with different centre frequencies but the same power density. Five subjects were tested in all 28 conditions.

Each experiment was a series of forced choice trials. On each trial, the subject’s task was to determine which of two intervals contained a sinusoidal motion signal in addition to the random motion. The signal amplitude was adjusted using an adaptive algorithm designed to converge on the level which yielded 76% correct responses. This level is referred to as the critical amplitude. It will be shown below that this definition of critical amplitude is consistent with that used in Chapter 5.

The experimental results show that random motion does mask the signal. The signal-to-noise ratio at the critical amplitude is almost constant for the three broadband noise conditions. Results for the narrow band spectra show that sinusoidal motion is masked primarily by noise components near the signal frequency. An energy detection model with an internal pre-detection filter provides a good qualitative explanation of the results.

Section 6.2 describes the experimental conditions and the forced choice procedure. Section 6.3 presents and discusses the results.
6.2 Experimental Method and Conditions

The specific objective of this experiment was to determine the critical amplitude for a variety of signal and noise conditions. The critical amplitude is the signal amplitude which has a detectability, \(d'_A\), of 1.0. The area under the ROC curve at this amplitude is 0.76.

To get a complete picture of the influence of noise on signal detectability, it would be necessary to find the entire psychometric curve for each condition. In the interest of efficiency, we limited our investigation to finding a single point on the curve. We assume that the shape, or slope of the curve is invariant with condition. The critical amplitude was chosen as the target point because it is near the middle of the psychometric curve. Since the slope of the curve is steepest near its midpoint, the critical amplitude is a good indicator of the curve’s position along the abscissa. Also, the critical amplitude is a level which is "just detectable," so it is more or less comparable to "thresholds" quoted in earlier studies.

A total of 28 different conditions were considered. This included all possible combinations of four signal conditions and seven noise spectra. Each run tested a subject’s performance in one condition. A full factorial experimental design was used in which each of five subjects was tested once in each of the 28 conditions. The order of presentation for each subject was randomized in such a way as to eliminate any linear correlation between condition and run number. Each subject took part in five runs per week over a period of about six weeks.

The four signal conditions covered two axes and two frequencies. As in the ROC experiment, we considered both pitch and surge, in order to obtain data for both a rotation and a translation. The signal frequencies considered were 0.2 and 0.6 Hz. This gives two points in the range of interest for flight simulation. The latter frequency is the same as that used in the ROC experiment. The duration of the gated sinusoidal signal was 10 seconds, excluding the fade-in and fade-out ramps. The signal amplitude was varied from trial to trial by an algorithm designed to converge on the critical amplitude.

The seven noise spectra were designed to provide some variation in both power density and frequency content. The first three, conditions A, B and C, were broadband spectra with different power levels. These were chosen to test the prediction that detectability is a function of signal to noise ratio. If this prediction holds, we would expect to find that...
the critical amplitude is proportional to the RMS noise amplitude, or the square root of the noise power density. The remaining four spectra, conditions D, E, F and G, were narrow band spectra with different centre frequencies. These were chosen to test the concept of an internal pre-detection filter. If the concept is valid, we would expect the spectrum nearest to the signal frequency to have the largest effect on detectability. The characteristics of the noise conditions are summarized in Table 6.1. Design spectra are shown in Figures 6.1 and 6.2. Appendix C describes the noise generation algorithm and shows some typical measured spectra.

The design spectra for noise conditions A, B and C are shown in Figure 6.1. The baseline condition, A, is identical to that used in the ROC experiments and described in Chapter 5. It consists of a low-power, broadband noise on all six axes of motion. The motion is uncorrelated between axes. For conditions B and C, a single axis broadband spectrum was added to the baseline condition, on the same axis as the signal. These spectra have the same shape as spectrum A. The acceleration power density is constant from 0.25 Hz to 4.0 Hz, and rolls off at 40 dB per decade above 4.0 Hz and 60 dB per decade below 0.25 Hz. The power densities of spectra B and C, respectively, are 0.001 and 0.005 (m/s^2)^2/Hz in surge, or 0.001 and 0.005 (rad/s^2)^2/Hz in pitch. Spectra A, B and C resulted in measured RMS motion levels of about 3.0, 6.5 and 13.0 milli-g in surge, and about 1.7, 4.0 and 8.6°/s^2 in pitch.

Spectra D, E, F and G are shown in Figure 6.2. Each of these conditions consisted of noise condition A plus a single axis, narrow band (one octave) spectrum on the same axis as the signal. The power density of these spectra is the same as that for spectrum C. Taken together, spectra D, E, F and G cover the same four octave band as spectrum C (0.25 to 4.0 Hz).

It is worthwhile to explain why 0.25 Hz was chosen as the lower bound for the noise spectra, when the lower signal frequency was 0.2 Hz. This was a purely practical consideration. Low frequency accelerations can result in substantial motion of the simulator cab. This is particularly true for rotations about a horizontal axis. The translational displacement required to cancel the change in specific force caused by a rotation is proportional to 1/\(\omega^4\), for a fixed level of angular acceleration. To avoid running into the displacement limits of the motion base, the lower corner frequency of the noise was set at 0.25 Hz.
The initial plans for this experiment also included a "zero-noise" case: a case which tested the detectability of sinusoidal motion in the absence of random motion. The results of such a case would provide an estimate of the characteristics of the internal noise relevant to motion perception. However, this case proved impractical due to the mechanical limitations of the simulator motion base. When no motion is commanded, no motion is generated. When a low-amplitude pure sinusoidal motion is commanded, the resulting simulator motion is a somewhat noisy sinusoid. For signal amplitudes which might be detectable in the absence of random motion, the signal-to-noise ratio of the motion base may be as low as 1.5 to 2.0. In a detection experiment like that described here, it would be difficult to determine if the subject's response was based on the intended sinusoidal signal or on the attendant (unintended) noise. To ensure that the noise on each axis was well defined and independent of the presence or absence of a signal, the simulator was kept continually in motion, disturbed by multi-axis, low-power broadband noise. This low-level disturbance, which is noise condition A (described earlier), was a component of all of the noise conditions considered.

All other conditions were identical to those in the ROC experiment. During each run, the subject was seated and restrained with a five-point aviation harness. The simulator cab was dark except for a dimly illuminated dashboard, and the subject kept his eyes open. Finally, the subject wore an audio headset which provided instructions and prompts and masked auditory cues such as hydraulic hiss.

Five male research workers in the flight simulation laboratory served as paid subjects. All were in good health and had no history of vestibular disorder. They ranged in age from 24 to 33. Four of the subjects had previous experience as subjects in motion perception experiments, amounting to about 40 hours of testing time for each. The fifth subject had no experience. Each subject received three to five hours of training prior to the start of production runs. The subject data are summarized in Table 5.1.

Each experimental run was a series of forced choice trials. Figure 6.3 shows the events in a typical trial. Each trial was divided into two intervals. During one of the intervals, a gated sinusoidal signal was superimposed on the random background motion. The background motion, or noise, was continuous throughout the run. At the end of the trial, the subject indicated which interval he thought the signal occurred in, by pressing a button on a lap-held response box. An auditory feedback tone told him if his response was right or wrong. Following the tone, there was a delay of five seconds before the start
of the next trial. When the signal frequency was 0.6 Hz, a run comprised 50 trials. For
the 0.2 Hz case, a run was 40 trials. Each run took about 35 minutes to complete.

The signal amplitude was adjusted from trial to trial by an adaptive algorithm designed to
home in on the level which yielded 76% correct responses. As shown below, the fraction
correct in a two alternative forced choice task is equal to the area under the ROC curve.
Therefore, the algorithm converges on the critical amplitude, which has a detectability of
1.0.

Numerous different adaptive algorithms have been developed for perceptual testing. The
best known of these are the staircase and double staircase (Cornsweet 1962, Clark and
Stewart 1968a) and PEST (Taylor and Creelman, 1967). Existing methods were not
entirely suitable for this experiment, so a new algorithm was developed. A full descrip­
tion of the adaptive algorithm and a comparison with other methods is provided in
Appendix D. Two different formulations of the algorithm are described in Appendix D;
the least squares algorithm was used in this study.

In this method, the critical amplitude is estimated after each trial. This estimate is
referred to as the target level. It is calculated by fitting a psychometric curve to the data
collected thus far, using a least squares method. The target level is the amplitude which
 corresponds to the appropriate probability (0.76) on the fitted curve. The signal
amplitude on the next trial is set equal to or near to the target level. This maximizes the
amount of information which each trial contributes to the final estimate. The target level
calculated after the last trial of the run is the experimental result.

In surge, the signal amplitude was constrained to be between 0.008 m/s² and 0.200 m/s².
It could be varied in steps of 0.004 m/s². For pitch, the minimum and maximum were
0.08°/s and 2.0°/s (angular velocity), and the step size was 0.04°/s. In both cases, the
step size was near the resolution of the motion base, and was small relative to the
uncertainty of each estimate of the critical amplitude.

A typical run history showing the variation of signal amplitude over a run is shown in
Figure 6.4. The circles represent the actual signal amplitudes presented on each trial. An
open circle indicates that the subject's response was correct; a closed circle indicates that
it was not. The solid line shows the target level after each trial.
We claimed above that the fraction correct in a two alternative forced choice task is equal to the area under the ROC curve. This is easy to demonstrate. It is assumed that the observer bases his decision on an evidence variable, e. For a given criterion, he responds "yes" if and only if e exceeds that criterion. Let $e_s$ be the value of e obtained from the interval containing signal plus noise, and $e_n$ be the value from the interval containing noise alone. The consistent and optimal response policy is to select the interval which yields the larger value of e. The probability of a correct response is therefore equal to the probability that $e_s$ is greater than $e_n$.

The conditional probability of a correct response given that $e_s$ equals some value k is:

$$P_2(C \mid e_s=k) = P(e_n<k)$$

$$= P_k(N \mid n)$$

$$= 1 - P_k(S \mid n)$$

(6.1)

$P_k(S \mid n)$ is the false alarm rate when the yes/no response criterion is k. The total probability of a correct response is found by multiplying $P_2(C \mid e_s=k)$ by the density function, $f(k \mid s)$, and integrating over all possible values of k.

$$P_2(C) = \int_{-\infty}^{\infty} \left[ 1 - P_k(S \mid n) \right] f(k \mid s) \, dk$$

(6.2)

$f(k \mid s)$ may be obtained by differentiating equation 3.12:

$$f(k \mid s) = \frac{-dP_k(S \mid s)}{dk}$$

(6.3)

with $P_k(S \mid s) \to 1$ as $k \to -\infty$, and $P_k(S \mid s) \to 0$ as $k \to \infty$. Replacing this in equation 6.2 yields:

$$P_2(C) = \int_{0}^{1} \left[ 1 - P_k(S \mid n) \right] dP_k(S \mid s)$$

(6.4)

Therefore, $P_2(C)$ is equal to the area to the right of the ROC curve, within the unit square. This is simply the area under the ROC curve, $P_A$. Note that this relationship is independent of the shape of the density functions $f(e \mid n)$ and $f(e \mid s)$. 

6.3 Experimental Results

As described in full above, this experiment was run as a 7x2x2x5 full factorial design; each subject was tested once in each condition. The experiment comprised 140 production runs, each of which tested one subject in a single condition. The main result from each run was an estimate of the critical amplitude for that subject and condition. The results of all runs are summarized in Table 6.2.

The data were converted to a decibel scale and submitted to a three stage analysis. First, the data were checked for consistency. The results were found to be uncorrelated with both run number and trial number. Second, an analysis of variance was carried out to test the significance of patterns in the data. The main effects of noise and subject were significant. Also, the noise x frequency and noise x axis interaction effects were significant. An energy detection model with an internal filter provides a reasonably good explanation of the results. Finally, the slope of the psychometric curve was estimated from the data. The data indicate that the psychometric curve is quite steep, similar to that predicted by the energy detector. The three stages of analysis are described in the following sections.

6.3.1 Consistency Checks

Before proceeding with the main data analysis, the data were checked for consistency. The purpose of this was to see if there were any trends with run number or trial number which might affect the conclusions. Two separate analyses were done. The first checked for consistency from run to run, and the second checked for consistency from trial to trial. For both analyses, the data were converted to a decibel scale and fit with the four factor model described in Section 6.3.2.

To check for consistency from run to run, an analysis of covariance was carried out, using critical amplitude and run number as the dependent variables. No significant correlation was found between amplitude and run number. In order to show this graphically, the critical amplitude data were normalized by removing the main effects of noise condition, axis, frequency and subject. This can be done without affecting the correlation. Obviously, subject is invariant with run number. Also, the order of presentation of conditions for each subject was randomized to eliminate any correlation of noise condition, axis or frequency with run number. The normalized data are plotted versus
run number in Figure 6.5. The data do not exhibit any obvious trend with run number. The coefficient of correlation (amplitude versus run number) is 0.01. Therefore, it was concluded that there was no significant learning effect or similar trend over the duration of the experiment.

The consistency of the data from trial to trial was checked by analyzing the variation of target level with trial number. As described in Section 6.2, the adaptive algorithm calculates a target level after each trial. This target level is an estimate of the critical amplitude based on the outcome of all previous trials of that run. The data for each run were normalized by removing the influence of all factors and interactions which were found to have a significant effect (see Section 6.3.2). This included the main effects of noise condition, axis and subject, as well as the noise × frequency and noise × axis interaction effects. Normalizing the data in this manner allowed us to compare and combine data from different runs. Note that the mean normalized target level after 50 trials (the end of a run) will be 0 dB.

The mean and standard deviation of the normalized target levels are shown as a function of trial number in Figure 6.6. The low means for trials 1 to 10 are an artifact of the adaptive algorithm. Early in the run, the target level is based on the results of a very small number of trials, and is susceptible to bias. Since the initial setting of the signal amplitude was usually greater than the critical amplitude, the subject had a high probability of responding correctly on the first two or three trials. As a result, the algorithm underestimates the critical amplitude, until there are some incorrect responses to balance the estimation procedure.

The mean value is almost constant at 0 dB from trial 16 to 50. On average, the critical amplitude estimated at the end of the run is equal to that estimated after the first third of the run. This implies that, on average, performance is roughly constant over the duration of the run. There is no evidence to suggest that performance degrades due to fatigue or boredom toward the end of a run. The standard deviation of the estimates decreases monotonically with trial number, and is about proportional to $1/\sqrt{N}$. This is the fastest possible convergence, indicating that the algorithm is functioning efficiently. The experimental error is the standard deviation after 50 trials, which is about 2.0 dB.
6.3.2 The Effect of Noise Condition

Consistency checks showed that the results are uncorrelated with both run number and trial number, and can be assumed to represent typical steady state performance. This section describes the second stage of the data analysis. The main purpose of this was to determine the effect of noise power and frequency content on the critical amplitude.

The raw data shown in Table 6.2 were converted to a logarithmic decibel scale prior to analysis. The 0 dB levels were defined as 0.01 m/s² for surge, and 0.1°/s (angular velocity) for pitch. The converted data are summarized in Table 6.3. The uncertainty of each estimate of critical amplitude is proportional to the estimate itself. The change to a logarithmic scale, recommended by Bartlett (1947), improves the homogeneity of variance of the data. The result is a more even weighting of the results for different conditions. While the F-test used in the analysis of variance is not overly sensitive to heterogeneity of variance (Edwards, 1985), it is prudent to reduce heterogeneity when it can be done by a simple transformation.

A four factor mixed-effect structural model was used as the basis for statistical analysis. According to this model, each observation may be described as:

\[
X_{ijkl} = \mu + N_i + A_j + F_k + S_l + NA_{ij} + NF_{ik} + NS_{il} + AF_{jk} + AS_{jl} + FS_{kl} + NAF_{ijk} + NAS_{ijl} + NFS_{ikl} + AFS_{jkl} + NAFA_{ijkl} + e_{ijkl}
\]

where

- \(X_{ijkl}\) is the observation for noise condition \(i\), axis \(j\), frequency \(k\), and subject \(l\),
- \(\mu\) is the mean of all observations,
- \(N_i\) is the main effect of noise condition \(i\),
- \(A_j\) is the main effect of axis \(j\),
- \(F_k\) is the main effect of frequency \(k\),
- \(S_l\) is the main effect of subject \(l\),
- \(NA_{ij}\) is the two-factor noise \(\times\) axis interaction effect,
- \(NAF_{ijk}\) is the three-factor noise \(\times\) axis \(\times\) frequency interaction effect,
- \(NAFS_{ijkl}\) is the four-factor interaction effect, and
- \(e_{ijkl}\) is the experimental error.
In this model, noise condition, axis and frequency were taken to be fixed-effect factors. Subject was taken to be a random-effect factor. A discussion of the mixed-effect model may be found in any textbook on experimental design, such as Edwards (1985) or Winer (1971).

Standard analysis of variance procedures were used to test the significance of patterns seen in the data. In considering the effect of each factor (or interaction), the null hypothesis is that the factor has no effect. This implies that the associated parameter in equation 6.5 is zero for all values of the indices. If the parameter is significantly different from zero for some values of the index, we say that the effect of that factor is significant. Significance is tested using an F-test based on the ratio:

\[
F_r = \frac{MS_{\text{factor}}}{MS_{\text{ref}}} \tag{6.6}
\]

If the null hypothesis is true, then \( F_r \) follows the F distribution with the appropriate number of degrees of freedom. If \( \alpha = P(F > F_r) \) is small, then we must reject the null hypothesis and conclude that the factor does have a significant effect.

The numerator of \( F_r \) is always the mean square associated with the effect or interaction to be tested. The choice of denominator depends on the structural model. The proper denominators for the mixed-effect model are given by Edwards (1985) and are summarized in Table 6.4. In cases where the error mean square is needed, the within treatment mean square is normally used. In this experiment, however, each subject was tested only once in each condition, so there is no within treatment mean square. In this case, Edwards (1985) suggests that the four-factor interaction mean square (N×A×F×S) may be used. Assuming that the four-factor interaction effect is negligible, the expected value of the N×A×F×S mean square is equal to the error mean square.

The results of the analysis of variance are summarized in Table 6.5. The analysis shows that the main effects of noise condition and subject are both highly significant. These effects are shown in Figure 6.7, where the average result for each noise condition is plotted for each subject. The noise × frequency and noise × axis interaction effects are less pronounced, but still highly significant (\( \alpha < 0.002 \)). Finally, the main effect of axis is significant. However, the difference between axes is meaningless because the data for each axis were normalized arbitrarily.
Figure 6.7 shows that while the variation with condition is about the same for all subjects, some subjects did perform better than others. Averaged over all conditions, the most sensitive subject outperformed the least sensitive by 4.0 dB. It is interesting to note that the most sensitive subject was the experimenter, and the least sensitive was the subject with no previous experience in perception tests. Much of the difference between subjects is probably due to such factors as motivation and training.

Figure 6.7 also shows how the response varied with noise condition. The results are about what we would expect, based on the characteristics of the noise conditions. First consider noise conditions A, B and C, the three broadband spectra. On average, the critical amplitude is about 12 dB higher for condition C than it is for condition A. The difference in noise power density between conditions A and C is about 14 dB, so the signal energy at critical amplitude is roughly proportional to noise power density. The average critical amplitude for conditions A and C, for each axis and frequency, are given in Table 6.2.

Next consider noise conditions D, E, F and G, the narrow band spectra. The results for these conditions all lie between the limits defined by conditions A and C. The greatest shift in critical amplitude (as compared with condition A), is caused by conditions D and E. Thus, the lower frequency noise spectra provide the most effective signal masking.

Considering the two frequencies separately makes this observation more specific. The noise × frequency interaction is highly significant. This effect is shown in Figure 6.8. The data for surge are shown in Figure 6.8.a, and the data for pitch are shown in Figure 6.8.b. There appear to be two main differences associated with frequency. First, the difference between conditions A and C is slightly less for the 0.2 Hz case than for the 0.6 Hz case. Second, the highest value found among the narrow band conditions (D to G) is with condition D for the 0.2 Hz case, and E for the 0.6 Hz case. Condition D comprises noise in the 0.25 to 0.5 Hz band; condition E covers the 0.5 to 1.0 Hz band. Therefore, the data show that the noise spectrum nearest to the signal frequency causes the greatest shift in critical amplitude. This observation holds for both axes.

The other significant effect found was the noise × axis interaction. This effect is shown in Figure 6.9. Two main differences can be seen. First, the difference between the results for conditions A and C is greater for pitch than it is for surge. Second, the pattern of results for the narrow band conditions are different for the two axes. For pitch, as
compared to surge, the effect of the noise is greater at low frequencies (condition D), and drops off more quickly at high frequencies (conditions F and G). This pattern is consistent for both signal frequencies. The implication of this will be discussed later.

Let us summarize our findings. The data show two results very clearly. First, the signal energy at critical amplitude is roughly proportional to noise power density, for the broadband spectra. This implies that signal detectability is a function of signal to noise ratio. Second, all of the narrow band spectra cause some shift in critical amplitude, with the largest effect observed for the noise spectrum which is nearest in frequency to the signal. This finding supports the energy detection model with an internal pre-detection filter. All noise tends to mask the signal, but the observer can filter out noise which is far from the signal frequency.

A detailed comparison of the data to the model predictions is deferred until Chapter 8. However, it is useful at this point to do a rough comparison, to see if the model provides at least approximate agreement with experiment. A loose interpretation of equation 3.66, assuming that the integration time T is constant for all cases, is that $d_A'$ is a function of the ratio of signal power to the post-filter noise power. Since $d_A' = 1$ for all of the critical amplitudes found in this experiment, the signal to noise power ratio should be approximately the same for all cases.

For this comparison, the detector was assumed to have the form shown in Figure 4.1. The internal filter was assumed to be a second order band pass filter with a transfer function of the form:

$$H(s) = \frac{2\zeta \omega_s s}{s^2 + 2\zeta \omega_s s + \omega_s^2}$$

(6.7)

where $\omega_s$ is the signal frequency, and $\zeta$ is a parameter which controls the filter bandwidth. For each noise condition, the post-filter noise power was calculated as:

$$P_n = \int_{0}^{\infty} |H(j\omega)|^2 \Phi_{nm}(\omega) \, d\omega$$

(6.8)

where $\Phi_{nm}(\omega)$ is the input noise power spectrum on the same axis as the signal. The
signal to noise power ratios based on the experimental data may then be expressed:

\[ R_{cr} = \frac{1}{2} \frac{A_{cr}^2}{P_n} = \frac{\text{signal power}}{\text{noise power}} \quad (6.9) \]

where \( A_{cr} \) is the critical amplitude, and \( P_n \) is the post-filter noise power for the appropriate condition. If the detection model holds, then \( R_{cr} \) should be approximately constant for all conditions.

The bandwidth of the internal filter was estimated by adjusting the damping parameter, \( \zeta \), to minimize the variability of \( R_{cr} \) between conditions. For surge, both signal and noise power were expressed in terms of acceleration. The noise spectra, \( \Phi_{nn}(\omega) \), were based on averages of measured spectra. The contribution of the multi-axis baseline condition, \( A \), to all of the noise conditions was increased by 50% to account for the surge masking caused by the pitch component of condition \( A \). This is consistent with the findings of Chapter 7. For pitch, both signal and noise power were expressed in terms of angular velocity. No adjustment was made for inter-axis masking.

For a signal frequency of 0.2 Hz, the best fit to the data was obtained by setting \( \zeta = 1.2 \). This yields a filter with a bandwidth of 0.48 Hz. The pass band, defined by the \(-3\, \text{dB} \) points, is 0.07 to 0.55 Hz. For a signal frequency of 0.6 Hz, the best fit was obtained by setting \( \zeta = 0.6 \). The resulting bandwidth is 0.72 Hz (0.34 to 1.06 Hz pass band). Bode plots of the two filters are shown in Figure 6.10.

The variation of \( P_n \) with noise condition, for each signal axis and frequency, is shown in Figure 6.11. Also shown in Figure 6.11 is the signal energy based on the average critical amplitude found for each condition. The comparison is not perfect, but the curves do have a similar shape. The signal to noise power ratio, \( R_{cr} \) (expressed in dB), is simply the vertical distance between the curves. The average value of \( R_{cr} \) over all conditions is about 0.75 dB. This value is consistent for both axes and frequencies. In other words, a 10 second sinusoidal motion signal is "just detectable" (\( d_A' = 1 \)) when the ratio of signal power to "post-filter" noise power is about 1.2. This is reminiscent of the rule proposed by Fletcher (1940) for auditory detection, which states that a tone is just audible when its energy is equal to the noise energy in a "critical band."

In the analysis above, the signal and noise power for surge were expressed in terms of acceleration, while those for pitch were expressed in terms of angular velocity.
Recalculating the pitch power in terms of angular acceleration yields dramatically different results, as illustrated in Figure 6.12. In this case, the average value of $R_{cr}$ is $-9.0$ dB for the 0.2 Hz case, and $-1.3$ dB for the 0.6 Hz case. Therefore, expressing the pitch $R_{cr}$ in terms of angular acceleration introduces a disparity in the results for the two frequencies. Furthermore, the general shape of the curves shown in Figure 6.12 does not agree as well with the experimental data as those shown in Figure 6.11. Expressing the surge $R_{cr}$ in terms of linear velocity has a similar effect. The data collapse best when $R_{cr}$ is expressed in terms of acceleration for surge, and angular velocity for pitch.

The implication is that the surge detector works in terms of acceleration, while the pitch detector works in terms of angular velocity. This is consistent with the vestibular models proposed in Chapter 4. It also helps to explain the noise $\times$ axis interaction effect observed in the data. Both the surge and pitch noise power spectra were flat in terms of acceleration (or angular acceleration). But the pitch detector appears to work as a velocity detector. If the pitch noise spectra are expressed in terms of angular velocity, they are no longer flat with frequency. Spectrum C, for example, has a peak near 0.25 Hz, and rolls off at 20 dB per decade with increasing frequency. Similarly, the noise power density of the narrow band spectra, in angular velocity, is highest for spectrum D, and falls off at higher frequencies. Figure 6.9 shows that for pitch, as compared to surge, the masking effectiveness of noise is greater at low noise frequencies, but falls off more quickly with increasing frequency. The formulation of the pitch detector as a velocity detector explains this effect.

The entire data set was normalized using equation 6.9 and the values of $P_n$ given in Table 6.6. For this analysis, surge signal and noise power was expressed in terms of acceleration, while pitch power was expressed in terms of angular velocity. The resulting values of $R_{cr}$ are summarized in Table 6.6. An analysis of variance was carried out on the normalized data. The results are shown in Table 6.7. It is clear from Table 6.7 that this approximate model does not fit the data perfectly. The main effect of noise condition, and the noise $\times$ frequency and noise $\times$ axis interaction effects are still significant. However, the model does account for most of the variance in the data. The mean square associated with noise condition has been reduced by 92%. It is also encouraging to note that there is no significant difference between axes. This comparison is meaningful now, since the data are normalized in terms of the model, rather than arbitrarily. Based on these observations, it is concluded that the data are in reasonably good agreement with the qualitative predictions of the energy detection model.
6.3.3 The Psychometric Curve

In the last section, we explored the way in which the critical amplitude varies with signal and noise conditions. The result for each subject and condition is a single point near the middle of the psychometric curve. This gives us the location of the curve along the abscissa, but tells us nothing about its shape. The purpose of the analysis presented below was to estimate the slope of the psychometric curve.

The slope estimates were based on the record of responses for each run, rather than the final target level. Each run consisted of either 40 or 50 trials at different signal amplitudes. This number of trials was sufficient to obtain a low-variance estimate of the critical amplitude, but not sufficient to provide an estimate of the slope. However, if we increase the sample size by combining the data from a number of runs, it is possible to obtain a good estimate of the slope. The implicit assumption, of course, is that the slope of the psychometric curve is equal for all of the data sets combined. In order to combine data from different runs, the amplitude data for each run were normalized by the final target level for that run. Therefore, the normalized amplitude for each trial is the signal amplitude relative to the critical amplitude for the appropriate subject and condition. The other datum for each trial was the response (correct or not).

The combined data were fit with a generic form of psychometric curve, the logistic function, which was discussed in Chapter 5. The parameter estimation was done by varying both the midpoint and slope of the function to achieve the best fit to the data. This procedure is similar to that used by the adaptive algorithm, except that it uses maximum likelihood estimation rather than least squares. The standard deviation of the slope estimates was obtained from the second derivative of the likelihood function.

Figure 6.13 shows the fitted curve based on the entire data set (6300 trials). Each point represents the results of all trials within a range of relative amplitude. The abscissa for each point is the relative amplitude in dB, and the ordinate is the fraction of trials in which the response was correct. As discussed in Section 6.2, this is equal to the area under the ROC curve. The error bars for each point are the binomial standard deviation,

\[
\sigma_p = \sqrt{\frac{P(1-P)}{N}}
\]  

(6.10)

where \( N \) is the number of trials included in the point. Note that the majority of trials had a relative amplitude of between \(-4 \, \text{dB} \) and \(4 \, \text{dB} \). Points far from 0 dB are each based on
a much smaller number of trials. Consequently, the standard deviation is greater for points near the tails of the curve.

As seen from Figure 6.13, the data suggest a very steep psychometric curve. The slope, B, of the fitted logistic function is 0.45, with a standard deviation of 0.025. Of the three detection models considered, the energy detector is the closest to this, with B typically between 0.35 and 0.39. The envelope detector and the correlation detector both predict much shallower slopes than that shown by the data.

Figure 6.14 shows the data for each subject individually. In each case, the curve shown has a slope of 0.45. A maximum likelihood estimate of B was calculated separately for each subject. Each of these fits was based on 1260 trials. The resulting values are given in Table 6.8. The values found range from 0.38 to 0.60 with a standard deviation of 0.088. The ratio of the variance between subjects to the theoretical variance of each estimate is 2.4. An F-test based on this ratio shows that the difference between subjects is significant ($\alpha < 0.05$). In particular, the data for subject BW follow a steeper curve than the data for the other subjects.

Finally, the data were considered separately by noise condition. Each of these fits was based on 900 trials. The results of this analysis are summarized in Table 6.9. An F-test showed that the difference in slope between noise conditions is not significant ($\alpha > 0.10$).
Table 6.1  Power density and bandwidth of the masking noise spectra used in experiment 2.

