ACCURACY BOUNDS FOR NORMAL-INCIDENT
ACOUSTIC STRUCTURE ESTIMATION

by

KENNETH B. THERIAULT

B.S., Massachusetts Institute of Technology
(1968)

M.S., E. Eng., Massachusetts Institute of Technology
(1975)

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
and the
WOODS HOLE OCEANOGRAPHIC INSTITUTION
August, 1977

Signature of Author

M.I.T./W.H.O.I. Joint Program in Oceanographic Engineering

Certified by

Thesis Supervisor

Accepted by
Chairman, M.I.T./W.H.O.I. Joint Committee for
Oceanographic Engineering
ACCUARCY BOUNDS FOR NORMAL-INCIDENCE ACOUSTIC
STRUCTURE ESTIMATION

by

Kenneth B. Theriault


ABSTRACT

Determination of the structure of a medium from normal-incidence acoustic reflection data is a basic problem in fields as diverse as medical technology and the earth sciences; this research examines the accuracy with which quantitative structure estimates can be made from noise-corrupted measurements of reflected energy. Two classes of simple physical models, which exclude geometrical spreading and attenuation, are developed: one in which the properties of the medium change continuously with depth, and one in which they change discretely. Given these reasonable models, estimation accuracy is studied by computing a statistical lower bound on estimator performance, the Cramer-Rao bound, for three cases of interest. (1) The bound is computed for the estimation of unknown, non-random reflection coefficients in a medium containing only discrete reflectors; special attention is given to the one- and two-reflector situations. The bound's ability to predict estimator performance is demonstrated, as is the inadequacy of a particular ad-hoc estimation method based on the Wiener-Levinson algorithm of stochastic filtering theory. (2) The bound is developed for estimation in a continuous medium whose structure (acoustic impedance, for example) parametrized by a set of unknown, non-random coefficients, and for which the reflection response may be computed in closed form. The problem of estimating the parameters of a single, isogradient velocity layer of known depth is studied in detail. It is demonstrated that one can identify the parameters of such a layer from normal-incidence measurements given an appropriate source and experiment geometry. (3) A unique extension of some known results in random process estimation is used to derive a pointwise bound for estimation in a continuous medium whose structure (reflection coefficient density) is a random process. Again we give special consideration to the problem
of identifying a single isolated layer structure. We demonstrate that for a weakly scattering structure, estimation accuracy is independent of the mean or nominal structure.

Thesis Supervisor: Arthur B. Baggeroer

Title: Associate Professor of Electrical Engineering
Associate Professor of Ocean Engineering
ACKNOWLEDGEMENTS

Few people can learn, or perform research, in solitude; I am not an exception to this rule. The daily give-and-take with two of my fellow graduate students, Dick Kline and Tom Marzetta, has been invaluable in helping me clarify my understanding of this problem. My thesis readers, Prof. Ira Dyer and Prof. Timothy Johnson, of M.I.T., and Dr. Robert Porter of W.H.O.I., have provided the detailed criticism so necessary to a research effort. Prof. Johnson deserves special distinction for service above and beyond that normally expected of a thesis reader, when he allowed me to use him as a de facto thesis advisor, during Prof. Baggeroer’s six-month Italian sabbatical.

I have known Prof. Arthur Baggeroer since my return to M.I.T. in 1972, both as a friend, and as a thesis advisor in two theses. It is because of his insistence that this work is as readable and as accurate as it is. The educational experience I have had in graduate school is due in large measure to his efforts.

While performing this research, I was supported by a variety of sources: a National Science Foundation Graduate Fellowship, a Vinton Hayes Graduate Fellowship, and, as a Research Assistant, by the Office of Naval Research.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title Page</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Abstract</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Table of Contents</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>CHAPTER 1:</td>
<td>INTRODUCTION AND BACKGROUND</td>
<td>7</td>
</tr>
<tr>
<td>1.1</td>
<td>Motivation and Problem Statement</td>
<td>7</td>
</tr>
<tr>
<td>1.2</td>
<td>Survey of Existing Estimation Methods</td>
<td>16</td>
</tr>
<tr>
<td>1.3</td>
<td>Preview of Later Chapters</td>
<td>28</td>
</tr>
<tr>
<td>CHAPTER 2:</td>
<td>MODELS FOR ELASTIC MEDIA</td>
<td>34</td>
</tr>
<tr>
<td>2.1</td>
<td>Introduction and Assumptions</td>
<td>34</td>
</tr>
<tr>
<td>2.2</td>
<td>Continuous Model</td>
<td>38</td>
</tr>
<tr>
<td>2.3</td>
<td>Discrete Model</td>
<td>55</td>
</tr>
<tr>
<td>CHAPTER 3:</td>
<td>BOUNDS AND ESTIMATORS FOR A DISCRETE MEDIUM</td>
<td>70</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>70</td>
</tr>
<tr>
<td>3.2</td>
<td>The Cramer-Rao Bound</td>
<td>72</td>
</tr>
<tr>
<td>3.3</td>
<td>An Optimal Estimator</td>
<td>93</td>
</tr>
<tr>
<td>3.4</td>
<td>The PEF Estimator</td>
<td>102</td>
</tr>
<tr>
<td>3.5</td>
<td>Monte Carlo Tests</td>
<td>111</td>
</tr>
<tr>
<td>CHAPTER 4:</td>
<td>BOUNDS FOR PARAMETRIC STRUCTURE ESTIMATION IN CONTINUOUS MEDIA</td>
<td>122</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>122</td>
</tr>
<tr>
<td>4.2</td>
<td>Theoretical Results</td>
<td>126</td>
</tr>
<tr>
<td>4.3</td>
<td>Example: Structure Estimation for a Bottom Sediment Layer</td>
<td>130</td>
</tr>
<tr>
<td>CHAPTER 5:</td>
<td>BOUNDS FOR RANDOM STRUCTURE ESTIMATION IN CONTINUOUS MEDIA</td>
<td>167</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>167</td>
</tr>
<tr>
<td>5.2</td>
<td>The Cramer-Rao Bound</td>
<td>173</td>
</tr>
<tr>
<td>5.3</td>
<td>A Model for the Reflection Response of Weakly Scattering Media</td>
<td>188</td>
</tr>
<tr>
<td>5.4</td>
<td>Cramer-Rao Bound Computation for a Weakly Scattering Medium</td>
<td>198</td>
</tr>
<tr>
<td>APPENDIX A:</td>
<td>NUMERICAL COMPUTATION OF THE FISHER INFORMATION MATRIX</td>
<td>223</td>
</tr>
<tr>
<td>APPENDIX B:</td>
<td>PARTIAL DERIVATIVES OF A FUNCTIONAL IN TERMS OF THE VARIATIONAL DERIVATIVE</td>
<td>229</td>
</tr>
<tr>
<td>APPENDIX C:</td>
<td>REGULARITY PROPERTIES OF THE REFLECTION RESPONSE OF A RANDOM MEDIUM</td>
<td>232</td>
</tr>
<tr>
<td>C.1</td>
<td>Introduction</td>
<td>232</td>
</tr>
<tr>
<td>C.2</td>
<td>Constraints on g(ξ)</td>
<td>233</td>
</tr>
<tr>
<td>C.3</td>
<td>Continuity of the Response and Partial Derivative</td>
<td>237</td>
</tr>
<tr>
<td>C.4</td>
<td>Existence and Continuity of the Variational Derivative</td>
<td>244</td>
</tr>
<tr>
<td>References</td>
<td>259</td>
<td></td>
</tr>
<tr>
<td>Biographical Note</td>
<td>264</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION AND BACKGROUND

1.1 Motivation and Problem Statement

A common method of non-invasively studying the internal structure of a medium is to illuminate the medium with a source of wave energy, either acoustic or electromagnetic, and to observe the energy scattered back from the medium's interior. The observation is then analyzed to infer the medium's structure, a procedure often referred to as solution of the "inverse problem" (inverse to solution for the medium's reflection response given its structure). Such methods are used when direct, physical examination of medium structure is uneconomical, impossible, or undesirable. The "structure" which we may identify with such techniques is, of course, limited to those medium properties, such as acoustic impedance, which affect or control wave propagation within the medium.

Our particular concern will be with acoustic wave probing, although the results we obtain will be applicable to electromagnetic probing as well. Acoustic probing is important in fields as diverse as the earth sciences and medicine. In pure earth science, a very general problem is the global determination of the structure of the earth's crust [3]; on a more applied plane, wide-area structural studies are crucial for oil and mineral exploration [4]. Applications in medical technology include the detection and evaluation of
arteriosclerotic plaque in major arteries [1], and of pulmonary emboli [2]. The specific physical problem which motivates this research is from the earth sciences, that of determining seafloor sediment structure, e.g., acoustic propagation velocity as a function of depth in the sediment layer. A knowledge of sediment layer structure is required for propagation studies in underwater communication links and active sonar systems which utilize near-bottom ducting, or upward-refractive ray paths involving the bottom [5], [26], for the selection of offshore facility construction sites [6], [7], as well as for purely geophysical purposes.

Our objective in this research is to examine the accuracy with which quantitative structural determinations can be made from noise-corrupted measurements of reflected energy. In the absence of noise, it is generally possible to invert reflection observations to obtain medium structure exactly; the presence of observation noise, however, limits the accuracy with which we may estimate medium structure. We will not examine the estimation problem per se, in that we will not emphasize estimator design, or dwell on the performance (estimate variance) of specific estimators. Rather, we will pursue a more general and more useful course of action: we will develop, for a class of reasonable physical models for both media and measurement process, lower bounds on the accuracy with which the structure of a medium can be estimated (lower bounds on the estimate variance), bounds which are
independent of the actual estimator employed, depending solely upon the models used for medium response and observation noise. Insofar as the models are accurate, such lower bounds provide: (1) an ultimate constraint on our ability to extract structural information from noisy reflection observations, and hence a benchmark against which we can evaluate the performance of any estimator; (2) qualitative physical insight into the way in which medium structure itself affects estimation accuracy; and (3) approximate performance predictions for actual estimators when the noise level is sufficiently low. This amount of information is simply not available from an analysis of a particular estimator, hence our emphasis on general performance bounds rather than the exact performance of specific estimators.

It is surprising that, despite the bulk of work published on the general acoustic structure estimation problem, no authors have addressed the topic of estimation accuracy, although some have considered structure identification schemes for which performance could easily have been calculated (see, for instance, [8]). In the seismic area, which is responsible for the bulk of the published research on the practical solution of the general reflection problem, the question of estimator performance, if considered at all, is usually discussed in a deterministic context, in terms of resolving two closely-spaced reflectors in the medium [9]. Even the theoretical literature dealing with the identification of a
distributed parameter system (of which medium structure
identification is a specific example) contains few results on
estimator performance [10]. Our research, then, represents a
unique and significant contribution, both to the specific,
practical problem of sediment layer structure estimation, and
also, as a case study, to the general theoretical problem.

To make the structure identification problem more con-
crete, and to place estimation accuracy in context within the
problem, we will consider a brief, hypothetical, and very
qualitative example. Suppose that the acoustic source, and
the receiver, are in contact with, or just below the surface
of, the medium of interest, illustrated schematically in
Figure 1.1a. The source emits infinite parallel plane
(pressure) waves, and the medium's properties (the acoustic
impedance, a(z)) are functions of depth in the medium, z,
alone. These assumptions reduce our nominally three-dimen-
sional medium identification problem to a one-dimensional wave-
guide identification problem. Figure 1.1b shows the acoustic
impedance profile for our hypothetical (waveguide) medium.
There is no response from the region where a(z) is constant, a
discrete reflection from the step change in a(z) at z = z₁, and
distributed scattering from the region below z₁ where
there is a continuous change in a(z). We have also plotted
the reflection coefficient density, g(ξ), as a function of
two-way travel time in the medium; g(ξ) is related to a(ξ)
(acoustic impedance as a function of two-way travel time in
Figure 1.1 -- (a) Schematic view of a medium to be identified; (b) Acoustic impedance profile of the medium as a function of depth, \( z \); (c) Reflection coefficient density corresponding to \( a(z) \), \( g(\xi) \), as a function of two-way travel time in the medium, \( \xi \).
the medium) by

\[ g(\xi) = \frac{a'(\xi)}{2a(\xi)}. \]  

(1.1)

Figure 1.1c shows the \( g(\xi) \) corresponding to the \( a(z) \) of Figure 1.1b.

The acoustic impedance is a very direct function of the physical properties of the medium, being the product of medium density and acoustic propagation velocity. The reflection process, however, is much more closely related to the reflection coefficient density; to a first approximation, the reflection response of the medium is the convolution of \( g(\xi) \) with the source waveshape (the actual relationship is a non-linear functional of \( g(\xi) \)). Figure 1.2b shows the noiseless response of the medium of Figure 1.1 to a short pulse of the form \( t \exp(-\psi t) \); \( g(\xi) \) is reproduced in Figure 1.2a for comparison. As well as the discrete reflection associated with the impulse in \( g(\xi) \) (the step change in \( a(z) \)), and the distributed scattering from the continuously-varying region of \( g(\xi) \) \( a(z) \), we also see the effect of the interface between the medium and the atmosphere, energy reverberating between the free surface and the deep structure. The noise-corrupted reflection response is shown in Figure 1.2c. As well as accounting for the actual presence of corrupting noise, which places an essential limit on the estimation accuracy we can attain in practice, observation noise represents, in a crude way, the
Figure 1.2 -- (a) Reflection coefficient density; (b) Noiseless reflection response; (c) Noisy reflection response; (d) Hypothetical estimate of acoustic impedance as a function of two-way travel time, showing estimate standard deviation, $\sigma(t)$. 
inaccuracies of the model, and also serves the useful theoretical purpose of preventing any attempts to use estimation schemes which are unstable or singular. Some of the estimators we will review in the next section are suspect because they have not been designed under the assumption of noise-corrupted measurements, i.e., they work only for very high signal-to-noise ratios.

Figure 1.2d shows an estimate of acoustic impedance derived from the noise-corrupted response (by an estimator as hypothetical as \( a(z) \) itself); the standard deviation of the estimate is indicated by the dashed lines surrounding the solid line plot of the estimate itself. The purpose of the estimator, in this case, is to reconstruct the acoustic impedance (Figure 1.1b) from the noisy reflection response (Figure 1.2c); our objective is to develop methods by which we can compute a lower bound, \( \sigma_{\text{BND}}^2(t) \), on the estimate variance, \( \sigma^2(t) \), i.e., we want to define \( \sigma_{\text{BND}}^2(t) \) such that

\[
\sigma^2(t) \geq \sigma_{\text{BND}}^2(t). \tag{1.2}
\]

In particular, we will define bounds which are substantially estimator-independent, so that our \( \sigma_{\text{BND}}^2(t) \) will limit the performance of an arbitrary estimator.

Before continuing to the next section, we note that we will consider only "normal-incidence" reflection observations, in which source and receiver are in close horizontal proximity,
so that we have available reflection information from only one angle of incidence \((0^\circ)\). In "wide-angle" or array methods, source and receiver are separated so that we have reflection or refraction information from multiple angles. Although these methods have been employed in the medium identification problem with considerable success, \([26], [11]-[13]\), we will not treat them here. The simpler normal-incidence case provides a substantially challenging problem, and we leave the even more complicated wide-angle analysis to future investigators.

We emphasize from the outset that the normal-incidence method provides limited information about the medium. In particular, the only medium property which can be measured is acoustic impedance, the product of density and propagation velocity, or the related quantity, reflection coefficient density. Moreover, we can determine acoustic impedance not as a function of depth in the medium, but rather as a function of the time required for the source energy to penetrate to a given depth and return to the surface ("two-way travel time"). Without an independent measurement of propagation velocity or density, it is impossible to transform this travel time into a spatial coordinate.

In closing this section, we note that the process of estimating medium structure from normal-incidence reflection observations is sometimes termed "deconvolution." Unfortunately, it is also used to describe various other processing
tasks, such as the recovery of the medium impulse response from the observation (source waveshape compression), or the suppression of free surface multiples (particularly in marine seismic applications). We consider, therefore, that "deconvolution" is an insufficiently specific term, and hence have avoided its use.

In this section, we have motivated and described the general objectives of this research, and also indicated its uniqueness. In the remaining two sections of this chapter, we will survey the existing methods of solution of the structure determination problem, and then give a preview of the topics to be considered in later chapters.

1.2 Survey of Existing Estimation Methods

In this section, we will discuss the major categories of estimation methods which have been used, or proposed, for the medium identification problem. The choice of an estimator is not, of course, the point of this research, so that this section will serve primarily as a survey of various modelling efforts, and only secondarily as a survey of estimator structures. We also hope to indicate the range of approaches which have been applied to this problem, and the degree to which each has succeeded in producing usable information about media structures.

Attempts to invert normal-incidence reflection observations of a medium to obtain a structural description have
ranged from the very simple and ad hoc to the very complicated and rigorous, and from the very qualitative to the very quantitative. We have attempted to present a codification scheme for such estimation methods in Figure 1.3; this scheme is certainly not hard and fast, but will provide a framework for classifying estimation methods for the purposes of this discussion. Figure 1.3 does not, of course, indicate the relative number of estimation methods within any of the classification categories. In particular, virtually all developed methods are based on the assumption of noise-free observations (they are inherently deterministic): their accuracy, stability and sensitivity in the presence of noise are not considered in their design.

The major division indicated in Figure 1.3 is that between quantitative and qualitative methods. While we are concerned in this research exclusively with quantitative methods of determining the structure of media, a substantial amount of the total effort in this area, particularly in the seismic field, is devoted to qualitative methods. We will discuss these briefly, for the sake of completeness. Qualitative analyses of reflection response are essentially limited to the detection and accurate location of discrete reflectors within the medium. These discrete reflectors are associated with regions where the acoustic impedance changes abruptly, such as the step change in the $a(z)$ of Figure 1.1b at $z_1$. Regions where the gradient of $a(z)$ is large but finite will
Medium structure identification problem

Qualitative methods

Quantitative methods

Visual analysis of reflection response

Partial description of structure

Complete description of structure

Noise corrupted

Deterministic

Noise corrupted

Deterministic

Estimation of parameters of a few prominent features

"Bright spot" methods

(None known)

Fitting of data to a model for total medium response

Figure 1.3 -- A classification scheme for methods which have been applied to the medium structure identification problem.
also produce a reflection response which may be approximated as a discrete reflection.

Characterization of a continuously-varying medium in terms of its most prominent "discrete" reflectors alone obviously obliterates the details of the medium's structure. Such simple analyses are inadequate, for example, to provide the continuous velocity information required to describe ducting in horizontal, near-bottom acoustic propagation [5], where the gradient in propagation velocity causes upward refraction. Discrete-reflector characterizations are used quite successfully in seismic prospecting, however, where a continuous determination of properties is not required. Although considerable effort is expended to process normal-incidence data for qualitative analysis, the usual objectives are to perform noise suppression, source pulse compression, and generally "clean up" the record so that a human interpreter can pick out the discrete reflectors. While this processing is often quite sophisticated (see [4] and [27] for overviews of current processing methods), it does not produce a truly quantitative description of the entire medium.

Of course, insofar as it produces an estimate of the arrival time of the return from a "discrete" reflector, the qualitative method does provide quantitative results. What is missing is an estimate of the strength of the reflection, an estimate of the reflection coefficient of the discrete reflector. Such estimates are provided, for isolated portions of
the response, by a relatively new method commonly known as "bright spot" [14]-[16], which we have classified in Figure 1.3 as a quantitative technique, albeit one which provides only a partial description of medium structure. We note, moreover, that this is essentially a deterministic approach. In this method, the received signal is subjected to a time-varying gain, tailored to compensate for geometrical spreading, earth attenuation, and reflection losses in succeeding layers. The time-varying gain ideally applies a correction which makes the qualitatively-displayed reflection coefficient (the amplitude of a waveform plot) correspond to its true value. This enables visual detection of reflection coefficients, and phase shifts which are indicative of rock/gas or rock/oil interfaces. More intensive and sophisticated processing is applied to shorter segments of the response to obtain true quantitative estimates of reflection coefficients associated with specific structures (gas sands, for example).

In the hands of a skilled interpreter, this simple method has enjoyed considerable success in locating petroleum-bearing structures. This success, however, depends upon the empirical manipulation of the parameters which control the time-varying gain. Recent studies [16], [17] have shown that "bright spot" is quite sensitive in this regard, and can produce misleading results if applied uncritically.

We now consider three quantitative methods which, unlike bright spot, provide estimates of the complete structure
of a medium, rather than of just a single feature or section. These methods are based on deterministic models for the response of the medium (they ignore the presence of observation noise), and proceed by simply fitting the observed reflection response to the response model.

The first method [18] is based on a medium modelled as a collection of homogeneous, equal travel-time layers (this model is discussed in detail in Chapter 2). By allowing a large number of such layers, we can approximate a continuous acoustic impedance profile, and its response, to any desired accuracy, so that this model is capable of representing considerably more detail than those which include only prominent reflectors. Briefly, a state-variable representation of the model is constructed, and used to define a deterministic, algebraic processing of the received signal. That is, the model is used to express the (noiseless) received signal in terms of the model parameters (the set of reflection coefficients for the layer interfaces). The experimental received signal is then inverted to obtain reflection coefficients by solution of these equations using the observed data.

The cited article presents no example of such an inversion in the presence of noise, nor has a promised future paper treating the noisy case appeared. The method, however, has a hint of statistical optimality to it: the authors suggest the fitting of a model to the observed data by solving a set of equations relating observed and noiseless response
values. What is lacking, however, is an appropriate specification of the sense in which the parametric model is to be fitted to the data: minimization of the mean-square error between the observed and postulated system responses, or minimization of the maximum difference between the two, are examples of reasonable specifications.

In addition, the processing scheme is not presented in a usable, algorithmic form. The authors simply indicate, by means of a few examples, the deterministic relationship between successive received data samples. Such an unsystematic computation is virtually useless as a practical analysis tool.

A second method which provides a complete, quantitative, but deterministic structure estimate is one based on the close mathematical relationship between scattering theory and linear filtering theory [19]. We have termed this the prediction-error filter (PEF) method, since the procedure involved is identical to that required to compute the prediction-error filter of linear filtering theory. This method has been developed by several authors [20]–[22], and mentioned implicitly by at least one other [23]. It has seen extensive practical application in the estimation of discrete reflector models for vocal tract shape [20], [21] (predictive encoding of speech). This application is similar to our medium identification problem, except that, in the vocal tract case, the receiver is on the opposite side of the medium from the source, and hence observes transmitted rather than
reflected energy. Recently, it has been shown that the PEF method is applicable not only to a noiseless, lossless medium model consisting of discrete equal-travel time layers, but also to a model whose properties vary continuously [22].

The PEF method of inverting reflection observations is exact in the absence of noise, and provides a quantitative description of all the scatterers in the medium. Moreover, it can be expressed in a convenient computational form: the discrete-layer estimator is based on the solution of the normal equations of linear filtering theory by the well-known Wiener-Levinson algorithm; the continuous estimator, on solution of a time-varying Wiener-Hopf integral equation, for which a well-defined series solution, the Neumann series, exists. Indeed, the Wiener-Levinson algorithm is exactly the algorithmic procedure which was lacking in [18].

One disadvantage of the PEF method is that the estimator requires the medium's response to an impulse source. Sources are generally not impulsive, and the problem of removing the non-impulsive character of the source waveshape from the observation is generally a difficult one in its own right [24]. This limits the practical utility of the PEF method when the source is non-impulsive.

The most telling criticism of the PEF method is that it is theoretically ill-defined when observation noise is present, which, as a practical matter, restricts its use to situations where the signal to noise ratio is very large. This obtains
because the method is based on the fact (proved in [22]) that the reflection response of a one-dimensional, lossless medium due to an impulse source has a particular mathematical property: it is a positive-definite function (i.e., the class of functions which can be the impulse reflection response of a medium is a restricted class). The addition of noise to the response destroys its positive-definite character and causes the estimation algorithm to produce non-physical estimates (reflection coefficients greater than 1, or negative acoustic impedances). The method is successful for vocal tract shape estimation because the PEF estimator is applied to the correlation function of the transmission observation, which is positive definite by construction. In the reflection case, it must be applied to the raw data. We will discuss the discrete-layer form of the PEF method in greater detail in Chapter 3, where we demonstrate, by Monte Carlo simulation, its unstable behavior in the presence of noisy measurements.

The third method for complete deterministic medium structure estimation is based on an approximation for the reflection response. It has been shown [29] that, in the absence of free-surface multiples, and when the total change in acoustic impedance is small, the impulse reflection response of a medium can be approximated by the reflection coefficient density, \( g(\xi) \), defined in (1.1). This approximation is the basis for another deterministic estimator [7] which has been successfully used to extract near-bottom acoustic impedance
information when water-column multiples are not significant, and the noise level is low. The impulse response of the medium, \( h(t) \), is obtained by filtering the received signal to compress the source waveshape, and the estimate, \( \hat{a}(\xi) \), is formed by

\[
\hat{a}(\xi) = a(0) \exp\left(\frac{2}{\pi} \xi \int h(t) \, dt\right).
\]

(1.3)

The approximation \( h(t) \approx g(t) \) is a truncation of an infinite integral series solution for the response of the medium, the Bremmer series (described in more detail in [28], or Chapter 2). We will use this approximation in Section 5.5.

We now consider three estimators which provide, in the nomenclature of Figure 1.3, only a partial description of medium structure, but which are designed to deal with noise-corrupted observations. These are generally motivated by a common problem in (qualitative) marine seismology: the presence of water column multiples, energy reverberating between the ocean surface and ocean bottom, obscures the response of deeper structures, and must be removed so the structures can be visually detected and interpreted. Signal processing methods which remove the multiples commonly involve modelling the ocean bottom as a discrete reflector, characterized by a reflectivity coefficient and an arrival time. The reflectivity coefficient is only approximately the acoustic reflection coefficient of the bottom, since it includes the
effects of geometrical spreading, attenuation, and a non-planar sea surface, as well as the actual bottom reflection mechanism. Nonetheless, design of multiple-removal processors is a common, practical problem which requires a partial, quantitative, albeit phenomenological, characterization of a medium.

Reference [8] contains two parameter estimation methods for the design of multiple-removal processors. The method treated most extensively, and which is applied successfully to real data, involves parameter estimation for a single-reflector model by optimization of an ad hoc index of merit involving the received signal. Although this scheme apparently works well, at least for the limited data presented, it is not clear how the estimator could be extended to describe an entire medium, nor is it clear that this scheme has any statistically desirable properties (e.g., robustness, consistency).

The other identification method considered in [8] is maximum-likelihood parameter estimation. This approach is similar to the preceding, except that the ad hoc index of merit is now replaced by the statistically well-defined maximum-likelihood criterion. This is a statistically robust method, and allows representation of several discrete reflectors. Although the approach was rejected by the authors because of its computational complexity (in favor of the ad hoc index-of-merit scheme), it has the distinct advantage of being
statistically well-motivated, and easily generalized, at least theoretically, to the complete characterization of a medium. Also, computation of a Cramer-Rao performance bound for this estimator is so straightforward, it is surprising that the authors did not calculate it. We will examine a form of this estimator in Section 3.3.

The third estimator [25] is a particularly powerful adaptive method for removing water column multiples; it is quite unique, in that it employs a linear filter model for the reflection response of the ocean bottom. That is, rather than characterizing the bottom as a discrete reflector, described by a single reflection coefficient, the bottom is assumed to have a response which is frequency-dependent, and which may be modelled as a linear filter. The frequency response of the model filter is estimated from second-order (correlation) statistics of the observations. This technique gives a quantitative estimate of structure, is noise-tolerant and has been used successfully in practice. It is limited in scope, however, since it does not easily lend itself to structure determination for more than one layer or reflector structure.

In this section we have reviewed a number of techniques used to infer media structures from reflection acoustic observations. Our purpose was not to evaluate the estimators themselves so much as to gain an appreciation of the degree of sophistication required in the model to yield acceptable estimator behavior. We draw the following conclusions: to be
generally useful, the medium model must allow quantitative estimation of the entire medium in question, not simply the prominent features (layers or reflectors) of the medium; it is also essential that the observation model explicitly include noise, both to allow for its actual presence, and also to ensure a non-singular estimator structure.

1.3 Preview of Later Chapters

Before describing the structure and contents of the body of this research, let us restate our main objective: we seek to derive statistical bounds on the accuracy with which the structure of a medium can be estimated from noise-corrupted acoustic reflection measurements. The utility of such a bound is threefold: (1) it represents an ultimate performance benchmark which we may use to evaluate the performance of various estimators; (2) we may draw from the analytic behavior of the bound qualitative insight into the way in which medium structure itself affects estimation accuracy, i.e., why some structures are inherently more difficult to identify than others; and (3) the performance (estimate variance) for a specific estimator is generally impossible to compute directly, but, when the signal-to-noise ratio is large, estimator performance is generally well-approximated by the value of the lower bound. We have noted that this research is unique, not because we have approached the problem using performance bounds, but rather because, surprisingly, the
question of estimation accuracy for acoustic structure estimation has not been addressed in the past.

We have given practical motivation for the estimation problem, in terms of exploration seismology in particular, and theoretical motivation, expressing the problem as one of identification of a distributed-parameter system. By examining some conventional estimator structures, we have gained insight into the methods used to solve the estimation problem, and the pitfalls and limitations of the models involved.

Because of the extreme variety of situations which we will encounter, it is unlikely that one model, or one solution technique, will suffice to treat all cases of interest. While the basic assumptions will remain intact, we must be prepared to make approximate additional specifications as the need arises, to choose from the collection of mathematical models at our disposal. Chapter 2 contains the basic physical assumptions and the general medium model which we will use in this research. We have developed two models, a continuous model expressed in terms of partial differential equations, which serves to put us on firm theoretical ground, and a discrete model, based on a finite-element formulation, which is suitable for simulation purposes, and also allows us to directly model media characterized in terms of their prominent reflectors. We also demonstrate that the discrete model converges to the continuous in the continuum limit.
We consider the computation of estimation bounds, specifically the Cramer-Rao bound, for three distinct classes of media: those whose structure may be modelled in terms of a finite set of reflection coefficients (media with discrete, prominent reflectors); those whose structure is basically continuous, but may be parametrized by a finite set of coefficients; and those whose structure is so complicated that it may best be estimated as though it were a continuous random process. Sample profiles of acoustic impedance for each class are illustrated in Figure 1.4.

In Chapter 3, we treat the case of a medium composed of discrete reflectors. Since this case is substantially simpler than the other two, we use this medium model to present a comprehensive example of the estimation problem, so that the role of the Cramer-Rao bound may be viewed in context. In this regard, we examine two simple, but important examples which correspond to the problem of estimating ocean bottom reflectivity, and the estimation of the reflection coefficients of two closely-spaced layers. Parametric structure estimation for continuous media is considered in Chapter 4. As well as presenting the appropriate formulae, we consider an example based on the problem of estimating ocean bottom sediment structure. We demonstrate that computation of the Cramer-Rao bound provides us with information which is quite valuable in the design of an experiment (choice of source, for instance). In particular, we demonstrate that it is possible to identify
Figure 1.4 -- Acoustic impedance profiles corresponding to the models employed in Chapters 3-5 respectively: (a) a medium consisting of discrete layers; (b) a medium with a continuous, parametric profiles; (c) a medium whose structure is sufficiently complicated that it is best characterized as a random process.
the parameters of an isogradient layer structure, such as the one illustrated in Figure 1.4b, given only reflection data; it is commonly believed that such structures can be determined only by studies utilizing refractive (wide angle) observations. Estimation bounds for continuous structures which are typified as random processes are considered in Chapter 5. Such models are appropriate when the structure is sufficiently complicated that an adequate parametric description cannot be achieved using a reasonable number of parameters. This chapter entails a substantial amount of theoretical development (presented in Appendices), since it is necessary to extend some known results to derive a form of the Cramer-Rao bound appropriate for this problem. We obtain an expression for the Cramer-Rao bound, and evaluate it for an example based on the problem of identifying deep ocean bottom sediment structures. We obtain the interesting results that, when the structure is weakly scattering (neither internal reverberation nor water column multiples are significant components of the response) the estimation accuracy is substantially independent of the mean profile, depending only upon the bandwidth of the source and of the random portion of the profile.
and "dry" equipment (including the estimation algorithm), but also upon the implicit assumption which we have made throughout, that the model we have used for the layer to be identified corresponds to the true structure. If the layer structure does not correspond to the model we have employed, we have incurred an inherent global error in our analysis, rendering it, if not useless, then of questionable integrity.
2.1 Introduction and Assumptions

The objective of this research is to gain insight into those factors which influence medium identification accuracy, rather than to produce exact, generally applicable formulae. The choice of model requires, therefore, a compromise between the precision with which the model represents the physics of the problem, and the degree of insight into the physical problem which the model affords us. We require a model which retains the salient aspects of the physical problem, but is still simple enough to allow us to draw meaningful and incisive physical conclusions. A judicious choice of model is crucial to the successful solution of the problem at hand.

We believe that the model to be used in this research, although rather simple, contains the essential elements of the reflection problem. This model can be briefly, and qualitatively, described by four basic assumptions: (1) the medium to be identified is represented as an acoustic waveguide; (2) propagation within the medium is lossless, with no dissipative or frictional effects; (3) the interface between the medium and its surroundings is stress-free (a free surface); and, of course, (4) the observations are corrupted by additive broadband noise.
Our first two assumptions, restriction of the problem to one spatial dimension, and the exclusion of dissipative effects, are common ones, as we saw throughout our survey of estimators in Chapter 1. They are made, primarily, because they enable considerable mathematical simplification. The model which results is a reasonable one, however, and a fundamental understanding of the identification problem for this simplified model is necessary before considering more complicated ones. Also, it is expected that, for a given structure, identification is inherently less accurate in the presence of geometrical spreading and attenuation than in their absence. Given this, and recalling that we are developing lower bounds on accuracy, we conclude that the bounds we will derive for our simple model are still valid lower bounds when spreading and attenuation are present.

