A MULTIVARIATE APPROACH
TO THE ANALYSIS OF
GEOMAGNETIC ARRAY DATA

by

Gary D. Egbert

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Geophysics Program

University of Washington

April 1987

University of Washington

Abstract

A MULTIVARIATE APPROACH TO THE ANALYSIS OF GEOMAGNETIC ARRAY DATA

by Gary David Egbert

Chairperson of the Supervisory Committee: Professor John R. Booker

Graduate Program in Geophysics

In this dissertation we develop new methods for the multiple station analysis of geomagnetic array data. The methods are based on the observation that the space of external sources, and hence the space of observable electromagnetic fields at the earth's surface (the 'response space'), can be well approximated by spaces of very low dimension p. We show that the simplifying assumption of a finite dimensional external source space allows for a rigorous justification of transfer function methods, and we demonstrate that the response space is equivalent to all possible interstation and intercomponent transfer functions. The response space can be estimated from the p dominant eigenvectors of the spectral density matrix, the matrix of averaged cross products of Fourier coefficients for all components in the array. We also develop a simple stochastic model for random sources. This model, and results from a number of 5 station arrays, show that for arrays which are small compared to typical source scales that the sources can be approximated as plane waves ($p = 2$); a first order correction to this approximation allows for three source gradient terms ($p = 5$). We also develop statistical methods for the estimation of the response space which allow for (potentially correlated) noise in all measured field components. For the plane wave source case, the covariance structure for the noise (including source effects) can be parametrized in a way which allows for the separation of coherent source "noise" from the desired signal. In statistical terms, the model is a (complex) multivariate errors-in-variables model

with a generalized, but identifiable error covariance parametrization. Applications of these methods to a series of small arrays reveal that much of the misfit to the plane wave source model is highly coherent over hundreds of kilometers and can be ascribed to violations of the uniform source field assumption. In particular, we show that source gradient effects on electric field measurements in 2- and 3-dimensional situations can be significant. We also briefly consider methods for combining a series of small overlapping arrays to synthesize the plane wave response for a large region.

Table of Contents

List of Figures

Acknowledgments

I would first like to take this opportunity to thank the faculty in the Geophysics Program for their inspiration and assistance; they, in general, convinced me that geophysics was "for me". More particularly, I would like to thank Professor John Booker, my thesis advisor, for his support throughout this endeavor, for providing me with the opportunity to get fully involved in electromagnetic induction studies, and for reading this manuscript through with care. Professor Ronald Merrill, a fellow denizen of the coffee shops, was always willing to talk about geomagnetism, geophysics, science or life in general, and gave enough advice to qualify as a co-advisor. Professor Marcia Baker, offered me financial support when I first entered the program, and provided a more general sense of support until the end. I would also like to thank J. Torquil Smith for many hours of shared discussions, and arguments, about geomagnetic induction, inverse theory, statistics and whatever else came up. Finally, I would like to acknowledge my family, particularly my wife, Mina Ossiander, for the support and patience she has offered (especially through the last few months of writing), but also my son Carl, who did not exactly help with the completion of this dissertation, but at least made me happier at times while I was writing it.

Chapter 1
Introduction

Electric currents in the earth's ionosphere and magnetosphere give rise to time varying magnetic fields which induce electric currents in the earth. The spatial pattern and frequency dependence of the resulting induced currents depends both on the temporal and spatial distribution of external sources, and on the spatial variations of electrical conductivity within the earth. Measurements of the magnetic (and electric) fields at the earth's surface can thus be used to study both the external sources and the electrical conductivity of the earth. Arrays of magnetometers have been used for both of these purposes (e.g. Schmuker, 1964; Porath and Gough, 1971; Banister and Gough, 1977, 1978; Kupper, *et al.*, 1979; Lilley *et al.*, 1981). The focus in this dissertation is on the latter purpose - we consider some aspects of the problem of extracting information about the electrical conductivity within the earth from measurements of time varying electromagnetic fields at the surface. Because the data is the sum of external and internal components it is, in fact, impossible to completely ignore the nature of the external sources. Indeed, much of this thesis is devoted to studying the way in which the external sources effect the geological interpretation of the data. Nonetheless, our treatment of the external sources is extremely schematic, and our methods are motivated by a desire to eliminate, or at least reduce, the complications in the external sources so that the interpretation of the data in terms of variations in electrical conductivity of the earth will be clearer.

As with most physical properties of the earth, the principal variation of electrical conductivity is radial. The conductivity increases steeply in a depth range from 400-800 km to values in excess of 1 S m^{-1} (Banks, 1969; Schultz and Larsen, 1987). This high mid-mantel conductivity limits the diffusion of

externally induced electromagnetic fields to the crust and upper mantle. Deeper penetrations are possible with longer period variations, but periods of a few weeks or more are required to penetrate to mid-mantle depths. As a result, information about electrical conductivity obtained from most geomagnetic array studies is limited to the crust and upper mantle.

There are also substantial lateral variations of conductivity, and these are usually the focus of geomagnetic array studies. Such lateral conductivity contrasts are particularly pronounced in the crust, where several orders of magnitude of variation in electrical conductivity are not uncommon. The most significant surface conductivity contrast occurs at the coast where the highly conductive oceans abut the relatively resistive crust (Parkinson, 1962; Schmuker, 1964; Lilley and Bennett, 1972). In addition significant variations in crustal conductivity occur. Numerous major crustal conductivity 'anomalies' have been mapped by geomagnetic induction methods (e.g. Vozoff and Swift, 1968; Camfield *et al.*, 1971; Porath and Dziewonski, 1971; Babour and Mosnier, 1979; Law *et al.*, 1980; Banks and Beamish, 1985). Some of these crustal anomalies have been related to geothermal activity (Ingham, *et al.*, 1983). Others have been interpreted as suture zones or old plate boundaries (e.g. Camfield and Gough, 1977; Stanely *et al.*, 1987); electromagnetic induction methods may thus be useful for studying the tectonic history of a region.

Induction methods are also useful in the study of ongoing tectonic processes. Significant conductivity anomalies in the lower crust and/or upper mantle have also been mapped in continental rift zones (e.g. Schmuker, 1964, 1970; Porath and Gough, 1971; Banks and Ottey, 1974; Hermance, 1982). In a similar extensional setting, variations in the depth to a layer of high conductivity in the uppermost mantle (Porath *et al.*, 1970) beneath the basin and range have been correlated with seismic parameters and heat flow (Gough, 1974). In an application to a different

sort of tectonic problem Helferty *et al.* (1987) have used geomagnetic induction methods to locate the edge of the subducting Gorda plate in Northern California. At greater depths, Schultz and Larsen (1987) have inferred variations in mid-mantle conductivity from their analysis of very long period magnetic observatory data. Such long period global studies may ultimately provide significant new constraints on the physical properties of the upper mantle. This extremely brief review is meant only to give the reader some idea of the sort of variations in conductivity that have been detected with geomagnetic induction methods, and to indicate in a very general way the relevance of these observations to geological and geophysical problems. More thorough reviews are given, for example, by Gough (1974, 1983), and Hermance (1982, 1983).

This dissertation is about the analysis of data from geomagnetic arrays. By a geomagnetic array we mean a set of simultaneously operating instruments which measure the magnetic and/or electric fields at n sites. The geomagnetic arrays referred to above are quite variable in size - both in spatial extent and in the number of instruments involved. At the largest spatial scales we have the globally distributed geomagnetic observatories; at the smallest we may have a maximum station separation on the order of kilometers or less (a single instrument, or an instrument with a nearby remote reference, may be usefully considered as an array for that matter, so the smallest spatial scale is essentially zero). The largest regional arrays cover several thousand kilometers (Camfield *et al.*, 1970; Porath and Gough, 1971; The EMSLAB Group, 1987). Most of the results referred to above are from regional arrays of an intermediate size - arrays whose size is on the order of hundreds of kilometers. Alabi (1983) reviews many of the more significant regional arrays.

The number of stations in an array is also quite variable. The smallest arrays may consist only of a few instruments (e.g. Schmuker, 1964; 1970; Law *et al.*,

1980; Banks and Beamish; 1985), while the more or less continually operating global observatory array consists of several hundred (independently operated, variable quality) stations. For regional arrays, the largest array to date (as far as number of instruments) is the EMSLAB array (Booker *et al.*, 1987). This array included stations on land and on the seafloor and consisted of 76 sites with magnetic field measurements only and an additional 27 magnetotelluric (MT) sites with both magnetic and electric field measurements. We will consider data from five of the land based MT stations in this dissertation. Previous large regional arrays have included as many as 46 (magnetic fields only) stations (Porath and Gough, 1971).

In fact all of the arrays that we consider in this dissertation consist of 5 or fewer stations. These arrays are also all of relatively small spatial extent - array sizes are on the order of 50 - 200 kilometers. While many details of our array analysis techniques have been strongly effected by the size of the arrays that we have worked with, we believe that many of the techniques described here will be, with suitable modification, relevant to geomagnetic arrays in general. In particular, the methods described in this dissertation not depend in any way on the number of stations being small (although the way we have numerically implemented the procedures may be usefully improved for larger arrays). For arrays of much greater spatial extent the methods developed in this dissertations will require significant modifications. Nonetheless, we believe that the general perspective developed here will be quite relevant to the case of spatially larger arrays, including the global array of geomagnetic observatories. For these reasons we have kept much of our discussion as general as possible.

Due to the relatively limited number of instruments, all geomagnetic arrays are either of limited spatial extent and/or are sparse. For large arrays (i.e. the global observatory array (Chapman and Bartels, 1940; Berdichevsky and Zhdanov,

1984), or the very largest regional arrays (Porath, *et al.*,1970; Richmond and Baumjohann, 1983)) a separation of the variations into internal and external parts is possible, at least in theory. For arrays which are smaller, either in number or in spatial extent, such a separation is not feasible. In fact, Gough (1973), based on his experience with several large regional arrays in North America, has questioned the value of a formal separation of the fields in any regional arrays. Even for the largest such arrays the spatial scales of the external sources are much larger than the array and the internal/external separation is problematic (Gough, 1973; Berdichevsky and Zhdanov, 1984). In general, then, the quantitative interpretation of array data requires some simplifying assumptions about the spatial structure of the external sources. This is particularly true for the very small (5 station) arrays which we analyze here.

In regional array studies, where array dimensions are smaller than typical source scales, it is usually assumed that the external source fields can be approximated by plane waves of infinite horizontal extent. If this assumption holds exactly and measurements are made without error, it is easy to show that all field components can be related by frequency domain linear transfer functions to the two horizontal components of the magnetic fields at any chosen reference station. These transfer functions indirectly define the response of the earth to plane wave sources and can be used to make inferences about conductivity in the earth. The transfer function approach was originally developed for the case of very sparse arrays, where measurements are made simultaneously at only a few stations, often only one (e.g. Schmuker, 1964). The transfer function method has, however, become a standard tool in the quantitative interpretation of much larger (i.e. denser) arrays (e.g. Beamish, 1977; Alabi 1983; Gough and Ingham, 1983; Ingham *et al.*, 1983). The methods of data analysis developed in this thesis are essentially a generalization of the standard transfer function methods. We will show that our

approach is more reasonable both physically and statistically.

The usual interstation and intercomponent transfer functions have typically been estimated, one at a time, from the Fourier transformed time series using least squares methods. To do this, it is assumed that the reference fields are measured without error, and that the discrepancy between predictions and observations due to the failure of source field assumptions can be treated as incoherent noise in the predicted field component. Clearly, neither assumption is reasonable. There will, in general be noise in all field measurements. As is well known (e.g. Jenkins and Watts, 1968), noise in the input channels causes biases in the transfer function estimates. This problem has received considerable attention in the context of magnetotelluric impedance estimation (Gamble *et al.*, 1979; Clarke *et al.*, 1983; Pedersen and Svennekjaer, 1984) where the seriousness of this problem has led to the development of the remote reference technique. The treatment of the effect of violations of source assumptions as incoherent noise is potentially even more dangerous. Since the failure of these simplifying assumptions can result in coherent and systematic errors, an uncritical application of standard statistical methods can lead to extremely misleading results.

The data processing methods developed in this dissertation allow us to confront both of these problems in a much more natural fashion. Our approach, which is based on a statistical analysis of the eigenvalues and eigenvectors of the spectral density matrix (SDM), is a true multivariate statistical technique which allows for (potentially correlated) noise in all measured field components. In addition to providing a more physically reasonable statistical model for transfer function estimation, our approach has also proved extremely useful as a tool for exploring the ways in which violations of the simplifying plane wave source assumption effect the data. These results about source effects while interesting in their own right, will allow us to formulate a more physically reasonable statistical

model for the structure of the 'noise' (including source effects). This more reasonable model in turn, can be used to improve estimates of transfer functions.

In Chapters 2-4 we develop these ideas in detail. We illustrate the ideas with results obtained from a series of small five station arrays. Many details of our methods are geared to this specific situation, but wherever possible we have kept the treatment fairly general, and hence slightly abstract.

Chapter two develops the physical background and the basic idea behind our approach to geomagnetic array analysis. This chapter provides a rigorous connection between the physical situation and the statistical model that we will use. In the actual application we have in mind, the external magnetic source fields can be approximated as plane waves of infinite horizontal extent. We treat this as a special case of the general situation where the source fields are assumed to lie in a finite dimensional space, and we show how this assumption forms the basis for a generalized transfer function approach. No significant discussion of the statistical model is given in Chapter 2; rather we seek to heuristically justify the use of the dominant eigenvectors (i.e. principal components) of the SDM as estimates of the response of the earth to simple (e.g. plane wave) external source fields. The chapter closes with some examples of some exploratory data analysis for a number of small arrays. In addition to demonstrating the use of the eigenvectors of the SDM as estimators of the plane wave response of the earth we explore the structure of the 'noise' in the data. We find that, in fact, much of the 'noise' is highly coherent across the array and can be ascribed to violations of the uniform source field assumption.

Chapter 3 develops a stochastic model for random external sources. This model is then used to study the properties of the the eigenvalues and eigenvectors of the SDM's for several synthetic arrays. In this chapter we make the connection between our approach and a sort of discrete spatial spectral analysis of the fields.

The results of this chapter provide a strong justification for the approach developed heuristically in Chapter 2.

In Chapter 4 we develop the statistical model. In statistical terms, the model is a (complex) multivariate errors-in-variables model with a generalized, but identifiable error covariance parametrization. Adapting results from the statistical literature in factor analysis and errors-in-variables regression, we develop asymptotic estimation errors and goodness of fit statistics for our model parameters. In addition, using results from Chapters 2 and 3, we develop a physically based parametrization for the error covariance structure which allows for violations of the plane wave source assumption. We demonstrate that, provided the signal to noise ratio is large enough, it is possible to uniquely separate the signal and noise components in this model. The model for the noise covariance matrix is linear; we have developed an iterative generalized least squares scheme for estimation of the parameters of this model. We discuss our approach and relate this to a similar estimation scheme proposed (for the case of real covariance matrices with linear structure) by Anderson (1969, 1970, 1973).

We apply the statistical techniques developed in Chapter 4 to four five station arrays - three 3 component (magnetic fields only) arrays and the 5 component magnetotelluric array form the EMSLAB experiment. One significant result of this analysis is that a substantial fraction of the misfit between local electric and magnetic fields in the EMSLAB array can be ascribed to source effects. In particular we find that there are electric fields, coherent across the array, which are coherent with gradients in the horizontal fields and with source vertical fields, which are not coherent with the local horizontal magnetic fields. This shows that in 2- and 3-dimensional environments finite wavenumber sources can have a first order effect on impedances. This effect is not predicted by an analysis of the simple one dimensional case, and indicates that the analysis of source effects given by

Dmitriev and Berdichevsky (1979) is overly optimistic. We also show that in some circumstances there are small current flows in conductive anomalies which are not coherent with the local magnetic fields (but which are coherent with local gradients). Although this effect is small it demonstrates that the extreme form of current channeling suggested by Babour and Mosnier (1979) does occur, at least to a small extent.

Finally in a very brief Chapter 5 we outline some extensions and areas for future work. In particular, we give a very brief indication of how our estimation scheme can be generalized to optimally combine a number of small overlapping arrays to produce a large synthetic array. We illustrate this technique by synthesizing a 40 station array from a series of 3 - 5 station GDS arrays run in southwestern Washington over the past seven years.

Chapter 2

The Response Space Approach to Geomagnetic Array Analysis

The goal of this chapter is to establish a firm theoretical foundation for the array analysis schemes which will be developed in this thesis. The methods described here are closely related to transfer function (TF) methods which have been frequently applied to the analysis of data from geomagnetic arrays both large and small (e.g. Schmuker, 1970; Frazer, 1974; Beamish, 1977; Alabi, 1982; Ingham *et al.* , 1983; Gough and Ingham, 1983). We will thus begin with a fairly abstract and general treatment of the TF approach in geomagnetic induction studies. We will show that the TF method can be justified with complete rigor provided that the set of all external current sources used in an analysis are restricted to a fixed finite dimensional space. This development leads quite naturally to the general concept of the 'response space', which effectively defines the response of the earth to all possible source configurations in this restricted finite dimensional space. This response space will be seen to be equivalent to, but in many ways more fundamental than, the usual intercomponent and interstation transfer functions.

The response space formulation will be used throughout this thesis. This formulation unifies and simplifies the various transfer functions which have been used in the interpretation of geomagnetic induction data. More importantly, form the standpoint of this thesis, this formulation leads to new methods of TF estimation. These estimates are based on the eigenvalues and eigenvectors of the full spectral density matrix (SDM) (i.e. the matrix of the auto and cross spectral powers of all channels at all stations in the array). In contrast to the standard multiple input single output model used in the classical TF estimation approach (Jenkins and Watts, 1968), the statistical model which forms the basis for our

estimation scheme allows for noise in all channels. This is certainly more realistic than the standard TF model which assumes that there is noise only in the predicted, or output, channel. Furthermore, our model for noise can be generalized to allow for noise which is correlated between channels (which we shall see is commonly the case with geomagnetic induction data). We formulate this statistical model in this chapter; the estimation of the model parameters and computation of error estimates are discussed more fully in Chapter 4.

We will finish this chapter by examining the eigenvalues and eigenvectors (i.e. the principal components) of the spectral density matrices computed from several five station 3-component magnetovariational (MV) arrays from southwestern Washington, and a five station, 5-component magnetotelluric (MT) array from the EMSLAB long period MT profile (The EMSLAB Group, 1987). We will show that the data from these small arrays is largely consistent with the plane wave source assumption which justifies the standard TF approach. Using this data, we will demonstrate the computation of the appropriate plane wave response space. In addition, we shall demonstrate how a more careful study of the eigenvectors and eigenvalues of the SDM can reveal information about the structure of the noise. We will see that at periods greater than a few hundred seconds the noise is mostly coherent across the array and can be attributed to the violation of the simplifying plane wave source assumption.

Before proceeding some general comments on the time series aspects of this problem are in order. The raw data which is collected in a geomagnetic array consists of a number of simultaneously observed time series which are (analogue) low pass filtered and sampled at interval ΔT. Although natural source geomagnetic data has been interpreted with time domain techniques (Babour and Mosnier, 1979; McMechan and Barrodale, 1985) it is more common to use frequency domain techniques (see Beamish (1983) for a comparison; Jones (1983) for a

critique of some shortcomings of time domain methods). Our approach is of the latter sort. While the transformation between the raw time series and the frequency domain Fourier coefficients requires some care (cf. Thomson *et al.* 1976; Thomson, 1982), we will not discuss these aspects of the data analysis problem here. The time series processing we have used is essentially the same as that outlined in Egbert and Booker (1986). In very brief outline, we take very long ($N \approx 10^5$) time series for each channel, pre-whiten and window the time series with a series of relatively short overlapping pi-prolate spheroidal windows (Thomson 1982), and Fourier transform the resulting data segments. Longer time windows are used for lower frequencies, and we use a cascade decimation type scheme to keep the number of points which must be transformed small. This procedure results in a series of Fourier coefficients for each frequency. In this thesis we will, for the most part, treat this sequence of Fourier coefficients as if they were the raw data. When we refer to "the data", this is what we will mean, unless we explicitly state otherwise.

2.1: Physical Background

Assuming a time dependence of $e^{i\omega t}$ we may write the quasi-static approximations to Maxwell's equations in a region of arbitrary conductivity $\sigma(\mathbf{x})$ as (Rokityansky, 1982)

$$\nabla \times \mathbf{B} = \mu\sigma\mathbf{E} + \mathbf{J}_{ext} \tag{2.1.1a}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.1.1b}$$

$$\nabla \times \mathbf{E} = -i\omega\mathbf{B} \tag{2.1.1c}$$

$$\nabla \cdot \mathbf{E} = q/\varepsilon \tag{2.1.1d}$$

where **B** and **E** represent the magnetic and electric fields respectively, μ is the magnetic permeability and ε the dielectric constant, q is the electric charge density and $\mathbf{J}_{ext}$ represents external (non-induced) current sources. Note that with the quasi-static approximation we have omitted displacement currents (Maxwell's contribution to the equations which have come to bear his name); this approximation can be justified by standard scaling arguments (c.f. Rokityansky, 1982) for the frequencies, length scales and conductivities encountered in geomagnetic induction studies.

Here we will treat the surface of the earth as a plane of infinite horizontal extent. The region $z < 0$ represents the exterior of the earth and the region $z > 0$ the interior. The basic results described can be generalized to other geometries (such as a spherical geometry) when appropriate. In the interior region we assume that $\mathbf{J}_{ext} = 0$. This means that currents in the interior of the earth are induced ohmic ($\mathbf{J} = \sigma\mathbf{E}$) currents only. In particular, this formulation does not allow for fields produced in the earths core due to the hydromagnetic dynamo. For the frequency ranges we are interested in (up to a few hours) these internal fields are negligible due to the shielding effect of the (relatively) highly conducting lower mantle (e.g. Merrill and McElhinny, 1983). We also assume that in the exterior region the conductivity is zero. Again, this approximation is well justified for the frequencies and earth conductivities of interest to us (Rokityansky, 1982).

The physical picture here is that the currents $\mathbf{J}_{ext}$ in the ionosphere and magnetosphere (the exterior region) are sources which induce electromagnetic fields inside the earth (the interior region). In the absence of such external currents the fields are zero everywhere. This can be made precise via standard existence and uniqueness results for Maxwell's equations (c.f. Berdichevsky and Zhdanov, 1984). Specifically, for $\mathbf{J}_{ext}$ restricted to a bounded region and with the magnetic fields satisfying the boundary condition $\mathbf{B} \rightarrow 0$ as $|x| \rightarrow \infty$ there is a unique

solution **B**, **E** to the system of equations (2.1.1a-d). The uniqueness of the solution implies that the fields seen at the surface of the earth ($z = 0$) are determined by the configuration of external currents. In fact we can make a somewhat stronger statement: The solution in the interior and on the surface is determined uniquely by the fields at the surface which are due to external sources.

With $\sigma = 0$ in the exterior region ($z < 0$) we have $\nabla \times \mathbf{B} = 0$. Together with (2.1.1b) this implies that **B** can be written as the gradient of a magneto-static scalar potential

$$\mathbf{B} = \nabla\phi \qquad \text{where } \nabla^2\phi = 0$$

Using standard results (e.g. Schmuker, 1970; Berdichevsky and Zhdanov, 1984) we can write the potential ϕ as the sum of internal and external potentials

$$\phi = \phi_i + \phi_e$$

where

$$\phi_i \rightarrow 0 \quad \text{as } z \rightarrow -\infty$$

$$\phi_e \rightarrow 0 \quad \text{as } z \rightarrow +\infty$$

Note that the magnetic fields can be obtained from these potential functions only in free space in a region near the surface $z = 0$ (e.g. for z satisfying $0 > z > -z_0$) where we have both $\sigma = 0$ and $\mathbf{J}_{ext} = 0$.

We will now show that the internal potential ϕ_i is determined uniquely (up to a constant, of course) by the external potential ϕ_e. This is quite obvious from a physical standpoint - with our assumption that the currents in the earths interior are all induced by the time varying magnetic fields it is clear that the internal fields must be caused by the external fields. Furthermore, it is intuitively obvious that the exact distribution of currents is not important. The internal fields will depend only on the pattern of external fields impinging on the surface of the earth.

The proof of this is straightforward. Let $\mathbf{J}_1$ and $\mathbf{J}_2$ be two external current systems and let ϕ_{i1}, ϕ_{e1}, ϕ_{i2}, and ϕ_{e2} be the corresponding internal and external potentials (defined for $0 > z > -z_0$). We wish to show that if

$$\phi_{e1} = \phi_{e2} \tag{2.1.2}$$

then

$$\phi_{i1} = \phi_{i2} \tag{2.1.3}$$

Suppose, then, that (2.1.2) holds and let

$$\mathbf{J}_\Delta = \mathbf{J}_1 - \mathbf{J}_2$$

$$\phi_{i\Delta} = \phi_{i1} - \phi_{i2}$$

$$\phi_{e\Delta} = \phi_{e1} - \phi_{e2} = 0$$

The linearity of equations (2.1.1a-d) implies that for external currents $\mathbf{J}_\Delta$ the internal and external potentials at $z = 0$ will be $\phi_{i\Delta}$ and $\phi_{e\Delta}$ respectively. Let $\mathbf{B}_\Delta$ be the magnetic fields for the solution to (2.1.1) with current source $\mathbf{J}_\Delta$. In the region $0 > z > z_0$ we may write

$$\mathbf{B}_\Delta = \nabla\phi_\Delta = \nabla\phi_{i\Delta} \tag{2.1.4}$$

where

$$\phi_{i\Delta} \to 0 \quad \text{as} \quad z \to -\infty \tag{2.1.5}$$

Define

$$\mathbf{B}' = \begin{cases} \mathbf{B}_\Delta & z > 0 \\ \nabla\phi_\Delta & z \le 0 \end{cases} \tag{2.1.6}$$

By definition $\mathbf{B}'$ satisfies equations (2.1.1a-d) for $z > 0$; by (2.1.4) this holds for the larger region $z > -z_0$. In the exterior region, $z < 0$, the gradient of any scalar

potential (and hence $\mathbf{B}'$) satisfies the system (2.1.1) with $\mathbf{J}_{ext} = 0$. Finally, note that by (2.1.5) $\mathbf{B}'$ satisfies the boundary conditions $\mathbf{B}' \to 0$ as $|z| \to \infty$. Hence, $\mathbf{B}'$ is a solution to the system (2.1.1) with $\mathbf{J}_{ext} = 0$. Since $\mathbf{B} = 0$ satisfies (2.1.1) with no external currents, we conclude from the uniqueness properties of solutions to (2.1.1) that

$$\mathbf{B}' \equiv 0$$

Hence we see that

$$\phi_{i1} = \phi_{i2}$$

(modulo an additive constant) and we are done.

With this result we may write the internal potential as a function of the external potential:

$$\phi_i = L_{\sigma,\omega}(\phi_e)$$

where the operator $L_{\sigma,\omega}$ depends on the frequency ω and the conductivity distribution σ. From the linearity of Maxwell's equations it is clear that this operator is linear.

As an explicit example, we consider the simple and well understood 1-dimensional (1-d) case where conductivity varies only with depth $\sigma(x,y,z) = \sigma(z)$. For the 1-d case the operator $L_{\sigma,\omega}$ can be most easily expressed in the spatial wave number domain. In general we may write ϕ_i and ϕ_e as Fourier integrals (Schmuker, 1970):

$$\phi_e(\mathbf{x},z) = \frac{1}{4\pi^2} \iint d^2\nu\, \tilde{\phi}_e(\nu) \exp[i(\nu \cdot \mathbf{x}) - \nu z]$$

$$\phi_i(\mathbf{x},z) = \frac{1}{4\pi^2} \iint d^2\nu\, \tilde{\phi}_i(\nu) \exp[i(\nu \cdot \mathbf{x}) - \nu z]$$

where $\nu = |\nu| = \nu_x^2 + \nu_y^2$. The operator $L_{\sigma,\omega}$ can be written in the wavenumber domain as

$$\tilde{\phi}_i(\nu') = \iint d^2\nu\, \tilde{L}_{\sigma,\omega}(\nu,\nu')\, \tilde{\phi}_e(\nu') \qquad (2.1.7)$$

In the 1-d case the kernel of the operator in (2.1.7) has the simple diagonal form

$$\tilde{L}_{\sigma,\omega}(\nu,\nu') = \delta(\nu - \nu')\beta(\omega,\nu) \qquad (2.1.8)$$

where $\beta(\omega,\nu)$ can be written in terms of the wavenumber dependent impedance $\zeta(\omega,\nu)$ (Rokityansky, 1982):

$$\beta(\omega,\nu) = \frac{\omega - i\nu\zeta(\omega,\nu)}{\omega + i\nu\zeta(\omega,\nu)} \qquad (2.1.9)$$

In the 1-d case then, we see that $L_{\sigma,\omega}$ is equivalent to the most general 1-d response function, the wavenumber dependent impedance. More generally, it can be seen that all information about electrical conductivity σ which can be obtained from measurements on the earth's surface is contained in the operators $L_{\sigma,\omega}$ as ω varies.

The relationship between the data (observations of time varying magnetic and electric fields on the earths surface) and the conductivity in the interior of the earth σ can be expressed abstractly by the mappings:

$$\sigma \rightarrow L_{\sigma,\omega} \rightarrow data$$

The quantitative interpretation of geomagnetic data requires inverting the above mappings from a finite number of observations. This task can be broken into the same two steps.

$$data \rightarrow L_{\sigma,\omega} \rightarrow \sigma$$

The work described in this thesis is concerned with the first step of the data interpretation process. Specifically we study ways to best use measurements obtained from small arrays of simultaneously operating instruments to obtain information about the response of the earth, represented completely, and somewhat abstractly here by the operator $L_{\sigma,\omega}$.

2.2: A Rigorous Justification of Transfer Functions

The formulation of the geomagnetic data interpretation problem given above is completely general and will be useful to us in the sequel. However, this general formulation is not, by itself, particularly practical. $L_{\sigma,\omega}$ defines the response of the earth to all possible external source potentials, and is thus an operator defined on an infinite dimensional function space. The computation of $L_{\sigma,\omega}$ would thus require in general the numerical solution of the system of equations (2.1.1) for a large number of possible source configurations (literally an infinite number). As a result $L_{\sigma,\omega}$ cannot be practically computed except in the simple 1-d case where analytic solutions are possible (i.e (2.1.7 - 2.1.9)). At the same time, the direct estimation of $L_{\sigma,\omega}$ from the sort of data which are typically available is not practical. Measurements are made simultaneously at a relatively small number of stations which are sparse and/or of limited spatial extent. In order to interpret natural source geomagnetic data some simplifying assumptions are thus required.

The approach to this problem which has proved most useful is to assume a specific form for the external source fields and then try to choose data which is reasonably consistent with the assumed source morphology. In the best known example of this approach one assumes that the sources consist of vertically propagating plane waves of infinite horizontal extent. The plane wave source assumption is central to most quantitative data interpretation in local and regional geomagnetic induction studies. As is well known, and as we shall discuss in more detail below, the plane wave source assumption implies the existence of intercomponent and interstation transfer functions. Specifically, the classical magnetotelluric model first proposed by Tikhonov (1950) and Caignard (1953), the single station vertical field transfer function approach used in magnetovariational (MV) studies (Schmuker, 1964; 1970; Everett and Hyndman, 1967; Gregori and Lanzerotti, 1980) and the interstation transfer function method used for the computation of

anomalous horizontal fields from magnetometer array data (Schmuker, 1970; Beamish, 1977; Beamish and Banks (1983)) all make this simplifying assumption about the external sources.

Another example of a simplifying source assumption can be found in the work of Schultz and Larson (1983, 1987). In this global scale study of long period ($T \geq$ several days) geomagnetic variations it was assumed that external source potentials were all proportional to the P_1^0 spherical harmonic. Again, this assumption implies the existence of a transfer function, which in this case gives a linear relation between the north-south and vertical magnetic field components. More generally numerous global conductivity studies (see Rokityansky (1982) for a general review), have assumed that the external source potentials are linear combinations of a few of the lowest order spherical harmonics.

More specific source assumptions have been proposed for specific types of geomagnetic disturbances whose source current systems are reasonably well understood. For instance, Schmuker (1970) suggested that magnetic bay type disturbances should be analyzed separately with consideration given to the general spatial structure of bays. Another example is provided by the analysis of solar quiet day (*Sq*) records (e.g. Price and Wilkins, 1962) where the detailed geometry of the fields can be used for interpretation of the data (Campbell, 1986).