<table>
<thead>
<tr>
<th>Noise Condition</th>
<th>Pass band (Hz)</th>
<th>Power density in pass band</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.25 - 4.0</td>
<td>0.0002</td>
</tr>
<tr>
<td>B</td>
<td>0.25 - 4.0</td>
<td>0.0010</td>
</tr>
<tr>
<td>C</td>
<td>0.25 - 4.0</td>
<td>0.0050</td>
</tr>
<tr>
<td>D</td>
<td>0.25 - 0.5</td>
<td>0.0050</td>
</tr>
<tr>
<td>E</td>
<td>0.50 - 1.0</td>
<td>0.0050</td>
</tr>
<tr>
<td>F</td>
<td>1.00 - 2.0</td>
<td>0.0050</td>
</tr>
<tr>
<td>G</td>
<td>2.00 - 4.0</td>
<td>0.0050</td>
</tr>
</tbody>
</table>

* units: (m/s²)²/Hz for surge
(rad/s²)²/Hz for pitch
Table 6.2  Critical amplitudes for each subject and condition, from experiment 2.
Surge data are expressed in milli-g (0.01 m/s²).
Pitch data are expressed in degrees per second.

<table>
<thead>
<tr>
<th>Axis</th>
<th>Freq</th>
<th>Noise cond.</th>
<th>Subject</th>
<th>Geom. mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>DG</td>
<td>PG</td>
<td>GG</td>
</tr>
<tr>
<td>Surge</td>
<td>0.2 Hz</td>
<td>A</td>
<td>3.19</td>
<td>2.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>4.25</td>
<td>5.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>8.14</td>
<td>8.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D</td>
<td>7.79</td>
<td>4.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E</td>
<td>4.25</td>
<td>4.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>3.54</td>
<td>1.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>2.83</td>
<td>4.25</td>
</tr>
<tr>
<td>Surge</td>
<td>0.6 Hz</td>
<td>A</td>
<td>2.78</td>
<td>2.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>5.10</td>
<td>4.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>10.67</td>
<td>10.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D</td>
<td>6.03</td>
<td>5.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E</td>
<td>6.03</td>
<td>7.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>5.10</td>
<td>3.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>3.71</td>
<td>2.78</td>
</tr>
<tr>
<td>Pitch</td>
<td>0.2 Hz</td>
<td>A</td>
<td>.456</td>
<td>.326</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>.554</td>
<td>.554</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>1.174</td>
<td>1.239</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D</td>
<td>1.206</td>
<td>.717</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E</td>
<td>.652</td>
<td>.326</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>.424</td>
<td>.359</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>.522</td>
<td>.424</td>
</tr>
<tr>
<td>Pitch</td>
<td>0.6 Hz</td>
<td>A</td>
<td>.339</td>
<td>.212</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>.594</td>
<td>.890</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>1.696</td>
<td>1.442</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D</td>
<td>.933</td>
<td>.848</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E</td>
<td>.848</td>
<td>1.102</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>.466</td>
<td>.424</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>.339</td>
<td>.382</td>
</tr>
</tbody>
</table>
Table 6.3  Critical amplitudes for each subject and condition, expressed as dB.
For surge, 0 dB = .01 m/s². For pitch, 0 db = 0.1 °/s.

<table>
<thead>
<tr>
<th>Axis</th>
<th>Freq</th>
<th>Noise cond.</th>
<th>Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>DG</td>
<td>PG</td>
</tr>
<tr>
<td>Surge</td>
<td>0.2 Hz</td>
<td>A</td>
<td>10.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>12.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>18.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D</td>
<td>17.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E</td>
<td>12.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>10.98</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>9.04</td>
</tr>
<tr>
<td>Surge</td>
<td>0.6 Hz</td>
<td>A</td>
<td>8.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>14.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>20.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D</td>
<td>15.61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E</td>
<td>15.61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>14.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>11.39</td>
</tr>
<tr>
<td>Pitch</td>
<td>0.2 Hz</td>
<td>A</td>
<td>13.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>14.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>21.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D</td>
<td>21.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E</td>
<td>16.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>12.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>14.35</td>
</tr>
<tr>
<td>Pitch</td>
<td>0.6 Hz</td>
<td>A</td>
<td>10.61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>15.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>24.59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D</td>
<td>19.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E</td>
<td>18.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>13.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>10.61</td>
</tr>
</tbody>
</table>
Table 6.4  Appropriate denominators for tests of significance for the mixed-effect model.

<table>
<thead>
<tr>
<th>Numerator Mean Square</th>
<th>Denominator Mean Square*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise</td>
<td>Noise ( \times ) Subj</td>
</tr>
<tr>
<td>Axis</td>
<td>Axis ( \times ) Subj</td>
</tr>
<tr>
<td>Noise ( \times ) Axis</td>
<td>N ( \times ) A ( \times ) S</td>
</tr>
<tr>
<td>Frequency</td>
<td>Freq ( \times ) Subj</td>
</tr>
<tr>
<td>Noise ( \times ) Freq</td>
<td>N ( \times ) F ( \times ) S</td>
</tr>
<tr>
<td>Axis ( \times ) Freq</td>
<td>A ( \times ) F ( \times ) S</td>
</tr>
<tr>
<td>N ( \times ) A ( \times ) F</td>
<td>N ( \times ) A ( \times ) F ( \times ) S</td>
</tr>
<tr>
<td>Subject</td>
<td>Error</td>
</tr>
<tr>
<td>Noise ( \times ) Subj</td>
<td>Error</td>
</tr>
<tr>
<td>Axis ( \times ) Subj</td>
<td>Error</td>
</tr>
<tr>
<td>N ( \times ) A ( \times ) S</td>
<td>Error</td>
</tr>
<tr>
<td>Freq ( \times ) Subj</td>
<td>Error</td>
</tr>
<tr>
<td>N ( \times ) F ( \times ) S</td>
<td>Error</td>
</tr>
<tr>
<td>A ( \times ) F ( \times ) S</td>
<td>Error</td>
</tr>
<tr>
<td>N ( \times ) F ( \times ) A ( \times ) S</td>
<td>Error</td>
</tr>
</tbody>
</table>

* In cases where the error mean square is needed, the N \( \times \) A \( \times \) F \( \times \) S interaction mean square was used.
Table 6.5  Results of the analysis of variance for the experimental data.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Degrees of Freedom</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Ratio</th>
<th>P(x&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise</td>
<td>6</td>
<td>1806.7</td>
<td>301.12</td>
<td>61.08</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Axis</td>
<td>1</td>
<td>108.6</td>
<td>108.59</td>
<td>11.66</td>
<td>0.0269</td>
</tr>
<tr>
<td>Noise × Axis</td>
<td>6</td>
<td>121.4</td>
<td>20.24</td>
<td>5.01</td>
<td>0.0019</td>
</tr>
<tr>
<td>Frequency</td>
<td>1</td>
<td>82.6</td>
<td>82.55</td>
<td>6.64</td>
<td>0.0615</td>
</tr>
<tr>
<td>Noise × Freq</td>
<td>6</td>
<td>169.2</td>
<td>28.19</td>
<td>6.05</td>
<td>0.0006</td>
</tr>
<tr>
<td>Axis × Freq</td>
<td>1</td>
<td>9.5</td>
<td>9.47</td>
<td>2.42</td>
<td>0.1328</td>
</tr>
<tr>
<td>N × A × F</td>
<td>6</td>
<td>16.4</td>
<td>2.73</td>
<td>0.70</td>
<td>0.6536</td>
</tr>
<tr>
<td>Subject</td>
<td>4</td>
<td>250.2</td>
<td>62.55</td>
<td>15.99</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Noise × Subj</td>
<td>24</td>
<td>118.3</td>
<td>4.93</td>
<td>1.26</td>
<td>0.2876</td>
</tr>
<tr>
<td>Axis × Subj</td>
<td>4</td>
<td>37.2</td>
<td>9.31</td>
<td>2.38</td>
<td>0.0799</td>
</tr>
<tr>
<td>N × A × S</td>
<td>24</td>
<td>97.0</td>
<td>4.04</td>
<td>1.03</td>
<td>0.4679</td>
</tr>
<tr>
<td>Freq × Subj</td>
<td>4</td>
<td>49.7</td>
<td>12.43</td>
<td>3.18</td>
<td>0.0314</td>
</tr>
<tr>
<td>N × F × S</td>
<td>24</td>
<td>111.9</td>
<td>4.66</td>
<td>1.19</td>
<td>0.3350</td>
</tr>
<tr>
<td>A × F × S</td>
<td>4</td>
<td>11.0</td>
<td>2.75</td>
<td>0.70</td>
<td>0.5980</td>
</tr>
<tr>
<td>N × A × F × S</td>
<td>24</td>
<td>93.9</td>
<td>3.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>139</td>
<td>3083.6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 6.6  Signal-to-noise power ratios at critical amplitude, for each subject and condition. The values shown are the data of Table 6.3, normalized by $P_n$.

<table>
<thead>
<tr>
<th>Axis &amp; Freq.</th>
<th>Noise Cond.</th>
<th>$P_n$</th>
<th>Signal-to-noise power, $R_{cr}$ (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>DG</td>
</tr>
<tr>
<td>Surge, 0.2 Hz.</td>
<td>A</td>
<td>5.30</td>
<td>4.76</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>10.94</td>
<td>1.62</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>16.86</td>
<td>1.35</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>13.81</td>
<td>4.02</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>13.42</td>
<td>-0.86</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>11.84</td>
<td>-0.86</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>9.89</td>
<td>-0.85</td>
</tr>
<tr>
<td>Surge, 0.6 Hz.</td>
<td>A</td>
<td>8.31</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>13.95</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>19.87</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>13.83</td>
<td>1.78</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>17.06</td>
<td>-1.45</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>15.79</td>
<td>-1.63</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>13.40</td>
<td>-2.01</td>
</tr>
<tr>
<td>Pitch, 0.2 Hz.</td>
<td>A</td>
<td>7.47</td>
<td>5.72</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>15.25</td>
<td>-0.38</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>21.62</td>
<td>-0.23</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>21.94</td>
<td>-0.31</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>16.36</td>
<td>-0.08</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>10.82</td>
<td>1.72</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>8.13</td>
<td>6.22</td>
</tr>
<tr>
<td>Pitch, 0.6 Hz.</td>
<td>A</td>
<td>7.54</td>
<td>3.07</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>15.32</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>21.69</td>
<td>2.90</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>19.97</td>
<td>-0.57</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>19.26</td>
<td>-0.69</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>13.56</td>
<td>-0.18</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>8.95</td>
<td>1.66</td>
</tr>
</tbody>
</table>
Table 6.7  Results of the analysis of variance for the normalized data.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Degrees of Freedom</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Ratio</th>
<th>P(x&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise</td>
<td>6</td>
<td>135.2</td>
<td>22.53</td>
<td>4.57</td>
<td>0.0032</td>
</tr>
<tr>
<td>Axis</td>
<td>1</td>
<td>0.2</td>
<td>0.20</td>
<td>0.02</td>
<td>0.8914</td>
</tr>
<tr>
<td>Noise × Axis</td>
<td>6</td>
<td>105.7</td>
<td>17.61</td>
<td>4.36</td>
<td>0.0041</td>
</tr>
<tr>
<td>Frequency</td>
<td>1</td>
<td>2.0</td>
<td>2.00</td>
<td>0.16</td>
<td>0.7087</td>
</tr>
<tr>
<td>Noise × Freq</td>
<td>6</td>
<td>135.0</td>
<td>22.51</td>
<td>4.83</td>
<td>0.0023</td>
</tr>
<tr>
<td>Axis × Freq</td>
<td>1</td>
<td>11.9</td>
<td>11.92</td>
<td>3.05</td>
<td>0.0937</td>
</tr>
<tr>
<td>N × A × F</td>
<td>6</td>
<td>15.7</td>
<td>2.62</td>
<td>0.67</td>
<td>0.6744</td>
</tr>
<tr>
<td>Subject</td>
<td>4</td>
<td>250.2</td>
<td>62.55</td>
<td>15.99</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Noise × Subj</td>
<td>24</td>
<td>118.3</td>
<td>4.93</td>
<td>1.26</td>
<td>0.2876</td>
</tr>
<tr>
<td>Axis × Subj</td>
<td>4</td>
<td>37.2</td>
<td>9.31</td>
<td>2.38</td>
<td>0.0799</td>
</tr>
<tr>
<td>N × A × S</td>
<td>24</td>
<td>97.0</td>
<td>4.04</td>
<td>1.03</td>
<td>0.4679</td>
</tr>
<tr>
<td>Freq × Subj</td>
<td>4</td>
<td>49.7</td>
<td>12.43</td>
<td>3.18</td>
<td>0.0314</td>
</tr>
<tr>
<td>N × F × S</td>
<td>24</td>
<td>111.9</td>
<td>4.66</td>
<td>1.19</td>
<td>0.3350</td>
</tr>
<tr>
<td>A × F × S</td>
<td>4</td>
<td>11.0</td>
<td>2.75</td>
<td>0.70</td>
<td>0.5980</td>
</tr>
<tr>
<td>N × A × F × S</td>
<td>24</td>
<td>93.9</td>
<td>3.91</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total        139   1175.0
Table 6.8  Maximum likelihood estimates of the psychometric curve slope, B, for each subject.

<table>
<thead>
<tr>
<th>Subject</th>
<th>B</th>
<th>$\sigma_B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DG</td>
<td>.452</td>
<td>.052</td>
</tr>
<tr>
<td>PG</td>
<td>.385</td>
<td>.054</td>
</tr>
<tr>
<td>GG</td>
<td>.391</td>
<td>.048</td>
</tr>
<tr>
<td>MN</td>
<td>.500</td>
<td>.062</td>
</tr>
<tr>
<td>BW</td>
<td>.597</td>
<td>.075</td>
</tr>
<tr>
<td>All data</td>
<td>.449</td>
<td>.025</td>
</tr>
</tbody>
</table>

Table 6.9  Maximum likelihood estimates of the psychometric curve slope, B, for each noise condition.

<table>
<thead>
<tr>
<th>Noise Condition</th>
<th>B</th>
<th>$\sigma_B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>.411</td>
<td>.061</td>
</tr>
<tr>
<td>B</td>
<td>.623</td>
<td>.085</td>
</tr>
<tr>
<td>C</td>
<td>.488</td>
<td>.082</td>
</tr>
<tr>
<td>D</td>
<td>.415</td>
<td>.069</td>
</tr>
<tr>
<td>E</td>
<td>.495</td>
<td>.074</td>
</tr>
<tr>
<td>F</td>
<td>.435</td>
<td>.064</td>
</tr>
<tr>
<td>G</td>
<td>.375</td>
<td>.055</td>
</tr>
</tbody>
</table>
Figure 6.1  Power spectra for noise conditions A, B and C.
For surge, $N_{\text{ref}} = 0.005 (\text{m/s}^2)^2/\text{Hz}$.
For pitch, $N_{\text{ref}} = 0.005 (\text{rad/s}^2)^2/\text{Hz}$.

Figure 6.2  Power spectra for noise conditions D, E, F and G.
For surge, $N_{\text{ref}} = 0.005 (\text{m/s}^2)^2/\text{Hz}$.
For pitch, $N_{\text{ref}} = 0.005 (\text{rad/s}^2)^2/\text{Hz}$.
Figure 6.3  Main events in a forced choice trial.
Figure 6.4  Typical history of an experimental run showing how signal amplitude is controlled by the adaptive algorithm.
Figure 6.5  Normalized target levels as a function of run number.
Figure 6.6   Mean and standard deviation of the normalized target levels, plotted as a function of trial number.
Figure 6.7  Average result for each subject, plotted against noise condition.
Figure 6.8 Noise × Frequency interaction. The values plotted are the mean for all subjects, for each axis and frequency (NAF_{ijk}). Error bars indicate ± 1 standard deviation of each mean.
Figure 6.9 Noise × Axis interaction. The values plotted are the mean for all subjects, for each axis and frequency (NAF<sub>ijk</sub>). Error bars indicate ±1 standard deviation of each mean.
Figure 6.10  Response characteristics of the assumed internal filters.
Figure 6.11  Comparison of signal power (at $d_A'=1$) to the filtered noise power. Signal and noise power are expressed in terms of acceleration for surge, and angular velocity for pitch. The error bars indicate ± 1 standard deviation of each mean.
Figure 6.12  Comparison of signal power (at \(d_A' = 1\)) to the filtered noise power, for pitch. Signal and noise power are expressed in units of angular acceleration. The error bars indicate \(\pm 1\) standard deviation of each mean.
Figure 6.13  Psychometric curve based on all data from experiment 2 (6300 trials). For the logistic function shown, $B = 0.45$. 
Figure 6.14  Response data for each subject. For the logistic function shown, \( B = 0.45 \).
Chapter 7

Inter-Axis Effects

7.1 Introduction

This chapter describes the third and final experiment of this study. The experiments described in Chapters 5 and 6 were designed to test the applicability of the signal detection model to the human motion perception process. This involved deriving ROC curves and psychometric curves, and estimating the critical bandwidth, for comparison with model predictions. The purpose of the third experiment was to answer a more practical question: does random motion on one axis mask signals on other axes?

Altogether, 100 different cases were considered. This covered four signal axes (surge, pitch, sway and roll), five signal frequencies (0.2 to 1.0 Hz), and five noise conditions. One noise condition was a low-power, multi-axis baseline condition. In the other four, a high-power broadband noise spectrum was added on one of the above four axes. Six subjects were tested in all 100 conditions.

Each experimental run was a series of 80 to 90 trials which tested four different conditions. The noise condition and signal frequency were fixed throughout each run. The signal axis, on the other hand, was selected at random for each trial. The subject’s task on each trial was to identify the axis of motion. The signal amplitude for each axis was adjusted from trial to trial by an adaptive algorithm designed to home in on the critical amplitude.

The experimental results can be divided into two groups. The first includes the cases with signal and noise on the same axis; the second includes cases with signal and noise on different axes. The same-axis data confirm the findings of Chapter 6. The inter-axis data show that random motion on one axis does mask signals on other axes. The effectiveness of inter-axis masking depends on the degree of similarity between the signal and the noise. The strongest effects were observed for a surge signal, masked by pitch noise, and for a sway signal, masked by roll noise.
The experimental conditions and procedure are discussed in Section 7.2. The results are presented in Section 7.3.

7.2 Experimental Method and Conditions

The specific objective of this experiment was to determine the extent to which random motion on one axis affects the detectability of a signal on another. We proposed to answer this by locating the psychometric curve along the abscissa for a number of different noise conditions. As with the single axis masking experiment, we limited our investigation to finding a single point on the psychometric curve: the critical amplitude.

A total of 100 different conditions were considered. This included all possible combinations of four signal axes, five signal frequencies and five noise conditions. A full factorial experimental design was used in which each of 6 subjects was tested once in each condition. The experiment comprised 25 runs for each subject. The noise condition and signal frequency were fixed throughout each run, and all four signal axes were tested in each run. The order of presentation (of noise condition and signal frequency) for each subject was randomized using a Latin square procedure (see Edwards, 1985) to eliminate any linear correlation between condition and run number. Each subject took part in five runs per week over a period of five weeks.

The 20 signal conditions covered four axes and five frequencies. The four axes considered were two translations, surge and sway, and two rotations, pitch and roll. These were selected in discussion with engineers from CAE Electronics Ltd., because they are associated with two maneuvers which are difficult to simulate: sustained longitudinal acceleration and coordinated turn entry. The five signal frequencies tested were 0.2, 0.4, 0.6, 0.8 and 1.0 Hz. The signal duration was 10 seconds for all frequencies. The signal amplitude varied from trial to trial, under control of the adaptive algorithm.

The five noise conditions were chosen to include all four signal axes, plus a low-power case for comparison. The first noise condition, condition A, was the same low-power, multi-axis broadband noise used in the two previous experiments. This served as a baseline case. The other conditions (X, Q, Y and P, respectively) consisted of condition A plus an uncorrelated high-power broadband spectrum on one of four axes: surge, pitch, sway or roll. The design power spectrum for conditions X, Q, Y and P is flat, in
terms of acceleration or angular acceleration, from 0.25 Hz to 4.0 Hz. The noise power density is 0.004 (m/s²)²/Hz for surge and sway, and 0.004 (rad/s²)²/Hz for pitch and roll. Note that these conditions are identical to condition C described in Chapter 6, except that the power density is slightly lower in the present case.

All conditions relating to the subjects' posture and the visual and auditory environment were identical to those in the previous experiments. These are discussed in Chapter 4.

Six men between the ages of 23 and 31 served as subjects. Five were research workers in the flight simulation laboratory; the sixth (DS) was a student working with another group. All were in good health and had no history of vestibular disorder. Table 5.1 provides a summary of the relevant data for each subject. Chronologically, this experiment was run first. Therefore, none of the subjects had any previous experience in motion perception experiments or other perceptual tests. Each subject received four to six hours of training prior to the start of data collection. This consisted of 6 to 8 practice runs. Each practice run took the same form as an actual production run. The practice runs were chosen to provide each subject with some exposure to all noise conditions and all signal frequencies.

The initial plan was to run this experiment using the forced choice procedure described in Chapter 6. However, with 100 conditions and six subjects, this would entail some 400 hours of testing and would take five or six months to complete. In the interest of efficiency, an alternate procedure was devised. In this procedure, the noise condition and signal frequency were constant throughout a run. The signal axis, on the other hand, was selected randomly for each trial. The subject's task on each trial was to identify the axis of motion: a four alternative forced choice task. The procedure is relatively fast. Using it reduced the size of the experiment by a factor of 4, to 150 runs. However, it does have some weaknesses, which are discussed below.

Each run was a series of single-interval trials similar to that shown in Figure 5.1. As with the other experiments, the noise motion was continuous throughout the run. A warning tone denoted the start of the trial. After a short delay, a gated sinusoidal motion signal was superimposed on the random motion for a duration of about 10 seconds. A signal was present on every trial. The axis of signal motion was selected randomly and independently for each trial. At the end of the trial, the subject identified the axis of motion by pushing a button on a lap-held response box. A feedback tone indicated if his
response was right or wrong, but did not identify the correct answer. The run continued until 20 trials had been completed for each axis, or until a total of 90 trials were complete, whichever came first. A complete run took about 35 minutes.

As in the single axis masking experiment, the signal amplitude was adjusted from trial to trial by the adaptive algorithm described in Appendix D. The maximum and minimum signal levels and the step size were the same as those given in Chapter 6. The target level for each axis (and the subsequent presentation amplitude) was based only on trials which contained a signal on that axis. Therefore, the amplitude on each axis could vary independently of the amplitudes on the other axes. The target level for each axis was the critical amplitude for that particular signal and noise condition. For the four alternative forced choice task, this is the amplitude which results in 56% correct responses.

In Chapter 6 it was shown that the fraction correct in a two-alternative forced choice task equals the area under the ROC curve. This argument may be extended to relate the fraction correct in a four alternative forced choice task to the ROC curve. The derivation here follows the one given by Green and Swets (1974). Assume that the distributions \( f(e \mid n) \) and \( f(e \mid s) \) are the same for all four alternatives, and that the alternatives are independent. Then the optimal strategy is to select the alternative which yields the largest value of the evidence variable, \( e \). The probability of a correct response is equal to the probability that \( e_s \), the value of \( e \) resulting from the correct alternative, is greater than three independent samples of \( e_n \). The conditional probability of a correct response, given that \( e_s = k \), is:

\[
P_4(C \mid e_s=k) = [P(e_n < k)]^3
\]

\[
= [1 - P_k(S \mid n)]^3
\]

Thus the total probability of a correct response is:

\[
P_4(C) = \int_0^1 [1 - P_k(S \mid n)]^3 \, dP_k(S \mid s)
\]

Assuming that the ROC curve is based on Gaussian distributions with equal variance, then \( P_4(C) = 0.56 \) at the critical amplitude, that is, when \( d' = 1.0 \).
This argument is based on the assumption that the signal and noise distributions are equal and independent for all four alternatives. This is probably valid when the alternatives are similar, for example in the detection of a signal in one of four equal intervals. But how valid is it here? It is not immediately obvious how an observer would compare evidence variables from different types of motion, which might stimulate different sensors.

The assumption of equal distributions is justifiable if the underlying detection process is the same for all alternatives, and if all alternatives have the same index of detectability, \( d' \). This is ensured by adjusting the signal amplitude for each axis independently to keep it near the critical amplitude. The assumption of independence implies that the presence or absence of a signal on one axis (alternative) does not affect the signal and noise distributions on the other axes. This is on a much weaker footing. Indeed, this is what we are trying to find out in this experiment. We cannot take this assumption to be true, but we can probably assume that its failure will affect all of the data in a similar way, so that comparisons are still valid. We therefore concluded that performance in this task could be described at least approximately by equation 7.2.

The weakness of the identification task is that it is susceptible to bias. Suppose that a subject tends to select one response, say pitch, much more often than the others. Then his probability of responding correctly on a pitch trial is greater than that predicted by equation 7.2, and his probability of correctly identifying the other axes is reduced. The adaptive algorithm responds by lowering the target level for pitch, and raising it for the other axes. Clearly, this would tend to skew the results.

Swets and Pickett (1982) discuss the multi-choice detection and classification problem. They conclude that no adequate method has yet been developed to account for bias in this situation. They do suggest, however, that bias is less of a problem in a multi-choice case than in a simple yes/no decision, because the difference between responses is perceived to be less extreme. In the present experiment, we monitored the total frequency of responses for the four alternatives, and admonished subjects when they developed a noticeable bias. Even so, there is probably some effect of bias in the data, as discussed in Section 7.3.2.
7.3 Experimental Results

The experiment described above was run as a 5x5x4x6 full factorial design in which each subject was tested once in each condition. The experiment comprised 150 individual runs, each of which tested one subject in four different conditions. The main result from each run was an estimate of the critical amplitude for each of the four conditions. These results are tabulated in Appendix H.

The analysis of the data followed very closely the procedures described in Chapter 6. The analysis included three stages: consistency checks, statistical analysis to determine significant effects, and estimation of the slope of the psychometric curve. These stages are described in Sections 7.3.1 through 7.3.3.

7.3.1 Consistency Checks

The data were checked for consistency following the procedures discussed in Chapter 6. The purpose of this analysis was to find any trends with run number or trial number which might affect the conclusions. No trends were found, so it was concluded that the data are representative of typical steady state performance.

An analysis of covariance revealed no significant correlation between critical amplitude and run number. A plot of relative amplitude versus run number is shown in Figure 7.1. The relative amplitude for each condition is the critical amplitude, normalized by removing the effect of conditions and interactions which are uncorrelated with run number. This included the main effects of noise condition, axis, frequency and subject, and the noise x axis and axis x subject interaction effects. The coefficient of correlation between relative amplitude and run number is -0.02, indicating that there is no significant run-to-run trend in the data.

Figure 7.2 shows the mean and standard deviation of the relative target level as a function of trial number. The target level is an estimate of the critical amplitude which is calculated by the adaptive algorithm after each trial. The target levels for each subject and condition were normalized by removing the influence of all pure factors and all interactions which had a significant effect. This included the main effects of noise condition, axis, frequency and subject, as well as the noise x axis and axis x subject
interaction effects. Figure 7.2 is based on the normalized values for all subjects and conditions.

The mean is flat within 0.5 dB from trial 7 through trial 29. This indicates that, on average, performance is constant over the duration of a run. The low means on trials 1 and 2 are an artifact of the estimation procedure. The variability of the mean past trial 25 is due to the fact that each of these averages is based on a small number of runs. The experimental design called for only 20 trials per condition. However, because the axis is selected randomly and independently for each trial, a larger number of trials was run for some conditions.

The standard deviation decreases monotonically with trial number, indicating convergence as more data are obtained. The decrease is at a rate slower than $1/\sqrt{N}$, however, which suggests that the algorithm is not converging optimally. It seems likely that observer bias contributed extra variability to the data, and is responsible for this sub-optimal performance. The experimental error is the standard deviation after 20 or 25 trials: about 3.4 dB. This is considerably larger than the experimental error in the single axis experiment (2.0 dB).

### 7.3.2 Effect of Noise Condition

This section describes the statistical analysis of the data and discusses the significant effects. As with the single axis data, standard analysis of variance techniques were applied. The most interesting feature of the data is the noise × axis interaction effect, which shows the extent to which a motion signal is masked by noise on other axes.

A four factor mixed-effect model was used as the basis for statistical analysis. The form of the model is given in equation 6.5. Signal axis, signal frequency and noise condition were taken to be fixed effects; subject was taken to be a random effect. The data were converted to a decibel scale prior to analysis. A discussion of the analysis procedure is given in Chapter 6.

The results of the analysis of variance are summarized in Table 7.1. The random experimental error is quite large, with a standard deviation of 3.4 dB. As expected, the main effect of noise condition and the noise × axis interaction effect were found to be significant. There are also significant differences between subjects. Both the main effect
of subject and the axis \times subject interaction effect are highly significant. Neither frequency nor any interaction involving frequency has a significant effect. This implies that, within the confidence of the experiment, performance is invariant with frequency over the range tested. This finding is consistent with the single axis results of Chapter 6. Finally, the analysis shows the effect of axis to be significant. However, the difference between axes is meaningless because the translations and rotations were normalized by arbitrary factors.

The differences between subjects are summarized in Figure 7.3. This figure shows the performance of each subject in each noise condition, averaged over all axes and frequencies. The variation with noise condition is roughly the same for all subjects, but there is a significant difference in average performance. The most sensitive subject outdid the least sensitive by 2.4 dB, on average. Training is not a factor in this difference, since all subjects started the experiment with no experience. Motivation, on the other hand, may have been a factor. The author and his assistant were the most sensitive; the subject who was not a member of the simulation group was the least sensitive.

Figure 7.3 also shows the main effect of noise condition. This display format is not overly informative, since the results for all axes are averaged together. However, the results for condition A are the lowest, and the results for conditions Q and P are the highest. This implies that noise conditions X, Y, Q and P all had some masking effect, and that on average, Q and P had a larger effect than X and Y.

It is more informative to look at the results for each axis separately, as shown in Figure 7.4. This shows the noise \times axis interaction effect. In all cases, the lowest value was obtained for noise condition A. The values for all other axes were higher. This shows that adding noise on any axis tends to drive up the critical amplitude for signals on all other axes. In all cases but one, the highest value was found when the noise was on the same axis as the signal. The exception was sway: roll noise caused the greatest shift in critical amplitude for sway.

Before proceeding with the analysis, it is worthwhile to consider the effects of observer bias. In Section 7.2, we noted that the identification procedure is susceptible to bias. This can affect the data in two ways. First, if observer bias is present, but varies from run to run, the result will be greater random variability in the results. This may partially explain the large experimental error in this data set (\sigma = 3.4 \text{ dB}). Second, if observer bias
is correlated in some way with experimental condition, it could influence the overall results.