Our model is, of course, appropriate when geometrical spreading and attenuation are not important factors. For instance, we can successfully ignore spreading in cases where the source is not omnidirectional, but rather has a very narrow spatial beamwidth.

We saw, in the qualitative example of Figure 1.1, that reverberation due to the presence of the free surface can be a significant portion of the medium's reflection response. This is not surprising, since, for the most solid media, the density contrast between the medium and the atmosphere is
very large, leading to a reflection coefficient at the interface of magnitude $\approx 1$. This feature is thus one of the strongest in the medium. As we shall see in Chapter 3, the presence of such multiple energy can be used by a processor to improve estimate accuracy. We therefore include the free surface effect in our general model, although we will find that there are situations in which it may be ignored. It is possible to formulate the problem in a somewhat more general manner by allowing the medium's boundary to have an arbitrary reflection coefficient; the extension is straightforward, and will not be pursued here.

The inclusion of additive, white observation noise in our model serves three purposes: it models, very straightforwardly, the actual presence of broadband corrupting noise, such as the streamer tow noise present in marine seismic applications; it absorbs, in a very crude way, the errors which result from model inaccuracies, such as our exclusion of attenuation; and, it ensures that estimation schemes based on this model will not be singular, or yield unstable estimates.

Specific physical assumptions which support these
general, qualitative assumptions are:

1. The medium is an infinite, isotropic, linearly elastic half space.
2. The medium is lossless.
3. The properties of the medium, \( \lambda \), the elastic parameter, and \( \rho \), the density, are functions of depth in the medium, \( z \), alone.

4. All waves are infinite, parallel plane infinitesimal acoustic compressional waves, propagating along the \( z \) axis.

5. The interface between the medium and its surroundings, the plane \( z=0 \), is a free surface (the reflection coefficient there has magnitude 1).

It would be fortunate if these assumptions led to a mathematical model from which we could obtain the desired results analytically. This, however, is not the case. Despite the considerable simplifications made above, the model is still quite complicated, due to the extremely complex relationship between a medium's properties and its reflection response. It is not sufficient, therefore, to develop only a continuous model (partial-differential equation formulation) for this problem; we must also develop a discrete model (difference equation formulation) for numerical simulation of the medium's response.

We will not be concerned with detailed models for the source and receiver, or for their coupling to the medium. Such studies are beyond the scope of this treatment, and we therefore use the simplest possible representations for these devices.

In this Section, we have discussed the criteria for choosing a model, stated and explained the general assumptions which underly the model, and listed the specific, physical
assumptions which we will employ. In the remainder of this Chapter, we will develop the two versions of the medium model referred to above, a continuous model, based on a partial differential equation for wave propagation, and a discrete model, based on a difference equation, and will demonstrate the convergence of the discrete model to the continuous.

2.2 Continuous Model

Development of the continuous-medium model begins with a derivation of the equation of propagation (the wave equation) for the medium. Since the medium is linearly elastic and isotropic, and since the wave is infinitesimal, propagation may be described using the linear stress-strain relationship

\[ p(t, z) = -\lambda(z) \frac{\partial u(t, z)}{\partial z}, \]  

(2.1)

and the force balance equation

\[ \rho(z) \frac{\partial^2 u(t, z)}{\partial t^2} = \frac{\partial}{\partial z} (\lambda(z) \frac{\partial u(t, z)}{\partial z}), \]  

(2.2)

where

- \( p(t, z) \) = acoustic pressure
- \( u(t, z) \) = \( z \)-component of acoustic particle displacement
- \( \lambda(z) \) = bulk modulus (elastic parameter)
- \( \rho(z) \) = density
\[ z = \text{depth in the medium (positive down)} \]
\[ t = \text{time.} \]

Equations (2.1) and (2.2) are simplified versions of the most general forms (see, for instance, Chapter 2 of [30]), made possible by our assumptions that propagation is along the z-axis only, and that media properties are a function of z alone. If we introduce the acoustic velocity, \( v(t, z) \) given by

\[
v(t, z) = \frac{\partial u(t, z)}{\partial t} \tag{2.3}
\]

and use (2.1), (2.2) can be written

\[
p(z) \frac{\partial v(t, z)}{\partial t} = \frac{\partial p(t, z)}{\partial z}. \tag{2.4}
\]

Taking \( \partial / \partial t \) of (2.1), we obtain

\[
\frac{\partial p(t, z)}{\partial t} = -\lambda(z) \frac{\partial v(t, z)}{\partial z} \tag{2.5}
\]

We could obtain wave equations for \( p(t, z) \) and \( v(t, z) \) from (2.4) and (2.5), but first we choose to make a change of independent variable. Recall that the local velocity of wave propagation, \( c(z) \), is

\[
c(z) = \left( \frac{\lambda(z)}{\rho(z)} \right)^{1/2} \tag{2.6}
\]
We define a new spatial variable, $\xi$, by

$$
\xi = \xi(z) = \int_0^z \left( \frac{\rho(s)}{\lambda(s)} \right)^{1/2} ds = 2 \int_0^z \frac{ds}{c(s)}, \quad (2.7)
$$

so that it has the units of time. In fact, $\xi$ is the two-way travel time associated with the depth $z$, the time required for a wave to penetrate to depth $z$ and return to the surface ($z = 0$). Since, as mentioned in Chapter 1, we can only determine structure as a function of two-way travel time, $\xi$, this substitution is a suitable one. Rather than introduce a new family of dependent variables which are functions of $\xi$, we will retain the same symbols, indicating the new dependence by the use of $\xi$ as an argument (e.g., $c(\xi)$ is propagation velocity as a function of two-way travel time, while $c(z)$ is propagation velocity as a function of depth). *

The inverse relation to (2.7) is

$$
z = z(\xi) = \frac{1}{2} \int_0^\xi \left( \frac{\lambda(\tau)}{\rho(\tau)} \right)^{1/2} d\tau = \frac{1}{2} \int_0^\xi c(\tau) dz. \quad (2.8)
$$

Using (2.7) and (2.8), we obtain from (2.4) and (2.5)

$$
\frac{\partial v(t,\xi)}{\partial \xi} = - \frac{1}{2a(\xi)} \frac{\partial p(t,\xi)}{\partial t}. \quad (2.9)
$$

*We assure the reader that, once the substitution is made, we plan to work in the $\xi$-domain almost exclusively, so that he need not prepare for a succession of algebraic sleight-of-hand tricks.
\[ \frac{\partial p(t, \xi)}{\partial \xi} = \frac{a(\xi)}{2} \frac{\partial v(t, \xi)}{\partial t}, \]  

(2.10)

where \( a(\xi) \), the acoustic impedance, is

\[ a(\xi) = c(\xi) \rho(\xi) = (\lambda(\xi) \rho(\xi))^{1/2}. \]  

(2.11)

We will refer to (2.9) and (2.10) as the acoustic Telegrapher's equations, to emphasize their structural parallel with the equations for electromagnetic waveguide propagation. We can combine (2.9) and (2.10) to obtain wave equations for \( v(t, \xi) \) and \( p(t, \xi) \):

\[ \frac{1}{2} \frac{\partial^2 p(t, \xi)}{\partial t^2} = \frac{\partial^2 p(t, \xi)}{\partial \xi^2} - 2g(\xi) \frac{\partial p(t, \xi)}{\partial \xi}, \]  

(2.12)

\[ \frac{1}{2} \frac{\partial^2 v(t, \xi)}{\partial t^2} = \frac{\partial^2 v(t, \xi)}{\partial t^2} + 2g(\xi) \frac{\partial v(t, \xi)}{\partial \xi}, \]  

(2.13)

where \( g(\xi) \) is the reflection coefficient density,

\[ g(\xi) = \frac{1}{2} \frac{d}{d\xi} \ln(a(\xi)) = \frac{a'(\xi)}{2a(\xi)}. \]  

(2.14)

We have implicitly assumed that the dependent variables \( p(t, \xi) \) and \( v(t, \xi) \) are sufficiently smooth (have the appropriate number of derivatives) such that (2.12), (2.13) and their antecedents make sense. We also must place similar
restrictions on the smoothness of the physical parameters of the medium. In particular, we assume that \( a(\xi) \) is continuous, and differentiable everywhere except at a finite number of points. This means that, at worst, \( g(\xi) \) of (2.12) and (2.13) is piecewise continuous, containing a finite number of step singularities of finite amplitude. The continuity of \( a(\xi) \), or the piecewise continuity of \( g(\xi) \) are generally sufficient to ensure the existence of a solution to the propagation equations. We can, of course, handle step discontinuities in \( a(\xi) \), which give rise to impulse discontinuities in \( g(\xi) \), by matching the solutions for the continuous regions adjacent to the discontinuity across the boundary.

We do not expect that this regularity restriction on \( a(\xi) \) will prevent us from modelling the observed response of a physically reasonable structure to a realistic source. Even though the true \( a(\xi) \) may contain step discontinuities, both the source and receiver must have finite bandwidth; it should be adequate, therefore, to approximate the "true" \( a(\xi) \) by one which itself has a large, but finite, bandwidth.

This is an appropriate point at which to specify precisely the parameter or characteristic of the medium which we have heretofore referred to as "structure." Even though we are not concerned with estimators per se, we must select the feature to be estimated, so that we may define estimation bounds. Given its central role in the basic wave equations,
(2.12) and (2.13), and because it leads to simpler estimator structures, we choose the reflection coefficient density, \( g(\xi) \), as the medium characteristic to be estimated. We note that the one processor we have examined which estimates \( a(\xi) \) [7] does this by first forming an estimate of \( g(\xi) \), \( \hat{g}(\xi) \); the estimate of \( a(\xi) \), \( \hat{a}(\xi) \), is obtained by

\[
\hat{a}(\xi) = a(0) \exp \left( 2 \int_{0}^{\xi} \hat{g}(\tau) d\tau \right). \tag{2.15}
\]

A decomposition of \( p(t, \xi) \) and \( v(t, \xi) \) into upward- and downward-travelling wave components will assist our specification of boundary conditions and source/observation processes considerably. Although such a decomposition is well-known for the homogeneous waveguide \( (g(\xi) = 0) \), there is no commonly-accepted decomposition for our inhomogeneous waveguide. We will use a decomposition [22] which has been successfully applied to this problem, employing the definitions:

\[
p(t, \xi) = p^+(t, \xi) + p^-(t, \xi) \tag{2.16a}
\]

\[
v(t, \xi) = v^+(t, \xi) + v^-(t, \xi) \tag{2.16b}
\]

\[
p^+(t, \xi) = a(\xi)v^+(t, \xi) \tag{2.16c}
\]

\[
p^-(t, \xi) = -a(\xi)v^-(t, \xi). \tag{2.16d}
\]
It has been shown [22] that the components of this decomposition \( p^+(t,\xi), p^-(t,\xi), v^+(t,\xi), \) and \( v^-(t,\xi) \) can be interpreted as travelling waves; we will verify this in Section 2.3, while showing that the discrete, equal travel-time layer model converges to the continuous case in the limit, as the layer thickness goes to zero. There we will see that \( p^+(t,\xi) \) is the component travelling in the \( +\xi \) (down) direction, and \( p^-(t,\xi) \), in the \( -\xi \) (up) direction.

For compactness, and because of our underlying marine seismic motivation (estimation of sediment structure), we hereafter restrict our attention to pressure waves. Parallel results can also be derived for velocity waves.

Substitution of (2.16a) and (2.16b) into the acoustic Telegrapher's equations, (2.9) and (2.10), and elimination of \( v^+(t,\xi) \) and \( v^-(t,\xi) \), gives the coupled propagation equations for pressure waves, in terms of travelling wave components:

\[
\frac{1}{2} \frac{\partial p^+(t,\xi)}{\partial t} = - \frac{\partial p^+(t,\xi)}{\partial \xi} + g(\xi)(p^+(t,\xi) - p^-(t,\xi)) \quad (2.17a)
\]

\[
\frac{1}{2} \frac{\partial p^-(t,\xi)}{\partial t} = \frac{\partial p^-(t,\xi)}{\partial \xi} + g(\xi)(p^+(t,\xi) - p^-(t,\xi)) \quad (2.17b)
\]

This set of coupled equations may be reduced to the single wave equation for \( p(t,\xi) \), (2.12).

Like (2.12), (2.17) requires two boundary conditions.
We can obtain one boundary condition from the assumption that the medium boundary at $\xi = 0$ is a stress-free surface: this implies that $p(t,0) = 0$, or

$$p(t,0) = p^+(t,0) + p^-(t,0) = 0.$$  \hspace{1cm} (2.18)

The second boundary condition is a radiation condition, arising from our assumption that the medium is semi-infinite, and the implicit assumption that there are no sources other than at the surface: no upcoming waves from $\xi = -\infty$. Since we will assume that the observation interval is $T$ seconds in length, starting at $t = 0$, the radiation condition may be applied at any depth $\xi \geq T$, without affecting the received signal. Hence, our second boundary condition is

$$p^-(t,T) = 0.$$ \hspace{1cm} (2.19)

We must also specify initial conditions. A sufficient initial condition is initial rest:

$$p^+(0,\xi) = 0$$ \hspace{1cm} (2.20a)

$$p^-(0,\xi) = 0.$$ \hspace{1cm} (2.20b)

Note that these are stronger than $p(0,\xi) = p^+(0,\xi) + p^-(0,\xi) = 0$,
since they ensure there is no propagation present in the system prior to $t = 0$. They also ensure $\partial p^\pm(0,\xi)/\partial \xi = 0$.

We now have a well-posed problem with propagation equations (2.17), boundary conditions (2.18), (2.19), and initial conditions (2.20): it remains to define the source and observation mechanisms. As mentioned previously, we do not intend to model the source/medium or receiver/medium coupling in detail, since we consider this portion of the problem outside our area of concern. We have chosen coupling mechanisms which can be criticized as unrealistic, but which give usable results. Because of their simplicity, we argue, as we did when excluding geometrical spreading and attenuation, that while the results we obtain may not be completely accurate, they are certainly still valid lower bounds for estimation performance in the presence of more complicated phenomena.

We will model the source by modifying the boundary condition at $\xi = 0$, (2.18), so that the interface is no longer a free surface, but is subject to an externally-imposed, time-varying stress:

$$p(t,0) = p^+(t,0) + p^-(t,0) = d(t), \quad (2.21a)$$

where the external stress, $d(t)$, is the source waveform. Equation (2.21a) can be interpreted as modelling the injection
of downward-travelling pressure waves into the medium through a free surface, since

\[ p^+(t,0) = d(t) - p^-(t,0). \quad (2.21b) \]

Here \( p^+(t,0) \) is composed of a source term plus the upcoming pressure wave inverted by its interaction with the free surface. In order to ensure that the boundary condition (2.21), evaluated at \( t = 0 \), is compatible with the initial conditions (2.20), evaluated at \( \xi = 0 \), we must require \( d(0) = 0 \). We additionally assume that the medium is excited beginning at \( t = 0 \), so both restrictions may be summarized as \( d(t) = 0, \quad t \leq 0 \).

We model the deterministic portion of the observation, \( s(t) \), as the upcoming component of the pressure wave field, with

\[ s(t) = p^-(t,0), \quad 0 \leq t \leq T. \quad (2.22) \]

The random portion of the observation, the noise corruption, \( n(t) \), is modelled by additive zero-mean white Gaussian noise, so that the received signal, \( r(t) \), is

\[ r(t) = s(t) + n(t) = p^-(t,0) + n(t), \quad 0 \leq t \leq T \quad (2.23) \]
where

\[ E[n(t)] = 0, \quad E[n(t)n(\tau)] = \sigma^2 u_0(t-\tau). \]

A more general model for additive noise might include non-stationary (time-varying) behavior, or a colored (correlated) component, and our results may be extended to include these cases using well-known methods (see, for instance, [31]). Such phenomena are outside the central thrust of our problem, however, and we therefore leave their treatment to the interested reader.

This completes our model for a continuous medium, and we collect the relevant equations below.

Propagation equations:

\[ \frac{1}{2} \frac{\partial p^+(t, \xi)}{\partial t} = \frac{\partial^2 p^+(t, \xi)}{\partial \xi^2} + g(\xi)(p^+(t, \xi) - p^-(t, \xi)) \]  \hspace{1cm} (2.17a)

\[ \frac{1}{2} \frac{\partial p^-(t, \xi)}{\partial t} = \frac{\partial^2 p^-(t, \xi)}{\partial \xi^2} + g(\xi)(p^+(t, \xi) - p^-(t, \xi)) \]  \hspace{1cm} (2.17b)

Boundary conditions (including source):

\[ p^-(t, 0) + p^+(t, 0) = d(t) \]  \hspace{1cm} (2.21)

\[ p^-(t, T) = 0 \]  \hspace{1cm} (2.19)
Initial conditions:

\[ p^+(0,\xi) = p^-(0,\xi) = 0 \]  \hspace{1cm} (2.20)

Observation model:

\[ r(t) = p^-(t,0) + n(t), \quad 0 \leq t \leq T \]  \hspace{1cm} (2.23)

Figure 2.1 is a schematic view of this model, in which the boundary conditions, source, and observation models are imposed on a conceptual lumped version of the medium.

Although the development of such a model has the desirable effect of putting us on firm theoretical ground, it is of limited practical utility, since it does not possess simple analytical solutions for any reasonably diverse family of reflection coefficient densities and source waveshapes. A slightly different problem formulation is based on the assumption that the density, \( \rho(z) \), is approximately constant, specifically that

\[ \frac{\partial^2 \rho(z)}{\partial z^2} \ll \rho(z), \]

and

\[ \left( \frac{\partial \rho(z)}{\partial z} \right)^2 \ll \rho(z). \]
Figure 2.1 -- A schematic rendering of the model developed for the reflection response of a continuous acoustic waveguide. Note that $a(\xi)$ is constant for $\xi > T$. 

\begin{align*}
\frac{1}{2} \frac{\partial p^+(t, \xi)}{\partial t} &= -\frac{\partial p^+(t, \xi)}{\partial \xi} + g(\xi)(p^+(t, \xi) - p^-(t, \xi)) \\
\frac{1}{2} \frac{\partial p^-(t, \xi)}{\partial t} &= \frac{\partial p^-(t, \xi)}{\partial \xi} + g(\xi)(p^+(t, \xi) - p^-(t, \xi))
\end{align*}

Propagation Equations
The pressure wave equation which results is

\[ \frac{\partial^2 p(t,z)}{\partial z^2} = \frac{1}{c^2(z)} \frac{\partial^2 p(t,z)}{\partial t^2} \tag{2.24} \]

(for a derivation, see [32], Chapter 3). Suppose a medium obeying (2.24) is overlain by an infinite homogeneous half-space, with acoustic impedance \( a(0) \), and the lower half-space is illuminated by a sinusoidal \( e^{-j2\pi ft} \) travelling wave incident from the upper half space. The reflection response of the medium to this continuous-wave source, measured at \( z = 0 \), is given by the Bremmer series [28], a generalization of the well-known WKB perturbation solution of (2.24):

\[ p^-(t,0) = \sum_{m=1}^{\infty} b_m(f,\xi) \bigg|_{\xi=0} = e^{-j2\pi ft} B(f,\xi) \bigg|_{\xi=0} \tag{2.25} \]

where

\[ b_1(f,\xi) = e^{j2\pi f \xi/2} \int_0^\infty g(\tau)e^{-j2\pi f \tau} d\tau \tag{2.26a} \]

\[ b_{m+1}(f,\xi) = -e^{j2\pi f \xi/2} \int_0^\infty g(\tau)g(\eta) b_m(f,\eta) e^{j2\pi f \eta/2} d\eta e^{-j2\pi f \tau} d\tau, \tag{2.26b} \]

\[ B(f,\xi) = \sum_{m=1}^{\infty} b_m(f,\xi) \tag{2.26c} \]
and $g(\xi)$ is the now-familiar reflection coefficient density (as a function of two-way travel time). While several series solutions are possible for this problem (Peano-Baker series, Neumann series, and a special form developed in [22], among others) the Bremmer series has an advantage in that, by its derivation, each succeeding term models the effect of a distinct, progressively higher-order, internal scattering process. Hence, in the Bremmer series, we can discard higher order terms knowing exactly what phenomena we are throwing away.

Although the initial derivation of the Bremmer series [28] was heuristic, it has been theoretically justified and examined in detail by subsequent investigators. Of particular interest are the conditions derived in [33] under which the Bremmer series $B(\xi, \xi)$ of (2.26c) converges uniformly and absolutely in $\xi$:

1. $a(\xi)$ is continuously differentiable ($g(\xi)$ is continuous)
2. $f \neq 0$
3. $\int_{0}^{\infty} |a'(\xi)| \, d\xi < \infty$
4. $\int_{0}^{\infty} \left| \frac{a'(\xi)}{a(\xi)} \right| \, d\xi < \pi$

These are sufficient conditions for absolute, uniform convergence of (2.26c), except for (4), which is both necessary and sufficient. We expect some convergence prob-
lems near \( f = 0 \), since this is a singular point of (2.24).

The Bremmer series solution of (2.24) contains an implicit radiation condition at \( \xi = -\infty \). It does not (and cannot, if it is to converge) model interaction with a free surface, where \( a(\xi) \) would contain a step, and condition (1) above would be violated. We can view the medium, in the absence of free surface, as a lumped-parameter linear system with \( p^+(t,0) \) as input, and \( p^-(t,0) \) as output (this is implicit in our conceptual system diagram of Figure 2.1). The transfer function of this lumped model is given by the Bremmer series (in its region of validity):

\[
\frac{P^-(f,0)}{P^+(f,0)} = B(f,0) = \sum_{m=1}^{\infty} b_m(f,0) \tag{2.27}
\]

where

\[
P^-(f,0) \leftrightarrow p^-(t,0)
\]

\[
P^+(f,0) \leftrightarrow p^+(t,0)
\]

This observation allows us to define the model using the linear system diagram of Figure 2.2; it is identical to that of Figure 2.1, save that the medium is now modelled by its impulse response, \( b(t) \), the Fourier transform of the Bremmer series,
Medium, modelled by the Fourier transform of its Bremmer series $d(t)$, $s(t)$, Noiseless response $r(t)$, Received signal $n(t)$, Additive noise $b(t)$, Free surface interaction $-1$.

Figure 2.2 -- System diagram model for the reflection response of an acoustic waveguide, based on a Bremmer series model for propagation.
This conceptual separation of a medium's reflection response into an internal, continuous scattering model, plus a discrete model for the free surface interaction will be used extensively in Chapters 4 and 5, albeit not with the Bremmer series response model.

The Bremmer series will play a role in Chapter 5, as the basis for a physically-motivated approximation for the response of a medium whose reflection coefficient density function is a random process. We have presented it, not because we intend to make extensive use of it, but rather to demonstrate that a general method exists for computing the reflection response for a wide class of \( g(\xi) \).

In this Section, we have developed the basic continuous model to be used for the reflection response of a medium, in the process amplifying and adding to the assumptions stated in Section 2.1. Recognizing that this model is, in and of itself, of limited practical value, we have presented another model, based on a slightly simplified physical situation, which is applicable to a wide variety of structures, and whose series form has a valuable physical interpretation.

2.3 Discrete Model

The simulation model we develop in this Section is
based on the intuitive assumption (which we will verify) that we can approximate the response of a medium whose properties vary continuously by that of a medium consisting of a large number of discrete, parallel, planar, homogeneous layers. The thickness of each layer is proportional (by the same proportion for all layers) to the velocity of wave propagation within that layer. This constraint ensures that the structure has "equal travel-time layers": the time required for a wave to propagate through a layer is the same for all layers. The layer thicknesses are unequal if measured in the spatial variable, $z$, but uniform if measured in the travel time variable, $\xi$. A schematic diagram for such a model, indicating the layer and interface numbering scheme, is shown in Figure 2.3.

Since the layers are homogeneous, we can, without ambiguity, decompose the (pressure) wave field within each layer into upward- and downward-travelling wave components, and can model the propagation of each component through a layer by a simple time delay. It is sufficient, therefore, to represent the wave field in each layer by the travelling wave components observed just below the overlying interface. We will denote these components, for the $i$th layer, by $p^+_i(t)$ and $p^-_i(t)$, respectively.

The components in adjacent layers are related by the reflection and transmission processes at the layer interfaces,
Figure 2.3 -- Layer, interface, and parameter numbering scheme for a discretely-layered medium.
which are defined by the interface reflection coefficients. We denote the reflection coefficient for the ith interface as $g_i$; it is given by the well-known formula

$$g_i = \frac{a_{i+1} - a_i}{a_{i+1} + a_i}$$

(2.29)

where $a_i$ is the acoustic impedance of the ith layer. A system model for an interface and the overlying layer is shown in Figure 2.4, relating $p_i^+(t)$, $p_i^-(t)$, with $p_{i+1}^+(t)$, $p_{i+1}^-(t)$. The one-way delay in the layer is $\Delta \xi/2$, so that the total time required for a wave to penetrate layer $i$, interact with interface $i$, and return to layer $i-1$, is $\Delta \xi$ seconds (the two-way travel time thickness of each layer is $\Delta \xi$). A complete system model is obtained by interconnecting the appropriate number of such layer-interface models, with boundary conditions applied at the top and bottom.

Before proceeding further, we change from a continuous-time to a discrete-time system model. Our motivation for doing this is threefold: (1), digital computer-based simulations are inherently discrete-time algorithms, and we desire to build this fact into our model at an early stage; (2), the model has a particularly simple discrete-time state-variable formulation; and (3), we will be able to simulate experimentally-observed reflection data directly, since these data are typically collected in sampled-data form.
Figure 2.4 -- System model for pressure wave propagation in a single layer-interface pair of a discrete equal travel-time layer medium. Note that $p_i^+(t)$ and $p_i^-(t)$ are the wave components at the top of the $i$-th layer, just below the $(i-1)$st interface.
With the selection of a discrete-time model, one issue which must be considered is the choice of sampling interval, $\Delta t$, which, by construction, must also be the one-way layer travel time, $\Delta t/2$. It is clear that the sampling rate $(1/\Delta t)$ must be above the Nyquist frequency for $d(t)$ (which we assume to be bandlimited). The selection of $\Delta t$, however, also influences the accuracy with which the continuous acoustic impedance, $a(\xi)$, which we assume underlies our simulation, can be modelled by a step-function (simple function) approximation, so that our selection of $\Delta t$ cannot be based upon the bandwidth of $d(t)$ alone. We have successfully used an unverified but reasonable rule of thumb: the minimum allowable $\Delta t$ is defined by the larger of the bandwidths of $d(t)$ and $a(2t)$ (acoustic impedance as a function of one-way travel time). If extensive efforts involving this discrete model are to be undertaken, a more detailed study of the problem of selecting a sampling interval should be pursued.

We use the notation $p_i^+(k)$ and $p_i^-(k)$ for our discrete-time pressure variables, which are related to $p_i^+(t)$ and $p_i^-(t)$ by

$$p_i^+(k) = p_i^+(k\Delta t)$$

$$p_i^-(k) = p_i^-(k\Delta t),$$

and $d(k)$, $s(k)$, $n(k)$, and $r(k)$ for the discrete-time versions
of the other time functions which enter into the model.

The boundary conditions (including the source term), initial conditions, and observation mechanism for the discrete-layer model are the sampled-data analogs of those for the continuous model:

**Boundary conditions (including source):**

\[
p^+_1(k) + p^-_1(k) = d(k) \tag{2.30}
\]

\[
p^-_{N+1}(k) = 0 \tag{2.31}
\]

**Initial conditions:**

\[
p^+_1(0) = p^-_1(0) = 0 \tag{2.32}
\]

**Observation model:**

\[
r(k) = p^+_1(k) + n(k), \quad 1 \leq k \leq 2N \tag{2.33}
\]

where \( N \) is the number of interfaces in the model, and is related to the length of the observation interval, \( T \), by

\[
N = T/2\Delta t = T/\Delta \xi \tag{2.34}
\]

Note that \( k \), the new time index, runs from 1 to 2N. The additive noise term is still zero-mean, uncorrelated (white)
and Gaussian, with the discrete-time correlation structure

\[ \mathbb{E}[n(k)n(l)] = \sigma^2 \delta_{k,l}. \]  

(2.35)

Figure 2.5 shows a three-interface discrete-layer, discrete-time model which utilizes these boundary conditions and observation model (the initial conditions are implicit). At the risk of confusing the reader, we have indicated the layer delays using the conventional z-transform notation, \( z^{-1} \).

Such a model may be implemented in several ways; we have chosen a state-variable formulation because of its simple, compact canonic form. Neither the discrete-layer model itself, nor its representation in state-variable form are unique to our study (see, for instance, [8]); our choice of state variables, however, yields a very simple system matrix (compare that in [8]), and has not appeared before in the literature.

It is well-known that 2N state variables are required to represent a system model such as ours, which contains 2N delays. The elements of the 2N-dimensional state vector, \( \mathbf{x}(k) \), are selected as shown in the example of Figure 2.5: \( x_{2m+1}(k) \) is given by the upward-travelling wave component just below the \( m \)th interface (\( p_m^-(k) \)); \( x_{2m}(k) \), by the downward-travelling wave component just above the \( m \)th interface.
Figure 2.5 -- System model for the reflection response of a three-interface discrete, equal travel-time layer medium with time-sampled signals. Nodes which define state variables are indicated with the appropriate variable name, as \( x_1(k) \).
Two basic difference equations describe the propagation of the system in terms of these state variables:

\[ x_{2m+1}(k+1) = g_{m+1} x_{2m+2}(k) + (1-g_{m+1}) x_{2m+3}(k), \quad 0 \leq m \leq N-2 \]  
\( (2.36a) \)

\[ x_{2m}(k+1) = (1+g_{m-1}) x_{2m-2}(k) - g_{m-1} x_{2m-1}(k), \quad 1 \leq m \leq N. \]  
\( (2.36b) \)

Special cases of these equations hold for \( x_2(k) \) and \( x_{2N-1}(k) \), where the boundary conditions enter:

\[ x_2(k+1) = -x_1(k) + d(k) \]  
\( (2.37a) \)

\[ x_{2N-1}(k+1) = g_N x_{2N}(k). \]  
\( (2.37b) \)

Equations (2.36) and (2.37) are summarized by the state vector propagation equation

\[ x(k+1) = \Phi(g)x(k) + \Gamma d(k), \]  
\( (2.38) \)

with the initial condition \( x(0) = 0 \), where

\[ \Gamma = (0, 1, 0, \ldots, 0)^T \]  
\( (2.39) \)

\[ g = (g_1, \ldots, g_N)^T \]  
\( (2.40) \)
and the system matrix, \( \phi(g) \), has the very sparse band matrix structure (shown here for three interfaces)

\[
\phi(g) =
\begin{bmatrix}
0 & g_1 & 1-g_1 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & g_2 & 1-g_2 & 0 \\
0 & 1+g_1 & -g_1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & g_3 \\
0 & 0 & 0 & 1+g_2 & -g_2 & 0
\end{bmatrix}
\] (2.41)

(In practice, the system (2.38) is propagated using the difference equations (2.36) rather than matrix multiplication.)

The observation equation is

\[
r(k) = x_1(k) + n(k), \quad 1 \leq k \leq 2N
\]

or

\[
r(k) = M^T x(k) + n(k), \quad (2.42)
\]

where

\[
M = (1, 0, \ldots, 0)^T \quad (2.43)
\]

This completes the development of the system model for
the discrete, equal travel-time layer system model per se. Before proceeding to the next Chapter, however, we will demonstrate the convergence of the difference equations describing the system propagation in the discrete system to the partial differential equations we derived for the continuous system.

The discrete-layer, discrete-time simulation model was developed under the intuitive assumption that, in the limit as the layer thickness decreases to zero, the discrete model's response would converge to that of the continuous. By our choice of boundary and initial conditions, and of an observation mechanism, we have ensured that these aspects of the discrete model will converge in the continuum limit. We will verify our intuition by showing that the discrete propagation equations (2.36) converge to the continuous equations (2.17).

Let $p^+(t, \xi)$ and $p^-(t, \xi)$ be analytic functions in the region $[0, T] \times [0, T]$. We require that

$$p^+(k\Delta t, (m-1)\Delta \xi) = p_m^+(k),$$

(2.44)

where $p_m^+(k)$ are the travelling wave components in our discrete medium. Thus $p^+(t, \xi)$ and $p^-(t, \xi)$ take on values specified by the discrete wave propagation process over a mesh of points in the plane. Similarly, let $a(\xi)$ be a continuous function
on \([0,T]\) differentiable everywhere except at a finite number of points, which must satisfy

\[
a((m-1)\Delta \xi) = a_m. \tag{2.45}
\]

Since

\[
g_i = \frac{a_{i+1} - a_i}{a_{i+1} + a_i} \tag{2.29}
\]

we can also write \(g_i\) (for small \(\Delta \xi\)) as

\[
\begin{align*}
g_i &= \frac{a(i\Delta \xi) - a((i-1)\Delta \xi)}{a(i\Delta \xi) + a((i-1)\Delta \xi)} \\
&= \frac{1}{2a(i\Delta \xi)} \Delta \xi = g(i\Delta \xi) \Delta \xi \tag{2.46}
\end{align*}
\]

(and we see the motivation for referring to \(g(\xi)\) as the "reflection coefficient density").