Finally, the magnetic spatial gradient technique (Schmuker, 1970; Kuckes, 1973; Lilley *et al.*, 1976; Jones, 1980) in which the statistical relationship between the vertical fields and the gradients of the horizontal fields is used to compute an equivalent 1-d magnetotelluric impedance makes the simplifying assumption that the fields can be represented as uniform plane wave sources plus uniform gradients of the horizontal field components. Kuckes *et al.*, (1985) have discussed this aspect of the spatial gradient method fairly explicitly, and they have shown that the vertical fields will in general be linearly related to two terms representing

the two polarizations of the plane wave sources plus three terms representing the gradients in the horizontal fields. We will return to this point in some detail in Chapter 3.

All of the simplifying source assumptions discussed above are special cases of the general model which will be central to this thesis. In all of these examples the possible external fields are restricted to a finite (and usually very low) dimensional subspace of the space of all possible source potentials. As we shall now show, it is this assumption which allows for the rigorous justification of the transfer function approach.

We assume then, that the space of possible source potentials Φ is of finite dimension p:

$$\Phi = \left\{ \phi_e : \phi_e = \sum_{j=1}^{p} \alpha_j \phi_j \right\} = \mathbf{Sp} \left\{ \phi_j : j = 1, p \right\} \qquad (2.2.1)$$

where **Sp** { · } denotes the span (i.e. the linear space consisting of all linear combinations) of the vectors in brackets. In the case where we assume that sources are plane waves the space of external potentials Φ is of dimension p=2

$$\Phi = \mathbf{Sp} \left\{ \phi_1, \phi_2 \right\} \quad \text{where} \quad \phi_1 = x \quad \phi_2 = y \qquad (2.2.2)$$

The potentials in the space Φ correspond to all possible polarizations of uniform vertically propagating plane wave sources. For the other examples discussed above the space Φ has dimension ranging from $p = 1$ (e.g. for the case where source potentials were assumed proportional to the P_1^0 spherical harmonic) to $p = 5$ (for the generalized spatial gradient model described by Kuckes *et al.*, 1985).

The assumption that the external sources lie in a fixed finite dimensional space implies a set of linear relations between measured field components. A

variant on this idea has been developed in great detail in the Russian literature. Berdichevsky and Zhdanov (1984) and Svetov and Shimelevich (1982) give an abstract treatment of this idea and discuss a number of specific examples. Svetov (1986) summarizes the principle points. The Russian approach assumes that the external currents $\mathbf{J}_{ext}$ lie in a finite dimensional space; our approach is slightly different in that we make the weaker assumption that the external source potentials lie in a finite dimensional space, but the basic idea is the same.

Suppose then that (2.2.1) holds. Then any observable magnetic field at the surface $z = 0$ is the gradient of the scalar potential $\phi = \phi_e + \phi_i$ where

$$\phi_e = \sum_{j=1}^{p} \alpha_j \phi_j$$

Then we may write

$$\mathbf{B} = \nabla\left(\phi_e + L_{\sigma,\omega}(\phi_e) \right) =$$

$$\sum_{j=1}^{p} \alpha_j \nabla\left(\phi_j + L_{\sigma,\omega}(\phi_j) \right) = \sum_{j=1}^{p} \alpha_j \mathbf{B}_j \qquad (2.2.3)$$

Thus we see that the observable magnetic fields lie in the space of finite dimension p spanned by the p basis functions $\mathbf{B}_j = \nabla\left(\phi_j + L_{\sigma,\omega}(\phi_j) \right)$ which give the total magnetic fields (internal plus external) resulting from the external source polarizations ϕ_j , $j = 1, p$. The electric fields can be added without changing this situation. Proceeding as in Dmitriev and Berdichevsky (1979) (see also the discussion in Chapter 3) it can be shown that a linear operator L_E (which also depends on σ and ω) relates the electric fields to the horizontal magnetic fields $\mathbf{B_h}$ on the surface $z = 0$. Thus:

$$\mathbf{E} = L_E(\mathbf{B_h}) = L_E\left(\sum_{j=1}^{p} \alpha_j \mathbf{B_{hj}} \right) = \sum_{j=1}^{p} \alpha_j L_E(\mathbf{B_{hj}}) = \sum_{j=1}^{p} \alpha_j \mathbf{E}_j(\mathbf{x}) \qquad (2.2.4)$$

where $\mathbf{E}_j$ gives the electric fields associated with ϕ_j.

Together (2.2.3) and (2.2.4) imply that all observable 5 component electromagnetic fields are linear combinations of p basis functions:

$$\mathbf{F}(\mathbf{x}) = \sum_{j=1}^{p} \alpha_j \mathbf{F}_j(\mathbf{x}) \qquad (2.2.5)$$

where

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} \mathbf{B}(\mathbf{x}) \\ \mathbf{E}(\mathbf{x}) \end{bmatrix} \qquad \mathbf{F}_j(\mathbf{x}) = \begin{bmatrix} \mathbf{B}_j(\mathbf{x}) \\ \mathbf{E}_j(\mathbf{x}) \end{bmatrix} \qquad j = 1, p$$

Suppose now that we measure these five components simultaneously at n stations, whose locations are given by $\mathbf{x}_i$, $i = 1, n$. After windowing and Fourier transforming the measured time series we obtain the fundamental frequency domain complex data vectors of dimension 5n. For now we assume that these measurements are noise free. For a fixed frequency ω we will denote this vector by $\mathbf{b}$, and use the similar notation $\mathbf{u}_j$ for the $5n$ dimensional complex vector giving the fields associated with the j^{th} external potential basis function ϕ_j at the station locations $\mathbf{x}_i$. Explicitly,

$$\mathbf{b} = \begin{bmatrix} \mathbf{F}(\mathbf{x}_1) \\ \mathbf{F}(\mathbf{x}_2) \\ \cdot \\ \cdot \\ \mathbf{F}(\mathbf{x}_n) \end{bmatrix} \qquad \mathbf{u}_j = \begin{bmatrix} \mathbf{F}_j(\mathbf{x}_1) \\ \mathbf{F}_j(\mathbf{x}_2) \\ \cdot \\ \cdot \\ \mathbf{F}_j(\mathbf{x}_n) \end{bmatrix}$$

By (2.2.5) $\mathbf{b}$ is a linear combination of the p basis functions $\mathbf{u}_j$

$$\mathbf{b} = \sum_{j=1}^{p} \alpha_j \mathbf{u}_j \qquad (2.2.6)$$

Put another way, $\mathbf{b}$ lies in the p dimensional subspace of $\mathbf{C}^{5n}$ spanned by the vectors $\mathbf{u}_j$:

$$\mathbf{b} \in \boldsymbol{R} = \mathrm{Sp}\left\{ \mathbf{u}_j : j=1, p \right\}$$

i.e. any observable frequency domain data vector **b** is a linear combination of the p basis functions $\mathbf{u}_j$. In general all of the five components will not be measured at some (or all) stations. Assuming that a total of $m \leq 5n$ channels of data are recorded, the definitions and dimensions of **b** and $\mathbf{u}_j$ and $\boldsymbol{R}$ are modified in the obvious way. We will refer to the space $\boldsymbol{R}$, a p dimensional subspace of $\mathbf{C}^m$, as the response space. This space contains all possible observable fields, and as we shall see knowledge of this space is equivalent to knowledge of all possible inter-component and interstation transfer functions. In this sense this space defines the response of the earth to the assumed space of possible external sources.

In general, with the vector **b** restricted to the p dimensional subspace $\boldsymbol{R}$ we would expect that fixing the values of p components of **b** would linearly determine the remaining $m - p$. To be specific, suppose the first p components of **b** are fixed (more generally, rearrange the order of the components so this is true). Let **U** be the matrix whose j^{th} column is the basis vector $\mathbf{u}_j$ and partition the vector **b** and the matrix **U** in the obvious way:

$$\mathbf{b} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix} \qquad \mathbf{U} = \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix} \quad \begin{matrix} p \times p \\ m-p \times p \end{matrix}$$

If the matrix $\mathbf{U}_1$ is non-singular and if $\mathbf{b} \in \boldsymbol{R}$, it is easy to check that

$$\mathbf{b}_2 = \mathbf{U}_2\mathbf{U}_1^{-1}\mathbf{b}_1 = \mathbf{T}\mathbf{b}_1. \qquad (2.2.7)$$

Thus provided $\mathbf{U}_1$ is non-singular the $m - p$ components of $\mathbf{b}_2$ may be linearly related to the p components of $\mathbf{b}_1$ by the matrix of transfer function coefficients $\mathbf{T} = \mathbf{U}_2\mathbf{U}_1^{-1}$. It is this fact which forms the basis for the transfer function method.

There are several well known examples of transfer functions in electromagnetic induction which we may site as specific cases of this general formulation.

The best known is the standard magnetotelluric impedance tensor ζ computed from magnetic and electric fields at a single site. In this case, as noted above, one assumes that the external sources are plane waves of infinite horizontal extent so that $p = 2$. We take basis vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ which correspond to two orthogonal source polarizations

$$\mathbf{u}_i = \begin{bmatrix} B_{xi} \\ B_{yi} \\ E_{xi} \\ E_{yi} \end{bmatrix} \qquad i = 1, 2$$

and form the single station impedance

$$\zeta = \begin{bmatrix} E_{x1} & E_{x2} \\ E_{y1} & E_{y2} \end{bmatrix} \begin{bmatrix} B_{x1} & B_{x2} \\ B_{y1} & B_{y2} \end{bmatrix}^{-1}$$

Other examples, such as the vertical field transfer function or the interstation anomalous horizontal field transfer function, arise in a similar fashion.

There are several remarks that are in order at this point. First, the matrix $\mathbf{U}_1$ defined above will not always be non-singular for all possible choices of p components as the fixed or independent variables. A very simple example of this is contained in the single component transfer function used by Schultz and Larson (1987). Here $p = 1$ and $m = 3$ - three components of the magnetic field are measured at a single station. With the P_1^0 assumption the (geomagnetic) east-west component of the magnetic field is zero. If this component of the field were chosen as the independent variable then we would have $\mathbf{U}_1 = 0$. In this particular case only a transfer function between the north-south and vertical field components is possible. It is easy to construct slightly more complex examples from the MT impedance tensor case. Here with $p = 2$ and $m = 4$ we would normally choose two components of the magnetic fields or two components of the electric field as independent variables and treat the two components of the other field type as the

dependent variable. If we used one electric (say E_y) and one magnetic (say B_x) field component as the independent variables we would not be assured that $\mathbf{U}_1$ would be non-singular.

Second, we have referred to the coefficients $\mathbf{T}$ of the linear relationship between the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ as transfer functions. This use of the term is rather careless in some ways. Our discussion here has focused on the fields observed for a fixed frequency ω and has ignored the time domain aspects of this problem. A transfer function, however, has both frequency and time domain representations. The time domain representation of the transfer function, the impulse response function, is just the inverse Fourier transform of the frequency domain representation that we consider here

$$\tilde{T}(t) = \int d\omega \, e^{2\pi i \omega t} \, T(\omega)$$

Two functions $\tilde{u}_1(t)$, $\tilde{u}_2(t)$ which are linearly related in the frequency domain by the transfer function $u_2(\omega) = T(\omega)u_1(\omega)$ are related in the time domain via the convolution $\tilde{u}_2(t) = \tilde{T} * \tilde{u}_1(t)$.

In the usual situation, $\tilde{u}_1(t)$ is the input variable, $\tilde{u}_2(t)$ is the output variable, and the transfer function $\tilde{T}(t)$ is causal ($\tilde{T}(t) = 0$ for $t < 0$) and stable ($\int_{-\infty}^{\infty} |\tilde{T}(t)|^2 dt < \infty$). These are properties that we expect from a physically realizable system - there can be no output before the input and if the input has finite power the output must also. In our situation it is *not* proper to consider the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ of (2.2.6) as, respectively, input and output variables. The decision as to which components are to be treated as input and which as output is totally arbitrary. If we consider the time domain relationship between the internal and external potentials, simple physical arguments suggest that we may consider ϕ_i, ϕ_e, respectively, as the output and input variables. The total fields observed at different points, however, should more properly be thought of as joint outputs

which may be causally and stably related to the input ϕ_e. In fact, Svetov (1986) claims that causality of the transfer functions of (2.2.6) can be established under very general circumstances, but that these transfer functions need not be stable. (No proof or reference is given for the first statement, although in very specialized cases causality can be easily established (e.g. Jones, 1980); the second statment is fairly obvious). At any rate, one must be careful here; the simple physical input-output model of the transfer function is not exactly valid for geomagnetic transfer functions for the simple reason that the measured quantities always contain a mixture of the external and internal fields (i.e. the inputs and outputs).

Note also that for $\mathbf{T}$ to be a transfer function, the response space $\boldsymbol{R}$ must not depend on frequency. For $\boldsymbol{R}$ to be independent of frequency the spatial and temporal dependence of the sources must be separable (i.e. $\phi_e(\mathbf{x},t) = f(\mathbf{x})g(t)$). In general, current systems which move relative to the earth (e.g. *Sq*) won't satisfy this constraint (Mareschal, 1986), so this is a fairly restrictive assumption. The existence of the frequency domain linear relationships of (2.2.6) for any fixed frequency ω, however, does not depend on $\mathbf{T}$ being causal, stable or, for that matter even a transfer function. In this thesis we study such frequency domain linear relationships even if they do not really constitute transfer functions. We will, however, generally call these relationships transfer functions, even though in some cases this will be an abuse of terminology.

2.3: The Response Space Formulation

Model Parameterization

In the previous section we demonstrated that the transfer function approach could be justified rigorously if the external source potentials could be restricted to a subspace of small finite dimension. In this section we will demonstrate that there is a slightly different way to use this simplifying assumption about the

sources which is more natural and more satisfying from both the physical and statistical perspectives. This approach is particularly useful for the analysis of data from small, regional geomagnetic arrays, and it is this application that we have in mind. Initially however, we will keep the discussion fairly general.

The transfer function approach described above is rather awkward in two distinct ways. First, as discussed above, it is not really physically correct to treat some channels as independent inputs and and others as dependent outputs. This introduces an asymmetry between the different field components which has no physical basis. In some cases, such as in the estimation of the MT impedance tensor from single station data, this asymmetry may seem fairly natural. Even here, where there is a fairly clear distinction between the magnetic fields and the electric fields, there is no physical basis for choosing one set of fields as inputs and the other as outputs (although there may be good statistical reasons for making such a choice for the estimation problem). In other cases, such as in the estimation of interstation horizontal field transfer functions, the asymmetry is even more unnatural. Here, one resorts to notions such as a "normal" reference station (e.g. Schmuker, 1970; Beamish, 1977) to justify the choice of independent input variables for the transfer functions. While the choice of a normal reference station with careful reference to other geological information may be useful for interpretation purposes, the concept of a special normal station is difficult to justify physically.

The choice of independent input components becomes even more awkward when the number of predictor variables p is not equal to two. For instance in the spatial gradient model proposed by Kuckes *et al.* (1985) with $p = 5$ some of the "input" variables are gradient terms which cannot be identified with any particular channel, but rather are linear combinations of various channels (with the method of estimation proposed in Kuckes *et al.*). In general, for array data, where the

total number of components becomes larger, and more complex models may be useful, the simple multiple input single output transfer function approach becomes increasingly awkward.

The standard TF approach is also difficult to justify on statistical grounds. The statistical model used in the standard TF model assumes that noise is restricted to the output channel. Specifically, using the notation from above, one assumes that each observed output channel X_i, $i = p+1, m$ is a linear combination of the p input channels plus an error term:

$$X_i = b_i + e = \sum_{j=1}^{p} T_{ij} b_j + e$$

The standard approach to estimation of the transfer function parameters is based on least squares (LS) estimation of each row $\mathbf{T}_{.i}$ of the matrix $\mathbf{T}$ separately. This is accomplished in the standard way by minimizing the residual sum of squares from N sets of Fourier coefficients of the input and output variables

$$\sum_{k=1}^{N} (X_{ik} - \sum_{j=1}^{p} T_{ij} b_{jk})^2$$

Note that the N independent input variable sets b_{ik}, $(k = 1, N; i = 1, p)$ are assumed to be measured without error. All measurement error, as well as any inadequacies in the model assumptions (i.e. about the external source structure) are treated as noise in the output variable.

Recently, robust alternatives to LS have been applied to the estimation of geomagnetic transfer functions (Egbert and Booker, 1986; Chave *et al.*, 1987). While these techniques have been shown to preform better than the standard LS estimate, they are essentially based on the same sort of model assumptions that are intrinsic to the transfer function approach - some variables are treated as independent (and noise free) and others as dependent (and noisy).

Clearly, it is seldom reasonable to assume that some channels are measured exactly, while others are noisy. Again, this assumption may be approximately valid for some frequency ranges in the single station MT impedance estimation problem, where one type of field may be measured much more accurately than the other. However, for the estimation of interstation horizontal magnetic field transfer functions, where the measurement process is the same for both dependent and independent variables, such an assumption seems absurd. The biasing effect of noise in the input variables is well known (Jenkins and Watts, 1968, Gamble *et al.*, 1979; Pedersen and Svennekjaer, 1984). The seriousness of this bias problem for MT impedance estimates led to the development of the remote reference technique (Gamble *et al.*, 1979).

It is much more reasonable to use a model which allows for noise in all channels. Let $\mathbf{X}_k$ denote the fundamental m-dimensional complex data vector consisting of Fourier coefficients from the k^{th} data segment for all m measured channels of data. If we assume that the external sources are elements of the finite dimensional space Φ defined above, and we allow for noise in all channels, our model for $\mathbf{X}_k$ is

$$\mathbf{X}_k = \mathbf{b}_k + \mathbf{e}_k = \sum_{j=1}^{p} \alpha_{jk}\mathbf{u}_j + \mathbf{e}_k = \mathbf{U}\alpha_k + \mathbf{e}_k \tag{2.3.1}$$

In (2.3.1) $\mathbf{u}_j$ gives the fields at the observation sites caused by the j^{th} external source potential function ϕ_j. We will refer to these vectors ($j = 1, p$) as the fundamental response vectors (FRV's). $\mathbf{U}$ is the $m \times p$ matrix whose j^{th} row is $\mathbf{u}_j$. The complex p-vectors α_k give the magnitude and phase of the p source functions ϕ_j, $j = 1, p$ for the k^{th} data segment. In the two dimensional case $p = 2$, the vectors α_k are just the polarization vectors from the standard wave propagation theory (e.g. Jackson, 1975). Following Samson (1983), we extend this nomenclature to refer to the vectors α_k as polarization vectors for the general case.

This sort of model, with noise in all components, has been applied to geomagnetic data analysis by Jupp (1978) and Park and Chave (1984). Both of these papers considered the use of a model with measurement noise in all channels for estimating the magnetotelluric impedance tensor for single station data. More generally, models of this sort have been extensively studied in the case where the data is real. These have been variously referred to in the statistical literature as *linear functional equation models, structural relationship models, factor analysis models* and *multivariate errors-in-variables models*. All of these linear statistical models are thoroughly reviewed in Anderson (1984) who discusses the distinctions between these models. Without worrying about these distinctions for now, we will refer to the model of (2.3.1) as a multivariate errors-in-variables (MEV) model. We defer discussion of model differences, together with other statistical aspects of the model, until Chapter 4.

The model given in (2.3.1) is overparametrized. Let $\mathbf{A}$ be any non-singular $p \times p$ matrix. Then

$$\mathbf{X}_k = \mathbf{U}\alpha_k + e_k = \mathbf{U}\mathbf{A}\mathbf{A}^{-1}\alpha_k + e_k = \tilde{\mathbf{U}}\tilde{\alpha}_k + e_k \tag{2.3.2}$$

where

$$\tilde{\mathbf{U}} = \mathbf{U}\mathbf{A} \qquad \tilde{\alpha}_k = \mathbf{A}^{-1}\alpha_k$$

Thus we see that the parameters $\mathbf{U}$ and α_k, $k = 1, N$ cannot be uniquely determined from the data. The parameters $\tilde{\mathbf{U}}$ and $\tilde{\alpha}_k$, $k = 1, N$ give rise to exactly the same data values. On the other hand, letting $\tilde{\mathbf{u}}_j$ be the j^{th} column of $\tilde{\mathbf{U}}$, it is easy to check that

$$\boldsymbol{R} = \mathbf{Sp}\left\{\mathbf{u}_j : j = 1, p\right\} = \mathbf{Sp}\left\{\tilde{\mathbf{u}}_j : j = 1, p\right\} \tag{2.3.3}$$

so that the response space $\boldsymbol{R}$ is uniquely determined.

The physical significance of this non-uniqueness can be readily understood. To be definite we consider the specific case where $n = 2$ and the external sources are assumed to be plane waves. If we take

$$\phi_1 = x \qquad \phi_2 = y$$

then the FRV's $\mathbf{u}_1$, $\mathbf{u}_2$ are the fields which would be seen at all stations in the array for plane wave sources of unit magnitude that are linearly polarized in the x and y directions respectively. While ideally we would like to determine the FRV's from the data, this will not generally be possible without some additional information. The problem stems from the fact that we cannot directly observe the external source polarizations for any of the data.

Consider a simple case of an array consisting of n stations of 3 component (magnetic field only) data and ignore the effects of noise. Then the components of a (noise free) observation $\mathbf{b}$ are

$$\mathbf{b} = \begin{bmatrix} B_x(\mathbf{x}_1) \\ B_y(\mathbf{x}_1) \\ B_z(\mathbf{x}_1) \\ \hline \cdot \\ \cdot \\ \cdot \\ \hline B_x(\mathbf{x}_n) \\ B_y(\mathbf{x}_n) \\ B_z(\mathbf{x}_n) \end{bmatrix}$$

If the electrical conductivity of the earth varies with depth only, then the FRV's take a particularly simple form. For this particular case the internal horizontal fields are equal (in amplitude and phase) to the external fields (this can be shown to hold in general for plane wave sources over a 1-d earth; see Rokityansky, 1982, pp 43-45).

From (2.2.2) and (2.2.3), then, we have

$$\mathbf{u}_1 = \begin{bmatrix} 2 \\ 0 \\ 0 \\ \hline \cdot \\ \cdot \\ \cdot \\ \hline 2 \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 2 \\ 0 \\ \hline \cdot \\ \cdot \\ \cdot \\ \hline 0 \\ 2 \\ 0 \end{bmatrix} \tag{2.3.4}$$

In this case we can unambiguously identify the external source polarization associated with any particular data vector $\mathbf{b}$. This is only because we knew ahead of time the form of the vectors $\mathbf{u}_i$. More generally lateral variations in conductivity will cause local concentrations of currents which will locally distort the magnetic fields, perturbing amplitudes, phases and directions at each station in an unknown manner (indeed these perturbations are precisely what we wish to estimate). In this case the FRV's will not have the simple form of (2.3.4), and a precise determination of the source polarization associated with $\mathbf{b}$ is not possible. The ambiguity in the parameters stems from the fact that both the source polarizations parameters α_i and the field basis vectors $\mathbf{u}_i$ are unknown, and uncertainty in one leads to uncertainty in the other.

While (2.3.2) says that mathematically we cannot unambiguously identify specific vectors in the response space with specific external source potentials, we can use other physical or geological information to make such an identification, at least approximately. In the context of interstation transfer functions the resolution of the ambiguity is equivalent to the problem of defining the "normal" reference fields. The standard approach to this problem is to assume that the fields at one particular station (station 1, say) are not perturbed by lateral conductivity variations. This assumption imposes a constraint on the first two components of the FRV's,

$$u_{ij} = 2\delta_{ij} \qquad i, j = 1, 2 \tag{2.3.5}$$

where δ_{ij} is the Kroenecker delta - i.e. it is assumed that (2.3.4) holds for the horizontal components at station one. This constraint allows us to identify the FRV's in terms of any pair of linearly independent vectors $\tilde{\mathbf{u}}_1$, $\tilde{\mathbf{u}}_2$ contained in $\boldsymbol{R}$. If $\mathbf{U}$ is the matrix whose i^{th} column is the i^{th} FRV, $\tilde{\mathbf{U}}$ the corresponding matrix of the $\tilde{\mathbf{u}}_i's$, $\tilde{\mathbf{U}}_1$ the 2×2 matrix consisting of the first two rows of $\tilde{\mathbf{U}}$, and the constraint (2.3.5) holds then it is easy to check that

$$\mathbf{U} = 2\tilde{\mathbf{U}}\tilde{\mathbf{U}}_1^{-1} = 2\begin{bmatrix} \mathbf{I}_2 \\ \mathbf{T} \end{bmatrix} \tag{2.3.6}$$

where $\mathbf{T}$ is the matrix of transfer functions defined above in (2.2.7). Note that $\mathbf{T}$ gives the values of all components of the fields seen at all other stations relative to the values of the horizontal fields observed at station 1. This interpretation is valid independent of any assumptions about the "normality" of this reference station. Only the interpretation of the vector given in (2.3.6) as an FRV depends on the assumption that the reference fields are normal.

In statistical terms, the parametrization of the MEV model of (2.3.1) is not *identifiable* (Bickel and Doksum, 1977 p. 60) (i.e. distinct parameters give rise to identical data). To obtain a parametrization that is identifiable we must impose some constraints on the FRV's and/or the polarization vectors. Many sorts of constraints can be imposed to make the parametrization of the MEV model identifiable. Anderson (1984) discusses a number of these. For our purposes the FRV's, which characterize the response of the earth to the external sources, represent the parameters of interest. The polarization vectors (which characterize the sources) are nuisance parameters. We thus concentrate our attention on the unique identification of the FRV's. As we have already discussed, one way to make the MEV model parametrization identifiable is to make the normal station assumption of (2.3.5). We will adopt a more general point of view.

As we have noted in (2.3.3) the response space $\boldsymbol{R}$ is uniquely determined. Thus, one way to resolve this non-uniqueness is to adopt the viewpoint that the fundamental parameter to be estimated is $\boldsymbol{R}$. Treating $\boldsymbol{R}$ as the fundamental parameter emphasizes the indeterminacy which is fundamental to our problem. When noise is present in the data, estimates $\hat{\boldsymbol{R}}$ of the response space will be randomly perturbed from $\boldsymbol{R}$. We can use a statistical model to compute estimation errors which serve to characterize the degree of uncertainty in our knowledge of $\boldsymbol{R}$ which is due to this noise. Our uncertainty in the FRV's within $\boldsymbol{R}$ is of an entirely different nature. This uncertainty is present even for an infinite amount of noise free data. It is not possible to quantify the degree of this uncertainty from the data alone; additional information or assumptions must be provided. The assessment of this uncertainty is essentially independent of the statistical parameter estimation problem.

There are other reasons to treat $\boldsymbol{R}$ as the fundamental parameter. As we have indicated above, all interstation and intercomponent transfer functions, can be computed from the FRV's which span $\boldsymbol{R}$. More generally, it is not hard to check that if $\{ \tilde{\mathbf{u}}_i : i = 1, p \}$ is another basis for $\boldsymbol{R}$, then the transfer functions can be computed in an identical manner from these vectors. Thus, all interstation and intercomponent transfer functions are determined by $\boldsymbol{R}$; the response space formulation thus subsumes the transfer function formulation.

While in theory we can think of the set of p-dimensional subspaces of $\mathbf{C}^m$ as the parameter space, in practice it will be necessary to consider explicit representations of these subspaces. Such explicit representations will be useful for the presentation of estimates of $\boldsymbol{R}$, and, they will be necessary for the derivation of statistical properties (such as estimation errors) of these estimates. We will make the representation explicit in two different ways. The first approach maintains our general formulation and represents the response space as the projection matrix $\mathbf{P}_R$

which projects $\mathbf{C}_m$ onto $\boldsymbol{R}$. We will discuss and apply this approach in Chapter 4.

The second approach is a generalization of the normal station constraint described above - i.e. the non-identifiability of $\mathbf{U}$ is resolved by imposing more general linear constraints on the columns of $\mathbf{U}$. We will consider constraints of the general form

$$\mathbf{WU} = \mathbf{C} \tag{2.3.7}$$

where $\mathbf{W}$ is a $p \times m$ matrix of constraint coefficients and $\mathbf{C}$ is a $p \times p$ matrix. If, as above, the columns of $\tilde{\mathbf{U}}$ form a basis for $\boldsymbol{R}$ then we can find $\mathbf{U}$ satisfying the constraints (2.3.7) as

$$\mathbf{U} = \tilde{\mathbf{U}}(\mathbf{W}\tilde{\mathbf{U}})^{-1}\mathbf{C}$$

For this to be possible, of course, $\mathbf{W}\tilde{\mathbf{U}}$ must be invertible. Note that this condition does not depend on the choice of the basis $\tilde{\mathbf{U}}$. If the columns of both $\mathbf{U}$ and $\tilde{\mathbf{U}}$ form a basis for $\boldsymbol{R}$, then there is a non-singular matrix $\mathbf{A}$ such that $\mathbf{UA} = \tilde{\mathbf{U}}$ so

$$\mathbf{W}\tilde{\mathbf{U}} = \mathbf{WUA}$$

Since $\mathbf{A}$ is invertible, $\mathbf{W}\tilde{\mathbf{U}}$ is non-singular if and only if $\mathbf{WU}$ is.

The normal station constraint discussed above is a special case of (2.3.7) obtained by setting

$$\mathbf{W} = \left[\mathbf{I}_2 \mid 0\right] \qquad \mathbf{C} = 2\mathbf{I}_2$$

This general constraint formulation allows some useful alternative approaches to the definition of normal fields. One which we will use in particular for the plane wave source case $p = 2$ defines the average horizontal fields as normal - i.e. we assume that the fields averaged over the array are roughly the same as would be seen for a 1-d earth. This is accomplished (for 3 component data at n stations) with

$$\mathbf{W} = \frac{1}{n}\begin{bmatrix} 1\,0\,0 & \cdots & 1\,0\,0 \\ 0\,1\,0 & \cdots & 0\,1\,0 \end{bmatrix} \qquad \mathbf{C} = 2\mathbf{I}_2$$

In the absence of other information, we would argue that this constraint is probably the most reasonable for geomagnetic induction problems.

2.4: Estimation of the Response Plane

Eigenvectors of the Spectral Density Matrix

We now consider the estimation of the response plane $\boldsymbol{R}$ at frequency ω from a series of N frequency domain data vectors $\mathbf{X}_k$. We describe the basic idea behind our approach here but postpone the full development of the statistical methodology until Chapter 4.

The sample spectral density matrix (SDM) at frequency ω is

$$\mathbf{S} = \frac{1}{N}\sum_{k=1}^{N}\mathbf{X}_k\mathbf{X}_k^* \tag{2.4.1}$$

where here, and subsequently, the superscript * represents the complex conjugate transpose. The SDM is basically the multivariate analogue of the power spectral density for stationary time series and gives a frequency domain description of the joint second moment properties for a multivariate time series. $\mathbf{S}$ provides a basic method of moments estimator of the true SDM. A thorough treatment of the SDM and many applications are given by Brillinger (1981). From our point of view, in which we concentrate on a single frequency at a time and ignore most of the time series aspects of this problem, it is most useful to consider the vectors $\mathbf{X}_k$ as independent identically distributed complex random m vectors. Then $\mathbf{S}$ is just the sample covariance matrix. Note that $\mathbf{S}$ is an $m \times m$ positive semi-definite Hermitian matrix.

Consider first the case of perfect, noise free data where the external sources are restricted to the p dimensional space Φ. Then the data vectors satisfy (2.2.6) which we may write in matrix notation as

$$\Xi = \mathbf{U}\alpha \tag{2.4.2}$$

where Ξ is the $m\times N$ matrix whose k^{th} column is the k^{th} data vector $\mathbf{X}_k$; $\mathbf{U}$ is, as above, the $m\times p$ matrix of response vectors; and α is the $p\times N$ matrix whose ik^{th} element is α_{ik}. Then we have

$$\mathbf{S} = N^{-1}\,\Xi\Xi^* = N^{-1}\mathbf{U}\alpha\alpha^*\mathbf{U}^* = \mathbf{U}\Sigma_\alpha\mathbf{U}^* \tag{2.4.3}$$

where we define $\Sigma_\alpha = N^{-1}\alpha\alpha^*$ Since $\mathbf{S}$ is Hermitian it can be diagonalized by a unitary matrix

$$\mathbf{S} = \mathbf{V}\Gamma\mathbf{V}^*$$

where $\Gamma = \mathbf{diag}(\gamma_1, \cdots, \gamma_m)$ is a diagonal matrix and $\mathbf{V}$ is an $m\times m$ unitary matrix ($\mathbf{V}^*\mathbf{V} = \mathbf{I}_m$). The i^{th} column of $\mathbf{V}$, $\mathbf{v}_i$ is an eigenvector of $\mathbf{S}$ and γ_i the corresponding eigenvalue. We now will show that the number of non-zero eigenvalues is p and that the corresponding eigenvectors determine the response plane $\boldsymbol{R}$.