This second possibility is potentially serious, so we proposed a method of checking the data for bias and making corrections if necessary. It was not possible to correct individual runs, because each run had too few trials to provide a reliable estimate of the bias. Instead, the data were grouped by axis and noise condition, and the data in each group were adjusted by an average correction factor. The rationale for this grouping was twofold. First, it seems plausible that the addition of noise on one axis might bias a subject toward that axis and away from other axes, or vice versa. Second, the noise x axis interaction is the most interesting feature of the data. Correcting for bias might bring that feature into clearer focus.

The correction for each axis was based on the false alarm rate for that axis. Here, the false alarm rate, \( P(FA) \), is the probability of selecting a particular axis when the stimulus is on a different axis. For the unbiased case,

\[
P(FA) = \frac{1 - P(Hit)}{3}
\]

\[
= 0.147
\]  

(7.3)

since the adaptive algorithm adjusts the amplitude to obtain \( P(Hit) = P_4(C) = 0.56 \). The actual false alarm rates for each group, averaged over frequency and subject, are given in Table 7.2. Clearly, some correlated bias is present.

To calculate the correction, \( P(Hit) \) and \( P(FA) \) were plotted in forced-choice ROC space as shown in Figure 7.5. Green and Swets (1974) introduce the forced-choice ROC and note that the value of \( d' \) based on the forced-choice ROC is proportional to that based on the simple detection ROC. A symmetrical forced-choice ROC based on Gaussian distributions was plotted through the point, and the corresponding index, \( d_4' \), was determined. This value was divided by \( d_{4u} \), the value of \( d_4' \) for the unbiased case. This ratio is an estimate of the true index of detectability, \( d_{A'} \), of the signal. We want to estimate the signal amplitude for which \( d_{A'} \) equals 1.0. Assuming that \( d_{A'} \) is approximately proportional to signal energy, the square root of \( d_4'/d_{4u} \) is the appropriate correction factor to
apply to the estimated critical amplitude. Expressed in dB, the correction was calculated by:

\[ X_c = X - 10 \log(\frac{d_4'}{d_{4u}}) \]  

(7.4)

where \( X \) is the original estimate of critical amplitude (dB),
\( X_c \) is the corrected estimate of critical amplitude (dB),
and \( d_{4u} = 1.20 \) is the value of \( d_4' \) for the unbiased case.

The corrections used, and the new mean values for each noise condition × axis case, are summarized in Table 7.2. It must be emphasized that these are approximate corrections only. The corrections are generally small and may not fully compensate for the effects of bias. However, they are in the correct direction; trends shown by the corrected data are probably more accurate than those shown by the uncorrected data.

An analysis of variance was run on the corrected data. The only changes, of course, relate to noise, axis, and the noise × axis interaction. The results of the analysis are given in Table 7.3. The revised noise × axis interaction results are shown in Figure 7.6. These results are substantially the same as those shown in Figure 7.4, except that the same axis masking effect is slightly more pronounced.

In Figure 7.7, the noise × axis interaction data have been rearranged to show how the masking effect depends on the similarity between the signal axis and the noise axis. Proceeding from left to right along the abscissa of Figure 7.7, the relationship of the noise axis to the signal axis follows this progression:

<table>
<thead>
<tr>
<th>Relation of noise to signal</th>
<th>Signal axis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X</td>
</tr>
<tr>
<td>1. same axis</td>
<td>X</td>
</tr>
<tr>
<td>2. another axis in the same plane</td>
<td>Q</td>
</tr>
<tr>
<td>3. out of plane, similar motion</td>
<td>Y</td>
</tr>
<tr>
<td>4. out of plane, dissimilar motion</td>
<td>P</td>
</tr>
<tr>
<td>5. baseline condition A.</td>
<td>A</td>
</tr>
</tbody>
</table>

Thus, the similarity of the noise to the signal decreases from left to right. The critical amplitude found for each axis also decreases from left to right, implying that the inter-axis effect depends on the degree of similarity between the noise and signal axes.
Slightly different trends are seen for translations and rotations. Translational signals are masked more by in-plane rotations than rotational signals are by in-plane translations. This may have a simple explanation. The rotational motion of the simulator did not cause any change in specific force at the vestibular sensors, but geometry dictates that some specific force would be generated at other parts of the body. Therefore, rotational noise causes translational noise at the distributed motion sensors, such as the somatic and kinesthetic receptors, which could help to mask a translational signal.

The significance of the inter-axis effects was tested by averaging the results over all axes according to the above categories. Normalized to 0 dB for condition A, the means are:

1. same axis 8.3 dB  
2. different axis in same plane 4.1 dB  
3. out of plane, similar motion 2.8 dB  
4. out of plane, dissimilar motion 1.4 dB  
5. condition A 0.0 dB

The standard error of each mean is 0.31 dB. A DMR test (Duncan’s minimum range test, see Edwards, 1985) shows that the difference between each pair of means is highly significant (α < 0.005). Therefore, all of the inter-axis effects are significant, including masking by noise in category 4 above.

In an attempt to quantify the noise-axis interaction effect, we hypothesized that the "effective noise power" on a given axis could be written as a sum of the noise on all axes, with appropriate weights. For example, the effective surge noise power could be written as:

\[ N_{xe} = \gamma_{xx}N_x + \gamma_{xq}N_q + \gamma_{xy}N_y + \gamma_{xp}N_p \]  

(7.5)

where \( N_i \) is the noise power on axis \( i \), and \( \gamma_{ij} \) is the inter-axis weighting factor.

If the experimental results (critical amplitudes) are assumed to represent a fixed ratio of signal power to effective noise power then the weights are given approximately by:

\[ \gamma_{ij} = \frac{P_{ij} - P_{io}}{P_{ii} - P_{io}} \times \frac{N_{ih}}{N_{jh}} \times \gamma_{ii} \]  

(7.6)
where $P_{ij}$ is the signal power at critical amplitude for axis $i$ with noise on axis $j$, $P_{io}$ is the signal power at critical amplitude for noise condition A, and $N_{ih}$ is the power of the component of noise condition $i$ which is not included in condition A.

The same-axis terms $\gamma_{ii}$ were all taken to be 1. The calculated values of $\gamma_{ij}$ are given in Table 7.4. It is clear from these values that although the inter-axis effects are statistically significant, most of them are fairly small. For example, six times as much noise power is required to mask a pitch signal with roll noise rather than with pitch noise. The notable exceptions are $\gamma_{yp}$ and $\gamma_{xq}$. Roll noise effectively masks sway signals, and pitch noise effectively masks surge signals. A possible explanation is provided above. It should be remembered that these weighting factors are based only on the noise spectra used here. No attempt is made to predict the effect for different spectral shapes.

The idea that critical amplitude depends on the effective noise power helps to explain another feature of the results. On average, the difference in critical amplitude between condition A and the same-axis masking case is 8.3 dB, while the difference in same-axis noise power is 13 dB. However, a significant portion of the signal masking in condition A is due to inter-axis effects. The difference in effective noise power between conditions A and the same-axis masking case is 11.3 dB. Thus, the comparison is improved by almost 2 dB. There is still a 3 dB discrepancy, which may be attributed to several factors. First, we have not accounted for the masking effect of the yaw and heave components of condition A. Second, we have not accounted for "internal" noise, which might have a more pronounced effect when the external noise motion is small. Finally, we may not have accounted fully for the effects of observer bias. This may well be a factor, since the discrepancy here is larger than that found in the single axis experiment, where bias was eliminated by procedure.

The final effect that must be discussed is that of axis. It was noted above that the mean levels for rotations and translations could not be compared because they are normalized arbitrarily. However, we can compare surge to sway, and pitch to roll. The results which are most directly comparable are those for condition A and the same-axis masking condition. The means for these conditions, averaged over all frequencies and corrected for bias, are summarized in Table 7.5. The standard error of each mean is 0.62 dB. The results for surge and sway are not significantly different. The mean for roll, on the other hand, is almost 4 dB higher than that for pitch. This difference is highly significant.
(\(\alpha < 0.005\)). Since the noise spectra used for the two axes were the same, and the inter-axis effects were small for both axes, there is no obvious explanation for this difference.

In summary, it must be noted that the quantitative results presented in this section are approximate. They are the best that can be derived from an imperfect data set. The qualitative results, however, are correct and appear to make sense. Detectability is a function of the ratio of signal power to the effective noise power, which includes noise on the same axis as well as contributions from other axes. Some inter-axis masking is evident, but it is usually not as effective as same-axis masking. The effectiveness of inter-axis masking depends on the degree of similarity between the signal and noise axes.

### 7.3.3 Psychometric Curves

In Chapter 6, the slope of the psychometric curve was estimated from the forced-choice response data. The data from all runs were combined to form a large data set relating \(P_4(C)\) to relative signal amplitude. A logistic function was fit to the data. This section describes a similar analysis for the identification data.

The detailed experimental results comprise a record of signal presentation amplitudes and observer responses. The amplitude data for each subject and condition were normalized by the final target level for that subject and condition, then the data from all runs were pooled. In this experiment, the observers' task was a four alternative forced-choice task, so the pooled data set relates \(P_4(C)\) to relative amplitude. This is in contrast to the previous experiments, in which the data related \(P_2(C)\) or \(P_A\), the area under the ROC curve, to relative amplitude.

The data were fit with the logistic function, which was introduced in Chapter 5. For the four alternative case, the form of the function is:

\[
P_4(C) = 0.25 + \frac{0.75}{1 + \exp[-B(X-M)]}
\]

(7.7)

where \(X\) is the signal amplitude in dB, and \(M\) and \(B\) are the midpoint and slope of the function.
A maximum likelihood procedure was used to estimate $M$ and $B$ and to calculate the variance of the estimate of $B$. The value of $B$ found here may be compared directly with the values estimated in Chapters 5 and 6.

Figure 7.8 shows the fitted curve based on the entire data set (13203 trials). The abscissa is the relative amplitude in dB, and the ordinate is $P_4(C)$. Each point represents all of the data available within a given range of relative amplitude. The error bars for each point show the binomial standard deviation defined in equation 6.10. The relative amplitude on most trials was between $-4$ and $+4$ dB, so the variance of $P_4(C)$ is large for points near the tails of the curve. The maximum likelihood estimate of $B$ based on all the data is 0.46. The standard deviation of the estimate is 0.014. This value of $B$ is in good agreement with that found in Chapter 6.

The data for each subject is shown individually in Figure 7.9. In each case, the curve shown has a slope of 0.46. A maximum likelihood estimate of $B$ was determined separately for each subject. Each estimate was based on about 2200 trials. The results are given in Table 7.6. The standard deviation of $B$ between subjects was 0.034. An F-test showed that the difference in slope between subjects is not significant.

The data were also considered separately by noise condition, following the relational categories introduced in Section 7.3.2. The results of this analysis are given in Table 7.7. The difference in slope between noise conditions was not found to be significant.
Table 7.1 Results of the analysis of variance for the data of experiment 3.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Degrees of Freedom</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Ratio</th>
<th>P(x&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise</td>
<td>4</td>
<td>1981.3</td>
<td>495.32</td>
<td>39.11</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Axis</td>
<td>3</td>
<td>1956.5</td>
<td>652.18</td>
<td>16.75</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Noise × Axis</td>
<td>12</td>
<td>2697.9</td>
<td>224.83</td>
<td>12.70</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Frequency</td>
<td>4</td>
<td>27.3</td>
<td>6.82</td>
<td>0.46</td>
<td>0.7623</td>
</tr>
<tr>
<td>Noise × Freq</td>
<td>16</td>
<td>185.5</td>
<td>11.59</td>
<td>1.02</td>
<td>0.4356</td>
</tr>
<tr>
<td>Axis × Freq</td>
<td>12</td>
<td>145.3</td>
<td>12.11</td>
<td>1.07</td>
<td>0.3899</td>
</tr>
<tr>
<td>N × A × F</td>
<td>48</td>
<td>470.2</td>
<td>9.80</td>
<td>0.86</td>
<td>0.7262</td>
</tr>
<tr>
<td>Subject</td>
<td>5</td>
<td>503.1</td>
<td>100.63</td>
<td>8.86</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Noise × Subj</td>
<td>20</td>
<td>253.3</td>
<td>12.66</td>
<td>1.11</td>
<td>0.3349</td>
</tr>
<tr>
<td>Axis × Subj</td>
<td>15</td>
<td>584.0</td>
<td>38.93</td>
<td>3.43</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>N × A × S</td>
<td>60</td>
<td>1062.0</td>
<td>17.70</td>
<td>1.56</td>
<td>0.0107</td>
</tr>
<tr>
<td>Freq × Subj</td>
<td>20</td>
<td>294.9</td>
<td>14.74</td>
<td>1.30</td>
<td>0.1812</td>
</tr>
<tr>
<td>N × F × S</td>
<td>80</td>
<td>645.1</td>
<td>8.06</td>
<td>0.71</td>
<td>0.9632</td>
</tr>
<tr>
<td>A × F × S</td>
<td>60</td>
<td>509.6</td>
<td>8.49</td>
<td>0.75</td>
<td>0.9099</td>
</tr>
<tr>
<td>N × A × F × S</td>
<td>240</td>
<td>2726.9</td>
<td>11.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>599</strong></td>
<td><strong>14042.9</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 7.2  Corrections applied to the data based on the observed false alarm rates. Means values shown are in milli-g for surge and sway, and in degrees per second for pitch and roll.

<table>
<thead>
<tr>
<th>Signal Axis</th>
<th>Noise Cond.</th>
<th>P(FA)</th>
<th>$d_4'$</th>
<th>$10 \log(d_4'/d_{40})$</th>
<th>Corrected Geom. mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surge</td>
<td>X</td>
<td>.223</td>
<td>.913</td>
<td>-1.187</td>
<td>6.84</td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.163</td>
<td>1.133</td>
<td>-0.250</td>
<td>4.95</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>.113</td>
<td>1.362</td>
<td>0.550</td>
<td>3.96</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>.196</td>
<td>1.007</td>
<td>-0.762</td>
<td>3.64</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>.149</td>
<td>1.191</td>
<td>-0.033</td>
<td>3.00</td>
</tr>
<tr>
<td>Pitch</td>
<td>X</td>
<td>.166</td>
<td>1.121</td>
<td>-0.296</td>
<td>0.393</td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.204</td>
<td>.979</td>
<td>-0.884</td>
<td>0.855</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>.179</td>
<td>1.070</td>
<td>-0.498</td>
<td>0.315</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>.132</td>
<td>1.268</td>
<td>0.239</td>
<td>0.427</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>.170</td>
<td>1.105</td>
<td>-0.358</td>
<td>0.290</td>
</tr>
<tr>
<td>Sway</td>
<td>X</td>
<td>.124</td>
<td>1.306</td>
<td>0.368</td>
<td>4.06</td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.148</td>
<td>1.196</td>
<td>-0.015</td>
<td>3.76</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>.203</td>
<td>.982</td>
<td>-0.871</td>
<td>6.67</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>.121</td>
<td>1.321</td>
<td>0.417</td>
<td>6.68</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>.168</td>
<td>1.113</td>
<td>-0.327</td>
<td>3.03</td>
</tr>
<tr>
<td>Roll</td>
<td>X</td>
<td>.108</td>
<td>1.389</td>
<td>0.635</td>
<td>0.499</td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>.089</td>
<td>1.498</td>
<td>0.963</td>
<td>0.599</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>.095</td>
<td>1.463</td>
<td>0.861</td>
<td>0.594</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>.175</td>
<td>1.086</td>
<td>-0.434</td>
<td>1.342</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>.105</td>
<td>1.406</td>
<td>0.688</td>
<td>0.435</td>
</tr>
</tbody>
</table>
### Table 7.3 Results of the analysis of variance for the corrected data from experiment 3.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Degrees of Freedom</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Ratio</th>
<th>P(x&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise</td>
<td>4</td>
<td>2058.3</td>
<td>514.56</td>
<td>40.63</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Axis</td>
<td>3</td>
<td>1250.1</td>
<td>416.71</td>
<td>10.70</td>
<td>0.0005</td>
</tr>
<tr>
<td>Noise × Axis</td>
<td>12</td>
<td>3387.4</td>
<td>282.28</td>
<td>15.95</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Frequency</td>
<td>4</td>
<td>27.3</td>
<td>6.82</td>
<td>0.46</td>
<td>0.7623</td>
</tr>
<tr>
<td>Noise × Freq</td>
<td>16</td>
<td>185.5</td>
<td>11.59</td>
<td>1.02</td>
<td>0.4357</td>
</tr>
<tr>
<td>Axis × Freq</td>
<td>12</td>
<td>145.3</td>
<td>12.11</td>
<td>1.07</td>
<td>0.3899</td>
</tr>
<tr>
<td>N × A × F</td>
<td>48</td>
<td>470.2</td>
<td>9.80</td>
<td>0.86</td>
<td>0.7262</td>
</tr>
<tr>
<td>Subject</td>
<td>5</td>
<td>503.1</td>
<td>100.63</td>
<td>8.86</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Noise × Subj</td>
<td>20</td>
<td>253.3</td>
<td>12.66</td>
<td>1.11</td>
<td>0.3349</td>
</tr>
<tr>
<td>Axis × Subj</td>
<td>15</td>
<td>584.0</td>
<td>38.93</td>
<td>3.43</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>N × A × S</td>
<td>60</td>
<td>1062.0</td>
<td>17.70</td>
<td>1.56</td>
<td>0.0107</td>
</tr>
<tr>
<td>Freq × Subj</td>
<td>20</td>
<td>294.9</td>
<td>14.74</td>
<td>1.30</td>
<td>0.1811</td>
</tr>
<tr>
<td>N × F × S</td>
<td>80</td>
<td>645.1</td>
<td>8.06</td>
<td>0.71</td>
<td>0.9632</td>
</tr>
<tr>
<td>A × F × S</td>
<td>60</td>
<td>509.6</td>
<td>8.49</td>
<td>0.75</td>
<td>0.9100</td>
</tr>
<tr>
<td>N × A × F × S</td>
<td>240</td>
<td>2726.9</td>
<td>11.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>599</td>
<td>14103.0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 7.4  Estimates of the weighting factors for inter-axis masking.

<table>
<thead>
<tr>
<th>Signal Axis</th>
<th>Noise Axis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X</td>
</tr>
<tr>
<td>X</td>
<td>1.00</td>
</tr>
<tr>
<td>Q</td>
<td>0.11&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>Y</td>
<td>0.21</td>
</tr>
<tr>
<td>P</td>
<td>0.04&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Factors are dimensionless unless noted.

Otherwise, units are:

1. (m/rad)<sup>2</sup> for γ<sub>xq</sub>, γ<sub>xp</sub>, γ<sub>yq</sub> and γ<sub>yp</sub>
2. (rad/m)<sup>2</sup> for γ<sub>xq</sub>, γ<sub>yq</sub>, γ<sub>xp</sub> and γ<sub>yp</sub>
Table 7.5  Summary of the critical amplitudes for surge, sway, pitch and roll in two different noise conditions. Data for surge and sway are in milli-g; data for pitch and roll are in degrees per second.

<table>
<thead>
<tr>
<th>Noise Cond.</th>
<th>Subject</th>
<th>Surge</th>
<th>Sway</th>
<th>Pitch</th>
<th>Roll</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition A</td>
<td>DG</td>
<td>2.5</td>
<td>2.5</td>
<td>0.25</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>PG</td>
<td>3.0</td>
<td>2.7</td>
<td>0.24</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>3.6</td>
<td>3.0</td>
<td>0.31</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>MN</td>
<td>3.0</td>
<td>3.6</td>
<td>0.25</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>PR</td>
<td>2.8</td>
<td>3.2</td>
<td>0.38</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>DS</td>
<td>3.2</td>
<td>3.4</td>
<td>0.33</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>Geom. mean</td>
<td>3.0</td>
<td>3.0</td>
<td>0.29</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>Std. dev</td>
<td>0.40</td>
<td>0.42</td>
<td>0.056</td>
<td>0.069</td>
</tr>
<tr>
<td>Same axis as signal</td>
<td>DG</td>
<td>11.7</td>
<td>8.1</td>
<td>0.69</td>
<td>1.30</td>
</tr>
<tr>
<td></td>
<td>PG</td>
<td>4.9</td>
<td>5.3</td>
<td>1.10</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>6.9</td>
<td>8.0</td>
<td>0.54</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>MN</td>
<td>7.4</td>
<td>9.9</td>
<td>0.72</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>PR</td>
<td>6.1</td>
<td>4.8</td>
<td>1.58</td>
<td>2.17</td>
</tr>
<tr>
<td></td>
<td>DS</td>
<td>5.8</td>
<td>5.3</td>
<td>0.83</td>
<td>1.53</td>
</tr>
<tr>
<td></td>
<td>Geom. mean</td>
<td>6.8</td>
<td>6.7</td>
<td>0.86</td>
<td>1.34</td>
</tr>
<tr>
<td></td>
<td>Std. dev</td>
<td>2.41</td>
<td>2.06</td>
<td>0.377</td>
<td>0.423</td>
</tr>
</tbody>
</table>
Table 7.6  Maximum likelihood estimates of the psychometric curve slope, B, for each subject.

<table>
<thead>
<tr>
<th>Subject</th>
<th>B</th>
<th>σ_B</th>
</tr>
</thead>
<tbody>
<tr>
<td>DG</td>
<td>.458</td>
<td>.035</td>
</tr>
<tr>
<td>PG</td>
<td>.406</td>
<td>.027</td>
</tr>
<tr>
<td>GG</td>
<td>.456</td>
<td>.034</td>
</tr>
<tr>
<td>MN</td>
<td>.494</td>
<td>.036</td>
</tr>
<tr>
<td>PR</td>
<td>.497</td>
<td>.036</td>
</tr>
<tr>
<td>DS</td>
<td>.484</td>
<td>.034</td>
</tr>
<tr>
<td>All data</td>
<td>.462</td>
<td>.014</td>
</tr>
</tbody>
</table>

Table 7.7  Maximum likelihood estimates of the psychometric curve slope, B, for each noise condition. Data were grouped according to the similarity between signal and noise axes of motion.

<table>
<thead>
<tr>
<th>Noise Condition</th>
<th>B</th>
<th>σ_B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same axis</td>
<td>.460</td>
<td>.033</td>
</tr>
<tr>
<td>Same plane</td>
<td>.501</td>
<td>.034</td>
</tr>
<tr>
<td>Same mode</td>
<td>.495</td>
<td>.030</td>
</tr>
<tr>
<td>Other</td>
<td>.541</td>
<td>.036</td>
</tr>
<tr>
<td>Condition A</td>
<td>.417</td>
<td>.030</td>
</tr>
</tbody>
</table>
Figure 7.1  Normalized target levels as a function of run number, for experiment 3.
Figure 7.2  Mean and standard deviation of the normalized target levels, plotted as a function of trial number, for experiment 3.
Figure 7.3  
Average result for each subject, plotted against noise condition. Each data point is the mean over all frequencies and axes for one subject and noise condition, minus the grand mean.
Figure 7.4 Noise × axis interaction effects, based on the uncorrected data for experiment 3. Each data point is the mean over all frequencies for one subject, axis, and noise condition, minus the appropriate subject × axis mean.
Figure 7.5  Determination of the forced-choice ROC curve from the false alarm rate.
Figure 7.6 Noise x axis interaction effects, based on the data of experiment 3 after correction for response bias. Each data point is the mean over all frequencies for one subject, axis, and noise condition, minus the appropriate subject x axis mean.
Figure 7.7 Noise × axis interaction data showing the dependence of masking effectiveness on the similarity between the signal and noise axes.
Figure 7.8  Psychometric curve based on all data from experiment 3 (13203 trials). For the logistic function shown, \( B = 0.46 \).
Figure 7.9  Response data for each subject. For the logistic function shown, $B = 0.46$. 
Chapter 8

Discussion of Experimental Results

8.1 Introduction

Chapters 5, 6 and 7 described three experiments which were run to assess human sensitivity to sinusoidal motion in a background of random motion. Each chapter presents the data from one experiment, and compares the data with some aspects of the proposed signal detection models. The purpose of this chapter is to look at all the data together, in order to summarize the main findings and provide an overall comparison of the results to the detection model predictions. In the following discussion, the experiments described in Chapters 5, 6 and 7 are referred to as experiments 1, 2 and 3, respectively.

Section 8.2 compares the results of the three experiments and finds them to be consistent. Section 8.3 reviews the main findings, and compares the data to the three ideal detection models. Overall, the energy detector agrees best with the data. A more detailed comparison of the experimental data to the energy detector, based on the results of a computer simulation, is presented in Section 8.4. Section 8.5 compares the findings of this study to previous results reported in the literature. Section 8.6 presents a simplified model which can be used to estimate signal detectability for arbitrary signal and noise conditions.

Finally, Section 8.7 closes the chapter with a discussion of the implications of this study for flight simulation.

8.2 Comparison Between Experiments

As discussed in previous chapters, the data set for each experiment is fairly consistent within itself. Analysis of performance as a function of run number showed that learning effects over the course of each experiment were negligible. The mean performance was approximately constant with trial number, indicating that, on average, there was no significant change in sensitivity over the duration of a single run. The results were
probably affected to some extent by variations in subject motivation and alertness. However, these factors did not cause any specific trends in the data, but simply resulted in an increase in the experimental error (scatter).

It is useful to compare the results of the three experiments with each other. There are three main points of comparison. First, all three experiments provided an estimate of the critical amplitude for pitch and surge signals, at 0.6 Hz, with noise condition A. The values obtained (and the respective standard deviations) are as follows:

<table>
<thead>
<tr>
<th>Surge (milli-g)</th>
<th>Pitch (%/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>experiment 1:</td>
<td>2.6 (0.37)</td>
</tr>
<tr>
<td>experiment 2:</td>
<td>2.8 (0.76)</td>
</tr>
<tr>
<td>experiment 3:</td>
<td>3.0 (0.40)</td>
</tr>
</tbody>
</table>

The agreement between experiments is quite good; the largest difference seen is about 15%.

Second, experiments 2 and 3 each give an estimate of the masking effectiveness of broadband noise on the same axis as the signal. The observed effect was similar in both experiments, although the sensitivity shift was numerically larger in experiment 2 (single-axis) than in experiment 3 (multi-axis). Changes in subjective bias with noise condition probably reduced the apparent masking effect in the multi-axis experiment.

Finally, all three experiments provide an estimate of the slope of the psychometric curve. The slope parameter, B, is based on the logistic function, which is defined in equation 5.17. The maximum likelihood estimates of B based on all data from each experiment are:

- experiment 1: $B = 0.28 \ (\sigma_B = 0.031)$
- experiment 2: $B = 0.45 \ (\sigma_B = 0.025)$
- experiment 3: $B = 0.46 \ (\sigma_B = 0.014)$

Clearly, the data from experiment 1 suggest a much shallower curve than is indicated by experiments 2 and 3. Each curve fit is based on data from several thousand trials, so the variance of the slope estimates is small. The difference in slope between experiments is highly significant ($\alpha < 0.005$). No explanation for this disparity was found.
The mechanisms of detection were similar in all three experiments. For high amplitude translational motion (surge and sway), subjects identified body sway, head movement, and neck proprioception as major contributors to their perception of motion. For lower amplitude surge and sway, near the critical amplitude, these postural cues were much less obvious. Subjects reported sensing motion, but were unable to locate the source of the sensation. Vestibular cues probably dominated at these low levels.

Postural cues were not reported to be important in the detection of angular motion. Subjects reported feeling an empty or weightless sensation in the pit of the stomach, and the sensation of being on a swing. Most subjects reported that they used the apparent motion of the illuminated dashboard to help them detect the motion. This illusion of apparent motion is referred to as the oculogyral illusion.

Finally, it is worthwhile commenting on the issue of individual differences. Small, but significant differences in performance between subjects were found in all three experiments. For each experiment, the standard deviation of the variability in critical amplitude between subjects was as follows:

- experiment 1: 1.65 dB
- experiment 2: 1.50 dB
- experiment 3: 1.00 dB

These differences are much smaller than the value reported by Rodenburg et al (1981a). Based on an analysis of data collected by Clark and Stewart (1969), they found a standard deviation of 5.5 dB for 53 subjects. It is not clear if the low value found here indicates that individual differences are smaller for signal-in-noise detection than for simple motion detection, or if it is simply an artifact of the small size of the subject pool.

8.3 Comparison with Models

It was proposed earlier that motion perception might be modelled by one of three ideal detection models. The experimental data provide four points for comparison with the models:

1. the ROC curve shape
2. the signal-to-noise ratio at critical amplitude ($d_{A'} = 1.0$)
3. the slope of the psychometric curve
4. the variation of critical amplitude with noise bandwidth.

These four points, taken together, enable us to draw some solid conclusions about the applicability of the detection models to motion perception.

First, consider the ROC curve shape. In the first experiment, ROC data were obtained for 5 subjects detecting 0.6 Hz surge and pitch signals with noise condition A. The data were satisfactorily fit by ROC curves based on Gaussian distributions, with $\sigma_n < \sigma_s$. The average slope parameter, $b$, of the fitted ROC curves was 0.90. Two different methods of analysis showed that the data were unlikely to result from a detector for which the ROC slope parameter is outside the range (0.85,0.95).

The correlation detector results in symmetrical Gaussian ROC curves with $b = 1.0$. On the basis of the experimental data, this model was rejected with a significance level of 0.025. The envelope detector results in slightly skewed curves which provide a good fit to the data. For the energy detector, it is necessary to estimate the slope parameter predicted by the model, using equation 3.68. Estimates of $E_s/N_0$ and WT can be derived from the results of experiment 2. For surge at 0.6 Hz with noise condition C, $E_s/N_0 = 10$ at the critical amplitude. If we let W equal the bandwidth of the internal filter and T equal the duration of each presentation interval, then $WT \approx 12$. Using these values in equation 3.68 yields $b = 0.88$, which is well within the range indicated by the data. Therefore the envelope and energy detectors fit the ROC data well, while the correlation detector does not.

The second point to consider is the signal-to-noise ratio at critical amplitude. Based on the results of Chapter 6, it was concluded that signal detectability is a function of signal-to-noise ratio. This finding is consistent with all three detection models. However, the three detectors predict markedly different performance. The signal-to-noise ratio at $d_{A'} = 1.0$ therefore provides another useful point of comparison between experiment and theory. For this comparison, it is best to use the experimental results found with noise spectrum C. This permits more accurate calculation of the signal-to-noise ratio, since the noise entering the detector in this condition is almost entirely external noise. Figure 8.1 shows the signal-to-noise ratios (at $d_{A'} = 1.0$) for all subjects, for pitch and surge motion at 0.6 Hz. Also shown in Figure 8.1 are psychometric curves based on the three detectors. It is clear that the human observers are less sensitive than predicted by any of
the three models. The energy detector is the closest, and differs from the data by about 3 to 4 dB.