With these preliminary results, we first consider the difference equation for the propagation of odd-numbered state variables:

\[
x_{2m+1}(k+1) = g_{m+1}x_{2m+2}(k) + (1-g_{m+1})x_{2m+3}(k). \tag{2.36a}
\]

Using the definition of the state variables in terms of travelling wave components, (2.36a) becomes
\[ p_m^{-(k+1)} = g_m^{m+1} p_{m+1}^{+(k-1)} + (1-g_m^{m+1}) p_m^{-(k)} \]  

(2.47)

Let \( k\Delta t = t \) and \( (m-1)\Delta \xi = \xi \); then we can write (2.47) in terms of \( p^-(t,\xi) \) as

\[ p^-(t+\Delta t, \xi) = g_{m+1} p^+(t-\Delta t, \xi+\Delta \xi) + (1-g_{m+1}) p^-(t, \xi+\Delta \xi) \]  

(2.48)

Expanding in Taylor series about \( (t, \xi) \) to first order in \( \Delta t \) and \( \Delta \xi \), we have

\[ p^-(t, \xi) + \frac{\partial p^-(t, \xi)}{\partial t} \Delta t = g_{m+1} (p^+(t, \xi) - \frac{\partial p^+(t, \xi)}{\partial t} \Delta t) + \frac{\partial p^+(t, \xi)}{\partial \xi} \Delta \xi \]  

(2.49)

Substituting for \( g_{m+1} \) from (2.46), and again eliminating terms higher than first order, we obtain

\[ p^-(t, \xi) + \frac{\partial p^-(t, \xi)}{\partial t} \Delta t = g(\xi+\Delta \xi) \Delta \xi p^+(t, \xi) + (1-g(\xi+\Delta \xi) \Delta \xi) p^-(t, \xi) + \frac{\partial p^-(t, \xi)}{\partial \xi} \Delta \xi \]  

(2.50)

Cancelling \( p^-(t, \xi) \), and dividing through by \( \Delta \xi \), (2.50) becomes

\[ \frac{1}{2} \frac{\partial p^-(t, \xi)}{\partial t} = \frac{\partial p^-(t, \xi)}{\partial \xi} + g(\xi+\Delta \xi) (p^+(t, \xi) - p^-(t, \xi)) \]
Letting $\Delta \xi \to 0$, we have the desired result

$$
\frac{1}{2} \frac{\partial p^- (t, \xi)}{\partial t} = \frac{\partial p (t, \xi)}{\partial \xi} + g(\xi) \left( p^+ (t, \xi) - p^- (t, \xi) \right) \quad (2.17a)
$$

By a similar approach, we can obtain (2.17b) from (2.36b).

We have demonstrated, therefore, that in the continuum limit, the discrete-layer, discrete-time propagation model described by the state-variable difference equations (2.36) converges to the continuous propagation model, the acoustic Telegrapher's equations (2.17). Our intuition is thereby verified, and we see that the discrete model we have developed in this Section is a valid approximation for the continuous model of Section 2.2.

In closing, we note that there are many discrete models which might be used to simulate the reflection response of a continuous medium. The model presented here, however, is unusual, in that it has a physical interpretation in and of itself -- our discrete model corresponds exactly to a real physical system. Discrete models obtained by the straightforward application of finite-difference approximations to the continuous system equations generally do not have such an interpretation, and hence do not allow the same degree of insight as does our model.
CHAPTER 3
BOUNDS AND ESTIMATORS FOR A DISCRETE MEDIUM

3.1 Introduction

We have elected to study the medium structure identification problem in terms of lower bounds, specifically Cramer-Rao lower bounds, on estimation accuracy (estimate variance). As discussed in Chapter 1, such performance bounds are useful for three basic reasons: their analytic form provides insight into the manner in which media properties influence estimation accuracy; they are a performance benchmark against which we may evaluate the effectiveness of practical (sub-optimal) estimation schemes; and, they predict the performance of a large class of nonlinear estimators, generally impossible to compute directly, when the estimator is operating in its linear region (for high signal-to-noise ratios). In this Chapter, we will introduce the Cramer-Rao performance bound for a simple, but important medium structure model, and will demonstrate its utility in each of the three categories described above.

This Chapter contains the only truly comprehensive study we will make of the structure estimation problem; Chapters 4 and 5, in which we utilize more sophisticated structure models for the medium, are devoted exclusively to the consideration of performance bounds. We have started
our treatment with such a detailed analysis in order to present the Cramer-Rao bound in the context of a complete estimation problem, so that the bound may be viewed in proper perspective, with both its utility and its limitations demonstrated.

We will consider a class of media which have a particularly simple structure: they consist of discrete, equal travel-time layers, so that the discrete simulation model developed in Section 2.3 describes the model physics exactly. The "structure" to be estimated is \( g \), the vector of reflection coefficients of the medium, which we will assume to be unknown, non-random parameters. Despite our criticism of such simple, discrete reflector models in Chapter 1, and our emphasis on the distributed-parameter (continuous) model in Chapter 2, there are compelling reasons, both pedagogical and practical, for considering such a medium model at the outset of our treatment.

The comparative simplicity of the model makes it an ideal vehicle for introducing the Cramer-Rao bound, and demonstrating its properties. On a more practical level, it is appropriate to treat the discrete, equal travel-time layer model, both because it is a natural one to consider when initially selecting a model, and also because it is widely used. Moreover, this model will allow us to analyze two meaningful physical problems in detail, estimation of ocean
bottom reflection coefficients, and identification of two closely-spaced reflectors. These two simple case studies provide a clear and concrete demonstration of the physical insights which may be obtained by examination of the Cramer-Rao bound.

In Section 3.2, we introduce the Cramer-Rao bound in a form appropriate for the problem of structure estimation in a discrete medium, i.e., estimation of a set of discrete parameters given a set of discrete measurements. We define models for the one- and two-interface identification problems, and compute the associated Cramer-Rao bounds, which we show to have physically meaningful interpretations. We present an optimal (maximum-likelihood) estimator for \( g \) in Section 3.3, where we also discuss the notion of global estimation error, not predicted by the Cramer-Rao bound, and the use of ambiguity functions to visualize such errors. In Section 3.4, we present a sub-optimal estimator, the PEF estimator described in Chapter 1. These two estimators are compared in Section 3.5, by means of Monte Carlo tests; these tests demonstrate the ability of the Cramer-Rao bound to predict and bound estimator performance. We also observe the effect which bias may have on the performance of an estimator.

3.2 The Cramer-Rao Bound

The Cramer-Rao bound places a theoretical lower bound
on the performance (estimate variance) of an unbiased estimator. In particular, suppose we estimate a vector of unknown, non-random parameters, $\mathbf{g}$, given a vector of noisy observations, $\mathbf{r}$. If the estimate, $\hat{\mathbf{g}}(\mathbf{r})$, is unbiased,

$$E[\hat{\mathbf{g}}(\mathbf{r})] = \mathbf{g}, \quad (3.1)$$

then the Cramer-Rao inequality is valid, and places a lower bound on the variance of each element of $\hat{\mathbf{g}}(\mathbf{r})$:

$$E[(\hat{g}_i(\mathbf{r}) - g_i)^2] \geq (J^{-1})_{ii}. \quad (3.2)$$

$J$ is called the Fisher information matrix, and is given by

$$J_{ij} = E\left[\frac{\partial^2 \ln(p(\mathbf{r}|\mathbf{g}))}{\partial g_i \partial g_j}\right] = -E \left[\frac{\partial^2 \ln(p(\mathbf{r}|\mathbf{g}))}{\partial g_i \partial g_j}\right] \quad (3.3)$$

where

$$p(\mathbf{r}|\mathbf{g}) = \text{probability density function of the observation vector } \mathbf{r}, \text{ conditioned on the parameter vector } \mathbf{g}. \quad J$$

is a square matrix with dimension $\text{dim}(\mathbf{g})$. We will not derive (3.2) and (3.3) here, but refer the interested reader to [31] or [34]. Note that $J$ is generally a function both of the parameter vector, $\mathbf{g}$, and of the noise statistics.
of the observation vector, $\mathbf{r}$.

The bound expression (3.2) is satisfied with equality if and only if the estimation error can be expressed as a weighted sum of the partial derivatives of $\ln(p(\mathbf{r}|\mathbf{g}))$,

$$\hat{g}_i(\mathbf{r}) - g_i = \sum_{j} w_{ij}(\mathbf{g}) \frac{\partial \ln(p(\mathbf{r}|\mathbf{g}))}{\partial g_j}$$

(3.4)

for all $\mathbf{r}$ and $i$. An estimator for which (3.4) is true (whose performance is given exactly by the Cramer-Rao lower bound) is called an **efficient** estimator.

Equation (3.4) does not hold in general for nonlinear estimation problems, for which, in the absence of noise, $\mathbf{r}$ is a nonlinear function of $\mathbf{g}$; hence, nonlinear estimators cannot be efficient in general. Such estimators, however, usually exhibit performance essentially equal to that predicted by the Cramer-Rao bound when operating in their so-called **linear region**. We may typify this region as one of high signal-to-noise ratio, SNR (or, in some cases, long observation time). The linear region of operation is defined by that value of SNR below which estimate variance exceeds that predicted by the Cramer-Rao lower bound by some minimum amount, commonly 1 dB. This value of SNR is called the estimator performance **threshold**. The linear region, and the associated threshold, are specific to a particular estimator, and define the range of SNR (or other statistical parameter) over which estimator performance may be predicted by the Cramer-
Rao bound (over which the estimator is effectively efficient). Identification of the linear region and the threshold are important aspects of the study of any nonlinear estimator.

A more general version of the bound expression, (3.2), will be useful for our development in Chapter 4. Define the error covariance matrix, \( Q \), as

\[
(Q)_{ij} = E[ (\hat{g}_i(r) - g_i)(\hat{g}_j(r) - g_j)]; \quad (3.5)
\]

then the Cramer-Rao inequality may be stated as

\[
Q \geq J^{-1}, \quad (3.6)
\]

where the inequality sign implies that the matrix \((Q-J^{-1})\) is positive semi-definite.

We have presented the Cramer-Rao bound expressions for the problem of estimating an unknown, non-random parameter, the situation we are presently concerned with. A bound also exists for the situation in which the parameter to be estimated is a random variable from a specified distribution. In this case, we have available both the observation, \( r \), and a prior statistical description of \( g \). Define

\[
(J_0)_{ij} = E[ \frac{\partial \ln(p(r|g))}{\partial g_i} \frac{\partial \ln(p(r|g))}{\partial g_j} ] \quad (3.7)
\]
where

\[ p(g) = \text{a priori probability density function of } g. \]

The Cramer-Rao inequality is now

\[ Q \geq (J_o + J_p)^{-1}, \]  \hspace{1cm} (3.9)

and the equivalent of (3.2) follows directly. We will use these results in Chapter 5.

For the structure model used in this Chapter, \( g \) is non-random, so (3.2) and (3.3) are appropriate. The vector of noisy observations, \( r \), for the discrete, equal-travel-time layer model is given by

\[ r = (r(1), r(2), \ldots r(2N))^T \]  \hspace{1cm} (3.10)

where,

\[ r(k) = M^T x(k) + n(k), \quad 1 \leq k \leq 2N. \]  \hspace{1cm} (2.42)

For the purposes of this Chapter, we extend (2.42) by assuming that the medium consists of \( N \) discrete reflectors overlying a homogeneous halfspace. After time \( k = 2N \),
therefore, we observe only the internal reverberations within the first N layers, which propagate according to the system equation, (2.38). There are no structures below the Nth interface, and hence no "new" returns after k = 2N. As a result, (2.42) is valid for all k, 1 ≤ k ≤ ∞, and is not restricted to 1 ≤ k ≤ 2N. Also, (2.34) is no longer a constraint on T and N. We denote the number of points in the observation sequence, previously 2N, by L, and write the observation equation for a structure which overlies an infinite halfspace as

\[ r(k) = M^T x(k) + n(k), \quad 1 \leq k \leq L, \quad (3.11) \]

We may allow an infinite observation interval; the modification to (3.11) is obvious.

Since n(k) is a zero-mean, white Gaussian sequence, with

\[ E[n(k)n(\ell)] = \sigma^2 \delta_{k\ell}, \quad (2.35) \]

the elements of \( r \) are jointly Gaussian, independent random variables, and we can write \( p(r|g) \) as

\[ p(r|g) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2}(r-s(g))^T(r-s(g))\right) \]

\[ = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{k=1}^{L} (r(k) - s(k|g))^2\right) \quad (3.12) \]
where

\[(s(g))_k = s(k|g) = N^T x(k).\]

We have written \(s(g)\) and \(s(k|g)\) to emphasize the functional dependence of the deterministic portion of the received signal on the parameter vector. Note that the dimension of \(r\) is \(L\), and of \(g\), \(N\). We now evaluate the bound expression, (3.3). We have

\[
\frac{\partial \ln(p(r|g))}{\partial g_i} = \frac{1}{\sigma^2} \sum_{k=1}^{L} \frac{\partial s(k|g)}{\partial g_i} (r(k) - s(k|g)), \quad (3.13)
\]

so that

\[
E\left[\frac{\partial \ln(p(r|g))}{\partial g_i} \frac{\partial \ln(p(r|g))}{\partial g_j}\right] = \frac{1}{\sigma^2} \sum_{k=1}^{L} \frac{\partial s(k|g)}{\partial g_i} \frac{\partial s(k|g)}{\partial g_j}, \quad (3.14)
\]

The Fisher information matrix for our system is

\[
J = \frac{1}{\sigma^2} \nabla_g (s(g)) (\nabla_g (s(g)))^T, \quad (3.15)
\]

where we have introduced the gradient operator \(\nabla_g(\cdot)\),

\[
(\nabla_g (s(g)))_{lk} = \frac{\partial s(k|g)}{\partial g_l} \quad (3.16)
\]
These expressions, (3.14), and (3.15), must generally be evaluated numerically; an algorithm for this computation is developed in Appendix A.

We will illustrate the bound with two simple examples. In the first, the medium to be identified contains only one interface. This example corresponds to the problem of estimating the reflection coefficient of the ocean bottom, modelled as a discrete reflector, when there are no significant sub-bottom reflectors. The medium's response, therefore, consists of the primary (direct) bottom reflection, together with water column multiples (pressure waves reverberating between ocean bottom and ocean surface). We assume that the one-way travel time in the water layer, \( T_w \), is known exactly. The source waveform, \( d(t) \), is also known, and has duration less than \( T_w \) seconds, so successive multiples do not overlap. The (continuous) received signal is

\[
r(t) = - \sum_{k=1}^{L/2} (-g)^k d(t-2kT_w) + n(t),
\]

where \( g \) is the reflection coefficient of the bottom, and the receiver is turned off after \((L/2-1)\) multiples (and the one primary reflection) have been received (we assume \( L \) is even).

We obtain a discrete sequence of observations from this continuous received signal by processing \( r(t) \) with the matched-filter receiver structure of Figure 3.1, in which
\[ s(t) \rightarrow r(t) \rightarrow d(T_w - t) \rightarrow r(k) \]

\[ t = (k+1)T_w, \quad k = 1, \ldots, L \]

\[ n(t) = \sum_{\ell=1}^{\infty} (-g)\ell d(t-2\ell T_w) \]

Figure 3.1 -- A matched-filter/sampler receiver for transforming the continuous observation, \( r(t) \), into the discrete sequence, \( r(k) \).
the output of a filter, matched to \( d(t) \), is sampled every \( T_w \) seconds, beginning at \( t = 2T_w \). The matched filter output, evaluated at \( t = (k+1)T_w \), is

\[
\begin{align*}
    r(k) &= (r(t) * d(T_w - t))_{t=(k+1)T_w} \\
    &= \int_0^{\infty} r(\tau) d(\tau - kT_w) d\tau \\
    &= -\frac{L}{2} \sum_{m=1}^{\infty} (-g)^m \int_0^{\infty} d(\tau - 2mT_w) d(\tau - kT_w) d\tau \\
    &\quad + \int_0^{\infty} n(\tau) d(\tau - kT_w) d\tau. \quad (3.18)
\end{align*}
\]

Since \( d(t) \) is of duration less than \( T_w \), we have, for \( 1 \leq k \leq L \),

\[
\begin{align*}
    r(k) &= \begin{cases} 
    -(-g)^m \int_0^{T_w} d^2(t) dt + n(k), & k = 2m \\
    n(k), & k = 2m + 1
    \end{cases} \quad (3.19)
\end{align*}
\]

where \( n(k) \) is a zero-mean, white Gaussian sequence, with

\[
E[n(k)n(\ell)] = \sigma^2 \int_0^{T_w} d^2(t) dt \delta_{k\ell}, \quad (3.20)
\]

\( \sigma^2 \) being the power spectral density of the white process \( n(t) \).
Equation (3.19) is precisely the response of a single-interface, discrete, equal travel-time layer medium to a discrete impulse of amplitude \( \int_0^T w^2(t) \, dt \), with sampling interval \( \Delta t = T_w \).

For the single-parameter case, the Cramer-Rao bound is given, using (3.14) in (3.2),

\[
\mathbb{E}[(\hat{g}(r) - g)^2] \geq \frac{\sigma^2 \int_0^T w^2(t) \, dt}{\sum_{k=1}^L \left( \frac{\partial s(k|g)}{\partial g} \right)^2} \tag{3.21}
\]

The bound is a non-increasing function of \( L \): the longer the observation, the better the performance. We saw, in (3.19), that \( s(k|g) \) is zero for \( k \) odd; therefore, it is sufficient to take the sum in (3.21) over even values of \( k \) only. Let

\[
\tilde{s}(k|g) = s(2k|g) = -(-g)^k \int_0^T w^2(t) \, dt \tag{3.22}
\]

so

\[
\sum_{k=1}^L \left( \frac{\partial s(k|g)}{\partial g} \right)^2 = \frac{L/2}{\sum_{k=1}^L \left( \frac{\partial \tilde{s}(k|g)}{\partial g} \right)^2} = \frac{T_w^2}{(\int_0^T d^2(t) \, dt)^2} \sum_{k=1}^L k^2 g^{2k-2} \tag{3.22}
\]

We see immediately that the bound value is insensitive to the
sign of \( g \). Equation (3.21) can now be written

\[
E[(\hat{g}(r) - g)^2] \geq \frac{1}{\text{SNR}} \frac{1}{\sum_{k=1}^{L/2} k^2 g^{2k-2}}
\]  

(3.23)

where we have introduced \( \text{SNR} \), the signal-to-noise ratio,

\[
\text{SNR} = \frac{\int_0^T w^2(t) dt}{\sigma^2}
\]

(3.24)

The summation in (3.23) can be found in [35]:

\[
\sum_{k=1}^{L/2} k^2 g^{2k-2} = \frac{1+g^2 - g^{L+2} - g^L (1+L(1-g^2)/2)^2}{(1-g^2)^3}
\]

(3.25)

So that

\[
E[(\hat{g}(r) - g)^2] \geq \frac{1}{\text{SNR}} \frac{(1-g^2)^3}{1+g^2 - g^{L+2} - g^L (1+L(1-g^2)/2)^2}.
\]

(3.26)

For an infinite observation interval, \( L \to \infty \), the bound is

\[
E[(\hat{g}(r) - g)^2] \geq \frac{1}{\text{SNR}} \frac{(1-g^2)^3}{(1+g^2)}
\]

(3.27)

Defining

\[
\sigma_{\text{CRB}}^2 (g, L) = \frac{(1-g^2)^3}{1+g^2 - g^{L+2} - g^L (1+L(1-g^2)/2)^2}
\]

(3.28)
we have

\[ E[(\hat{g}(x) - g)^2] \geq \frac{\sigma_{\text{CRB}}^2(g, L)}{\text{SNR}}. \quad (3.29) \]

We therefore characterize the bound's behavior in terms of \( \sigma_{\text{CRB}}^2(g, L) \).

In Figure 3.2, we have plotted \( \sigma_{\text{CRB}}^2(g, L) \) as a function of \( L \), for various values of \( |g| \) (recall that \( \sigma_{\text{CRB}}^2(g, L) = \sigma_{\text{CRB}}^2(-g, L) \), as noted above). Recall that the smaller \( \sigma_{\text{CRB}}^2(g, L) \), the better the potential estimator performance.

We now examine the insights which may be drawn from the Figure. As predicted below (3.21), the bound is a monotone decreasing function of \( L \). This is physically reasonable, since the longer the observation, the more information is available. As \( L \to \infty \), however, \( \sigma_{\text{CRB}}^2(g, L) \) is bounded away from zero, approaching the asymptotic value (3.27). This occurs because the signal, proportional to \( |g|^k \), becomes progressively smaller with increasing \( k \), while the noise level remains constant; eventually, the signal component is so small relative to the noise that no significant improvement in performance can be obtained by considering a longer observation. We also see that performance improves as \( |g| \) increases, verifying our intuition: the larger, \( |g| \), the more energy is trapped in the water layer at any given time. Alternatively, we can say that the larger \( |g| \), the more
Figure 3.2 -- Cramer-Rao bound values, $\sigma_{\text{CRB}}^2(g,L)$, for a single-interface system.
slowly source energy leaks away into the lower half-space, and is lost to the measurement process.

The theoretical bound implies, therefore, that improved performance may be obtained as (1) L increases, and (2), as \(|g|\) increases. These are physically reasonable results, and we expect that the performance of an actual estimator, designed to take full advantage of the physical situation, should follow the same trends.

The second example is based on the next more complicated discrete reflector structure which we might encounter: the medium to be identified is assumed to have two equi-spaced interfaces. This example corresponds to the practical problem of estimating the seafloor reflection coefficient, \(g_1\), in the presence of a strong sub-bottom reflector whose reflection coefficient, \(g_2\), must be estimated as well. We again assume that there are no significant reflectors aside from those we have postulated, and that the source waveform is known and has duration less than \(T_w\) seconds. We further assume that the layers are equi-spaced, and that the one-way travel time of each layer is \(T^*_w\).

\[\text{This assumption is somewhat restrictive, but suits our main purpose, to demonstrate the physical insight which may be obtained from the Cramer-Rao bound. We could, of course, allow layer thicknesses which are arbitrary integer multiples of a common time interval; if this is done, the duration of } d(t) \text{ should be less than the one-way travel time of the thinnest layer. This example, as it stands, may be considered a worst case, in the sense that we are not able to separate different components of the response on the basis of time structure, as we might do if the layer travel times were incommensurate.}\]
The continuous received signal is

\[ r(t) = \sum_{k=1}^{L} s(k|g_1, g_2)d(t-2kT_w) + n(t) \quad (3.30) \]

where \( s(k|g_1, g_2) \) denotes the response of a discrete, two-interface system to a unit impulse. We will not write out \( s(k|g_1, g_2) \) explicitly, as we did for \( s(k|g) \) in the previous example; it is sufficiently complicated to be confusing rather than enlightening. We process \( r(t) \) with a matched-filter/sampler receiver, similar to that shown in Figure 3.1, to obtain the discrete sequence, \( r(k) \):

\[
\begin{align*}
r(k) &= \begin{cases} 
  s(k|g_1, g_2) \int_{0}^{T_w} d^2(t)dt, & k=2m \\
  n(k), & k=2m+1 
\end{cases} 
\end{align*}
\quad (3.31)
\]

So, again, the resultant discrete sequence is exactly the response of the analogous discrete, equal travel-time layer model, with sampling interval \( \Delta t = T_w \), to a discrete impulse of amplitude \( \int_{0}^{T_w} d^2(t)dt \).

Using (3.31), (3.20), and (3.15) in (3.2), the Cramer-Rao inequality for \( g_i \) is

\[
E[(\hat{g}_i(r) - g_i)^2] \geq \frac{\sigma^2}{\int_{0}^{T_w} d^2(t)dt} \cdot [Vg(g)]^{-1} \quad (3.32)
\]
or, using (3.24),

\[
E[ (\hat{g}_1(t) - g_1)^2 ] \geq \frac{1}{\text{SNR}} \left[ \nabla_g (s(g)) \, \nabla_g (s(g))^T \right]_{ii}^{-1} \tag{3.33}
\]

Let

\[
\sigma^2_{\text{CRB}}(g, L, i) = \left[ \nabla_g (s(g)) \, \nabla_g (s(g))^T \right]_{ii}^{-1} \tag{3.34}
\]

we characterize the bound's behavior in terms of \( \sigma^2_{\text{CRB}}(g, L, i) \), as we did for the single-interface example. Although quite complicated analytic expressions may be derived for \( \sigma^2_{\text{CRB}}(g, L, i) \), we have chosen to compute the bound values numerically, using the technique described in Appendix A.

Contour plots of \( \sigma^2_{\text{CRB}}(g, L, i) \) for the two-interface example are shown in Figures 3.3, 3.4, and 3.5, for \( L = 4, 6, \) and 20 respectively. Both \( \sigma^2_{\text{CRB}}(g, L, 1) \) and \( \sigma^2_{\text{CRB}}(g, L, 2) \) were found to be even functions of \( g_1 \), so values are shown for positive \( g_1 \) only. It is evident that the performance bound improves for both \( \hat{g}_1(t) \) and \( \hat{g}_2(t) \) as \( L \) increases. The performance for an infinite observation length is well-approximated (over the indicated range of \( g \)) by the bound values in Figure 3.5, where \( L = 20 \). This behavior of the two-interface bound as a function of \( L \) corresponds to that observed for the single-interface example; it may be explained in the same physically reasonable way.
Figure 3.3a -- Contours of $\sigma_{CRB}^2(g, 4, 1)$.

Figure 3.3b -- Contours of $\sigma_{CRB}^2(g, 4, 2)$.
Figure 3.4a -- Contours of $\sigma^2_{\text{CRB}}(g, 6, 1)$.

Figure 3.4b -- Contours of $\sigma^2_{\text{CRB}}(g, 6, 2)$. 
Figure 3.5a -- Contours of $\sigma_{\text{CRB}}^2(g, 20, 1)$.

Figure 3.5b -- Contours of $\sigma_{\text{CRB}}^2(g, 20, 2)$.
The bound's behavior as a function of the reflection coefficients is substantially more complicated here than in the previous example. Two significant trends may be observed in the Figures: (1) as $|g_1|$ increases, the estimation bound for $g_2$ gets worse; (2) as $|g_2|$ increases, the estimation bounds for both $g_1$ and $g_2$ improve. The first observation is attributed to the "masking" effect of the upper interface; the larger $|g_1|$, the smaller the amplitude of any response which returns to the surface from the deep interface, since such a response must be scaled by $(1-g_1^2)$, the product of the downward and upward transmission coefficients of the upper interface. This effectively lowers the signal-to-noise ratio for the observation of the deep layer, even though all the source energy does eventually illuminate the lower interface as it leaks away into the lower halfspace. The explanation for our second observation is two-fold: increasing $|g_2|$ not only increases the strength of the return from the lower interface, but also retards the penetration of energy to the lower halfspace, trapping the source energy in the layer structure for a longer time. The asymmetry of the bound results as a function of $g_2$ is apparently due to interaction between the interface responses, as the asymmetry is linked to the sign of $g_2$.

In this Section, we have presented the appropriate Cramer-Rao bound formulae for the discrete structure model,
and we have illustrated the bound using two simple but meaningful examples. In particular, we have demonstrated that we may draw physically incisive conclusions about the effect of problem parameters on estimation accuracy by examining the behavior of the Cramer-Rao bound.

3.3 An Optimal Estimator

In this Section, we present a statistically optimal estimator for the reflection coefficients of a discrete, equal-travel-time layer medium: the maximum-likelihood (ML) estimator. This estimator is formed by choosing that \( g \) which maximizes \( p(r|g) \) for a given observation vector, \( r \). We emphasize that this estimator is optimal only in the sense that it employs a well-defined statistical criterion for selecting \( \hat{g} \) given an observation. In particular, it is not optimal in the sense of giving an estimate which has minimum variance. Indeed, we will see in Section 3.5 that, for some values of signal-to-noise ratio, the ad hoc estimator of Section 3.4 performs better (gives a smaller estimate variance) than does the ML estimator.

Maximization of \( p(r|g) \) is equivalent to maximization of \( \ln(p(r|g)) \), since \( \ln(.) \) is a monotone increasing function. Using the expression for \( p(r|g) \), (3.12), the maximum likelihood estimate, \( \hat{g}(r)_{\text{ML}} \), may be obtained by
\[ \hat{g}(r)_{\text{ML}}: \min_{\hat{g}} I(\hat{g}) = \min_{\hat{g}} \sum_{k=1}^{L} (r(k) - s(k|\hat{g}))^2 \]  

(3.35)

It is one thing to write down such an expression for the estimator, but quite a different matter to obtain an estimate given some real data. The quantity to be minimized in (3.35), \( I(\hat{g}) \), is, in all but the most trivial cases, a complicated nonlinear function of \( \hat{g} \), making it virtually impossible to obtain a closed-form expression for the estimate in terms of the data. In practice, \( \hat{g}(r)_{\text{ML}} \) is obtained by numerical minimization, using any one of a number of algorithms (conjugate gradient, Fletcher-Powell, etc. -- see, for instance, [36], or [37]). Note that it is possible to calculate the gradient, \( \nabla_g I(\hat{g}) \), using the methods of Appendix A. The necessity of using numerical optimization techniques is considered to be a major but unavoidable disadvantage of the ML method.

In the two examples presented in Section 3.2, we assumed that some global properties of the structure to be identified were known à priori, i.e., the number of reflectors in the medium, \( N \), and the layer spacing, \( T_w \). Such information, of course, is not generally available to us. A reasonable approach in this case might be to sample the continuous system response, \( r(t) \), at an appropriate rate, and fit the resulting sequence, \( r(k) \), to a discrete model by estimating reflection coefficients. Our development of
Section 2.3 assures us that, if we take $\Delta t$ sufficiently small, the discrete model will be a good approximation to the continuous, and this approach will be valid. Unfortunately, in the absence of prior structural information, we must allow one reflection coefficient for every two data points, as initially indicated in (2.34). For an observation length of a few hundred points (samples) this would require on the order of a hundred reflection coefficients. Even without observation noise, however, minimization of an expression such as (3.35) over a hundred independent variables is almost surely numerically ill-conditioned; too many parameters are competing to describe a limited amount of data. The dimensionality required in the model, together with the ML optimization (3.35), restrict the usefulness of model and estimator. One solution is to reduce the number of independent variables by making the reflection coefficients parametrically dependent on a coefficient vector of reduced dimension.

On the positive side, of course, the ML estimator has the desirable property of being, by design, noise tolerant. Moreover, an ML estimator is both asymptotically efficient, and asymptotically unbiased as the number of independent observations increases. For our problem, this is equivalent to $\text{SNR} \rightarrow \infty$.

\[
\lim_{\text{SNR} \rightarrow \infty} \frac{\mathbb{E}[(\hat{g}_i(r)_{\text{ML}} - g_i)^2]}{\sigma^2_{\text{CRB}}(g,L,i)/\text{SNR}} = 1
\]

(3.36)
and

$$\lim_{\text{SNR} \to \infty} E[\hat{g}(r)_{\text{ML}}] = g.$$  \hspace{1cm} (3.37)

Also, $\hat{g}(r)_{\text{ML}}$ is asymptotically Gaussian, with mean $g$, and covariance $J^{-1}$. For large SNR, the Cramer-Rao bound accurately predicts the performance of the ML estimator, which is then operating in its linear region, as discussed in Section 3.2.

The Cramer-Rao bound depends upon the analytical behavior (the partial derivatives) of the log likelihood function, $\ln(p(r|g))$, in the neighborhood of $\hat{g} = g$, and hence is referred to as a local performance measure. We may, however, incur global estimation errors which cannot be predicted by a local performance bound. We will examine in detail one source of such global error in the ML estimator, the possible existence of subsidiary local minima in $I(\hat{g})$ in addition to the global minimum which we are attempting to locate. This error source is studied by examining the minima of the ambiguity function, $A(g, \hat{g}, L)$, defined to be $I(\hat{g})$ in the absence of noise:

*Our definition is somewhat different from that employed for the delay-Doppler ambiguity function used in radar/sonar signal processing[38]. This difference is due to the fact that the value of $g$ affects the total energy in the response, $s(k|g)$, so that we must consider the complete sum of squares (3.38). On the other hand, for the delay-Doppler ambiguity function signal energy is independent of the parameters to be estimated, so that only the cross-correlation term is significant.
\[ A(q, \hat{q}, L) = I(\hat{q}) \bigg|_{\sigma^2=0} \]

\[ = \sum_{k=1}^{L} (s(k|q) - s(k|\hat{q}))^2 \]

(3.38)

If \( A(q, \hat{q}, L) \) is not unimodal, that is, if \( A(q, \hat{q}, L) \) has local minima at \( \hat{q} \neq q \), a global estimation error may occur even in the absence of noise, as a poor initial guess may cause the numerical algorithm which implements the minimization to converge to the wrong extremum. The objective of our analysis of the ambiguity function is, therefore, to determine whether or not such local minima exist; in their absence, the estimator is unambiguous, in that we are assured that we have chosen the true global minimum of \( I(q) \) as \( \hat{q}(r)_{ML} \) (hence the name "ambiguity function").