First

$$\gamma_i\mathbf{v}_i = \mathbf{S}\mathbf{v}_i = (\mathbf{U}\Sigma_\alpha\mathbf{U}^*)\mathbf{v}_i$$

So, provided $\gamma_i \neq 0$ we may write

$$\mathbf{v}_i = \gamma_i^{-1}\mathbf{U}\,(\Sigma_\alpha\mathbf{U}^*\mathbf{v}_i)$$

which shows that $\mathbf{v}_i$ is a linear combination of the columns of $\mathbf{U}$ (the response vectors) and thus $\mathbf{v}_i \in \boldsymbol{R}$ whenever $\gamma_i \neq 0$. Let l be the number of non-zero eigenvalues of $\mathbf{S}$, and order the eigenvectors $\mathbf{v}_i$ so that the first l of them correspond to these eigenvalues. Then we have

$$\mathbf{Sp}\left\{\mathbf{v}_i,\, i = 1,\, l\right\} \subseteq \boldsymbol{R}$$

Since the eigenvectors $\mathbf{v}_i$ are linearly independent and since the dimension of $\boldsymbol{R}$ is p we have $l \leq p$. Note also that l is the rank of $\mathbf{S}$. From the representation of $\mathbf{S}$ in (2.4.3) and the fact that the matrix $\mathbf{U}$ is of full rank ($= p$) we see that the rank of $\mathbf{S}$ is equal to the rank of the $p\times p$ matrix Σ_α. If this is of full rank (i.e. non-

singular) then we have $l = p$ and

$$\mathbf{Sp}\left\{\mathbf{v}_i:\, i = 1,\, p\right\} = \boldsymbol{R} \tag{2.4.5}$$

The matrix Σ_α is essentially the covariance matrix of the polarization vectors α_k, and can thus be used to compute the average power for any chosen polarization. If Σ_α is singular then some of the possible source polarizations allowed for in the model have not been sampled. Note that if the sample size is too small $N < p$, then this situation must occur. In general, however, with sufficient data we expect that all polarizations will be sampled so that Σ_α will be non-singular. Otherwise, a reduced model for the external sources would be called for. Hence, it is reasonable to assume that Σ_α is non-singular so that for perfect (noise free) data, the eigenvectors associated with the non-zero eigenvalues of $\mathbf{S}$ exactly span the response space $\boldsymbol{R}$.

For real data, which will always be contaminated by some noise, the situation will be more complicated. We consider first the expected value of the sample SDM in the presence of noise. Now with the data vectors $\mathbf{X}_k$ satisfying (2.3.2) a simple calculation shows that, for errors which are uncorrelated with the signal

$$\Sigma = \boldsymbol{E}(\mathbf{S}) = \mathbf{U}\Sigma_\alpha\mathbf{U}^* + \boldsymbol{E}(\mathbf{e}\mathbf{e}^*) = \Sigma_S + \Sigma_N \tag{2.4.5}$$

In (2.4.5) $\boldsymbol{E}$ is the expectation, Σ_S represents the noise free signal covariance, and Σ_N is the covariance matrix of the error (noise) vectors $\mathbf{e}_k$. The situation is by far simplest if the errors are uncorrelated and of equal magnitude in all components so that the error covariance has the simple isotropic form

$$\Sigma_N = \sigma^2\,\mathbf{I}_m \tag{2.4.6}$$

If $\mathbf{v}_i$ is the i^{th} eigenvector of of the signal covariance matrix Σ_S then

$$\Sigma\,\mathbf{v}_i \;=\; (\Sigma_S + \Sigma_N)\,\mathbf{v}_i = (\gamma_i + \sigma^2)\,\mathbf{v}_i$$

so that $\mathbf{v}_i$ is also an eigenvector of the total (signal + noise) covariance matrix $\Sigma = \boldsymbol{E}\mathbf{S}$. Note that now all eigenvalues are non-zero; the $m - p$ smallest eigenvalues of Σ are now all equal to σ^2. Also the eigenvectors associated with the largest p eigenvalues remain identical; these p "dominant" eigenvectors still precisely define the response space $\boldsymbol{R}$.

The matrix Σ is the expectation or mean of the sample SDM **S** and represents (we hope) the SDM for an infinite amount of data. When we have only a finite sample (always) there will be random deviations of **S** from it's mean. In particular, the distribution of eigenvalues will be perturbed from the simple form described above. The signal-to-noise ratio can be roughly characterized by the ratio γ/σ^2 where γ is a typical size for one of the non-zero eigenvalues of Σ_S. If this signal-to-noise ratio is reasonably large we would expect that the distribution of eigenvalues will approximate the ideal case described above in a useful way. Specifically, we would expect that there will be p large eigenvalues, roughly of magnitude γ, and m–p small eigenvalues with approximate magnitude σ^2. We would also expect the eigenvectors associated with the p largest eigenvalues to define a p dimensional subspace which is perturbed from, but reasonably close to the response space $\boldsymbol{R}$. Thus, with a reasonable signal-to-noise ratio we would expect that a good estimate of $\boldsymbol{R}$ would be

$$\hat{\boldsymbol{R}} = \mathbf{Sp}\left\{ \mathbf{v}_j : j = 1, p \right\} \tag{2.4.7}$$

In fact this estimate can be justified by several explicit statistical criteria. We will discuss these issues, as well as the estimation of errors for these estimates in Chapter 4.

When the errors vectors do not have the simple isotropic form of (2.4.6), the situation can be far more complicated. If the signal to noise ratio is reasonably large, the eigenvalues will still divide into two fairly distinct groups, consisting of

p large, and m–p smaller eigenvalues. Now however, the distribution of of the smaller eigenvalues will generally be more complicated. While in the isotropic errors case the m–p smallest eigenvalues are all roughly equal, for the general case these eigenvalues may be quite spread out. More significantly, the eigenvectors of the total expected covariance matrix Σ will not in general be the same as the eigenvectors of the signal covariance matrix Σ_S. Thus the estimate of the response space $\hat{\boldsymbol{R}}$ given in (2.4.7) will in general be perturbed from the true response space even in the idealized infinite data case. In statistical terms, $\hat{\boldsymbol{R}}$ defined by (2.4.7), will in general be an asymptotically biased estimate of $\boldsymbol{R}$.

An asymptotically unbiased estimate of $\boldsymbol{R}$ can be computed if the the structure of the noise covariance matrix Σ_N is known or can be consistently estimated up to a multiplicative constant (e.g. Gleser, 1981; Park and Chave, 1984). Roughly, this estimate is computed by using the matrix Σ_N to transform the sample SDM **S** so that the isotropic error assumption holds. We describe this in more detail in Chapter 4, where we also consider approaches to the estimation of the form of the noise covariance.

For a number of reasons we should expect that the error covariance will often fail to be isotropic. If electric field measurements are included in the analysis then we are faced with the problem of chosing the units that the fields are expressed in. Whether or not the magnitude of the noise power in the electric and magnetic field components are equal clearly depends on the relative choice of units for these fields. Even for fields of one type, measured with instruments with theoretically uniform noise characteristics, the assumption of equal error magnitudes in all channels may often be questionable. In electric field measurements, for instance, electrode noise may be the dominant noise. The magnitude of this electrode noise will depend in part on soil type and moisture content, and will thus vary from site to site. In all channels, the magnitude of local cultural noise is highly site

dependent, and the effects often vary from channel to channel at a fixed site. There are other sources of noise in the magnetic fields which can effect different components differently. For example, vibrations of the magnetometer sensor in the earth's magnetic field results in non-isotropic noise in the measured magnetic fields (Pedersen, 1986). The effects of variations in temperature on instrument response also can effect the magnetic components differently.

As long as the noise remains incoherent between channels, the noise covariance matrix Σ_N is a diagonal matrix, with potentially different entries on the diagonal for each channel. More generally, some forms of noise may be coherent between channels so that Σ_N need not be diagonal. This coherence may occur between components at a single site or on an array wide basis. The first case can easily arise, for instance, with local cultural noise in the magnetic fields. The second case can arise from several distinct causes. For instance, some sources of cultural noise (e.g. DC power lines or DC electric railways; Schnegg *et al.*, 1986)) can extend over large regions and can thus lead to noise which is coherent between measurements made at widely separated sites. There are other sources of coherent noise which are not anthropogenic. For instance, at daily variation periods the affect of diurnal temperature variations on instrument responses may result in noise which is coherent over large areas. As a further example, electric fields on land induced by oceanic tidal motions may be considered a coherent source of noise. This last example, may seem to strain the definition of noise a bit, since tidally induced electric fields are a part of the true electric field signal. Indeed, for Junge (1986), this was the desired signal. In a sense the same point could be made for cultural noise, both local and large scale - it is not really noise but is part of the true electromagnetic signal.

The problem of deciding what is noise and what is signal will arise often in this dissertation. We will often treat a portion of the measured electromagnetic

fields as noise even when this signal is coherent between stations and arises from ionospheric sources. The distinction between noise and signal may often seem rather arbitrary. To a large extent, the distinction depends on what one is trying to do with the data. If one is interested in an exploratory analysis designed to detect what sorts of information the data contains, it could well be argued that anything which is coherent across the array is signal. If, on the other hand, we make specific model assumptions about the signal, and the goal is to make formal inferences about specific parameters (e.g. the response of the earth to plane wave sources), then it is possible, and we would argue reasonable to treat the effect of deviations from the model assumptions as noise. From this perspective it is the model assumptions which define what is noise and what is not. Those fields which arise from the assumed sources are signal, those which do not are noise. This point of view is most reasonable if the deviations from the simplifying model assumptions are relatively small perturbations - i.e. if the assumed model provides a reasonable fit to the data. If this is not the case, then one could argue that it would be more reasonable to expand the model. At any rate, we shall see that the violation of our simplifying assumptions about sources will generally result in complications which can be reasonably treated as coherent noise.

From the above discussion we see that, in general, unbiased estimation of the response space will require knowledge of the noise covariance matrix. In addition it is clear that estimation errors and other properties of response space estimators, will depend on the noise structure in a fundamental way. Thus, an important aspect of the estimation of the response space is the estimation of Σ_N. Our approach to this problem will be described in detail in Chapter 4. In the remainder of this chapter we will adopt a more exploratory point of view and look at some specific examples.

2.5: Some Examples

We will now present some example applications of the ideas which have been discussed in this chapter. We will discuss in particular results from a series of 3-5 station, 3-component (m = 9–15) MV (magnetic fields only) arrays which were run in Southwestern Washington in 1982 (figure 2.1), and a single five station 5-component (electric plus magnetic fields) (m = 25) MT linear array that was run in Western Oregon as part of the EMSLAB project (The EMSLAB Group, 1987; figure 2.2). All of these arrays were located at geomagnetic mid-latitudes.

We give first a brief description of some experimental procedures and general data characteristics which may be useful for interpreting some of the results described below. Magnetic field measurements were made with EDA fluxgate magnetometers, which were either partially or fully (EMSLAB array only) buried to limit temperature effects. Electric field measurements were made using lead-lead chloride electrodes spaced roughly 250 meters apart and differential input amplifiers. Electric fields were high pass filtered (with a cut off period of roughly two days) and all components were low pass filtered to reduce aliasing. Data was digitally recorded with either an 8 or 16 second sampling interval. The total time series length used in the analyses described here ranged from 10 days for most of the 3-component MV arrays to about 4 weeks for the EMSLAB MT line. The periods which we will consider are in the range 25 - 10000 seconds. Time series processing methods used are outlined at the beginning of this chapter and described in more detail in Egbert and Booker (1986).

All of the arrays discussed here (and indeed all arrays studied in this thesis) are small, both in the sense that the number of stations involved are few, and (more importantly) in the sense that the spatial extents of the arrays are small. (Station spacings range from roughly 10 to a maximum of 50 kilometers and total array sizes from a few tens to less than 200 kilometers.) At geomagnetic mid-

latitudes external source length scales are typically on the order of several thousands of kilometers (e.g. Porath *et al.*, 1971), so that source fields are usually fairly uniform over arrays of this small size. We would thus expect that the external sources could typically be treated as plane waves of infinite horizontal extent. Indeed, as we have discussed above, it is this assumption which rigorously justifies the usual intercomponent and interstation transfer functions which have become a standard tool for data analysis and interpretation in regional array studies (e.g. Beamish, 1977; Ingham *et al.*, 1983; Alabi, 1983; Gough and Ingham, 1983).

The plane wave source assumption corresponds to a response space of dimension p = 2. Thus, if the plane wave source assumption is reasonable, the SDM **S** should have two dominant eigenvalues. In figure 2.3 we plot the ordered eigenvalues for the 25×25 SDM computed from 5 MT stations on the EMSLAB profile for a period of 1000 seconds. As expected, there are indeed two dominant eigenvalues corresponding to the two plane wave source polarizations. Note that while the first two eigenvalues are, as expected, much larger than all of the others, there is a significant amount of structure in the remaining eigenvalues - the isotropic noise model which leads to a simple estimation theory for the plane wave response space will clearly not be appropriate. There are several additional eigenvalues which are reasonably well above background noise levels. We will explore the significance of these more fully in below and we will demonstrate that they can be related to the violation of our simplifying source assumptions. We first consider the dominant two eigenvalues and the associated eigenvectors which define the plane wave response space in somewhat more detail.

The occurrence of two dominant eigenvalues of the SDM is ubiquitous. In our experience with over 500 SDM's from roughly 30 arrays, the matrices **S** are always approximately of rank two. The fraction of the total power which is fit by

the plane wave source model (more precisely: a two dimensional response space) is given by

$$R^2 = (\gamma_1 + \gamma_2) / \sum_{i=1}^{m} \gamma_i = (\gamma_1 + \gamma_2) / \mathbf{Tr\ S} \quad . \tag{2.5.1}$$

Here γ_i, $i = 1, m$ are the eigenvalues of **S**, ordered largest to smallest, so that $\gamma_1 + \gamma_2$ represents the power in the two dimensional space spanned by the two dominant eigenvectors, and **Tr S**, the trace of **S**, gives the total power in all channels. In figure 2.4 we plot R^2 as a function of frequency for a set of 9 three component arrays. Note that at all frequencies R^2 is greater than .95 and often R^2 is nearly one.

For the EMSLAB array each eigenvector of **S** is a 25 dimensional complex vector whose components correspond to the five field components B_{xi}, B_{yi}, B_{zi}, E_{xi} and E_{yi}, at each of the five stations (i=1,5). We can display 10 of the 25 components corresponding to the horizontal magnetic (or electric) components by plotting the real and imaginary parts of the 2 component field vectors (B_{xi}, B_{yi}) (or (E_{xi}, E_{yi})) for each station on a map of station locations. In figure 2.5 we exhibit such plots for the eigenvectors corresponding to the two dominant eigenvalues from figure 2.3. Note that on the magnetic field plot the components corresponding to vertical fields B_{zi} are printed under the station locations so all 25 components are plotted.

The magnetic field plots (figures 2.3(a) and (c)) show that the horizontal field components of both of the dominant eigenvectors are fairly uniform over the array. While some variations are apparent, the approximate uniformity of these fields indicates that the plane wave source assumption is at least a reasonable approximation. We thus interpret the fields plotted here as estimates of the magnetic fields which would be seen at the five stations in the array for two distinct plane wave source polarizations. Although, as we have discussed above, the exact

external source polarizations corresponding to these total field vectors are indeterminate, it is reasonable to assume that the source polarizations are roughly the same as the polarizations of the total field vectors. Note that since the real and imaginary field vectors are not co-linear the source polarizations are of the general elliptic form. The polarization of the fields for the first eigenvector represent the source polarization of maximum power, while the second eigenvector gives the polarization of minimum power. Note that at this period the dominant polarization has most of it's power in the direction of geomagnetic north. This feature is frequency dependent; it is typically quite noticeable at a period of 100 seconds but disappears at longer ($\approx$ 3000 seconds) periods.

The vertical magnetic and horizontal electric fields which are produced by these two two elliptically polarized sources are displayed in figures 2.5(a) and (c) and are plotted in figure 2.5(b) and (d), respectively. Note that these components do not show the same uniformity as the horizontal magnetic fields.

While the two dominant eigenvectors plotted above completely define the response space, they are somewhat difficult to interpret since they correspond to elliptical source polarizations. Interpretation is much easier if the source polarizations are linear and if the coordinate system is chosen to match the dominant geologic strike of the region. As discussed in section 2.3, the unique determination of the total fields corresponding to any given source polarization requires some additional assumptions to define the "normal" fields. In figure 2.6 we plot the magnetic and electric fields (again at a period of 1000 seconds) corresponding to source fields linearly polarized in the east-west and north-south directions. For these plots we have assumed that the average fields are normal and can be used to uniquely define the source polarizations (see section 2.3). Note that we have chosen geographic coordinates here because the dominant geologic strike (most significantly the continental margin) is roughly north-south.

The magnetic fields for the north-south polarization have nearly uniform horizontal components and very small vertical components. In contrast, there are relatively large variations in the horizontal components and significant vertical field components for the east-west polarization. This reflects the dominant north-south strike to the conductivity variations in this region, with the Pacific coast, the Willamette Valley and Cascade ranges all trending north-south. The most obvious feature of the electric fields is the large variation in both direction and magnitude of the fields. This can largely be ascribed to the distortion of the electric fields by near surface conductivity inhomogeneities.

The geologically interesting part of the horizontal fields is the "anomalous" part - i.e the deviation from the normal fields. These can easily be obtained from the total fields displayed above (once a normal field definition has been decided upon) by subtracting the normal fields. In figure 2.7, we display the anomalous horizontal fields from the EMSLAB array for three periods with east-west inducing magnetic fields. For this calculation we have taken the fields at the western-most station as normal. The effect of the Willamette valley, clearly seen at 100 seconds, weakens rapidly with increasing period. A deeper anomaly is seen in the Eastern part of the array (in the High Cascades). This is still significant at 3000 seconds. A much more thorough (but still quite preliminary) interpretation of the data from the EMSLAB experiment is given by The EMSLAB Group (1987).

We now turn our attention to the smaller eigenvalues of the SDM shown in figure 2.3. In the context of the plane wave source assumption these eigenvalues represent noise. As we shall now show, a significant fraction of this "noise" is coherent across the array and reflects the fact that the sources are not exactly plane waves. The clearest demonstration of this involves the eigenvector associated with the third largest eigenvalue.

If the plane wave source assumption holds exactly, then over a 1-d earth the vertical magnetic fields B_z will be zero. If in fact, this assumption holds only approximately, with the source length scales r_0 being finite but still large compared to the skin depth of the magnetic fields δ, the vertical fields will be proportional to δ/r_0 (Schmuker, 1970). These vertical fields will vary slowly over source length scales; hence, for a small (compared to typical source length scales) array, this component of source related B_z will be approximately equal and in phase at all stations in the array. The fields seen in the array due to this effect will thus be approximately proportional to the unit vector

$$\mathbf{u_z} = \frac{1}{\sqrt{n}}\left[\,0\,0\,1\,0\,0\,\middle|\,\cdots\,\middle|\,0\,0\,1\,0\,0\,\right]^T \qquad (2.5.2)$$

which is zero for all horizontal components and equal and in phase for the vertical magnetic components for all stations. In figure 2.8 we plot the fraction of power in the third eigenvector $\mathbf{v}_3$ which is parallel to this "source Z" vector $\mathbf{u_z}$ for the EMSLAB MT array. For periods above 1000 seconds, the third eigenvector is essentially coincident with $\mathbf{u_z}$. This demonstrates graphically that at longer periods violations of simplifying source assumptions are the dominant source of "noise" in fitting the vertical field response to plane wave sources. This has been noted previously in the context of single station vertical field transfer function estimation by Beamish (1979) and Egbert and Booker (1986).

Figures 2.9 and 2.10 summarize similar results from 9 three component MV arrays. Again, for periods beyond about 500 seconds $\mathbf{v}_3$ is essentially parallel to the source Z vector $\mathbf{u_z}$ (figure 2.9) and the power in this eigenvector makes up a significant fraction of the total noise power. We illustrate this in figure 2.10 where we plot (for 9 MV arrays)

$$\gamma_3 / \sum_{i=3}^{10} \gamma_i$$

- the fraction of the total residual power (power which is not in the two dominant plane wave eigenvectors) which is in the third eigenvector.

While the coincidence of the third eigenvector with the source Z vector $\mathbf{u_z}$ explains a significant fraction of the noise power, there is clearly substantial additional structure in the eigenvalues plotted for the EMSLAB array in figure 2.3. In contrast to the first three eigenvectors, which can be readily related to specific aspects of geomagnetic array data, the subsequent eigenvectors are not so clearly interpretable by themselves. This is because the additional structure in the data is due to several effects (e.g. unusually noisy stations, differences in noise power between electric and magnetic field measurements, complications to the simple source effects model discussed above in connection with the third eigenvector). These are all mixed up in the remaining eigenvectors. To understand the additional structure in the data some alternative approaches are thus required.

One approach which has proved useful is to analyze the spectral density matrices of the horizontal magnetic and electric fields separately. The SDM's formed from only the horizontal magnetic or electric field components are 10×10 complex matrices which we denote by $\mathbf{S_H}$ and $\mathbf{S_E}$ respectively. In figure 2.11(a) and (b) we plot the ordered eigenvalues of $\mathbf{S_H}$ and $\mathbf{S_E}$. In both cases there are two dominant eigenvalues (corresponding to the plane wave sources), but the patterns of the smaller eigenvalues are quite different for the two matrices. For $\mathbf{S_H}$ the eight smallest eigenvalues break into two groups - two intermediate eigenvalues (the third and fourth largest) are roughly an order of magnitude larger than the six smallest eigenvalues which are all of comparable size. For $\mathbf{S_E}$ there is no clear break into distinct groups and the variation in the size of the smallest eigenvalues is much greater.

The eigenvectors associated with the intermediate eigenvalues of $\mathbf{S_H}$ for periods of 1000 and 3000 seconds are plotted in figure 2.12. For the third

eigenvector $\mathbf{v_3}$ at both periods the components of B_x and B_y vary linearly along the array transect. The vectors $\mathbf{v_3}$ thus represent gradients of the horizontal magnetic fields. The results for the fourth eigenvector $\mathbf{v_4}$ are somewhat less clear, but the plotted fields for $\mathbf{v_4}$ contain a substantial gradient component, particularly at 3000 seconds.. Note that $\mathbf{v_3}$ and $\mathbf{v_4}$ correspond to gradients in two roughly perpendicular source polarizations. This result is not particularly surprising - a natural refinement of the uniform plane wave source model would allow for gradients in the source field components in addition to the uniform fields. For a linear array, we can observe gradients only in the direction of the array transect. These can occur independently for each component B_x and B_y; hence the two gradient terms observed are exactly what we would expect. Note that this result is unusually clear. For arrays with more uneven station spacings, or with one or more unusually noisy channels, the third and fourth eigenvectors are difficult to associate directly with gradients (although some gradient component is usually evident).

This result suggests that the dimension of the assumed external source space Φ, and hence of the response space $\boldsymbol{R}$ should be expanded to treat the gradient terms as signal. Since the relationship between horizontal gradients and vertical fields can be used to estimate an equivalent 1-d MT impedance (Schmuker, 1970; Kuckes, 1973; Lilley *et al.*, 1976; Jones, 1980; Kuckes *et al.*, 1985) the gradients can clearly be used to obtain geologically interesting information. We will explore this idea further in Chapters 3 and 4.

The uniformity of the six smallest eigenvalues of $\mathbf{S_H}$ indicates a relatively uniform source of noise is present in the magnetic field measurements (see also the plateau in the ordered eigenvalues in figure 2.3). We interpret this as being indicative of a magnetometer system noise level which is at a relatively constant level at all stations. In contrast the variation in size of the smaller electric field eigenvalues suggests that the telluric system noise is highly variable from site to

site. This may well indicate that there is a great deal of interstation variability in electrode noise levels and that electrode noise dominates the residuals from the plane wave model fit. Note however, that as a result of surface distortion effects electric field signal amplitudes vary greatly both between stations and between polarizations. Features in the electric field SDM arising from source effects (such as gradients) will also be distorted and of variable amplitude from station to station. This can contribute to the variability in the smaller eigenvalues.

As a result of this interstation variability in the amplitude of both signal and noise the smaller eigenvalues and associated eigenvectors of $\mathbf{S_E}$ are difficult to interpret directly. Nevertheless we can demonstrate here that source related effects are probably significant in the electric field measurements also. We can estimate the fraction of power in the magnetic (or electric) fields which is inconsistent with the plane wave source assumption in terms of the ordered eigenvalues γ_i, $i = 1, 10$ of $\mathbf{S_H}$ (or $\mathbf{S_E}$). Specifically we define

$$\sigma_{\mathbf{H}}^2 = \left[\frac{1}{8} \sum_{i=3}^{10} \gamma_i \right] \left[\frac{(\gamma_1 + \gamma_2)}{10} \right]^{-1} \qquad (2.5.3)$$

with $\sigma_{\mathbf{E}}^2$ defined similarly in terms of the ordered eigenvalues of $\mathbf{S_E}$. The relative noise parameters $\sigma_{\mathbf{E}}^2$ and $\sigma_{\mathbf{H}}^2$ represent the average total noise power (i.e. the power in the data which is not consistent with the plane wave source assumption) in the electric and magnetic fields normalized by average signal power. Note that these estimates of noise power are only roughly correct since some noise power will be present in γ_1 and γ_2. This can be refined easily, but for the purposes of this discussion these simple estimates of noise power are sufficient.

In figure 2.11(a) we plot $\sigma_{\mathbf{H}}^2$ and $\sigma_{\mathbf{E}}^2$ for five periods from 100 to 10000 seconds. The plot shows that the relative noise for the magnetic fields decreases with increasing period, while the relative noise increases with increasing period for the electric fields. This result is clearly relevant only to the particular types of

instruments used in this array. For instance, the high relative noise level in the magnetic fields at short periods (i.e. hundreds of seconds) reflects the relatively high system noise levels for EDA fluxgate magnetometers at these periods.

We will now demonstrate that at longer periods a substantial fraction of the misfit between the electric and magnetic fields is not due to local noise. Instead this misfit is dominated by a source of 'noise' which is coherent across the array. Note that with the definition of (2.5.3) the relative noise parameters $\sigma_{\mathbf{H}}^2$ and $\sigma_{\mathbf{E}}^2$ include all deviations from the plane wave source assumption, including both local noise (i.e. instrumental noise plus local cultural noise which is incoherent between stations) and the horizontal magnetic field gradients discussed above. The parameters $\sigma_{\mathbf{H}}^2$ and $\sigma_{\mathbf{E}}^2$ thus represents an *upper bound* on the (relative) local noise levels and the sum $\sigma_{\mathbf{H}}^2 + \sigma_{\mathbf{E}}^2$ gives an upper bound on the total (electric plus magnetic) noise which is of a strictly local nature. The actual level of local noise will in general be smaller than this.

Using electric and magnetic field data from each station separately we can estimate the impedance tensors for each station using standard least squares transfer function estimation procedures. In the standard way, we obtain an estimate of the fractional misfit between predicted and observed electric field components for each station - i.e. $\eta_{ij}^2 = 1 - R_{ij}^2$ where R_{ij}^2 is the multiple coherence for the j^{th} electric field component ($j = 1, 2$) at the i^{th} station. If all of the misfit between electric and magnetic fields is due to noise of a strictly local nature then η_{ij}^2, $j = 1, 2$ should represent the total (electric + magnetic) fractional noise powers at the i^{th} station. Let η^2 be the average of the η_{ij}'s weighted by the electric field signal power in the j^{th} component at the i^{th} station. If all of the misfit between the electric and magnetic fields fit to a single station transfer function were due to local noise we can show that we would have $\eta^2 < \sigma_{\mathbf{H}}^2 + \sigma_{\mathbf{E}}^2$.

In figure 2.12) we plot $\sigma_H^2 + \sigma_E^2$ along with η^2. For periods beyond 1000 seconds η^2 exceeds $\sigma_H^2 + \sigma_E^2$; at 8000 seconds η^2 is larger than $\sigma_H^2 + \sigma_E^2$ by a factor of roughly five. This implies that for fitting MT impedances at these periods a significant fraction of the noise cannot be of a local nature. While some of this non-local noise could represent large scale cultural noise or temperature effects on the instruments, we will show with a more careful analysis in chapter four that most of this non-local noise is source related.

The analysis of the noise structure for the EMSLAB MT array described here is rather tentative and exploratory. Methods which we will develop in the following chapters will allow us to make more definitive statements about the character of this noise and its relation to source complications. The results presented do however give a feeling for some of the complications which should be allowed for in any statistical estimation scheme. We will use this as guidance in the development of our statistical model. These results also give some indication of the power of multivariate methods. With a univariate transfer function estimation approach the misfit is characterized by a single number (the mean square residual, say) which will in general include instrumental and cultural measurement noise as well as effects due to violations of model assumptions. With the multivariate approach these sources of noise can be distinguished.

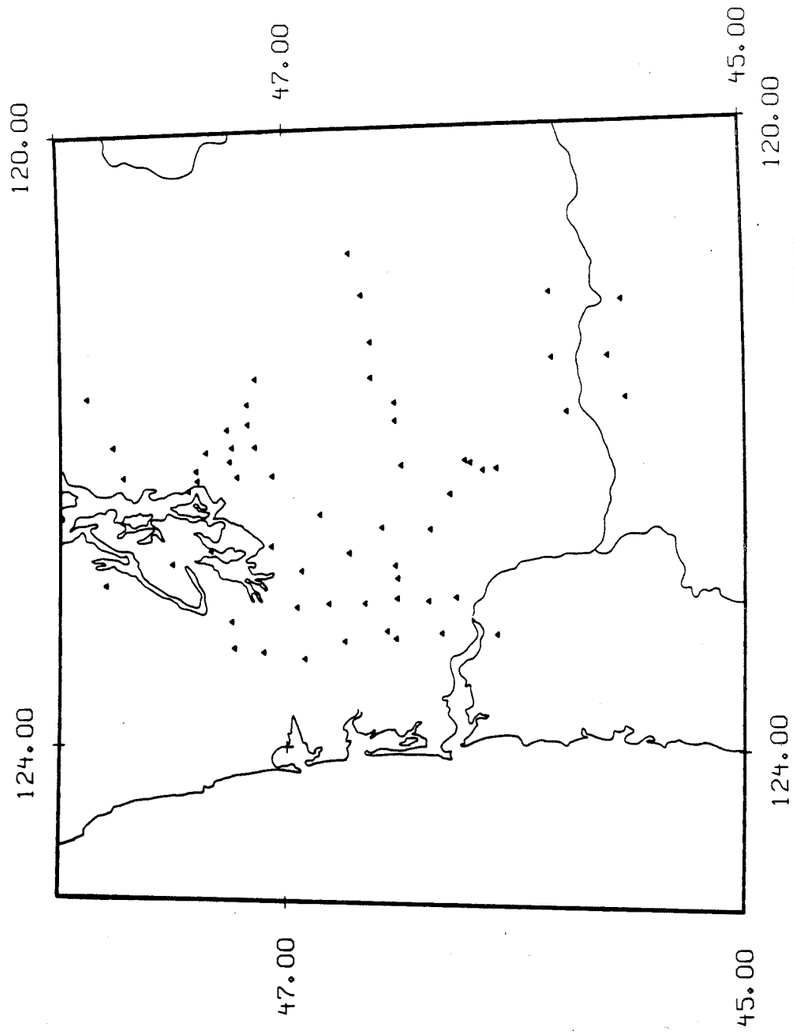


Figure 2.1 Map of station locations for small MV arrays discussed in text. The arrays consisted of a maximum of five simultaneously operating station. Arrays have been overlapped so that all stations may be connected together.