The third point of comparison between experiment and theory is the slope of the psychometric curve, B. Each of the three experiments provided an estimate of the slope. As discussed in Section 8.2, there is a discrepancy in the estimates. Experiment 1 yields a value of 0.28, while experiments 2 and 3 yield a value of 0.45. A "best" estimate of B based on the results of all three experiments is about 0.35 to 0.40. Thus the data follow a steeper curve than that predicted by the correlation detector (B = 0.20), and agree reasonably well with curves predicted by the envelope detector (B = 0.33) and the energy detector (B = 0.35 to 0.39).

A comparison of the data of experiment 2 with the psychometric curve predicted by the energy detector (with WT = 12) is shown in Figure 8.2. The data have been shifted 4 dB to the left to coincide with the theoretical curve. The general shape of the curve is in good agreement with the trend shown by the data.

The final point to be considered is the variation in detectability with noise bandwidth. According to equations 3.38 and 3.48, the signal detectability predicted by the correlation and envelope detectors is independent of the noise bandwidth. It depends only on the noise power at the signal frequency. For the energy detector, signal detectability decreases as the noise bandwidth increases, until the bandwidth exceeds that of the internal filter. The data presented in Chapter 6 provide some evidence on this point.

The relevant data are the critical amplitudes found for 0.2 Hz signals with noise spectra D (narrow band) and C (broadband), and for 0.6 Hz signals with spectra E (narrow band) and C (broadband). These data are summarized in Table 8.1. The critical amplitudes are always lower for the narrow band case than for the broadband case. The mean difference is 2.5 dB; the standard deviation of the difference, based on the experimental error for experiment 2, is 0.63 dB. Therefore, the difference is statistically significant (α < 0.005), indicating that detectability does vary with noise bandwidth. This finding supports the energy detector over the other two.

On the basis of the foregoing discussion, it is apparent that, of the three ideal detectors considered, the energy detector agrees best with the data. A detailed comparison of the experimental results with the energy detector is presented in the next section.
8.4 Detailed Comparison with Energy Detector

As discussed above, some of the experimental results compare well with the energy detection model. However, it is difficult to compare all of the results directly with theory, because the experimental conditions do not meet some of the simplifying assumptions made in the derivation of the theory. The main assumption made in Chapter 3 concerned the shape of the noise spectrum. The noise spectrum entering the detector was assumed to be a rectangular spectrum spanning a frequency band which contained the signal frequency.

The experimental conditions deviate from this assumption for three reasons. First, the observer's internal filter is unlikely to be a rectangular filter. The simplified analysis in Chapter 6 suggested that the internal filter might be approximated by a second order bandpass filter. Second, the narrow band noise spectra D, E, F and G do not always contain the signal frequency. Finally, the pitch spectra are flat (in their band) in angular acceleration, while our model of the pitch detector is an angular velocity detector. The power density of the pitch spectra, expressed in terms of velocity, is proportional to $1/w^2$ over the appropriate "band."

No attempt was made to derive theoretical expressions for the energy detector performance in these more complex situations. Instead, a Monte Carlo simulation of the energy detector was set up to facilitate comparison between experiment and theory. The simulation served as a basis for estimating some of the model parameters, and provided numerical results for comparison with the experimental data.

8.4.1 Description of the Simulation

A block diagram of the energy detector simulation is shown in Figure 8.3. The simulation was designed to match the experimental conditions as closely as possible. The left hand side of Figure 8.3 shows the four motion inputs to the human observer:

1. $s(t)$, the gated sinusoidal signal.

2. $r_A(t)$, the low-power, broadband noise (spectrum A) which was common to all experimental runs.
3. \( r_M(t) \), the additional high-power random noise defined by spectrum B, C, D, E, F or G, as appropriate.

4. \( r_Z(t) \), the inter-axis masking noise which accounted for the fact that noise condition A comprised random motion on all 6 axes. In a very approximate application of equation 7.5, using the data in Table 7.4, \( r_Z(t) \) was set to zero for simulations of pitch detection, and set equal to 0.5 \( r_A(t) \) for simulations of surge detection.

The sum of these inputs, plus the observer's internal noise \( r_I(t) \), made up the entire input to the detector. The model used for \( r_I(t) \) is discussed in Section 8.4.2. The first component of the detector was a second order bandpass filter. The form and characteristics of the filter are discussed in Section 8.4.3. The filtered inputs were squared and integrated to obtain an evidence variable, \( e \), equal to the energy of the filtered waveform. This energy variable is defined as:

\[
e = \int_0^T x^2(t) \exp\left[-\frac{(T-t)}{T_m}\right] dt \tag{8.1}
\]

where \( x(t) \) is the filtered waveform:

\[
x(t) = \int_{-\infty}^t h(t-\tau) y(\tau) d\tau
\]

where \( y(t) = s(t) + r_A(t) + r_M(t) + r_Z(t) + r_I(t) \), and \( h(t) \) is the impulse response of the internal filter. The exponential decay in equation 8.1 models an observer's limited memory. This term is discussed briefly below.

All of the signal and noise inputs were calculated using the same algorithms used to generate signal and noise motion commands for the experiments. Simulations were run using the 28 different combinations of signal and noise considered in experiment 2.

Each simulation was run as a series of trials similar to those used in the experiments. Each simulation trial was a single interval, during which a signal \( s(t) \) was superimposed on the combined random inputs. As in experiment 2, the interval duration was 23 seconds for 0.2 Hz signals, and 16.3 seconds for 0.6 Hz signals. The program integrated the square of the input over the duration of the trial to obtain a single estimate of the
evidence variable, e. Repeating the simulation for a large number of trials yielded the conditional probability density functions \( f(e \mid s) \) and \( f(e \mid n) \). The signal detectability and ROC shape were calculated directly from the conditional densities.

Three model parameters influence the detector performance: \( \Phi_n \), the internal noise power; the bandwidth of the internal filter; and the exponential decay constant, \( T_m \), in the energy convolution integral. The first two parameters were investigated in detail, as discussed below. \( T_m \) models an observer's limited ability to remember or integrate the input over a long period of time. Decreasing \( T_m \) tends to reduce the detector sensitivity for all signal and noise conditions. The effect of this parameter was not considered in detail. For most of the simulation runs, \( T_m \) was set to positive infinity.

The first objective of the analysis presented here was to find the combination of model parameters which yielded the best agreement with experiment. The quality of the agreement obtained then provides a basis for evaluating the validity of the energy detection model for motion detection.

8.4.2 Estimation of the Internal Noise Power

We concluded from the experimental data presented in Chapter 6 that signal detectability is a function of the signal to noise ratio, \( E_s/N_0 \). However, the experimental data do not follow this rule exactly. The shift in critical amplitude (expressed in dB) between noise conditions A and C was always slightly lower than the shift in the power density of the noise motion. This suggests that there might be some constant noise source for which we have not yet accounted. We assume that this constant noise source is the internal noise, \( r(t) \).

In this section, an approximate method is developed to estimate the internal noise power. For this analysis, it was necessary to assume a particular spectral shape for the internal noise. In the absence of any neurological data, we assumed that the internal noise is broadband with a spectral shape similar to that of noise condition C. This simplified the analysis. The analysis yielded estimates of the internal noise power, which were then used as inputs to the energy detector simulation program.
According to the signal detection model, detectability is a function of signal-to-noise ratio, all else being equal. For noise spectra with similar shapes, and a given value of $d_A':$

$$\frac{E_s}{N_o} = \text{constant} \tag{8.2}$$

This should hold for noise conditions A, B and C because the spectral shape of the noise is the same in each case. Compare the results found for conditions A and C. The critical amplitude corresponds to $d_A' = 1.0$ in both cases. Therefore,

$$\left(\frac{E_s}{N_o}\right)_A = \left(\frac{E_s}{N_o}\right)_C \tag{8.3}$$

In each case, $E_s$ is proportional to $A_{cr}^2$, the critical amplitude squared. $N_o$ is the noise power density at the signal frequency, $f_s$. For condition A,

$$N_o = \Phi_{\Pi} + \Phi_{AA} + \Phi_{ZZ} \tag{8.4}$$

and for condition C,

$$N_o = \Phi_{\Pi} + \Phi_{AA} + \Phi_{ZZ} + \Phi_{MM} \tag{8.5}$$

where $\Phi_{\Pi}$, $\Phi_{AA}$, $\Phi_{ZZ}$ and $\Phi_{MM}$ represent the spectral density at $f_s$ due to $r_I(t)$, $r_A(t)$, $r_Z(t)$ and $r_M(t)$, respectively. Combining equations 8.3, 8.4 and 8.5 and solving for $\Phi_{\Pi}$ yields the following expression:

$$\Phi_{\Pi} = \frac{\Phi_{MM}}{\Omega^2 - 1} - \Phi_{AA} - \Phi_{ZZ} \tag{8.6}$$

where $\Omega = A_{cr}(C)/A_{cr}(A)$. For each signal axis and frequency, the value of $\Omega$ can be calculated from the results of experiment 2. The values of $\Phi_{MM}$ and $\Phi_{AA}$ are known, and $\Phi_{ZZ}$ can be estimated using equation 7.5. Therefore, equation 8.6 provides an estimate of the internal noise power, $\Phi_{\Pi}$. Estimates of $\Phi_{\Pi}$ were made separately for each axis/frequency combination, based on the data from experiment 2. These values are tabulated in Table 8.2.

These estimates can be used to extrapolate the experimental results to the "zero-noise" condition, that is, detection of sinusoidal motion in the absence of any external noise. In
this case, \( N_o = \Phi_{II} \). Denoting the critical amplitude in this case as \( A_o \), it follows from equation 8.2 that

\[
\frac{A_o}{A_{cr}(A)} = \left( \frac{\Phi_{II}}{\Phi_{II} + \Phi_{AA} + \Phi_{ZZ}} \right)^{1/2} \tag{8.7}
\]

Estimated values of \( A_o \) for all four signal conditions, based on this expression, are given in Table 8.2. It should be noted that the estimated values of \( \Phi_{II} \) and \( A_o \) are very approximate.

The neural signal-to-noise power ratio for each case was computed based on \( A_o \) and \( A_{cr} \) and the estimates of RMS neural noise given in Chapter 2. The estimated ratios, which are tabulated in Table 8.2, seem fairly reasonable as indicators of the limits of detectability. No explanation is offered for the factor of two difference between the values obtained for surge and pitch.

As seen from Table 8.2, this analysis shows that internal noise had a greater effect on the data at 0.2 Hz than at 0.6 Hz. Consequently, estimates of the internal noise power are higher for 0.2 Hz than for 0.6 Hz. The reason for this is not clear, but the following tentative explanation is offered. As formulated above, the internal noise includes not only neural noise, but also such postural sources as spontaneous body sway. It is hypothesized that the main power of such motions is at very low frequencies. Breathing, for example, results in small body motions and changes in muscular tension at the breathing frequency, which is typically on the order of 0.2 Hz. Several subjects reported that, for 0.2 Hz surge at very low amplitudes, body motion due to breathing occasionally interfered with their attempts to detect the motion. When the signal to be detected is at a higher frequency, much of this low frequency internal noise may be attenuated by the observer's internal filter.

### 8.4.3 The Shape of the Internal Filter

The internal filter was assumed to be a second order bandpass filter with a transfer function of the form:

\[
H(s) = \frac{2\zeta\omega_n s}{s^2 + 2\zeta\omega_n + \omega_n^2} \tag{8.8}
\]
where $\omega_s$ is the signal frequency in radians per second. The filter bandwidth is dependent on the damping parameter, $\zeta$. The primary purpose of the parameter estimation runs was to estimate $\zeta$.

The parameter estimation procedure was rather involved. Each signal axis and frequency combination was analyzed separately, although with the constraint that the final selection of filter bandwidth must be the same for both axes. For a given value of $\zeta$, a series of simulations was run to generate results for noise conditions C, D, E, F and G. For each noise condition, the signal detectability, $d_A'$, was determined for several different signal amplitudes. A quadratic fit on these values yielded the critical amplitude (for which $d_A' = 1.0$). Thus the series of simulations generated estimates of the critical amplitude for the given signal axis and frequency, for noise conditions C, D, E, F and G. Critical amplitudes for noise conditions A and B were inferred from the value found for condition C, using equation 8.2.

The variation of the critical amplitude with noise condition was compared with that observed in the experimental data. The value of $\zeta$ for each signal frequency was chosen to minimize the variation of $\Delta A$ between noise conditions, where:

$$\Delta A = 20 \log \left( \frac{A_{cr}(\text{experiment})}{A_{cr}(\text{simulation})} \right)$$  \hspace{1cm} (8.9)

Simulations of 200 trials were run for each signal and noise condition for various values of $\zeta$ between 0.1 and 10. Because the noise inputs are random, the critical amplitude estimated from each simulation is also a random variable. Two consecutive runs using the same model parameters may yield slightly different results. Consequently, it was not possible to use a standard minimization procedure to determine the precise best value of $\zeta$. The variance of $\Delta A$ is approximately minimized by using the following values of $\zeta$: $\zeta = 1.2$ for signals at 0.2 Hz, and $\zeta = 0.6$ for signals at 0.6 Hz. The fit is insensitive to changes of $\pm 0.1$ to 0.2 in the value of $\zeta$. It is interesting to note that these values are identical to those estimated from the simple analysis presented in Chapter 6. The resulting critical bandwidth is 0.48 Hz, for signals at 0.2 Hz, and 0.72 Hz for signals at 0.6 Hz. The response characteristics of the internal filters are shown in Figure 6.10.
8.4.4 Simulation Results

For the internal filter models selected, detailed simulations comprising 2000 trials were run for each signal condition, for noise conditions C, D, E, F and G. As described above, density functions $f(e|s)$ and $f(e|n)$ were obtained from each simulation. Estimates of the critical amplitude were based on analysis of the density functions. The purpose of running 2000 trials was to obtain relatively low-variance estimates of the detector performance for comparison with the experimental results.

Typical histograms of $f(e|s)$ and $f(e|s)$ are shown in Figures 8.4 and 8.5. The histograms of $f(e|n)$ shown are fit by $\chi^2$ distributions with the indicated number of degrees of freedom. The histograms of $f(e|s)$ are fit by non-central $\chi^2$ distributions with the same number of degrees of freedom. The fit to the histograms is generally good, indicating that the assumption of $\chi^2$ distributions is reasonably valid.

In each case the number of degrees of freedom was calculated using the fact that the mean of a $\chi^2$ distribution is the number of degrees of freedom, and the variance is twice the number of degrees of freedom. Therefore, the number of degrees of freedom, $N$, is given by:

$$N = 2 \frac{\mu^2}{\sigma^2} \quad (8.10)$$

For the surge case, the value of $N$ estimated from the simulation results may be compared with that predicted by the theory. For a rectangular input spectrum, equation 3.59 shows that $2z_n/N_0$ is $\chi^2$ with $2WT$ degrees of freedom. The work of Mathews and Pfafflin (1965) suggests that for white noise passed through a single tuned bandpass filter, $4z_n/N_0$ is approximately $\chi^2$ with $4WT$ degrees of freedom. For the surge case shown in Figure 8.4, $W = 0.72$ and $T = 16.3$, so $4WT = 47$. This is in reasonably good agreement with the value found. Because the shape of the pitch spectrum does not match that assumed in the derivation of the theory, it is not possible to make a similar comparison between simulation and theory for the pitch case.

An ROC curve was derived from each histogram by calculating $P(S|s)$ and $P(S|n)$ at 6 different points, and fitting these data with a Gaussian ROC curve. The ROC points and fitted curves are shown in Figures 8.6 and 8.7. The ROC slope parameters for the two
cases are 0.89 and 0.93. These values compare very well with the average ROC slope of 0.90 found for the experimental data.

Figure 8.8 shows the predicted variation in critical amplitude with noise condition. Also shown in Figure 8.8 are the mean experimental results. The selection of filter parameters mainly affects the shape of the right hand side of each curve, that is, the predicted values for noise conditions D, E, F and G as compared with that for condition C. There is an overall shift of about 3 dB between the experimental data and the simulation results. Neglecting the shift, however, the experimental data follow the general trend predicted by the model.

To demonstrate the important effect of the internal filter on the detector performance, it is useful to consider the case in which no internal filter is present. The detector performance for this case is plotted in Figure 8.9. Without the internal filter, noise components of all frequencies are equally effective at masking the signal. For surge, the predicted critical amplitudes in this case increase from condition D through condition G. It is clear from Figure 8.9 that this is inconsistent with the data. For pitch, the disagreement is not as obvious, because the power density of the narrow band spectra (expressed in terms of angular velocity) decreases with increasing frequency. However, the variance of \( \Delta A \) for this case is almost twice that found with the internal filter included. It would be worthwhile confirming these estimates of the internal bandwidth for pitch, by repeating the experiment using noise conditions for which the spectra are flat in terms of angular velocity.

As shown in Figure 8.8, the energy detection model predicts critical amplitudes which are, on average, about 3 dB lower than those found for the human observers. In other words, the humans are less sensitive than predicted by the model. Some possible reasons for this are discussed below. For the moment, it is sufficient to note that this shift should be roughly consistent for all signal conditions. This provides a way to test one of our assumptions about the form of the detectors. It was assumed that the surge detector calculates acceleration energy, while the pitch detector calculates the angular velocity energy. A simulation was run in which the pitch detector was assumed to be an angular acceleration detector. The results are shown in Figure 8.10. The sensitivity differences between experiment and model are no longer consistent. The experimental subjects were, on average, 3 dB less sensitive than this detector for signals at 0.6 Hz, but 4 dB more sensitive than the model for signals at 0.2 Hz. Also, the variance of \( \Delta A \) between
noise conditions, based on this model, is 4 to 5 times as high as that found for the data of Figure 8.8. Because this model does not help to collapse the data, it was rejected. The most consistent comparison between experiment and model is obtained if we assume that the pitch channel detects angular velocity. A similar check for surge supports the assumption of a linear acceleration detector.

In order to obtain a quantitative assessment of the model fit to the data, the experimental data were normalized by the predicted values shown in Figure 8.8. Each experimental result was normalized by the predicted value for the appropriate signal and noise condition. An analysis of variance was carried out on the normalized data. The results of this analysis are shown in Table 8.3. By comparing Table 8.3 with Table 6.5, it can be seen that the variance associated with noise condition in the normalized data is only 6% of that found in the raw data. In other words, the model accounts for 94% of the variance associated with noise condition. It also accounts for most of the variance associated with the noise × axis and noise × frequency interactions. Also noteworthy is the fact that the data for both axes collapse to approximately the same value. The mean value of ΔA is 3.9 dB for surge, and 2.4 dB for pitch. While the difference is fairly small, the analysis of variance showed it to be significant (α < 0.05). Altogether, the model accounts for over 90% of the variance associated with the main experimental variables, excluding all interactions with subject. The fit is not perfect; the main effect of noise condition is still significant. However, in view of the large reduction in overall variance, the model fit is judged to be good.

One final point remains to be discussed: the overall error in predicted sensitivity noted above. The mean difference between predicted and measured performance over all conditions is 3.1 dB. This is equivalent to a factor of 1.43 in signal amplitude. It is not surprising that the human observers do not perform as well as predicted by theory. The theory describes the performance of an optimal detector. At least two factors may contribute to the humans' sub-optimal performance: imperfect memory and inattentiveness.

Imperfect memory can be modelled approximately using the memory time constant T_m shown in Figure 8.3. A few exploratory simulation runs showed that reducing T_m from infinity (perfect memory) to 5 seconds increased the predicted critical amplitudes for the cases considered by about 2 dB. Reducing T_m further would increase the predicted values even more. Therefore, this factor could account for much of the difference in sensitivity between theory and experiment. Unfortunately, this study did not examine the
effect of signal duration on detectability. Therefore, we have no reliable way to estimate $T_m$.

The experimental data may also have been affected to some extent by inattention on the part the subjects. Consider the forced choice task. Suppose that a subject is inattentive on some fraction of the trials, $P_1$. On these trials, he guesses with $P(C)=0.5$. On the remainder of the trials, he detects the correct interval with $P(C) = P^*$. His total fraction of correct responses is therefore:

$$P_2(C) = 0.5 P_1 + P^* (1 - P_1) \quad (8.11)$$

The adaptive algorithm converges on the amplitude for which $P_2(C)$ is 0.76. Suppose that $P_1 = 0.20$. Then, from equation 8.11, $P^*$ is 0.825. Therefore, the target level estimated by the adaptive algorithm has a true probability of detection of 0.825 rather than 0.76. By interpolating from the psychometric curve, we find that the target level overestimates the critical amplitude in this case by 1.2 to 1.4 dB.

These factors are probably influenced by training and motivation. Earlier in this chapter we discussed the small but significant individual differences in performance, and noted that the most highly motivated subjects often achieved the highest performance. In experiment 2, for example, the most sensitive subject (the author) was 1.9 dB more sensitive than the mean. Averaged over all conditions, therefore, this subject achieved within 1.2 dB of the optimal performance predicted by the theory.

In closing this comparison, it is interesting to note that the estimates of the internal filter bandwidth made here are identical to those found using the simple and very approximate analysis in Chapter 6. The analysis of variance on the normalized data shows that the detailed model fits the data more closely than the simplified model of Chapter 6. However, for complex noise spectra, evaluation of the detailed model is difficult because no simple theoretical expressions are available. For engineering applications, the simple approach taken in Chapter 6 is probably sufficient.
8.5 Comparison with Previous Studies

The results of this study provide two main points for comparison with previous studies in the literature. These are: the absolute threshold levels, and the masking effectiveness of random motion and vibration.

8.5.1 Absolute Threshold Levels

Table 8.4 shows a comparison of the motion perception "thresholds" found in this study with those from previous studies. The critical amplitude is similar to some definitions of threshold used in the literature. The value quoted for each axis is the mean critical amplitude found with noise condition A. On the whole, the critical amplitudes found in this study are near the lower bound of the thresholds reported in the literature. The values found for translational acceleration (surge and sway) agree with the very low levels (2 milli-g) found by Walsh (1961).

The detection of pitch and roll in this study was aided by the oculogyral illusion (OGI). Therefore, the results of this study are compared with reported thresholds for the OGI. The critical amplitudes for pitch and roll with noise condition A are comparable to the thresholds for the OGI reported by Doty (1969) and Clark et al (1980).

Higher critical amplitudes were found for roll than for pitch. This difference was found to be statistically significant. Clark and Stewart (1970), on the other hand, found higher thresholds for pitch than roll, but concluded that the difference was not significant. There is no obvious explanation for this finding.

It was suggested in Section 8.4.2 that the results of this study could be extrapolated to the "zero-noise" case. These extrapolations are very approximate, but the results shown in Table 8.2 indicate that the critical amplitudes in the "zero-noise" case might be on the order of 0.5 to 0.7 times the values obtained for condition A. The resulting values would be among the lowest motion perception thresholds quoted in the literature.

8.5.2 Masking Effectiveness

The results of this study show unequivocally that background motion and vibration does mask sinusoidal signals. At first glance, this seems to contradict the findings of previous
8.17

studies by Coleman (1979), Bury (1980) and Clark et al (1980), which are discussed in Chapter 2. These studies found that the addition of 1 Hz, 2 Hz, or 5 Hz sinusoidal yaw vibration, or 2 to 5 Hz quasirandom yaw vibration had little effect on the detection of yaw angular acceleration steps or pulses.

However, the typical pulse duration was 0.7 to 2.5 seconds. Interpreting a pulse as half of a cycle, a significant portion of the energy of such pulses is in the frequency range 0.2 to 0.7 Hz. The current study found that sinusoidal motion is masked primarily by noise components near the signal frequency. In particular, pitch signals at 0.2 and 0.6 Hz were not effectively masked by noise above 1 Hz. Therefore, the earlier findings are consistent with this study.

We might expect that vibration at 1 Hz would be close enough in frequency to have some masking effect. Indeed, Bury’s data for detectability of step yaw accelerations in the presence of 1 Hz vibration show a decrease in sensitivity with increasing vibration intensity, although his analysis does not show this effect to be significant. It might be argued that a pure sine wave is so predictable that the observer can filter it completely, and that it therefore does not mask the signal.

8.6 Prediction of Performance for Different Conditions

This study has shown that human perception of motion can be modelled, at least approximately, by an energy detection model with an internal bandpass filter. The experimental data are limited, but using the model, it is possible to make some tentative predictions of detectability for signal and noise conditions other than those tested. However, care must be taken in applying the results to human performance in the cockpit, where conditions may differ significantly from those in the laboratory. Section 8.7 discusses some of the major factors which limit the applicability of the results to performance in the cockpit.

To predict signal detectability for signal and noise conditions other than those tested, there are three options available. First, if the noise spectrum is flat and the noise band spans the signal frequency, it may be possible to use the simplified formula for $d'_{A}$ given in equation 3.66, or similar results found for a single-tuned internal bandpass filter by Pfafflin and Mathews (1962). This method is not applicable when the noise condition is more complex. The second approach is to run a Monte Carlo simulation of the energy
detector, with the inputs selected to match the conditions under study. The final possibility is to use a semi-empirical model based on the energy detector and the results of Chapter 6.

All three methods require some assumptions about model parameters, and thus are subject to error. The magnitude of the estimation error is judged to be similar for all three methods. Therefore, it seems reasonable to use the semi-empirical model, which is the simplest of the three, for engineering purposes.

The foundation of the empirical model is the concept introduced in Chapter 6, that detectability depends on the ratio of signal power to the noise power in a "critical band" near the signal frequency. Factors were also included to describe how detectability varies as a function of the signal duration and the observation period. The model may be expressed as follows:

\[ d_A' = k \frac{P_s T_s}{P_N \sqrt{T}} \]  

(8.12)

where \( P_s \) is the average signal power,
\( P_N \) is the noise power in the "critical band,"
\( T_s \) is the signal duration (where \( T_s \leq T \)),
\( T \) is the observation period, and
\( k \) is an empirical constant.

The behaviour of this model is similar to that of the true energy detector. The signal energy is simply \( E_s = P_s T_s \). For a rectangular noise spectrum, \( P_N = N_o W \), and equation 8.12 may be rewritten as:

\[ d_A' = k \frac{E_s}{N_o W \sqrt{T}} \]  

(8.13)

With the exception of a factor of \( \sqrt{W} \), this expression is equivalent to equation 3.66 for \( WT >> E_s/N_o \). For non-rectangular spectra, the value of \( W \) is not well defined. Also, the empirical analysis presented in Chapter 6 showed that, for \( d_A' = 1 \), \( kP_s/P_N \) is nearly constant for the seven noise spectra used in experiment 2. Therefore, the simplified model is justified on an empirical basis.
For given values of $P_s$ and $P_N$, detectability depends on $T_s$, $T$ and $k$. In the laboratory, $T_s$ and $T$ are both defined by the temporal structure of the experiment. In a more realistic situation, the value of $T$ is not well defined. An experienced subject would probably adopt a sampling period on the same order as his estimate of $T_s$, say $T_s \leq T \leq 2T_s$. The maximum value of $T$ is limited by a human’s inability to integrate a signal over a prolonged period. No data are available to estimate the maximum value of $T$, but it is probably no more than about 20 seconds, and may be much less.

The value of $k$ may be estimated from the results of Chapter 6. For signals at 0.6 Hz, $T_s = 11.1$ seconds (10 second signal duration plus a third of the ramp time), $T = 16.3$ seconds, and $P_s/P_N = 1.2$ when $d_A' = 1.0$. Solving for $k$ yields the value $k = 0.30$.

In the most general case, we want to determine the detectability of a signal on axis $M$, in the presence of noise on all 6 degrees of freedom. In this case, the major inputs to the model are:

1. $s(t)$, the signal on axis $M$,
2. $r_j(t)$, $j=1,6$, the noise on axes 1 through 6,
3. $r_i(t)$, the internal noise.

The power spectra associated with $r_j(t)$ and $r_i(t)$ are denoted $\Phi_j(\omega)$ and $\Phi_i(\omega)$, respectively. The signal power is defined as:

$$P_s = \frac{1}{T_s} \int_0^{T_s} s^2(t) \, dt \quad (8.14)$$

and the noise power is defined as:

$$P_N = \int_0^\infty |H(j\omega)|^2 \Phi_{TT}(\omega) \, d\omega \quad (8.15)$$

where $H(s)$ is the transfer function of the internal filter:

$$H(s) = \frac{2\zeta\omega_0 s}{s^2 + 2\zeta\omega_0 s + \omega_0^2} \quad (8.16)$$
\( \Phi_{TT}(\omega) \) is the total noise power at frequency \( \omega \):

\[
\Phi_{TT}(\omega) = \Phi_{II}(\omega) + \sum_{j=1}^{6} \gamma_{Mj} \Phi_{ji}(\omega)
\]  

(8.17)

where \( \gamma_{Mj} \) is the inter-axis weighting factor which describes the "effective" noise power on axis \( M \) due to noise on axis \( j \). Estimates of \( \gamma_{ij} \) are provided in Table 7.4.

The internal noise, \( r_i(t) \), is small enough to be ignored in most cases. In the simulation runs described earlier, \( \Phi_{II}(\omega) \) was simply modelled as a broadband noise spectrum, flat from 0.25 to 4.0 Hz, with the power density given in Table 8.2. This is not an accurate model of the spectral shape, but it does serve to describe the limits of detectability for very low intensity external noise.

The remaining model parameters affecting performance are the bandwidth and centre frequency of the internal filter, \( H(s) \). If the signal frequency is known, then an experienced observer will adopt an internal bandpass filter centred at \( \omega_o = \omega_s \), the signal frequency. Only two estimates of bandwidth are available from this study. The results of experiment 2 indicate that, for a signal frequency of 0.2 Hz, \( \zeta = 1.2 \), and the bandwidth is 0.48 Hz. For a signal frequency of 0.6 Hz, \( \zeta = 0.6 \) and the bandwidth is 0.72 Hz. Therefore, the bandwidth of the internal filter tends to increase with frequency. However, the ratio of the bandwidth to the centre frequency appears to decrease with increasing frequency. These conclusions are fairly tenuous. Further experimentation is needed to improve our basis for estimating the internal filter bandwidth. If the signal frequency is unknown, or if the signal energy is spread over a wide frequency band, the internal filter bandwidth will probably be larger than the values found here.