Unfortunately, there does not appear to be a simple, general analytic form for \( A(q, \hat{q}, L) \), so that the ambiguity function must be computed numerically for each case of interest. This problem is exacerbated by the dimensionality of the domain: \( A(q, \hat{q}, L) \) is a function of \( 2N+1 \) independent variables. We will examine the ambiguity function for the two examples considered in Section 3.2; because of the number of variables involved, the results are easy to present.

For the single-interface example, we can write \( A(q, \hat{q}, L) \) in simple form, using
We then have

$$A(g, g, L) = \sum_{k=1}^{L} \left( s(k|g) - s(k|\hat{g}) \right)^2 \tag{3.40}$$

Although we have been unable to prove analytically that this ambiguity function is unimodal about $g=\hat{g}$, we have obtained conclusive numerical results which show this to be true. $A(g, g, L)$ is plotted in contour form for $L=2, 4,$ and $16$ in Figures 3.6, 3.7, and 3.8, respectively.

The ambiguity function for the two-interface example is also unimodal about $g=\hat{g}$. We obtained this result, as in the previous example, by numerical computation of the ambiguity function; a general analytical proof is even further beyond our grasp in this example than in the previous. Because five independent variables, $g_1, g_2, \hat{g}_1, \hat{g}_2,$ and $L,$ are involved, we are unable to present as comprehensive a graphical demonstration of the unimodal character of the two-interface ambiguity function as we were for the single-interface case. As an example, we have plotted $A(g, \hat{g}, L)$ in
Figure 3.6 -- Contours of the single-interface ambiguity function, $A(g, \hat{g}, 2)$
Figure 3.7 -- Contours of the single-interface ambiguity function, $A(g, g, 4)$. 
Figure 3.8 -- Contours of the single-interface ambiguity function, $A(g,\hat{g},16)$. 
contour form for \( q = (0.1, 0.1)^T \), and \( L = 4, 8, \) and \( 16 \), in Figures 3.9, 3.10, and 3.11 respectively; the ambiguity function is evidently unimodal in these cases.

As we have seen, it is difficult at best to analyze the ambiguity, or global error properties of an estimator, particularly when more than a few parameters are involved. Nonetheless, when selecting an estimator for use in a practical situation, one must consider both its local behavior as characterized, in the high SNR regime, by the Cramer-Rao bound, as well as its global behavior, as typified by the ambiguity function.

We add, in closing our discussion of global estimator error, that, although we have shown that both the single- and double-interface ambiguity functions are unimodal, this cannot be taken to imply that the ambiguity functions for higher-order models are necessarily unimodal as well.

In this Section, we have presented an optimal (ML) estimator for the reflection coefficients of a discretely-layered medium, discussed some of its general properties, and introduced the concept of global estimator accuracy, as typified by the ambiguity function, in counterpoint to local estimator behavior, as described by the Cramer-Rao bound.

3.4 The PEF Estimator

It is a commentary on the complexities associated with
Figure 3.9 -- Contours of the two-interface ambiguity function, \( A(g, g, 4) \), for \( g = (0.1, 0.1)^T \).
Figure 3.10 -- Contours of the two-interface ambiguity function, $\Lambda(g_1, g_2, g_3)$, for $g_3 = (0.1, 0.1)^T$. 

\[ g_1, g_2, g_3 \]
truly optimal signal processing or estimation schemes that they are often not implemented in practice, but are supplanted by simpler, sub-optimal techniques. In this Section, we will present such a sub-optimal scheme for the estimation of reflection coefficients in a discrete, and equal travel-time layer medium. We have termed this estimator the "prediction-error filter" (PEF) estimator, since it is implemented using a modification of the same algorithm (the Wiener-Levinson algorithm) as that used in stochastic filtering theory to compute a discrete-time prediction-error filter from covariance information. We emphasize, however, that this estimator is not directly related to stochastic filtering theory, and, in particular, possesses none of the optimal properties associated with the Wiener filter.

We will not derive the estimator, but will simply present it; the interested reader is directed to references [19]-[23]. Suppose that there is no observation noise (the problem is deterministic) and that a discrete, equal travel-time layered medium is excited by an impulse source, \( d(k) = d(0) \delta_{k0} \). We know that the response of the medium, \( s(k|g) \) obeys

\[
s(k|g) = 0, \quad k = 2m+1,
\]

so that it is sufficient to consider \( \hat{s}(k|g) = s(2k|g) \), equiv-
alent to using a sampling interval of $\Delta \xi$, rather than $\Delta t$. We may obtain the sequence of reflection coefficients of the medium, $g_1, \ldots, g_N$, by processing $\tilde{s}(1|g), \ldots, \tilde{s}(N|g)$ according to the following algorithm:

$$g_k = \left[ \tilde{s}(k|g) - \sum_{m=1}^{k-1} \tilde{s}(k-m|g) \cdot \gamma_m \right] / E_{k-1}$$

(3.41a)

$$\gamma_{(k)} = -g_k$$

(3.41b)

$$\gamma_{(k)} = -g_k \gamma_{(k-1)}, \quad 1 \leq m \leq k-1$$

(3.41c)

$$E_k = (1-g_{(k-1)}^2) E_{k-1}$$

(3.41d)

where $E_0 = d(0)$.

This is a modification of the Wiener-Levinson algorithm for the iterative solution of the normal equations,

$$\begin{bmatrix}
d(0) & \tilde{s}(1|g) & \ldots & \tilde{s}(N-1|g) \\
\tilde{s}(1|g) & d(0) & \ldots & \tilde{s}(N-2|g) \\
\vdots & \vdots & \ddots & \vdots \\
\vdots & \vdots & \ddots & d(0)
\end{bmatrix} \begin{bmatrix}
\gamma_1^{(N)} \\
\vdots \\
\gamma_N^{(N)}
\end{bmatrix} = - \begin{bmatrix}
\tilde{s}(1|g) \\
\vdots \\
\tilde{s}(N|g)
\end{bmatrix}$$

(3.42)

which arise in the computation of a prediction-error filter.
[39], [40], hence our choice of estimator name. A modified version of the algorithm is required because we are considering pressure wave propagation, in which the surface reflection coefficient is \(-1\); for velocity wave propagation, the algorithm we require and the Wiener-Levinson algorithm are identical. We now have a deterministic relationship, (3.41), which allows us to recover the reflection coefficients of a medium from the medium's noiseless impulse response.

This deterministic relationship defines our sub-optimal (PEF) estimator: given noisy observations of the response of a discrete medium to an impulse of known amplitude, we apply the algorithm (3.41) to the noise-corrupted data, \(r(k)\):

\[
\hat{g}_k(\tau)_{PEF} = \frac{\sum_{m=1}^{k-1} r(k-m|\tau) \gamma_m}{\sum_{m=1}^{k-1} r(k-m|\tau) \gamma_m} / E_{k-1}
\]  

(3.43a)

\[
\gamma_{k} = -\hat{g}_k(\tau)_{PEF}
\]  

(3.43b)

\[
\gamma_m = \gamma_m - \hat{g}_k(\tau)_{PEF} \gamma_{k-m}, \quad 1 \leq m \leq k-1
\]  

(3.43c)

\[
E_k = (1 - \hat{g}_{k-1}(\tau)_{PEF}^2) E_{k-1}
\]  

(3.43d)

where \(E_0 = d(0)\), and \(\tilde{r}(k) = r(2k)\). In other words, the estimator deals with the problem of observation noise by ignoring it, or by assuming that it is sufficiently small that it will not affect the estimate.
Such a cavalier approach is sure to cause problems, and it is easy to see that the PEF estimator is sub-optimal in a very explicit sense: it may be interpreted as a conditional maximum-likelihood estimator. The received signal, \( \tilde{r}(k) \), contains the first return from interface \( k \), and has the form

\[
\tilde{r}(k) = d(0)g_k \prod_{i=1}^{k-1} (1-g_i^2) + F_{k-1}(2k) + n(2k) \tag{3.44}
\]

where \( F_{k-1}(2k) \) is the response at time \( 2k \) given \( g_1, \ldots, g_{k-1} \) alone. The ML estimate of \( g_k \), given \( \tilde{r}(k) \), conditioned on a knowledge of \( g_1, \ldots, g_{k-1} \), is

\[
\hat{g}_k = \frac{\tilde{r}(k)-F_{k-1}(2k)}{d(0) \prod_{i=1}^{k-1} (1-g_i^2)} \tag{3.45}
\]

But this is the exact expression (3.43a) evaluated by the PEF estimator to obtain \( \hat{g}_k(\tilde{r})_{\text{PEF}} \) at the kth step of the algorithm, and our contention is proven. Observe that the estimate's computation, however, is conditioned on the previously obtained estimates, rather than on the true reflection coefficients. In this procedure of sequential, conditional ML estimation used by the PEF estimator, each successive reflection coefficient estimate is affected by the errors in all the previous estimates, so we expect that
the accuracy of $\hat{g}_k(r)_{\text{PEF}}$ will deteriorate as $k$ increases.

We see as well that the PEF estimator is suboptimal because of its sequential nature, noted above. The estimate $\hat{g}_k(r)_{\text{PEF}}$ is formed at step $k$ of the algorithm, based on $\tilde{r}(1), \ldots, \tilde{r}(k)$, and is never recomputed. The PEF estimator, therefore, does not take advantage of the information which may be obtained from intra-layer reverberations, as does the ML estimator. We might typify the PEF estimator as a "filtering," or real-time processor, and the ML as a "smoothing," or off-line processor.

In addition to its sub-optimal character, a practical stumbling block to the use of this algorithm is its requirement for the medium impulse response. It may well happen that the source waveform does not have a causal, stable inverse, so that it is impossible to deconvolve source and medium impulse response, except in some approximate sense, which will, in turn, distort the resulting estimate in an unknown way. The ML estimator, on the other hand, may be used directly with an arbitrary source waveshape.

The one evident advantage of the PEF estimator over the ML is that it is algorithmic. Given an observation, we can be assured of obtaining the PEF estimate for the reflection coefficients in a fixed number of computational steps; we cannot make such a general statement about the ML estimator, founded as it is on a gradient search algorithm.
The PEF estimator does answer one technical question which we have heretofore avoided: (3.41) demonstrates that it is indeed possible to invert the (impulse) reflection response of a medium to obtain its reflection coefficients (its structure). We have, up to this point, implicitly assumed that this was true; now it is proven. Although we are treating discrete media, there is a parallel result for continuous media [22], in which the Wiener-Hopf integral equation plays the same role as the normal equations do here. Hence, both the discrete and continuous medium models of Chapter 2 are seen to be identifiable from acoustic reflection data.

In this Section, we have presented a sub-optimal, ad-hoc, estimator for the discrete medium structure estimation problem, and have discussed its potential performance and limitations. Although its practical utility seems in doubt, it has provided us with a proof that a medium's structure can indeed be estimated from reflection observations, in the limited sense in which we use the term "structure" in this research.

3.5 Monte Carlo Tests

In many cases, the performance (estimate variance) of a reasonably-structured nonlinear estimator, operating in the high SNR region, will approach that predicted by the Cramer-
Rao bound; such estimators are termed asymptotically efficient. In particular, we saw in Section 3.3 that the ML estimator is asymptotically efficient. Indeed, one of the major reasons for considering the Cramer-Rao bound was its ability to provide (asymptotically) accurate performance measures for a broad class of nonlinear estimators, whose performance was at best difficult to compute directly. In this Section, we will present the results of some limited Monte Carlo experiments in order to demonstrate the asymptotic properties of the ML and PEF estimators, and to compare their performance.

We will also be interested in determining the threshold SNR for each estimator, that SNR above which the estimator's performance is well-approximated by the bound value. This threshold value is usually defined [41] as that SNR at which the estimator variance exceeds the bound value by 1 dB:

\[ 10 \log_{10} \left[ \frac{\mathbb{E}[(\hat{g}_i(r) - g_i)^2]}{\sigma^2_{\text{CRB}}(g_i, L, i) / \text{SNR}} \right] = 1 \]  

(3.46)

or

\[ \frac{\mathbb{E}[(\hat{g}_i(r) - g_i)^2]}{\sigma^2_{\text{CRB}}(g_i, L, i) / \text{SNR}} = 1.26 \]  

(3.47)
Clearly, threshold is a function of $g$ and $L$, and depends upon the particular element of $g$ (reflection coefficient) being estimated. Although it is sometimes possible to compute a value for threshold analytically, both the ML and PEF estimators are sufficiently nonlinear that this generally cannot be done, and Monte Carlo techniques are required.

Monte Carlo tests were performed for specific cases of the two examples we have scrutinized previously, the one- and two-reflector discrete medium models. For the single reflector problem, we set $g=0.1$, and simulated the medium's response to a discrete impulse of unit amplitude (so $\text{SNR} = 1/\sigma^2$). We used $L=8$ (three multiple observations), effectively an infinite observation interval for this reflection coefficient value. Using (3.28),

$$\sigma^2_{\text{CRB}} (g,L) = \sigma^2_{\text{CRB}} (0.1,8) \approx 0.96$$

From (3.43), we see that

$$\hat{g}_1(r)_{\text{PEF}} = \bar{r}(1) = r(2), \quad (3.48)$$

so that the PEF estimate variance is equal to $1./\text{SNR}$; hence the PEF estimator never exhibits threshold in the sense described above, since the ratio (3.47) is not a function of $\text{SNR}$, and in fact equals $1./\sigma^2_{\text{CRB}} (g,L)$. In view of this, and
the results of Figure 3.2, we see that the PEF estimator
cannot be efficient if \( L \geq 4 \).

In Figure 3.12, we have plotted the results of the
single-interface Monte Carlo experiment; both bias, defined
as \( \mathbb{E}[\hat{g} - g] = \mathbb{E}[\hat{g}] - 0.1 \), and estimate variance, relative to
the value predicted by the Cramer-Rao bound, are shown. We
see that, as predicted above, the PEF variance is a constant,
greater than the Cramer-Rao bound. The behavior of the ML
estimate variance is quite disturbing in view of our previous
comments concerning the Cramer-Rao lower bound: although the
ML estimator achieves the Cramer-Rao bound almost exactly for
SNR greater than 50, below that point, it performs better
than the Cramer-Rao bound, rather than worse. This striking
behavior is due to the presence of bias in the estimator.
Recall that we introduced the Cramer-Rao bound in Section 3.2
as a lower bound on the performance of an unbiased estimator.
If an estimator is indeed biased, we should not expect the
bound to hold, and we are guaranteed only that the ML
estimator is asymptotically unbiased (3.37). It is not sur-
prising, therefore, that we obtain unexpected results in the
low SNR region.

A form of the Cramer-Rao bound does exist, however,
for the instance in which the estimator has a known bias.
For the problem of estimating a single parameter, this is [34]
Figure 3.12 -- Monte Carlo test results for a single-interface medium, with $g = 0.1$ and $L = 8$. These results are based on 454 simulation runs.
\[
E[(\hat{g}(x) - g)^2] \geq \sigma^2 \int_0^T d^2(t) dt \cdot \frac{(1 + \frac{db(g)}{dg})^2}{\sum_{k=1}^L \frac{3s(k|g)}{\delta g}} \tag{3.49}
\]

where we have used the notation of (3.21), and

\[
b(g) = E[\hat{g} - g] = E[\hat{g}] - g \tag{3.50}
\]

Note that \(b(g)\) may be a function of SNR as well as of \(g\). We see that the value of the bound is affected by the derivative of the bias, rather than the bias itself.

Equation (3.49) is not useful unless we have an analytic form for \(b(g)\), which we generally do not. We have attempted to estimate \(db(g)/dg\) for the case of Figure 3.12 using Monte Carlo tests, but, in the interesting region of low SNR, have been unable to obtain a bias estimate sufficiently stable to enable accurate computation of the derivative.\(^*\) The estimates we did obtain for the derivative were less than zero, however, which is consistent with (3.49) and the ML estimator behavior observed in Figure 3.12. It is interesting that the effect of bias in this instance was to make the ML estimator more robust in the low SNR region, at

\(^*\) This is due solely to the short cycle length of the random number generator available to us. There is no fundamental reason why \(b(g)\) and its derivative cannot be estimated as accurately as desired.
small cost in actual bias. We note that this is most certainly not generally true for this estimator, although some estimators have been designed to trade bias for reduced variance [59].

Monte Carlo results for the two-interface example are shown in Figure 3.13a, and 3.13b, for the upper and lower interfaces, respectively. For this example, we set $g = (0.1, 0.1)^T$, used an observation length of 16 samples (effectively infinite), and, again, simulated the medium's response to a discrete impulse of unit amplitude. Using (3.34), we find

$$\sigma^2 \left( g, 16, 1 \right) = 0.97$$

$$\sigma^2 \left( g, 16, 2 \right) = 0.98$$

Since $\hat{g}_1(r)_{\text{PEF}}$ is still given by (3.48), its variance equals $1./\text{SNR}$, and is independent of SNR; hence, for $L \geq 6$, the PEF estimator cannot be efficient (recall the bound results of Figures 3.3-3.5). We see in Figure 3.13a that the PEF estimator for the upper reflection coefficient behaves as predicted. The ML estimator, on the other hand, while it is asymptotically efficient, exhibits unexpected behavior for low SNR. As SNR decreases, estimator variance diverges from the bound value, but barely exceeds threshold before it
Figure 3.13a -- Monte Carlo test results for a two-interface medium, with \( q = (0.1, 0.1)^T \), and \( L = 16 \). This is the data for estimation of \( q_1 \). These results are based on 454 simulation runs.
Figure 3.13b — Monte Carlo test results for a two-interface medium, with $g = (0.1, 0.1)^T$ and $L = 16$. This is the data for estimation of $g_2$.

These results are based on 454 simulation runs.
begins to improve, eventually falling below the Cramer-Rao lower bound. Here we see a combination of estimator divergence, which we expect in a nonlinear estimator, and bias effects, which we observed in the single-interface ML estimator. Again, the result is an estimator which is quite robust. We see graphic evidence, however, that the ML estimator is not optimal in the sense of being the minimum-variance estimator for $g_1$ (as we discussed at the outset of Section 3.3): for intermediate values of SNR, the ad-hoc PEF estimator has smaller variance.

The estimate $\hat{g}_2(r)_{\text{PEF}}$ is given by

$$\hat{g}_2(r)_{\text{PEF}} = \frac{\tilde{r}(2) + \hat{g}_1(r)_{\text{PEF}} \tilde{r}(1)}{d(0) (1 - \hat{g}_1(r)^2_{\text{PEF}})}$$

(3.51)

which may be verified by working through the algorithm (3.43). In Figure 3.13b, we see that this nonlinear estimator does exhibit divergence, with threshold occurring at SNR $= 50$. The variance of $\hat{g}_2(r)_{\text{ML}}$ has the same qualitative behavior we noted in Figure 3.13a for $\hat{g}_1(r)_{\text{ML}}$; although this estimate is robust, the bias which is incurred for low SNR is probably unacceptable for most applications. We note that both PEF and ML estimators are less robust for the deep reflector than they are for the shallow.

It is interesting that, for the two-interface example, the PEF estimator performs approximately as well as, or better than, the ML estimator, for SNR values as small as 30.
This indicates that, if we have a moderately large SNR, we are probably best advised to implement the simple, PEF estimator rather than the more complicated "optimal" ML estimator. Of course, when SNR is small, we have little choice but to employ the ML estimator, with its robust behavior, but also its attendant bias.

In this Section, we have used Monte Carlo experiments to demonstrate the ability of the Cramer-Rao bound to predict the performance of nonlinear estimators operating in the high SNR region. We have also compared the performance of the ML and PEF estimators for the single- and double-interface examples. We have seen that the ML estimator is not optimal in the minimum variance sense, as we noted in Section 3.3. Although the ML estimator is more robust than the PEF for these examples for small SNR, this robustness is achieved because of the presence of estimate bias, which may be unacceptably large. As a practical matter, one would surely want to operate in the asymptotic region (high SNR) if possible, to ensure a small estimate variance, and a small bias. In this region, however, the advantage of the ML over the PEF estimator is slight (depending upon the value of $g$), so that it may well be, in such a circumstance, that the simpler suboptimal estimator is preferable to the more complicated optimal estimator.
4.1 Introduction

In Chapter 3, we considered estimation bounds for media whose structure could be modelled as a set of (equally spaced) discrete reflectors: the medium was identified by estimating the interface reflection coefficients. As we have discussed previously, such a model, in which only the prominent discrete reflectors of the medium are included, is inadequate when a continuous description of medium properties is required. The application for which structural information is being sought, of course, must be the determining factor in deciding whether a discrete or continuous model is appropriate.

For example, consider the continuous medium structure of Figure 4.1, where propagation velocity is plotted as a function of depth (actual depth, not two-way travel time), and we assume that density is constant. This layer structure, termed an isogradient layer, is a simplified model for thick, shallow-water sediment layers [43]. We might represent this structure as a single reflector with reflection coefficient $g$ (the single-interface example of Chapter 3 would be directly applicable). This model is in fact used when the objective is to characterize the seafloor in terms of its
Figure 4.1 -- Velocity profile of an isogradient layer. Note that propagation velocity is a function of actual depth in the medium, not two-way travel time.
reflectivity ([8], for instance), a simple model for describing the structure's interaction with normally incident waves. Acoustic waves non-normally incident on the layer, however, do not experience the simple specular reflection predicted by the discrete-reflector model. Rather, they are refracted upwards, away from the region of higher velocity. Such refraction phenomena are important, for instance, when studying acoustic propagation paths in which there is significant interaction with the bottom [5], [26]. We see that, for this application, the simple discrete reflector model does not represent the phenomenon of interest, and hence is not useful.

In this Chapter, we will consider estimation bounds for the class of continuous media whose structure may be parametrized by a finite-dimensional set of unknown, non-random parameters, \( \alpha \). Our results are directed toward those structures for which the wave equation may be solved analytically so that \( s(t) \), the noiseless medium response of (2.22), is parametrized by \( \alpha \). The number of parametric structures (which have some physical meaning) for which the propagation equations may be solved in closed form is quite limited [42], but does include several of practical interest. In particular, the isogradient layer structure's response may be obtained analytically, parametrized by the layer gradient, \( w \), and the layer "reflection coefficient," \( g \).
In some applications, such as predicting refraction effects in sediment layers modeled as isogradient, it is sufficient to obtain estimates of the structure parameters themselves (and hence to consider estimation bounds for them). On the other hand, one might wish to obtain an estimate of the medium structure as a continuous function of depth, as might be the case if one were surveying for the placement of some offshore facility [6], [7]. In this case, it is appropriate to ask for the pointwise variance of the estimated function, e.g., \( \text{Var}[\hat{g}(\xi)] \). In Section 4.2, we will obtain expressions for the Cramer-Rao bound for \( \alpha \), given a continuous observation of the form (2.23), when \( s(t) \) has a known, parametric dependence on \( \alpha \). We will also indicate how a pointwise bound may be obtained for the continuous structure which depends parametrically on \( \alpha \), given a Cramer-Rao estimation bound for \( \alpha \).

In Section 4.3, we will examine the problem of estimating the parameters of the isogradient layer of Figure 4.1, using the theoretical results of Section 4.3. In particular, we will demonstrate that the Cramer-Rao bound may be used to select an appropriate source waveform, and to define an experimental geometry for actually performing such a measurement. We also reach the interesting conclusion that improved estimation accuracy is achieved by using a low-frequency, rather than a high-frequency source. This is in
contrast to the conventional philosophy which holds that layer structures are invisible to waves of sufficiently low frequency.

4.2 Theoretical Results

For a continuous medium whose structure is parametrized by $\alpha$, we rewrite the observation equation (2.23) to emphasize the parametric dependence of the received signal on $\alpha$:

$$r(t) = s(t|\alpha) + n(t), \quad 0 \leq t \leq T,$$

(4.1)

where we have used

$$s(t|\alpha) = p(t, 0)$$

(4.2)

in (2.23)*. The Fisher information matrix for the estimation of $\alpha$, an unknown, non-random parameter vector, given the continuous observation $r(t)$ is [31]

$$(J)_{ij} = E \left[ \frac{\partial \ln \Lambda(r(t), \alpha)}{\partial \alpha_i} \cdot \frac{\partial \ln \Lambda(r(t), \alpha)}{\partial \alpha_j} \right]$$

(4.3a)

$$= -E \left[ \frac{\partial^2 \ln \Lambda(r(t), \alpha)}{\partial \alpha_i \partial \alpha_j} \right]$$

(4.3b)

*The reader should be careful not to confuse $s(t|\alpha)$, the noiseless response of a parametrized continuous medium with $s(k|g)$, the noiseless response of a discrete, equal travel-time layer medium. The two may be most easily distinguished by their time parameter.
where

\[ \ln \Lambda(r(t|\alpha)) = \frac{1}{\sigma^2} \left[ \int_0^T r(t)s(t|\alpha) dt - \frac{1}{2} \int_0^T s^2(t|\alpha) dt \right] \tag{4.4} \]

is the log-likelihood ratio, which plays the same role in (4.3) as \( \ln(p(r|g)) \) does in (3.3). Using (4.4) in (4.3), and recalling the statistical properties of \( n(t) \), described in Chapter 2, we find

\[ (J)_{ij} = \frac{1}{\sigma^2} \int_0^T \frac{\partial s(t|\alpha)}{\partial \alpha_i} \frac{\partial s(t|\alpha)}{\partial \alpha_j} dt \tag{4.5} \]

which is analogous to (3.14), our result for the discrete case. The Cramer-Rao inequality is then

\[ E[(\hat{\alpha}_i - \alpha_i)^2] \geq (J^{-1})_{ii}. \tag{4.6} \]

As discussed in Section 4.1, we might well be concerned with estimating the continuous structure of the medium rather than the parameters upon which it depends. We will, for the sake of concreteness, assume that the continuous "structure" of concern is the reflection coefficient density (2.14), although our result is equally applicable to any function parametrized by \( \alpha \). To emphasize that the continuous structure is parametrically dependent on \( \alpha \), we will use the notation \( g(\zeta; \alpha) \), rather than \( g(\zeta) \). A reasonable question is,
then, if we are given some estimate of the parameter vector, \( \hat{a} \), what can we say about the reflection coefficient density, the continuous structure of the medium? First, it is well known that maximum-likelihood estimation commutes over non-linear operations [31]. Hence, if we are given \( \hat{a}(r(t))_{ML} \), we may form \( \hat{g}(\xi)_{ML} \) as

\[
\hat{g}(\xi)_{ML} = g(\xi; \hat{a}(r(t))_{ML})
\]  

(4.7)

regardless of the functional dependence of \( g(\xi; a) \) on \( a \).

For the purpose of this research, a more significant result is one which allows us to place a Cramer-Rao lower bound on the variance of \( \hat{g}(\xi) = g(\xi; \hat{a}) \), regardless of the sense in which \( \hat{a} \) is formed (in particular \( \hat{a} \) need not be a maximum-likelihood estimate). The bound which results will be a pointwise bound on \( \text{Var}[\hat{g}(\xi)] \). If we define the gradient vector

\[
\nabla_a (g(\xi; a)) = \left[ \frac{\partial g(\xi; a)}{\partial a_1}, \ldots, \frac{\partial g(\xi; a)}{\partial a_q} \right]^T
\]  

(4.8)

then

\[
\mathbb{E}[(g(\xi; \hat{a}) - g(\xi; a))^2] \geq \nabla_a (g(\xi; a))^T J^{-1} \nabla_a (g(\xi; a))
\]  

(4.9)

where \( J \) is given by (4.5).
Equation (4.9) is true without regard to the type of estimate, \( \hat{a} \), involved (although, of course, we assume unbiased estimators throughout), while (4.7) is true only for ML estimators. We emphasize, however, that the threshold of \( \hat{g}(\xi)_{\text{ML}} \) certainly need not be the same as that of \( \hat{a}(r(t))_{\text{ML}} \), i.e., the region over which estimator performance is predicted by the Cramer-Rao bound is not invariant under the transformation (4.7).

We note that (4.9) assumes a particularly simple form when we represent \( g(\xi:a) \) as a linear sum of basis functions:

\[
g(\xi:a) = \sum_{k=1}^{q} a_k \phi_k(\xi) = a^T \phi(\xi)
\]  

where \( \phi(\xi) = [\phi_1(\xi), \phi_2(\xi), \ldots, \phi_q(\xi)] \). Then (4.9) becomes

\[
E[(g(\xi:a_\hat{a}) - g(\xi:a))^2] \geq a^T \phi(\xi) J^{-1} \phi(\xi)
\]

The choice of basis functions in (4.10) is completely arbitrary: they need not be orthogonal, normalized, etc. for (4.11) to hold. Of course, in our problem, \( g(\xi:a) \) must correspond to a structure for which the propagation equations have an analytic solution, which does restrict the representations which we may use.
In this Section, we have presented the Cramer-Rao bound formulae appropriate for the problem of structure estimation in continuous media when the medium response is a known function of a parameter vector, \( \mathbf{a} \), whose components are unknown but nonrandom. We have considered bounds both for the problem of estimating \( \mathbf{a} \), and for the problem of estimating the continuous physical structure of the medium, assumed to depend on \( \mathbf{a} \), such as \( g(\xi; \mathbf{a}) \).

4.3 Example: Structure Estimation for a Bottom Sediment Layer

We now present an example of parametric structure estimation for a case in which the reflection response of the medium may be computed analytically, the isogradient layer model discussed in Section 4.1, and illustrated in Figure 4.1. This example is directly applicable to the practical problem of identifying the structure (the propagation velocity) of an ocean bottom sediment layer, a problem which was a primary motivation for this research. We will compute performance bounds for the estimation of this commonly-used parametric sediment layer model, and will use these computed bound values in the design of a measurement experiment. We believe our results provide an accurate indication of our true ability to estimate the parameters of such structures in the ocean environment.

Before proceeding, we briefly reiterate our discussion
of the model. Models for near-bottom sediment layer velocity structures have been presented in [43]; we will base our investigation on a simplified version of one of these models, appropriate for thick, shallow-water sediment layers. Note that the propagation velocity in Figure 4.1 is expressed as a function of depth in meters, not two-way travel time in the medium. While a simplification, our model is an appropriate one for modelling shallow ducting phenomena, where the properties of the deep structure do not enter. We believe, therefore, that the simple model of Figure 4.1 contains the essential elements of the identification problem, and is valuable in its own right, as well as being a reasonable and practical preliminary step in the analysis of a more complete, but more complicated, model.

In addition to our specification of the form of c(z), we will make several additional assumptions. We first assume that the density is constant throughout the medium; this allows us to solve for the system response analytically. Since [43] indicates that density is essentially constant throughout each layer, it would not be difficult to include realistic density variations in more sophisticated models. The layer is assumed to be situated at a known depth, Z, and is assumed to be isolated, that is, there are no significant deeper reflectors. By assuming a known layer depth, we have ensured that the bound values we will compute are for coher-
ent processors, i.e., the processors do not have to estimate the arrival time of the layer response. This is somewhat unrealistic, but we expect that estimator performance will not be greatly affected by this omission, e.g., 3dB is a typical improvement in performance for coherent over incoherent processing. Finally, we retain the simple source and receiver coupling models of Chapter 2.

The model of Figure 4.1 may be parametrized in several ways; we have chosen to estimate the velocity gradient in the layer, \( w \), and the layer "reflection coefficient," \( g \), given by

\[
g = \frac{c_{\infty} - c_0}{c_{\infty} + c_0},
\]

(4.12)

where \( c_0 \), the propagation velocity in the water layer, is known. We have termed the parameter \( g \) a reflection coefficient because it is the peak amplitude of the layer response; this terminology, to our knowledge, has not been used in the literature. In the notation of Section 4.2, the elements of the parameter vector \( \alpha \) are

\[
\alpha_1 = w,
\]

(4.13a)

\[
\alpha_2 = g.
\]

(4.13b)

We will consider slope values in the range \( 0.5 \, \text{sec}^{-1} \leq w \leq 1.5 \, \text{sec}^{-1} \), and reflection coefficient values in the range \( 0.02 \leq g \leq 0.08 \);
these correspond to \( c_\infty \) in the range 1600-1800 m/sec, and a layer thickness of 40-530 m. These values of \( g \) are representative of actual sediment structures; a more realistic range for \( w \) might be \( 0.8 \text{ sec}^{-1} \leq w \leq 1.2 \text{ sec}^{-1} \), but we have used an extended range to emphasize the qualitative behavior of the bound as a function of \( w \).

We will use \( c_0 = 1535 \text{ m/sec} \), and \( z = 150 \text{ m} \) for our analysis. The impulse response of the layer structure of Figure 4.1, \( G(\nu | \alpha) \), not including the effects of the free surface (without multiples) is well known ([44], Section 148b), and may be conveniently expressed in the frequency domain as

\[
G(\nu | \alpha) = e^{-j2\pi f \tau} \begin{cases} 
\frac{\sinh(\mu \ln \beta)}{2}, & 2\pi f < \frac{W}{2} \\
\frac{\sin(\mu \ln \beta)}{2}, & 2\pi f > \frac{W}{2} 
\end{cases}
\]

(4.14)

where

\[
\mu = \left( \frac{1}{4} - \left( \frac{2\pi f}{W} \right)^2 \right)^{1/2} \quad (4.15)
\]

\[
\beta = \frac{c_0}{c_\infty} \quad (4.16)
\]
\[ \tau_w = \frac{2Z}{c_0}. \] (4.17)

Note that

\[ \lim_{f \to 0} G(f | \alpha) = \frac{c_\infty - c_0}{c_\infty + c_0} = g \] (4.18)

which was our motivation for terming \( g \) a reflection coefficient. We have plotted \( |G(f | \alpha)|/g \), on a normalized frequency scale, in Figure 4.2. The zeroes of \( G(f | \alpha) \) occur at

\[ f_0 = \frac{w}{2\pi} \left( \frac{1}{4} + \left( \frac{k\pi}{\ln \beta} \right)^2 \right)^{1/2}, \quad k=1,2,\ldots \infty \] (4.19)

which, for the parameter values which we are considering, may be written

\[ f_0 \approx \frac{kw}{2|\ln \beta|}, \quad k=1,2,\ldots \infty \] (4.20)

so the layer resonances are approximately uniformly spaced.