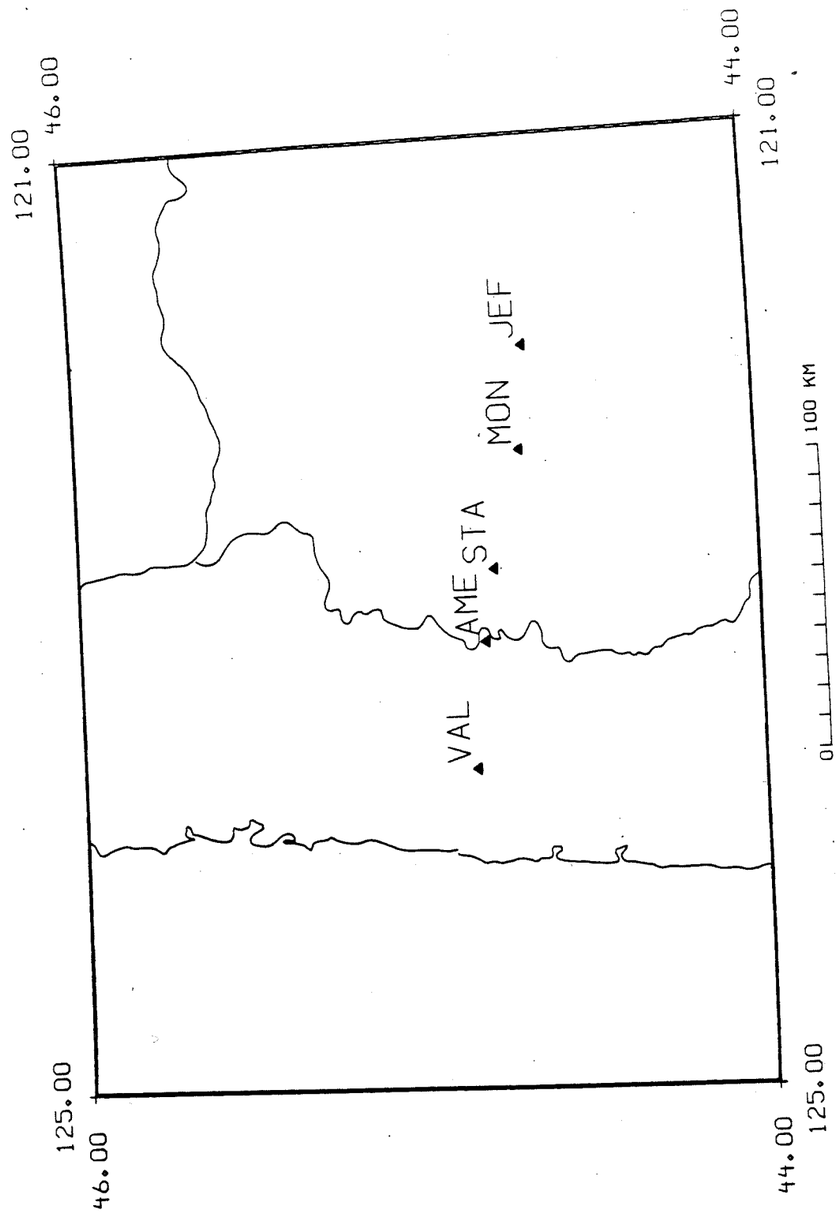


Figure 2.2 Map of EMSLAB long period MT line.

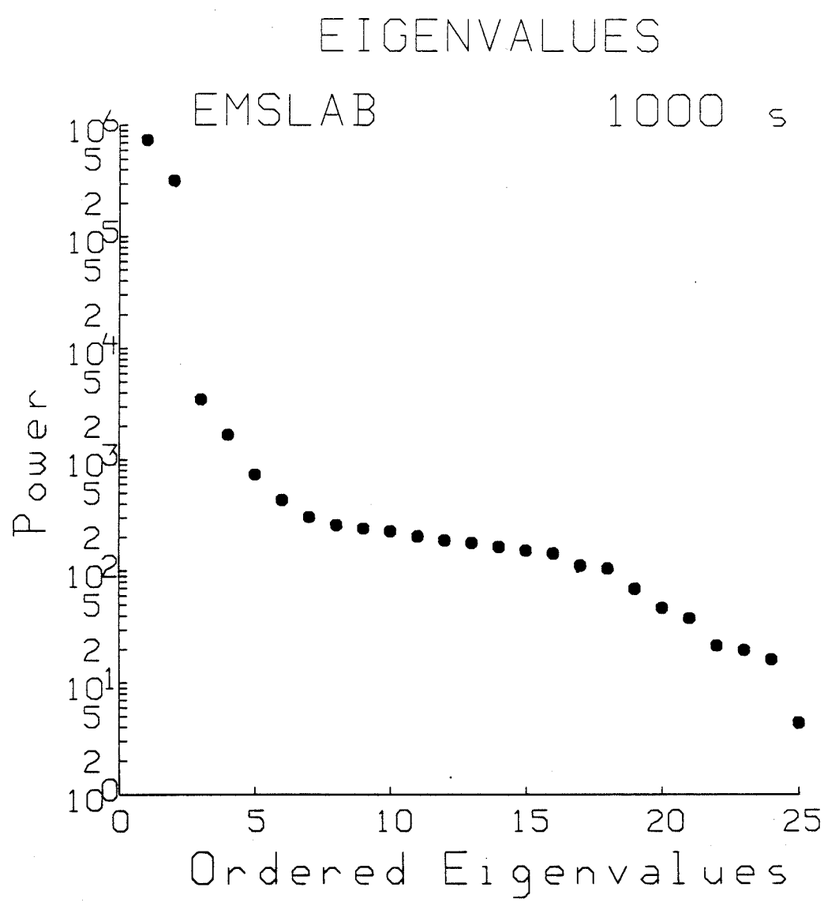


Figure 2.3 Ordered eigenvalues for the 25×25 SDM S from the five station EMSLAB MT profile: period = 1000 *s*.

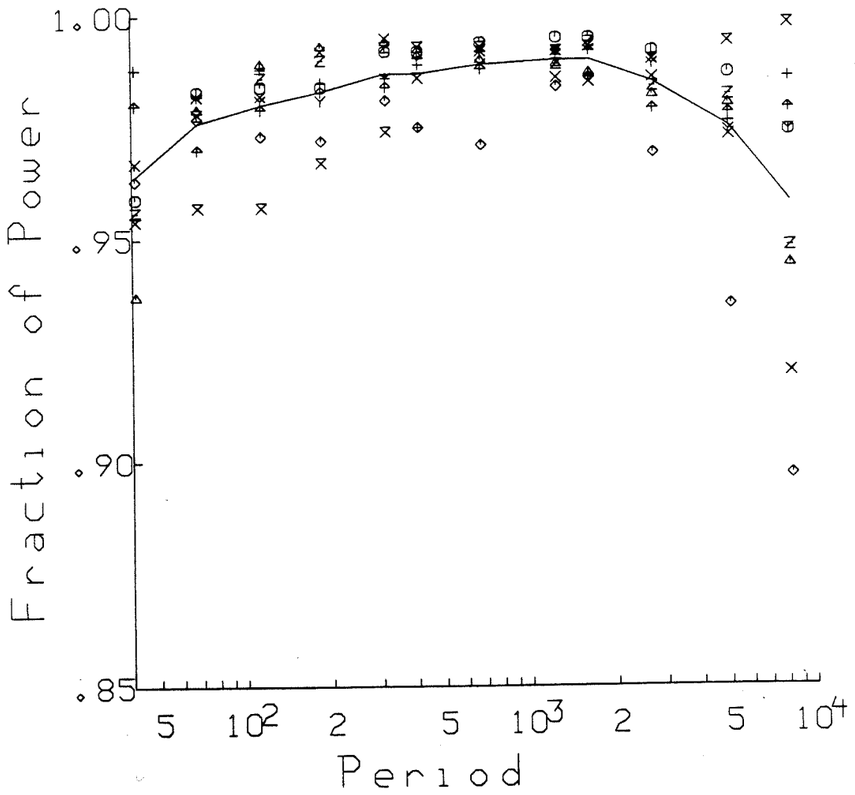


Figure 2.4 Fraction of total power in first two eigenvalues (R^2) *vs.* frequency: results for nine 3 component MV arrays. Solid line is average value of R^2. Except for the longest period, R^2 is always above .95, and very often it is above .99.

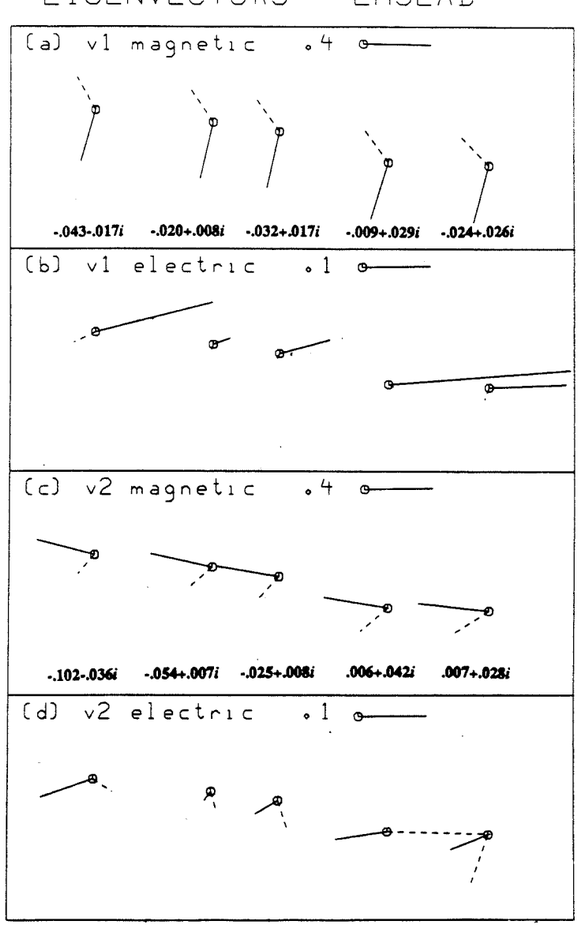


Figure 2.5 Dominant eigenvectors for EMSLAB MT profile: period = 1000 *s*. The eigenvector is of unit magnitude. Scales for magnetic and electric field vectors are given at the top center for each figure.

(a) $\mathbf{v}_1$ horizontal magnetic fields
(vertical field values are printed under station location).
(b) $\mathbf{v}_1$ electric fields
(c) $\mathbf{v}_2$ horizontal magnetic fields
(vertical field values are printed under station location).
(d) $\mathbf{v}_2$ electric fields

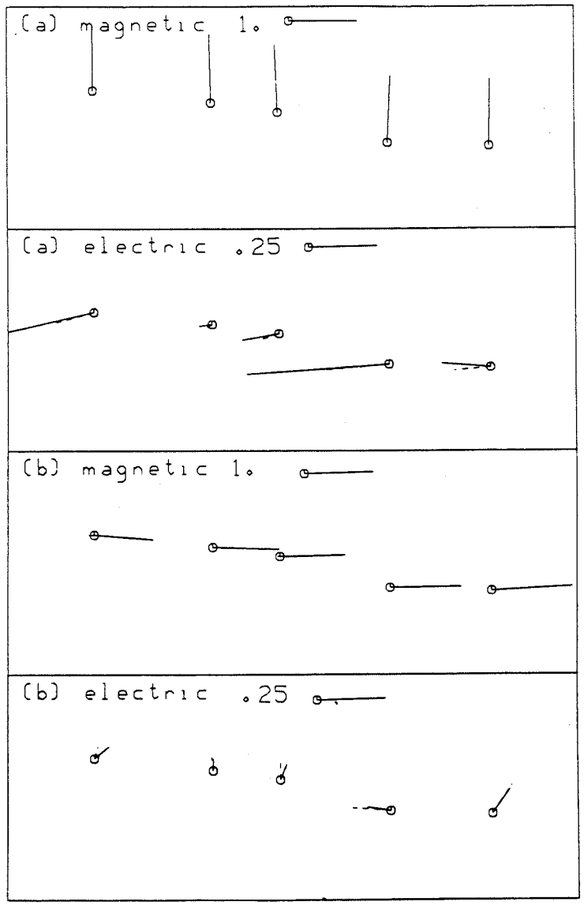


Figure **2.6** Magnetic and electric fields corresponding to linear polarizations of unit magnitude for EMSLAB MT profile. For this plot, the average fields are assumed to be 'normal'. Scale is at top center of each figure. Period is 1000 *s* as in figure 2.5. The (25-dimensional) vectors plotted here are linear combinations of the vectors plotted in figure 2.5.

(a) north-south magnetic field polarization - magnetic fields
(b) north-south magnetic field polarization - electric fields
(c) east-west magnetic field polarization - magnetic fields
(d) east-west magnetic field polarization - electric fields

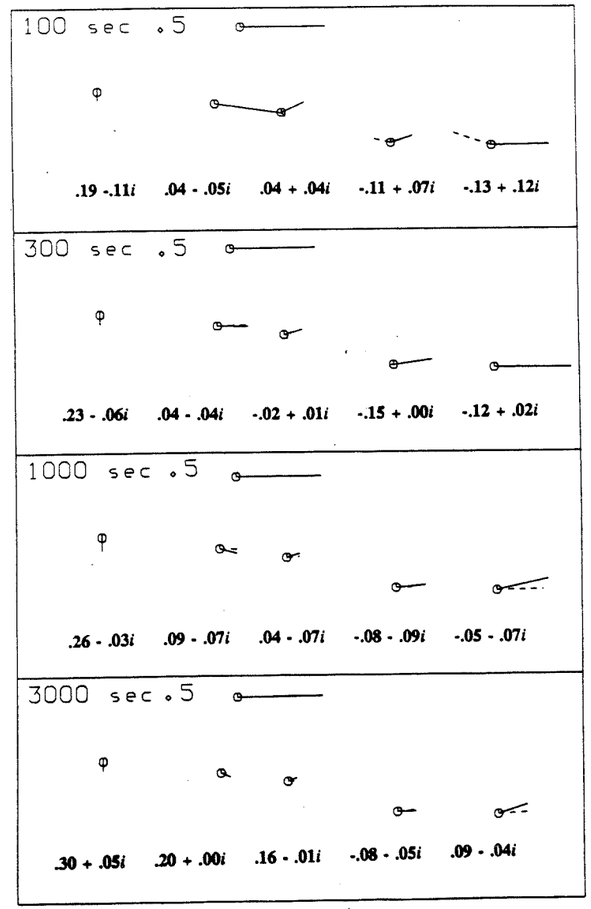


Figure **2.7** Anomalous horizontal magnetic fields for EMSLAB profile: (a) 100 *s* (b) 300 *s* (c) 1000 *s* (d) 3000 *s*. Vertical field values are printed under station location. Scale is at upper right on figure (a). A significant anomaly (40% of average field values) is seen at 100 seconds at station AME in the center of the Willamette Valley. A deeper anomaly is seen at the eastern-most station. This anomaly peaks at around 300 *s* and is still present at 3000 *s*. These conductivity anomalies are also seen in the vertical fields, although at longer periods these fields are dominated by the coast effect.

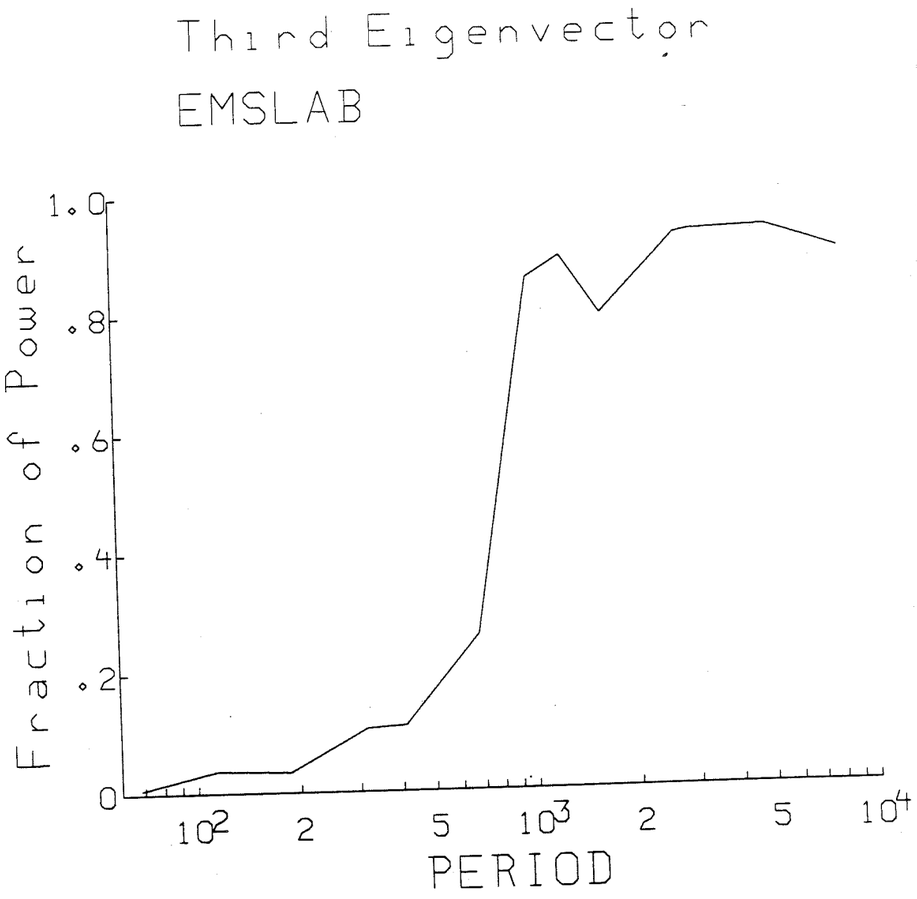


Figure 2.8 Fraction of power in third eigenvector $\mathbf{v}_3$ which is parallel to source Z vector $\mathbf{u}_z$ for EMSLAB profile.

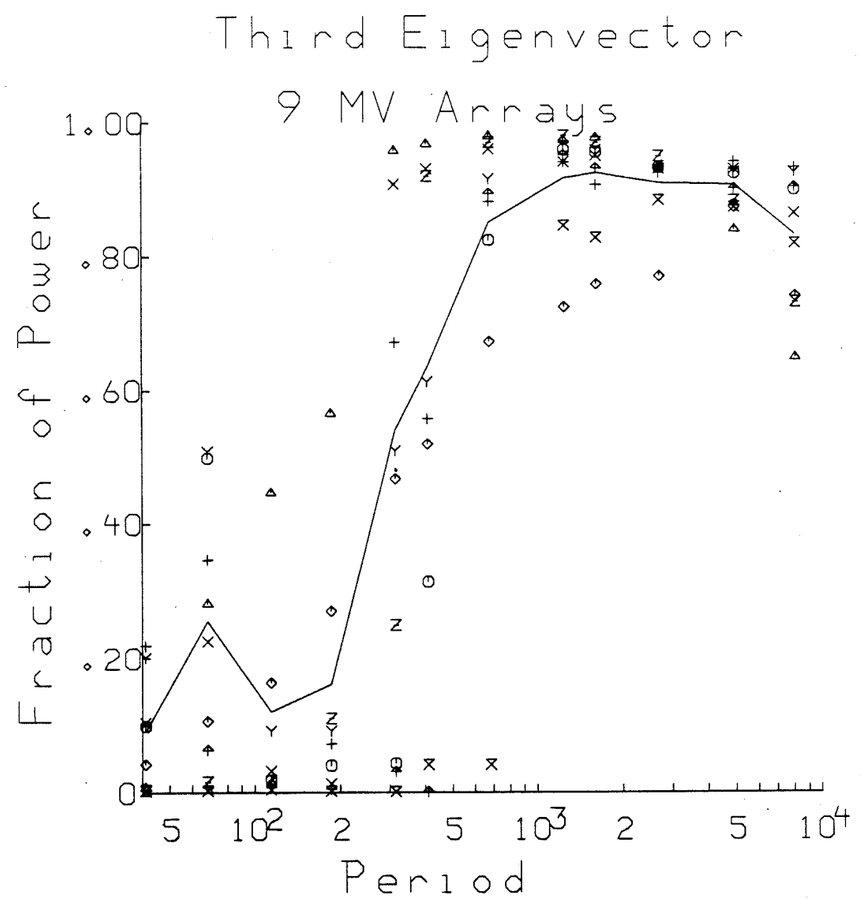


Figure 2.9 Fraction of power in third eigenvector $\mathbf{v}_3$ which is parallel to source Z vector $\mathbf{u}_z$: summary of results for 9 three component MV arrays. Solid line is average.

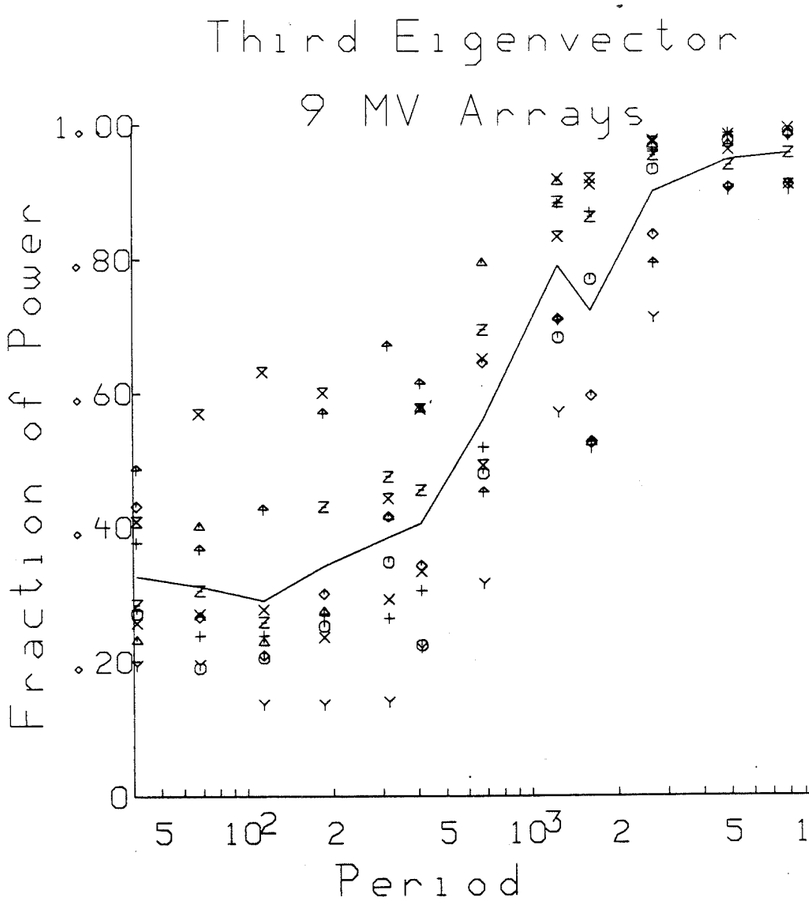


Figure 2.10 Fraction of residual power in third eigenvalue $\gamma_3 \left[\sum_{i=3}^{10} \gamma_i\right]^{-1}$: summary of results for 9 three component MV arrays.

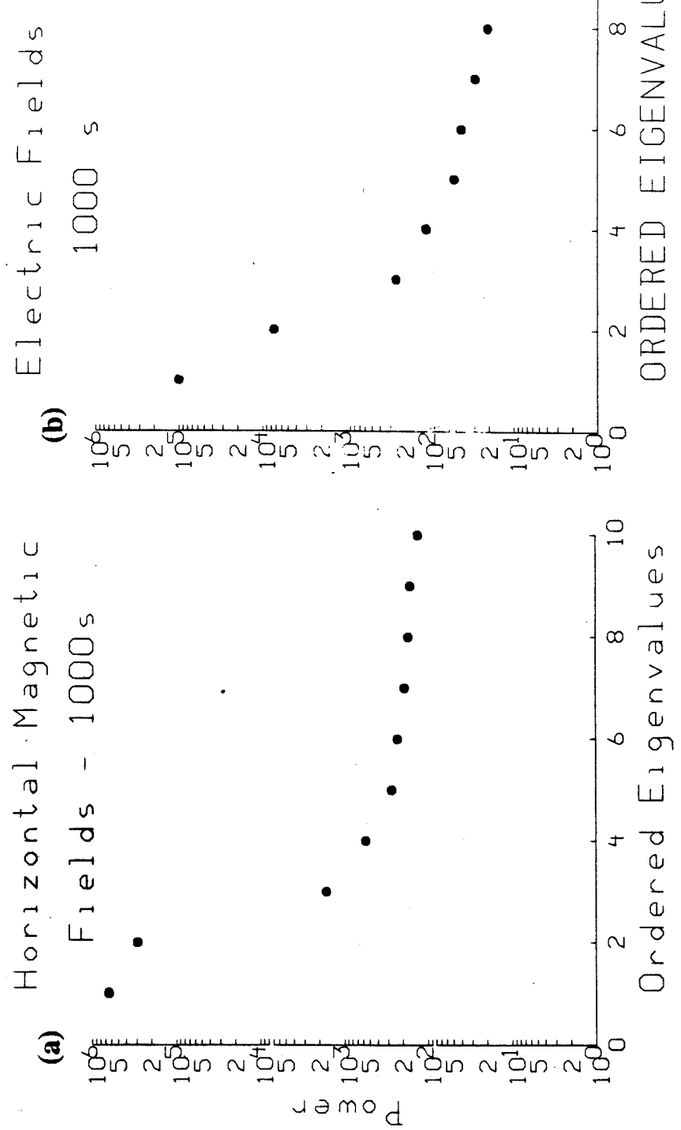


Figure 2.11 Ordered eigenvalues for 10×10 submatrices $\mathbf{S_H}$, $\mathbf{S_E}$ for EMSLAB MT profile: (a) horizontal magnetic fields only $\mathbf{S_H}$ (1000 *s*) (b) electric fields only $\mathbf{S_E}$ (1000 *s*).

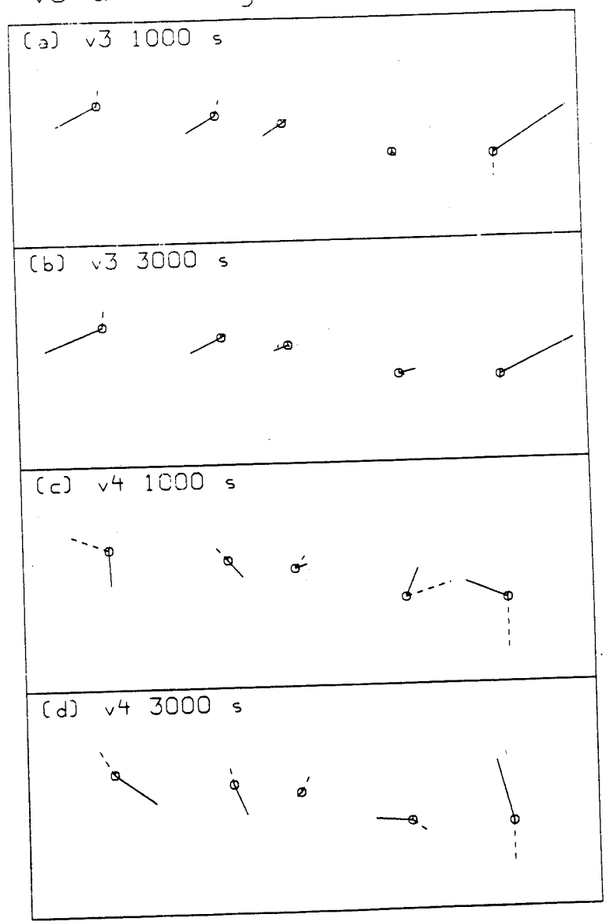


Figure 2.12 At longer periods the eigenvectors associated with the third and fourth largest eigenvalues of the horizontal magnetic field spectral density matrix S_H can be associated with gradients in the source fields. This result is clearest for v_3.

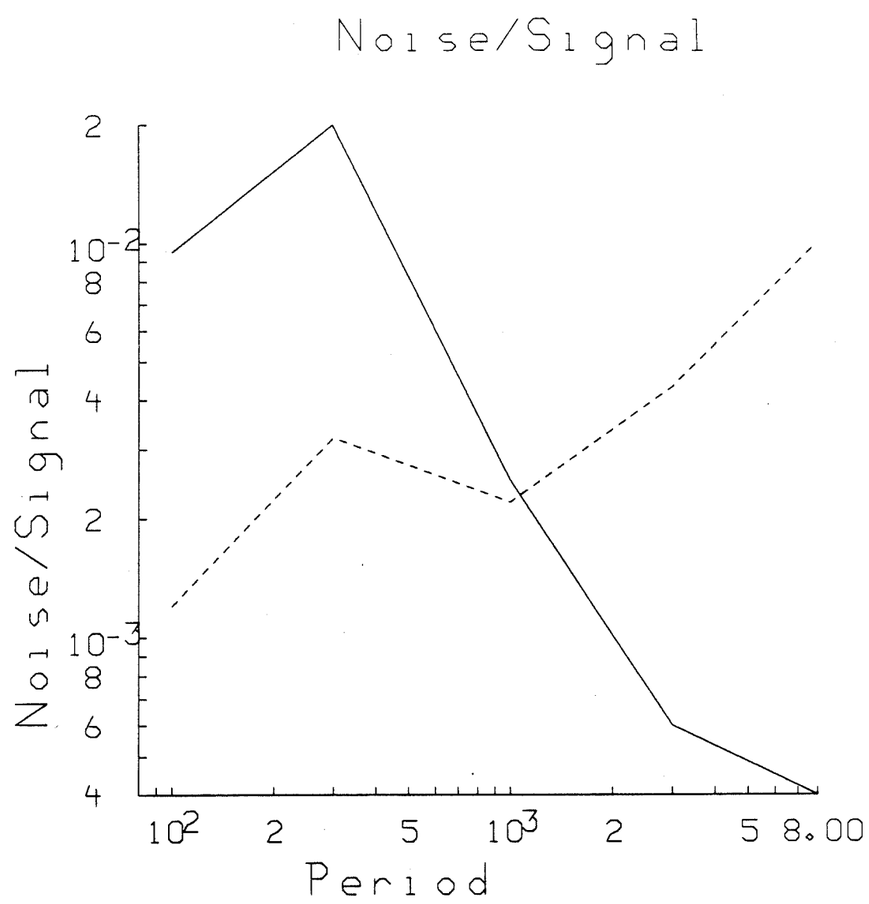


Figure 2.13 Noise-to-signal ratios for magnetic and electric fields from EMSLAB MT profile (σ_H^2 and σ_E^2) vs. period. Solid line is for magnetic fields, dashed line is for electric fields.

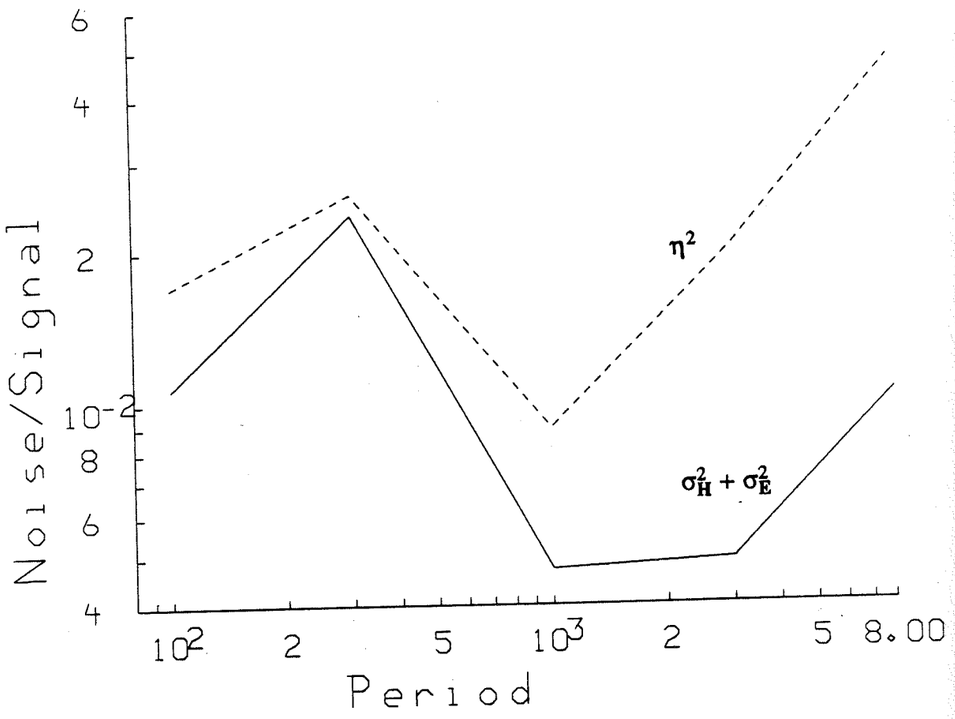


Figure 2.14 *Upper bound* on local noise (normalized by signal) $\sigma_H^2 + \sigma_E^2$ plotted with actual average fractional misfit η^2 computed from single station results. At longer periods, the total misfit is much greater than the upper bound on local noise, indicating that a significant fraction of the 'noise' is not local; we will relate this to source effects in chapter four.

Chapter 3

A Random Source Model

In Chapter 2 we developed the notion of the response space and related this idea to the usual transfer functions used in geomagnetic data analysis. We found that a rigorous justification of these methods depended on the assumption that the external source potentials were restricted to a space of finite dimension p. In fact, for any realistic situation, the set of all realizable source potentials Φ, and hence the response space $\boldsymbol{R}$, will be an infinite dimensional space. The transfer function/response space approach works in practice because this infinite dimensional space can often be well approximated by a space of low dimension. The examples considered at the end of Chapter 2, where we demonstrated that the data vectors from a series of small MV and MT arrays could be well approximated by elements of a two dimensional (plane wave source) response space, provide a graphic illustration of this idea. Our preliminary analysis of these arrays has also demonstrated that there are additional coherent features in the data which are not consistent with the simple plane wave source interpretation. We have interpreted these features in terms of violations of the plane wave source assumption.