After estimating values for \( \omega_o, \zeta, P_N, \) and \( T \), the signal detectability may then be estimated using equation 8.12. It must be emphasized that predictions based on this method are very approximate. Due to the number of assumptions and estimates which must be made, the potential for error is large. This method is probably most useful for comparing signal detectability in different conditions, rather than producing absolute predictions of performance.
8.7 Implications for Flight Simulation

It seems appropriate to conclude this discussion with a few remarks relating the results of this study to flight simulation and the design of motion systems. It is clear that what a pilot feels depends on the motion history of the simulation run. During periods of intense motion and vibration, a pilot may not detect specific cues or washout motions that he would feel under more quiescent conditions. Alternatively, he may simply interpret the "signals" (cues or washouts) as part of the background motion in the simulator. This raises the possibility of modifying the parameters of the motion-drive or cue-shaping algorithms on-line, in an attempt to provide the best cues possible for a given motion condition.

The potential for this is limited by the frequency dependence of the masking effects observed. Ideally, we would like to mask low-frequency, high-displacement washout motions with high-frequency (i.e. small displacement) vibration or buffet. This would yield the greatest gains in terms of actuator travel. However, this study showed that signal motion is masked mainly by noise in a frequency band near the signal frequency. High frequency buffet will not effectively mask low-frequency motion.

The current study is a significant step toward understanding the factors influencing human sensitivity to low-amplitude motion. However, the results were obtained in strictly controlled laboratory conditions which do not completely match those in the cockpit. As a result, the critical amplitudes or "thresholds" found here are very small, demonstrating that humans are highly sensitive to motion. The values reported here are much too small to be used as practical limits in motion-drive design. For example, the roll velocity limit commonly used in tilt-coordination algorithms is typically 2 or 3°/s, a factor of 4 to 6 larger than the roll "threshold" reported here.

Due to a combination of factors, the levels reported here are probably lower than motion perception thresholds for pilots flying the simulator. The first of these factors is the predictability of signal and noise. The use of a sinusoid of known frequency probably aids detection. In terms of the signal detection model, knowing the signal frequency allows the observer to set up a narrow internal filter which screens out noise inputs with frequencies far from the signal frequency. If the signal frequency is unknown, all input must be considered. Therefore, the effective noise level is higher and, for a given signal energy, detectability is reduced. Green and Swets (1974, Chapter 10) discuss this issue.
with respect to auditory detection, and note that the threshold shift attributable to uncertain frequency is typically a few dB.

Another factor influencing the results is the structure of the experiment. The experimental procedure used in this study had a very strict temporal structure. The observer monitored the motion over a well-defined interval, and attempted to ascertain the presence or absence of a signal in that interval. By contrast, in the cockpit we want to know if the pilot will detect a cue for which the onset time is very uncertain, such as an engine failure or wind gust. The available evidence indicates that signal detectability decreases as temporal uncertainty increases. Therefore, the effective sensitivity of the pilot is lower than that of an observer in a controlled experiment. With reference to the signal detection model, integrating the noise over a longer time period increases $\sigma_n$, thereby reducing the relative strength of the signal. Again, Green and Swets (1974, Chapter 9) discuss this with reference to auditory detection.

The final two factors, which may be the most important, are the visual field and the pilot workload. As discussed in Chapter 2, both affect the perception of motion and may be expected to raise the critical amplitude for any given noise condition. It may be possible to incorporate these effects into the signal detection model. Visual inputs could be modelled either as attenuating the vestibular signals or increasing the internal noise. The effect of pilot workload might be represented by an attention-sharing model. Increasing the workload would reduce the attention directed toward motion perception, thereby reducing the effective observation time and reducing sensitivity. These suggestions are hypothetical and have not yet been tested.

Clearly, great care must be taken in applying the results of this study to flight simulation. Further work is still required to determine motion detection performance in more realistic environments.
Table 8.1  Mean critical amplitudes for surge (milli-g) and pitch (°/s) as a function of noise bandwidth. Broadband noise is condition C in all cases. Narrow band noise is:
condition D when signal frequency is 0.2 Hz,
condition E when signal frequency is 0.6 Hz.

<table>
<thead>
<tr>
<th>Noise Case</th>
<th>Signal Axis</th>
<th>Freq. (Hz)</th>
<th>Broad band</th>
<th>Narrow band</th>
<th>Ratio (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Surge</td>
<td>0.2</td>
<td>7.60</td>
<td>5.24</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.6</td>
<td>10.78</td>
<td>9.11</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Pitch</td>
<td>0.2</td>
<td>1.19</td>
<td>0.95</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.6</td>
<td>1.53</td>
<td>1.05</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td>2.5</td>
</tr>
</tbody>
</table>
Table 8.2  Estimates of the internal noise power density, $\Phi_\Pi$, and estimates of the critical amplitude for surge (milli-g) and pitch ($^\circ$/s) for noise condition A and for the "zero-noise" condition.

Note: the value quoted for $\Phi_\Pi$ is an artificial quantity. It is the estimated power density of the internal noise at 1.0 Hz, assuming that the internal noise is broadband with a spectral shape similar to that for noise condition C. The units of $\Phi_\Pi$ are (m/s$^2$)$^2$/Hz for surge, and (rad/s$^2$)$^2$/Hz for pitch.

<table>
<thead>
<tr>
<th>Signal Axis</th>
<th>Freq. (Hz)</th>
<th>$\Phi_\Pi$</th>
<th>$A_{cr}(A)$</th>
<th>$A_o$</th>
<th>$A_{cr}(A)$</th>
<th>$A_o$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surge</td>
<td>0.2</td>
<td>$3.3 \times 10^{-4}$</td>
<td>2.7</td>
<td>1.9</td>
<td>1.62</td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>$7.5 \times 10^{-5}$</td>
<td>2.8</td>
<td>1.3</td>
<td>1.68</td>
<td>0.78</td>
</tr>
<tr>
<td>Pitch</td>
<td>0.2</td>
<td>$2.5 \times 10^{-4}$</td>
<td>0.34</td>
<td>0.25</td>
<td>0.84</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>$5.0 \times 10^{-5}$</td>
<td>0.31</td>
<td>0.14</td>
<td>0.76</td>
<td>0.35</td>
</tr>
</tbody>
</table>
### Table 8.3: Results of the analysis of variance of $\Delta A$, where

\[
\Delta A = 20 \log \left( \frac{A_c^{\text{experiment}}}{A_c^{\text{simulation}}} \right)
\]

<table>
<thead>
<tr>
<th>Effect</th>
<th>Degrees of Freedom</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Ratio</th>
<th>P($x&gt;F$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise</td>
<td>6</td>
<td>111.5</td>
<td>18.58</td>
<td>3.77</td>
<td>0.0087</td>
</tr>
<tr>
<td>Axis</td>
<td>1</td>
<td>77.7</td>
<td>77.70</td>
<td>8.34</td>
<td>0.0446</td>
</tr>
<tr>
<td>Noise x Axis</td>
<td>6</td>
<td>17.0</td>
<td>2.83</td>
<td>0.70</td>
<td>0.6520</td>
</tr>
<tr>
<td>Frequency</td>
<td>1</td>
<td>0.056</td>
<td>0.056</td>
<td>0.005</td>
<td>0.9496</td>
</tr>
<tr>
<td>Noise x Freq</td>
<td>6</td>
<td>32.3</td>
<td>5.38</td>
<td>1.15</td>
<td>0.3627</td>
</tr>
<tr>
<td>Axis x Freq</td>
<td>1</td>
<td>11.0</td>
<td>10.97</td>
<td>2.81</td>
<td>0.1069</td>
</tr>
<tr>
<td>N x A x F</td>
<td>6</td>
<td>9.5</td>
<td>1.58</td>
<td>0.40</td>
<td>0.8691</td>
</tr>
<tr>
<td>Subject</td>
<td>4</td>
<td>250.2</td>
<td>62.55</td>
<td>15.99</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Noise x Subj</td>
<td>24</td>
<td>118.3</td>
<td>4.93</td>
<td>1.26</td>
<td>0.2876</td>
</tr>
<tr>
<td>Axis x Subj</td>
<td>4</td>
<td>37.2</td>
<td>9.31</td>
<td>2.38</td>
<td>0.0799</td>
</tr>
<tr>
<td>N x A x S</td>
<td>24</td>
<td>97.0</td>
<td>4.04</td>
<td>1.03</td>
<td>0.4679</td>
</tr>
<tr>
<td>Freq x Subj</td>
<td>4</td>
<td>49.7</td>
<td>12.43</td>
<td>3.18</td>
<td>0.0314</td>
</tr>
<tr>
<td>N x F x S</td>
<td>24</td>
<td>111.9</td>
<td>4.66</td>
<td>1.19</td>
<td>0.3350</td>
</tr>
<tr>
<td>A x F x S</td>
<td>4</td>
<td>11.0</td>
<td>2.75</td>
<td>0.70</td>
<td>0.5980</td>
</tr>
<tr>
<td>N x A x F x S</td>
<td>24</td>
<td>93.9</td>
<td>3.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>139</strong></td>
<td><strong>1028.3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 8.4  Comparison of the critical amplitudes found in this study (for noise condition A) with thresholds reported in the literature. All data for translational motion are in milli-g; those for angular motion are in degrees/sec.

<table>
<thead>
<tr>
<th>Signal Axis</th>
<th>$A_{\alpha}(A)$</th>
<th>Thresholds from previous studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surge, Sway</td>
<td>2.6 - 3.0</td>
<td>3.0 - 9.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sinusoidal heave (1)</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>sinusoidal surge (2)</td>
</tr>
<tr>
<td></td>
<td>3.0 - 20.0</td>
<td>sinusoidal surge (3,4)</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>step acceleration (5)</td>
</tr>
<tr>
<td>Pitch</td>
<td>0.29 - 0.34</td>
<td>0.22 - 0.53</td>
</tr>
<tr>
<td>Roll</td>
<td>0.44</td>
<td>0.22 - 0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sinusoidal pitch (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sinusoidal roll (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>step yaw acceleration (6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td>step yaw acceleration (7)</td>
</tr>
</tbody>
</table>

Sources:

5. Young and Meiry (1968).
Figure 8.1  Signal-to-noise ratios for surge signals at 0.6 Hz (noise condition C) compared with theoretical psychometric curves based on the three detection models.
Figure 8.2  Variation of detectability with signal amplitude for the data from experiment 2, compared with the theoretical psychometric curve based on the energy detector with WT = 12. Note that the data have been shifted 4 dB to the left to coincide with the theoretical curve.
Figure 8.3  Block diagram of the energy detector simulation.
Figure 8.4 Typical condition probability density functions $f(e|n)$ and $f(e|s)$ for surge signals at 0.6 Hz, with spectrum C, based on simulation of 2000 trials. Central and non-central $\chi^2$ distributions with 55 degrees of freedom are shown for comparison with the simulation results.
Figure 8.5  Typical condition probability density functions $f(e \mid n)$ and $f(e \mid s)$ for pitch signals at 0.6 Hz, with spectrum C, based on simulation of 2000 trials. Central and non-central $\chi^2$ distributions with 26 degrees of freedom are shown for comparison with the simulation results.
Figure 8.6  ROC curve for surge, based on the conditional distributions shown in Figure 8.4.
Figure 8.7  ROC curve for pitch, based on the conditional distributions shown in Figure 8.5.
Figure 8.8  Variation in critical amplitude with noise condition. Each experimental point shown is the mean over all subjects for the appropriate axis, frequency, and noise condition. Simulation results are based on the internal filter parameters described in the text.
Figure 8.9  Variation in critical amplitude with noise condition. Each experimental point shown is the mean over all subjects for the appropriate axis, frequency, and noise condition. The simulation results shown were obtained with the internal filter removed.
Figure 8.10 Variation in critical amplitude with noise condition. Each experimental point shown is the mean over all subjects for the appropriate axis, frequency, and noise condition. The simulation results shown were calculated assuming the pitch detector to be an angular acceleration detector.
Chapter 9

Summary and Conclusions

9.1 Summary

This study investigated human sensitivity to low-amplitude sinusoidal motion cues superimposed on a background of random motion or vibration. The general purpose of the study was to help define the minimum perceptible levels of motion under various conditions in a flight simulator, for application to the design of motion-drive algorithms. A literature review showed that very few data are available on this topic.

Two specific aims were set. The first was to collect sensitivity data in a variety of vibration conditions representative of conditions in a flight simulator. The second was to formulate a model which would improve our understanding of the motion perception process, and facilitate extrapolation of the results to other conditions. It was proposed that, for the conditions studied, motion sensation could be represented by a signal-in-noise model and analyzed using signal detection theory. Three ideal detector models were borrowed from the literature on auditory detection.

Three experiments were run to test the applicability of these models to motion perception. In the first experiment, ROC curves were obtained for 5 subjects detecting sinusoidal pitch and surge signals in a background of low-amplitude broadband random motion (noise). The second experiment tested the detectability of sinusoidal pitch and surge signals with several different noise spectra. The effects of both noise power and frequency content were investigated. In this experiment, noise and signal motion were on the same axis. The final experiment tested the masking effect of noise on a different axis than the signal. All interactions of surge, sway, pitch and roll were studied.

The experimental results were compared with predictions of the three ideal detector models. Of these, an energy detector gave the best fit. For this initial comparison, the energy detector performance was calculated using equations which are based on several simplifying assumptions. To make a more valid comparison, a computer simulation of
the energy detector was set up which more closely reflected the experimental conditions. The simulation was used as a basis for model parameter estimates.

The major findings of the study are summarized in the next section.

9.2 Conclusions and Contributions

The work presented in this report contributes to three distinct fields of knowledge: motion perception, psychophysical methodology, and simulation engineering.

The primary focus of this study was motion perception, and the major contributions are in this field. The major findings of the study are as follows:

1. The average critical amplitudes or "thresholds" for perception of sinusoidal motion signals in the presence of low intensity broadband random motion are as follows:

   - surge: 3.0 milli-g
   - sway: 3.0 milli-g
   - pitch: 0.29 °/sec
   - roll: 0.44 °/sec

   These are near the lower bound of the range of values quoted in the literature. Extrapolation of these results to the "zero-noise" case yields values which are 30 to 50% lower.

2. The ROC curves obtained in experiment 1 show that signal detection theory is more appropriate than the traditional "high threshold" models. Gaussian ROC curves with a slope of 0.9 provide a satisfactory fit to the data.

3. Data from the second experiment show that:
   a) Detection performance may be expressed as a function of the signal-to-noise ratio. Random motion does mask a sinusoidal signal.
   b) Sinusoidal motion is masked primarily by noise components near the signal frequency. This supports the idea of an internal filter or "critical band."
c) The psychometric curve is very steep, rising from near chance to near perfect performance over a 10 dB rise in signal amplitude.

4. Data from the third experiment show that random motion on one axis can mask a signal on another axis. The masking effectiveness depends on the degree of similarity between the signal and noise motion (i.e. axes). Inter-axis masking is generally much less effective than same-axis masking.

5. The data are limited, but give good qualitative agreement with an energy detection model. For a given noise condition, human subjects are, on average, about 3 dB less sensitive than predicted by the model. Several reasons for this disparity are discussed.

This study is a step toward a model which would allow prediction of motion cue detectability in arbitrary conditions. The results improve our understanding of the motion perception system, and have applications to flight simulation and prediction of pilot performance.

In the area of psychophysical methodology, this study led to several developments that will be of interest to researchers studying human perceptual performance. The 4 main points are as follows:

1. A least squares adaptive algorithm for psychophysical testing was developed independently for this study. A subsequent review showed this to be similar to maximum likelihood algorithms described in the literature. The least squares approach seems somewhat less sensitive than the maximum likelihood algorithm to lapses in the subject’s attention. A simple algorithm suitable for computer implementation is presented.

2. The idea of using a payoff/penalty matrix to influence an observer’s response criterion was extended to the rating task.

3. Maximum likelihood procedures for fitting ROC curves to rating task data (Dorfman and Alf, 1969) were extended to the following cases:
   a) Gaussian ROC curves with a predetermined slope
   b) Rayleigh distribution ROC curves based on the envelope detector.
4. The approximate formula given by Green and Swets (1974) to describe the shape of energy detection ROC curves was found to be significantly in error. A more accurate approximation was proposed.

Finally, a new simulator motion-drive algorithm was developed in the course of this study. This algorithm can provide essentially uncoupled pitch and roll motion, that is, rotations about a horizontal axis which do not generate any change in specific force which might stimulate the otolith receptors. The primary application of such an algorithm is in studies of vestibular function. Also, an open loop filtering system was developed to reduce simulator crosstalk.

9.3 Directions for Future Study

There are many promising avenues for future research in this area. These fall under two main headings: basic studies in motion perception, and application of the detection model to motion-drive design.

Further perception studies are needed to flesh out the signal detection model proposed in this report. The results of this study are based on a small subject pool and cover a limited range of conditions. Future studies could examine in more detail the effects of such variables as signal frequency and noise condition, as well as the effect of parameters not tested explicitly in this study, such as signal duration.

The next step is to abandon sinusoidal stimuli in favour of transient waveforms more similar to flight simulator washout motions. The best detection model for this case would probably be a hybrid, somewhere between the energy detection model used here and the statistical cue model formulated by Ormsby (1974). Initial work with transient waveforms should follow a structured experimental procedure such as the forced choice procedure described in Chapter 6. Later experiments could use a more realistic procedure in which the waveform shape and onset time is less predictable.

Finally, it would be very useful to study the effect of visual cues and pilot workload on motion sensitivity, in various noise conditions. The visual field, in particular, can have a profound influence on subjective sensation of motion. A comprehensive motion detection model must eventually include the effect of the visual field. A recent study by Reid et al (1987) showed that, in certain cases, visual display motion could reduce
sensitivity to congruent physical motion by a factor of up to 3. The obvious next step is to study the case in which the visual and motion cues are contradictory. This remains a very rich field for future study.

On the practical side, there is also potential for further work in applying the signal detection model to washout filter design and evaluation. The eventual goal is an algorithm which could modify the washout filter parameters on-line, based on current motion levels in the simulator, to maximize the detectability of congruent cues and minimize the detectability of spurious cues and washout motions.

The first step is to develop a probabilistic approach to evaluating washout performance. This would involve setting up a subroutine, called by the motion-drive software, which would calculate the probability that a pilot "felt" specific motions, that is, discriminated particular motion cues from the background motion. The probabilities calculated would depend on the current motion level in the simulator. A scoring routine would add points for congruent cues detected, and dock points for congruent cues missed and spurious cues detected. Run off-line, the routine could be used to evaluate motion-drive performance. The pilot evaluations of 10 motion-drive configurations reported by Reid and Nahon (1986) would provide a useful check on the results.
Appendix A

A Motion Drive Algorithm for Uncoupled Motion

A.1 Introduction

This appendix describes the algorithm used to drive the flight simulator motion base for these experiments. The aim in developing this algorithm was to generate uncoupled motion: commanded motion on one axis should not result in a motion cue on a different axis. Specifically, we wanted to ensure that rotation about a horizontal axis does not cause any change in specific force which might be detected by the otoliths. Simple rotation about a fixed horizontal axis through the pilot’s head results in a change in specific force due to the rotation of the gravity vector relative to the pilot’s body frame.

A.2 Reference Frames

For this analysis it is convenient to use two reference frames. The first is a body frame, $F_H$, illustrated in Figure 2.1. It is attached to the pilot’s (or subject’s) head with its origin midway between his right and left vestibular labyrinths. The $X$ axis points forward, the $Y$ axis points to the right, and the $Z$ axis points downward along the spine.

The second frame used is an inertial frame, $F_I$, fixed to the earth. The $Z$ axis is aligned with the gravity vector. It is convenient to define a reference condition in which $F_H$ is stationary and aligned with $F_I$. Let $S_{lo}$ denote the initial location of the origin of $F_H$ in frame $F_I$.

A.3 Notation

It is now necessary to define some terms. The notation used here is similar to that used by Etkin (1972).

Let $\omega = [p \ q \ r]^T$ be the angular velocity of $F_H$ relative to $F_I$, expressed in $F_H$ components.
Let $\beta = [\phi \ \theta \ \Psi]^T$ be the Euler angles which define the attitude of $F_H$ relative to $F_I$.

Let $S_I$ be the position of $F_H$ relative to $F_I$, expressed in $F_I$ components. That is, $S_I$ is the vector from the origin of $F_I$ to the origin of $F_H$.

Let $a = [a_x \ a_y \ a_z]^T$ be the acceleration of $F_H$ relative to $F_I$. This will be subscripted as $a_H$ or $a_I$, depending on the frame of reference in which the components are expressed.

Let $g = [0 \ 0 \ g]^T$. The gravity vector written in $F_I$ components is always equal to $g$ and will not be subscripted.

Let $g_{Hi}$ be the gravity vector expressed in $F_H$ components.

Let $f_{Hi} = a_H - g_{Hi}$ be the specific force in $F_H$ components, and let $f_{I} = a_I - g$ be the specific force in $F_I$ components.

Let $\Delta f_{Hi} = f_{Hi} + g$ be the change in specific force (in $F_H$ components) from that in the reference condition (where $f_{Hi} = -g$).

Finally, let $L_{Hi}$ be the matrix for transforming vectors from the frame $F_H$ to $F_I$. This is defined by Etkin (1972) as:

$$L_{Hi} = \begin{bmatrix}
\cos \theta \cos \Psi & \sin \phi \sin \theta \cos \Psi - \cos \phi \sin \Psi & \cos \phi \sin \theta \cos \Psi + \sin \phi \sin \Psi \\
\cos \theta \sin \Psi & \sin \phi \sin \theta \sin \Psi + \cos \phi \cos \Psi & \cos \phi \sin \theta \sin \Psi - \sin \phi \cos \Psi \\
-sin \theta & \sin \phi \cos \theta & \cos \phi \cos \theta 
\end{bmatrix}$$  \hspace{1cm} (A.1)

This is called the transformation matrix. Note that the inverse transformation matrix is $L_{HI} = L_{Hi}^{-1} = L_{Hi}^T$.

A.4 Actuator Commands

The motion base is driven by specifying the length and second derivative of length for each hydraulic actuator on each time step. The actuator lengths can be determined from the position and attitude of the simulator in the inertial frame, $F_I$. We assume that the pilot's head is fixed with respect to the simulator, so $F_H$ serves as a simulator-fixed body.
frame. The length of the i-th actuator is given by:

\[ l_i = \sqrt{\mathbf{1}_i^T \mathbf{1}_i} \]  (A.2)

where \( \mathbf{1}_i = \mathbf{S}_i + \mathbf{L}_{ih} \mathbf{A}_{ih} - \mathbf{B}_{il} \)  (A.3)

In equation A.3, \( \mathbf{S}_i \) and \( \mathbf{L}_{ih} \) are as defined above, \( \mathbf{B}_{il} \) is the location of the lower pivot of the i-th actuator in \( \mathbf{F}_i \) components, and \( \mathbf{A}_{ih} \) is the location of the upper pivot of the i-th actuator in \( \mathbf{F}_h \) components. Note that \( \mathbf{B}_{il} \) and \( \mathbf{A}_{ih} \) are constant.

The lengths of the six actuators were found using equations A.2 and A.3 for \( i = 1 \) to 6. The second derivative of length for the i-th actuator was calculated by numerically differentiating \( l_i \) twice, using backward differences. The actuator commands were updated every 40 milliseconds.

### A.5 Determination of Motion in \( \mathbf{F}_i \)

#### A.5.1 Unfiltered Motion

Because the pilot's motion sensors are fixed in \( \mathbf{F}_h \), it is natural to define the commanded simulator motion in terms of \( \mathbf{F}_h \). The command motion is completely specified by two vectors: \( \omega \), the angular velocity, and \( \Delta f_h \), the change in specific force.

The motion of the simulator in the inertial frame may be found using the transformations described by Etkin (1972). The rate of change of the Euler angles is:

\[ \dot{\beta} = T \omega \]  (A.4)

where \( T = \begin{bmatrix} 1 & \sin \phi \tan \theta & \cos \phi \tan \theta \\ 0 & \cos \phi & -\sin \phi \\ 0 & \sin \phi \sec \theta & \cos \phi \sec \theta \end{bmatrix} \)  (A.5)

The Euler angles are found by integrating \( \dot{\beta} \):

\[ \beta(t) = \beta_0 + \int_0^t \dot{\beta}(\tau) \, d\tau \]  (A.6)
where by our assumptions above, \( \beta_0 = 0 \). \( L_{III}(t) \) is determined from \( \beta(t) \) using equation A.1. The inertial acceleration of \( F_H \) in \( F_I \) components is given by:

\[
a_I = f_I + g \\
= L_{III} f_H + g \\
= L_{III} \Delta f_H + [I - L_{III}] g
\]  

(A.7)

For the case where \( \Delta f_H = 0 \), the inertial acceleration of the simulator is sufficient to cancel the change in specific force caused by rotation of the gravity vector in \( F_H \). The position of \( F_H \) relative to \( F_I \) is obtained by integrating twice:

\[
S_I = S_{I_0} + \int_0^t \int_0^\tau a_I(v) \, dv \, d\tau
\]  

(A.8)

where we have assumed that \( \dot{S}_I(0) = 0 \).

A.5.2 Filtered Motion: Usual Solution

The simulator motion described by equations A.4 to A.8 will result in completely uncoupled motion cues. Unfortunately, the resulting motion cannot generally be realized. From equation A.7, the inertial acceleration needed to compensate for the rotation of the gravity vector is \( [I - L_{III}] g \). The vertical component of this is \( g(1 - \cos \phi \cos \theta) \), which is greater than zero except when \( \phi = \theta = 0 \). Therefore, the simulator will accelerate downward until it encounters the displacement limits.

If \( \phi(t) \) and \( \theta(t) \) are known sinusoidal motions, then the simulator drift can be eliminated by specifying a small constant vertical component of \( \Delta f_H \) which is just sufficient to cancel the constant component of \( [I - L_{III}] g \). For example, suppose that \( \phi(t) = \phi_0 \cos \omega_p t \), and \( \theta(t) = \theta_0 \cos \omega_q t \), and that \( \phi_0 \) and \( \theta_0 \) are both much less than 1. It may be shown using small angle approximations and some algebra that the constant vertical component of \( \Delta f_H \) required to eliminate drift is approximately \( -(\phi_0^2 + \theta_0^2)g/4 \). In the current study, however, there was always some random angular motion present. As a result, it was not possible to calculate the required drift correction to \( \Delta f_H \). Therefore, a different solution was sought.
The simulator drift can also be controlled by filtering the commanded acceleration. Given an oscillatory roll or pitch, the resulting vertical component of $a_t$ can be represented by a constant component plus an oscillatory component at twice the frequency of the angular motion. Application of the final value theorem shows that, to obtain a steady-state displacement of zero with a constant acceleration input, we must filter the motion using at least a third order high-pass filter.

In theory, it should only be necessary to filter the vertical component of $a_t$. In practice, however, it was necessary to filter all three components of the translational motion in order to keep the simulator within its displacement limits. The transfer function of the translational washout filter used in this study is:

$$H(s) = \frac{s^3}{(s + \omega_o)(s^2 + 2\zeta\omega_1 s + \omega_1^2)} \quad (A.9)$$

where $\omega_o = \omega_1 = 0.5$ rad/sec and $\zeta = 0.475$. A Bode plot of the filter response is shown in Figure A.1.

The phase shift introduced by this filter results in coupling between the rotational and translational motion. This may be demonstrated as follows. Let $\Delta f_{H,C}$ be the commanded change in specific force. From equation A.7, the "commanded" inertial acceleration in $F_1$ is:

$$a_{l,C} = L_{H}\Delta f_{H,C} + [1 - L_{H}]g \quad (A.10)$$

Assume that the same high-pass filter is applied to all three components of $a_{l,C}$. Then from the theory of linear differential equations, the filtered inertial acceleration is:

$$a_l(t) = \int_0^t h(t - \tau) a_{l,C}(\tau) d\tau \quad (A.11)$$
where \( h(t) \) is the impulse response of the filter. The actual change in specific force obtained is:

\[
\Delta f_H = f_H + g \\
= L_{HI} [a_I - g] + g \\
= L_{HI} a_I + [I - L_{HI}]g 
\]  
(A.12)

Combining equations A.10, A.11 and A.12 and expanding yields:

\[
\Delta f_H(t) = \int_0^t h(t-%20\tau) L_{HI}(t) L_{II}(%20\tau) \Delta f_{H,C}(\tau) \, d\tau + \Delta f_{H,g} 
\] 
(A.13)

where

\[
\Delta f_{H,g}(t) = \int_0^t h(t-%20\tau) [I - L_{HI}(\tau)]g \, d\tau + [I - L_{HI}(t)]g 
\] 
(A.14)

For small angles \( L_{HI}(t) L_{II}(\tau) = I \), so the first term of equation A.13 is a filtered version of \( \Delta f_{H,C} \). The second term, \( \Delta f_{H,g} \), is the change in specific force associated with rotation. For uncoupled motion, \( \Delta f_{H,g} \) must be zero. If \( \phi, \theta \) and \( \Psi \ll 1 \), we can use a small angle approximation for \( L_{IH} \):

\[
L_{IH} = \begin{bmatrix}
1 & -\Psi & \theta \\
\Psi & 1 & -\phi \\
-\theta & \phi & 1 
\end{bmatrix} 
\] 
(A.15)

and \( L_{IH} = L_{IH}^T \). Using this approximation, it follows that to the first order in the Euler angles, the last term of equation A.14 is zero. Combining equations A.14 and A.15, setting second order terms to zero, and taking the Laplace transform yields:

\[
\mathcal{L}\{\Delta f_{H,g,s}\} = \begin{bmatrix}
g[1 - H(s)]\theta(s) \\
-g[1 - H(s)]\phi(s) \\
0
\end{bmatrix} 
\] 
(A.16)
The gain of $H(s)$ is approximately 1 in the frequency range of interest; however, as shown in Figure A.1, the phase of $H(s)$ is generally not zero. Therefore $|1 - H(s)| > 0$, and a rotation does cause a change in specific force. Because of the phase shift introduced by the filter, the simulator rotation and the compensatory acceleration are no longer in phase, so the resultant specific force is not zero. For oscillatory pitch or roll at 0.5 Hz with the filter specified in equation A.9, the phase of $H(s)$ is $18^\circ$ and $|1 - H(s)| = 0.3$.

A.5.3 Filtered Motion: New Solution

The structure of the motion drive algorithm was altered slightly to correct this problem. In the new algorithm, the translations and rotations are both determined by applying the same high-pass filter to the "commanded" motion. As a result, both the rotations and the inertial accelerations are affected by the same phase shift, so the resultant change in specific force due to rotation is zero. A block diagram of the new algorithm is shown in Figure A.2.