We add the water layer multiples to the layer response in the same manner as we did in Chapter 2 for the Bremmer series model. A system block diagram is shown in Figure 4.3. Working out the algebra, we have

\[ S(f | \alpha) = \frac{D(f)G(f | \alpha)}{1+G(f | \alpha)} \] (4.21)
Figure 4.2: Frequency response of an isogradient layer, plotted on normalized scales.
Figure 4.3 -- System model for the reflection response of a medium containing a single isogradient layer. Propagation delay is included in \( G(f|\alpha) \).
where

$$S(f|a) \leftrightarrow s(t|a) \quad (4.22)$$

We have plotted $|S(f|a)| \cdot (1+g)/g$, on a normalized frequency scale, for $g = 0.08$, in Figure 4.4. The evident flattening of the main lobe in comparison with the response of the layer alone (Figure 4.2) is attributed to resonance in the water layer.

We saw in Section 4.2 that the Fisher information matrix for this problem is given by

$$\begin{align*}
(J)_{ij} &= \frac{1}{\sigma^2} \int_0^T \frac{\partial s(t|a)}{\partial a_i} \frac{\partial s(t|a)}{\partial a_j} \, dt \\
&= (4.5)
\end{align*}$$

To simplify our computations, we will assume that the observation interval is infinite ($T \to \infty$). The resulting infinite interval bound value will certainly be a lower bound on the performance attainable with a finite $T$. As well, since the amplitude of successive multiples is proportional to $g^k$, and $g<<1$, the multiples contain little energy in comparison to the direct reflection from the layer, so that the infinite interval bound should differ little from the finite interval bound (as long as the finite interval contains substantially all the direct reflection). With this assumption, we may write
Figure 4.4 -- Total response of medium containing an isogradient layer. Here, $g = 0.08$, $\tau_w = 0.2$, $w = 1.0$. 

$$f \times \frac{2|\ln \beta|}{w}$$
(J)_{ij} = \frac{1}{2 \sigma^2} \int^\infty_0 \frac{\partial S(t|\alpha)}{\partial \alpha_i} \frac{\partial S(t|\alpha)}{\partial \alpha_j} dt

= \frac{1}{\sigma^2} \int^\infty_{-\infty} \frac{\partial S(f|\alpha)}{\partial \alpha_i} \frac{\partial S^*(f|\alpha)}{\partial \alpha_j} df, \hspace{1cm} (4.23)

where we made use of \( d(t) = 0, \; t<0 \), and Parseval's Theorem.

Using the conjugate symmetry of \( \frac{\partial S(f|\alpha)}{\partial \alpha_i} \), (4.23) becomes

\[ (J)_{ij} = \frac{2}{\sigma^2} \text{Re} \left[ \int^\infty_0 \frac{\partial S(f|\alpha)}{\partial \alpha_i} \frac{\partial S^*(f|\alpha)}{\partial \alpha_j} df \right] \hspace{1cm} (4.24) \]

Using (4.21), we find

\[ \frac{\partial S(f|\alpha)}{\partial \alpha_i} = \frac{D(f)}{(1+G(f|\alpha))^2} \cdot \frac{\partial G(f|\alpha)}{\partial \alpha_i} \hspace{1cm} (4.25) \]

so that, at last, (4.24) is

\[ (J)_{ij} = \frac{2}{\sigma^2} \text{Re} \left[ \int^{\infty}_{0} \frac{D(f)D^*(f)}{|1+G(f|\alpha)|^4} \frac{\partial G(f|\alpha)}{\partial \alpha_i} \frac{\partial G^*(f|\alpha)}{\partial \alpha_j} df \right] \hspace{1cm} (4.26) \]

We will evaluate \( \bar{J} \) by integrating the right-hand side of (4.26) numerically, given analytic formula for \( D(f), G(f|\alpha), \) and \( \frac{\partial G(f|\alpha)}{\partial \alpha_i} \).

Simple differentiation of (4.14) gives the required partial derivatives,
where we have suppressed the leading phase terms in each expression, \( \exp(-j2\pi f_\text{w} t) \). Note that while \( |G(f|a)| \to 0 \) as \( f \to \infty \), \( |\partial G(f|a)/\partial a_1| \to 0 \) as \( f \to \infty \). Hence if we attempt to evaluate (4.26) for an impulse source (which is unrealistic, since it has infinite energy), the integral would diverge.
The implication here is that we may obtain arbitrarily accurate estimation if we use a source which has infinite energy (or, equivalently, an infinite bandwidth). We must, therefore, evaluate performance bounds using realistic (bandlimited) source models. We will compute the performance bound for a suite of three sources: a 3.5 kHz echo sounder (a tone burst), an explosive source (an exponential waveshape), and an airgun (a damped sinusoid). We have selected these three sources for analysis because they are commonly available, and could realistically be applied to the problem at hand.

Before presenting the numerical details, we will highlight one important aspect of the results. It is evident from the layer response plot of Figure 4.2 that the layer is transparent to high frequencies (\(|G(f)\alpha|\) falls off as \(1/f\)) so that sources whose energy is concentrated in a high frequency band, such as the 3.5 kHz echo sounder, will give rise to a weaker layer response than will sources which have substantial low-frequency components, such as an airgun. In addition, our numerical results indicate that estimation accuracy improves as source energy is concentrated in the low-frequency region. It is interesting that both these observations run counter to a generally-accepted rule of thumb in exploration seismology, that layers whose thickness is less than a quarter of a source wavelength are too "thin," and cannot be detected. This statement is obviously a sim-
plification, since the "hardness" of the layer must also be taken into account, e.g., if our "layer" were an air bubble, we should expect to be able to detect it without difficulty, regardless of thickness. Our approach may provide a method for quantitatively answering the question, "How thin must a 'thin' layer be?"

Before we compute the Cramer-Rao bound for each of the three sources, we will introduce some notation. We saw in Chapter 3 (in (3.29), for instance) that the Cramer-Rao bound is inversely proportional to the signal-to-noise ratio, SNR. In this case, we define

\[
\text{SNR} = \frac{\int_0^\infty d^2(t) dt}{\sigma^2}, \tag{4.29}
\]

so that the Cramer-Rao bound matrix $\mathbf{J}^{-1}$, may be written as

\[
\mathbf{J}^{-1} = \frac{\mathbf{J}^{-1}}{\text{SNR}}, \tag{4.30}
\]

where $\mathbf{J}$ is the Fisher information matrix of (4.24) computed for the case of unit signal to noise ratio (SNR=1). We will compute $\tilde{\sigma}_w^2$ and $\tilde{\sigma}_g^2$, defined as

\[
\tilde{\sigma}_w^2 = (\mathbf{J}^{-1})_{11}, \tag{4.31a}
\]

\[
\tilde{\sigma}_g^2 = (\mathbf{J}^{-1})_{22}. \tag{4.31b}
\]
The Cramer-Rao bound variances for a given SNR are then obtained from the computed values, \( \tilde{\sigma}_w^2 \) and \( \tilde{\sigma}_g^2 \) by dividing by SNR. For instance,

\[
(\mathbf{J}^{-1})_{11} = \frac{(\mathbf{J}^{-1})_{11}}{\text{SNR}} = \frac{\tilde{\sigma}_w^2}{\text{SNR}}.
\]  

(4.32)

As well, this allows easy comparison of the three sources, since the values we will compute are indicative of the relative performance of each source for any given source energy and background noise power.

The waveform used for the 3.5kHz echo sounder source was a sinusoidal tone burst

\[
d(t) = \begin{cases} 
  d_b \sin(2\pi(3500)t), & 0 \leq t \leq \tau_d \\
  0, & \text{elsewhere}
\end{cases}
\]  

(4.33)

where \( \tau_d \) is the pulse length of the tone burst, either 0.5 msec or 1.0 msec (typical values for commercial equipment), and \( d_b \) is chosen so that \( d(t) \) has unit energy. (We also assume \( \sigma^2 = 1 \) throughout, so that SNR = 1, as desired).

The Cramer-Rao bound variances for estimation of layer slope, \( \tilde{\sigma}_w^2 \), and reflection coefficient, \( \tilde{\sigma}_g^2 \), were computed for \( 0.5 \text{ sec}^{-1} \leq w \leq 1.5 \text{ sec}^{-1} \), and \( c_\infty = 1600 \text{ m/sec}, 1700 \text{ m/sec}, \) and \( 1800 \text{ m/sec} \) (corresponding, approximately, to \( g = 0.02 \), \( g = 0.05 \), and \( g = 0.08 \), respectively), and \( \tau_d = 0.5 \text{ msec} \) and
1.0 msec. We have plotted the bound standard deviations, e.g., $\tilde{\sigma}_w$, in Figures 4.5, 4.6, and 4.7. (The reader is warned that we have used split scales, one for $\tilde{\sigma}_w$, and a quite different one for $\tilde{\sigma}_g$, in these and subsequent plots.)

We note that performance is poorer for the shorter (0.5 msec) tone burst, often used for "high resolution" discrimination of adjacent layers.*

The behavior of $\tilde{\sigma}_g$ as a function of $c_\infty$ is somewhat counter-intuitive, as we might expect that, as the medium got "harder" (as $g$ increased), the increase in response amplitude would give improved performance. For fixed $w$, however, as $g$ increases, so does the thickness of the layer, and the less the layer response resembles that of a step change in velocity (the more smeared out it is). This effect dominates the amplitude increase. On the other hand, as $w$ increases for a fixed $g$ ($c_\infty$), the thinner the layer, the more it resembles a step, and the smaller (better) is $\tilde{\sigma}_g$.

We may explain the behavior of $\tilde{\sigma}_w$ as a function of $g$ and $w$ in a similar manner. The thinner the layer (the larger

---

*One aspect of the problem which is not evident from these and subsequent numerical results is that the estimates $\hat{w}$ and $\hat{g}$ are very strongly correlated, so that the Fisher information matrix may lose its positive definite character because of numerical error. It was found that, for the tone-burst source (4.33), double-precision computation was required in the numerical integration of (4.26) in order to preserve the positive definite character of the matrix. The use of double precision improved the accuracy of the results for the other two sources somewhat, but the definiteness of the matrices was never a problem.
Figure 4.5 -- Cramer-Rao bound for isogradient layer, tone burst source, $c_\infty = 1600 \text{ m/sec}$. 

$3.5 \text{ kHz tone burst}$
Figure 4.6 -- Cramer-Rao bound for isogradient layer, tone burst source, $c_\infty = 1700$ m/sec.
\( c_\infty = 1800 \text{ m/sec} \)

3.5 kHz tone burst

**Figure 4.7 -- Cramer-Rao bound for isogradient layer, tone burst source, \( c_\infty = 1800 \text{ m/sec} \).**
w), the more step-like its response, the less distributed scattering we observe, and the larger (worse) is \( \tilde{\sigma}_w \). As \( c_\omega(g) \) increases for fixed \( w \), the thicker the layer, the more scattering occurs, and the smaller (better) is \( \tilde{\sigma}_w \).

These trends in \( \tilde{\sigma}_w \) and \( \tilde{\sigma}_g \) as functions of \( g \) and \( w \) are characteristic of the layer structure itself, rather than of the source involved. We will observe the same results in our examination of the performance associated with the other two sources.

The explosive source is modelled by the exponential waveform

\[
d(t) = d_e \exp(-2\pi f_e t)u_-(t)
\]  

(4.34)

where \( f_e \) is a bandwidth measure (the 3-dB-down point of \( D(f) \)), and \( d(t) \) is scaled by \( d_e \) to have unit energy. This is a crude model for an explosive source, since it ignores the existence of the bubble pulse phenomenon; it is substantially correct, however, and is adequate for our purposes. We will consider a range of values for \( f_e \), 500-2000 Hz, which is typical for charge of small to moderate size [53].*

The bound variances \( \tilde{\sigma}_w \) and \( \tilde{\sigma}_g \) are shown as functions

\[\text{In fact, } f_e \text{ varies with range, as well as with depth of explosion and charge size, but we will ignore this. The range of } f_e \text{ quoted is sufficient to include, in broad-brush fashion, most cases of interest.}\]
of \( w \) in Figures 4.8, 4.9, and 4.10 respectively. The behavior of \( \tilde{\sigma}_w \) and \( \tilde{\sigma}_g \) as functions of \( g \) and \( w \) is the same as we noted for the tone burst source, substantiating our contention that this behavior is related to the layer structure rather than the source used to illuminate the layer.

As we discussed above, performance improves as the available energy is concentrated in the lower frequencies (as \( f_e \) decreases). Over the range of frequencies considered, in fact, both \( \tilde{\sigma}_w^2 \) and \( \tilde{\sigma}_g^2 \) are proportional to \( f_e \). We conjecture that the relationship obtains because the source spectrum is approximately constant, with amplitude proportional to \( 1/\sqrt{f_e} \), over the range of frequencies where the system response is significant. Thus \( \mathcal{I} \), as given in (4.29), is proportional to \( 1/f_e \), and \( \mathcal{I}^{-1} \) to \( f_e \). This relationship, of course, breaks down for small \( f_e \), where the full bandwidth of the layer response is not excited by the source.

Quantitatively, we see, on comparison with the results for the tone-burst source, that, for a given SNR, use of the explosive source gives a performance improvement of about an order of magnitude (10 dB) for a given \( w \) and \( g \).

The airgun source is modelled by the damped-sinusoidal waveform

\[
d(t) = d_0 e^{-5t} \sin(2\pi(10)t) \mu_{-1}(t)
\]

(4.35)
Figure 4.8 -- Cramer-Rao bound for isogradient layer, explosive source, $c_\infty = 1600$ m/sec.
Figure 4.9 -- Cramer-Rao bound for isogradient layer, explosive source, $c_\infty = 1700$ m/sec.
Figure 4.10 -- Cramer-Rao bound for isogradient layer, explosive source, $c_\infty = 1800$ m/sec.
where, again, $d_0$ is chosen so that $d(t)$ has unit energy. The bubble pulse center frequency of 10 Hz is perhaps a bit low, but is physically reasonable. Equation (4.35) is a smooth approximation to what is generally a rather less continuous waveform, but it retains the essential spectral character typical of an airgun.

The performance bounds for the airgun source are shown in Figures 4.11, 4.12, and 4.13. We first note that $\tilde{\sigma}_w$ and $\tilde{\sigma}_g$ are no longer monotone functions of $w$, as they were for the tone burst and explosive sources. This behavior may be directly related to the interaction of the peak of the source spectrum, at 10 Hz, with the zeroes of the layer response, $G(f|\alpha)$,

$$f_0 = \frac{kw}{2|\ln \beta|}, \quad k=1,2,\ldots,\infty \quad (4.20)$$

When a response zero is in the vicinity of 10 Hz, the source is less effective at illuminating the layer, and performance is degraded. For instance, in Figure 4.11, $\tilde{\sigma}_w$ has a local peak (poor performance) at $w \approx 0.85$. Substituting in (4.20), we find that the first response zero ($k=1$) occurs at

$$f_0 = \frac{1 \cdot 0.85}{2|\ln \frac{1535}{1600}|} \approx 10.2 \text{ Hz}$$

which is as we predicted. Of course, despite the non-monotone
Figure 4.11 -- Cramer-Rao bound for isogradient layer, airgun source, $c_\infty = 1600$ m/sec.
\( c_\infty = 1700 \text{ m/sec} \)

Airgun source

Figure 4.12 -- Cramer-Rao bound for isogradient layer, airgun source, \( c_\infty = 1700 \text{ m/sec} \).
Figure 4.13 -- Cramer-Rao bound for isogradient layer, airgun source, \( c_\infty = 1800 \text{ m/sec} \).
behavior of the curves, the same functional trends of \( \tilde{c}_w \) and \( \tilde{c}_g \) with respect to \( g \) and \( w \) are seen here as were observed in the preceding cases.

Performance for the airgun source is improved by about 10 dB over the explosive source, and about 20 dB over the echo sounder. We emphasize that this performance ordering is due solely to the spectral content of the source wave- shapes, since we have taken care to use a unit SNR in all our calculations. We note parenthetically that if our objective were to estimate \( g \) alone, assuming that \( w \) was known, none of the sources has much to recommend it over the others in terms of estimation performance: \( \langle \tilde{J} \rangle_{22} \) was within 15% of 1.0 for all the sources and all the cases considered.

In the preceding paragraphs, we have presented and discussed the Cramer-Rao bounds for estimation of the parameters of an isogradient layer, using three typical source waveforms. We now will use these Cramer-Rao bound results to perform a preliminary design study for the mechanical or "wet" components of a system to measure layer parameters experimentally. Our objectives are to select a source from one of the three we have considered, and define the system geometry.

For our analysis, we will use the "sonar equations" formulation, in which the sound pressure levels, signal-to-noise ratio, spreading loss, etc. are expressed in decibels.
This methodology is well known, and is described in detail in [53]; we will not elaborate on it here, but will simply state the equation to be used, and define its components. This is

\[ CRB = SL - NL - 2TL + AG \]  

(4.36)

where

\[ CRB = \text{signal-to-noise ratio, in dB, required to achieve the desired estimation accuracy, as determined from the computed Cramer-Rao bound value for the source in question.} \]

\[ SL = \text{source sound pressure level, in dB re } 1 \mu b, \text{ re } 1 \text{ m, measured on the main acoustic axis of the source.} \]

\[ NL = \text{background noise level, in dB re } 1 \mu b. \]

\[ TL = \text{transmission loss due to spherical spreading and attenuation, for one-way travel between source and target, in dB re } 1 \text{ m.} \]

\[ AG = \text{array gain, in dB, indicating the processing gain achieved because of the directional character of the receiver when processing against a particular noise field.} \]

In (4.36), TL is multiplied by a factor of two to
Figure 4.14 -- Ambient noise spectrum, shallow water average traffic noise, wind force 3 (Beaufort).
From Figure 14 of [60].
indicate the fact that loss occurs as the source pulse propagates both to and from the target (the bottom). Since we have assumed that multiples are not significant, it is appropriate to correct for the transmission loss incurred by the direct reflection alone. Significant multiples would complicate computation of the effects of total transmission path loss considerably, and render (4.36) inaccurate, if not misleading.

We will use (4.36) to compute the required source sound pressure level, so that a more useful form is

$$SL = CRB + NL + 2TL - AG$$

(4.37)

Values for NL, the noise level, will be obtained from Figure 4.14 (based on Figure 14 of [60]); this spectral shape is valid in shallow water, with average levels of shipping traffic present, and with a force 3 wind (7-10 kts). Since our bound results are based on a white noise model, clearly not consistent with the ambient noise spectrum of Figure 4.14, we will select NL to overbound the colored spectrum over the bandwidth of each source. This approach, while not completely accurate, will give meaningful and useful results.

We will initially assume that the source and receiver are near the water surface, either towed, or hull-mounted.
(or dropped!). If the water depth is 150 m (corresponding to about 0.2 msec of two-way travel time), then the transmission loss is essentially due to spreading alone.

Attenuation per unit distance increases with frequency, so that, of the three sources we are evaluating, the 3.5 kHz echo sounder will suffer the most from attenuation. At 3.5 kHz, however, attenuation is less than 1 dB per km [54], which, over a 300-meter two-way travel distance is less than 1 dB, and hence negligible in comparison to the other components of (4.37) (as we shall see). The transmission loss is then given by the spherical spreading effect only. For a 150 m path length, TL is

\[
TL = 20 \log_{10}(150) \approx 44 \text{ dB re } 1 \text{ m}
\]

Before we can select CRB values for each source, we must define the desired estimate accuracy, which will be the same for all sources. We would probably like to obtain standard deviations of about 0.1 for \(\hat{w}\), and 0.01 for \(\hat{g}\). These accuracies are consistent with the spread of parameter values associated with typical sediment structures; they represent a 20% error at the midpoint of the parameter ranges. For the sake of concreteness, we will evaluate the Cramer-Rao bound at these midpoint values: \(w = 1.0 \text{ sec}^{-1}\).
and $g = 0.05$ ($c_\infty = 1700$ m/sec).

We consider first an airgun source. On examination of Figure 4.12, we find

$$\sigma_g^2 \approx 35$$

$$\sigma_w^2 \approx 1.4 \times 10^4$$

The SNR required to achieve the desired overall accuracy is obtained by taking the largest of the SNR's required for each parameter, so that $\text{SNR} \approx 1.4 \times 10^6$ (SNR is computed using (4.32)). Then $\text{CRB} \approx 61$ dB. We assume that the receiver is omnidirectional, simply a hydrophone with array gain $AG = 0$. NL is chosen to be 6 dB re 1 $\mu$B. The required SL is, using (4.37),

$$\text{SL} \approx 61 + 6 + 88 - 0 = 155 \text{ dB re 1 } \mu\text{B, re 1 m}$$

This is an unrealistically large peak sound pressure level. For comparison, a moderately large airgun, with a 300 in$^3$ chamber, operating at a pressure of 1600 psi, produces a peak sound pressure level of about 120 dB re 1 $\mu$B re 1 m [55].

*Because of the large SNR we must require for all of the sources, we are virtually assured that our (ML) estimator will be operating in its linear region, that is, the estimates should be (almost) efficient and unbiased.
We use the same analysis for an explosive source, again assuming that the receiver is a single hydrophone, with AG = 0. We find, using the most optimistic bound values, those for 500 Hz, and the same noise level as for the pre-airgun examples, NL = 6 dB re 1 μb, and the required source level is about 165 dB re 1 μb, re 1 m. We note that a 1000-lb TNT charge (rather more than one might care to use) has a peak sound pressure level of only 145 dB [53]. The explosive source, therefore, appears to be less than useful in this situation.

We will also work the problem through for the 3.5 kHz echo sounder, with a 1 msec pulse width, although, as we might guess, our efforts will indicate that it too is impractical. If we allow an array gain of 18 dB (corresponding to a 12 degree beamwidth, measured from the beam center to the 3 dB-down point), and if we recognize the relatively narrow-band, high-frequency nature of the source by using a -20 dB noise level, the required source level is 137 dB re 1 μb, re 1 m. This is far above what we could expect to achieve in practice, unless we had a large military sonar system available.

The nub of the problem we have encountered is the effect of spherical spreading, and the associated 88 dB of transmission loss. We conclude, therefore, that the most appropriate course of action is to reduce the range to the
target, by mounting both transmitter and receiver in a body to be towed close to the seafloor. Such an arrangement has been used to observe sediment properties in a less quantitative sense than we are employing here [56]. If we tow 10 m above the mottom, transmission loss is reduced to 20 dB for the body-to-bottom leg of the sound travel path. Therefore, the required sound pressure levels for each source are reduced by about 48 dB, which brings us into the realm of reasonable sound pressure levels. We note that, since multiples are not significant in this problem, and since the bound expression (4.26) is independent of delay, that the Cramer-Rao bounds computed for source and receiver at the surface are still valid when source and receiver are near the bottom.

This geometry, however, is not compatible with the use of an explosive source, and the use of an airgun, while certainly possible, is probably rather complicated mechanically. All things being equal, the simplest solution would be to use a 3.5 kHz echo sounder mounted on the towed body. The required sound pressure level for such a system would be 89 dB. The radiated peak source power required, $P$, is related to the peak sound pressure level, $SL$, by

$$10 \log_{10} P = SL - DI - 71$$

(4.38)
where

$$DI = \text{directivity index, in dB, a measure of the narrowness of the beampattern of the transmitting array.}$$

For this source, the same array is to be used for both transmitting and receiving, and, since we are processing against isotropic noise, $DI = AG$. This gives $P \approx 1 \text{ w}$, which is quite reasonable. We note that, for a 1 msec pulse length, the largest SNR required to achieve the desired estimation accuracy over the full range of parameters is only about 6 dB greater than required for our nominal values, corresponding to $P \approx 4 \text{ w}$. The array gain or directivity index we have assumed corresponds to the field of a baffled circular piston 42 inches in diameter. This is rather large, but we can easily trade off a smaller transducer size against more power (the lateral homogeneity of the bottom to be surveyed will affect this decision as well).

This completes our design study of a system for the measurement of the parameters of an isogradient layer, in which we used the Cramer-Rao bound results to define the signal-to-noise ratio required to achieve an estimate of given accuracy. We emphasize that the success of such an experiment depends not only upon the proper choice of "wet"
and "dry" equipment (including the estimation algorithm), but also upon the implicit assumption which we have made throughout, that the model we have used for the layer to be identified corresponds to the true structure. If the layer structure does not correspond to the model we have employed, we have incurred an inherent global error in our analysis, rendering it, if not useless, then of questionable integrity.
CHAPTER 5

BOUNDS FOR RANDOM STRUCTURE ESTIMATION IN CONTINUOUS MEDIA

5.1 Introduction

In the previous two Chapters, we have considered situations in which the medium to be identified, whether discrete (Chapter 3), or continuous (Chapter 4), had a very explicit structure, one which allowed us to specify the medium's reflection response as a function of a finite number of structural parameters. Such models allow a comprehensive and detailed analysis of the estimation problem, including performance bounds, and we were indeed able to treat three simple but important examples in considerable depth. On the other hand, the form of the models, which simplified the analysis so much, simultaneously placed a severe restriction on the type of problems which we could address. We found, for the discrete layer model, which can in principle approximate the response of any medium with a continuous acoustic impedance to any desired accuracy, that we were limited in the general complexity of the model we could represent by the dimensionality of the required parameter set. The continuous formulation of Chapter 4 is even more limited, since it was appropriate only for those structures for which the medium reflection response could be computed in closed form, a small set indeed.
In this Chapter, in contrast, we consider a much more
general structural model which will allow us to treat a class
of problems substantially larger than we could address using
parametric models. Given this model, our objective is to
derive a pointwise Cramer-Rao lower bound on the variance of
\( \hat{g}(\xi) \), an arbitrary, unbiased estimate of the reflection co-
efficient density of the medium. We will model \( g(\xi) \) as a
covariance-stationary Gaussian random process, with a
spatially-varying mean, \( \bar{g}(\xi) \). Note that this is purely a
spatially-indexed random process, not a time-indexed one.
That is, were we to probe the medium several times, its
deterministic response, \( s(t) \), would be unchanged from obser-
vation to observation. The received signal would be
different each time, however, because we receive a different
sample function of \( n(t) \) at each observation.

Primarily, such a model is useful because it allows
us to obtain general results valid for a wide variety of
structures. This is possible because, by our choice of
mean, \( \bar{g}(\xi) \) and covariance function, \( K(\tau) \), we have con-
strained \( g(\xi) \) to lie in an ensemble of possible functions
whose elements may be quite different in their individual
appearance, and quite complicated (possessing a large number
of degrees of freedom). For instance, we constrain the
bandwidth of \( g(\xi) \) (or, more properly, the bandwidth of
\( \tilde{g}(\xi) = g(\xi) - \bar{g}(\xi) \)) in a second-order sense, by specifying
the correlation distance of its covariance function (or, more directly, the bandwidth of the corresponding power spectral density function). The constraint, however, is not hard-and-fast: a particular sample function, or realization, of \( g(\xi) \) might have a bandwidth well in excess of that implied by the bandwidth of the correlation function.* Specification of the covariance function bandwidth, then, does not exclude the occurrence of sample functions of (arbitrarily) large bandwidth, but ensures that such sample functions occur with small probability (perhaps with probability zero). In short, no reasonable function is strictly excluded from consideration. In this sense, a random process model for \( g(\xi) \) is robust, since our results are decoupled from the detailed (sample function) behavior of \( g(\xi) \).

A random process model for \( g(\xi) \) also is sufficient to ensure that both the estimation problem and the bound computation are well-posed and stable. For example, one might consider modelling \( g(\xi) \) as a continuous deterministic function to be estimated in, say, the ML sense, using some variational technique. Such an approach has been developed quite completely in [48], under very general assumptions. The difficulty with this analysis, which is not brought out

*We assume here, of course, that the process spectral density function is not strictly bandlimited.
in [48], is that the estimator performance, and indeed its stability, is very sensitive to the choice of observation noise model. In particular, the estimator of [48] is not well-posed if the observation noise is white. In a qualitative sense, these problems arise because no constraint has been placed on the number of degrees of freedom of \( g(\xi) \): the space of functions allowed for \( g(\xi) \) is too large, including some which are quite unreasonable. A maximum bandwidth specification for \( g(\xi) \) is an example of the type of constraint which would alleviate this problem. As discussed above, however, our random process model may be interpreted as just such a bandwidth constraint, with the additional advantage that the constraint is not a rigid one so that a large class of functions may be considered.

The random process model requires us to incorporate prior information about structure in the estimation problem; this is quite useful, since, despite how complicated the structure may be (or, in another sense, how little we know about it à priori), as a practical matter we always have some prior knowledge about the structure. In Chapters 3 and 4, for instance, we assumed very explicit prior knowledge of medium structure, in which structure was known up to a finite set of parameters. The prior information which we must provide for the random structure model is considerably weaker: we need specify only a mean, \( \bar{g}(\xi) \), and covariance,
**K_g(τ)**. The process mean may be interpreted as our à priori best estimate of \( g(\xi) \); the magnitude of the covariance at zero lag, as the variance of our à priori information; and the correlation distance of the covariance, as a measure either of coherency in the structure, or as a bandwidth constraint. Observe that if we consider \( \tilde{g}(\xi) \) to be the nominal structure, and \( \tilde{g}'(\xi) \) to be a perturbation about it, we may deal successfully with the problem of modelling error discussed briefly in Section 4.3. In short, we may allow, by an appropriate choice of \( K_g(0) \), a sufficiently large variation in \( \tilde{g}(\xi) \) to absorb any errors in \( \tilde{g}(\xi) \). This is another sense in which this model may be considered robust.

It is interesting that our model has not been used in the literature for the structure estimation problem. Wave propagation in random media has received considerable attention in the literature (for instance, [45] - [47]), but most effort has been placed on solution of the wave equation (the direct problem), rather than structure estimation from measurements of scattered energy (the inverse problem). In the seismic exploration literature, however, the earth's reflection response is sometimes modelled as a doubly stochastic white Poisson process (for example, [58]), but this model is not as strongly coupled to the physics of the medium as is ours.

In Section 5.2, we develop the Cramer-Rao bound in a
form appropriate for our problem; we have emphasized the
generality of the derivation, since the result is applicable
to a wide class of problems, and is a significant extension
of established results. We consider a simplified model for
medium response in Section 5.3, a model which allows rela-
tively easy solution of the integral equation which defines
the bound derived in Section 5.2. This model assumes that
the medium is weakly scattering, so that \( s(t) = d(t) * g(t) \)
(neither internal scattering nor free surface multiples
are significant in the response). In Section 5.4, we
evaluate the bound for the case of a weakly scattering
medium, using simple models for source and process covar-
iance. In particular, we are able to define an optimum
source bandwidth, one which minimizes the Cramer-Rao bound
for this example. This result is in contrast to conventional
wisdom, which holds that the larger the source bandwidth,
the better estimation accuracy we may obtain.

A considerable amount of important, but tedious,
mathematical detail has been relegated to Appendices B and C.
In Appendix B, we prove a result required in the derivation
of the general Cramer-Rao bound form in Section 5.2. In
Appendix C, we demonstrate that the medium model of Chapter
2, together with appropriate restrictions on \( g(\xi) \), ensures
that the general bound result of Section 5.2 is indeed
applicable to our problem; this entails proving that the
reflection response of the medium obeys certain regularity conditions assumed in the derivation of the bound.

5.2 The Cramer-Rao Bound

In this Section, we will derive a form of the Cramer-Rao bound appropriate for estimation of a continuous medium structure, \( g(\xi) \), which is characterized as a Gaussian random process. In this derivation, we will draw upon some of the basic relationships which we quoted in Chapter 3, in particular (3.9). Our approach follows that of [49], and Section 5.3 of [31]: we first write the Cramer-Rao bound expression for a finite-dimensional (random) parameter approximation to \( g(\xi) \) (a truncated form of the Karhunen-Loeve expansion); we then manipulate this result into an appropriate form, and take the limit as the order of the approximation goes to infinity. The result is a definition of the bound as the solution of a Fredholm integral equation.

Our development is a significant extension of existing results, in that it allows us to treat a much broader class of problems. To demonstrate this, recall that our observation is of the form

\[
\mathbf{r}(t) = \mathbf{s}(t|g) + \mathbf{n}(t), \quad 0 \leq t \leq T \tag{5.1}
\]

where we have written \( s(t|g) \) to emphasize that the reflection
response is a mapping of the reflection coefficient density function into the time domain. More technically, we may write

\[ s(\cdot | \cdot) : \mathcal{X}[0, \infty) \rightarrow \mathcal{F}[0, T] \]  

(5.2)

where \( \mathcal{X}[0, \infty) \) is the space sample functions of \( g(\xi) \) (we have changed the upper limit of the domain to obtain a more general discussion), and \( \mathcal{F}[0, T] \) is the space of noiseless reflection responses. That is, \( s(t|g) \) is a functional of the function \( g \).