To better understand how deviations from the assumed finite dimensional source space effect the data, we now drop this assumption and consider more general stochastic models for randomly varying external sources. In general, we will assume that the source potentials ϕ are realizations of a random spatial process and study the statistical properties of the magnetic and electric fields observed at a finite number of points $\mathbf{x}_i$, $i = 1, n$. Although we can make some general qualitative statements about the most general case (three dimensional conductivity, arbitrary source distributions), to be quantitative we will have to make very restrictive assumptions. Specifically, we will for the most part assume that the conductivity is one dimensional (1-d), varying with depth only, and that the sources are

spatially homogeneous. This model will allow us to compute the second moments of the random electromagnetic fields $\mathbf{F}(\mathbf{x})$ - i.e. the spatial covariance tensor $\mathbf{K}(\mathbf{x},\mathbf{x}') = E(\mathbf{F}(\mathbf{x})\mathbf{F}(\mathbf{x}')^*)$. The elements of the expected SDM $\mathbf{\Sigma} = E(\mathbf{S})$, for synthetic arrays consisting of stations in any given locations, can in turn be computed from $\mathbf{K}(\mathbf{x},\mathbf{x}')$. We can thus use the model to explore the behavior of the eigenvalues and eigenvectors of $\mathbf{\Sigma}$ under assumptions about sources that are somewhat more realistic than the finite dimension source assumption considered in Chapter 2.

We will show that most of the features we observed in our examples at the end of Chapter 2 can be readily explained by this simple model. Ultimately, we will see that this model provides strong justification for our response space formulation, particularly for arrays which are small compared to typical source length scales. For such small arrays we will see that the eigenvalues break into distinct, well separated clusters. The first cluster consists of two eigenvalues, and the associated eigenvectors correspond approximately to plane wave sources. The second cluster consists of three eigenvalues. For regular arrays with uniform station spacing these eigenvectors correspond to a set of canonical horizontal field gradients. In theory, there will be additional clusters of eigenvalues; in practice these will be lost in the background noise.

We will also see, that for arrays of fixed spatial extent, the features of the first few eigenvalues and eigenvectors of $\mathbf{\Sigma}$ are insensitive to the density of sampling stations. Even with a large number of stations the same clusters of eigenvalues occur, and the associated eigenvectors correspond approximately to the same pattern of fields sampled on a denser grid. It is interesting to consider the limiting case where the fields are sampled at all points in a bounded region S. In this limiting case the analogue of the matrix $\mathbf{\Sigma}$ is just the covariance tensor $\mathbf{K}(\mathbf{x},\mathbf{x}')$, and the matrix eigenvalue problem is replaced by the integral equation

analogue

$$\int_S \mathbf{K}(\mathbf{x},\mathbf{x}')\mathbf{v}(\mathbf{x}')d\mathbf{x}' = \gamma\mathbf{v}(\mathbf{x}) \qquad (3.0.1)$$

Using standard results from the theory of linear operators, we will show that the eigenvalues of (3.0.1) form a discrete set. Furthermore, the eigenvectors associated with positive eigenvalues correspond to fields in the region S which are realizable (i.e. are consistent with the physics and the electrical conductivity distribution). Any observable fields can be written as a linear combination of these positive eigenvalue modes. In the context of our random source model, the coefficients of the expansion are uncorrelated random variables whose variance is given by the corresponding eigenvalue. The eigenvalues of (3.0.1) thus form a sort of discrete spatial power spectrum for the fields restricted to a bounded region.

Our simple 1-d model shows that for small arrays most all of the power is in the modes associated with the first few eigenvalues; for this case a finite dimensional response space model is thus a good approximation. The considerations outlined above indicate that this property should hold in general for geomagnetic arrays, although the interpretation of the eigenvectors in terms of known simple sources may be more difficult than in the small array case considered here. This is an interesting problem for further study.

3.1: A 1-d Random Source Model
Horizontal Magnetic Fields

In this section we develop a very simple random source model which will allow us to calculate the spatial covariance tensor $\mathbf{K}(\mathbf{x},\mathbf{x}')$. To make the problem tractable we assume that the electrical conductivity is 1-d, varying with depth only. We also assume that at a fixed frequency ω the external source potential on

the surface $z = 0$ is a realization of a spatially homogeneous complex random field ϕ_e. By spatial homogeneity we mean that the statistical properties of the fields are invariant under translations of the coordinate system (cf. Ripley, 1981). Spatial homogeneity is thus the random spatial process analogue of the familiar stationarity assumption for time series. This is a strong assumption which can only be approximately valid for the real earth since the distribution of ionospheric and magnetospheric current sources is strongly dependent on geomagnetic latitude. The consequences of this non-homogeneity will be discussed briefly below. For now we note that this simple model has pedagogic value and that, over spatial scales of a few hundred kilometers at geomagnetic mid-latitudes, the assumption of homogeneity may be expected to be at least approximately valid.

Since we have also assumed that the electrical conductivity is 1-d, it is clear that the statistical properties of the total potential $\phi = \phi_e + \phi_i$ defined on the surface $z = 0$ will also be translationally invariant; thus the total potential ϕ will also be a spatially homogeneous complex random field. The second moment properties of ϕ are given by the spatial covariance function

$$K_\phi(\mathbf{x}) = \boldsymbol{E}(\ \phi(\mathbf{x})\phi(0)^*\ ) \tag{3.1.1}$$

where $\boldsymbol{E}$ represents the expectation operator. Note that without loss of generality we may assume that $\boldsymbol{E}(\phi(\mathbf{x})) \equiv 0$ since only the gradients of the potential are physically meaningful. Note also that homogeneity implies

$$\boldsymbol{E}(\ \phi(\mathbf{x})\phi(\mathbf{x}')^*\ ) = K_\phi(\mathbf{x}-\mathbf{x}')$$

so K_ϕ can be used to compute the covariance between values of the random potential at any two points.

Assuming this model, let the magnetic (and perhaps the electric) fields be observed at n locations $\mathbf{x}_i$, $i = 1, n$, for a series of N independent realizations of the random external sources. As in Chapter 2 let $\mathbf{X}_k$, $k = 1, N$ be the fundamental

frequency domain data vectors for these observations, and form the SDM $\mathbf{S}$ as in (2.4.2). Our goal here, and in the next section, is to show how the expected value of $\mathbf{S}$

$$\boldsymbol{\Sigma} = \boldsymbol{E}(\mathbf{S}) = \boldsymbol{E}(\ \mathbf{X}_1\mathbf{X}_1^*\ )$$

can be computed in terms of the spatial covariance function $K_\phi(\mathbf{x})$ and the wave number dependent impedance of (2.1.9).

Initially we consider horizontal magnetic fields B_x, B_y only. We will refer to the SDM constructed from the components B_x, B_y as $\mathbf{S_H}$. Individual realizations of the fields B_x, B_y can be computed directly as derivatives of the random source potentials and $\boldsymbol{\Sigma}_\mathbf{H} = \boldsymbol{E}(\mathbf{S_H})$ can be computed in terms of K_ϕ alone. Specifically, for $u, v = x, y$ let

$$K_{uv}(\mathbf{x}) = \boldsymbol{E}\left[B_u(\mathbf{x})B_v^*(0)\right]$$

Then for any set of station locations $\mathbf{x}_i$, $i = 1, n$ the elements of $\mathbf{S_H}$ are of the form $K_{uv}(\mathbf{x}_i-\mathbf{x}_j)$, and, as we will now show

$$K_{uv}(\mathbf{x}) \ = \ -\frac{\partial^2}{\partial u \partial v} K_\phi(\mathbf{x})\Big|_\mathbf{x} \qquad u, v = x, y \tag{3.1.2}$$

Equation (3.1.2) is most easily derived by Fourier transforming into the wavenumber domain. In general we will denote the Fourier transform of the complex valued function $f(\mathbf{x})$ defined on $\mathbf{R}^2$ by $\tilde{f}(\boldsymbol{\nu})$

$$\tilde{f}(\boldsymbol{\nu}) = \iint d^2\mathbf{x}\, f(\mathbf{x})\, \exp(-i\mathbf{x}\cdot\boldsymbol{\nu}) \tag{3.1.3a}$$

where $\boldsymbol{\nu} = (\nu_x,\nu_y)$ is the wave vector. Note that with this definition of the Fourier transform the inverse transform has the form

$$f(\mathbf{x}) = (2\pi)^{-2}\iint d^2\mathbf{x}\, \tilde{f}(\boldsymbol{\nu})\, \exp(+i\mathbf{x}\cdot\boldsymbol{\nu}) \tag{3.1.3b}$$

These definitions of the Fourier transform are consistent with the usual convention

used in the electromagnetic induction literature but they are different from the convention used in the spectral analysis of stochastic processes; as a result we will have a few extra factors of 2π in some of the formulae given here. The assumption of stationarity implies that we may write the random potential as the Fourier transform of a wavenumber domain orthogonal increments process (e.g. Doob, 1953) which, with the Fourier transform defined as in (3.1.3) takes the form

$$\phi(\mathbf{x}) = (2\pi)^{-2} \iint d\tilde{\phi}(\boldsymbol{\nu}) \exp(i\boldsymbol{\nu} \cdot \mathbf{x}) \tag{3.1.4}$$

where

$$E\left[d\tilde{\phi}(\boldsymbol{\nu}) d\tilde{\phi}^*(\boldsymbol{\nu}')\right] = (2\pi)^2 S_\phi(\boldsymbol{\nu}) \delta(\boldsymbol{\nu}-\boldsymbol{\nu}') d^2\boldsymbol{\nu} d^2\boldsymbol{\nu}'$$

where δ is the (2-dimensional) Dirac delta function and $S_\phi(\boldsymbol{\nu})$ is the spatial power spectral density of the random potential ϕ. With the definitions given here a simple calculation shows that the power spectrum is just the Fourier transform of the spatial covariance function $S_\phi = \tilde{K}_\phi$ as in the time series case.

Then using (3.1.4) and noting

$$B_u(\mathbf{x}) = \frac{\partial}{\partial u}\phi\Big|_{\mathbf{x}} \qquad u, v = x, y$$

we calculate for $u, v = x, y$

$$E\left[\frac{\partial\phi}{\partial u}\Big|_{\mathbf{x}} \frac{\partial\phi^*}{\partial v}\Big|_0\right] =$$

$$E\left[\left[(2\pi)^{-2}\iint i\nu_u \exp(i\boldsymbol{\nu}\cdot\mathbf{x}) d\tilde{\phi}(\boldsymbol{\nu})\right]\left[(2\pi)^{-2}\iint -i\nu_v' d\tilde{\phi}^*(\boldsymbol{\nu}')\right]\right]$$

$$= (2\pi)^{-4}\iiiint \nu_u\nu_v' \exp(i\boldsymbol{\nu}\cdot\mathbf{x}) E\left[d\tilde{\phi}(\boldsymbol{\nu}) d\tilde{\phi}^*(\boldsymbol{\nu}')\right]$$

$$= (2\pi)^{-2}\iint \nu_u\nu_v' \exp(i\boldsymbol{\nu}\cdot\mathbf{x}) S_\phi(\boldsymbol{\nu}) d^2\boldsymbol{\nu}$$

$$= -\frac{\partial^2}{\partial u \partial v}\left[(2\pi)^{-2}\iint S_\phi(\boldsymbol{\nu})\exp(i\boldsymbol{\nu}\cdot\mathbf{x})\, d^2\boldsymbol{\nu}\right] = -\frac{\partial^2}{\partial u\partial v} K_\phi(\mathbf{x})$$

which establishes (3.1.2).

Remark: We may use the symmetries in the definition of $K_{uv}(\mathbf{x})$ to show that

$$K_{uv}(\mathbf{x}) = K_{vu}^*(-\mathbf{x}) \tag{3.1.5}$$

while (3.1.2) implies

$$K_{uv}(\mathbf{x}) = K_{vu}(\mathbf{x}) \tag{3.1.6}$$

Together (3.1.5) and (3.1.6) imply

$$K_{xy}(0) = K_{xy}^*(0)$$

Thus, the assumption of spatial homogeneity implies that $K_{xy}(0)$ must be real. Since an estimate of $K_{xy}(0)$ can be computed from data at a single station it is actually possible to prove that the fields *aren't* spatially homogeneous from data at a single station.

To apply this model we must choose a specific form for the spatial covariance function K_ϕ. This function must be a positive definite function (i.e. the Fourier transform must be strictly positive; cf. Miller, 1974; Ripely, 1981) to be a valid spatial covariance. For the calculations described below we have chosen K_ϕ to have the simple form

$$K_\phi(\mathbf{x}) = \sigma_0^2 \exp\left[\frac{-\|\mathbf{x}\|^2}{r_0^2}\right] \tag{3.1.7}$$

In (3.1.7) the parameter σ_0 gives the total amplitude of the potential variations while r_0 gives a characteristic correlation length scale for the potentials.

A simple calculation shows for $u, v = x, y$

$$K_{uv}(\mathbf{x}) = -\frac{\partial^2}{\partial u\partial v} K_\phi(\mathbf{x})\Big|_{\mathbf{x}} = \frac{2}{r_0^2}\left[\delta_{uv} - \frac{2uv}{r_0^2}\right] K_\phi(\mathbf{x}) \tag{3.1.8}$$

Note that since the magnetic fields are the derivatives of the potential they are

rougher and have a characteristic correlation length scale somewhat shorter than r_0. As we see from (3.1.8) the correlation of the fields is reduced to zero at a distance $r_0/\sqrt{2}$.

3.2: A 1-d Random Source Model

Electric and Vertical Magnetic Fields

We now consider the incorporation of vertical magnetic and horizontal electric components into our random source model. To do this we use the space domain impedance filter approach described in Dmitriev and Berdichevsky (1979), hereafter DB. We will refer to many features of this approach in the sequel. We will thus outline the key results described in DB. This approach is based on the observation that for a given conductivity distribution in the region $z > 0$ the horizontal magnetic fields on the surface $z = 0$, together with the appropriate boundary condition at $z = \infty$, uniquely determine the magnetic and electric fields for $z \geq 0$. This, together with the linearity of the governing equations, implies that a linear operator relates the horizontal electric fields on the surface $z = 0$ to the horizontal magnetic fields on $z = 0$

$$\mathbf{E}(\mathbf{x}) = \iint d^2\mathbf{x}'\, \mathbf{G}_E(\mathbf{x},\mathbf{x}')\mathbf{B}_H(\mathbf{x}') \tag{3.2.1}$$

where the 2×2 tensor $\mathbf{G}_E(\mathbf{x},\mathbf{x}')$ is the kernel of the operator, which depends both on frequency and the conductivity distribution.

In the special case of 1-d conductivity which we consider here, the kernel takes the simpler form

$$\mathbf{G}_E = \begin{bmatrix} 0 & G_{xy} \\ G_{yx} & 0 \end{bmatrix}$$

where G_{xy} satisfies

$$G_{xy}(\mathbf{x},\mathbf{x}') = -\,G_{yx}(\mathbf{x},\mathbf{x}') = G_{xy}(\mathbf{x}-\mathbf{x}',0) = G_E(\mathbf{x}-\mathbf{x}') \tag{3.2.2}$$

The function G_E is called the spatial impedance filter. From (3.2.1) and (3.2.2) we see that the electric fields can be written as the convolution of the impedance filter with the horizontal magnetic fields.

$$E_x = G_E * B_y \qquad E_y = -\,G_E * B_x \tag{3.2.3}$$

where $*$ denotes convolution, i.e.

$$f * g(\mathbf{x}) = \iint d^2\mathbf{x}'\, f(\mathbf{x}-\mathbf{x}')\, g(\mathbf{x}')$$

Note that as we have assumed that the atmospheric conductivity is zero so that no currents flow across the surface $z = 0$, G_E is just the inverse spatial Fourier transform of the usual wavenumber dependent impedance $\zeta(\omega,\nu)$ (see DB):

$$G_E(\mathbf{x}) = (2\pi)^{-2} \iint d^2\nu\zeta(\omega,\nu)\exp(i\mathbf{v}\cdot\mathbf{x})$$

The exact form of G_E is derived for a half-space of constant conductivity σ in DB. The result is

$$G_E(\mathbf{x}) = \frac{-i\omega}{2\pi}\frac{\exp(ik\|\mathbf{x}\|)}{\|\mathbf{x}\|} \qquad k^2 = i\omega\sigma \qquad \text{Re } k > 0 \tag{3.2.4}$$

Here $k = \delta(1 + i)$ where δ is the skin depth. We thus see that the impedance filter goes to zero very rapidly beyond distances of a few skin depths. This implies that the electric fields are locally determined - values of the magnetic fields more than a few skin depths away will have very little effect on the electric fields.

Note also that the impedance filter for the half-space is radially symmetric. This will clearly be a general feature of the impedance filters for all 1-d conductivities (see DB). If the horizontal magnetic fields are expanded in a Taylor's series around a point $\mathbf{x}$ then the symmetry of the impedance filter implies that the electric fields at $\mathbf{x}$ are sensitive to only the even order terms. In particular, the local electric fields are not sensitive to linear gradients of the magnetic fields. This fact was used by DB to demonstrate that, at least for a 1-d earth, the MT

impedance determined from the ratio of local electric to magnetic fields will be the correct plane wave impedance even in the presence of large lateral variations in the sources, provided these variations are linear gradients of the fields. (See also the discussion in Mareschal, 1986.)

The vertical magnetic fields can also be obtained from the impedance filter formulation. Since from (2.1.1)

$$B_z = \frac{i}{\omega}(\nabla \times E)_z = \frac{i}{\omega}\left[\frac{\partial}{\partial x}E_y - \frac{\partial}{\partial y}E_x\right]$$

$$= -\frac{i}{\omega}\left[\frac{\partial}{\partial x}(G_E * B_x) + \frac{\partial}{\partial y}(G_E * B_y)\right] \qquad (3.2.5)$$

We can put (3.2.5) into a more revealing form by using the fact that the order of convolution and differentiation may be interchanged (e.g. Folland, 1976)

$$\frac{\partial}{\partial x}(f * g) = \frac{\partial f}{\partial x} * g = f * \frac{\partial g}{\partial x}$$

Then we may write (3.2.5) as

$$B_z = -\frac{i}{\omega}\left[\left[\frac{\partial}{\partial x}G_E\right] * B_x + \left[\frac{\partial}{\partial y}G_E\right] * B_y\right]$$

$$= -\frac{i}{\omega}\, G_E * \left[\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y}\right] \qquad (3.2.6)$$

From (3.2.5) we see that the vertical fields B_z can also be written as a vector convolution of the horizontal fields. The alternative formula (3.2.6) shows that B_z can be written as a convolution of the scalar field

$$\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y}$$

Furthermore, (3.2.6) shows that the convolution function is the spatial impedance filter G_E times a constant. This gives a neat formal justification for the spatial

gradient approach (Schmuker, 1970; Kuckes, 1973; Lilley, *et al.*, 1976) to the estimation of the 1-d MT impedance.

With the above results in hand we can calculate the elements of $\mathbf{\Sigma}$ corresponding to covariances involving electric and vertical magnetic fields. We first establish a general result which we subsequently apply. Let $\mathbf{f} = (f_1, f_2)$ be a complex vector valued spatially homogeneous random field with covariance tensor $K_{ij}(\mathbf{x}) = \boldsymbol{E}(f_i(\mathbf{x})\, f_j^*(0))$. Then for any sufficiently regular function G

$$\boldsymbol{E}\Big[(G * f_i)(\mathbf{x})\, f_j^*(0)\Big] = (G * K_{ij})(\mathbf{x}) \qquad (3.2.7a)$$

$$\boldsymbol{E}\Big[f_i(\mathbf{x})\, (G * f_j)^*(0)\Big] = (G^* * K_{ij})(\mathbf{x}) \qquad (3.2.7b)$$

These are established by a simple calculation

$$\boldsymbol{E}\Big[(G * f_i)(\mathbf{x})\, f_j^*(0)\Big] = \boldsymbol{E}\Big[\iint d^2\mathbf{x}'\, G(\mathbf{x}'-\mathbf{x})\, f_i(\mathbf{x}')\, f_j^*(0)\Big]$$

$$= \iint d^2\mathbf{x}'\, G(\mathbf{x}'-\mathbf{x})\boldsymbol{E}(f_i(\mathbf{x}')f_j^*(0)) = \iint d^2\mathbf{x}' G(\mathbf{x}'-\mathbf{x})\, K_{ij}(\mathbf{x}') = (G*K_{ij})(\mathbf{x})$$

This establishes (3.2.7a); (3.2.7b) is proved by a similar calculation.

Then consider the calculation of elements of $\mathbf{\Sigma}$ such as $E(E_u B_v)$ where $u, v = x, y$. To present a unified single formula we adopt the convention that $\tilde{u}$ will denote the index y when $u = x$ and the index x when $u = y$ (and similarly for $\tilde{v}$). We also define

$$\varepsilon(u,v) = \begin{cases} +1 & \text{if } u = v \\ -1 & \text{if } u \neq v \end{cases}$$

Then using (3.2.3) and (3.2.7a) we calculate

$$\boldsymbol{E}(E_u(\mathbf{x})B_v(0)) = \boldsymbol{E}\Big[\ \varepsilon(x,u)\ (G_E * B_{\tilde{u}})(\mathbf{x})\ B_v(0)\Big] = \varepsilon(x,u)\ (G_E * K_{\tilde{u}v})(\mathbf{x})$$

$$= \varepsilon(y,u)\, \frac{\partial^2}{\partial\tilde{u}\partial v}\, (G_E * K_\phi)\Big|_{\mathbf{x}} \qquad (3.2.8)$$

In a similar fashion, using (3.2.7b) in addition, we may calculate the covariances between two electric components at **x** and 0 as

$$E(E_u(\mathbf{x})E_v(0)) = \varepsilon(v,u)\,\frac{\partial^2}{\partial\tilde{u}\partial\tilde{v}}\,((G_E * G_E^*) * K_\phi)\Big|_{\mathbf{x}} \qquad (3.2.9)$$

The calculation of covariances involving vertical fields proceeds in a similar fashion using (3.2.6) and (3.2.7). We find for $u = x, y$

$$E(B_z(\mathbf{x})B_u(0)^*) = \sum_v \frac{i}{\omega}\frac{\partial^3}{\partial u\partial v^2}(G_E * K_\phi)\Big|_{\mathbf{x}} \qquad (3.2.10)$$

$$E(B_z(\mathbf{x})B_z(0)^*) = \sum_v\sum_u \frac{1}{\omega^2}\frac{\partial^4}{\partial u^2\partial v^2}((G_E * G_E^*) * K_\phi)\Big|_{\mathbf{x}} \qquad (3.2.11)$$

$$E(B_z(\mathbf{x})E_u(0)) = \frac{\varepsilon(u,y)\,i}{\omega}\sum_v \frac{\partial^3}{\partial\tilde{u}\partial v^2}((G_E * G_E^*) * K_\phi)\Big|_{\mathbf{x}} \qquad (3.2.12)$$

Using (3.1.2) and (3.2.8) - (3.2.12) we can calculate all of the entries in the matrix Σ for any set of station locations $\mathbf{x}_i$, $i = 1, n$. All of these quantities are calculated as derivatives of the convolution of the spatial impedance filter with the spatial covariance function of the random potential

$$F_1 = G_E * K_\phi \qquad F_2 = (G_E * G_E^*) * K_\phi$$

The computation of F_1 and F_2 for a specific choice of K_ϕ and G_E can be done in the wavenumber domain using the fact that convolution in the space domain is replaced by multiplication in the wavenumber domain. Thus, for instance

$$\tilde{F}_1(\mathbf{v}) = \tilde{G}_E(\mathbf{v})\,\tilde{K}_\phi(\mathbf{v}) = \zeta(\omega,\mathbf{v})\,S_\phi(\mathbf{v}) \qquad (3.2.13)$$

where $\zeta(\omega,\mathbf{v})$ is the usual wave number dependent impedance and $S_\phi(\mathbf{v})$ is the power spectrum of the random potential. F_1 can then be calculated as the inverse Fourier transform of the right hand side of (3.2.13). If the random potential is

assumed to be isotropic so that the spatial covariance K_ϕ is radially symmetric $K_\phi(\mathbf{x}) = K_\phi(r)$ (where $r = \|\mathbf{x}\|$), then $\tilde{F}_1(\mathbf{v})$ is also radially symmetric and the calculation of the inverse Fourier transform can then be reduced to the evaluation of a one-dimensional integral. With a limited amount of experimentation we were not able to find any (non-trivial) cases which allow for computation of F_1 and F_2 in closed form. We thus turn to methods of approximating these functions.

In general $\zeta(\omega,\mathbf{v})$ is defined as a function of depth z and satisfies the Ricatti equation (Rokityanski, 1982)

$$\frac{d}{dz}\zeta(\omega,\mathbf{v}) + \frac{(\mathbf{v}^2 - k^2)}{i\omega}\zeta(\omega,\mathbf{v})^2 = i\omega \qquad (3.2.14)$$

with

$k = (i\omega\mu\sigma)^{1/2} = (1 + i)/\delta$ where δ is the skin depth. Setting $\varepsilon = \mathbf{v}\delta$, expand the impedance in the parameter ε to obtain an approximation valid when ε is small. Write

$$\zeta(\omega,\mathbf{v}) = \zeta_0(\omega) + \varepsilon\zeta_1(\omega) + \varepsilon^2\zeta_2(\omega) + \cdots \qquad (3.2.15)$$

Substituting (3.2.15) into (3.2.14) and collecting terms in like powers of ε we find that

$$\frac{d}{dz}\zeta_i(\omega) = 0 \qquad \text{for } i \text{ odd}$$

With the boundary condition $\zeta(\omega,\mathbf{v}) \to 0$ as $z \to 0$ this implies that $\zeta_i(\omega) = 0$ for i odd. Thus we may generally write

$$\zeta(\omega,\mathbf{v}) = \zeta_0(\omega) + \zeta_2(\omega)\delta^2\mathbf{v}^2 + \cdots$$

where $\zeta_0(\omega)$ is the standard plane wave impedance.

As an explicit example consider the case of a half space of conductivity σ. Here we have

$$\zeta(\omega,\nu) = \frac{-i\omega}{(\nu^2 - k^2)^{1/2}} = \frac{\omega}{k}\,(\,1 - \nu^2/k^2)^{-1/2} = \zeta_0(\omega)\,(\,1 - \nu^2/k^2\,)^{-1/2}$$

$$= \zeta_0(\omega)\left[\,1 + \frac{1}{2}\frac{1}{2i}\nu^2\delta^2 + \frac{1\cdot 3}{2\cdot 4}\left(\frac{1}{2i}\right)^2\nu^4\delta^4 + \cdots\right] \qquad (3.2.16)$$

where we have used the Taylor series expansion for $(1 - \nu^2/k^2)^{-1/2}$ and $\nu^2/k^2 = \nu^2\delta^2/2i$ in the last step. For the half space, then, we can give an explicit form for the coefficients $\zeta_i(\omega)$. Specifically, the first few are

$$\zeta_2(\omega) = \frac{1}{4i}\,\zeta_0(\omega) \qquad \zeta_4(\omega) = -\frac{3}{32}\,\zeta_0(\omega)$$

With this series expansion for the wave number dependence of the impedance we can calculate approximations to the functions F_1 and F_2 for the the error function type spatial covariance function given in (3.1.5). First note that for this covariance the power spectrum has the similar form

$$S_\phi(\mathbf{v}) = \pi r_0^2\sigma_0^2\exp\left[-\frac{\|\mathbf{v}\|^2 r_0^2}{4}\right]$$

Then we have for F_1

$$F_1(\mathbf{x}) = \frac{1}{4\pi^2}\iint d^2\mathbf{v}\,\exp(i\mathbf{x}\cdot\mathbf{v})\,S_\phi(\mathbf{v})\left[\sum_{j=0}^{\infty}\zeta_{2j}\,\delta^{2j}\,\nu^{2j}\right]$$

$$= \frac{r_0^2\sigma_0^2}{4\pi}\sum_{j=0}^{\infty}\zeta_{2j}\delta^{2j}\iint d^2\mathbf{v}\,\|\mathbf{v}\|^{2j}\exp(-\|\mathbf{v}\|^2 r_0^2/4 + i\mathbf{x}\cdot\mathbf{v}\,)$$

$$= K_\phi(\mathbf{x})\left[\sum_{j=0}^{\infty}\zeta_{2j}\,M_{2j}(\mathbf{x})(\delta/r_0)^{2j}\right] \qquad (3.2.17)$$

where the terms

$$M_j(\mathbf{x}) = \frac{1}{2\pi}\iint d^2\mathbf{u}\,\|\mathbf{u}\|^j\exp\left[-\frac{1}{2}\,\|\mathbf{u} - i\sqrt{2}\mathbf{x}/r_0\|^2\right]$$

will be polynomials of degree j. The first two of even order are

$$M_0(\mathbf{x}) = 1 \qquad M_2(\mathbf{x}) = 2\left[\,1 - \frac{\|\mathbf{x}\|^2}{r_0^2}\,\right]$$

Expansions for F_2 can be obtained by similar calculations. We find for the first few terms

$$F_2(\mathbf{x}) = K_\phi(\mathbf{x})\left[\zeta_0\zeta_0^* + 2\mathrm{Re}(\zeta_0\zeta_2)(\delta/r_0)^2 M_2(\mathbf{x}) + \cdots\right]$$

These formulae give explicit expressions for F_1 and F_2 as power series in (δ/r_0). For natural magnetospheric and ionospheric sources at geomagnetic mid-latitudes the typical source length scale r_0 is on the order of some thousands of kilometers (e.g. Porath *et al.*, 1970), while penetration depths δ are at most a few hundreds of kilometers (until periods on the order of days are reached when δ may approach a thousand kilometers). Thus $\delta/r_0 \ll 1$ and the series will converge very rapidly so that only one or two terms are needed to get a very accurate approximation.

Keeping only the first term will in fact be valid for most conditions (see discussion in DB). In this case we see that the convolution operator $G_E *$ reduces to multiplication by a constant ζ_0. This is equivalent to ignoring the wave number dependence of the impedance, since if the impedance is independent of wave number the corresponding space domain filter is proportional to a delta function. In the calculations presented below we have for the most part used the zero order approximation. We have also done limited calculations with the second order correction for the case of a half space to verify the validity of the zero order approximation.

Explicitly then, for the zero order approximation, we have for the vertical field terms for $u = x, y$

$$K_{zu}(\mathbf{x}) = E(B_z(\mathbf{x})B_u(0)) = \frac{i\zeta_0}{\omega}\frac{16}{r_0^4}\left[\,1 - \frac{\|\mathbf{x}\|^2}{2r_0^2}\right]K_\phi(\mathbf{x}) \qquad (3.2.18)$$

$$K_{zz}(\mathbf{x}) = E(B_z(\mathbf{x})B_z(0)) = \frac{|\zeta_0|^2}{\omega^2}\frac{16}{r_0^4}\left[1 - \frac{4\|\mathbf{x}\|^2}{r_0^2} + \frac{\|\mathbf{x}\|^4}{r_0^4}\right] K_\pi(\mathbf{x}) \quad (3.2.19)$$

We can express the electric field covariances in terms of the magnetic field covariances K_{uv}, $u, v = x, y, z$ which are given above in (3.1.6), (3.2.18) and (3.2.19):

$$E(E_u(\mathbf{x})B_v) = -\varepsilon(y,u)\,\zeta_0\,K_{\tilde{u}v}(\mathbf{x}) \qquad u = x, y \quad v = x, y, z \quad (3.2.20)$$

$$E(E_u(\mathbf{x})E_v) = \varepsilon(u,v)\,|\zeta_0|^2\,K_{\tilde{u}\tilde{v}}(\mathbf{x}) \qquad u, v = x, y \quad (3.2.21)$$

We can now use (3.1.6) and (3.2.18) - (3.2.21) to calculate $\mathbf{\Sigma}$ for synthetic arrays.

3.3: Results From Synthetic Arrays I: Symmetric arrays

We have shown above how to calculate the spatial auto and cross covariance functions for all combinations of field components (horizontal and vertical magnetic and horizontal electric fields) for a 1-d earth subject to spatially homogeneous random sources. These spatial covariances can be used to calculate the elements of the expectation of the SDM $E(\mathbf{S}) = \mathbf{\Sigma}$ for any set of components measured at any number of sites. We will now apply these results to study the properties of $\mathbf{\Sigma}$ for four idealized synthetic arrays. Arrays 1-3 are regular symmetric arrays with uniform station spacings, while array 4 has a more realistic non-uniform station spacing. Array one consists of 5 stations laid out in a cross, array two is 25 stations in a uniform 5×5 grid, and array three is a linear five station array. The configurations of the arrays are summarized in figure 3.1.

There are three length scales involved here: r_0 is the typical source length scale, δ is the skin depth of the electromagnetic fields in the earth, and x is the array size. The properties of $\mathbf{\Sigma}$ for these synthetic arrays will depend on the two non-dimensional parameters $\delta_0 = \delta/r_0$ and $x_0 = x/r_0$. We are primarily interested here in the case where $\delta_0, x_0 \ll 1$. For the initial discussion we consider the

particular case $x_0 = .1$, $\delta_0 = .05$. We will latter consider the effect of varying these parameters. Note that we do all calculations for a half space of constant conductivity. Since we consider only a single fixed frequency this assumption only effects the relative phases of field components.