The changes to the algorithm may be summarized in a few lines. Let $\omega_c$ be the commanded angular velocity, and let $\Delta f_{HC}$ be the commanded change in specific force. Two sets of Euler angles are carried in the calculations: $\beta_C$, the "command" Euler angles, and $\beta$, the actual (filtered) Euler angles. $\beta_C$ is calculated from $\omega_c$ following equations A.4 and A.6. The filtered angles are determined by:

$$\beta(t) = \int_{0}^{t} h(t-\tau) \beta_C(\tau) d\tau$$  \hspace{1cm} (A.17)

Two transformation matrices are calculated: $L_{IH,C}$ is based on $\beta_C$, and $L_{IH}$ is based on $\beta$. The angular rate transformation matrix $T$ (equation A.5) is calculated based on $\beta_C$.

The "commanded" inertial acceleration in $F_I$ is now:

$$a_{IC} = L_{IH} \Delta f_{HC} + [I - L_{IH,C}]g$$  \hspace{1cm} (A.18)

Note the slight change from equation A.10. The filtered inertial acceleration, $a_f$, is then found using equation A.11. Following the same steps outlined in Section A.5.2, we can
evaluate the actual change in specific force in $F_H$. From equation A.12, the actual change in specific force is:

$$\Delta f_H = L_{IH} a_I + [I - L_{IH}]g$$

(A.19)

Combining equations A.11, A.18 and A.19 and expanding yields:

$$\Delta f_H(t) = \int_0^t h(t-\tau) L_{IH}(t) L_{IH}(\tau) \Delta f_{H,C}(\tau) d\tau + \Delta f_{H,s}$$

(A.20)

where in this case,

$$\Delta f_{H,s}(t) = \int_0^t h(t-\tau) [I - L_{IH,C}(\tau)]g d\tau + [I - L_{IH}(t)]g$$

$$- \int_0^t h(t-\tau) [I - L_{IH}(t)] [I - L_{IH,C}(\tau)]g d\tau$$

(A.21)

Using the small angle approximation for $L_{IH}$ given in equation A.15, the last term of equation A.21 is second order in the Euler angles and can be neglected. Furthermore, using the small angle approximation for $L_{IH}$ and $L_{IH,C}$, it follows that:

$$\int_0^t h(t-\tau) [I - L_{IH,C}(\tau)]g d\tau = [I - L_{IH}(t)]g$$

(A.22)

Therefore,

$$\Delta f_{H,s}(t) = [I - L_{IH}(t)]g + [I - L_{IH}(t)]g$$

$$= 0$$

(A.23)

Therefore, to first order in the Euler angles, the new algorithm eliminates the coupling between pitch and roll rotations and specific force. Note that in the case where $\Delta f_{H,C} = 0$, the vertical component of the actual specific force, $\Delta f_H$, will be non-zero, but it will be second order in the Euler angles. In the current study, $\Delta f_{H,C}$ was set equal to the desired translational stimulus motion (noise alone or signal plus noise). In this case, the error in $\Delta f_H$ due to concurrent pitch and roll is second order in the Euler angles.
A.6 Further Reduction of Coupling

The algorithm described in Section A.5.3 was implemented on the UTIAS Flight Research Simulator. Numerical checks of the algorithm showed that, for the motion conditions used in this study, coupling between all axes was minimal. However, measurements of the simulator motion (see Appendix B) did reveal some coupling, or crosstalk. This crosstalk was attributed to small asymmetries in the simulator geometry and in the analog components of the motion-base control hardware. The crosstalk was small enough to be of little concern for typical flight simulation applications, but it seemed prudent to minimize it for the current study.

Two possible crosstalk reduction strategies were considered: one open-loop, and the other closed-loop. In the closed-loop method, the motion base was instrumented with a 6-channel motion sensing package. The recorded motions were fed back to the motion-drive program in a discrete-time implementation of an analog feedback loop. This method appeared promising for two reasons. First, it would reduce nonlinear as well as linear coupling effects. Second, its performance should be independent of the open-loop crosstalk transfer functions, which tend to change slightly whenever the motion-base hardware is tuned. However, it was not possible to obtain a satisfactory motion bandwidth except by sampling and updating the motion commands at 100 Hz. Updating at this rate overloaded the motion-drive computer. Therefore, this method was rejected as impractical and the open-loop method was used.

In the open-loop method, some coupling was introduced in the motion cue commands prior to sending them to the motion drive algorithm. The gain and phase of the coupling terms was chosen to just cancel the crosstalk. For example, the X-component of the measured specific force was found to have the following form (expressed in the Laplace domain):

$$
\Delta f_X(s) = H_{xx}(s)\Delta f_{xc}(s) + H_{xy}(s)\Delta f_{yc}(s) \\
+ H_{xp}(s)p_c(s) + H_{xq}(s)q_c(s)
$$

(A.24)

where $\Delta f_{xc}$ and $\Delta f_{yc}$ are the commanded changes in specific force, $p_c$ and $q_c$ are the commanded angular rates, and the $H_{ik}$'s are the crosstalk transfer functions. The X-
component of the commanded specific force was altered to the following:

\[
\Delta f_{X,*}(s) = \Delta f_{X,C}(s) + G_{XY}(s)\Delta f_{Y,C}(s) + G_{XP}(s)p_c(s) + G_{XQ}(s)q_c(s)
\] (A.25)

Replacing \(\Delta f_{X,C}(s)\) in equation A.24 with \(\Delta f_{X,*}(s)\) yields:

\[
\Delta f_X(s) = H_{XX}\Delta f_{X,C}(s) + [H_{XX}G_{XY} + H_{XY}]\Delta f_{Y,C}(s) + [H_{XX}G_{XP} + H_{XP}]p_c(s) + [H_{XX}G_{XQ} + H_{XQ}]q_c(s)
\] (A.26)

where we have dropped the arguments \((s)\) for clarity. Other components of the commanded motion cues can be altered in the same way. To eliminate crosstalk, each of the crosstalk reduction filters, \(G_{ik}(s)\), must be chosen such that:

\[
G_{ik}(s) = -\frac{H_{ik}(s)}{H_{ii}(s)}
\] (A.27)

The open-loop transfer functions were obtained by measuring the motion base response to a single axis sinusoidal input at a range of frequencies between 0.1 and 2.0 Hz. The transfer function \(H_{ik}(j\omega)\) is defined:

\[
H_{ik}(j\omega) = \frac{Z_i(\omega)}{X_k(\omega)}
\] (A.28)

where \(Z_i(\omega)\) is the Fourier transform of the output signal on the \(i\)-th axis, and \(X_k(\omega)\) is the Fourier transform of the input signal (on the \(k\)-th axis).

In practice it was not possible to satisfy equation A.27 exactly. Instead, we chose physically realizable filters \(G_{ik}(s)\) which minimized, rather than eliminated, the crosstalk. The problem of choosing the parameters of \(G_{ik}(s)\) was formulated as a multivariate minimization problem. The penalty function was defined:

\[
J = \sum_m \left| H_{ii}(j\omega_m)G_{ik}(j\omega_m) + H_{ik}(j\omega_m) \right|^2
\] (A.29)

where the summation is over a range of discrete frequencies between 0.1 and 2.0 Hz. In each case the form of \(G_{ik}(s)\) was defined in advance, and the parameters of \(G_{ik}(s)\) were
chosen to minimize the penalty function. J was minimized using the robust conjugate-gradient methods described by Fletcher and Powell (1963) and Davison and Wong (1975).

A block diagram of the crosstalk reduction filters for surge and sway is shown in Figure A.3. The corresponding parameter values are given in Table A.1. No crosstalk reduction filters were used for pitch, roll, yaw or heave. Measurements of the simulator motion with these filters active showed that this method reduced crosstalk by 6 to 20 dB. Transfer functions including the effect of the crosstalk reduction filters are given in Table 4.2.

The chief disadvantage of the open-loop method is that the crosstalk transfer functions of the motion base change somewhat whenever the base is tuned, and whenever a major component such as a servovalve is replaced. After every change to the system, the crosstalk reduction filters must also be reprogrammed. Therefore, this method is not completely satisfactory as a long term solution to the crosstalk problem.
Table A.1 Parameters used for the crosstalk reduction filters shown in Figure A.3.

<table>
<thead>
<tr>
<th>Filter</th>
<th>$K$</th>
<th>$\omega$</th>
<th>$\zeta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_{XY}(s)$</td>
<td>0.22</td>
<td>1.0</td>
<td>1.00</td>
</tr>
<tr>
<td>$G_{XPA}(s)$</td>
<td>-10.80</td>
<td>8.5</td>
<td>1.90</td>
</tr>
<tr>
<td>$G_{XPB}(s)$</td>
<td>1.30</td>
<td>1.2</td>
<td>0.50</td>
</tr>
<tr>
<td>$G_{XQA}(s)$</td>
<td>2.35</td>
<td>16.0</td>
<td>0.25</td>
</tr>
<tr>
<td>$G_{XQB}(s)$</td>
<td>-2.25</td>
<td>3.0</td>
<td>0.65</td>
</tr>
<tr>
<td>$G_{VP}(s)$</td>
<td>2.65</td>
<td>4.3</td>
<td>1.10</td>
</tr>
<tr>
<td>$G_{YO}(s)$</td>
<td>0.57</td>
<td>1.13</td>
<td>0.30</td>
</tr>
</tbody>
</table>
Figure A.1 Transfer function of the washout filter defined in equation A.9.
Figure A.2  Block diagram of the revised motion-drive algorithm.
Figure A.3  Block diagram of the crosstalk reduction filters for surge and sway specific force.
Appendix B

Measurement of Simulator Motion

The characteristics of the signal and noise motion conditions were verified by directly measuring the motion of the simulator. Because of the large volume of data involved, the simulator motion was not recorded continuously throughout the experiments. Instead, the motion-base transfer function and selected noise spectra were measured about once per week. The motion characteristics did not vary substantially over the duration of the experimental program.

A special purpose program was developed to drive the motion base and record its motion. Signal and noise motion was generated using the same algorithms used to generate simulator motion during the actual experiments. As in the experiments, the drive signals for the motion base were updated every 40 milliseconds. The simulator motion, however, was sampled every 20 milliseconds. The program was capable of producing signal alone, noise alone, or any combination of signal and noise.

A brief description of the instrumentation and the signal processing procedure is given below. The instrumentation was set up for a previous study of the motion-base dynamics, and is described in more detail by Grant (1986). The signal processing procedures used were also similar to those used by Grant, with one major difference. In Grant's study, the simulator motion was determined relative to an earth-fixed reference frame. In this study, the motion was determined simply as angular rates and specific forces in the body-fixed reference frame, \( F_H \), introduced in Chapter 2.

B.1 Instrumentation and Measurement

The simulator motion was measured using a 6-channel instrument package consisting of three Sundstrand Data Control model 303 translational accelerometers and three Honeywell GG440 Gnat miniature rate gyros. The instruments were rigidly mounted on the motion base with their sensitive axes aligned with the principal axes of the simulator cab. The rate gyros provided a measurement of the angular rates of the simulator, \( p_s, q_s \)
and $r_s$. The accelerometers measured the specific force vector, $f_s$, at the instrument package.

For the accelerometers, the effect of instrument dynamics was negligible in the frequency range of interest (0 to 10 Hz). The dynamics of the rate gyros may be represented as a second order lowpass filter with a break frequency of 40 Hz and a damping ratio of 0.5. The instrument dynamics do affect the angular measurements in the frequency range of interest, mainly by introducing a few degrees of phase lag. The analog signals from both the accelerometers and the rate gyros were passed through a fourth order Butterworth filter with a break frequency of 30 Hz in order to remove high frequency electrical and instrument noise and 60 Hz hum. This reduced aliasing effects in the sampled data. These "anti-aliasing" filters also affected the motion measurements. As with the rate gyros, the main effect was to introduce a small phase lag.

The analog signals were sampled every 20 milliseconds (i.e. at 50 Hz) using a 12 bit analog-to-digital converter. The resolution of the digital data is 0.002 m/s² for the translational data, and 0.005°/s for the angular rate data. The sampled data were recorded on disk for processing at a later time.

### B.2 Signal Processing

The sampled data were processed off-line to provide time histories and power spectra of the simulator motion. Normally, only a portion of each record was analyzed. When measuring a sample of random motion (noise), the period used for analysis was chosen to avoid transients in the first few seconds of the motion record. For estimates of the motion base transfer function, the base was driven sinusoidally at a single frequency. In this case, the period used for analysis was chosen to include an integer number of full cycles at the driven frequency.

The recorded data provided a discrete time history of the specific force and angular rate vectors at the instrument package. As noted above, the data were affected by instrument dynamics and the anti-aliasing filter. The main purpose of the signal processing stage was to remove these effects, and transform the data to the reference frame $F_H$.

This procedure comprised several stages. First, the DC offset and linear instrument drift was subtracted from the measured time histories, and the data were converted to
engineering units. Next, a frequency-domain representation of the data was generated by calculating the discrete Fourier transform (DFT) of each time history. Using the same notation as that used by Grant (1986), the DFT for a record of length $T_0$ is defined as:

$$X(\omega) = \text{DFT}[x(t),\omega] = \frac{\Delta t}{2\pi} \sum_{n=0}^{N-1} x(n\Delta t) \exp(-j\omega n\Delta t)$$

(B.1)

where $\Delta t$ is the time increment between samples and $N = T_0/\Delta t$ is the total number of samples.

The data were then adjusted to remove the effect of the anti-aliasing filter. The correction was made in the frequency domain, by multiplying the Fourier transform of the time history by the inverse of the transfer function of the anti-aliasing filter. For the angular data, a similar correction was made for the rate gyro dynamics. The corrected time history was obtained by calculating the inverse Fourier transform of the adjusted frequency domain data. The inverse transform of $X(\omega)$ is:

$$x(t) = \text{DFT}^{-1}[X(\omega),t] = \Delta \omega \sum_{k=0}^{N-1} X(k\Delta \omega) \exp(jk\Delta \omega t)$$

(B.2)

where $\Delta \omega = 2\pi/T_0$ is the frequency increment in rad/sec.

The resulting six time histories describe the specific force and angular rates in terms of reference frame $F_s$, a body frame which is fixed to the simulator with its origin located at the instrument package. A geometric transformation was applied to determine the time histories of the motion in terms of reference frame $F_H$, which has its origin at the pilot's or observer's head. Reference frame $F_H$ is shown in Figure 2.1.

For this analysis, it was assumed that the pilot's head is fixed in $F_s$, and that the axes of $F_H$ are parallel to the corresponding axes of $F_s$. The transformation required, therefore, is a simple translation. Since $F_H$ does not rotate relative to $F_s$, and because the axes are parallel, the angular rates at the pilot's head ($p$, $q$ and $r$) are equal to the measured values ($p_s$, $q_s$ and $r_s$).

Expressions for the acceleration of a point in a moving frame are given by Etkin (1972, pp. 122-124). Also recall that specific force is the vector difference between inertial
acceleration and the acceleration due to gravity:

\[ f = a - g \] (B.3)

Based on these expressions, it may be shown that the specific force vector at the pilot's head is:

\[ f_H = f_s - \tilde{\omega} r_s - \tilde{\omega} \tilde{\omega} r_s \] (B.4)

where \( f_s \) is the measured specific force vector at the instrument package,

\[
\tilde{\omega} = \begin{bmatrix}
0 & -r & q \\
 r & 0 & -p \\
-\dot{q} & \dot{p} & 0
\end{bmatrix}, \quad \dot{\omega} = \begin{bmatrix}
0 & \dot{r} & \dot{q} \\
\dot{r} & 0 & \dot{p} \\
-\dot{q} & \dot{p} & 0
\end{bmatrix},
\]

and \( r_s = \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix} \) is the location of the instrument package in \( F_H \).

The angular acceleration terms, \( \dot{p}, \dot{q} \) and \( \ddot{r} \), were obtained by differentiating \( p, q \) and \( r \) using a standard central difference algorithm.

Power spectra were calculated after the data had been corrected for instrument and filter dynamics and transformed to reference frame \( F_H \). As discussed by Grant (1986), the one-sided power spectrum of \( x(t) \) may be expressed approximately as:

\[ \Phi_{xx}(j\omega) = \frac{4\pi}{T_o} X^*(\omega) X(\omega) \] (B.5)

where \( X(\omega) \) is the discrete Fourier transform of \( x(t) \), and \( (\cdot)^* \) indicates the complex conjugate. Note that \( \Phi_{xx}(j\omega) \) is strictly defined only at \( N/2+1 \) discrete frequencies: \( \omega = k\Delta\omega, k = 0,1,\ldots,N/2 \).

A flowchart of the signal processing procedure is provided in Figure B.1.
Figure B.1  Flowchart of the motion analysis program.
Appendix C

Generation of Random Motion

The random background motion used in the experiments was generated by passing digital white noise through a fifth order bandpass filter in real time. This yielded a random time series with the desired frequency characteristics. This appendix briefly describes the noise generation algorithm.

C.1 Gaussian Noise

Frequency limited Gaussian noise, denoted \( w(t) \), was generated as a time series of independent standard normal random numbers. The time increment between points in the time series, \( \Delta t \), was 40 milliseconds. It may be shown that the one-sided power spectrum of \( w(t) \) at frequency \( f \), expressed as power/Hz, is:

\[
\Phi_{ww}(f) = 2\Delta t, \quad 0 \leq f \leq f_N \tag{C.1}
\]

where \( f_N = 1/(2\Delta t) \) is the Nyquist frequency expressed in Hertz. For the current study, \( f_N = 12.5 \text{ Hz} \).

C.2 Shaping Filters

The Gaussian noise was passed through a fifth order filter in order to produce a random time series with the desired frequency characteristics. The transfer function of the filter was:

\[
H(s) = K H_1(s) H_H(s) H_L(s) \tag{C.2}
\]

where \( K \) is a pure gain, \( H_L(s) \) is a second order low-pass filter:

\[
H_L(s) = \frac{\omega_L^2}{s^2 + 2\zeta_L\omega_Ls + \omega_L^2} \tag{C.3}
\]
and $H_I(s)$ is a second order high-pass filter:

$$H_I(s) = \frac{s^2}{s^2 + 2\omega_I s + \omega_I^2} \quad (C.4)$$

The final component, $H_F(s)$, is defined differently for translations than for rotations. The noise spectra used in this study were the same for rotation (in terms of angular acceleration) as for translation (in terms of acceleration). However, for compatibility with the motion drive algorithm described in Appendix A, the output time series is defined to be an acceleration time history for translation, and an angular velocity time history for rotation. For translations, $H_I(s)$ is a first order high-pass filter:

$$H_I(s) = \frac{s}{s + \omega_I} \quad (C.5)$$

For rotations, the high-pass filter and the integration (from angular acceleration to angular velocity) were combined into a single function. Therefore, in this case, $H_I(s)$ is a first order low-pass filter:

$$H_I(s) = \frac{\omega_I}{s + \omega_I} \quad (C.6)$$

and the gain, $K$, is divided by $\omega_I$ to maintain the correct power spectral density.

**C.3 Filter Implementation in the Time Domain**

The filters were implemented in the time domain using the Tustin method described by Knese (1979). Reid and Nahon (1985) investigated seven different real-time integration methods for differential equations. They recommended that the Tustin method be used for low order linear equations, because of its speed and accuracy. For the current case, the best results were obtained by integrating each filter separately. The value of the output time series at time step $n$, $Z_n$, was calculated by the following equations. In these equations, $W$, $X$ and $Y$ are intermediate time series. The subscripts on $W$, $X$, $Y$ and $Z$ indicate the time step. $X_n$ is the value of $X$ at time step $n$; $X_{n-1}$ and $X_{n-2}$ are previous values of $X$. 
Gaussian Noise

\[ W_n = K N(0,1) \]  \hspace{1cm} \text{(C.7)}

where \( N(0,1) \) is a standard normal random number.

Second order low pass filter

\[ X_n = \left( \frac{\gamma_L^2}{1 + 2\zeta_L\gamma_L + \gamma_L^2} \right) (W_n + 2W_{n-1} + W_{n-2}) + \left( \frac{2 - 2\gamma_L^2}{1 + 2\zeta_L\gamma_L + \gamma_L^2} \right) X_{n-1} \]

\[ - \left( \frac{1 - 2\zeta_L\gamma_L + \gamma_L^2}{1 + 2\zeta_L\gamma_L + \gamma_L^2} \right) X_{n-2} \]  \hspace{1cm} \text{(C.8)}

where \( \gamma_L = \omega_L \Delta t/2 \)

Second order high pass filter

\[ Y_n = \left( \frac{1}{1 + 2\zeta_H\gamma_H + \gamma_H^2} \right) (X_n - 2X_{n-1} + X_{n-2}) + \left( \frac{2 - 2\gamma_H^2}{1 + 2\zeta_H\gamma_H + \gamma_H^2} \right) Y_{n-1} \]

\[ - \left( \frac{1 - 2\zeta_H\gamma_H + \gamma_H^2}{1 + 2\zeta_H\gamma_H + \gamma_H^2} \right) Y_{n-2} \]  \hspace{1cm} \text{(C.9)}

where \( \gamma_H = \omega_H \Delta t/2 \)

First order low pass filter (rotations only)

\[ Z_n = \left( \frac{\gamma_i}{1 + \gamma_i} \right) (Y_n + Y_{n-1}) + \left( \frac{1 - \gamma_i}{1 + \gamma_i} \right) Z_{n-1} \]  \hspace{1cm} \text{(C.10)}

where \( \gamma_i = \omega_i \Delta t/2 \)
First order high pass filter (translations only)

\[
Z_n = \left( \frac{1}{1 + \gamma_1} \right) (Y_n - Y_{n-1}) + \left( \frac{1 - \gamma_1}{1 + \gamma_1} \right) Z_{n-1}
\]  

(C.11)

where \( \gamma_1 = \omega_1 \Delta t/2 \)

C.4 Filter Parameters Used in This Study

The values of the filter parameters used to generate translational noise conditions A through G are summarized in Table C.1. The parameters used for conditions A and B are identical to those used for condition C, except for the gain, K. Identical parameters were used to generate rotational random motion, except that K was divided by \( \omega_1 \), and \( H_i(s) \) was a low-pass filter, rather than a high-pass filter.

Typical measured power spectra for conditions C through G are shown in Figures C.1 through C.5. Each spectrum is based on 2000 points (40 seconds, sampled at 50 Hz). The theoretical spectra are shown for comparison with the data. Finally, typical time histories for noise conditions C through G are shown in Figures C.6 through C.10.

The measured spectra shown in Figures C.1 through C.5 are generally in good agreement with the theoretical design spectra. However, for frequencies above 4 Hz, the measured spectra drop off with frequency much more rapidly than the theoretical spectra. This disagreement is apparent in Figures C.1 and C.5. This high frequency attenuation results from the discrete-time implementation of the filters. Reducing \( \Delta t \) eliminates the problem (or, more accurately, it shifts the effect to higher frequencies). For the current study, we were concerned primarily with the noise power below 4 Hz, and the spectra obtained with \( \Delta t = 40 \) ms were judged to be satisfactory.
Table C.1  Filter parameters used for noise conditions A through G.

<table>
<thead>
<tr>
<th>Noise Condition</th>
<th>K</th>
<th>$\omega_l$</th>
<th>$\omega_H$</th>
<th>$\zeta_H$</th>
<th>$\omega_L$</th>
<th>$\zeta_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.010</td>
<td>0.25</td>
<td>0.25</td>
<td>0.475</td>
<td>4.00</td>
<td>0.64</td>
</tr>
<tr>
<td>B</td>
<td>0.050</td>
<td>0.25</td>
<td>0.25</td>
<td>0.475</td>
<td>4.00</td>
<td>0.64</td>
</tr>
<tr>
<td>C</td>
<td>0.250</td>
<td>0.25</td>
<td>0.25</td>
<td>0.475</td>
<td>4.00</td>
<td>0.64</td>
</tr>
<tr>
<td>D</td>
<td>0.064</td>
<td>0.05</td>
<td>0.275</td>
<td>0.20</td>
<td>0.455</td>
<td>0.20</td>
</tr>
<tr>
<td>E</td>
<td>0.064</td>
<td>0.10</td>
<td>0.55</td>
<td>0.20</td>
<td>0.91</td>
<td>0.20</td>
</tr>
<tr>
<td>F</td>
<td>0.064</td>
<td>0.10</td>
<td>1.10</td>
<td>0.20</td>
<td>1.82</td>
<td>0.20</td>
</tr>
<tr>
<td>G</td>
<td>0.064</td>
<td>0.10</td>
<td>2.20</td>
<td>0.20</td>
<td>3.64</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Notes:  
1. The quoted value of $K$ was used for translational noise (units m/s$^2$).  
   For rotations, use $K/\omega_l$.  
2. This value of $K$ is based on a time step of 40 ms.  
   For other values of $\Delta t$, $K$ is proportional to $1/\sqrt{\Delta t}$.  

Figure C.1  Typical measured power spectrum of surge specific force for noise condition C.
Figure C.2  Typical measured power spectrum of surge specific force for noise condition D.

Figure C.3  Typical measured power spectrum of surge specific force for noise condition E.
Figure C.4  Typical measured power spectrum of surge specific force for noise condition F.

Figure C.5  Typical measured power spectrum of surge specific force for noise condition G.
Figure C.6  Typical measured time history of surge specific force for noise condition C.
Figure C.7  Typical measured time history of surge specific force for noise condition D.

Figure C.8  Typical measured time history of surge specific force for noise condition E.
Figure C.9  Typical measured time history of surge specific force for noise condition F.

Figure C.10  Typical measured time history of surge specific force for noise condition G.
Appendix D

An Adaptive Algorithm for Perceptual Testing

D.1 Introduction

Many natural processes can be modeled as Bernoulli processes in which the Bernoulli parameter, p, represents the probability of occurrence of a particular outcome. We often wish to investigate how p varies with some independent variable. For example, a structural engineer may want to know how the probability of structural failure varies with the applied load. A toxicologist may want to know how the probability of a fish dying varies with the dose administered. In the current study, we want to determine the psychometric curve, which describes how the probability of a correct response varies with signal amplitude.

The most direct approach in such investigations is the fixed-level method. In psychophysics this is often called the method of constant stimuli. Using this method, a large number of trials are run at each of several predetermined signal amplitudes. The value of p is then calculated for each amplitude, and a curve is fitted to these data. The value of p for any amplitude can then be estimated by interpolating along the curve.

The fixed-level method requires a large number of trials, and as a result is usually time consuming and expensive. Often, it may be sufficient to determine a single point on the curve: the signal level corresponding to a given probability of correct response. The given probability and the associated signal amplitude are called the target probability (P_t) and the target level, respectively.

Numerous methods, known as adaptive algorithms, have been developed to home in on particular target levels. One feature is common to all adaptive algorithms: the signal amplitude for each trial is set based on the results of one or more previous trials. The information provided by each trial is maximized by setting the signal amplitude to be near the target level. Adaptive algorithms can therefore determine specific target levels in far fewer trials than the fixed-level method. The disadvantage, of course, is that only one point is obtained, rather than the entire psychometric curve.
One of the earliest and simplest adaptive algorithms is the staircase method described by Cornsweet (1962). In this method, the signal amplitude is lowered after each correct response, and raised after each incorrect response. The step size is adjusted from time to time following a set of rules. The method is normally terminated after some predetermined number of reversals. This method converges to the amplitude which yields 50% correct responses. A modification of the staircase method which is suitable for two-alternative forced choice tests is described by Clark and Stewart (1968a) and by Levitt (1971). In the modified method, the signal amplitude is decreased after two correct responses, and increased after each incorrect response. This method converges to the level which yields 75% correct responses.

A more general form of staircase algorithm, known as PEST, was developed by Taylor and Creelman (1967). In PEST, the decision to change amplitude is based on a Wald sequential test. Testing continues at a particular signal amplitude until the deviation between the observed number of correct responses and the expected number of correct responses exceeds a predefined value. The signal amplitude is then increased or decreased, as appropriate. The step size is varied following a set of formal rules. The algorithm stops when the step size is reduced to a predefined minimum value. PEST is very flexible, because it can be used to home in on any target probability. Simulations run by Taylor and Creelman (1967) show that the method is fairly efficient; for a given number of trials, it produces fairly low variance estimates of the target level.

This appendix describes an adaptive algorithm which was developed for the current study. This method assumes that the form of the psychometric curve is known, at least approximately, in advance. After each trial, a curve with the assumed form is fitted to all of the data collected so far. The target level is then estimated as the signal amplitude which corresponds to a probability of \(P_t\) on the fitted curve. The signal amplitude on the next trial is set near to the estimated target level. The estimated target level calculated after the last trial of the run is the experimental result.

The new algorithm was very suitable for the present study, and should be useful in other applications. Like PEST, it can be used to converge on any target probability. Simulation results show the new algorithm to be more efficient than PEST.
D.2 General Description of the Algorithm

Assume that each trial is an independent Bernoulli trial, where the Bernoulli parameter, \( p \), represents the probability of a correct response, \( P(C) \). Assume that \( p = p(x) \) is a strictly increasing continuous function of the signal amplitude, \( x \). It follows that \( p \) is bounded by \( p_{\min} \) and \( p_{\max} \), where \( p_{\min} = p(0) \) and \( p_{\max} \) is the limit of \( p(x) \) as \( x \) approaches infinity. Further assume that the process is stationary: \( p(x) \) does not vary from trial to trial. In perceptual studies, \( p(x) \) is called the psychometric curve.

Each trial yields a data pair \((x_i, y_i)\), where \( x_i \) is the signal amplitude and \( y_i \) is the response. Consistent with our definition of a trial as a Bernoulli process, let \( y_i = 1 \) for a correct response, and let \( y_i = 0 \) for an incorrect response. The data from \( N \) trials may be summarized in vector form as \( x \) and \( y \).

The purpose of the experiment is to find the signal amplitude, \( \theta \), for which \( p(\theta) = P_t \), where \( P_t \) is some predefined target probability. This amplitude is referred to as the target level. Because \( p(x) \) is continuous, \( \theta \) exists for every value of \( P_t \) for which \( p_{\min} \leq P_t \leq p_{\max} \).

Assume that the generic form of the function is known. Specifically, define a function \( \tilde{p}(\cdot) \) such that:

\[
\tilde{p}(x/\theta) = p(x) \tag{D.1}
\]

and assume that \( \tilde{p}(\cdot) \) is a known function. In the current problem this is equivalent to assuming that the form of the psychometric curve is that defined by, say, the correlation detector. Note that in all cases, \( \tilde{p}(1) = P_t \).