A lower bound on the estimate variance, \( \text{Var}[\hat{g}(\xi)] \), has been developed in [49] for two restrictive classes of functional, which are, nonetheless, quite important in communications: (1) \( s(t|g) \) is a memoryless, nonlinear function of \( g(\cdot) \), such as \( s(t|g) = \sin(2\pi f_0 t + w g(t)) \) (phase modulation); and, (2), \( s(t|g) \) is a memoryless nonlinear function of the output of a linear system driven by \( g(t) \), such as \( s(t|g) = \sin(2\pi f_0 t + \int g(\tau) d\tau) \) (frequency modulation). We require, however, a form which is applicable to nonlinear functionals in which the nonlinearity has memory: for example, \( s(t|g) = \int_{0}^{t} \sin(g(\tau)) d\tau \).

We can motivate the need for considering such an extended class of functionals by recalling that, regardless of the reflection response of the medium without the free
surface, the presence of the free surface required a feedback model, as indicated in (4.21):

\[
S(f|\alpha) = \frac{D(f)G(f|\alpha)}{1+G(f|\alpha)}.
\]  

(4.21)

The time-domain equivalent of this, independent of the nature of \(G(f|\alpha)\), is a nonlinear functional of the class we have identified above.

Before beginning the derivation, we note that [49] develops a bound on the integrated mean-square error:

\[
\mathbb{E} \left[ \int_{0}^{\xi} \left( (\hat{g}(\xi) - g(\xi))^2 \right) d\xi \right] \geq \int_{0}^{\xi} J^{-1}(\xi, \xi) d\xi.
\]

(5.3)

This was subsequently extended in [50] to a pointwise bound on estimate variance,

\[
\mathbb{E}[\left( (\hat{g}(\xi) - g(\xi))^2 \right) \geq J^{-1}(\xi, \xi).
\]

(5.4)

This is the form of the bound which we shall derive.

Let us begin by restating the properties of \(g(\xi)\):

\(g(\xi)\) is a covariance-stationary Gaussian random process, with mean \(\bar{g}(\xi)\), and covariance \(K_g(\tau)*\):

*This derivation could also be carried through in precisely the same manner without the assumption of covariance stationarity; the extension is straightforward.
\[ E[g(\xi)] = \bar{g}(\xi), \]  
\[ (5.5) \]

\[ E[(g(\xi+\tau)-\bar{g}(\xi+\tau))(g(\xi)-\bar{g}(\xi))] = K_g(\tau). \]  
\[ (5.6) \]

We will assume that \( g(\xi) \) has been specified in such a way (by our choice of \( \bar{g}(\xi) \) and \( K_g(\tau) \)) that \( s(t|g) \) is well-defined. This issue must be addressed separately for each application, since it depends upon the exact form of \( s(\cdot|\cdot) \); it is discussed in depth in Appendix C for the current application, when \( g(\xi) \) is the reflection coefficient density of a medium, and \( s(t|g) \) is its reflection response.

We want to form a Cramer-Rao bound for the estimation problem defined by the observation equation (5.1). We will approach the problem by first considering the finite-dimensional estimation problem based on a \( K \)-term approximation of \( g(\xi) \). This approximation is based on the Karhunen-Loeve expansion (see, for instance, [31], or [52]),

\[ g(\xi) = \sum_{k=1}^{\infty} g_k \phi_k(\xi). \]  
\[ (5.7) \]

Here \( \{\phi_k(\xi)\} \) is a complete orthonormal set, the eigenvectors of the integral equation

\[ v_i \phi_i(\xi) = \int K_g(\xi-\tau) \phi_i(\tau) d\tau, \]  
\[ (5.8) \]
and the $g_k$ are statistically independent Gaussian random variables, with variance $\nu_k$ ($\nu_k$ is the eigenvalue associated with $\phi_k(\xi)$). The sum in (5.7) converges in mean square. The prime advantage of this representation is the independence of the $\{g_k\}$. We assume that $\bar{g}(\xi)$ has finite energy (we will, in fact, specify boundedness in Appendix C), so that we may write the mean value of each $g_k$ as

$$\bar{g}_k = \int_0^\infty \bar{g}(\xi) \phi_k(\xi) \, d\xi$$  \hspace{1cm} (5.9)

We will define $g_K(\xi)$ to be the Kth partial sum of the Karhunen-Loeve expansion, (5.7):

$$g_K(\xi) = \sum_{k=1}^K g_k \phi_k(\xi).$$  \hspace{1cm} (5.10)

We may now consider an approximate version of the observation equation, (5.1),

$$r(t) = s(t|g_K) + n(t), \quad 0 \leq t \leq T.$$  \hspace{1cm} (5.11)

This approximate observation implies a finite-parameter estimation problem, in which our objective is to estimate the $K$ random parameters $g_1, g_2, \ldots, g_K$. Not only can we compute the Fisher information matrix for this problem, $J_K$, but we may also obtain a pointwise bound on $g_K(\xi)$, using the result (4.11):
\[ E[(g_k^*(\xi) - g_k(\xi))^2] \geq \phi_k(\xi) J^{-1}_K \phi_k(\xi), \]  

(5.12)

where \( \phi_k(\xi) = [\phi_1(\xi), \phi_2(\xi), \ldots, \phi_K(\xi)]^T. \)

Since we are estimating random variables, rather than unknown, nonrandom parameters, the appropriate form for the Fisher information matrix is as noted in (3.7)-(3.9), which is, in our current notation,

\[ J_K = J_0 + J_p \]

(5.13)

where

\[ (J_0)_{ij} = E\left[ \frac{\partial \ln \Lambda(r(t)|g_K)}{\partial g_i} \cdot \frac{\partial \ln \Lambda(r(t)|g_K)}{\partial g_j} \right] \]

(5.14)

\[ (J_p)_{ij} = E\left[ \frac{\partial \ln p(g_K)}{\partial g_i} \cdot \frac{\partial \ln p(g_K)}{\partial g_j} \right] \]

(5.15)

\[ \ln \Lambda(r(t)|g_K) = \frac{1}{\sigma^2} \int_0^T (r(t) - \frac{1}{2}s(t|g_K)s(t|g_K))dt \]

(5.16)

\[ \ln p(g_K) = \sum_{k=1}^K \left[ \frac{(g_k^* - g_k)^2}{2\nu_k} - \frac{1}{2} \ln(2\pi\nu_k) \right] \]

(5.17)

and
$$g_K = (g_1, g_2, \ldots, g_K).$$  \hspace{1cm} (5.18)

Evaluating (5.14) and (5.15), we find that

$$\left( J^*_K \right)_{ij} = \frac{\delta_{ij} \nu_i}{v_i} + \frac{1}{\sigma^2} \mathbb{E} \left[ \int_0^T \frac{\partial s(t|g_K)}{\partial g_i} \frac{\partial s(t|g_K)}{\partial g_j} dt \right].$$  \hspace{1cm} (5.19)

where we have assumed that \( \frac{\partial s(t|g_K)}{\partial g_i} \) exists.

This is an appropriate point at which to require that the functional \( s(\cdot | \cdot) \) satisfy certain regularity conditions, chosen to ensure that the limiting operation \( K \to \infty \) yields a well-defined result. Specifically, we require that:

1. \( s(t|g) \) is a continuous functional;
2. \( \frac{\partial s(t|g)}{\partial g_i} \) exists, and is a continuous functional.

Note that (2) implies (1), we state them both for completeness. The appropriate metrics for continuity are

$$||s(t|g_1), s(t|g_2)|| = |s(t|g_1) - s(t|g_2)|$$  \hspace{1cm} (5.20)

$$||g_1(\xi), g_2(\xi)|| = \int_0^\infty |g_1(\xi) - g_2(\xi)| d\xi$$  \hspace{1cm} (5.21)

That is, we require \( s(t|g) \) and \( \frac{\partial s(t|g)}{\partial g_i} \) to be bounded in \([0,T]\), and restrict the admissible class of \( g(\xi) \) to be those which are absolutely integrable on \([0,\infty]\).

These are minimal constraints; in particular, depending upon the nature of the functional
s(t|g) (depending upon the application) it may be necessary to place more stringent requirements on the class of admissible g(\xi). We will indeed do this in Appendix C, as a prelude to demonstrating the above continuity conditions for our problem.

We now introduce the covariance function for \( g_K(t) \),

\[
K_{g_K}(t,z) = \sum_{k=1}^{K} n_k \phi_k(t) \phi_k(z).
\]

(5.22)

Although we have specified that \( g(t) \) is covariance-stationary, \( g_K(t) \) need not be. We note that as \( K \to \infty \), \( K_{g_K}(t,z) \) converges uniformly to \( K_{g}(t-z) \) on \([0,\Xi] \times [0,\Xi] \). We may also define the inverse kernel, \( V_K(t,z) \), as

\[
V_K(t,z) = \sum_{k=1}^{K} \frac{\phi_k(t) \phi_k(z)}{n_k}.
\]

(5.23)

Because \( g(t) \) is not modelled as including a white noise component, the sum in (5.23) will not exist in general as \( K \to \infty \). Thus we must eliminate \( V_K(t,z) \) from our solution before taking the limit. We note that

\[
\int_{0}^{\Xi} V_K(t,z) K_{g_K}(z,\tau) dz = \sum_{k=1}^{K} \phi_k(t) \phi_k(\tau),
\]

(5.24)

We may write the first term on the right hand side of
(5.19) in terms of the inverse kernel, recalling that the basis functions \( \{ \phi_k(\xi) \} \) are mutually orthogonal:

\[
\frac{\epsilon_{ij}}{\nu_i} = \int \int \phi_i(\xi) \nu_k(\xi,z) \phi_j(z) d\xi dz. \tag{5.25}
\]

We now define \( \mathbf{J}^{-1}_K(\xi,z) \),

\[
\mathbf{J}^{-1}_K(\xi,z) = \mathbf{T} \phi_K(\xi) \mathbf{J}^{-1}_K \phi_K(z). \tag{5.26}
\]

We note, from our result in (5.12), that \( \mathbf{J}^{-1}_K(\xi,\xi) \) is a lower bound on \( \text{Var} [\hat{g}_K(\xi)] \). Similarly, we define \( \mathbf{J}_K(\xi,z) \) as

\[
\mathbf{J}_K(\xi,z) = \mathbf{T} \phi_K(\xi) \mathbf{J}_K \phi_K(z), \tag{5.27}
\]

and note that*

\[
\int \mathbf{J}^{-1}_K(\xi,\eta) \mathbf{J}_K(\eta,z) d\eta = \sum_{k=1}^{K} \phi_k(\xi) \phi_k(z). \tag{5.28}
\]

As we did in (5.25) with the inverse kernel, \( \nu_K(\xi,z) \), we may write the elements of the Fisher information matrix, the left-hand side of (5.19), as

\[
(\mathbf{J}_K)_{ij} = \int \int \nu_i(\xi) \mathbf{J}_K(\xi,z) \phi_j(z) d\xi dz. \tag{5.29}
\]

*The reader should note carefully that \( \mathbf{J}^{-1}_K(\xi,z) \) is not \( 1/\mathbf{J}_K(\xi,z) \), but \( \mathbf{J}^{-1}_K \) is the matrix inverse of \( \mathbf{J}_K \).
Using (5.29) and (5.25) in (5.19), we obtain

\[ J = \int_0^\infty \int_0^\infty \phi_i(\xi) \mathcal{J}_K(\xi, z) \phi_j(z) d\xi dz \]

\[ = \int_0^\infty \int_0^\infty \phi_i(\xi) V_K(\xi, z) \phi_j(z) d\xi dz \]

\[ + \frac{1}{a^2} \mathbb{E} \left[ \int_0^T \frac{\partial s(t | g_k)}{\partial g_i} \frac{\partial s(t | g_j)}{\partial g_j} dt \right]. \quad (5.30) \]

We would now like to write the partial derivatives in (5.30) in terms of an integral operation on the eigenfunctions, \( \{ \phi_k(\xi) \} \), so that both sides of (5.30) will be in similar form, and we may identify an expression for \( J_k(\xi, z) \).

We now introduce the concept of the variational derivative, [51]; this will allow us to obtain the desired form in (5.30).

**Definition:** Let \( s(\cdot | \cdot) \) be a functional with the regularity properties specified above. Let \( h(\xi) \), an element of the domain of \( s(\cdot | \cdot) \) take on an increment, \( \delta h(\xi) \), which has compact support in a neighborhood of \( \xi_0 \). Consider the difference

\[ s(t | h + \delta h) - s(t | h). \]

If \( \Delta \) is the area under \( \delta h(\xi) \),

\[ \Delta = \int_0^\infty \delta h(\xi) d\xi \quad (5.31) \]
and if we can write

\[ s(t|h+\delta h) - s(t|h) = \left( \frac{\delta s(t, \xi_0)}{\delta h} + \varepsilon \right) \Delta \]  

(5.32)

where \( \varepsilon \to 0 \) as both \( \max |\delta h(\xi)| \) and the length of the interval in which \( \delta h(\xi) \) is nonzero go to zero, then \( \frac{\delta s(t, \xi_0)}{\delta h} \) is referred to as the **variational derivative** of \( s(t|h) \). We require, of course, that (5.32) hold for all functions \( h(\xi) \) in the domain of \( s(t|h) \), for all \( \xi_0 \in [0, \mathbb{E}] \), and for all \( t \in [0, T] \).

The variational derivative quantifies the changes we observe in the reflection response of the medium due to a small, localized perturbation in \( g(\xi) \). We may write the partial derivatives in (5.30) using the relationship (proven in Appendix B)

\[ \frac{\partial s(t|g_K)}{\partial g_i} = \int_0^\mathbb{E} \frac{\delta s(t, \xi)}{\delta g_K} \phi_i(\xi) d\xi. \]  

(5.33)

We will require strong-sense existence of the variational derivative, although it appears that we are only making use of it in an operator sense in (5.33). That is, we will require that the variational derivative be a continuous functional, in the same manner as we required \( s(t|g) \) and its partial derivative to be continuous. This continuity property is demonstrated, for the problem at hand, in Appendix C.
Using (5.33), we may write (5.30) as

\[
\begin{align*}
&\mathbb{E} \int_0^1 \int_0^1 \phi_i(\xi) J_K(\xi, z) \phi_j(z) d\xi dz \\
&= \int_0^1 \int_0^1 \phi_i(\xi) V_K(\xi, z) \phi_j(z) d\xi dz \\
&\quad + \frac{1}{\sigma^2} \int_0^1 \int_0^1 \phi_i(\xi) E\left[ \int_0^T \frac{\delta s(t, \xi)}{\delta g_K} \frac{\delta s(t, z)}{\delta g_K} dt \right] \phi_j(z) d\xi dz.
\end{align*}
\]

(5.34)

In (5.34), we have used the continuity of the variational derivative to exchange expectation and integration. We may now identify $J_K(\xi, z)$ as

\[
J_K(\xi, z) = V_K(\xi, z) + \frac{1}{\sigma^2} E\left[ \int_0^T \frac{\delta s(t, \xi)}{\delta g_K} \frac{\delta s(t, z)}{\delta g_K} dt \right].
\]

(5.35)

We may substitute (5.35) in (5.28) to obtain a relationship involving $J_K^{-1}(\xi, z)$, which will provide a lower bound on $\text{Var}[\hat{g}_K(\xi)]$:

\[
\begin{align*}
&\mathbb{E} \int_0^1 J_K^{-1}(\xi, \eta) [V_K(\eta, z) + \frac{1}{\sigma^2} E\left[ \int_0^T \frac{\delta s(t, \eta)}{\delta g_K} \frac{\delta s(t, z)}{\delta g_K} dt \right]] d\eta \\
&= \sum_{k=1}^K \phi_k(\xi) \phi_k(z)
\end{align*}
\]

(5.36)

We would like to take the limit of (5.35) as $K \to \infty$, but
$V_K(\eta, z)$ does not exist in the limit; neither does the right-hand side of (5.36), as a function. The right-hand side does converge, in a distributional sense, however, to an impulse, so that for any continuous $\psi(\xi)$,

$$
\lim_{K \to \infty} \int \psi(\xi) \cdot \sum_{k=1}^{\infty} \phi_k(\xi) \phi_k(z) d\xi = \psi(z) \tag{5.37}
$$

We may alleviate both these problems by multiplying (5.36) through by $K(z, \tau)$ and integrating over $z$. Using (5.24) we have

$$
\left( \int J^{-1}(\xi, \eta) [ \sum_{k=1}^{\infty} \phi_k(\eta) \phi_k(\tau) \right)
\left( \begin{array}{c}
\sum_{k=1}^{\infty} \phi_k(\xi) \phi_k(z) K(\tau, \xi) \\
+ \frac{1}{\sigma^2} \int \int E[ \int \frac{\delta s(t, \eta)}{\delta g_K} \frac{\delta s(t, z)}{\delta g_K} dt] g_K(z, \tau) dz d\eta
\end{array} \right)

= \int \sum_{k=1}^{\infty} \phi_k(\xi) \phi_k(z) K(\tau, \xi) dz \tag{5.38}
$$

We now allow $K \to \infty$; we have required sufficiently strong regularity conditions that the limit exists. Using (5.37), we obtain
\[ J^{-1}(\xi, \tau) + \frac{1}{\sigma^2} \int_0^T \int_0^T J^{-1}(\xi, \eta) \mathbb{E} \left[ \int_0^T \frac{\delta s(t, \eta)}{\delta g} \frac{\delta s(t, \tau)}{\delta g} dt \right] \times K_g(z-\tau)dzd\eta \]

\[ = K_g(\xi-\tau) \quad (5.39) \]

where \( J^{-1}(\xi, z) \) is the function we seek,

\[ \text{Var}[g(\xi)] \geq J^{-1}(\xi, \xi) \quad (5.40) \]

The form of (5.39) is identical to that of the result derived in [49] except for the use of a kernel formed from the variational derivative in place of a partial derivative.

To summarize our result, drawing together all the assumptions we made through the derivation above, suppose we have an observation

\[ r(t) = s(t|g) + n(t), \quad 0 \leq t \leq T \quad (5.1) \]

where \( g(\xi) \) is a covariance-stationary Gaussian random process on \([0,\xi]\), with properties given in (5.5) and (5.6), and \( n(t) \) is a white, zero-mean stationary Gaussian random process with power spectral density \( \sigma^2 \). Let \( s(t|g) \) be a continuous functional, in the sense of the metrics (5.20) and (5.21), mapping sample functions of \( g(\xi) \) into time functions, which
has a continuous variational derivative, and for which
\[ \frac{\delta s(t | g)}{\delta g_i} \]
exists and is a continuous functional. We again emphasize that these are the properties required for our derivation; depending upon the physical problem which underlies \( s(t | g) \), more stringent requirements may be necessary, particularly on the metric for \( g(\xi) \). Then, for an unbiased estimate of \( g(\xi), \hat{g}(\xi) \), (5.40) holds, with \( J^{-1}(\xi, z) \) given by (5.39).

Although we have considered in this derivation that \( \xi \) and \( t \) lie in different domains, this need not be true in general. If \( \xi \) took values in \([0, T]\), the appropriate form for (5.39) would be

\[ J^{-1}(\xi, \tau) = K_g(\xi - \tau) \tag{5.41} \]

We close this Section on a critical note. Although we have indeed succeeded in defining the Cramer-Rao bound in a form appropriate for our problem, we have obtained it in implicit form, as the solution to the integral equation (5.39), or (5.41). In order to solve this equation, we must evaluate the kernel,

\[ + \frac{1}{\sigma} \int_0^T J^{-1}(\xi, \eta) \int_0^T \frac{\delta s(t, \eta)}{\delta g} \frac{\delta s(t, z)}{\delta g} dt K_g(z - \tau) dz d\eta \]
\[
\mathbb{E} \left[ \int_0^T \frac{\partial s(t, n)}{\partial g} \frac{\partial s(t, z)}{\partial g} dt \right] K_g(z-\tau) dz.
\]

Neither solution of the equation nor evaluation of the kernel are trivial tasks, and, in all but the simplest cases, cannot be accomplished analytically. We proceed, therefore, with the knowledge that, although we have an analytic definition of the result we seek, it is unlikely that we will be able to obtain a simple, closed-form, analytic solution for it. To obtain usable results, we must be prepared to make judicious simplifications in our model.

5.3 A Model for the Reflection Response of Weakly Scattering Media

In this Section, we will develop a model for the reflection response of a particular class of media, termed "weakly scattering" media, whose reflection coefficient density is a random process. This model will enable us to compute the kernel of the integral equation (5.41) in a relatively straightforward fashion, and hence to solve (5.41) for the pointwise Cramer-Rao bound, \( J^{-1}(\xi, \xi) \). As we have discussed in Section 5.2, evaluation of the kernel is a major stumbling block to the solution of (5.41); we will address, in particular, the problem of evaluating the expectation term in the kernel,
This portion of the kernel embodies the nonlinear dependence of $s(t|g)$ on $g(\xi)$, and, as well, requires the computation of an expected value: its potential complexity is evident.

While the medium model we have chosen is simple enough to allow evaluation of (5.42), it is not so simple as to have no practical import. We will demonstrate, in particular, that the isogradient layer medium used in Section 4.3 to model the response of an ocean bottom sediment layer is substantially a weakly scattering medium. In contrast to the rigorous methods of the preceding Section, we will proceed in a rather heuristic fashion; our objective is to specify a useful engineering approximation, not to obtain an exact analytical result.

In Chapter 2, we presented a representation for $s(t|g)$ in terms of the Bremmer series, (2.26), which had the property that each succeeding term modelled the effect of a distinct, progressively higher-order internal scattering process. As a result, we may truncate this series knowing exactly what physical phenomena we are discarding. We also presented criteria under which the series converged uniformly and absolutely. This series is therefore a very valuable one for obtaining approximations to $s(t|g)$, and we would like to use it in addressing the current problem. In order to use the
Bremer series when \( g(\xi) \) is a random process, however, we must generally ensure that the convergence criteria of the Bremer series are satisfied by almost every sample function of \( g(\xi) \). That is, the set of sample functions which violate the criteria must occur with probability zero. Unfortunately, this cannot be guaranteed for the class of random processes which we have selected. One of the series' convergence criteria is, for a finite interval,

\[
\int_0^T |g(\xi)| \, d\xi \leq \pi, \tag{5.43}
\]

which may be interpreted as a bandwidth constraint. We have specified that \( g(\xi) \) is to be a Gaussian random process, and assumption that is central to our development of Section 5.1. For such a process, we cannot ensure that (5.43) will be satisfied on a sample function basis. Therefore, we may not employ the Bremer series in a rigorous fashion.

We note that there are several other integral series representations for \( s(t|g) \), similar to the Bremer series. In Appendix C, we have made extensive use of one series, developed in [22]. This series is uniformly and absolutely convergent when \( g(\xi) \) is piecewise continuous on the closed interval \([0,T]\). We can ensure that the random portion of \( \tilde{g}(\xi), \tilde{g}(\xi) \), is almost everywhere sample function continuous by placing relatively mild restrictions on the process co-
variance function. We need only require, in addition, that \( \bar{g}(\xi) \) be at worst piecewise continuous. One might ask, then, why do we not employ this series representation to obtain an approximate medium model. The answer is that the successive terms in the integral series of [22] do not have the same physical interpretation as do those of the Bremmer series: if we discard terms, we have no indication of what physical phenomena we are discarding.

While we cannot use the Bremmer series rigorously, we will use it to motivate a reasonable approximate model. If the process variance is sufficiently small, and if the process mean is such that

\[
\int_{0}^{T} |\bar{g}(\xi)| \, d\xi \ll \pi,
\]

we may consider that the probability of observing a sample function which violates (5.43), while not zero, is so small that we may, in practice, employ the Bremmer series as though (5.43) were indeed true on a sample function basis. This argument, while certainly not mathematically valid, is adequate to motivate the engineering approximation which we desire.

Given that the Bremmer series holds, however tenuous that assumption may be, we may write the Fourier transform of the complete reflection response, \( S(f|g) \), including the
free surface multiples not represented by the Bremmer series, as

\[ S(f | g) = \frac{D(f)B(f,0)}{1 + B(f,0)} \tag{5.44} \]

where \( D(f) \) is the transform of the source waveform, and \( B(f,0) \) is the Bremmer series of (2.26), evaluated at the surface. Equation (5.44) is the transfer function of the feedback system shown in Figure 2.2.

Since our objective is to obtain an expression for the variational derivative of the reflection response, we formally take the variational derivative of each side of (5.44), and obtain its Fourier transform:

\[ \frac{\delta S(t, z)}{\delta g} + \frac{\delta S(f, z)}{\delta g} = \frac{D(f)}{(1 + B(f,0)^2} . \frac{\delta B(f, z)}{\delta g} \tag{5.45} \]

where \( \delta B(f, z)/\delta g \) is the variational derivative of \( B(f,0) \) due to a perturbation in \( g(\xi) \) at \( \xi = z \). The Bremmer series through the second term is

\[ B(f,0) = \int_{0}^{T} g(\tau) e^{-j2\pi ft} d\tau \tag{5.46} \]

\[- \int_{0}^{T} g(\tau) e^{-j2\pi ft} \int_{0}^{\tau} g(\eta) e^{j2\pi f\eta} \int_{0}^{\eta} g(z) e^{-j2\pi f\eta} dz d\eta d\tau \]
Without loss of generality, we may assume that the reflection coefficient density is zero outside $[0,T]$, so that the leading term in (5.46) is the Fourier transform of $g(t)$. If we formally take the variational derivative of (5.46), we obtain

\[
\frac{\delta B(f,\xi)}{\delta g} = e^{-j2\pi f\xi} - e^{-j2\pi f\xi} \int g(\tau) e^{j2\pi f\tau} \int g(\eta) e^{-j2\pi f\eta} d\eta d\tau
\]

\[
= -e^{-j2\pi f\xi} \int g(\tau) e^{-j2\pi f\tau} \int g(\eta) e^{j2\pi f\eta} d\eta d\tau
\]

\[
= -e^{-j2\pi f\xi} \int g(\tau) e^{j2\pi f\tau} d\tau \int g(\eta) e^{-j2\pi f\eta} d\eta
\]

\[
= -e^{j2\pi f\xi} \left( \int g(\eta) e^{-j2\pi f\eta} d\eta \right)^2
\]

We see that if

\[
\int_{0}^{T} |g(\xi)| d\xi << 1,
\]

which is compatible with the smallness assumptions we have made on $g(\eta)$ and $g(\xi)$ to justify using the Bremmer series, there is motivation for discarding all terms above first order.

We postulate, therefore, that
This form for the Bremmer series' variational derivative implies that

\[ B(f,0) \leftrightarrow g(t) \]  \hspace{1cm} (5.49)

which we might have obtained directly from (5.46), by using the same argument as we did to obtain (5.48) from (5.47).

This response model corresponds physically to the situation in which scattering internal to the medium is so small as to be negligible by comparison with the direct reflection from each point in the medium.*

We now use the approximate result (5.48) in (5.45) to obtain an expression for the variational derivative of the reflection response. If we express the quotient in (5.45) in terms of an infinite geometric series in \( B(f,0) \), we have

\[ \frac{\delta S(f,z)}{\delta g} = D(f)e^{-j2\pi fz}(1-2B(f,0)+3B^2(f,0)-\ldots). \]  \hspace{1cm} (5.50)

*In Appendix C, we have developed an exact expression, (C.47), for the variational derivative of the series expansion presented in [22]. With some efforts, (C.47) may be expanded in terms of regression filters (defined in Appendix C), and it may then be seen that the leading term of the expression is \( \exp(j2\pi fz) \), with higher-order terms being weighted integrals of \( g(\xi) \). We could, therefore, draw the same analytical conclusion, (5.48), from (C.47), but it would be lacking in physical motivation.
But, using the same argument as above, we can ignore higher order terms, and so obtain the simple expression

$$\frac{\delta S(f,z)}{\delta g} = D(f)e^{-j2\pi fz}. \quad (5.51)$$

By ignoring higher order terms in (5.50), we have discarded all free surface multiples. In our previous investigations, we have seen instances in which this was physically reasonable, e.g., in the single-interface example of Chapter 3, the presence of water column multiples has little effect on estimator performance when $|g|$ is small, as indicated in Figure 3.2. The time-domain equivalent of (5.51) is

$$\frac{\delta s(t,z)}{\delta g} = d(t-z). \quad (5.52)$$

We note that (5.52) is independent of $g(\xi)$, and is, in fact, deterministic, so that evaluation of the expectation in (5.42) is trivial. Also, the process mean value does not enter the kernel, so the nominal profile does not affect our ability to estimate the unknown portion of the profile, $\tilde{g}(\xi)$ (under the smallness assumptions which allowed us to obtain (5.52)).

Applying the same argument to (5.44) we find the reflection response for such a medium:

$$s(t|g) = d(t) * g(t) \quad (5.53)$$
If (5.53) and (5.52) hold for a given medium, we shall term it a "weakly scattering" medium, one in which the scattering is so weak that neither internal reverberation nor free surface multiples have a significant effect on the reflection response of the medium. Given that the medium is weakly scattering, we see that our observation below (5.52), that performance was independent of process mean, is a very reasonable one. Since (5.53) is linear in \( g(\xi) \), we may simply subtract off the (known) mean, \( \bar{g}(\xi) \), so that the estimation problem can become, very explicitly, one of estimating \( g(\xi) \).

The validity of (5.53) has been investigated numerically in [29] for a simple class of (deterministic) \( g(\xi) \), \( g(\xi) \) constant over an interval and zero elsewhere. Equation (5.53) was found to be a good approximation when the net change in acoustic impedance across the layer was as much as three times the impedance above the layer (i.e., the net "reflection coefficient" of the structure was as large as 0.5).

We may investigate this same example analytically using the isogradient layer model we developed in Chapter 4. We may easily compute \( g(\xi) \) for the profile in Figure 4.1, using (2.7) to convert depth to two-way travel time. We find that the reflection coefficient density corresponding to Figure 4.1 is
g(\xi) = \begin{cases} 
\frac{w}{4}, & \tau_w \leq \xi \leq \tau_w + \Delta \xi \\
0, & \text{elsewhere}
\end{cases} \quad (5.54)

where \Delta \xi, the layer thickness in seconds, is \frac{2\ln(c_\infty/c_0)}{w}. If we take the Fourier transform of (5.54) to obtain the leading term in the Bremmer series for this structure, we find

\[ g(\xi) \approx e^{-j2\pi f(\tau_w + \Delta \xi/2)} \frac{w\sin(\pi f \Delta \xi)}{4\pi f} \quad (5.55) \]

The magnitude of (5.55) is almost identical to that of the exact layer response, \( G(f|\alpha) \) of (4.14), as we may see by examining Figure 4.12, and noting that the zeroes of (5.55) are given exactly by (4.20). In addition, we noted in Chapter 4 that addition of free surface multiples to \( G(f|\alpha) \) produced virtually the same transfer function (see Figure 4.4). These statements are, of course, qualitative, and depend for their validity on the parameters of the isogradient layer. We have given, however, an analytical demonstration that one common, but important medium model, an isolated isogradient layer with parameters typical of an ocean bottom sediment layer, may be considered to be a weakly scattering medium.

In this Section, we have motivated and developed an heuristic model for the response of a medium in which both internal scattering and free surface multiples were negligible.
in comparison to the direct reflection response. This re-
response model, (5.53), gave rise to the model (5.52) for the
variational derivative of the medium reflection response.
This form of the variational derivative allows a straight-
forward evaluation of (5.42), both because of its simple
analytical form, and because it is deterministic, making the
expectation in (5.42) trivial. We have noted that the weakly
scattering medium model is of practical, as well as analytical,
importance; in particular, the isogradient layer model we used
in Chapter 4, with parameters typical of ocean bottom
sediments, may be substantially described by the weakly
scattering medium model.

5.4 Cramer-Rao Bound Computation for a Weakly Scattering
Medium

In this Section, we will compute the Cramer-Rao bound
for reflection coefficient density estimation in a medium
which is weakly scattering, that is, one whose reflection
response is well-approximated by (5.53). We will compute the
bound for two different sources, using a simple, first-order
model for \( K_g(\tau) \). Our particular objective is to determine
how estmimation performance is affected by the bandwidths of
source and reflection coefficient density.

We begin our development by specifying \( g(\xi) \). For the
covariance function, we select the first-order model
In Appendix C, we have demonstrated that (5.56) corresponds to a random process whose sample functions are continuous with probability one (almost everywhere). We saw in Section 5.3 that the mean of the process, $\bar{g}(\xi)$, does not affect estimation performance in the case of a weakly scattering medium. We therefore do not specify a process mean, except to require that it satisfy the smallness constraint necessary for the medium to be weakly scattering. Similarly, we assume that the zero-lag covariance, $\sigma_g^2$, be sufficiently small that (5.53) holds. For the sake of concreteness, we assume that the length of the observation interval, $T$, is 1 second.

Given our observation in Section 5.3, where we indicated that the sediment layer model of Chapter 4 could be considered a weakly scattering medium, we might think of $\bar{g}(\xi)$ as being given by (5.54). The problem we are addressing here might then be thought of as determining the accuracy with which we can estimate perturbations of the reflection coefficient density when the nominal profile corresponds to an isogradient sediment layer.