Note also that we have assumed that all field components have uniform uncorrelated measurement errors added. We have rather arbitrarily chosen the error variance to be .05 % of the typical horizontal magnetic field signal power. As noted in Chapter 2, the addition of isotropic noise with variance σ^2 just adds σ^2 to all of the eigenvalues; none of the eigenvectors are effected by this. the addition of this noise serves, in a sense, a cosmetic purpose. With noise added all of the original noise free eigenvalues which are small compared to σ^2 appear equal to the noise level when plotted on a logarithmic scale. The addition of noise thus obscures structure in the smaller eigenvalues and emphasizes structure in the larger eigenvalues. There are several reasons why we have done this. First, with real data there will always be noise, so the addition of noise results in a model which is at least qualitatively more realistic. In the presence of noise only the features with power reasonably above background noise levels will be resolvable. We thus want to focus attention on the larger eigenvalues and their associated eigenvectors: these are the features which will be relevant to real data. Second, as we shall discuss in detail below, the larger eigenvalues have a simple interpretation which is for the most part independent of the details of the array configuration (i.e. then number and position of stations). This will not generally be true for the smaller eigenvalues. Note also that numerical problems begin to arise when one tries to extract the smaller eigenvalues. Given the lack of relevance of these small eigenvalues to the analysis of real data, we have chosen to ignore these problems and concentrate our attention on the structure of the larger eigenvalues. Adding a constant noise term to the diagonal effectively

accomplishes this. It should, however, be born in mind for the following discussion that there is additional structure in the small eigenvalues.

We first consider the case of the horizontal field SDM $\mathbf{\Sigma_H} = E(\mathbf{S_H})$ for array one. The ten ordered eigenvalues γ_i, i = 1, 10 of $\mathbf{\Sigma_H}$ are plotted in figure 3.2. The eigenvalues break into three distinct groups - two dominant eigenvalues γ_1, γ_2 of equal magnitude, three intermediate eigenvalues γ_3, γ_4, γ_5 which are roughly two orders of magnitude smaller than the first two, and five nearly equal small eigenvalues which are at the assumed noise level. Note that among the intermediate group, $\gamma_5 = \gamma_4$ and $\gamma_3 = 2\gamma_4$. The eigenvectors $\mathbf{v}_i$, i = 1, 5 corresponding to the five eigenvalues above the noise level γ_i, i=1,5 are plotted in figure 3.3. The dominant eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ correspond to virtually uniform horizontal magnetic fields which are linearly polarized in the east-west and north-south directions (figures 3.3(a) and 3.3(b)). We note here that there are some very small deviations from complete uniformity; we will discuss this in more detail below when we consider the denser 25 station array.

Since the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ correspond to nearly uniform sources, the magnitude of γ_1 and γ_2 relative to all of the others eigenvalues indicates that the typical source fields can be well approximated by uniform plane waves. The remaining three eigenvectors $\mathbf{v}_3$, $\mathbf{v}_4$ and $\mathbf{v}_5$ are essentially a first order correction to this approximation. Consideration of figure 3.3 shows that these eigenvectors correspond to linear combinations of gradients in B_x and B_y. For example, $\mathbf{v}_3$ (figure 3.3.c) corresponds to gradients of B_x in the x direction and B_y in the y direction. Linear interpolation of the fields plotted in (3.3c-e) results in three simple horizontal field functions which correspond to pure linear gradients

$$\mathbf{g}_3(\mathbf{x}) = \begin{bmatrix} B_x(\mathbf{x}) \\ B_y(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \tag{3.3.1a}$$

$$\mathbf{g}_4(\mathbf{x}) = \begin{bmatrix} B_x(\mathbf{x}) \\ B_y(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix} \tag{3.3.1b}$$

$$\mathbf{g}_5(\mathbf{x}) = \begin{bmatrix} B_x(\mathbf{x}) \\ B_y(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix} \tag{3.3.1c}$$

The eigenvectors $\mathbf{v}_3$, $\mathbf{v}_4$, $\mathbf{v}_5$ correspond to the functions $\mathbf{g}_3$, $\mathbf{g}_4$, $\mathbf{g}_5$ measured at the five sites in the artificial array. As we shall discuss shortly, these functions represent canonical gradient terms.

First, however, we consider the effect of adding vertical fields to this synthetic array. We plot the ordered eigenvalues from the full 15×15 SDM $\mathbf{\Sigma}$ in figure 3.4. The pattern of eigenvalues of $\mathbf{\Sigma}$ is essentially the same as that seen for $\mathbf{\Sigma_H}$ in figure 3.1. The only substantial difference (aside from the existence of five more eigenvalues at the noise level) occurs for γ_3 which is a roughly a factor of three larger when the vertical fields are included. The horizontal components of $\mathbf{v}_i$, i = 1, 5 are essentially identical to those displayed in figure 3.3. With the exception of $\mathbf{v}_3$ the vertical components are all nearly zero. For $\mathbf{v}_3$ the vertical components are approximately equal and in phase for all five stations in the synthetic array. If the phase of $\mathbf{v}_3$ is chosen so that the horizontal gradient components are real then the vertical components have real and imaginary parts which are equal and of opposite sign (corresponding to a 45 degree phase lag of the vertical fields relative to the horizontal gradients). The magnitudes of the real and imaginary parts of the B_z components are equal to the magnitudes of the four non-zero B_x and B_y components of $\mathbf{v}_3$ (see figure 3.3).

The fact that only $\mathbf{v}_3$ has non-zero vertical field components can be understood in terms of the horizontal spatial gradient approach to impedance estimation (Schmuker, 1970; Kuckes, 1973). Recall that this technique is based on the approximation (see (3.2.6))

$$B_z = C(\omega) \left[\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} \right] = C(\omega)\, \nabla_{\mathbf{h}} \cdot \mathbf{B}_{\mathbf{h}}. \tag{3.3.2}$$

where $C(\omega)$, Schmuker's inductive length scale, is related to the usual 1-d plane

wave impedance $\zeta(\omega)$ via

$$C(\omega) = \frac{-i\zeta(\omega)}{\omega}$$

and ∇_h denotes the horizontal divergence. A simple calculation shows that the three linear gradient functions $\mathbf{g}_3$, $\mathbf{g}_4$ and $\mathbf{g}_5$ defined in (3.3.1) satisfy

$$\nabla_h \cdot \mathbf{g}_3 = 2 \qquad \nabla_h \cdot \mathbf{g}_4 = 0 \qquad \nabla_h \cdot \mathbf{g}_5 = 0$$

Thus, given our identification of the horizontal components of the eigenvectors $\mathbf{v}_i$ as the pure gradient functions $\mathbf{g}_i$ measured at the five sites in the synthetic array, it is not surprising that only $\mathbf{v}_3$ has non-zero B_z components.

Furthermore, as we will now show, the complex ratio of the horizontal gradient components to the vertical components of $\mathbf{v}_3$ can be used to recover the impedance $\zeta(\omega)$. Using the numbering of the stations given in figure 3.1 we rescale $\mathbf{v}_3$ so that

$$\mathbf{v}_3 = \left[\, 0 \;\; 0 \;\; B_{z1} \;\middle|\; x/2 \;\; 0 \;\; B_{z2} \;\middle|\; 0 \;\; x/2 \;\; B_{z3} \;\middle|\; -x/2 \;\; 0 \;\; B_{z4} \;\middle|\; 0 \;\; -x/2 \;\; B_{z5} \,\right]^T \qquad (3.3.3)$$

Since x is the total array size (measured in Km, say) the horizontal components of $\mathbf{v}_3$ now correspond to gradients of unit magnitude (eg. 1 nT/Km) in B_x in the x direction and B_y in the y direction. For these fields we thus have $\nabla_h \cdot \mathbf{B}_h = 2$ (nT/Km). Then, from (3.3.2) we expect that the vertical fields should satisfy $B_{zi} \equiv 2C(\omega)$. (Note that with $\mathbf{v}_3$ scaled as in (3.3.3) the units of B_{zi} are in Km - the proper units for Schmuker's $C(\omega)$.)

We may easily verify that this expectation is satisfied for the actual $\mathbf{v}_3$ computed for this synthetic array. We have assumed a half space of constant conductivity and, since we have chosen $\delta_0 = .05$, $x_0 = .10$ we have $\delta = x/2$. Then

$$C(\omega) = \frac{-i\zeta}{\omega} = -\frac{i}{\omega}\,\frac{\omega\delta(1+i)}{2} = \frac{(1-i)x/2}{2} \qquad (3.3.4)$$

Thus we expect $B_{zi} = (1-i)x/2$ so that $\mathbf{v}_z$ should be proportional to

$$\left[\, 0 \;\; 0 \;\; 1-i \;\middle|\; 1 \;\; 0 \;\; 1-i \;\middle|\; 0 \;\; 1 \;\; 1-i \;\middle|\; -1 \;\; 0 \;\; 1-i \;\middle|\; 0 \;\; -1 \;\; 1-i \,\right]^T$$

This is precisely the form of $\mathbf{v}_3$ described above.

We now turn to array two, the 25 station array on a regular grid. We use the same parameters as for array one. The array sizes are identical (relative to source scales and the skin depth) - the array is simply more densely sampled. For this array we include electric field measurements so the total number of components is 25×5 = 125. The 125 ordered eigenvalues of the SDM are plotted in figure 3.5. The basic pattern of eigenvalues is identical to that seen for the much smaller 5 station array, with $\gamma_1 = \gamma_2 \gg \gamma_3 > \gamma_4 = \gamma_5 \gg \gamma_j$, $j = 6, 125$. The dominant five eigenvectors $\mathbf{v}_i$, $i = 1, 5$ also show a similar, but more densely sampled, pattern for the magnetic fields. We will take advantage of this denser sampling to discuss some more subtle aspects of the eigenvectors.

The source gradient eigenvectors $\mathbf{v}_3$, $\mathbf{v}_4$ and $\mathbf{v}_5$, are shown in figure 3.6 where both electric and magnetic fields are plotted. The horizontal magnetic fields sampled with the denser station spacing confirms our characterization of $\mathbf{v}_3$,$\mathbf{v}_4$ and $\mathbf{v}_5$ as measurements of the pure gradient functions $\mathbf{g}_1$, $\mathbf{g}_2$, and $\mathbf{g}_3$ at discrete points. The corresponding electric field vectors are plotted on the right in figure 3.6. At each station the electric fields are proportional to the magnetic fields at the same station but are turned 90 degrees counter-clockwise. The electric field components of these eigenvectors thus also represent gradients (in the electric fields). The electric fields are in phase across the array and lead the magnetic fields by 45 degrees (only real parts of the electric fields are plotted in figure 3.6). Since the electric fields are tangent to the current streamlines at the surface, they are useful for visualizing current flows. Note that an equivalent external source current which will result in these field patterns on the surface will have the same spatial pattern

with current flowing in the opposite direction.

The third eigenvector $\mathbf{v}_3$ is distinctly different from $\mathbf{v}_4$ and $\mathbf{v}_5$. For $\mathbf{v}_3$ the current flows in concentric circles about the central station. Again, only for this eigenvector are the vertical field components B_z substantially different from zero. The vertical fields components for $\mathbf{v}_3$ are nearly uniform for all 25 stations in the array. As we have discussed above, the magnitude and phase of these vertical fields can be related to the horizontal components via the spatial gradient equation (3.3.2).

The electric fields for $\mathbf{v}_4$ and $\mathbf{v}_5$ show current flowing in a family of hyperbolas centered at the central station. Note that the field patterns for $\mathbf{v}_4$ and $\mathbf{v}_5$ differ only by a rotation of 45 degrees. Since $\mathbf{v}_4$ and $\mathbf{v}_5$ correspond to the degenerate eigenvalue $\gamma_4 = \gamma_5$, they are not uniquely determined. Any linear combination of $\mathbf{v}_4$ and $\mathbf{v}_5$ will also be an eigenvector of $\mathbf{\Sigma}$. It is easy to show however, that the fields for any linear combination $\mathbf{v} = c_1\mathbf{v}_4 + c_2\mathbf{v}_5$ are a rotation of the fields plotted in figure 3.6, so that this general pattern of fields and current flow is characteristic of the entire eigenspace. Indeed, any linear combination of the gradient functions $\mathbf{g}_4$ and $\mathbf{g}_5$ can be written as

$$c\,[\,\cos(\theta)\mathbf{g}_4(\mathbf{x}) + \sin(\theta)\mathbf{g}_5(\mathbf{x})\,] \;=\; c\begin{bmatrix}\cos(\theta)x + \sin(\theta)y\\ \sin(\theta)x - \cos(\theta)y\end{bmatrix} \;=\; c\,\mathbf{g}_4(\mathbf{x}') \quad (3.3.5)$$

for some constant c, where

$$\mathbf{x}' = \begin{bmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}$$

is just the coordinate vector $\mathbf{x}$ rotated by θ degrees. Thus the fields for the linear combination of (3.3.5) look like those for $\mathbf{g}_4$ rotated by θ degrees. The same holds for the finite dimensional vectors $\mathbf{v}_4$ and $\mathbf{v}_5$ which are just the functions $\mathbf{g}_4$, $\mathbf{g}_5$ sampled at discrete points.

Some properties of the eigenvectors v_i, $i = 1, 5$ are clarified by consideration of a Taylor's series expansion for the potential function ϕ about the origin

$$\phi(\mathbf{x}) = \phi(0) + x\frac{\partial\phi}{\partial x}(0) + y\frac{\partial\phi}{\partial y}(0) + x^2\frac{\partial^2\phi}{\partial x^2}(0) + y^2\frac{\partial^2\phi}{\partial y^2}(0) + xy\frac{\partial^2\phi}{\partial x\partial y}(0) + \cdots$$

We may then write for the horizontal magnetic field vector

$$\begin{bmatrix}B_x(\mathbf{x})\\ B_y(\mathbf{x})\end{bmatrix} = B_x(0)\begin{bmatrix}1\\0\end{bmatrix} + B_y(0)\begin{bmatrix}0\\1\end{bmatrix} \quad (3.3.6)$$

$$+ 2\frac{\partial B_x}{\partial x}(0)\begin{bmatrix}x\\0\end{bmatrix} + 2\frac{\partial B_y}{\partial y}(0)\begin{bmatrix}0\\y\end{bmatrix} + 2\frac{\partial B_x}{\partial y}(0)\begin{bmatrix}y\\x\end{bmatrix} + \cdots$$

which upon rearrangement of the linear terms is

$$\begin{bmatrix}B_x(\mathbf{x})\\ B_y(\mathbf{x})\end{bmatrix} = B_x(0)\begin{bmatrix}1\\0\end{bmatrix} + B_y(0)\begin{bmatrix}0\\1\end{bmatrix}$$

$$+ \left[\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y}\right](0)\begin{bmatrix}x\\y\end{bmatrix} + \left[\frac{\partial B_x}{\partial x} - \frac{\partial B_y}{\partial y}\right](0)\begin{bmatrix}x\\-y\end{bmatrix} + 2\frac{\partial B_x}{\partial y}(0)\begin{bmatrix}y\\x\end{bmatrix} + \cdots$$

The Taylor's series approximation represents an expansion of the fields in a set of polynomial basis functions. In the rearranged form of (3.3.6) we see that eigenvectors $\mathbf{v}_i$, $i = 1, 5$ correspond approximately to the two zero order and three first order terms, sampled at the station locations. Note also that (3.3.6) demonstrates that the three gradient functions $\mathbf{g}_i$, $i = 1, 3$ can be used to represent any possible linear gradient of the horizontal fields.

With the rearranged form of the Taylor's series in (3.3.6) the coefficients of the polynomial basis functions have "nice" properties under coordinate transformations. These coefficients are $\beta_1 = B_x(0)$ and $\beta_2 = B_y(0)$ for the zero order terms and

$$\beta_3 = \frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} \qquad \beta_4 = \frac{\partial B_x}{\partial x} - \frac{\partial B_y}{\partial y} \qquad \beta_5 = \frac{\partial B_x}{\partial y} + \frac{\partial B_y}{\partial x}$$

for the first order terms. The pair (β_1, β_2), of course, form a vector. The gradient coefficients β_i, $i = 3, 5$ also have nice transformation properties. The first, β_3 is just the divergence of the horizontal magnetic fields and is thus a scalar, invariant under coordinate rotations. For (β_4, β_5) we find that changing the coordinate system by rotating θ degrees

$$\mathbf{x}' = \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

transforms (β_4,β_5) to

$$\begin{bmatrix} \beta_4' \\ \beta_5' \end{bmatrix} = \begin{bmatrix} \dfrac{\partial B_{x'}}{\partial y'} + \dfrac{\partial B_{y'}}{\partial x'} \\ \dfrac{\partial B_{x'}}{\partial x'} - \dfrac{\partial B_{y'}}{\partial y'} \end{bmatrix} = \begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix} \begin{bmatrix} \beta_4 \\ \beta_5 \end{bmatrix} \qquad (3.3.7)$$

This property of the gradient coefficients has been previously noted by Kuckes *et al.* (1985). Our approach is somewhat different, in that these authors used a spherical harmonic expansion for the potential over a perfect conductor to demonstrate these transformation properties.

The three gradient terms described above form a natural extension to the plane wave source approximation. These gradient terms can be justified in terms of a Taylor's series expansion of the total potential, and they arise in a natural manner when the expected spectral density matrix $\mathbf{\Sigma}$ is diagonalized. Together with the plane wave sources, the gradients define a five dimensional space of external sources. While the correspondence between these ideal source terms and the eigenvectors $\mathbf{v}_i$, $i = 1, 5$ is very good it is not in general exact.

In figure (3.7) we examine the eigenvector $\mathbf{v}_1$ from the 25 station array two more closely. For a north-south linearly polarized plane wave source the amplitudes of the B_x and E_y field components should be constant (and non-zero); all other components should be zero. Although, the field components satisfy these

conditions approximately, there are deviations in all components. For the parameters used in this example ($\delta_0 = .05$, $x_0 = .1$), these deviations are small, having typical amplitudes of a few tenths of a percent of the total field amplitude.

In figure (3.7a-c) we contour the magnetic field components B_x, B_y and B_z respectively. For B_x the fields are larger in the center of the array, and are smallest in the four corners. Note that the fields are more nearly constant in the y direction. The other horizontal component B_y is nearly zero everywhere as expected, and is exactly zero on the north-south and east-west lines passing through the center station. The largest deviations are found in the four corners, with the contour plots of the fields forming a saddle centered at the origin. The vertical field components show a linear gradient in the east-west direction with a zero line going going east-west through the origin.

For perfect plane wave sources, the electric fields should be uniform and they should point in the east-west direction. The actual electric fields for $\mathbf{v}_1$ show small deviations from this ideal form. In figure (3.7d) we plot the vector *deviations* from the ideal uniform east-west fields. Again these are small - a few tenths of a percent of the total electric fields. The electric field perturbation vectors plotted in figure (3.7d) are tangent to the streamlines of the small electric currents which produce the magnetic field perturbations. These perturbations to the dominantly uniform currents consist of two current vortices of opposite sign, elongated in the east west direction. The two vortices together lead to to an enhancement of the current flow from west to east in the center of the array. These currents are thus consistent with the observed magnetic field perturbations including the reversal of sign in the vertical component, the enhancement of B_x in the region of enhanced current flow, and the reduction of B_x in the region of reversed current.

The other dominant eigenvector $\mathbf{v}_2$ shows analogous perturbations in all components. In addition, the gradient eigenvectors are slightly perturbed from the

ideal pure gradient terms discussed above.

In our experience perturbations of the plane wave response estimates due to random measurement errors are usually much larger than a few tenths of a percent. Hence, as long as array sizes are as small (compared to typical source length scales) as those considered here, the perturbations noted above will not be a serious problem. For larger arrays (e.g. large magnetometer arrays such as those described in Porath and Gough (1971), Lilley *et al.* (1981) or by The EMSLAB Group (1987)) the deviations of the eigenvectors from the simple forms described here will certainly be non-negligible. In such cases the direct interpretation of the eigenvectors of $\mathbf{\Sigma}$ in terms of the simple source morphologies described above will not be possible. As we will discuss in the next section, however, these eigenvectors are still meaningful - albeit harder to interpret.

3.4: The Eigenvalues of $\mathbf{\Sigma}$ as a Spatial Power Spectrum

The basic properties of the eigenvalues and eigenvectors of $\mathbf{\Sigma}$ are the same for the two symmetric arrays arrays considered in section 3.3 - the addition of stations to form the more densely sampled larger array just revealed finer details. At least in theory, we can consider the case of sampling at every point in a bounded region S (e.g. the square region which the 25 station grid covers). Consideration of random source models in this situation leads to an integral equation eigenvalue problem which we can relate to a sort of spatial spectral analysis. Before returning to a more detailed discussion of results from our simple model we consider some qualitative properties of this problem. We thus temporarily drop the restrictive assumptions of the previous sections and allow an arbitrary conductivity distribution and an arbitrary random source distribution. Although the results given here can be plausibly demonstrated with rather simple calculations, formal proofs of the results are mathematically technical and have not been attempted. We will

be quite casual; some discussion of the mathematical technicalities required to formalize these results can be found, in part, in Adler (1981).

Consider, then, the magnetic and electric fields $\mathbf{F}(\mathbf{x})$ restricted to a bounded region S (note that this could be the entire surface of the earth). We assume that the external source potentials are given by a sufficiently regular complex valued stochastic process. Then the fields $\mathbf{F}(\mathbf{x})$ will be random vector-valued complex fields (related to the random potential ϕ via the linear operators, $L_{\sigma,\omega}$ and L_E discused in section 2.1). Let

$$\mathbf{K}(\mathbf{x},\mathbf{x}') = E(\, \mathbf{F}(\mathbf{x})\mathbf{F}(\mathbf{x}')^* \,)$$

be the spatial cross covariance tensor for the magnetic and electric field components. Just as the matrix $\mathbf{\Sigma}$ defines the second moment properties of the magnetic fields measured at a finite number of stations, the function $\mathbf{K}(\mathbf{x},\mathbf{x}')$ for $\mathbf{x}, \mathbf{x}' \in S$ defines the second moment properties of the magnetic fields measured everywhere in S. The analogue of the matrix eigenvalue problem that we have considered above is the integral equation eigenvalue problem of (3.0.1).

To make some qualitative statements about this problem we use some standard mathematical results from the theory of linear operators. Define the operator

$$\boldsymbol{K}(\mathbf{u}) = \int_S d^2\mathbf{x}'\, \mathbf{K}(\mathbf{x},\mathbf{x}')\mathbf{u}(\mathbf{x}')$$

on the space of square integrable (complex, vector valued) functions defined on S (i.e. $\boldsymbol{L}^2(\boldsymbol{S})$). Then we have the eigenvalue problem

$$\boldsymbol{K}(\mathbf{v}) = \gamma\mathbf{v} \qquad (3.4.1)$$

Using results summarized, for example, in Folland (1976; pp33-43) it can be shown that:

(1) $\boldsymbol{K}$ is a compact operator, and will thus have a discrete spectrum with only a countable number of eigenvalues γ_i (only a finite number of which will have

magnitude greater than any finite $\varepsilon > 0$).

(2) The set of eigenfunctions $\mathbf{v}_i$ associated with the eigenvalues γ_i form a complete orthonormal set so any function $\mathbf{f}$ in $L^2(S)$ can be expanded

$$\mathbf{f} = \sum_i \langle \mathbf{v}_i, \mathbf{f} \rangle \mathbf{v}_i$$

where the inner product is defined in the usual way

$$\langle \mathbf{g}, \mathbf{f} \rangle = \int_S d^2\mathbf{x}\, \mathbf{g}^*(\mathbf{x})\mathbf{f}(\mathbf{x})$$

If we apply these results to our random field model we find that the random electric and magnetic field vector $\mathbf{F}(\mathbf{x})$ can be written

$$\mathbf{F}(\mathbf{x}) = \sum_i \beta_i \mathbf{v}_i(\mathbf{x}) \qquad \text{where} \qquad \beta_i = \langle \mathbf{v}_i, \mathbf{F} \rangle \tag{3.4.2}$$

The coefficients β_i are now random variables. Now

$$E(\beta_i\, \beta_j^*) = E\left[\left[\int_S \mathbf{v}_i^*(\mathbf{x})\mathbf{F}(\mathbf{x})\, d^2\mathbf{x} \right] \left[\int_S \mathbf{v}_j^*(\mathbf{x}')\mathbf{F}(\mathbf{x}')\, d^2\mathbf{x}' \right]^* \right]$$

$$\int\int_{S\times S} \mathbf{v}_i^*(\mathbf{x})\, E(\mathbf{F}(\mathbf{x})\mathbf{F}^*(\mathbf{x}'))\mathbf{v}_j(\mathbf{x}')d^2\mathbf{x}d^2\mathbf{x}' = \int_S d^2\mathbf{x}\, \mathbf{v}_i^*(\mathbf{x}) \int_S d^2\mathbf{x}'\mathbf{K}(\mathbf{x},\mathbf{x}')\mathbf{v}_j(\mathbf{x}')$$

$$= \int_S d^2\mathbf{x}\, \mathbf{v}_i^*(\mathbf{x})\, \gamma_j\, \mathbf{v}_j(\mathbf{x}) = \gamma_j \delta_{ij} \tag{3.4.3}$$

so these random coefficients are uncorrelated and have variances given by the eigenvalues. The eigenfunction expansion thus gives a representation of the random electromagnetic fields as a random sum of discrete uncorrelated (and orthogonal) modes. Note that a calculation similar to that of (3.4.3) shows that all eigenvalues are non-negative (so the operator $\boldsymbol{K}$ is non-negative definite).

Equation (3.4.2) gives a representation of any function in $L^2(S)$ in terms of the the eigenfunctions $\mathbf{v}_i(\mathbf{x})$. Not all functions in $L^2(S)$ will correspond to possible realizations of the fields. Let Φ be the linear space containing all realizable external potential functions (note that these will be a subspace of the space of all analytic functions $\phi(x,y,z)$ satisfying $\phi \to 0$ as $z \to +\infty$). Then as in (2.2.3) and (2.2.4) the linear operators $L_{\sigma,\omega}$ and L_E (both of which depend on the electrical conductivity distribution) determine the space of realizable electromagnetic fields $\boldsymbol{R}$ (i.e. the response space, which is now of infinite dimension). The vector valued functions in $\boldsymbol{R}$ are all consistent with the physical laws of electromagnetism in general, and with the specific distribution of electrical conductivity, $\sigma(x,y,z)$, $z \geq 0$, in particular.

The distribution of the random fields can be described abstractly in terms of a probability measure P defined on (some collection of measurable subsets of) $L^2(S)$ (Adler, 1981). Note that since all realizable fields are contained in $\boldsymbol{R}$ we have $P(\boldsymbol{R}) = 1$. Then we may write the covariance tensor as an abstract integral over the space $\boldsymbol{R}$

$$\mathbf{K}(\mathbf{x},\mathbf{x}') = E(\, \mathbf{F}(\mathbf{x})\mathbf{F}^*(\mathbf{x}')\,) = \int_{\boldsymbol{R}} dP(\mathbf{F})\, \mathbf{F}(\mathbf{x})\mathbf{F}^*(\mathbf{x}')$$

Then if $\mathbf{v}(\mathbf{x})$ is an eigenvalue of $\boldsymbol{K}$ corresponding to a positive (non-zero) eigenvalue we may write

$$\mathbf{v}(\mathbf{x}) = \gamma^{-1} \int_S d^2\mathbf{x}'\, \mathbf{K}(\mathbf{x},\mathbf{x}')\mathbf{v}(\mathbf{x}')$$

$$= \gamma^{-1} \int_S d^2\mathbf{x}' \int_{\boldsymbol{R}} dP(\mathbf{F})\mathbf{F}(\mathbf{x})\mathbf{F}^*(\mathbf{x}')\, \mathbf{v}(\mathbf{x}') \tag{3.4.4}$$

$$= \gamma^{-1} \int_{\boldsymbol{R}} dP(\mathbf{F}) \left[\int_S d^2\mathbf{x}'\, \mathbf{F}^*(\mathbf{x}')\mathbf{v}(\mathbf{x}') \right] \mathbf{F}(\mathbf{x}) = \gamma^{-1} \int_{\boldsymbol{R}} dQ_{\mathbf{v}}(\mathbf{F})\mathbf{F}(\mathbf{x})$$

Thus all eigenfunctions corresponding to positive (non-zero) eigenvalues are averages of elements of $\boldsymbol{R}$, weighted by the abstract measure Q_v, and thus are elements of $\boldsymbol{R}$. (A substantial amount of mathematics is required to really prove this; what we give here is nothing more than a plausibility argument.) This is an important point: all eigenfunctions corresponding to positive eigenvalues of the spatial covariance operator $\boldsymbol{K}$ are realizable.

The eigenvalues γ_i represent the power in the i^{th} mode $\mathbf{v}_i$. They are thus essentially a spatial power spectrum for the fields restricted to the bounded region S. If the random fields are approximated by truncating the expansion of (4.3.2) at the p^{th} mode, then, on average, the total error (integrated over S) due to this approximation is $\sum_{i=p+1}^{\infty} \gamma_i$. It can be shown that this error can be made as small as desired by choosing p large enough. For arrays which are small compared to typical source scales, we have seen that a very good approximation is obtained with $p = 5$. More generally, this argument provides a strong justification for the response space approach in a very general setting.

Note that when the region is unbounded, the spectrum of the operator $\boldsymbol{K}$ becomes continuous. For the case of fields defined on the full plane $z = 0$ with spatially homogeneous random sources with a 1-d conductivity distribution, the eigenvalues of $\boldsymbol{K}$ define the usual spatial power spectrum $S(\nu)$. The corresponding eigenvectors for this particular case are

$$\mathbf{u}_\nu(\mathbf{x}) = \begin{bmatrix} \nu_x \\ \nu_y \\ \nu_z(\zeta) \end{bmatrix} \exp(i\nu \cdot \mathbf{x})$$

where $\nu = (\nu_x, \nu_y)$ and $\nu_z(\zeta)$ depends on the 1-d impedance ζ.

For regional arrays where the array size is small compared to typical source scales, we would argue that for many purposes the discrete spectral problem for a

bounded region is far more relevant than the usual continuous spectral treatment. It is clear that as the number of stations increases and as we sample the region S more densely the finite matrix eigenvalue problem becomes a better and better discrete approximation to the integral equation eigenvalue problem. The results obtained above for the 25 station array thus give us some idea of the behavior of the dominant eigenvectors and eigenvalues of the operator $\boldsymbol{K}$ for the square region covered by the grid. For scales relevant to regional arrays, we have seen that almost all of the power lies in a small number of eigenvalues. These eigenvalues cluster in groups ($\gamma_1 = \gamma_2 = \gg \gamma_3 > \gamma_4 = \gamma_5 \gg \cdots$). The dominant five eigenvectors which we have examined correspond closely to simple, easy to interpret field morphologies - plane wave sources and pure gradients.

Comparison of the 5 and 25 station arrays, shows that the eigenvectors of $\boldsymbol{\Sigma}$ are essentially just the eigenfunctions of the operator $\boldsymbol{K}$ sampled at the station locations. This will hold only for modes which vary slowly enough over the array that the effect of spatial aliasing due to sparse sampling is small. All five modes we consider satisfy this. Nonetheless, as we shall see below, an uneven distribution of stations will perturb the eigenvectors of $\boldsymbol{\Sigma}$ from the simple form described above. More generally, the question of how estimates of the discrete spatial spectrum described in this section are effected by sampling at a finite set of points is an important one. Another good problem for future study.

3.5: Non-symmetric Arrays

The two arrays considered in section 3.3 were symmetric - stations were evenly spaced on a regular two-dimensional lattice. This fact, together with the assumed symmetries of the sources and of the electrical conductivity, is responsible for the regular form of the eigenvalues and eigenvectors. If any of these symmetries fail to hold (and in general all will fail for real data) the eigenvalues and

eigenvectors will be perturbed somewhat from the form described above.