At the end of the experiment, we can calculate \( \hat{\theta} = g(y) \), an estimate of \( \theta \) based on the data from all trials. \( \hat{\theta} \) can be calculated using maximum likelihood or least squares estimation methods. The variance of \( \hat{\theta} \) is reduced by choosing the signal amplitude for each trial near to the target level. This is particularly true when there is some uncertainty about the form of \( \tilde{p}(\cdot) \). Errors due to extrapolation along an uncertain curve are minimized by choosing \( x_i \) near to \( \theta \).

How is \( x_i \) chosen? After each trial, the target level is estimated based on all of the data collected thus far. The signal amplitude for the next trial is then set equal to (or near to)
the target level. This optimizes the information contribution of each trial. There is no specific stopping rule; the procedure simply ends after some predetermined number of trials. The final estimate of the target level is the experimental result.

Two different, but similar adaptive algorithms were developed using the structure outlined above. The first estimates \( \theta \) using the method of maximum likelihood, and is called the MLE algorithm. The second estimates \( \theta \) using the least squares method, and is called the LSQ algorithm. The use of the two estimation procedures is outlined in the next two sections.

**D.3 Maximum Likelihood Estimation**

Goodwin and Payne (1977) define the maximum likelihood estimate for a parameter \( \theta \) as the value \( \hat{\theta} \) which maximizes the conditional probability density of the data. That is:

\[
\hat{\theta} = \text{arg max}_{\theta} f(y | \theta)
\]

for any \( \theta \). The function \( L(\theta) = f(y | \theta) \) is called the likelihood function. In this method we simply choose the value of \( \theta \) which makes the observed data most probable. For a discrete random process, the likelihood function is expressed in terms of probability rather than probability density. As discussed by Goodwin and Payne (1977), the maximum likelihood estimator is a "good" estimator, because it is consistent, asymptotically normal, and efficient for large sample sizes.

In the current problem, each trial is a Bernoulli trial with \( p = \tilde{p}(x_i | \theta) \). Let \( y_i = 1 \) for a correct response, and let \( y_i = 0 \) for an incorrect response. Then for a single trial, the likelihood function is:

\[
L_i(\theta) = \tilde{p}(x_i | \theta)^{y_i} (1-\tilde{p}(x_i | \theta))^{(1-y_i)}
\]

(D.3)

For \( N \) independent trials, the likelihood function is:

\[
L(\theta) = \prod_{i=1}^{N} \tilde{p}(x_i | \theta)^{y_i} (1-\tilde{p}(x_i | \theta))^{(1-y_i)}
\]

(D.4)
We want to maximize \( L(\theta) \). Since \( \ln(x) \) is monotonic with \( x \), this is equivalent to maximizing \( \ln L(\theta) \). Thus the maximum likelihood estimate of \( \theta \) is the value which maximizes:

\[
\ln L(\theta) = \sum_{i=1}^{N} y_i \ln[p(x_i/\theta)] + (1-y_i) \ln[1-p(x_i/\theta)]
\]

The optimizing value can be found by differentiating \( \ln L(\theta) \) with respect to \( \theta \), setting the derivative equal to zero, and solving for \( \theta \). The resulting nonlinear equation is:

\[
\sum_{i=1}^{N} \frac{y_i - \hat{p}(x_i/\theta)}{\hat{p}(x_i/\theta)[1-\hat{p}(x_i/\theta)]} \frac{d}{d\theta} [\hat{p}(x_i/\theta)] = 0
\]

### D.4 Least Squares Estimation

In least squares estimation, we assume the the experimental result, \( y \), is an estimate of \( h(x,\theta) \) for some function \( h \). The least squares estimate of \( \theta \) is the value, \( \hat{\theta} \), which satisfies:

\[
|y - h(x,\hat{\theta})|_Q \leq |y - h(x,\tilde{\theta})|_Q
\]

for any \( \tilde{\theta} \). In this expression, \( |x|_Q = x^TQx \), and \( Q \) is positive definite. Thus the least squares estimator minimizes the weighted sum of squares of the vector \( y-h(x,\theta) \). A more general definition is given by Goodwin and Payne (1977). The least squares method is widely used for parameter estimation for two reasons. First, it is very simple to use. Second, if certain conditions are met, the least squares estimator is the minimum variance unbiased estimator.

In the current application, we have \( N \) independent data pairs \((x_i,y_i)\) representing the signal amplitude and response from each trial. Let \( y_i = 1 \) for a correct response, and let \( y_i = 0 \) for an incorrect response. Each \( y_i \) is an unbiased estimate of \( \tilde{p}(x_i/\theta) \), since \( E[y_i] = \tilde{p}(x_i/\theta) \). The least squares estimator may then be derived from the general form given in equation D.7. Letting \( Q = I \), the least squares estimate of \( \theta \) is the value which minimizes the
penalty function $J(\theta)$, where:

$$J(\theta) = \sum_{i=1}^{N} [y_i - \tilde{p}(x_i/\theta)]^2$$  \hspace{1cm} (D.8)

The minimizing value may be found by differentiating $J(\theta)$ with respect to $\theta$, setting the derivative equal to zero, and solving for $\theta$. The resulting expression is:

$$\sum_{i=1}^{N} [y_i - \tilde{p}(x_i/\theta)] \frac{d}{d\theta} \tilde{p}(x_i/\theta) = 0$$  \hspace{1cm} (D.9)

Note the similarity between equations D.6 and D.9. The two methods are equivalent except for the weighting of the individual data points. In the least squares method, all points are given equal weight. In the maximum likelihood method, each point is weighted by the inverse of the expected variance of $y_i$. As a result, the maximum likelihood method assigns a heavy weight to points from the tail of the psychometric curve, where $\tilde{p}(x/\theta)$ approaches 1.

As noted above, there is often some uncertainty about the form of $\tilde{p}(x/\theta)$. In this case, least squares estimation is probably preferable to maximum likelihood estimation, because it does not place as much weight on the results for signal amplitudes far from the target level.

### D.5 Computer Implementation

The MLE and LSQ adaptive algorithms can be easily implemented on a small computer. In general, $\tilde{p}(\cdot)$ may be a highly nonlinear function, so that finding an exact solution of equation D.6 or D.9 becomes a complex computational task. However, a close approximation to $\theta$ can be found by simply computing $J(\theta)$ for a number of discrete values of $\theta$, then simply picking off the minimum.

The LSQ algorithm can be used to drive a detection experiment by following the nine steps outlined below.

1. Define $M$ discrete signal values, $a_k$ ($k = 1$ to $M$), representing the possible values of $x$ and $\theta$. 
2. Define a vector $J_k$ ($k = 1$ to $M$), where $J_k = J(a_k)$ represents the penalty function defined by equation D.8. Set $J_k = 0$, for $k = 1$ to $M$.

3. Define a trial counter, $i$. Set $i = 0$.

4. Set the signal amplitude for the first trial, $x_1$, to an arbitrary starting value.

Repeat steps 5 to 9 for a predetermined number of trials.

5. Set $i = i + 1$. Run an experimental trial using signal amplitude $x_i$, and record the observer’s response, $y_i$.

6. Increment the penalty function for all values of $k$:

$$J_k = J_k + [y_i - \hat{P}(x_i/a_k)]^2$$

7. Find $k$ such that $J_k \leq J_k$ for $k = 1$ to $M$. Set $\hat{a} = a_k$.

8. Set the signal amplitude for the next trial: $x_{i+1} = a_k$.

9. If the change in amplitude from trial $i$ to trial $i+1$ exceeds a specified maximum step size $\Delta x$, recalculate the signal amplitude for the next trial: $x_{i+1} = x_i \pm \Delta x$.

The final value of $\hat{\theta}$ after all trials have been completed is the experimental result.

The algorithm is readily modified to use maximum likelihood estimation rather than least squares, by changing the definition of the penalty function $J_k$ used in step 6. For the maximum likelihood method, $J_k = -\ln L(a_k)$. This may be calculated incrementally on trial $i$ by:

$$J_k = J_k - y_i \ln \hat{P}(x_i/a_k) - (1-y_i) \ln[1-\hat{P}(x_i/a_k)]$$

Clearly, computation time and storage requirements are reduced if we minimize the number of discrete values considered. For the current application, it was sufficient to use 50 discrete values spanning a 28 dB range in signal amplitude. In many cases, the minimum and maximum level and the increment between levels may be restricted by the resolution and limits of the system which generates the signal.
D.6 Evaluation of Performance

Monte Carlo simulations were run to test the performance of the LSQ and MLE algorithms. The simulated experiment was a four alternative forced choice task, with \( P_t = 0.56 \). Four different methods were tested: the LSQ and MLE algorithms, PEST, and a simple adaptive staircase method. For the staircase method, \( P_t \) was changed to 0.50. The behaviour of the simulated observer was similar to that for a correlation detector. A total of 400 simulations were run for each method. For each method, the mean and standard deviation of the parameter estimates were calculated and compared with the true value of \( \theta \). The results are shown in Table D.1.

Both the LSQ and MLE algorithms can be stopped after any predetermined number of trials. Results for these methods are given for 20, 30, 40 and 50 trials. Both methods appear to be slightly biased: they both tend to underestimate \( \theta \). After 20 trials, the MLE method underestimates by about 11\%, and the LSQ method underestimates by about 4\%. In both cases, the error decreases as the number of trials increases. For both algorithms, the variance of \( \hat{\theta} \) is roughly proportional to \( 1/N \), where \( N \) is the number of trials. The LSQ method yielded slightly lower variances than the MLE method.

PEST and the staircase method stop after a certain number of reversals, or when the step size reaches some minimum value. Therefore, the number of trials required for convergence varies from run to run. The number of trials quoted for each case in Table D.1 is the average over all simulations. For a given number of trials, the variance of \( \hat{\theta} \) obtained with PEST and the staircase method is higher than that found with either the LSQ or MLE algorithm. Therefore, the simulation data indicate that the LSQ and MLE algorithms are more efficient than PEST or the staircase method.

Monte Carlo simulations were also used to investigate the effect of two other factors. First, step 8 of the algorithm was modified to introduce some random variability in the signal amplitudes presented. The modified rule is:

\[
x_{i+1} = a_{k+j}
\]

where \( j \) is a zero mean integer random variable with:

\[
P(j=m) = 1/7, \quad m = -3 \text{ to } +3
\]
In other words, $x_{i+1}$ was chosen at random from the neighbourhood of $\hat{\theta}$, rather than being set equal to $\hat{\theta}$. This measure served to reduce the predictability of the signal amplitude. Simulation results showed that this change did not significantly affect the performance of the adaptive algorithm. If anything, the efficiency was slightly improved. The reason for the improvement is not clear.

Second, we investigated the effect of a mismatch between the true shape of $p(x)$ (in the simulated observer) and the form of $p(x)$ assumed by the adaptive algorithm. The important independent parameter in this study is the slope of $p(x)$ at $\theta$. Performance was not substantially affected by slope mismatches of up to a factor of 3. Similar results have been reported by Shelton et al (1982). We conclude that the algorithm should perform well even if the model assumed for $p(x)$ is inaccurate.

**D.7 Discussion**

Adaptive algorithms have been a subject of debate in the recent literature. Numerous algorithms have been developed, each with its proponents and detractors. Findlay (1978) proposed some improvements to the original PEST. Pentland (1980) and Emerson (1984) propose MLE methods which are similar to the one presented above. Hall (1981) proposes a hybrid algorithm, in which PEST is used to control the signal amplitude, but the final estimate of $\theta$ is a maximum likelihood estimate based on the entire data set. Taylor et al (1983) discuss these algorithms and recommend the use of the original PEST. Finally, Shelton et al (1982) used three different algorithms in an auditory detection experiment: an adaptive staircase method, PEST, and the MLE method. They found little difference in performance between methods.

The three chief advantages of the LSQ and MLE methods over PEST and the adaptive staircase method are as follows. First, for an observer with stationary behaviour, the MLE and LSQ methods are almost optimally efficient. The improvement in efficiency over PEST is illustrated by the results shown in Table D.1. Second, the MLE and LSQ methods can provide an estimate of $\theta$ in a very small number of trials (10 or 15), although the variance of such estimates is high. This makes these methods suitable for studies where only limited trials can be run. Finally, the logistics of running an experiment are simplified by using these methods, since the number of trials in each experiment is fixed and known in advance.
Two main criticisms have been levelled against the MLE methods. The first is that a large amount of computation is required between trials. This may have been a problem ten or twenty years ago, when many detection experiments were run under manual control. However, it is not a serious concern today, since most such experiments are now automated. The calculations required for the algorithm presented above are well within the capacity of even the most limited microcomputer.

The second criticism, raised by Taylor et al (1983), is that the MLE method may be sensitive to bias caused by nonstationary observer behaviour, particularly lapses in attention. During a lapse, $P(C)$ is low regardless of the signal amplitude. An incorrect response at a high signal amplitude may bias $\hat{\theta}$ upward. With the MLE algorithm, an incorrect response at signal level $x$ severely reduces $L(\theta)$ for values of $\theta$ for which $\hat{p}(x/\theta)$ approaches 1. All future estimates of $\hat{\theta}$ include the contribution of the "lapse" and may be biased upward. Proponents of PEST suggest that, because PEST has a limited "memory," it can recover from a lapse with little or no bias. Thus experimenters must choose between the efficiency offered by MLE methods, and the robustness of PEST.

The LSQ algorithm presented here provides a third alternative which might be preferable to both of the others. As shown by the simulation results discussed above, the LSQ method is as efficient as the MLE method. Also, because the LSQ algorithm places relatively less weight on data from the tails of $p(x)$, it should be less sensitive than the MLE algorithm to lapses.

The LSQ algorithm was used in the current study.
Table D.1  Performance of four adaptive algorithms.

Note: $\sigma_\theta = \sqrt{E[(\hat{\theta} - \theta)^2]}$

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Number of Trials</th>
<th>Target Prob (P_t)</th>
<th>$E[\hat{\theta}]$</th>
<th>$\sigma_\theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum Likelihood (MLE)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.56</td>
<td>0.891</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.56</td>
<td>0.928</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>0.56</td>
<td>0.940</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.56</td>
<td>0.957</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>Least Squares (LSQ)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.56</td>
<td>0.963</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.56</td>
<td>0.973</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>0.56</td>
<td>0.976</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.56</td>
<td>0.982</td>
<td>0.22</td>
<td></td>
</tr>
</tbody>
</table>

PEST:

1. Wald crit. = 1, min. step = 0.20
   | 34 ± 16 | 0.56 | 1.075 | 0.37 |

2. Wald crit. = 1, min. step = 0.10
   | 53 ± 26 | 0.56 | 0.984 | 0.31 |

Adaptive Staircase:

|                      |                  |                   |                   |                 |
|                      | 10 reversals     | 20 ± 4            | 0.50              | 1.000           | 0.44 |
|                      | 20 reversals     | 40 ± 6            | 0.50              | 0.988           | 0.34 |
Appendix E

A Payoff Matrix for the Rating Procedure

In Chapter 5, we introduced the idea of using a utility function or payoff matrix to influence an observer’s response criteria in the rating procedure. This appendix shows how the optimal response criteria can be derived for a given payoff matrix, and discusses some restrictions which must be imposed on the matrix elements in order to obtain meaningful results.

In the rating procedure, the observer is exposed to a stimulus which is either noise alone (n) or signal plus noise (s), and must choose a discrete response category j, where \( j = 1, 2, \ldots, m \). Suppose that the observer is scored on his performance: let \( V_{nj} \) be the value of selecting response category j when the stimulus is noise alone, and let \( V_{sj} \) be the value of selecting category j when the stimulus is signal plus noise. The values \( V_{nj} \) and \( V_{sj} \) (for \( j = 1, 2, \ldots, m \)) constitute a matrix like that shown in Figure E.1. This matrix is referred to here as the payoff matrix. Assume that the observer is well aware of the values of the matrix elements, and that his objective is to maximize his score.

As in the general development of signal detection theory given in Chapter 3, we assume that an observation can be summarized by a scalar evidence variable, e. For a given observation, the expected value of the payoff for selecting category j is:

\[
V(j|e) = V_{sj}P(s|e) + V_{nj}P(n|e)
\]  

(E.1)

where \( P(s|e) \) and \( P(n|e) \) are the \textit{a posteriori} probabilities of s and n, given e. To achieve his goal, the observer must choose the value of j for which \( V(j|e) \) is a maximum.

The optimum value of j may be found by sequentially comparing pairs of categories. Consider the comparison between categories i and j. The observer should select category j in preference to i if and only if \( V(j|e) \geq V(i|e) \), that is, if \( V(j|e) - V(i|e) \geq 0 \). (If \( V(j|e) = V(i|e) \), it clearly does not matter which category is selected. We arbitrarily assume the response to be j in the case of equality.) The difference in expected payoff
It was shown in Chapter 3 that:

\[
V(j | e) - V(i | e) = (V_{sj} - V_{si})P(s | e) + (V_{nj} - V_{ni})P(n | e) \tag{E.2}
\]

where \( P(s) \) and \( P(n) \) are the \textit{a priori} probabilities of \( s \) and \( n \), and \( L(e) \) is the likelihood ratio. Combining equations E.2 and E.3 and rearranging gives:

\[
V(j | e) - V(i | e) = \frac{P(s)}{P(n)} \left[ (V_{sj} - V_{si})L(e) + (V_{nj} - V_{ni}) \frac{P(n)}{P(s)} \right] \tag{E.4}
\]

We may assume without loss of generality that \( V_{sj} \geq V_{si} \). Also assume that \( P(s) > 0 \), \( P(n) > 0 \), and \( P(n | e) > 0 \), since if any of these is zero, the solution is trivial. There are three cases to be considered:

**Case 1:** \( V_{sj} - V_{si} = 0 \), and \( V_{nj} - V_{ni} < 0 \)

In this case, \( V(j | e) - V(i | e) < 0 \) for all values of \( L(e) \). Therefore the observer should always choose category \( i \) in preference to \( j \).

**Case 2:** \( V_{sj} - V_{si} \geq 0 \), and \( V_{nj} - V_{ni} \geq 0 \)

In this case, \( V(j | e) - V(i | e) \geq 0 \) for all values of \( L(e) \). Therefore the observer should always choose category \( j \) in preference to \( i \).

**Case 3:** \( V_{sj} - V_{si} > 0 \), and \( V_{nj} - V_{ni} < 0 \)

In this case, \( V(j | e) - V(i | e) \geq 0 \) if and only if \( L(e) \geq \beta(i,j) \), where:

\[
\beta(i,j) = \frac{P(n)(V_{ni} - V_{nj})}{P(s)(V_{sj} - V_{si})} \tag{E.5}
\]
The observer should select category j in preference to category i if and only if $L(e) \geq \beta(i,j)$. This is the result given in equation 5.6.

Only in case 3 does the observation influence the decision. In cases 1 and 2, one category is always preferable to the other, independent of the observation. In both cases, one category should never be selected. Cases 1 and 2 can be avoided by imposing suitable restrictions on the values of the payoff matrix elements.

The first two restrictions are as follows: $V_{sj}$ must be strictly increasing with j, and $V_{nj}$ must be strictly decreasing with j. These restrictions ensure that $V_{nj}$ is strictly decreasing with $V_{sj}$, so that each pair of categories meets the conditions for case 3. They also ensure that for each comparison between categories i and j with $i < j$, the optimal decision is to choose j if and only if $L(e) \geq \beta(i,j)$. Thus the resulting response rule is consistent: $L(e)$ is strictly increasing with j.

One further restriction must be imposed in order to obtain responses in all categories: the criteria $\beta_j = \beta(j,j+1)$ must be strictly increasing with j. If this restriction is not met, some response categories will never be used. For example, consider the case where $\beta(2,3) < \beta(1,2)$. It is readily shown from the definition of $\beta(i,j)$ that $\beta(1,3)$ satisfies the inequality:

$$\beta(2,3) < \beta(1,3) < \beta(1,2)$$  \hspace{1cm} (E.6)

This leads to the following response strategy for a given observation, e. If $L(e) < \beta(1,3)$, the optimal response is 1. If $L(e) \geq \beta(1,3)$, the optimal response is 3. Therefore, according to the optimal response rule, category 2 should never be selected.

If the three restrictions on $V_{nj}$ and $V_{sj}$ are met, the optimal decision rule given an observation e is: respond j if and only if $\beta_{j-1} \leq L(e) < \beta_j$, where for completeness we define $\beta_0 = 0$ and $\beta_m = +\infty$. For an m-alternative rating task, the observer must maintain only m-1 criteria.
The payoff matrix for a 5 category rating task.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>$V_{n1}$</td>
</tr>
<tr>
<td>Signal plus noise</td>
<td>$V_{s1}$</td>
</tr>
</tbody>
</table>

**Figure E.1** The payoff matrix for a 5 category rating task.
Appendix F

Maximum Likelihood Estimation of ROC Parameters

F.1 Introduction

In an m-alternative rating task, the observer is exposed to a stimulus which may be either noise alone (n) or signal plus noise (s), and responds by selecting a response category $R_j$, where $j = 1, 2, ..., m$. As discussed in Chapter 5, the data from repeated trials may be interpreted to give $m-1$ points on the ROC curve. In addition to finding these points, we would like to determine the underlying form and parameters describing the ROC curve. In other words, we need a procedure for fitting ROC curves to rating-procedure data. Curve fitting by eye is usually unsatisfactory unless the data agree very well with the model. Also, such informal methods do not provide an assessment of the confidence or variability of the parameter estimates.

Dorfman and Alf (1969) developed a procedure for obtaining maximum likelihood estimates of signal detection parameters based on rating-procedure data. They assumed that the ROC curve is based on Gaussian distributions with unequal variance. A summary of the general approach is given in Section F.2. The procedure provides maximum likelihood estimates of the ROC parameters, a covariance matrix for the estimates, and a $\chi^2$ statistic which summarizes the deviation of the data from the maximum likelihood fitted model. The covariance matrix may be used to determine confidence intervals for the parameter estimates. The $\chi^2$ statistic provides a goodness-of-fit index with which to assess the overall agreement between model and data. The use of this statistic is discussed in Chapter 5.

The procedure outlined in Section F.2 is general in the sense that it is independent of the form of the underlying distributions, $f(e | n)$ and $f(e | s)$. As a result, it is easy to apply the method to specific cases, that is, specific types of detector. Three such cases are considered here. In the first case, discussed in Section F.3, the distributions were assumed to be Gaussian with unequal variances. This is the case considered by Dorfman and Alf. In the second case, the distributions were assumed to be Gaussian with the ratio of variances fixed and known. This case is discussed in Section F.4. In the third case,
the distributions were assumed to be those associated with the envelope detector described in Section 3.4.3. This case is discussed in Section F.5.

**F.2 General Form of the Estimation Procedure**

The development in this section is adapted from that given by Dorfman and Alf (1969). They assumed that the distributions underlying the ROC curve were Gaussian with unequal variance, and derived a parameter estimation procedure for this specific case. In the presentation here, we derive a general parameter estimation procedure which is easily applied to any specific case.

**F.2.1 Response Probabilities**

In the rating procedure, the observable events of each experiment include two stimulus classes, s and n, and a set of m response categories, \( R_j \) (\( j = 1 \) to \( m \)). The derivation of the solution is based on the following assumptions.

1. On each trial, the presentation of s or n leads to an event e on a unidimensional continuum. Note that e may or may not be monotonic with the likelihood ratio, \( L(e) \).

2. Over an infinite number of trials, the presentation of n yields a continuous distribution of e’s with probability distribution function \( F(e \mid n) \) and density \( f(e \mid n) \).

3. Over an infinite number of trials, the presentation of s yields a continuous distribution of e’s with probability distribution function \( F(e \mid s) \) and density \( f(e \mid s) \).

4. There exists a set of criteria or cutoffs \( X_k \) (\( k = 1 \) to \( m-1 \)) such that:
   
   i. if \( e < X_1 \), the response is \( R_1 \)
   
   ii. if \( e \geq X_{m-1} \), the response is \( R_m \)
   
   iii. if \( X_{k-1} \leq e < X_k \), the response is \( R_k \)

5. All trials are mutually independent.
The definition of the event continuum is somewhat arbitrary. In some cases it may be convenient to develop the solution in terms of a transformed variable, z. Define $F_n(z)$ and $F_s(z)$ such that:

$$F_n(z) = F(e | n) \quad (F.1)$$

$$F_s(z) = F(e | s) \quad (F.2)$$

where $z = h(e)$, and $h(e)$ is a monotonic function of e. (The set of allowable functions includes the identity function.) Let $f_n(z)$ and $f_s(z)$ denote the conditional densities of z.

The probability that response category $R_j$ is chosen when the stimulus is n may be written:

$$P(R_j | n) = P(X_j - e < X_j | n)$$

$$= P(Z_j - z < Z_j | n)$$

$$= F_n(Z_j) - F_n(Z_{j-1}) \quad (F.3)$$

where $Z_j = h(X_j)$, for $j = 1$ to $m-1$. Similarly,

$$P(R_j | s) = F_s(Z_j) - F_s(Z_{j-1}) \quad (F.4)$$

In equations F.3 and F.4, we define $F_n(Z_0) = F_s(Z_0) = 0$, and $F_n(Z_m) = F_s(Z_m) = 1$.

**F.2.2 Formulation of the Maximum Likelihood Estimator**

$P(R_j | n)$ and $P(R_j | s)$ depend on several unknown parameters: the underlying form and parameters defining the distributions $F_n(z)$ and $F_s(z)$, and the value of the cutoffs, $Z_k$ ($k = 1$ to $m-1$). Let the unknown parameters be represented by a vector, $\theta$. The unknown parameters may be estimated using the method of maximum likelihood (see Goodwin and Payne, 1977, chapter 3). For a given data sample $y$, the "best" estimate of $\theta$ is the value which maximizes the likelihood function, $L(\theta) = P(y | \theta)$.

Given a random sample of rating responses, the likelihood function of the sample is:

$$L(\theta) = \prod_{j=1}^{m} P(R_j | n)^{r_{nj}} P(R_j | s)^{r_{sj}} \quad (F.5)$$
where \( r_{nj} \) and \( r_{sj} \) are the number of responses in category \( R_j \) obtained when the stimulus is \( n \) or \( s \), respectively. We want to find the value of \( \theta \) which maximizes \( L(\theta) \). It is easier to find the maximum of \( \ln L(\theta) \), where:

\[
\ln L(\theta) = \sum_{j=1}^{m} r_{nj} \ln P(R_j \mid n) + r_{sj} \ln P(R_j \mid s)
\]

\[
= \sum_{j=1}^{m} r_{nj} \ln [F_n(Z_j) - F_n(Z_{j-1})] + r_{sj} \ln [F_s(Z_j) - F_s(Z_{j-1})]
\]

Since \( \ln(L) \) is monotonic with \( L \), the value of \( \theta \) which maximizes \( \ln[L(\theta)] \) also maximizes \( L(\theta) \). \( L(\theta) \) can therefore be maximized by differentiating \( \ln[L(\theta)] \) with respect to \( \theta \), setting the first partial derivatives equal to zero, and solving the resulting set of equations for \( \theta \).

Because the equations are nonlinear, it is necessary to use an iterative solution technique. The equations can be solved using the method of scoring (Rao, 1965), an adaptation of the Newton-Raphson method which is commonly used for statistical estimation. Given a vector of parameter estimates, \( \theta_0 \), an improved vector of estimates is given by:

\[
\theta_1 = \theta_0 + V^{-1}g
\]

where \( g \) is the vector of first partial derivatives of \( \ln L \) (with respect to \( \theta \)) at \( \theta_0 \):

\[
g = \nabla \ln L(\theta_0)
\]

and \( V = E[-g'(\theta_0)] \), where \( g' \) is the Jacobian of \( g \) and \( E[\cdot] \) denotes expected value. The elements of matrix \( V \) are defined as:

\[
v_{ij} = -E \left[ \frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j} \right]
\]

Equation F.7 is repeated until some convergence criterion is met. For all cases examined in this study, the procedure converged within four to six iterations. As discussed by Rao (1965), \( V^{-1} \) is the covariance matrix for the parameter estimates after the iterations are complete. The diagonal of \( V^{-1} \) gives the variance of each parameter.
F.2.3 Expressions for $g$ and $V$

Let us consider the general case in which $F_n(z)$ and $F_s(z)$ are distribution functions of a known (or assumed) form which depend on two unknown parameters, $\alpha$ and $\beta$. The parameter vector for this case is:

$$\theta = [\alpha \beta Z_1 Z_2 \ldots Z_{m-1}]^T \quad (F.10)$$

The derivation of expressions deriving $g$ and $V$ is tedious but mechanical. We will use the following notation:

$$F_{nj} = F_n(Z_j)$$

$$F_{sj} = F_s(Z_j)$$

$$f_{nj} = \frac{\partial F_n}{\partial z}(Z_j)$$

$$f_{sj} = \frac{\partial F_s}{\partial z}(Z_j)$$

$$F_{n\alpha,j} = \frac{\partial F_n}{\partial \alpha}(Z_j)$$

$$F_{s\alpha,j} = \frac{\partial F_s}{\partial \alpha}(Z_j)$$

$$F_{n\beta,j} = \frac{\partial F_n}{\partial \beta}(Z_j)$$

$$F_{s\beta,j} = \frac{\partial F_s}{\partial \beta}(Z_j)$$

Also, let:

$$N_n = \sum_{j=1}^{m} r_{nj}$$

and

$$N_s = \sum_{j=1}^{m} r_{sj}$$

be the total number of incidences of $n$ and $s$ respectively. Then the elements of $g$ (the first partial derivatives) are:

$$\frac{\partial \ln L}{\partial \alpha} = \sum_{j=1}^{m-1} F_{n\alpha,j} C_{nj} + F_{s\alpha,j} C_{sj} \quad (F.11)$$

$$\frac{\partial \ln L}{\partial \beta} = \sum_{j=1}^{m-1} F_{n\beta,j} C_{nj} + F_{s\beta,j} C_{sj} \quad (F.12)$$
\[
\frac{\partial \ln L}{\partial Z_j} = f_{nj} C_{nj} + f_{sj} C_{sj}
\]  
(F.13)

where \(C_{nj} = \frac{r_{nj}}{F_{nj} - F_{n,j-1}} - \frac{r_{n,j+1}}{F_{n,j+1} - F_{nj}}\)  
(F.14)

and \(C_{sj} = \frac{r_{sj}}{F_{sj} - F_{s,j-1}} - \frac{r_{s,j+1}}{F_{s,j+1} - F_{sj}}\)  
(F.15)

The expected values of the second partial derivatives are:

\[
E \left[ \frac{\partial^2 \ln L}{\partial \alpha^2} \right] = \sum_{j=1}^{m-1} \left( -N_n F_{n\alpha,j} C_{n\alpha,j} - N_s F_{s\alpha,j} C_{s\alpha,j} \right)
\]  
(F.16)

\[
E \left[ \frac{\partial^2 \ln L}{\partial \alpha \partial \beta} \right] = \sum_{j=1}^{m-1} \left( -N_n F_{n\beta,j} C_{n\beta,j} - N_s F_{s\beta,j} C_{s\beta,j} \right)
\]  
(F.17)

\[
E \left[ \frac{\partial^2 \ln L}{\partial \alpha \partial Z_j} \right] = -N_n f_{nj} C_{n\alpha,j} - N_s f_{sj} C_{s\alpha,j}
\]  
(F.18)

where \(C_{n\alpha,j} = \frac{F_{n\alpha,j} - F_{n\alpha,j-1}}{F_{nj} - F_{n,j-1}} - \frac{F_{n\alpha,j+1} - F_{n\alpha,j}}{F_{n,j+1} - F_{nj}}\)  
(F.19)

and \(C_{s\alpha,j} = \frac{F_{s\alpha,j} - F_{s\alpha,j-1}}{F_{sj} - F_{s,j-1}} - \frac{F_{s\alpha,j+1} - F_{s\alpha,j}}{F_{s,j+1} - F_{sj}}\)  
(F.20)

\[
E \left[ \frac{\partial^2 \ln L}{\partial \beta^2} \right] = \sum_{j=1}^{m-1} \left( -N_n F_{n\beta,j} C_{n\beta,j} - N_s F_{s\beta,j} C_{s\beta,j} \right)
\]  
(F.21)

\[
E \left[ \frac{\partial^2 \ln L}{\partial \beta \partial Z_j} \right] = -N_n f_{nj} C_{n\beta,j} - N_s f_{sj} C_{s\beta,j}
\]  
(F.22)

where \(C_{n\beta,j} = \frac{F_{n\beta,j} - F_{n\beta,j-1}}{F_{nj} - F_{n,j-1}} - \frac{F_{n\beta,j+1} - F_{n\beta,j}}{F_{n,j+1} - F_{nj}}\)  
(F.23)
and $C_{s\beta,j} = \frac{F_{s\beta,j} - F_{s\beta,j-1}}{F_{sj} - F_{sj-1}} - \frac{F_{s\beta,j+1} - F_{s\beta,j}}{F_{sj+1} - F_{sj}}$ (F.24)

$$E\left[ \frac{\partial^2 \ln L}{\partial Z_j^2} \right] = -N_n f_{nj}^2 \left( \frac{1}{F_{nj} - F_{nj-1}} + \frac{1}{F_{nj+1} - F_{nj}} \right)$$

$$-N_s f_{sj}^2 \left( \frac{1}{F_{sj} - F_{sj-1}} + \frac{1}{F_{sj+1} - F_{sj}} \right)$$ (F.25)

$$E\left[ \frac{\partial^2 \ln L}{\partial Z_j \partial Z_{j+1}} \right] = \frac{N_n f_{nj} f_{nj+1}}{F_{nj+1} - F_{nj}} + \frac{N_s f_{sj} f_{sj+1}}{F_{sj+1} - F_{sj}}$$ (F.26)

$$E\left[ \frac{\partial^2 \ln L}{\partial Z_j \partial Z_k} \right] = 0, \text{ if } |j-k| > 1$$ (F.27)

The equations for $g$ and $V$ for the case where there are more (or less) than two unknown parameters are readily inferred from the equations given above.