We first assume that the source, $d(t)$, is a unit impulse, so that (5.53) becomes

\[
K_g(\tau) = \sigma_g^2 \exp(-\gamma |\tau|). \tag{5.56}
\]

\[
s(t) = g(t). \tag{5.57}
\]
This is not a realistic source, since it has infinite energy, but it does lead to considerable mathematical simplification. We note, however, that (5.57) implies that the variational derivative of the reflection response is a delayed impulse; in this case, (5.42) is ill-defined, and we cannot apply the integral equation form (5.41) to compute performance. We may resolve this problem by noting that (5.57) implies that we may write the observation as

\[ r(t) = g(t) + n(t). \] (5.58)

The medium identification problem has been reduced, by the assumption of the model (5.57), to a problem of estimating a stationary, Gaussian random process observed in the presence of white noise. The optimal (minimum estimate variance) processor for performing this estimation is a time-varying Wiener filter, \( h_o(t, \tau) \), is defined by the Wiener-Levinson integral equation

\[
h_o(t, \tau) + \frac{1}{\sigma^2} \int_0^T K(g(t-u)h_o(u, \tau)du = \frac{1}{\sigma^2} K(g(t-\tau)). \tag{5.59}
\]

For the covariance function (5.56), (5.59) may be solved in closed form:

\[
h_o(t, \tau) = \frac{\gamma \sigma^2 [(\theta+\gamma)e^{\theta t}+(\theta-\gamma)e^{-\theta t}][(\theta+\gamma)e^{\theta(T-\tau)}+(\theta-\gamma)e^{-\theta(T-\tau)}]}{\sigma^2 \theta [(\theta+\gamma)^2 e^{\theta T}-(\theta-\gamma)^2 e^{-\theta T}]} \tag{5.60}
\]
where

\[ \theta = \gamma \left( 1 + 2 \sigma_g^2 / \gamma \sigma^2 \right)^{1/2}. \] (5.61)

Since the "message" portion of \( r(t), s(t|g) \), is linear in \( g(t) \), the optimal processor is efficient (its performance equals the Cramer-Rao bound). Moreover, we may write the optimal estimate variance using (5.60), since [31]

\[ \mathbb{E}[(g(t) - \hat{g}_o(t))^2] = \sigma^2 h_o(t,t), \] (5.62)

where \( \hat{g}_o(t) \) is the optimal estimate of \( g(t) \) formed by the Wiener filter. This gives

\[ \text{Var}[\hat{g}_o(t)] = \sigma^2 \frac{\gamma (\theta+\gamma) e^{\theta T} + (\theta-\gamma) e^{-\theta T} + 2(\theta^2 - 2\gamma^2) \cosh(\theta(T-2t))}{(\theta+\gamma) e^{\theta T} - (\theta-\gamma) e^{-\theta T}} \] (5.63)

We have sketched (5.63) in Figure 5.1; note that \( \text{Var}[\hat{g}_o(t)] \) is symmetric about the center of the observation interval. This behavior is reasonable, since in the absence of any definitive internal structure for the response (such as significant water-column multiples) we are estimating \( g(\xi) \) (or, more accurately, \( \hat{g}(\xi) \)) based on its correlation structure alone. In the center of the observation interval, we may obtain the best possible estimate of the correlation function.
Figure 5.1 -- Optimal estimator performance, weakly scattering medium model, impulse source.
of the received signal. Near the ends of the interval, less two-sided information is available to us, and estimate performance degrades.

We typify the performance of the processor for any set of parameter values in terms of $\text{Var}[\widehat{g}_o(0)] = \text{Var}[\widehat{g}_o(T)]$, and $\text{Var}[\widehat{g}_o(T/2)]$, the worst and best performance through the interval. In Figure 5.2, we have plotted $\text{Var}[\widehat{g}_o(0)]/\sigma_g^2$, and $\text{Var}[\widehat{g}_o(0.5)]/\sigma_g^2$, as functions of $\sigma_g^2/\sigma^2$, for various values of $\gamma$. We observe that estimate variance can be no worse than $\sigma_g^2$; if a processor did give a larger variance, we would be better off to simply use $\widehat{g}(t) = \overline{g}(t)$, since it has smaller variance. As $\gamma$ increases, $g(\xi)$ becomes less correlated, and hence should be more difficult to distinguish from the white background noise; as we see in Figure 5.2, performance indeed becomes poorer as $\gamma$ increases.

When $\theta T$ is large, as is the case in the plot of (5.63) we have sketched in Figure 5.1, we see that performance is substantially steady-state, invariant across the observation interval. If we know that our processor is operating in this asymptotic region, we may express performance in the considerably more convenient approximate form

$$\text{Var}[\widehat{g}_o(t)] = \frac{\sigma_g^2 \gamma}{\theta},$$

(5.64)

which holds away from the ends of the interval, and when $\theta T$ is
Figure 5.2 -- Optimal estimator performance, weakly scattering medium, impulse source. For each $\gamma$, the variance at $t = 0.$ is the upper curve, and the variance at $t = 0.5$, the lower.
large.

The second source waveshape which we will consider is the exponential

\[ d(t) = \sqrt{2\chi} \exp(-\chi t)u_1(t) \]  \hspace{1cm} (5.65)

where the leading coefficient has been chosen to make the source energy unity, so that SNR, in the sense of (4.29), is \( 1/\sigma^2 \). This choice will allow us to examine the effects of source bandwidth on estimation accuracy, something we could not do with the impulse source. As significantly, (5.65) allows relatively easy computation of the integral equation kernel.

Using the weakly scattering medium model for the variational derivative (5.52), and the definitions for source, process covariance, and a one-second observation interval, we may write the kernel, \( \tilde{K}(\eta,z) \), as

\[ \tilde{K}(\eta,z) = 2\chi \sigma^2 \int_0^1 \int_0^1 \exp(-2\chi t + \chi \eta + \tau) - \gamma |\tau - z|) \]

\[ \times u_1(t-\tau)u_1(t-\tau)dt\tau. \]  \hspace{1cm} (5.66)

This may be evaluated to give an analytical result, asymmetric in \( \eta \) and \( z \), whose form is more extensive than enlightening; we do not reproduce it here. We now must solve
Because of the complicated form of (5.66), exact analytic solution of (5.67) is virtually impossible, leaving us with two possible courses: obtain an "exact" numerical solution of (5.67), or an approximate analytic solution. We will pursue both options. First, we may reduce the number of cases we must consider by noticing that \( \tilde{K}(\eta, z) \) is proportional to \( \sigma^{-2} \),

\[
\tilde{K}(\eta, z) = \sigma^{-2} \kappa(\eta, z)
\]  

(5.68)

Also, since \( K_g(\tau) \) is proportional to \( \sigma^{-2} \), we may solve for \( J^{-1}(\xi, z) / \sigma^{-2} \), rather than for \( J^{-1}(\xi, z) \). The reparametrized equation is

\[
\frac{J^{-1}(\xi, z)}{\sigma^2} + \frac{\sigma^2}{\sigma^2} \int_0^1 J^{-1}(\xi, \eta) \kappa(\eta, z) d\eta = \exp(-\gamma|\xi-z|)
\]  

(5.69)

For the numerical solution of (5.69), we represent each of the functions in (5.69) by its values on an evenly-spaced grid of points in \([0,1] \times [0,1] \). The integration is replaced by a discrete sum, using an appropriate quadrature formula; we used the simplest possible, the rectangular method. This reduces (5.69) to a linear matrix equation which may be solved in a straightforward, but computationally
expensive manner, such as Gauss elimination. This integral equation solution technique is described in detail in Section 3.15 of [57].

We have solved (5.69) in this manner for various values of $\chi$, $\gamma$, and $\frac{\sigma^2_g}{\sigma^2} = \sigma^2_g \cdot \text{SNR}$. A typical solution for $J^{-1}(\xi, \xi)/\sigma^2_g$ is shown in Figure 5.3; estimation accuracy is best near the center of the observation interval and deteriorates near the end. Qualitatively, this is the same behavior which we observed for the impulse source case, as illustrated in Figure 5.1, and may be explained, in part, in the same way i.e., in terms of the amount of correlation information available for a given value of $\xi$. In contrast to the impulse case, however, the solution shown in Figure 5.3 is asymmetric; in fact, $J^{-1}(1.1.1.) = \sigma^2_g$. This is due to our use of a source which has non-zero duration. Near the end of the observation interval, or deeper into the medium, less reflected energy is observed from each point in the medium, since a portion of the point's response lies outside the observation interval. The observed response from $g(1)$ contains no energy at all, so that, in effect, $r(t)$ contains no information about $g(1)$. Hence, $J^{-1}(1.1.1.) = \sigma^2_g$ is to be expected. Because of attenuation, transmission loss, and incoherent scattering within the medium, we do not expect that this behavior is significant in practice.

Numerical solution of (5.69) is essential for insight
Figure 5.3 -- Cramer-Rao bound for reflection coefficient density estimation in a weakly scattering medium. Data obtained by numerical solution of (5.69), using a 50 x 50-point grid.

\[ J^{-1}(\xi, \xi) \frac{1}{\sigma^2} \]

\[ \gamma = \chi = 10. \]

\[ \text{SNR} \times \sigma_g^2 = 100. \]
into the detailed behavior of $J^{-1}(\xi, \xi)/\sigma_g^2$ across the observation interval. On the other hand, numerical methods provide only an empirical description of the performance bound as a function of the problem parameters; in this regard, an approximate analytic solution of (5.69), if sufficiently accurate, would be a desirable alternative. In addition, numerical methods become prohibitively expensive for large values of $\gamma$ and $\chi$, because the large associated bandwidths of $\mathcal{K}(\xi, z)$ and $K_g(\tau)$ require high-dimensional approximating matrices; solution of an $N\times N$ system of matrix equations requires on the order of $N^3$ multiples. We have found, as well, that if positive definiteness of the solution, as well as accuracy, are to be retained for large $\sigma_g^2$ SNR, double precision arithmetic is required. Analytical solution also provides a cross-check on the numerical solution, for which we have little prior intuition.

In Figure 5.3, we observe that $J^{-1}(\xi, \xi)/\sigma_g^2$ is constant over almost three-quarters of the observation interval. We have found that, for larger $\gamma$ or $\chi$, the region of constant $J^{-1}(\xi, \xi)$ is still larger. A reasonable approximate analytical method, then, is to compute the steady-state solution of (5.69) as we did for the impulse source case (5.64). In (5.69) and in (5.66), we allow the observation interval to become (doubly) infinite, so that both the kernel and the solution of the integral equation may be considered stationary.
Specifically, (5.69) becomes

$$\frac{J^{-1}_s(\xi-z)}{\sigma^2_g} + \frac{\sigma_g^2}{\sigma^2_g} \int_{-\infty}^{\infty} \frac{\mathcal{H}_s(\eta-z) d\eta = \exp(-\gamma|\xi-z|)}{\sigma^2_g} = (5.70)$$

where we have used the subscript "s" to denote steady-state values. We may solve (5.70) for $J^{-1}_s(\xi-z)/\sigma^2_g$ using Fourier transform methods; we require only $J^{-1}_s(\xi-\xi)/\sigma^2_g$, so that it is sufficient to compute $J^{-1}_s(0)/\sigma^2_g$, which is given by

$$J^{-1}_s(0) = \frac{1}{2\gamma} \frac{\int (2\pi f)^2 + \chi^2}{\int (2\pi f)^2 + \gamma^2} \frac{df}{\int (2\pi f)^2 + \chi^2 + 4\chi \gamma \sigma^2 g \cdot \text{SNR}}$$

This may be computed in closed form using the tabulated integral ([61], p. 300)

$$\int_{\alpha}^{\beta} \frac{u^{-1}}{t^2 + \alpha} \frac{dt}{(t^2 + \beta)} = \frac{\pi}{2} \frac{u^{-1}}{\beta - \alpha} \csc(u\pi/2)$$

$$J^{-1}_s(0) = \begin{cases} \frac{\gamma}{2\cos(\theta/2)} \left( \frac{1}{r^{1/2}} + \frac{\chi^2}{r^{3/2}} \right), (\gamma^2 + \chi^2)^2 \leq 4(\gamma^2 \chi^2 + 4\chi \gamma \sigma^2 \text{g} \cdot \text{SNR}) \\ \frac{1}{2\gamma} \frac{\int (2\pi f)^2 + \chi^2}{\int (2\pi f)^2 + \gamma^2} \frac{df}{\int (2\pi f)^2 + \chi^2 + 4\chi \gamma \sigma^2 g \cdot \text{SNR}} \end{cases}$$
where

\[ r = \sqrt{\gamma^2 + 4\chi \sigma_g^2 SNR}, \]

\[ r_+ = -\left(\gamma^2 + \chi^2\right) + \sqrt{\left(\gamma^2 + \chi^2\right)^2 - 4\left(\gamma^2 + 4\chi \sigma_g^2 SNR\right)} \]

and

\[ \theta = \tan \left[ \frac{(\gamma^2 + \chi^2)^2}{\gamma^2 + \chi^2} \right]^{1/2} \]

We have found that \( J_s^{-1}(0) \sigma_g^2 \) is an excellent approximation to the constant portion of the exact (numerical) solution of (5.69), for \( \chi, \gamma \geq 10T \); although we had strong motivation for believing that this would indeed be true, there was no way to verify it in advance. Of course, the larger \( \chi \) and \( \gamma \), the larger the constant region of \( J_s^{-1}(\xi, \xi)/\sigma_g^2 \), and the more uniform the approximation

\[ J_s^{-1}(\xi, \xi)/\sigma_g^2 \]

From this point forward, we will characterize the performance bound in terms of its steady state value, \( J_s^{-1}(0)/\sigma_g^2 \), bearing in mind that (5.74) holds only for large \( \chi \) and \( \gamma \geq 10T \), and, in any event, is not valid near the end-
points of the observation interval.

In Figure 5.4 we have plotted $J^{-1}_s(0)/\sigma_g^2$ as a function
of $\sigma_g^2$SNR for $\gamma = \chi = 10, 30, \text{ and } 100$. We expect that as the
response becomes more correlated, as $\gamma = \chi$ decreases, the
better our estimation performance should be, as we saw in the
impulse source case. Our intuition is verified by the com-
puted values. We may think of using a lowpass filter to
reject the white noise component which lies outside the band-
width of $d(t)^*g(t)$: the smaller this bandwidth, the more
narrowband a filter we may employ, and the more broadband
noise power we may reject.

We can reasonably expect that estimation performance
is affected by our choice of source bandwidth ($\chi$, in radian
measure) relative to the bandwidth of $g(\xi)(\gamma)$. Because the
source correlation function and $K_g(\tau)$ are of similar form
(exponential) we might expect that the best performance is
obtained when the two bandwidths are equal, i.e., $\chi = \gamma$.
This, however, is not true, as we see in Figure 5.5, where
we have plotted $J^{-1}_s(0)/\sigma_g^2$ against $\text{SNR}\sigma_g^2$, for $\gamma = 30$, and $\chi = 10, 30, \text{ and } 100$. We observe two regions of behavior:
for small $\sigma_g^2$SNR, a source more coherent than the reflection
coefficient density ($\chi < \gamma$) gives improved performance; for
large $\sigma_g^2$SNR, a source less coherent than the reflection
coefficient density ($\chi > \gamma$) gives improved performance. This
indicates that there is a distinct tradeoff to be made in
Figure 5.4 -- Steady-state performance bound with $X = Y$. 

$SNR \times \sigma_g^2$ vs. $\frac{d_g^{-1}(0)}{\sigma_g^2}$.
Figure 5.5 -- Steady-state performance bound with $\gamma = 30$, various $X$. 
choosing source bandwidth, \( \chi \): a small bandwidth (smaller than \( \gamma \)) reduces the amount of noise energy at the processor input, but reduces the resolution of the estimate; a large bandwidth (larger than \( \gamma \)) enables a high-resolution estimate of the detailed behavior of \( g(\xi) \), but also requires the processor to admit more noise energy. It is reasonable that the first case should be appropriate for small \( \sigma_g^2 \text{SNR} \), and the second, for large \( \sigma_g^2 \text{SNR} \).

Performance, however, is definitely not monotone in \( \chi \) in either region, e.g., we cannot obtain arbitrarily good performance, for large \( \sigma_g^2 \text{SNR} \), by using an arbitrarily large \( \chi \). Specifically, for large \( \sigma_g^2 \text{SNR} \), \( \sigma_g^2 \text{SNR} \gg (\gamma^2 - \chi^2)^2 / 16XY \) we may write (5.73) as

\[
\frac{J_s^{-1}(0)}{\sigma_g^2} = \frac{Y}{\sqrt{2}} \left( \frac{1}{(4XYg^2 \text{SNR})^{1/4}} \right) + \frac{\chi^2}{(4XYg^2 \text{SNR})^{3/4}} \tag{5.75}
\]

which indicates improving performance for \( \chi \) increasing and small, and deteriorating performance for \( \chi \) increasing and large. This result implies that it should be possible to define an optimal source bandwidth, \( \chi_{\text{opt}} \), one which yields the smallest (steady-state) Cramer-Rao bound value for a given \( \gamma \) and \( \sigma_g^2 \text{SNR} \):
\[
\frac{J_s^{-1}(0)}{\sigma_g^2} \leq \frac{J_s^{-1}(0)}{\sigma_g^2}, \quad X = X_{\text{opt}}
\]  
(5.76)

In (5.76), of course, we assume that the steady-state approximation (5.74) is valid.

Because of their restricted range of applicability, asymptotic expressions for \( \frac{J_s^{-1}(0)}{\sigma_g^2} \) such as (5.75) are not generally appropriate for computing \( X_{\text{opt}} \). Numerical minimization of (5.73) as a function of \( X \) is, however, straightforward, if tedious. In Figure 5.6, \( X_{\text{opt}} \) is plotted as a function of \( \sigma_g^2 \text{SNR} \) for \( \gamma = 30, 60, \) and 120. For large \( \sigma_g^2 \text{SNR} \), a good empirical fit was found to be

\[
X_{\text{opt}} \approx \left( \frac{\gamma \sigma_g^2 \text{SNR}}{e} \right)^{1/3},
\]

where \( e \) is the base of the natural logarithms. For smaller \( \sigma_g^2 \text{SNR} \), it appears that \( X_{\text{opt}} \approx (\sigma_g^2 \text{SNR})^{1/2} \), with no significant dependence on \( \gamma \). The accuracy of this approximation is in some doubt, since the extrema of \( \frac{J_s^{-1}(0)}{\sigma_g^2} \) in this region are extremely flat, reducing the accuracy with which \( X_{\text{opt}} \) may be computed.

It is of interest to define a threshold value for \( \sigma_g^2 \text{SNR} \), below which \( X_{\text{opt}} < \gamma \), and above which \( X_{\text{opt}} > \gamma \). Both direct empirical analysis of the data of Figure 5.6, as well
Figure 5.6 -- Optimal source bandwidth for various $\gamma$. 
as the approximate form (5.77) support the expression

\[(\sigma_{g v}^2)_{\text{thld}} = \gamma^2.\]  

(5.78)

Optimal performance, given by the steady-state bound evaluated at \(\chi = \chi_{\text{opt}}\), is shown in Figure 5.7, the companion result of Figure 5.6. This is, in effect, a "lower" lower bound on performance, since it is the value of the Cramer-Rao lower bound optimized for the choice of source bandwidth. We note that, for a given \(\sigma_{g v}^2\), performance improves with increasing coherence of \(g(\xi)\); this is the same result which we observed for the impulse source case, in Figure 5.2, and may be explained in the same way. An empirical fit to the data of Figure 5.7, for large \(\sigma_{g v}^2\), gave

\[\left(\frac{J^{-1}(0)}{\sigma_{g v}^2 \text{opt}}\right) \approx \left(\frac{2\gamma^2}{7\sigma_{g v}^2}\right)^{1/3}\]  

(5.79)

An important issue associated with the definition of an optimal source bandwidth, \(\chi_{\text{opt}}\), is that of sensitivity: how much is performance degraded if we use \(\chi \neq \chi_{\text{opt}}\)? This is an important practical aspect of the problem, since it will determine how zealous we must be in matching our source to a given set of conditions. In Figure 5.8, we have plotted \(J^{-1}(0)\sigma_{g v}^2\) as a function of \(\chi\), for several values of \(\sigma_{g v}^2\). It is evident that performance is not strongly affected by
Note: $\chi_{\text{opt}}$ is a function of both $\gamma$ and $\text{SNR} \times \sigma_g^2$.

Figure 5.7 -- Steady-state performance bound for various $\gamma$, with $\chi=\chi_{\text{opt}}$. 
Figure 5.8 -- Steady-state performance bound for various $\text{SNR} \times \sigma_g^2$, as a function of $\chi$. $\chi_{\text{opt}}$ is indicated for each curve.
moderate deviations from an optimal value for $\chi$: a $\pm 50\%$ error in $\chi$, for the cases considered, gives a performance reduction of at most on the order of 10%. We conclude, therefore, that performance is substantially robust with respect to our choice of $\chi_{\text{opt}}$, although we might expect sensitivity, expressed in percentage terms, to increase somewhat with increasing $\sigma_g^2 \text{SNR}$.

In this Section, we have applied the Cramer-Rao bound formulation derived in Section 5.2, (5.41), and the reflection response model for a weakly scattering medium developed in Section 5.3, (5.52) and (5.53), to the actual computation of performance bounds for structure estimation in a weakly scattering medium whose structure covariance function was given by the simple first order-model, (5.56). We computed the bound both for a unit impulse source, and for a low-pass (exponential) source, (5.65). For both cases, we noted that estimation accuracy improves as $g(\xi)$ becomes more correlated across the observation interval (as $\gamma$ decreases). For the low-pass source, using a combination of exact numerical methods and approximate analytical techniques, we showed that one could define an optimal source bandwidth, $\chi_{\text{opt}}$, for a given $\gamma$, $\sigma_g^2$, and SNR. While we also saw that estimator performance was relatively insensitive to moderate errors in the choice of source bandwidth about the optimal value, the existence of such an optimal bandwidth is at odds with conventional wisdom in the seismic community, which holds that
larger source bandwidths are uniformly advantageous, without regard for the extra broadband noise which must then be admitted by the receiver. We must qualify our result, however, in that noise is not generally white, but, in particular, falls off for higher frequencies (as in Figure 4.14, for instance), so that one may indeed gain some advantage by using a large-bandwidth source. Such a determination must be made, however, using a realistic noise model. Our results, if not completely accurate, do indicate that the question of choice of source bandwidth should be addressed in a more cautious manner than has heretofore been the rule.
APPENDIX A
NUMERICAL COMPUTATION OF THE FISHER INFORMATION MATRIX

Suppose we have a sequence of observations, \( r(k) \), of the form

\[
r(k) = s(k|g) + n(k), \quad k=1,2,...,L
\]  

(A.1)

where \( n(k) \) is a white, zero-mean, Gaussian random sequence, and \( s(k|g) \) is parametrically dependent on \( g \), an unknown, nonrandom vector of dimension \( N \). If we form an unbiased estimate of \( g \) from these observations, \( \hat{g}(r) \), the Cramer-Rao lower bound on the estimate variance is

\[
E[(g_i(r) - g_i)^2] \geq (J^{-1})_{ii}
\]  

(3.2)

where

\[
J = \frac{1}{\sigma^2} \nabla g(s(g))^T \nabla g(s(g))
\]  

(3.15)

or

\[
(J)_{ij} = \frac{1}{\sigma^2} \sum_{k=1}^{L} \frac{\partial s(k|g)}{\partial g_i} \frac{\partial s(k|g)}{\partial g_j}
\]  

(A.2)

and
We will be concerned with the evaluation of $J$ when $\{s(k|g)\}$ is the output sequence of a particular discrete linear system. The system is

$$x(k+1) = \Phi(g)x(k) + \Gamma d(k)$$  \hspace{1cm} (2.38)

$$s(k|g) = M^T x(k)$$ \hspace{1cm} (A.3)

with

$$x(0) = 0.$$ \hspace{1cm} (A.4)

Note that only the system matrix, $\Phi(g)$, is dependent on the unknown parameter vector, $g$. Because of the complicated dependence of $s(k|g)$ on $g$ for this problem, it is more useful to develop an iterative, numerical method for computing $J$ rather than to attempt to find a general analytical result.

We may express $x(k)$ using the discrete variation of constants formula:

$$x(k) = \Phi(g)^k x(0) + \sum_{\ell=0}^{k-1} \Phi(g)^{k-1-\ell} \Gamma d(\ell)$$

$$= \sum_{\ell=0}^{k-1} \Phi(g)^{k-1-\ell} \Gamma d(\ell),$$ \hspace{1cm} (A.5)
where we have used (A.4). So

$$s(k|g) = \sum_{\ell=0}^{k-1} M^T \phi(g)^{k-1-\ell} \Gamma d(\ell)$$

(A.6)

and

$$\frac{\partial s(k|g)}{\partial g_m} = \sum_{\ell=0}^{k-1} M^T \frac{\partial \phi(g)^k}{\partial g_m} \Gamma d(\ell).$$

(A.7)

First, we define a method for computing $\frac{\partial \phi(g)^{k-1-\ell}}{\partial g_m} \Gamma$. recursively. Let

$$\psi(m) = \phi(g)^n \Gamma$$

(A.8)

with $\psi(0) = 0$. Then

$$\psi(m) = \frac{\partial \phi(g)}{\partial g_m} \phi(g)^n \Gamma + \phi(g) \psi(m)$$

(A.9)

If we define

$$\Gamma_{n+1} = \phi(g) \Gamma_n$$

(A.10)

with $\Gamma_1 = \Gamma$, we may propagate $\psi(m)$ according to

$$\psi(m) = \frac{\partial \phi(g)}{\partial g_m} \frac{\Gamma}{n} + \phi(g) \frac{\psi(m)}{n}.$$
We then write

$$\frac{\partial s(k|g)}{\partial g_m} = M^T \sum_{\ell=0}^{k-1} \psi_{k-\ell} d(\ell).$$  \hspace{1cm} (A.12)

For the system of concern, we have a particularly simple form for $M^T$,

$$M^T = (1,0,...,0).$$  \hspace{1cm} (2.43)

This implies that, in order to compute $\frac{\partial s(k|g)}{\partial g_m}$ for any $k$, we must retain only the first entry of $\psi(\ell, l<k$. Moreover, if the source, $d(k)$ has length $K$ samples, we need only retain the most recent $K$ or $k$ values of $\psi(\ell, which ever is smaller, for use in the sum of (A.12). These observations allow us to reduce the amount of storage space and computation considerably, in comparison to that required to evaluate (A.12) in its most general form.

To simplify our notation, we define the matrix of column vectors $\psi$:

$$\psi_n = \begin{bmatrix} \psi^{(1)}_n \\ \vdots \\ \psi^{(N)}_n \end{bmatrix}.$$  \hspace{1cm} (A.13)

We now may write the propagation equations for $\psi^{(m)}_n$, $m=1,N$, in terms of one equation for $\psi_n$:
\[
\frac{\psi}{n+1} = \left[ \frac{\partial \phi}{\partial g_1} \frac{r}{n}, \ldots, \frac{\partial \phi}{\partial g_N} \frac{r}{n} \right] + \phi(g) \frac{\psi}{n},
\]

(A.14)

with \( \frac{\psi}{n} = 0 \). \( \frac{\psi}{n} \) has dimension \( \text{dim}(x(k)) \times N \).

Figure A.1 shows the algorithm in flowchart form.

Although we have explicitly indicated the algorithm for the computation of \( J \), it may be easily modified to compute other quantities involving \( \frac{\partial s(k|g)}{\partial g_m} \), such as the gradient of \( I(\hat{g}) \):

\[
\frac{\partial I(\hat{g})}{\partial \hat{g}_m} = \frac{\partial}{\partial g_m} \frac{L}{\sum_{k=1}^L} (r(k) - s(k|\hat{g}))^2
\]

\[
= -2 \sum_{k=1}^L \frac{\partial s(k|\hat{g})}{\partial \hat{g}_m} (r(k) - s(k|\hat{g}))
\]

which is of interest in maximum-likelihood parameter estimation schemes.
Figure A.1 -- Flowchart for computation of Fisher information matrix, $J$. 

\[ \frac{\partial s(k|g)}{\partial g_m} = \min(k,K)-1 \sum_{\ell=0}^{\min(k,K)-1} M_{\psi}^{T} \psi^{(\ell)} Min(k,K)-1-\ell \mathbf{d}(\ell) \]

FOR \( m = 1, \ldots, N \), USING STORED \( M_{\psi}^{T} \).

\[ J = J + \left[ \frac{\partial s(k|g)}{\partial g_1} \ldots \frac{\partial s(k|g)}{\partial g_N} \right] \cdot \left[ \frac{\partial s(k|g)}{\partial g_1} \ldots \frac{\partial s(k|g)}{\partial g_N} \right] \]

\[ \text{k = k + 1} \]

\[ \text{k = L?} \]

\[ J = J/\sigma^2 \rightarrow \text{DONE} \]

\[ \psi_{k+1} \quad (A.14) \]

\[ \Gamma_{k+1} \quad (A.10) \]
APPENDIX B

In this Appendix, our objective is to prove a result required in the derivation of (5.39):

\[
\frac{\partial s(t|g_K)}{\partial g_i} = \int_0^\infty \frac{\delta s(t,z)}{\delta g_K} \phi_i(z) \, dz,
\]

where both the partial derivative and the variational derivative are continuous functionals, as is \( s(t|g) \), and \( g_K(z) \) is given by

\[
g_K(z) = \sum_{k=1}^{K} g_k \phi_k(z).
\]

The set of functions \( \{\phi_k(z)\} \) are the eigenfunctions of the integral equation (5.8), and hence are (uniformly) continuous on \([0,\infty)\) ([52], p. 84). We will restrict our proof, therefore, to the case in which \( \phi_i(z) \) is continuous. We note that we will not require a strict-sense definition of the variational derivative, that is, it is sufficient that it be defined in the operator sense of (5.32).

Let \( \phi_i(z) \) be approximated by a sequence of simple functions, \( y^N_N(z) \):

\[
y^N_N(z) = \sum_{n=1}^{N} h^n_N(z) \phi_i(z^n_n)
\]
where \( h^N_n(z) \) has unit height, the support sets of \( h^N_n(z) \) and \( h^N_m(z) \) are disjoint, and \( z^N_n \) lies in the support set of \( h^N_n(z) \). We may define \( \{ h^N_n(z) \} \) so that \( y_N(z) \) converges uniformly to \( \phi_i(z) \), since \( \phi_i(z) \) is uniformly continuous.

Define

\[
g'_K(z) = g_i \phi_1(z) + \ldots + (g_i + \Delta g) \phi_i(z) + \ldots + g_k \phi_k(z)
\]  

We may write, since \( s(t|g) \) is continuous,

\[
s(t|g'_K) - s(t|g_K) = s(t|g_K(z)) + \Delta g \sum_{n=1}^{N} h^N_n(z) \phi_i(z^N_n))
- s(t|g_K(z)) + \Delta g \sum_{n=1}^{N-1} h^N_n(z) \phi_i(z^N_n))
+ s(t|g_K(z)) + \Delta g \sum_{n=1}^{N-1} h^N_n(z) \phi_i(z^N_n))
- \ldots - s(t|g_K).
\]

We now regroup terms, writing each successive pair in terms of the variational derivative of \( s(t|g) \)

\[
s(t|g') - s(t|g_K) = \frac{\delta s(t, z^N_N)}{\delta (g_K(z)) + \Delta g \sum_{n=1}^{N-1} h^N_n(z) \phi_i(z^N_n))} \Delta g \phi_i(z^N_N) \Delta z_N
+ \ldots + \frac{\delta s(t, z^N_1)}{\delta (g_K(z)) + \Delta g h^N_1(z) \phi_i(z^N_1))} \Delta g \phi_i(z^N_1) \Delta z_1,
\]  

(B.4)
where \( \Delta z_i^N \) is the length of the support of \( h_i(z) \). Now, we divide through by \( \Delta g \), and take the limit as both \( \Delta g \to \infty \) and \( N \to \infty \). Since the variational derivative is continuous, we recognize that the right-hand side converges to the desired integral expression, and that the left-hand side converges to the partial derivative. Hence, (5.33) is proven.

We may view the result (5.33) as the projection of a generalized gradient (the variational derivative) on a given coordinate in a space of functions. Note that our derivation requires only that the variational derivative be defined for a particularly simple class of perturbations, i.e., those which are constant across the region of perturbation.
C.1 Introduction

In Section 5.2, we derived a novel form for the Cramer-Rao lower bound, (5.39), appropriate for waveform estimation of a random process, \( g(\xi) \), given an observation of the form

\[
r(t) = s(t|g) + n(t),
\]

(5.1)

where \( s(t|g) \) is a nonlinear functional of \( g(\xi) \). In our derivation, we assumed that \( s(t|g) \), and several related quantities, were well-defined and obeyed certain regularity conditions; before we can apply the result (5.39) to a given problem, we must be assured that \( s(t|g) \) meets these criteria. In this Appendix, we will demonstrate that the reflection response of a random medium is a well-defined functional of the reflection coefficient density, and possesses the regularity properties required for (5.39) to hold. In short, we are demonstrating that the theoretical result of Section 5.2 indeed is applicable to our problem.

In Section C.2, we address the problem of choosing an appropriate class of random processes for the reflection coefficient density function. In the derivation of the
bound, we required that the reflection response, and a partial derivative, be continuous functionals; this is demonstrated in Section C.3. Finally, in Section C.4, we develop a form for the variational derivative of the reflection response with respect to the reflection coefficient density function, and show that, as required, it too is a continuous functional.