Array four consists of five stations in an irregular configuration. This array looks considerably more like a real array than the previous two examples. We consider only the case of magnetic fields for this array. As above, we assume values for the non-dimensional parameters of $\delta_0 = .05$, $x_0 = .1$ (although the definition of x_0 is not so clear here). The eigenvalues of $\mathbf{\Sigma_H}$ and $\mathbf{\Sigma}$ follow the same basic pattern that we have seen for the two regular arrays considered above. There are, however, a few subtle differences. Now we find that all eigenvalues are distinct, satisfying

$$\gamma_1 \simeq \gamma_2 \gg \gamma_3 > \gamma_4 > \gamma_5 \gg \gamma_i, \qquad i > 5$$

The first five eigenvectors for the 10×10 matrix $\mathbf{\Sigma_H}$ are plotted in figure 3.8. The two dominant eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are, as expected, still essentially constant over the array (the small deviations discussed above occur here also; they are of the same magnitude - a few tenths of a percent). The next three eigenvectors $\mathbf{v}_3$, $\mathbf{v}_4$, and $\mathbf{v}_5$ clearly still represent gradients in the horizontal magnetic fields. The individual eigenvectors, however, are not at all regular and cannot be associated with a specific simple gradient term of the sort discussed above. Note however, that $\mathbf{v}_3$ is similar to the horizontal divergence term which we identified with $\mathbf{v}_3$ for the regular arrays, and $\mathbf{v}_4$, $\mathbf{v}_5$ bear some resemblance to their regular array counterparts. The situation is similar when we include vertical fields. The eigenvectors $\mathbf{v}_3$, $\mathbf{v}_4$, and $\mathbf{v}_5$ for the 15×15 matrix $\mathbf{\Sigma}$ are plotted in figure 3.9 with the real and imaginary parts of the vertical components (nearly constant for all five stations for each eigenvector) printed in the upper left hand corner.

For this irregular array the vertical field components are now non-zero for two eigenvectors, $\mathbf{v}_3$ and $\mathbf{v}_5$. Note that the horizontal components of $\mathbf{v}_3$ and $\mathbf{v}_5$ change when vertical field components are included in the analysis. Note also that

the eigenvector v_4 of $\mathbf{\Sigma}$ has no vertical component associated with it and that the horizontal components are identical to the horizontal fields only case ($\mathbf{\Sigma_H}$). (It can be shown that for the isotropic source potential models considered here that (at least) one eigenvector will always have these properties. This eigenvector will correspond to the canonical gradient function g_4 (or $\mathbf{g}_5$) in a particular coordinate system (which is determined by the station distribution).)

Due to the complications of the gradient eigenvectors the inductive length scale $C(\omega)$ cannot be obtained for irregular arrays in the simple direct manner described above for regular arrays. Nonetheless, it is clear that the phase of the vertical fields relative to the horizontal gradients (45 degree lag) is consistent with our assumption of a half space of constant conductivity.

To recover $C(\omega)$ from the eigenvectors, we first note that the three horizontal field gradient eigenvectors plotted in figure 3.8 are essentially linear combinations of the canonical gradient functions $\mathbf{g}_3$, $\mathbf{g}_4$, $\mathbf{g}_5$ measured at the five station locations $\mathbf{x}_i$, i=1,5 (with the origin chosen so that the mean station location vector $\bar{\mathbf{x}} = 0$),

$$\mathbf{u}_3 = \begin{bmatrix} x_1 \\ y_1 \\ - \\ . \\ . \\ . \\ - \\ x_5 \\ y_5 \end{bmatrix} \qquad \mathbf{u}_4 = \begin{bmatrix} y_1 \\ x_1 \\ - \\ . \\ . \\ . \\ - \\ y_5 \\ x_5 \end{bmatrix} \qquad \mathbf{u}_5 = \begin{bmatrix} x_1 \\ -y_1 \\ - \\ . \\ . \\ . \\ - \\ x_5 \\ -y_5 \end{bmatrix} \qquad (3.5.1)$$

To demonstrate this we can find the linear combinations of the $\mathbf{u}_i$ which best fit the eigenvectors $\mathbf{v}_j$ in a least squares sense - i.e. we minimize

$$\| \mathbf{v}_j - \sum_{i=3}^{5} \beta_i \mathbf{u}_i \| \qquad j = 3, 5$$

The total misfits (with $\mathbf{v}_j$ normalized to unit magnitude) for the three vectors obtained from this exercise are 5.25×10^{-7}, 8.06×10^{-7} and 2.03×10^{-6}, so we see

that the eigenvectors $\mathbf{v}_j$, $j = 3, 5$ are very nearly linear combinations of the canonical gradient vectors $\mathbf{u}_i$. We can do the same sort of calculations for the eigenvectors $\mathbf{v}_3$, $\mathbf{v}_4$,$\mathbf{v}_5$ from the full 15 component matrix $\boldsymbol{\Sigma}$ which includes vertical components. To be precise (but somewhat casual about notation) let the $\mathbf{u}_i's$ now be the 15-dimensional vectors with horizontal components the same as in (3.5.1) and with vertical components all zero. If we solve the least squares minimization problem

$$min! \parallel \hat{\mathbf{u}}_i - \sum_{j=3}^{5} \beta_j \mathbf{u}_j \parallel \qquad i = 3, 5$$

we find the linear combinations of the eigenvectors of $\boldsymbol{\Sigma}$ whose horizontal components are closest to the canonical gradient vectors. Doing this we find that the vertical components of the vectors $\hat{\mathbf{u}}_i$, i=3,5 are (of course) nearly constant for all stations. Expressing the station coordinates (used in the definition of the $\mathbf{u}_i's$) in the non-dimensional units $\mathbf{x}/r_0$ we find that the vertical components for $\hat{\mathbf{u}}_4$, $\hat{\mathbf{u}}_5$ are approximately zero, and that the vertical components for $\hat{\mathbf{u}}_3$ are $\simeq .05(\ 1 - i\ ) = 2C(\omega)$ (expressed in the same non-dimensional length). This is essentially the same as the result obtained for the regular arrays (see (3.3.4)).

We thus see that while the eigenvectors of the SDM's for irregular arrays do not have the simple regular form seen for regular arrays, the situation is not substantially different. The first five eigenvectors can still be divided cleanly into two groups. The first two eigenvectors correspond (approximately) to plane wave sources, while the next three correspond to gradients in the horizontal fields. The irregular spacing of stations causes the canonical symmetric array eigenvectors from each of these two groups to be mixed together within each group, but, to a good approximation, not between groups. Put another way, the groups of eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$, $\mathbf{v}_4$, $\mathbf{v}_5$ span the same spaces for both regular and irregular arrays. Finally, the canonical eigenvectors, and hence the inductive length scale

$C(\omega)$ can be recovered by taking linear combinations of the actual eigenvectors. These points are not particularly startling, but are important nonetheless since all real arrays will be irregular.

Array three deviates from the ideal arrays considered initially in a different fashion. For this array the station spacing is uniform but all stations lie on a line. This sort of array is often used when a small number of stations are occupied simultaneously. Many of the actual arrays that we will consider in this thesis (such as the EMSLAB MT line) take this form. The eigenvalues for the 10×10 SDM for the horizontal fields only $\boldsymbol{\Sigma}_{\mathbf{H}}$ now satisfy $\gamma_1 = \gamma_2 \gg \gamma_3 > \gamma_4 \gg \gamma_i$, i=5,10. The plane wave source vectors $\mathbf{v}_1$, $\mathbf{v}_2$ are the same as in the previous cases, but now there are only two gradient eigenvectors $\mathbf{v}_3$, $\mathbf{v}_4$ corresponding to gradients of the field components which are respectively parallel to (B_y) and perpendicular to (B_x) the strike of the array.

If we now add vertical fields to the linear array we find that the eigenvalues of the SDM satisfy $\gamma_1 = \gamma_2 \gg \gamma_3 > \gamma_4 > \gamma_5 \gg \gamma_i$, $i = 6, 15$ so there are again three gradient terms. There are only two gradient eigenvalues for $\boldsymbol{\Sigma}_{\mathbf{H}}$ because gradients perpendicular to the line (in the x direction) cannot be observed. The unobservable gradients $\frac{\partial B_x}{\partial x}$ are, however, correlated with the vertical components B_z. The effects of the gradients $\frac{\partial B_x}{\partial x}$ are thus observed even when the gradients themselves **are not**. Thus, three gradient modes are seen when vertical field components **are** included in the analysis.

Unfortunately, it is not in general possible to recover the response function $C(\omega)$ from a linear array. This is because the gradient terms $\frac{\partial B_x}{\partial x}$ and $\frac{\partial B_y}{\partial y}$ are generally correlated (even for the very simple homogeneous, isotropic source model that we have considered). Both gradient terms result in vertical fields but

only one is observed and it is not generally possible to determine what portion of the vertical fields are due to the unobserved gradients. Estimation of $C(\omega)$ by the horizontal spatial gradient technique thus requires 2-dimensional arrays.

3.6: Effect of Varying Parameters

In all of the results presented above we have kept the non-dimensional parameters δ_0 and x_0 fixed. We now consider very briefly the effect of varying these parameters. For this discussion we consider only the symmetric five station array, array one. We will consider results for the non-dimensional skin depth δ_0 in the range .01–.1. For source length scales of a few thousands of kilometers, these correspond to skin depths ranging from tens to hundreds of kilometers. These values are representative of the skin depths in the real earth for the MT and MV data (25 - 10000 s) that we consider in Chapters 2 and 4.

We have shown above that while the dominant two eigenvectors approximately correspond to plane wave sources there are some small deviations from complete uniformity of the fields across the synthetic arrays. In figure 3.10 we plot the fractional root mean square (RMS) misfit between the fields computed from the eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$ and the uniform fields expected for perfect plane wave sources. The fractional misfit is plotted as a function of non-dimensional array size x_0 for three values of the non-dimensional skin depth δ_0 = .1, .03, .01. For $x_0 < .1$ the misfit is on the order of a few tenths of a percent or less for all three values of δ_0, and even for $x_0 = .5$ the RMS misfit is only 2-3%. With our model for the spatial covariance of the random potential $K_\phi = \sigma_0^2 \exp(- \|\mathbf{x}\|^2/r_0^2)$ the correlation for the horizontal fields drops to zero for station separations of .7 (in the same non-dimensional units). These results thus indicate that the plane wave approximation remains reasonable for arrays whose size is a substantial fraction of the typical source field correlation length scales. Note, however that as

$x_0 \rightarrow 1$ the plane wave source approximation becomes poor very quickly.

Note also that as δ_0 increases the plane wave approximation degrades. This effect is fairly weak however, at least for the limited range of δ_0 considered here.

In figure 3.11 we demonstrate how variations of the parameters δ_0, x_0 effect the relative magnitudes of the plane wave and gradient eigenvalues for three component (magnetic field only) data. We plot here the ratio γ_3/γ_1 against the non-dimensional array size x_0 for values of the non-dimensional skin depth δ_0 = .1, .03, .01 and 0.0. Recall that for the symmetric array that we are considering here, half of the horizontal gradient power and all of the vertical field power is in the third eigenvector. Thus γ_3 is the sum of half of the power in the horizontal gradients and all of the power in the vertical fields that is due to source gradients. For $\delta_0 = 0.0$ the vertical fields are zero. Thus results for this value of δ_0 are equivalent to results that would be obtained for the horizontal fields only SDM $\mathbf{\Sigma}_H$.

In general, the power in the horizontal gradients is controlled by the parameter x_0, while the power in the vertical fields is controlled by δ_0. This is illustrated by the curves plotted in in figure 3.11. As $x_0 \rightarrow 0$ the horizontal gradient power rapidly decreases. For arrays with $x_0 = .01$ (corresponding to arrays extending over tens of kilometers at geomagnetic mid-latitudes) the power in the gradients is on the order of 5×10^{-5} the typical field power. Even for $x_0 = .1$ (hundreds of kilometers) the gradient power is only 10^{-3} of the main signal. The measurement of gradients in small arrays will thus require very low noise instruments. On the other hand, for fixed δ_0 as $x_0 \rightarrow 0$, γ_3/γ_1 approaches a constant value. This corresponds to the fraction of total power which is in the (gradient caused) vertical fields; this is independent of x_0. For $\delta = .1$ the power in these vertical fields is fully a tenth of the typical field values.

When $\delta_0 \geq x_0$, most all of the power in γ_3 is due to the vertical fields. As the array size gets small the horizontal gradients vanish, but the gradient effects on the vertical fields remain constant. Since γ_4, γ_5 represent power in horizontal gradients only, these eigenvalues will be small for small arrays. Hence as the array shrinks to a point the three gradient eigenvalues will degenerate to a single eigenvalue. The corresponding eigenvector will have only power in the B_z components. This explains why, for the small arrays which we considered in Chapter 2, there was generally a single dominant third eigenvalue whose associated eigenvector could be approximated quite well as vertical source fields with no associated gradients.

3.7: Inhomogeneous Sources; 3-d conductivity

The model we have emphasized in this chapter is highly idealized. We have made two very restrictive assumptions - that the external sources are spatially homogeneous and isotropic, and that the conductivity is 1-d, varying only with depth. While these assumptions are not realistic, they result in a tractable model which reproduces many of the main features of the array data described in Chapter 2. In this section we consider qualitatively the effect of relaxing our restrictive assumptions.

The assumption of isotropic source fields can easily be generalized. Isotropy of the covariance of the random source potential implies that the properties of the random potentials are invariant under rotations of the coordinate system. For real sources we would expect a strong directional dependence of the character of the signal - the external source currents are strongly controlled by the earths main field which imposes a fixed preferred direction. Much of the symmetry of the eigenvectors and eigenvalues of the SDM's will be lost when this assumption is dropped. In particular the occurrence of pairs of degenerate eigenvalues

corresponding to plane wave sources and to the gradient functions $\mathbf{g}_4$, $\mathbf{g}_5$ depends on the rotational symmetry of the sources. The effect of anisotropy of the source potential covariance will be similar to the effect of irregular station spacing - the eigenvalues and eigenvectors will not have the nice regular forms discussed above but there will still generally be a pair of plane wave eigenvectors, and a cluster of three gradient eigenvectors. The spaces spanned by these groups of vectors will be the same as for the idealized case we have discussed above. Thus in a practical sense the effect of anisotropy of the random source potential is fairly trivial.

Homogeneity of the source potential covariance implies that the properties of the random potentials are invariant under translations. An important consequence of spatial homogeneity is that the gradients of the field components at any point and the values of the fields at that point are uncorrelated. (If the field component B_x at point $\mathbf{x}$ is positively correlated with the gradient of B_x in the x direction at this point, then the B_x component a short distance away in the x direction will, on average, have a larger amplitude, contradicting the translational invariance.) As a consequence horizontal source field gradients and the (nearly) uniform plane wave source fields are statistically uncorrelated for homogeneous sources. These two types of fields thus separate into distinct groups of modes corresponding to distinct groups of eigenvalues. If, on the other hand, the gradients are correlated with the plane wave sources the gradient and plane wave modes will be mixed in the individual eigenvectors.

Of course the real ionospheric and magnetospheric sources do occur in preferred locations (and at preferred times) so we do not expect the random potential to be spatially homogeneous even for a 1-d earth. As a result we should expect that the dominant eigenvectors of the SDM for real data will be a mixture of plane wave and gradient sources and will thus yield a biased estimate of the plane wave response space. In a sense the problem here is similar to the problem, discussed

above, of choosing a normal reference station. Now we have a five dimensional response space $\boldsymbol{R}$ resulting from plane wave and gradient sources. Regardless of the correlation of the different types of sources we get the correct answer for $\boldsymbol{R}$. We would like, however, to identify the vectors in $\boldsymbol{R}$ which correspond to specific sources. In general, this identification will require some additional assumptions. Certainly for a 1-d earth where the general form of the plane wave and gradient terms is known, the identification of these terms is simple (if uninteresting). More generally, assuming that the plane wave and gradient terms are "nearly" of the form expected for a 1-d earth the identification may be possible. For larger arrays the fields associated with the vectors in the response space may be separated into internal and external parts (Porath *et al.*,1970; Richmond and Baumjohann; 1983) to help separate plane wave and gradient terms. This problem needs further study and is beyond the scope of this thesis.

In Appendix A we show that the fractional bias in the plane wave response space due to the correlation of plane wave and gradient source terms is of the order $b = |\rho|\sigma_g/\sigma_p$ where ρ is the (complex) correlation between the two source terms, and where σ_g^2, σ_p^2 give the power in the gradient and plane wave source terms respectively. Approximating σ_g/σ_p by $(\gamma_3/\gamma_1)^{1/2}$ and referring to figure 3.11 for estimates of γ_3/γ_1, we can get a feel for how large these biases may be. For *horizontal fields* in an array of (non-dimensional) size $x_0 = .1$ we use the curve $\delta_0 = .00$ and find $b \approx .05\,|\rho|$. In this case then, the bias will be small unless the correlation is very high (close to one). For the *vertical fields*, however, the situation may be much worse. With $\delta_0 = .1$ (corresponding to skin depths of some hundreds of kilometers) we find $b = .3|\rho|$. Unless $|\rho|$ is fairly small this bias may be quite large. A detailed analysis of the importance of this effect for real data will not be attempted here, but we offer some general comments based on our experience with small arrays at geomagnetic mid latitudes. Although it is difficult

to be sure that estimates are or aren't biased, we find no evidence for consistent gradient related biases in the horizontal field components estimated for small arrays (c.f. figures 5.2 and 5.3 in Chapter 5). For vertical fields, on the other hand, biases due to correlations between plane wave and gradient components do appear to be a problem sometimes. Based upon our experience with single station vertical field transfer function estimation (Egbert and Booker, 1986) we believe that the correlation problem is probably mostly due to a few isolated high power events associated with geomagnetic storms. It should thus be possible to minimize these effects with robust versions of the statistical methods presented in Chapter 4.

The symmetries in the eigenvectors and eigenvalues for our synthetic arrays are also highly dependent on the assumption that the conductivity is 1-d. For the general case, where conductivity varies in two or three dimensions the eigenvectors of the SDM will be perturbed from the simple forms described above. Indeed, it is these perturbations from the 1-d form that are geologically interesting. The exact computation of the statistical properties of the electromagnetic fields induced in a 3-d earth by randomly varying finite wavenumber sources would be an extraordinarily difficult problem. We restrict ourselves here to the qualitative discussion of a few main points.

Although the 1-d conductivity model is little more than a caricature of the real earth the main results derived from this simple model are still very relevant to the true situation. Most importantly, we have seen that the external source fields can be very well-approximated as linear combinations of two (approximately) plane wave source terms and three gradient terms (the eigenvectors actually give the total fields but for a 1-d earth the external fields can be easily inferred from the total fields). This property of the sources does not depend in any way on the conductivity distribution in the earth. We may thus conclude that for a small array (i.e. $x_0 \ll 1$) the finite dimensional response space model of Chapter 2 is

applicable. For small arrays, the source space Φ will still be of low dimension ($p = 2$ or $p = 5$ depending on whether gradient terms are included) and will still consist of source fields with simple, large scale, morphologies. Note that this is true even though the total fields may have small spatial scale variations caused by internal currents.

There will in fact be some mixing of the plane wave and gradient eigenvectors (as for the case of inhomogeneous sources) but we show in Appendix A that this effect will typically be very small (again the effect will be most severe for the vertical field components). In general, the fields associated with the dominant two eigenvalues will correspond to the plane wave sources, and those associated with the next three will represent the gradient terms. Furthermore, it is usually the case (for the horizontal magnetic fields at least) that the total fields due to a plane wave or gradient source are only perturbed a relatively small amount from the form expected for a 1-d earth. It will thus be possible to identify, at least approximately, vectors in the response space with specific source fields.

While many features of the spectral representation of the SDM for the general 3-d conductivity case will be the same as for the simple 1-d case that we have discussed in detail above, there will often be some significant differences. Particularly significant deviations from the simple picture painted above may occur for the electric field components of the eigenvectors. The electric fields may be severely distorted by small scale local conductivity gradients. The locally distorted electric fields at station i can be expressed in terms of of the undistorted electric fields $\mathbf{E}(\mathbf{x})$ via $\mathbf{E}(\mathbf{x}_i) = \mathbf{D}_i\mathbf{E}_u(\mathbf{x}_i)$ where for low enough frequencies the 2×2 matrix $\mathbf{D}_i$ is real and independent of frequency (Larsen, 1977; Zhang *et al.*, 1987). This local surface distortion can lead to significant variations in the amplitudes of the electric field components of the plane wave source eigenvectors (see figure 2.4 for example). The same effect should be expected for the electric field

components of the gradient eigenvectors. In fact, for truly local distortion, the distorted electric field components of the gradient eigenvectors should be related to the undistorted electric field gradients by the same distortion matrix $\mathbf{D}_i$.

There are some other effects which may occur in the electric field components of the gradient eigenvectors. In the 1-d case, the gradients in the magnetic fields produce identical gradients in the electric fields (with the electric field components rotated by 90 degrees and shifted in phase; see figure 3.7). In fact, the electric field components of the gradient eigenvectors at any site are related to the magnetic field components at the same site by the standard plane wave 1-d impedance - e.g $E_x = \zeta(\omega)B_y$. In general the electric fields can be written (DB; see also (3.2.1))

$$E_u(\mathbf{x}) = \sum_{v=1}^{2} \iint d^2\mathbf{x}' G_{uv}(\mathbf{x}-\mathbf{x}',\mathbf{x}) B_v(\mathbf{x}') \qquad u, v = x, y \tag{3.7.1}$$

The electric fields are thus determined by a weighted local average of the horizontal magnetic fields (where the 'weighting functions' are the kernels of the impedance operator). For the 1-d case the kernels depend only on $\mathbf{x}-\mathbf{x}'$ and are rotationally symmetric; see (3.2.4). As a result of this symmetry the electric fields are insensitive to local gradients. When the horizontal fields are linear combinations of uniform plane wave fields and gradient fields the electric fields at any point are always linearly related via the plane wave impedance $\zeta(\omega)$ to the local value of the magnetic fields. Thus for the small arrays we consider here, the dependence of the impedance on wavenumber can be completely neglected; there are no finite wave number source effects in the electric fields (as there are in the vertical magnetic fields).

The situation may be quite different when the conductivity is not 1-d. The kernels $G_{uv}(\mathbf{x}-\mathbf{x}',\mathbf{x})$, $u, v = x, y$ of the integral impedance operator of (3.7.1) will still be centered at $\mathbf{x}$ but they will now depend on $\mathbf{x}$ and they will not generally be

rotationally symmetric about this point. As a result, the electric fields will not necessarily be insensitive to gradients in the magnetic fields. In this case there may be source effects analogous to that seen for the vertical magnetic fields. For instance, if the shape of the kernels $G_{uv}(\mathbf{x}-\mathbf{x}',\mathbf{x})$, $u, v = x, y$ (determined by the dependence on $\mathbf{x}-\mathbf{x}'$) varies slowly as a function of $\mathbf{x}$, the perturbation of the electric fields due to the asymmetry of the kernels will be nearly constant and in phase across the array (but note that the fields will generally be statically distorted as well). In Chapter 4 we will demonstrate with the EMSLAB data that such situations do indeed occur; there can be significant first order source related effects in electric field data.

The situation with gradient related vertical fields is somewhat similar. With 1-d conductivity all vertical fields are correlated with a single gradient function, g_3. This is dependent on the symmetries inherent in the 1-d problem. In the general case there can be vertical fields correlated with all gradients. We will also demonstrate this with some 3 component MV data in Chapter 4. There are many other ways in which the interaction of finite wave-number sources and 3-d conductivity distributions will result in deviations from the results obtained from the simple 1-d model. We will see evidence of some of these when we consider some specific arrays in more detail in Chapter 4.

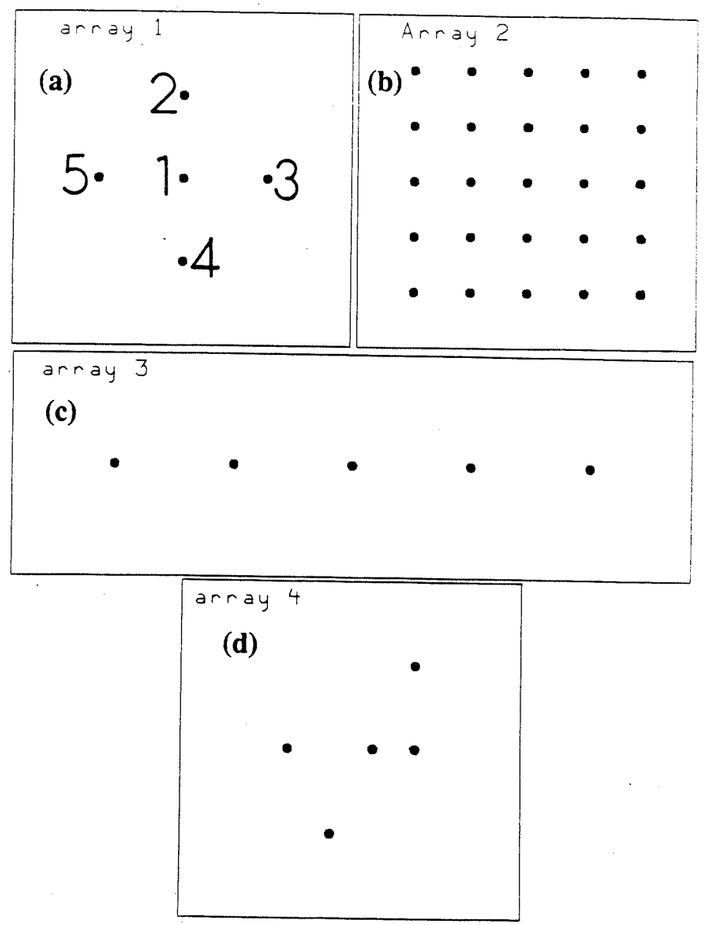


Figure 3.1 Summary of synthetic arrays:
(a) Array 1 - five stations in a cross
(b) Array 2 - 25 stations in a regular grid
(c) Array 3 - five stations in a straight line
(d) Array 4 - five stations in an irregular array

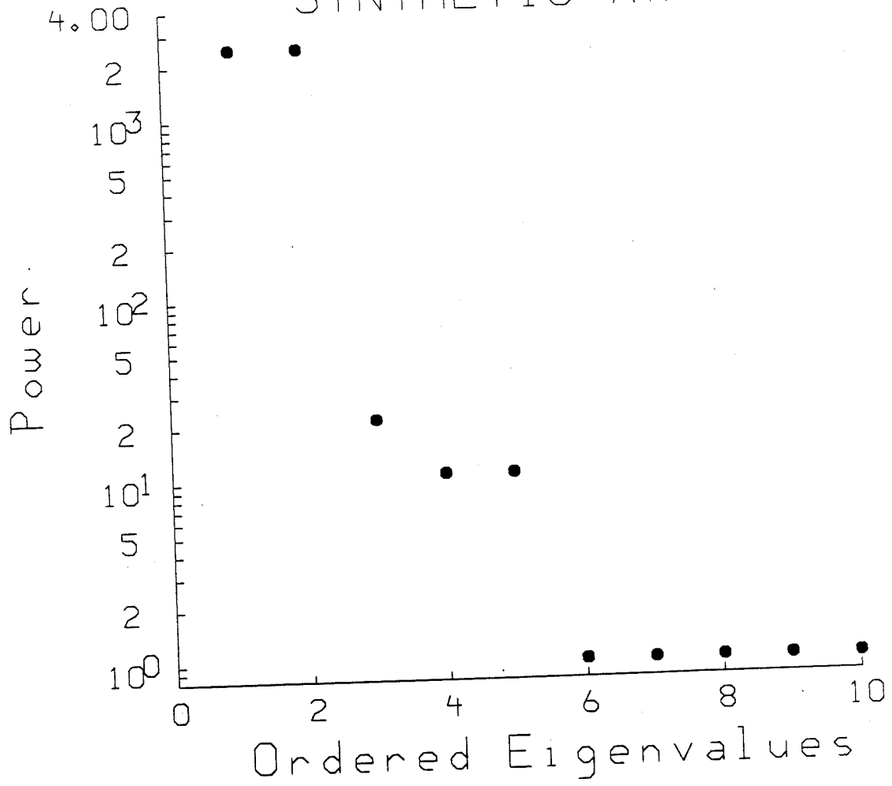


Figure 3.2 Ordered eigenvalues for Σ_H - array 1 with $x_0 = .1$. There are three distinct clusters of eigenvalues: $\gamma_1 = \gamma_2$, $\gamma_3 > \gamma_4 = \gamma_5$, and the remaining five eigenvalues at the noise level.

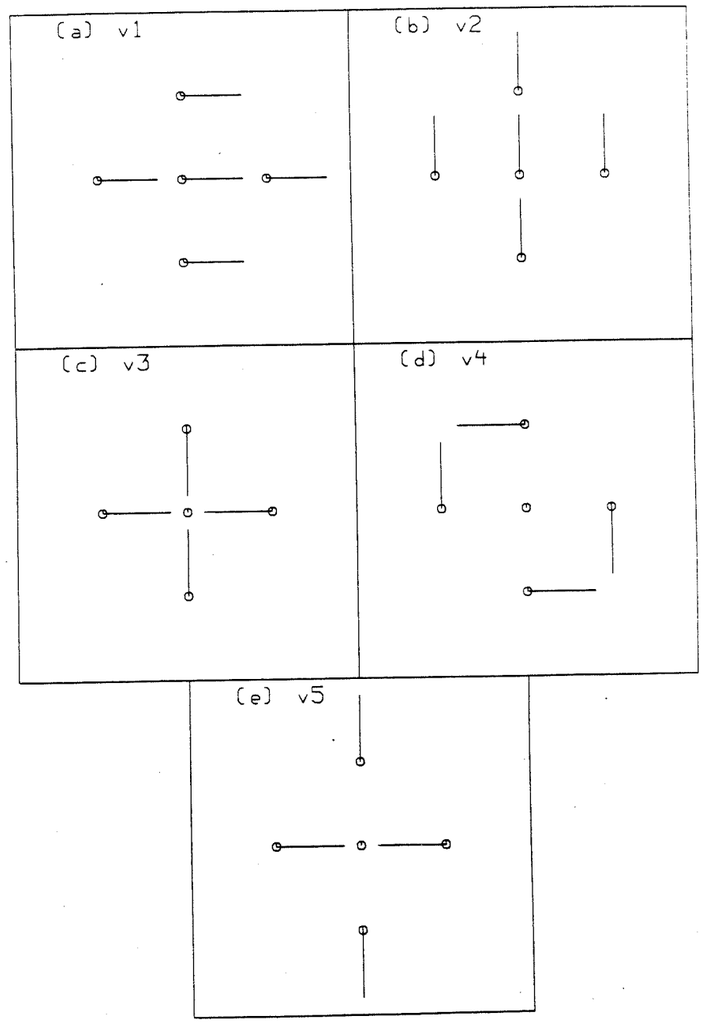


Figure 3.3 Five dominant eigenvectors $\mathbf{v}_i$, $i = 1, 5$ for Σ_H - array 1 with $x_0 = .1$. The eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ (a,b) correspond to nearly uniform source fields linearly polarized in the north-south and east-west directions respectively, while $\mathbf{v}_3$, $\mathbf{v}_4$ and $\mathbf{v}_5$ (c-e) correspond to gradients in the fields.

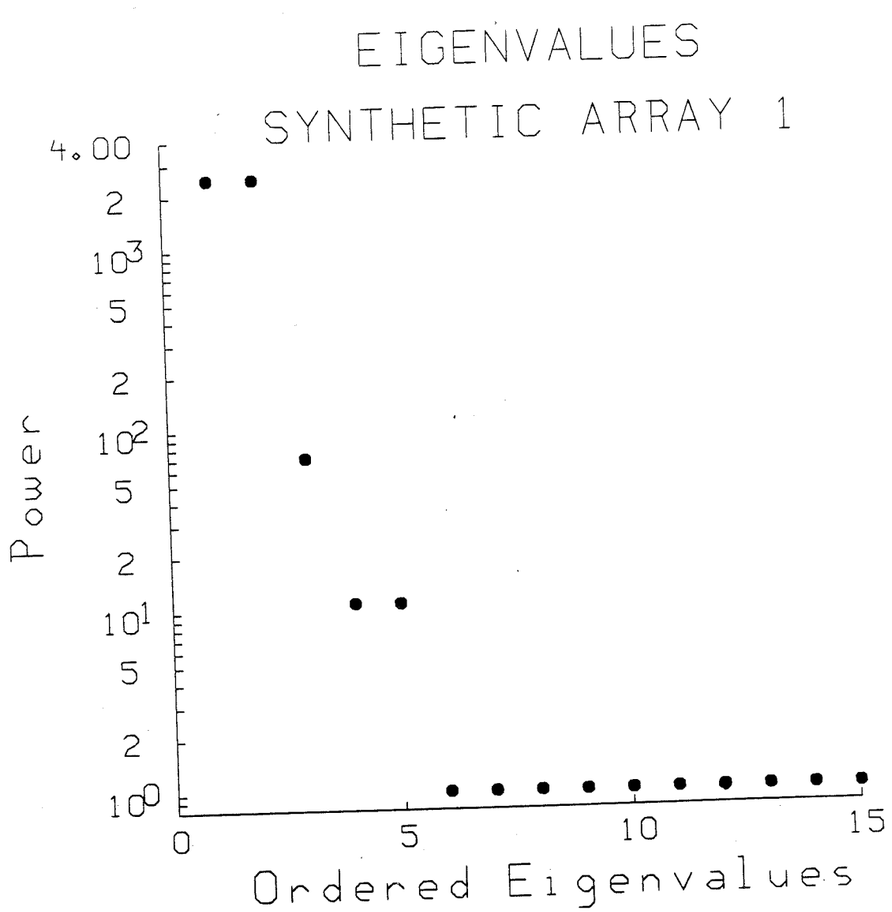


Figure 3.4 Ordered eigenvalues for Σ - array 1, with horizontal and vertical magnetic fields included. Here $x_0 = .1$ and $\delta_0 = .05$. The basic pattern is as in figure 3.2. The only difference is in γ_3 which is larger by a factor of roughly three. For a 1-d earth most all of the power in the vertical fields is in the third eigenvector.