**F.3 Gaussian Distributions with Unequal Variance**

Some commonly used detection models predict that the conditional distributions of $e$ are Gaussian with unequal variance. This section outlines the solution for this case.

Assume that $F(e\mid n)$ is Gaussian with mean $\mu_n$ and variance $\sigma_n^2$, and that $F(e\mid s)$ is Gaussian with mean $\mu_s$ and variance $\sigma_s^2$. Then:

$$F_n(z) = F(e\mid n) = F(z)$$ (F.28)

where $z = (e - \mu_n)/\sigma_n$ and $F$ is the standard normal distribution. Similarly:

$$F_s(z) = F(bs - A)$$ (F.29)

where $b = \sigma_s/\sigma_n$, and $A = (\mu_s - \mu_n)/\sigma_s$. In this case, therefore, $F_n(z)$ and $F_s(z)$ (and the resulting ROC) depend on two unknown parameters: $A$, a normalized difference of means, and $b$, the ratio of variances.
where \( b = \sigma_n/\sigma_s \), and \( A = (\mu_n - \mu_n)/\sigma_s \). In this case, therefore, \( F_n(z) \) and \( F_s(z) \) (and the resulting ROC) depend on two unknown parameters: \( A \), a normalized difference of means, and \( b \), the ratio of variances.

Maximum likelihood estimates of \( A \) and \( b \) may be found by a straightforward application of equations F.7 through F.9. The parameter vector for this case is given by equation F.10, where \( \alpha = A \), \( \beta = b \), and \( Z_j = (X_j - \mu_n)/\sigma_n \). The elements of \( g \) and \( V \) are described by equations F.11 through F.27. The necessary values of the probability distribution functions and their derivatives are as follows:

\[
F_{nj} = F(Z_j) \quad \text{(F.30)}
\]

\[
f_{nj} = f(Z_j) \quad \text{(F.31)}
\]

\[
F_{n\alpha,j} = 0 \quad \text{(F.32)}
\]

\[
F_{n\beta,j} = 0 \quad \text{(F.33)}
\]

\[
F_{sj} = F(bZ_j - A) \quad \text{(F.34)}
\]

\[
f_{sj} = b f(bZ_j - A) \quad \text{(F.35)}
\]

\[
F_{s\alpha,j} = -f(bZ_j - A) \quad \text{(F.36)}
\]

\[
F_{s\beta,j} = Z_j f(bZ_j - A) \quad \text{(F.37)}
\]

where \( F(\cdot) \) is the standard normal distribution function, and \( f(\cdot) \) is the standard normal density function. The maximum likelihood estimate for the index of detectability, \( d_A' \), can be found from the estimates of \( A \) and \( b \), as follows:

\[
d_A' = \frac{A}{\sqrt{(1+b^2)/2}} \quad \text{(F.38)}
\]
F.4 Gaussian Distributions Where $\sigma_n/\sigma_s$ is Fixed and Known

In some cases, the value of the ROC slope parameter, $b = \sigma_n/\sigma_s$, is fixed and known. An example of this is the correlation detector, for which $b = 1$. The solution for this case is identical to that outlined in Section F.3, except that the rank of the system of equations is reduced by one. Since $b$ is fixed, the equation for $b$ can be eliminated. In this case:

$$\theta = [\alpha \ Z_1 \ Z_2 \ \ldots \ Z_{m-1}]^T$$

(F.39)

where, as before, $\alpha = A$ and $Z_j = (X_j - \mu_n)/\sigma_n$. The second element of $g$ is eliminated (equation F.12), as are the second row and column of $V$ (equations F.17, F.21 and F.22).

F.5 Distributions Based on the Envelope Detector

The envelope detector is another case of interest in the current study. The distributions associated with the envelope detector were derived in Section 3.4.3. Assume that the conditional distributions of $e$ can be written as a function of the normalized variable, $z = e/\sigma_s$, for some $\sigma_s$. The conditional density functions are:

$$f_n(z) = z \exp(-z^2/2)$$

(F.40)

$$f_s(z) = z I_0(\alpha z) \exp(-\alpha^2/2) \exp(-z^2/2)$$

(F.41)

where $\alpha$ is an unknown parameter to be estimated. Note that $\alpha$ is the normalized signal amplitude, $A_\star$, of Section 3.4.3. As above, the maximum likelihood estimate of $\alpha$ may be found by applying equations F.7 through F.9. In this case, the parameter vector is:

$$\theta = [\alpha \ Z_1 \ Z_2 \ \ldots \ Z_{m-1}]^T$$

(F.42)

where $Z_k = X_k/\sigma_s$. The elements of $g$ and $V$ are described by equations F.11 through F.27. Since there is only one parameter to estimate (in addition to the $Z_j$'s), equations F.12, F.17 and F.21 through F.24 are unnecessary. The required values of the distribution functions and their derivatives are as follows:

$$F_{nj} = F_n(Z_j) = 1 - \exp(-Z_j^2/2)$$

(F.43)

$$f_{nj} = f_n(Z_j) = Z_j \exp(-Z_j^2/2)$$

(F.44)
Numerical evaluation of the equations was more cumbersome in this case. Because no simple expression for \( F_{\alpha j} \) or \( F_{s\alpha j} \) is available, it was necessary to numerically integrate \( f_s(z) \) to obtain \( F_s(z) \), then numerically differentiate \( F_s(z) \) to obtain \( F_{s\alpha j} \).
Appendix G

Results of Experiment 1:
Stimulus-Response Matrices and ROC Curves
DG: surge, 0.022 m/s². 436 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>50</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>14</td>
</tr>
</tbody>
</table>

DG: surge, 0.033 m/s². 351 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>104</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>4</td>
</tr>
</tbody>
</table>

DG: pitch, 0.30°/s. 443 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>41</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>9</td>
</tr>
</tbody>
</table>

DG: pitch, 0.42°/s. 436 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>70</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>14</td>
</tr>
</tbody>
</table>
PG: surge, 0.022 m/s^2. 350 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>25</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>12</td>
</tr>
</tbody>
</table>

PG: surge, 0.033 m/s^2. 434 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>77</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>12</td>
</tr>
</tbody>
</table>

PG: pitch, 0.30°/s. 431 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>64</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>17</td>
</tr>
</tbody>
</table>

PG: pitch, 0.42°/s. 433 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>57</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>6</td>
</tr>
</tbody>
</table>
GG: surge, 0.027 m/s². 430 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>36</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>10</td>
</tr>
</tbody>
</table>

GG: surge, 0.038 m/s². 429 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>74</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>7</td>
</tr>
</tbody>
</table>

GG: pitch, 0.30°/s. 421 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>48</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>15</td>
</tr>
</tbody>
</table>

GG: pitch, 0.42°/s. 430 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>68</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>12</td>
</tr>
</tbody>
</table>
MN: surge, 0.022 m/s². 426 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>41</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>12</td>
</tr>
</tbody>
</table>

MN: surge, 0.033 m/s². 427 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>79</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>17</td>
</tr>
</tbody>
</table>

MN: pitch, 0.24°/s. 429 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>56</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>11</td>
</tr>
</tbody>
</table>

MN: pitch, 0.36°/s. 423 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>91</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>9</td>
</tr>
</tbody>
</table>
PR: surge, 0.027 m/s². 423 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>66</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>24</td>
</tr>
</tbody>
</table>

PR: surge, 0.038 m/s². 426 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>115</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>28</td>
</tr>
</tbody>
</table>

PR: pitch, 0.30°/s. 424 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>61</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>25</td>
</tr>
</tbody>
</table>

PR: pitch, 0.42°/s. 434 trials.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Number of responses in category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noise alone</td>
<td>88</td>
</tr>
<tr>
<td>Signal+noise</td>
<td>29</td>
</tr>
</tbody>
</table>
DG: SURGE, 0.022 M/S². 436 TRIALS.

DG: SURGE, 0.033 M/S². 351 TRIALS.
DG: PITCH, 0.30 °/S. 443 TRIALS.

DG: PITCH, 0.42 °/S. 436 TRIALS.
PG: PITCH, 0.30°/S. 431 TRIALS.

PG: PITCH, 0.42°/S. 433 TRIALS.
GG: SURGE, 0.027 M/S². 430 TRIALS.

GG: SURGE, 0.038 M/S². 429 TRIALS.
GG: PITCH, 0.30 °/S. 421 TRIALS.

GG: PITCH, 0.42 °/S. 430 TRIALS.
MN: SURGE, 0.022 m/s². 426 TRIALS.

MN: SURGE, 0.033 m/s². 427 TRIALS.
MN: PITCH, 0.24°/S. 429 TRIALS.

MN: PITCH, 0.36°/S. 423 TRIALS.
Z IS IN J

PR: SURGE, 0.027 M/S^2, 423 TRIALS.

PR: SURGE, 0.038 M/S^2, 426 TRIALS.
PR: PITCH, 0.30°/S. 424 TRIALS.

PR: PITCH, 0.42°/S. 434 TRIALS.
Appendix H

Results of Experiment 3:
Inter-Axis Effects
Table H.1  Critical amplitudes for surge signals (milli-g), from experiment 3. Mean values shown are geometric means.

<table>
<thead>
<tr>
<th>Noise Cond.</th>
<th>Freq</th>
<th>DG</th>
<th>PG</th>
<th>GG</th>
<th>MN</th>
<th>PR</th>
<th>DS</th>
<th>Geom. Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surge</td>
<td>0.2</td>
<td>9.24</td>
<td>1.68</td>
<td>4.20</td>
<td>7.14</td>
<td>7.14</td>
<td>12.18</td>
<td>5.86</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>6.94</td>
<td>7.34</td>
<td>7.75</td>
<td>7.34</td>
<td>9.38</td>
<td>5.30</td>
<td>7.24</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>10.27</td>
<td>3.85</td>
<td>7.28</td>
<td>8.99</td>
<td>2.14</td>
<td>4.28</td>
<td>5.36</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>18.30</td>
<td>3.74</td>
<td>5.41</td>
<td>6.66</td>
<td>12.48</td>
<td>2.50</td>
<td>6.52</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>9.20</td>
<td>7.60</td>
<td>6.00</td>
<td>3.60</td>
<td>2.40</td>
<td>4.80</td>
<td>5.09</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>10.21</td>
<td>4.23</td>
<td>5.99</td>
<td>6.46</td>
<td>5.33</td>
<td>5.06</td>
<td>5.96</td>
</tr>
<tr>
<td>Pitch</td>
<td>0.2</td>
<td>5.88</td>
<td>3.78</td>
<td>5.46</td>
<td>2.94</td>
<td>5.04</td>
<td>5.46</td>
<td>4.63</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>7.34</td>
<td>2.04</td>
<td>6.12</td>
<td>3.67</td>
<td>3.67</td>
<td>5.71</td>
<td>4.38</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>9.42</td>
<td>4.28</td>
<td>4.71</td>
<td>3.85</td>
<td>5.56</td>
<td>7.28</td>
<td>5.56</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>5.82</td>
<td>4.16</td>
<td>5.41</td>
<td>4.16</td>
<td>2.91</td>
<td>4.99</td>
<td>4.46</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>4.40</td>
<td>4.40</td>
<td>2.80</td>
<td>5.60</td>
<td>8.40</td>
<td>7.20</td>
<td>5.14</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>6.36</td>
<td>3.60</td>
<td>4.74</td>
<td>3.96</td>
<td>4.79</td>
<td>6.06</td>
<td>4.81</td>
</tr>
<tr>
<td>Sway</td>
<td>0.2</td>
<td>2.10</td>
<td>4.20</td>
<td>5.88</td>
<td>4.20</td>
<td>5.46</td>
<td>5.04</td>
<td>4.26</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>1.63</td>
<td>3.26</td>
<td>4.90</td>
<td>4.08</td>
<td>5.71</td>
<td>4.49</td>
<td>3.74</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>2.57</td>
<td>4.71</td>
<td>4.71</td>
<td>4.71</td>
<td>5.14</td>
<td>3.00</td>
<td>4.00</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>2.50</td>
<td>4.16</td>
<td>7.07</td>
<td>4.16</td>
<td>4.16</td>
<td>5.82</td>
<td>4.41</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>3.20</td>
<td>4.40</td>
<td>6.40</td>
<td>4.40</td>
<td>5.20</td>
<td>5.60</td>
<td>4.75</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>2.34</td>
<td>4.12</td>
<td>5.72</td>
<td>4.30</td>
<td>5.10</td>
<td>4.67</td>
<td>4.22</td>
</tr>
<tr>
<td>Roll</td>
<td>0.2</td>
<td>2.10</td>
<td>5.04</td>
<td>7.56</td>
<td>7.98</td>
<td>3.78</td>
<td>4.20</td>
<td>4.65</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>3.26</td>
<td>4.08</td>
<td>3.26</td>
<td>1.63</td>
<td>4.49</td>
<td>5.30</td>
<td>3.45</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>3.42</td>
<td>2.57</td>
<td>2.57</td>
<td>1.71</td>
<td>2.57</td>
<td>3.00</td>
<td>2.58</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>1.66</td>
<td>3.74</td>
<td>2.91</td>
<td>2.91</td>
<td>4.58</td>
<td>4.16</td>
<td>3.17</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>3.20</td>
<td>2.80</td>
<td>2.80</td>
<td>2.80</td>
<td>3.20</td>
<td>4.40</td>
<td>3.16</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>2.63</td>
<td>3.54</td>
<td>3.49</td>
<td>2.83</td>
<td>3.64</td>
<td>4.14</td>
<td>3.34</td>
</tr>
<tr>
<td>Condition A</td>
<td>0.2</td>
<td>3.36</td>
<td>2.94</td>
<td>3.78</td>
<td>4.62</td>
<td>4.20</td>
<td>2.94</td>
<td>3.59</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>1.63</td>
<td>4.08</td>
<td>3.67</td>
<td>4.08</td>
<td>2.45</td>
<td>3.26</td>
<td>3.04</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>3.00</td>
<td>3.42</td>
<td>5.14</td>
<td>2.57</td>
<td>2.14</td>
<td>4.28</td>
<td>3.28</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>3.33</td>
<td>2.91</td>
<td>3.74</td>
<td>2.08</td>
<td>3.74</td>
<td>4.16</td>
<td>3.25</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>1.60</td>
<td>2.00</td>
<td>2.40</td>
<td>2.40</td>
<td>2.00</td>
<td>2.00</td>
<td>2.05</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>2.45</td>
<td>2.99</td>
<td>3.64</td>
<td>3.00</td>
<td>2.78</td>
<td>3.21</td>
<td>2.99</td>
</tr>
</tbody>
</table>
Table H.2 Critical amplitudes for pitch signals (°/s), from experiment 3.
Mean values shown are geometric means.

<table>
<thead>
<tr>
<th>Noise Cond.</th>
<th>Freq</th>
<th>DG</th>
<th>PG</th>
<th>GG</th>
<th>MN</th>
<th>PR</th>
<th>DS</th>
<th>Geom. Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.552</td>
<td>0.390</td>
<td>0.195</td>
<td>0.455</td>
<td>0.195</td>
<td>0.682</td>
<td>0.369</td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>0.311</td>
<td>0.506</td>
<td>0.311</td>
<td>0.467</td>
<td>0.467</td>
<td>0.467</td>
<td>0.413</td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>0.417</td>
<td>0.298</td>
<td>0.238</td>
<td>0.417</td>
<td>0.477</td>
<td>0.358</td>
<td>0.358</td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>0.301</td>
<td>0.481</td>
<td>0.301</td>
<td>0.361</td>
<td>0.421</td>
<td>0.602</td>
<td>0.398</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>0.298</td>
<td>0.417</td>
<td>0.298</td>
<td>0.179</td>
<td>0.417</td>
<td>0.834</td>
<td>0.363</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.364</td>
<td>0.411</td>
<td>0.265</td>
<td>0.356</td>
<td>0.377</td>
<td>0.564</td>
<td>0.380</td>
<td></td>
</tr>
<tr>
<td>Pitch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.195</td>
<td>1.170</td>
<td>0.260</td>
<td>0.715</td>
<td>1.170</td>
<td>0.682</td>
<td>0.569</td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>1.089</td>
<td>1.089</td>
<td>0.156</td>
<td>0.506</td>
<td>1.440</td>
<td>1.945</td>
<td>0.800</td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>0.417</td>
<td>0.953</td>
<td>1.311</td>
<td>1.132</td>
<td>1.251</td>
<td>0.834</td>
<td>0.922</td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>1.263</td>
<td>0.842</td>
<td>0.902</td>
<td>0.541</td>
<td>1.624</td>
<td>0.602</td>
<td>0.893</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>0.834</td>
<td>0.953</td>
<td>0.596</td>
<td>0.536</td>
<td>1.728</td>
<td>0.358</td>
<td>0.735</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.622</td>
<td>0.995</td>
<td>0.491</td>
<td>0.653</td>
<td>1.427</td>
<td>0.751</td>
<td>0.773</td>
<td></td>
</tr>
<tr>
<td>Sway</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.260</td>
<td>0.357</td>
<td>0.162</td>
<td>0.357</td>
<td>0.162</td>
<td>0.227</td>
<td>0.242</td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>0.272</td>
<td>0.389</td>
<td>0.311</td>
<td>0.117</td>
<td>0.311</td>
<td>0.117</td>
<td>0.228</td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>0.179</td>
<td>0.179</td>
<td>0.536</td>
<td>0.358</td>
<td>0.536</td>
<td>0.417</td>
<td>0.333</td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>0.241</td>
<td>0.301</td>
<td>0.180</td>
<td>0.361</td>
<td>0.662</td>
<td>0.481</td>
<td>0.338</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>0.179</td>
<td>0.358</td>
<td>0.179</td>
<td>0.775</td>
<td>0.596</td>
<td>0.536</td>
<td>0.376</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.222</td>
<td>0.306</td>
<td>0.245</td>
<td>0.334</td>
<td>0.403</td>
<td>0.310</td>
<td>0.298</td>
<td></td>
</tr>
<tr>
<td>Roll</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.260</td>
<td>0.390</td>
<td>0.227</td>
<td>0.455</td>
<td>0.487</td>
<td>0.975</td>
<td>0.413</td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>0.350</td>
<td>0.506</td>
<td>0.545</td>
<td>0.272</td>
<td>0.545</td>
<td>0.661</td>
<td>0.460</td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>0.477</td>
<td>0.298</td>
<td>0.358</td>
<td>0.655</td>
<td>0.715</td>
<td>0.417</td>
<td>0.464</td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>0.301</td>
<td>0.301</td>
<td>0.180</td>
<td>0.541</td>
<td>0.662</td>
<td>0.842</td>
<td>0.413</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>0.238</td>
<td>0.477</td>
<td>0.179</td>
<td>1.192</td>
<td>0.477</td>
<td>0.715</td>
<td>0.450</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.315</td>
<td>0.385</td>
<td>0.270</td>
<td>0.554</td>
<td>0.569</td>
<td>0.695</td>
<td>0.439</td>
<td></td>
</tr>
<tr>
<td>Condition A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.260</td>
<td>0.195</td>
<td>0.195</td>
<td>0.162</td>
<td>0.292</td>
<td>0.292</td>
<td>0.227</td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>0.311</td>
<td>0.233</td>
<td>0.428</td>
<td>0.195</td>
<td>0.389</td>
<td>0.311</td>
<td>0.300</td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>0.179</td>
<td>0.238</td>
<td>0.298</td>
<td>0.358</td>
<td>0.358</td>
<td>0.358</td>
<td>0.289</td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>0.241</td>
<td>0.241</td>
<td>0.301</td>
<td>0.180</td>
<td>0.541</td>
<td>0.241</td>
<td>0.273</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>0.238</td>
<td>0.238</td>
<td>0.298</td>
<td>0.417</td>
<td>0.298</td>
<td>0.417</td>
<td>0.309</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.242</td>
<td>0.228</td>
<td>0.295</td>
<td>0.243</td>
<td>0.366</td>
<td>0.318</td>
<td>0.278</td>
<td></td>
</tr>
</tbody>
</table>
### Table H.3  
Critical amplitudes for sway signals (milli-g), from experiment 3.

Mean values shown are geometric means.

<table>
<thead>
<tr>
<th>Noise Cond.</th>
<th>Freq</th>
<th>Subject</th>
<th>Geom. Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surge</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>DG</td>
<td>1.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PG</td>
<td>3.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GG</td>
<td>5.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MN</td>
<td>6.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PR</td>
<td>5.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DS</td>
<td>4.35</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td></td>
<td>3.10</td>
</tr>
<tr>
<td>Pitch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>DG</td>
<td>3.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PG</td>
<td>4.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GG</td>
<td>5.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MN</td>
<td>3.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PR</td>
<td>6.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DS</td>
<td>6.52</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td></td>
<td>3.69</td>
</tr>
<tr>
<td>Sway</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>DG</td>
<td>3.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PG</td>
<td>2.61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GG</td>
<td>6.52</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MN</td>
<td>7.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PR</td>
<td>10.44</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DS</td>
<td>5.65</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td></td>
<td>7.32</td>
</tr>
<tr>
<td>Roll</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>DG</td>
<td>8.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PG</td>
<td>1.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GG</td>
<td>8.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MN</td>
<td>9.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PR</td>
<td>6.52</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DS</td>
<td>11.74</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td></td>
<td>7.46</td>
</tr>
<tr>
<td>Condition A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>DG</td>
<td>2.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PG</td>
<td>3.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GG</td>
<td>4.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MN</td>
<td>4.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PR</td>
<td>3.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DS</td>
<td>5.22</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td></td>
<td>2.37</td>
</tr>
</tbody>
</table>
Table H.4  Critical amplitudes for roll signals (°/s), from experiment 3.  
Mean values shown are geometric means.

<table>
<thead>
<tr>
<th>Noise Cond.</th>
<th>Freq</th>
<th>Subject</th>
<th>Geom. Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>DG</td>
<td>PG</td>
</tr>
<tr>
<td>Surge</td>
<td>0.2</td>
<td>0.747</td>
<td>0.325</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.377</td>
<td>0.377</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.605</td>
<td>0.495</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.500</td>
<td>0.778</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.389</td>
<td>0.556</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>0.506</td>
<td>0.483</td>
</tr>
<tr>
<td>Pitch</td>
<td>0.2</td>
<td>0.520</td>
<td>0.390</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.415</td>
<td>0.415</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.550</td>
<td>0.935</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.556</td>
<td>0.945</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.500</td>
<td>0.667</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>0.505</td>
<td>0.562</td>
</tr>
<tr>
<td>Sway</td>
<td>0.2</td>
<td>0.617</td>
<td>0.455</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.528</td>
<td>1.169</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.660</td>
<td>0.605</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.667</td>
<td>0.778</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.778</td>
<td>0.889</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>0.645</td>
<td>0.740</td>
</tr>
<tr>
<td>Roll</td>
<td>0.2</td>
<td>1.007</td>
<td>1.624</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>1.847</td>
<td>0.829</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.935</td>
<td>0.880</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>1.223</td>
<td>1.278</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>1.334</td>
<td>0.667</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>1.232</td>
<td>1.002</td>
</tr>
<tr>
<td>Condition A</td>
<td>0.2</td>
<td>0.422</td>
<td>0.552</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.490</td>
<td>0.452</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.385</td>
<td>0.440</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.445</td>
<td>0.333</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.445</td>
<td>0.723</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>0.436</td>
<td>0.484</td>
</tr>
</tbody>
</table>
References


Grant, P.R. (1986). Motion Characteristics of the UTIAS Flight Research Simulator Motion-Base. UTIAS Technical Note No. 261, University of Toronto.


This report describes an investigation of human sensitivity to whole-body motion. Specifically, it discusses the ability of human subjects to detect a sinusoidal motion signal superimposed on a background of random motion. The purpose of the study was to determine the conditions in which motion cues are masked by concurrent random motion. The results have applications to flight simulation, and will also be of interest to other researchers working in the area of human perceptual performance. It is proposed that for the situation under study, motion perception is a signal-in-noise detection process which can be modelled using signal detection theory. A brief review of signal detection theory is provided. Three ideal detectors adapted from the literature on auditory perception are proposed as potential models for motion perception. Three motion perception experiments were run. In the first, a rating procedure was used to obtain receiver operating characteristic (ROC) curves for human subjects detecting sinusoidal motion in a background of low power broadband random motion. A good fit to the data was obtained using ROC curves based on Gaussian distributions of signal and noise. The second experiment used a 2-alternative forced choice task to determine the detectability of sinusoidal motion in a variety of noise conditions. The results show that detectability can be expressed as a function of signal-to-noise ratio, and that sinusoidal motion is masked primarily by noise components which are near the signal frequency. The third experiment tested the extent to which noise on one axis masks a signal on another axis. Inter-axis effects were found to be small, but significant. All three experiments provided an estimate of the slope of the psychometric curve. Of the three ideal detectors considered, the energy detector agrees best with the experimental data. The data are compared in detail with the predictions of the energy detector. A simplified method for estimating signal detectability in arbitrary conditions is presented. Finally, the implications of the results for flight simulation are discussed.

Available copies of this report are limited. Return this card to UTIAS, if you require a copy.
This report describes an investigation of human sensitivity to whole-body motion. Specifically, it discusses the ability of human subjects to detect a sinusoidal motion signal superimposed on a background of low power broadband random motion. The purpose of the study was to determine the conditions in which motion cues are masked or hidden by concurrent random motion. The results have applications to flight simulation, and will also be of interest to other researchers working in the area of human perceptual performance. It is proposed that for the situation under study, motion perception is a signal-in-noise detection process which can be modelled using signal detection theory. A brief review of signal detection theory is provided. Three ideal detectors adapted from the literature on auditory perception are proposed as potential models for motion perception. Three motion perception experiments were run. In the first, a rating procedure was used to obtain receiver operating characteristic (ROC) curves for human subjects detecting sinusoidal motion in a background of low power broadband random motion. A good fit to the data was obtained using ROC curves based on Gaussian distributions of signal and noise. The second experiment used a 2-alternative forced choice task to determine the detectability of sinusoidal motion in a variety of noise conditions. The results show that detectability can be expressed as a function of signal-to-noise ratio, and that sinusoidal motion is masked primarily by noise components which are near the signal frequency. The third experiment tested the extent to which noise on one axis masks a signal on another axis. Inter-axis effects were found to be small, but significant.

All three experiments provided an estimate of the slope of the psychometric curve. Of the three ideal detectors considered, the energy detector agrees best with the experimental data. The data are compared in detail with the predictions of the energy detector. A simplified method for estimating signal detectability in arbitrary conditions is presented. Finally, the implications of the results for flight simulation are discussed.

Available copies of this report are limited. Return this card to UTIAS, if you require a copy.

University of Toronto, Institute for Aerospace Studies (UTIAS) 4925 Dufferin Street, Downsview, Ontario, Canada, M3H 5T6

MASKING OF MOTION CUES BY RANDOM MOTION: COMPARISON OF HUMAN PERFORMANCE WITH A SIGNAL DETECTION MODEL

Greig, Glenn Lewis

1. Motion cues 2. Motion signal detection 3. Simulator motion

I. UTIAS Report No. 313 II. Greig, Glenn Lewis

Available copies of this report are limited. Return this card to UTIAS, if you require a copy.