C.2 Constraints on $g(\xi)$

Our first step in demonstrating the required properties for the reflection response of a medium with a random process reflection coefficient density is to consider what constraints must be imposed on the process model to ensure that the reflection response is well-defined. The propagation equations (2.17) of the general medium model introduced in Section 2.2 were valid for piecewise continuous $g(\xi)$. If we allow $g(\xi)$ to be a Gaussian random process, we must show that this continuity condition is still met with probability one, i.e., that the class of random processes we select has sample functions which are piecewise continuous, except perhaps for a set of functions which never occur (have probability zero). We cannot, for instance, allow the reflection coefficient density process to contain a white noise component, since no sample function of a white noise process is continuous (the sample functions are discontinuous with probability one). Let us decompose $g(\xi)$ as
\[ g(\xi) = \overline{g}(\xi) + \check{g}(\xi) \quad \text{(C.1)} \]

where \( \overline{g}(\xi) \) is the process mean, and \( \check{g}(\xi) \) a zero-mean random process. We allow \( \overline{g}(\xi) \) to be piecewise continuous, but require that \( \check{g}(\xi) \) be almost-surely sample function continuous; this ensures that \( g(\xi) \) is almost-surely sample-function piecewise continuous. We formalize the concept of an almost-surely sample function continuous random process in the following definition [52]:

**Definition:** A random process, \( g(\xi, \omega) \) is almost surely sample function continuous if it is separable* and if

\[ P_r \left( \bigcup_{\xi \in [0,T]} \{ \omega : \lim_{\Delta \to 0} |g(\xi+\Delta, \omega) - g(\xi, \omega)| \neq 0 \} \right) = 0 \quad \text{(C.2)} \]

where we have used the notation \( g(\xi, \omega) \) to indicate that the random process is a mapping from a probability space, \( \Omega \), into a space of functions indexed by \( \xi \).

That is, if the set of events corresponding to sample functions of the process which are discontinuous at any point in the index set have probability zero, we will (almost) never

---

*Separability is a technical property which, crudely speaking, allows us to characterize \( g(\xi, \omega) \) at the uncountably infinite number of points in the index set \([0,T]\) in terms of its behavior on a countably infinite set which is dense in the index set (in terms of the limit of its sequence of finite-dimensional distributions). For a separable process, it is sufficient to prove theorems for a countably infinite dense subset of the index set—a considerable simplification. We will impose separability on \( \check{g}(\xi) \).*
observe a discontinuous sample function. Thus, imposition of sample function continuity on $\tilde{g}(\xi)$ will satisfy the continuity requirements of (2.17).

This definition, however, is not useful as a practical means of determining if a process is sample function continuous. We require a condition equivalent to (C.2) which we may use to test a given process for sample function continuity. A sufficient condition which is of considerable utility is given by the following theorem [52]:

**Theorem:** (Kolmogorov's condition) Let $\{\tilde{g}(\xi, \omega), \xi \in [0, T]\}$ be a separable process on a finite interval. If there exist constants $\eta$, $\tau$, $\gamma$ such that

$$E[|\tilde{g}(\xi+\Delta) - \tilde{g}(\xi)|^\eta] \leq \gamma \Delta^{1+\tau} \quad (C.3)$$

with $\eta > 0, \tau > 0, \gamma > 0$, then almost every sample function is (uniformly) continuous on the interval.

This theorem reduces the problem of assuring the sample function continuity of $\tilde{g}(\xi)$ to one of specifying an appropriate covariance function for $g(\xi)$, $K_g(\tau)$, since, with the zero-mean Gaussian assumption, selection of $K_g(\tau)$ is the only remaining way we can influence the statistics of the process. The class of covariance functions which corresponds to sample function continuous Gaussian random
processes is, happily, rather large, so that no severe restrictions are placed on our choice of $K_g(\tau)$. In particular, a Gaussian random process with the correlation function

$$K_g(\tau) = \sigma^2 g \exp(-\gamma|\tau|)$$

is sample function continuous. This may be shown by evaluating the left-hand side of (C.3) for $n = 4$.

In addition to the requirements imposed on $g(\xi)$ by the nature of $s(t|g)$ itself, an additional requirement was specified in the derivation, in the definition of the metric to be used in the space of reflection coefficient density functions:

$$||g_1(\xi) - g_2(\xi)|| = \int_0^T |g_1(\xi) - g_2(\xi)| d\xi$$

We have required that $\tilde{g}(\xi)$ be sample function continuous on the compact interval $[0,T]$, and that $\bar{g}(\xi)$ be at worst piecewise continuous. This implies that $g(\xi)$ must be sample-function uniformly continuous on the interval, and must therefore be sample function bounded (the set of sample functions of $g(\xi)$ which are unbounded has probability zero). Almost every sample function of $g(\xi)$ is therefore absolutely integrable (since it is bounded), and the metric (5.21) is an appropriate one.

In this Section, we have shown that (almost everywhere) sample function continuity of $\tilde{g}(\xi)$, and piecewise continuity
of $\tilde{g}(\xi)$ are sufficient to ensure the existence of $s(t|g)$, using the requirements of the propagation model defined in Section 2.2. These continuity conditions, together with our restriction of $\xi$ to a compact interval, ensure as well that the metric (5.21) is defined on a sample function basis. We have also presented a sufficient condition for sample function continuity, (C.3), and have demonstrated a potentially useful form for $K_g(\tau)$, (C.4).

C.3 Continuity of the Response and Partial Derivative

In this Section, we will verify two properties required in the derivation of (5.39):

1. $s(t|g)$ is a continuous functional,
2. $\delta s(t|g + \alpha g')/\delta \alpha$ exists, and is a continuous functional.

Here, continuity is defined in terms of the metrics (5.20) and (5.21).

We first verify that the reflection response, $s(t|g)$ is a continuous function of $g(\xi)$. To do this, we make use of some results developed in [22], where a general solution method for the distributed system model of Figure 2.1 was developed.* In particular, the impulse response of a medium

*In [22], $g(\xi)$ was assumed to be uniformly continuous on $[0,T]$; we have required only piecewise continuity. An examination of the proofs of [22], however, indicates that piecewise continuity of $g(\xi)$ is sufficient to obtain the same results.
with reflection coefficient density \( g(\xi) \), \( s_0(t|g) \), must obey the integral equation ([22], equation (3.2), and Section 5.2)

\[
 h(t,\tau) + \int_0^t h(t,\mu)s_0(\mu-\tau|g)\,d\mu = s_0(t-\tau|g) \tag{C.5}
\]

where \( h(t,\tau) \) is given by the series ([22], (3.13))

\[
 h(x+\tau,\tau) = \sum_{m=0}^{\infty} (-1)^m f_m(x,\tau) \tag{C.6}
\]

and

\[
 f_0(x,\tau) = g(\tau) \tag{C.7a}
\]

\[
 f_m(x,\tau) = \int_0^x f_{m-1}(\tau,\mu)g(\mu+\tau)\,d\mu \tag{C.7b}
\]

The sum (C.6) is uniformly and absolutely convergent. Equation (C.5) is a Volterra integral equation, which may be solved for \( s_0(t|g) \) using the Neumann series [57], an integral series in \( h(t,\tau) \) which is also uniformly and absolutely convergent. This implies that \( s_0(t|g) \) may be expressed in terms of a single integral series in \( g(\xi) \) (the composition of (C.6) and the Neumann series solution for (C.5)) which is uniformly and absolutely convergent.

We therefore have an expression for \( s(t|g) = d(t)*s_0(t|g) \) as the integral of a uniformly, absolutely convergent integral series in \( g(\xi) \), and hence it is a continuous functional of \( g(\xi) \) in the sense of (5.20) and (5.21).
The proof of continuity for the partial derivative form is substantially more complicated, since we must derive a new series and demonstrate that it converges uniformly and absolutely. We formally take the partial derivative of (C.5):

\[
\frac{\partial h(t,\tau)}{\partial \alpha} + \int_0^t \frac{\partial h(t,\mu)}{\partial \alpha} s(\mu-\tau|g+ag')\,d\mu
\]

\[
+ \int_0^t h(t,\mu) \frac{\partial s(\mu-\tau|g+ag')}{\partial \alpha} \,d\mu = \frac{\partial s(t-\tau|g+ag')}{\partial \alpha}
\]  

(C.8)

As was (C.5), (C.8) is a Volterra integral equation, so that, if \( \partial h(x+\tau,\tau)/\partial \alpha \) exists, so does \( \partial s(t|g+ag')/\partial \alpha \), and, moreover, it may be found by summing the Neumann series for (C.8). If we can show that \( \partial h(x+\tau,\tau)/\partial \alpha \) is given by a uniformly, absolutely convergent integral series, as was \( h(x+\tau,\tau) \), then we may argue as we did for \( s(t|g) \) that \( \partial s(t|g+ag')/\partial \alpha \) is a continuous functional.

We will follow the method of proof employed in Chapter 3 of [22] to demonstrate the uniform and absolute convergence of (C.5). Consider the formal (term-by-term) derivative of (C.5):

\[
\frac{\partial h(x+\tau,x)}{\partial \alpha} = \sum_{m=0}^{\infty} (-1)^m \frac{\partial f_m(x,\tau)}{\partial \alpha}
\]  

(C.9)

We may write \( f_m(x,\tau) \) by successive application of (C.7) as
\[ f_m(x, \tau) = \int_x^\tau \int_{\mu_1}^{\mu_2} \ldots \int_{\mu_{m-2}}^0 g''(\tau+\mu_1)g''(\mu_1+\mu_2) \]
\[ \ldots g''(\mu_m)\,d\mu_1 \ldots \,d\mu_m \]  
(C.10)

where we have used \( g''(\xi) = g(\xi) + \alpha g'(\xi) \). Then

\[ \frac{\partial f_m(x, \tau)}{\partial \alpha} = \int_x^\tau \int_{\mu_1}^{\mu_2} \ldots \int_{\mu_{m-2}}^0 g'(\tau+\mu_1)g''(\mu_1+\mu_2) \]
\[ \ldots g''(\mu_m)\,d\mu_1 \ldots \,d\mu_m \]
\[ + \ldots + \int_x^\tau \int_{\mu_1}^{\mu_2} \ldots \int_{\mu_{m-2}}^0 g''(\tau+\mu_1)g''(\mu_1+\mu_2) \]
\[ \ldots g'(\mu_m)\,d\mu_1 \ldots \,d\mu_m \]  
(C.11)

and

\[ \left| \frac{\partial f_m(x, \tau)}{\partial \alpha} \right|^2 \leq (\int_x^\tau \int_{\mu_1}^{\mu_2} g'(\tau+\mu_1) \,d\mu \ldots \,d\mu_m)^2 \]
\[ + \ldots + (\int_x^\tau \int_{\mu_1}^{\mu_2} g''(\tau+\mu_1) \,d\mu \ldots \,d\mu_m)^2 \]  
(C.12)

We will bound each term of C.12 to obtain a bound on
\[ \left| \frac{\partial f_m(x, \tau)}{\partial \alpha} \right|^2 \]. First, we set up some notation. Let

\[ L^2(x, y) = \int_x^0 g''(\mu+y) \,d\mu \]  
(C.13)
\[ \tilde{L}^2(x,y) = \int_0^x g'(\mu+y)^2 d\mu \]  \hspace{1cm} (C.14)

Observe that
\[ \int_0^\mu L^2(v,x) \int_0^y L^2(x,y) dy dx = \frac{1}{2} \left[ \int_0^\mu L^2(v,y) dy \right]^2, \]  \hspace{1cm} (C.15)

and
\[ \tilde{L}^2(v,x) \leq \tilde{L}^2(T,0), L^2(v,x) \leq L^2(T,0). \]  \hspace{1cm} (C.16)

We consider one term of (C.12), and find, by repeated application of the Schwartz inequality,

\[ \int_0^\mu \int_0^{\mu-1} \int_0^{\mu-2} \cdots \int_0^{\mu-m+1} \left| g''(\mu_1+\mu_2) \right| \cdots \left| g''(\mu_{k+1}) \right| \]

\[ \cdots \left| g''(\mu_m) \right| d\mu_1 \cdots d\mu_m \]

\[ \leq \int_0^T g''(\mu)^2 d\mu L^2(x,\tau) \]

\[ \times \int_0^\mu \int_0^{\mu-1} \int_0^{\mu-2} \cdots \int_0^{\mu-m+1} L^2(\tau,\mu_1) \cdots \tilde{L}^2(\mu_{k+1}, \mu_{k+1}) \]

\[ \cdots L^2(\mu_{m-2}, \mu_{m-1}) d\mu_1 \cdots d\mu_{m-1} \]  \hspace{1cm} (C.17)

By successive application of (C.15), the right-hand side may be written

\[ L^2(x,\tau) \int_0^T g'(\mu)^2 d\mu \int_0^\mu \int_0^{\mu-1} \int_0^{\mu-2} \cdots \int_0^{\mu-m+1} L^2(\tau,\mu) \cdots \tilde{L}^2(\mu_{k+1}, \mu_{k+1}) \]

\[ \times \frac{1}{(m-k-2)!} \left[ \int_0^{\mu_k} L^2(\mu_{k+1}, y) dy \right]^{m-k-2} d\mu_1 \cdots d\mu_{k+1} \]  \hspace{1cm} (C.18)
Using (C.16), this is bounded above by

\[ L^2(T, 0) L^2(T, 0) \int_0^\infty \ldots \int_0^{\mu_{k-1}} L^2(\tau, \mu_1) \ldots L^2(\mu_{k-1}, \mu_k) \]

\[ \times \frac{1}{(m-k-2)!} [\int_0^{\mu_k} L^2(\mu_{k+1}, y) dy]^m-k-2 \, d\mu_1 \ldots d\mu_{k+1} \quad (C.19) \]

Observe that \( n \geq \mu_{k+1} \) implies

\[ \int_0^{\mu_k} L^2(\mu_{k+1}, y) dy \leq \int_0^{\mu_k} L^2(\eta, y) dy \quad (C.20) \]

so we may upper bound the integral

\[ \int_0^{\mu_k-1} \left[ \int_0^{\mu_k} L^2(\mu_{k+1}, y) dy \right]^{m-k-2} \, d\mu_{k+1} \]

\[ \leq \int_0^{\mu_k-1} \left[ \int_0^{\mu_k} L^2(\mu_{k-1}, y) dy \right]^{m-k-2} \, d\mu_{k+1} \]

\[ \leq T \left[ \int_0^{\mu_k} L^2(\mu_{k-1}, y) dy \right]^{m-k-2}. \quad (C.21) \]

Using (C.21) in (C.19), and applying (C.15) again, we may overbound (C.19) by

\[ T L^2(T, 0) L^2(T, 0) \int_0^{\infty} L^2(\tau, y) dy \]

\[ \frac{[\int_0^{\infty} L^2(\tau, y) dy]^{m-2}}{(m-2)!} \quad (C.22) \]

Finally, we see that (C.22) may be overbounded by
Equation (C.23) is an upper bound for each of the \( m \) terms on the right-hand side of (C.11), so we may write

\[
\left| \frac{\partial f_m(x, \tau)}{\partial \alpha} \right|^2 \leq \frac{m}{(m-2)!} \tilde{L}^2(T,0)(TL^2(T,0))^{m-1}
\]

or,

\[
\left| \frac{\partial f_m(x, \tau)}{\partial \alpha} \right| \leq \sqrt{\frac{m}{(m-2)!}} \tilde{L}(T,0)(L^2(T,0)T^{\frac{m-1}{2}})
\]

\[= F_m, \quad m \geq 2. \quad \text{(C.24)}\]

We required that every sample function of \( g(\xi) \) be piecewise continuous on \([0,T]\) with probability one; this implies each sample function is bounded on \([0,T]\) with probability one, so that \( L^2(T,0), \tilde{L}^2(T,0) < \infty \) with probability one. \( F_m \) is therefore of the form

\[
F_m = C_0 C_1^{m-1} \left( \frac{m}{(m-2)!} \right), \quad m \geq 2 \quad \text{(C.25)}
\]

where \( C_0 \) and \( C_1 \) are finite. We note that

\[
\left| \frac{\partial f_0(x,y)}{\partial \alpha} \right| = |g'(\tau)| \leq \sup_{\tau \in [0,T]} |g'(\tau)| = F_0 \quad \text{(C.26a)}
\]
and

\[
\left| \frac{\partial f_l(x, \tau)}{\partial \alpha} \right| = \left| \frac{\partial}{\partial \alpha} \int_0^x g(\mu_1 + \tau) g(\tau) d\mu_1 \right| \\
\leq \sqrt{2} L(T, 0) \tilde{L}(T, 0) = F_l
\]

where both \( F_0 \) and \( F_1 \) are finite by the above argument. Therefore, since \( \sum_{m=0}^{\infty} F_m \) converges uniformly for all \((x, \tau)\) (as may be seen by using the ratio test on (C.25), (C.9) converges uniformly and absolutely. This implies that \( \partial s(t|g)/\partial \alpha \) is a continuous functional of \( g(\xi) \), and we are done.

In this Section, we have demonstrated that the reflection response of a medium is a continuous functional of the reflection coefficient density function. We have also demonstrated the existence of the required partial derivative form, and have shown that it too is a continuous functional.

C.4 Existence and Continuity of the Variational Derivative

We will now consider the central element of our Cramer-Rao bound result (5.39), the variational derivative for the reflection response of a medium. We will compute a form for the variational derivative in the strong sense, that is, we will show that it is a functional of \( g(\xi) \) which takes values in the space of bounded functions on \([0, T]\) (the space implied by the metric (5.20)), rather than
Figure C.1 -- System model for the reflection response of a medium whose impulse response in the absence of the free surface is $G(f|g)$. 
an operator (e.g., it may not be an impulse). Moreover, we will demonstrate that it is a continuous functional. We will compute the variational derivative directly, by perturbing the reflection coefficient density of a medium and observing the effect this has on the reflection response.

Our problem becomes easier if we consider the medium reflection response in the absence of the free surface. We may achieve this by using a feedback model for the response, as shown in Figure C.1. We may then write

\[ s(t\mid g) \leftrightarrow S(f \mid g) = \frac{D(f)G(f\mid g)}{1 + G(f\mid g)}, \]  

(C.27)

where \( G(f\mid g) \) is the transform of the reflection impulse response of the medium in the absence of the free surface. We formally take the variational derivative of (C.22) to obtain

\[ \frac{\delta s(t, \xi)}{\delta g} \leftrightarrow \frac{\delta s(f, \xi)}{\delta g} = \frac{D(f)}{(1 + G(f\mid g))^2} \frac{\delta G(f, \xi)}{\delta g} \]  

(C.28)

(compare (4.21)), and observe that it is sufficient to show that \( \frac{\delta G(f, \xi)}{\delta g} \) exists and is a continuous functional of \( g(\xi) \).

We will solve for the impulse reflection response using a representation developed in [22]. First, we take the frequency transform of the propagation equations (2.17),
\[
\frac{s}{2} P_1(s, \xi) = - \frac{3P_1(s, \xi)}{\partial \xi} + g(\xi)(P_1(s, \xi) - P_2(s, \xi))
\]  
(C.29a)

\[
\frac{s}{2} P_2(s, \xi) = \frac{3P_2(s, \xi)}{\partial \xi} + g(\xi)(P_1(s, \xi) - P_2(s, \xi))
\]  
(C.29b)

where we have shifted to Laplace transform notation for compactness, and

\[
P_1(s, \xi) \leftrightarrow p^+(t, \xi) \quad \text{(C.30a)}
\]

\[
P_2(s, \xi) \leftrightarrow p^-(t, \xi) \quad \text{(C.30b)}
\]

We want to compute the response of the medium, \(P_2(s, 0)\), given the boundary conditions \(P_1(s, 0) = 1\) (a unit impulse source), and \(P_2(s, T) = 0\) (no upcoming waves from deep sources, a radiation condition); the situation is illustrated schematically in Figure C.2. We introduce the transformed variables

\[
X_1(s, \xi) = W(\xi) \left(-1e^{-s\xi/2}P_2(s, \xi)\right) \quad \text{(C.31a)}
\]

\[
X_2(s, \xi) = W(\xi) \left(-1e^{s\xi/2}P_1(s, \xi)\right) \quad \text{(C.31b)}
\]

where \(W(\xi) = (a(\xi)/a(0))^{1/2}\). We may write (C.29), using the state vector \(X(s, \xi) = (X_1(s, \xi), X_2(s, \xi))^T\), in the form
Figure C.2 -- Schematic illustration of the boundary conditions for the computation of the reflection response transform, for an impulse source, in the absence of a free surface.
\[
\frac{\partial X(s, \xi)}{\partial \xi} = \begin{bmatrix}
0 & -g(\xi) e^{-s\xi} \\
-g(\xi) e^{s\xi} & 0
\end{bmatrix} X(s, \xi) \tag{C.32}
\]

It has been shown in [22] that the solution of (C.32) may be written in the form

\[
X(s, \xi) = \phi(s, \xi) X(s, 0) \tag{C.33}
\]

The transition matrix, \( \phi(s, \xi) \), is

\[
\phi(s, \xi) = \begin{bmatrix}
Y_1(s, \xi) & Y_2(-s, \xi) \\
Y_2(s, \xi) & Y_1(-s, \xi)
\end{bmatrix} \tag{C.34}
\]

where

\[
Y_1(s, \xi) = \frac{1}{2} \left( R(s: \xi) + Q(s: \xi) \right) \tag{C.35a}
\]
\[
Y_2(s, \xi) = \frac{1}{2} \left( R(-s: \xi) - Q(-s: \xi) \right) \tag{C.35b}
\]

and

\[
R(s: \xi) = 1 - \int_0^\xi h_+ (\xi, \xi-\tau) e^{-s\tau} d\tau \tag{C.36a}
\]
\[
Q(s: \xi) = 1 - \int_0^\xi h_- (\xi, \xi-\tau) e^{-s\tau} d\tau \tag{C.36b}
\]

with

\[
h_+(x+\tau, x) = \sum_{m=0}^{\infty} (-1)^m f_m(x, \tau) \tag{C.37a}
\]
The $f_m(x, \tau)$ are given by (C.7). For example, applying the boundary conditions for the unperturbed problem, we have

$$
\begin{bmatrix}
X_1(s, \xi) \\
X_2(s, \xi)
\end{bmatrix} =
\begin{bmatrix}
Y_1(s, \xi) & Y_2(-s, \xi) \\
Y_2(s, \xi) & Y_1(-s, \xi)
\end{bmatrix}
\begin{bmatrix}
X_1(s, 0) \\
1
\end{bmatrix},
$$

which implies that

$$
X_1(s, \xi) = Y_1(s, \xi)X_1(s, 0) + Y_2(-s, \xi) \tag{C.38a}
$$

$$
X_2(s, \xi) = Y_2(s, \xi)X_1(s, 0) + Y_1(-s, \xi) \tag{C.38b}
$$

This relates the reflection response of the unperturbed medium, $X_1(s, 0) = P_2(s, 0)$, to the pressure field components at an arbitrary depth.

Now we perturb $g(\xi)$ by a small amount in the neighborhood of $\xi_0$, so that the reflection coefficient density is $g(\xi) + \delta g(\xi)$. Writing the propagation equations for the perturbed system, and subtracting those of the unperturbed, we find, to first order,

$$
\frac{s}{2} \delta P_1(s, \xi) = -\frac{\partial \delta P_1(s, \xi)}{\partial \xi} + g(\xi)(\delta P_1(s, \xi) - \delta P_2(s, \xi)) + \delta g(\xi)(P_1(s, \xi) - P_2(s, \xi)) \tag{C.39a}
$$
\[
\frac{s}{2} \delta P_2(s, \xi) = \frac{\partial \delta P_2(s, \xi)}{\partial \xi} + g(\xi)(\delta P_1(s, \xi) - \delta P_2(s, \xi)) \\
+ \delta g(\xi)(P_1(s, \xi) - P_2(s, \xi)) \tag{C.39b}
\]

Introducing transformations for the perturbed pressure field components

\[
\delta X_1(s, \xi) = w^{-1}(\xi)e^{-s\xi/2}\delta P_2(s, \xi) \tag{C.40a}
\]

\[
\delta X_2(s, \xi) = w^{-1}(\xi)e^{s\xi/2}\delta P_1(s, \xi) \tag{C.40b}
\]

we may write (C.39) in the combined matrix form, using

\[
\delta X(s, \xi) = (\delta X_1(s, \xi), \delta X_2(s, \xi))^T,
\]

\[
\frac{\partial \delta X(s, \xi)}{\partial \xi} = \begin{bmatrix}
0 & -g(\xi)e^{-s\xi} \\
-g(\xi)e^{s\xi} & 0
\end{bmatrix} \delta X(s, \xi) \\
+ \delta g(\xi)w^{-1}(\xi) \begin{bmatrix}
-e^{-s\xi/2} \\
e^{s\xi/2}
\end{bmatrix} (P_1(s, \xi) - P_2(s, \xi)) \tag{C.41}
\]

This may be interpreted as an equation for propagation in a system with distributed excitation. Note that \(\delta X_2(s, T) = 0\), and \(\delta X_1(s, 0) = 0\), so that the inhomogeneous term is the only source in (C.41). We solve (C.41) in a
simple manner by recalling that the support of $\delta g(\xi), \Delta \xi$, was to be a small interval containing $\xi_0$, so that we may divide the medium into three regions, as shown in Figure C.3.

In regions 1 and 3, $\delta g(\xi) = 0$, and the propagation equation (C.41) is homogeneous; in region 2, the full inhomogeneous equation is appropriate. Without taking the limit, however, we allow the thickness of the perturbed region, $\Delta \xi$, to become small, so that (C.41) may be written approximately as

$$\frac{\partial \delta \mathbf{X}(s, \xi)}{\partial \xi} = \begin{bmatrix} 0 & -g(\xi_0)e^{-s\xi_0} \\ -g(\xi_0)e^{s\xi_0} & 0 \end{bmatrix} \delta \mathbf{X}(s, \xi)$$

$$+ \delta g(\xi_0)W^{-1}(\xi_0) \begin{bmatrix} -s\xi_0/2 \\ -e \\ e^{s\xi_0/2} \end{bmatrix} (P_1(s, \xi_0) - P_2(s, \xi_0))$$

(C.42)

In writing (C.42), we have implicitly used the continuity of $g(\xi)$ and the unperturbed pressure field components as functions of $\xi$. In addition, the interval is chosen so that it does not contain a point of discontinuity of $g(\xi)$ in its interior. Equation (C.42) is a constant coefficient
Figure C.3 -- Region definition for solution of the perturbed reflection response of a medium.
differential equation in $\xi$, which we may use to relate
$\delta X(s, \xi^+)$, the state vector value above the perturbed region,
to $\delta X(s, \xi^-)$, the value below. The variation of constants
formula may now be employed, and gives

$$\delta X(s, \xi^-) = \tilde{\phi}(s, \Delta \xi) \delta X(s, \xi^-)$$

$$+ \int_0^{\Delta \xi} \tilde{\phi}(s, \eta) \delta g(\xi_0) \tilde{W}^{-1}(\xi_0) \left[ \begin{array}{c} -s \xi_0/2 \\ -e \\ s \xi_0/2 \\ e \end{array} \right] (P_1(s, \xi_0) - P_2(s, \xi_0)) d\eta$$

(C.43)

Note that we have assumed that the perturbation may
be approximated by a constant over its region of support.
While this is a somewhat restrictive assumption, it suf-
fices for our purposes; in particular, it suffices for the
derivation of (5.33), carried out in Appendix B, which is
our primary concern.

For small $\Delta \xi$, $\tilde{\phi}(s, \Delta \xi) = I$, the identity matrix, so

$$\delta X(s, \xi^+) = \delta X(s, \xi^-) + \Delta \xi \delta g(\xi_0) \tilde{W}^{-1}(\xi_0)$$

$$\times (P_1(s, \xi_0) - P_2(s, \xi_0))$$

(C.44)
This equation relates $\delta X(s, \xi)$ above and below a very thin region of perturbation, and may be viewed as modelling the point injection of signals into the medium, as illustrated in Figure C.4. We may solve this system in regions 1 and 3 using (C.33), and match across the regions using (C.44).

We find that

$$\begin{align*}
\frac{3\xi_0}{2} & - \frac{s_1}{2} - \frac{s_2}{2} \\
\left(\delta X_2(s, \xi_0) + e^{\frac{s_1}{2} U(s)} V(s) e^{\frac{s_2}{2}} U(s) e^{\frac{s_3}{2}}ight) = \delta X_1(s, \xi_0)
\end{align*}$$

(C.45)

where

$$U(s) = \Delta \delta g(\xi_0) W^{-1}(\xi_0) (P_1(s, \xi_0) - P_2(s, \xi_0))$$

and $V(s)$ is the pressure wave impulse reflection response transfer function of region 2, subject to the given boundary conditions. Using (C.33) we can write

$$\begin{bmatrix}
\delta X_1(s, \xi_0) \\
\delta X_2(s, \xi_0)
\end{bmatrix} =
\begin{bmatrix}
Y_1(s, \xi_0) & Y_2(-s, \xi_0) \\
Y_2(s, \xi_0) & Y_1(-s, \xi_0)
\end{bmatrix}
\begin{bmatrix}
\delta X_1(s, \xi_0) \\
\delta X_2(s, \xi_0)
\end{bmatrix}
\begin{bmatrix}
\delta X_1(s, \xi_0) \\
\delta X_2(s, \xi_0)
\end{bmatrix} = Y_1(s, \xi_0) \delta X_1(s, 0) = Y_1(s, \xi_0) \delta P_2(s, 0)
\delta X_2(s, \xi_0) = Y_2(s, \xi_0) \delta X_1(s, 0) = Y_2(s, \xi_0) P_2(s, 0)$$
\[ \delta P_2(s, 0) = \delta X_1(s, 0) \quad \delta X_2(s, 0) = 0 \]

\[ \xi = 0 \]

Region 1

\[ \xi = \xi_0 \]

Region 3

\[ \delta X_1(s, T) = 0 \]

Figure C.4 -- Schematic medium model to be solved to obtain the perturbation reflection response of the medium.
Solving (C.45), we have

\[ \delta P_2(s, 0) = \frac{V(s)e^{-s\xi_0/2} + e^{-s\xi_0/2}}{Y_1(s, \xi_0) - Y_2(s, \xi_0)V(s)e^{-s\xi_0}} \quad (C.46) \]

We now need an expression for \( P_1(s, \xi_0) - P_2(s, \xi_0) \) for the unperturbed system; this can easily be found using the same solution method as above. Doing this, and dividing through by \( \Delta\xi\delta g(\xi_0) \), we have

\[ \frac{\delta P_2(s, 0)}{\Delta\xi\delta g(\xi_0)} = e^{-s\xi_0/2} \theta \frac{2(V(s) + 1)}{(R(s, \xi_0) + Q(s, \xi_0)) - (R(-s, \xi_0) - Q(-s, \xi_0))e^{-s\xi_0}} \]

\[ \frac{-s\xi_0/2}{(R(s, \xi_0)Q(-s, \xi_0) + Q(s, T)R(s, \xi_0)) + e^{s\xi_0/2}} \frac{e^{s\xi_0/2}}{(Q(s, \xi_0)R(s, T) - R(s, \xi_0)Q(s, T))} \]

\[ \times \frac{R(s, T) + Q(s, T)}{R(s, T) + Q(s, T)} \quad (C.47) \]

If we take the limit of (C.47) as both the support of the perturbation and its amplitude go to zero, we have the strong-sense definition of the variational derivative in the frequency domain, that is, (5.32) expressed as a ratio rather than in operator form. The right-hand side of (C.47) is independent of \( \Delta\xi\delta g(\xi_0) \), so that the limit as \( \Delta\xi\delta g(\xi_0) \to 0 \) exists, and also, in the limit (C.47) becomes exact, so that the right-hand side of (C.47) is
\[ \delta G(f, \xi_0) / \delta g. \] The variational derivative exists, and, since the right-hand side of (C.47) is a continuous function of continuous functionals of \( g(\xi) \), the variational derivative is continuous, and we are done.

In this Section, we have shown both that the variational derivative of the reflection response of a medium exists in the strong sense (so that the integral in (5.42) makes sense), and that it is a continuous functional. We note that we have applied sufficiently strong conditions on the functional, and on \( g(\xi) \), that the expectation in (5.42) is also well-defined. Our achievement has been, however, more theoretical than practical, because the complexity of (C.47) (recall the definitions of its components) makes it unlikely that the variational derivative could be obtained using this formulation. We also note that our derivation is valid only for a class of perturbations (constant perturbations) which are adequate for our purposes, i.e., the proof of (5.33).
REFERENCES


15. "Bright spot in the energy picture", Ocean Industry, v. 9, no. 2, pp. 60-65 (February, 1974).


Kenneth B. Theriault was born in Boston, Massachusetts, on 9 August, 1946. He received the Bachelor of Science in Electrical Engineering from the Massachusetts Institute of Technology in June, 1968. He was subsequently employed as a systems engineer by the C.S. Draper Laboratory until April, 1969, when he entered Navy Officer Candidate School. Following commissioning, and various schools, he served aboard USS HARRY E. YARNELL (CG-17) until his release from active duty in May, 1972. In June, 1972, he returned to the Massachusetts Institute of Technology as a graduate student in the Department of Electrical Engineering and Computer Science. He received the degrees of Master of Science and Electrical Engineer in February, 1975. In January, 1977, he was admitted to the M.I.T./W.H.O.I. Joint Program in Oceanographic Engineering, in which program he completed his doctoral dissertation.

Dr. Theriault is employed by Bolt, Beranek, and Newman, Inc., of Cambridge, Massachusetts. He is a member of Eta Kappa Nu, Tau Beta Pi, Sigma Xi, the Institute of Electrical and Electronic Engineers, and the United States Naval Institute.

He is married to the former Marcie E. Abramson of Appleton, Wisconsin; they have two children, Seth and Emily.