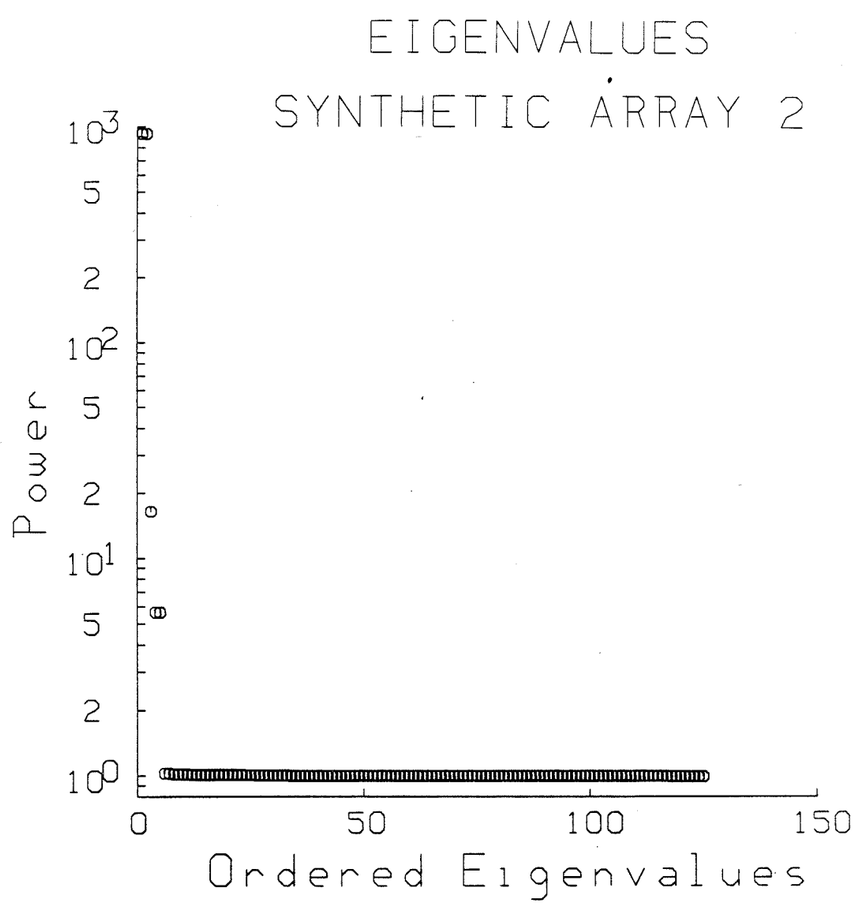
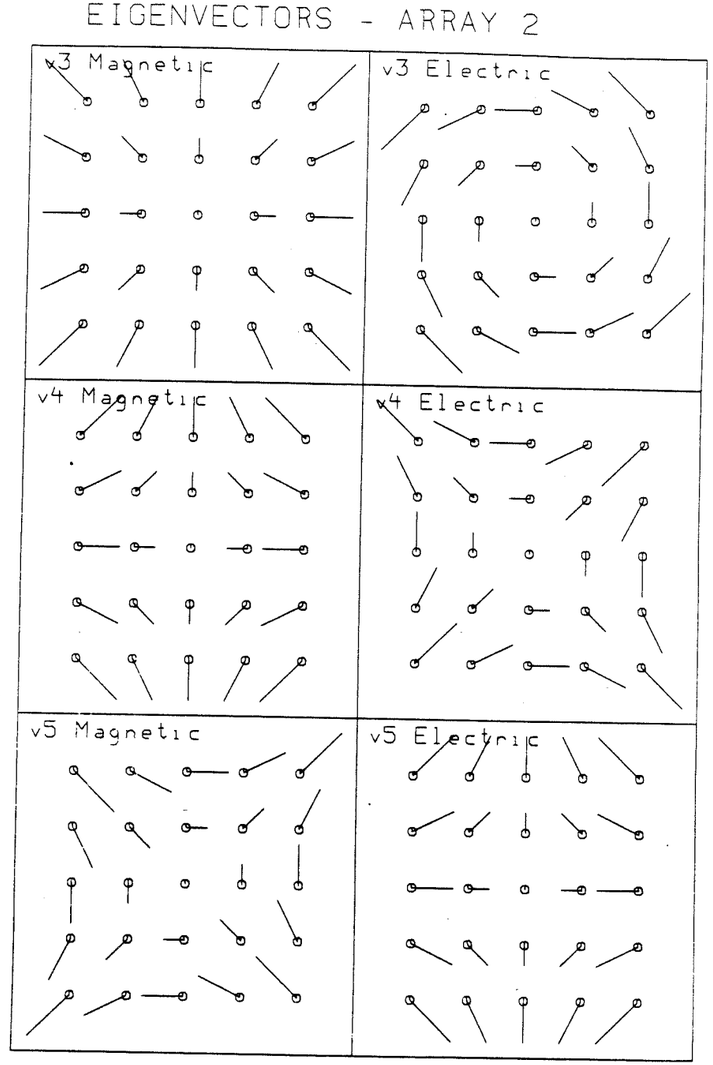


Figure 3.5 Ordered eigenvalues for Σ - array 2, five component data with $x_0 = .1$ and $\delta_0 = .05$. The pattern of eigenvalues is identical to that seen for the smaller arrays.

Figure 3.6 Magnetic and electric components for the three source gradient eigenvectors - array 2, five component data. The pattern of magnetic fields is as in figure 3.3 but the fields are sampled on a denser grid. The electric fields are tangent to streamlines of electric current flow in the earth. Note that in all cases the relative magnitudes of the electric field vectors are proportional to the magnetic field vectors at the same station, and that the electric field vectors are rotated 90 degrees counter-clockwise. The electric field components of these eigenvectors thus also represent gradients (in the electric fields). These three eigenvectors represent a set of canonical field gradients. The eigenvector $\mathbf{v}_3$ is distinctly different from $\mathbf{v}_4$ and $\mathbf{v}_5$, with the current streamlines forming a series of concentric circles about the central station. Vertical field components are non-zero only for this eigenvector. The latter two eigenvectors correspond to the degenerate eigenvalue $\gamma_4 = \gamma_5$ and $\mathbf{v}_5$ is just $\mathbf{v}_4$ rotated by 45 degrees.



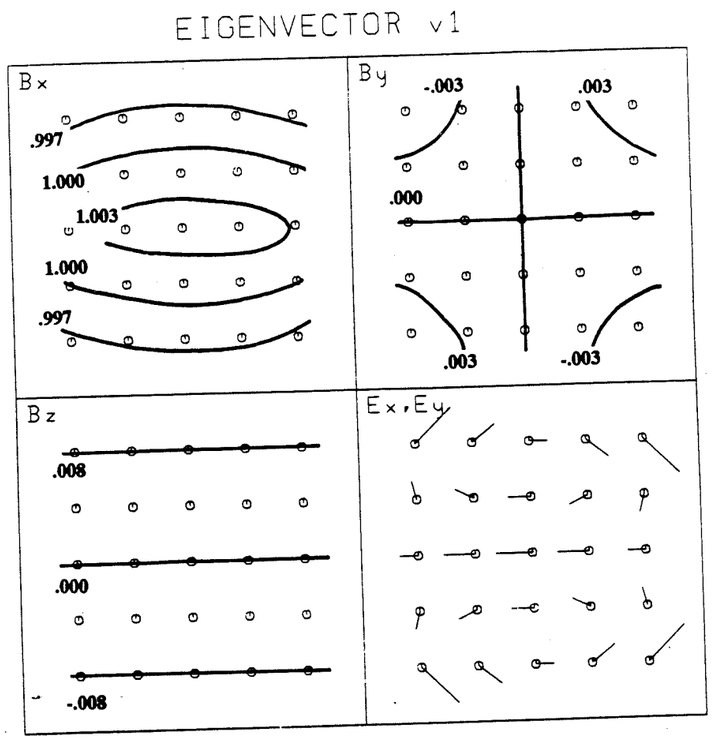


Figure 3.7 More detailed plot of eigenvector $\mathbf{v}_1$ for array 2. In figures 3.7(a), (b) and (c) the north (B_x), east (B_y) and vertical (B_z) components of the fields for $\mathbf{v}_1$ are plotted. The polarization is approximately north-south; There are, however some small perturbations (roughly .3 %) to the quasi-uniform north-south fields. To a good approximation the electric fields are uniform and point in the east-west direction. In 3.7(d) the deviations of the electric fields from a uniform east-west field are plotted. These are again on the order of .3% of the dominantly uniform fields. The fields correspond to two current vortices which result in an enhancement of the dominantly east-west current flow in the center of the array and a reduction of this current at the top and bottom of the array.

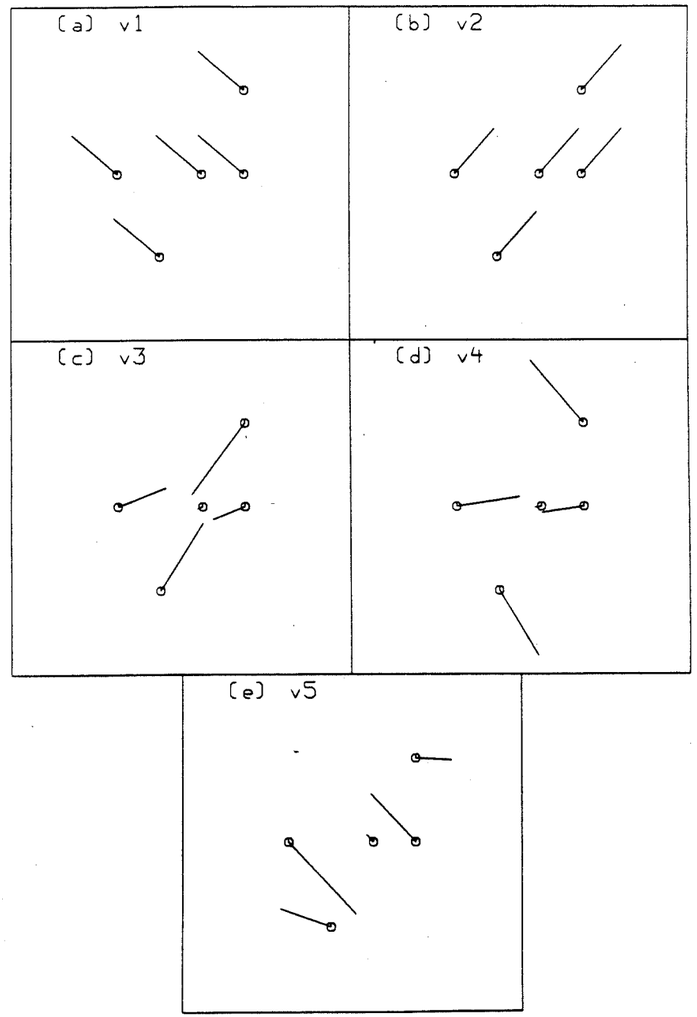


Figure 3.8 Eigenvectors $\mathbf{v}_i$, $i = 1, 5$ for $\Sigma_{\mathbf{H}}$ - irregular array 4. Figures 3.8(a) and (b) give plane wave source eigenvectors; 3.8(c)-(e) give gradient source eigenvectors. Although the eigenvectors do not have the simple symmetric form seen with regular arrays, the two clusters of eigenvalues still correspond to the two types of sources.

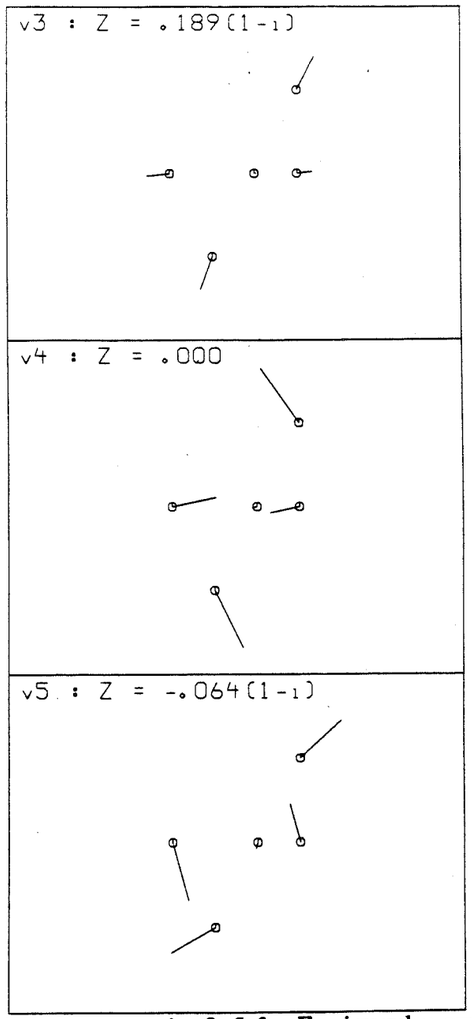


Figure 3.9 Gradient eigenvectors $\mathbf{v}_i$, $i = 3, 5$ for Σ - irregular array 4. The real and imaginary parts of the vertical fields are given at the top center of each figure. Note that unlike the symmetric arrays the vertical fields are non-zero for more than one eigenvector, and that the horizontal field components of the eigenvectors depend on whether vertical fields are included or not. (See figure 3.3.8). Note that by taking linear combinations of these eigenvectors the canonical eigenvectors may be recovered.

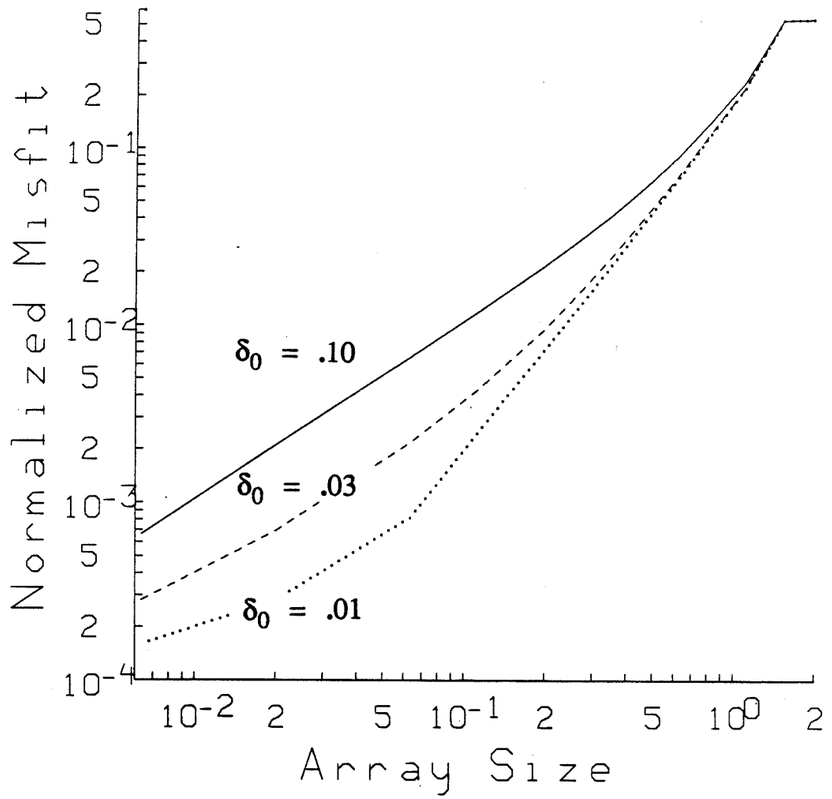


Figure 3.10 Fractional RMS deviation of $\mathbf{v}_1$, $\mathbf{v}_2$ from uniform source model for varying values of x_0 and δ_0. Solid line, $\delta = .1$, dashed line $\delta = .03$, dotted line $\delta = .01$. The fit remains reasonable until the array size is nearly equal to the source wavelength.

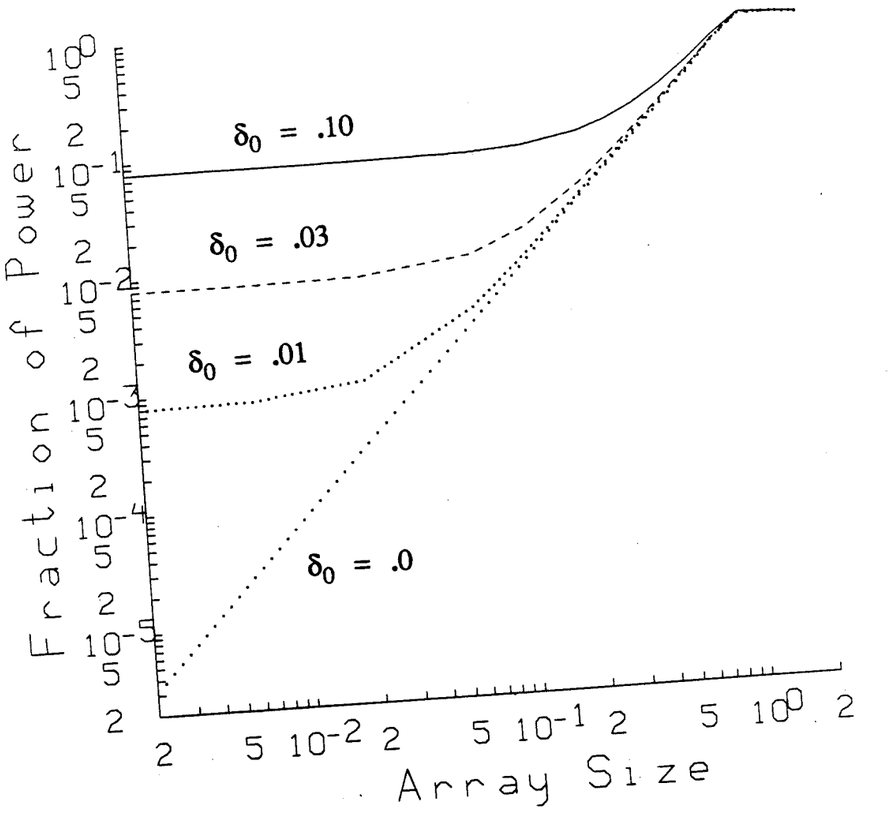


Figure 3.11 Ratio of gradient power to plane wave power γ_3/γ_1 from Σ for varying values of x_0 and δ_0. From top, $\delta = .1$, $\delta = .03$, $\delta = .01$. Dotted line at bottom is γ_3/γ_1 for Σ_H; Equivalently the bottom line corresponds to the case $\delta_0 = 0.0$.

Chapter 4

Statistical Theory: Modeling and Parameter Estimation

In Chapter 2 we showed that when the external source potentials were restricted to a p dimensional space Φ that the idealized, noise free, data vectors, consisting of Fourier transformed magnetic and electric field components measured at a finite number of sites, would lie in a finite dimensional space $\boldsymbol{R}$, the response space. In Chapter 3 we demonstrated that for sources varying randomly in space the external sources could be well approximated as linear combinations of a small number of fundamental modes. This result justifies the use of the response space formulation. In the particular case of interest to us here, where the array size is small compared to typical source dimensions, we have seen that the fields can be well approximated by plane wave sources of two polarizations ($p = 2$) and that this approximation can be improved by adding three terms representing pure gradients of the source fields ($p = 5$).

In Chapter 2 we heuristically justified the use of the dominant eigenvectors of the SDM to estimate the response space. In this chapter we give a more rigorous statistical treatment of the parameter estimation problem. Let $\mathbf{X}_k$ denote the fundamental m-dimensional complex data vector consisting of Fourier coefficients from the k^{th} data segment for all m measured channels of data. If we assume that the external sources are elements of the finite dimensional space Φ, and we allow for noise in all channels, our statistical model for $\mathbf{X}_k$ is

$$\mathbf{X}_k = \mathbf{b}_k + \mathbf{e}_k = \sum_{j=1}^{p} \alpha_{jk}\mathbf{u}_j + \mathbf{e}_k \qquad k = 1, N \qquad (4.0.1a)$$

The vectors $\mathbf{u}_j$, $j = 1, p$ in (4.1.0a) represent the fields seen at the n stations in the array for the j^{th} source basis function ϕ_j; the complex numbers α_{jk}, $j = 1, p$ $k = 1, N$ are the coefficients of the basis functions which define the source 'polarization' for the k^{th} time window (these give the usual source

polarization for the plane wave source case with $p = 2$), and $\mathbf{e}_k$ is the the error or noise vector for the k^{th} time window. In matrix notation we will write

$$\mathbf{X}_k = \mathbf{U}\boldsymbol{\alpha}_k + \mathbf{e}_k \tag{4.0.1b}$$

for a single data vector or

$$\boldsymbol{\Xi} = \mathbf{U}\boldsymbol{\alpha} + \mathbf{E} \tag{4.0.1c}$$

for all of the data. Here $\boldsymbol{\Xi}$ is the $m \times N$ data matrix, $\mathbf{U}$ is the $m \times p$ matrix whose j^{th} column is $\mathbf{u}_j$, $\mathbf{E}$ is the error matrix (whose k^{th} column is $\mathbf{e}_k$) and $\boldsymbol{\alpha}$ is the matrix of polarization parameters (whose k^{th} column is $\boldsymbol{\alpha}_k$).

This sort of model has been applied to geomagnetic data analysis by Jupp (1978) and Park and Chave (1984). Both of these papers considered the use of a model with measurement noise in all channels for estimating the magnetotelluric impedance tensor for single station data. As another geophysical application, Backus *et al.* (1981) developed a similar model and applied it to the problem of the driving forces of plate motions.

More generally, models of the form (4.0.1) for real valued data vectors $\mathbf{X}_k$ have been extensively described in the statistical, psychometric, and econometric literature. These have been referred to under a number of names including *linear functional equation* models (Sprent, 1966; Gleser and Watson, 1973), *structural relationship* models (Moran, 1958), *factor analysis* models (Thurstone, 1947; Anderson and Rubin, 1956; Morrison, 1967) and *multivariate errors-in-variables* models (Gleser, 1981). Anderson (1984) gives a thorough review of all of these (and other similar) models. Although these models all can be expressed in the general form of (4.0.1), they differ from each other in two principal ways. Most importantly, these models make different assumptions about the matrix of 'polarization' coefficients $\boldsymbol{\alpha}$. Specifically, for the linear functional equation and MEV models the coefficients $\alpha_{jk}'s$ are treated as fixed but unknown nuisance parameters,

while in the structural relationship and factor analysis models the $\alpha_{jk}'s$ are treated as random variables. Gleser (1983) discusses the distinctions between these two approaches and points out a large number of very close parallels. Although there are some important differences between these approaches which effect properties of maximum likelihood estimators (Anderson and Rubin, 1956; Mak and Chan, 1982) and various technical aspects of the large sample theory, for most practical purposes the distinctions are not particularly important. As Gleser (1982) points out, results for one sort of model correspond to a parallel (if not identical) result in the other.

In this chapter we will use results which have been obtained for both sorts of models. For most purposes it will be more convenient to treat the polarization parameters as random vectors. This assumption is consistent with the random source model of Chapter 3. Furthermore, with this point of view, and with the assumption of joint multivariate Gaussian distributions for the polarization vectors $\boldsymbol{\alpha}_k$ and the error vectors $\mathbf{e}_k$, the sample SDM $\mathbf{S}$ will have a complex Wishart distribution (Goodman, 1962; Brillinger, 1981). Many standard results from the theory of multivariate statistical analysis are available for the real analogue of this distribution; these can generally be easily adapted to the complex case considered here (Brillinger, 1981). At the same time, however, there are some nice results available on the asymptotic properties of the estimates for the case where the $\alpha_{jk}'s$ are treated as fixed parameters (Gleser, 1981). For our discussion of these results, then, we will adopt this alternative point of view. In fact, the results given in Gleser (1983) can be used to translate asymptotic results from one sort of model to analogous results for the other. Thus for instance, the large sample results of Gleser (1981) for the case of fixed (non-random) polarization parameters discussed in section 4.1, can be applied to the case of random $\alpha_{jk}'s$ if certain technical conditions are restated slightly. In practice then, the inconsistency in our point of

view does not present any serious problems.

The other distinction between the various models relates to the way in which the parameters are specified. As we have discussed in Chapter 2, the model of (4.0.1) is over parametrized. In order to uniquely identify the vectors $\mathbf{u}_j$ it is necessary to make some additional assumptions. The nature of these identifiability assumptions varies for the different model types. As we have discussed above, the restrictions imposed to uniquely identify the vectors $\mathbf{u}_j$ do not effect the response plane $\boldsymbol{R}$. From the standpoint of estimating $\boldsymbol{R}$ these distinctions are thus not important. Anderson and Rubin (1956) and Anderson (1984) discuss a number of ways to guarantee the unique identification of the model parameters. We have proposed one of these methods in Chapter 2 - i.e. requiring that the matrix of parameters $\mathbf{U}$ satisfy the constraints $\mathbf{WU} = \mathbf{C}$. The 'normal station' type constraint is a special case of this which specifies that $\mathbf{U}$ have the form

$$\mathbf{U} = \begin{bmatrix} \mathbf{I}_p \\ \mathbf{T} \end{bmatrix}$$

With this method of identification the statistical model is a multivariate errors-in-variables (MEV) model of the sort discussed by Gleser (1981, 1983). We will use this identifiability assumption for most purposes in this chapter, although here again, we will consider variants on this specific model at times.

Initially, then, we will discuss estimation for the MEV model in the standard case in which the error covariance has the isotropic form $\boldsymbol{\Sigma}_N = \boldsymbol{E}(\mathbf{e}_k\mathbf{e}_k^*) = \sigma^2\mathbf{I}_m$. We will generalize this first to the case where the form of the error covariance $\boldsymbol{\Sigma}_N$ is arbitrary but known and then to the case where $\boldsymbol{\Sigma}_N$ is unknown. For this last case we will see that we must assume a parametric form for $\boldsymbol{\Sigma}_N$ which allows for unique parameter identification. We will discuss in particular the case where $\boldsymbol{\Sigma}_N$ can be written as a linear combination of known matrices (with unknown coefficients), and we will describe a scheme which we have implemented for

estimating the parameters of this model.

The finite-dimensional response space model can only be approximately correct. Any deviation from this approximation must be included in the error vector in (4.0.1). In particular, if we assume a plane wave source model ($p = 2$) the gradient terms would be part of the noise. Although these gradient terms are clearly signal, there are situations in which it will be more reasonable to treat them as noise. For small arrays the horizontal field gradients will be small - the power in these gradients may not be substantially above the background system noise levels (although the vertical field components of the gradient vectors may still be large). In this case the estimation of the gradient vectors of the response space is difficult and of questionable value. Furthermore, when the array is approximately a straight line the meaningful estimation of the gradient response vectors is, as noted above, impossible. In these cases it will probably be best to treat the gradients as noise. Using the results from Chapter 3, we will show that when these gradient terms are treated as noise that the resulting error covariance matrix can be written approximately as a linear combination of known matrices so that our parameter estimation scheme is applicable to this situation. We will demonstrate this approach with data from the EMSLAB MT line and from the series of 5 station MV arrays described in Chapter 2.

The primary goal of this chapter is to develop statistical parameter estimation and inference procedures for the model of (4.0.1) that are relevant to geomagnetic array data. We will see, however, that to accomplish this we will require a specific, and fairly detailed, model for the structure in the noise (including gradient related source effects on vertical magnetic and electric fields). Finding a reasonable model for the noise requires careful exploratory data analysis (coupled with an analysis of the physics involved). This chapter develops some general concepts and statistical tools which are useful both for formal statistical inference

and for exploratory purposes. Many of the applications described at the end of this chapter are of an exploratory nature, and it would be fair to say that many of these results, particularly for the MT case, are still quite tentative. Additional work will be required to decide on the 'definitive yet practical' model for a multivariate analysis of geomagnetic array data.

Finally, we note that many of the statistical methods described below have been described in the statistical literature for the case of real data vectors $\mathbf{X}_k$. All of our applications involve complex data vectors. In general the extension of the results for the real case to the complex case are straightforward; as a result we will generally site the relevant results from the real case as justification for our methods without offering an explicit proof that the result still holds for the complex case. We have attempted to verify the validity of the extension in all relevant cases, and have made necessary small modifications. The complex multivariate normal distribution, which we will use on occasion in this discussion, is described in Goodman (1962) and in Giri (1977); many applications of this distribution, mostly in a time series context, are described in Brillinger (1981). Since these results are fairly standard we will use them freely without reference.

4.1: Parameter Estimation for the MEV Model

We first consider the estimation of the parameters of the MEV model

$$\Xi = \begin{bmatrix} \Xi_1 \\ \Xi_2 \end{bmatrix} = \mathbf{U}\boldsymbol{\alpha} + \mathbf{E} = \begin{bmatrix} \mathbf{I}_p \\ \mathbf{T} \end{bmatrix} \boldsymbol{\alpha} + \mathbf{E} \tag{4.1.1}$$

where the error vectors $\mathbf{e}_i$ (i.e. the columns of $\mathbf{E}$) are independent identically distributed random vectors with covariance $E(\mathbf{e}_i\mathbf{e}_i^*) = \sigma^2\mathbf{I}_m$, and where $\boldsymbol{\alpha}$ is a fixed (non-random) matrix. The unknown parameters in this problem are the elements of the complex matrices $\mathbf{T}$ and $\boldsymbol{\alpha}$ and the error variance σ^2. The elements of $\boldsymbol{\alpha}$ are nuisance parameters; it is the matrix $\mathbf{T}$ which we are really interested in

estimating. Our treatment of this problem closely follows Gleser (1981) who considered (4.1.1) for real (i.e. not complex) parameters.

There are three criteria which have been proposed to justify estimates for this problem. The first, the ordinary least squares estimate (LSE), is based on minimization of the norm of the matrix of residuals

$$\| R(\mathbf{T},\boldsymbol{\alpha};\Xi) \| = \| \Xi - \mathbf{T}\boldsymbol{\alpha} \| \tag{4.1.2}$$

where the matrix norm $\|\cdot\|$ is defined by $\| \mathbf{A} \| = (\mathbf{Tr}[\mathbf{A}\mathbf{A}^*])^{1/2}$. The second, proposed by Sprent (1966) is the generalized least squares estimate (GLSE) based on minimization of the norm of the normalized residual matrix

$$\| Q(\mathbf{T};\Xi) \| = \| (\mathbf{I}_{m-p} + \mathbf{T}\mathbf{T}^*)^{-1/2}(\Xi_2 - \mathbf{T}\Xi_1) \| \tag{4.1.3}$$

The third criterion, maximum likelihood estimation (MLE) assumes that the errors are Gaussian and maximizes the likelihood function, which for complex Gaussian errors can be written (Goodman, 1962)

$$L(\mathbf{T},\boldsymbol{\alpha},\sigma^2,\Xi) = (2\pi\sigma^2)^{-Nm} \exp\left\{ -\frac{1}{\sigma^2} \| R(\mathbf{T},\boldsymbol{\alpha};\Xi) \| \right\} \tag{4.1.4}$$

Gleser proves that the estimates of $\mathbf{T}$ obtained from all three of these criteria are identical. These proofs generalize readily to the complex parameter case considered here. The estimates are also identical to those suggested heuristically in Chapter 2. These can be given explicitly in terms of the eigenvectors of $\mathbf{S} = N^{-1}\,\Xi\Xi^*$. Let $\mathbf{V}$ be the matrix whose i^{th} column is the eigenvector $\mathbf{v}_i$ associated with the i^{th} largest eigenvalue of $\mathbf{S}$ and partition $\mathbf{V}$ as

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix} \qquad \text{where} \quad \mathbf{V}_{11} \text{ is } p\times p$$

Then the estimate of $\mathbf{T}$ is

$$\hat{\mathbf{T}} = \mathbf{V}_{21}\mathbf{V}_{11}^{-1} \tag{4.1.5}$$

To discuss the computation of estimation errors we use the expression for $\mathbf{X}_k$ given in (4.0.1b) and expand the cross power matrix for the k^{th} data vector

$$\mathbf{X}_k\mathbf{X}_k^* = \mathbf{U}\boldsymbol{\alpha}_k\boldsymbol{\alpha}_k^*\mathbf{U}^* + \mathbf{U}\boldsymbol{\alpha}_k\mathbf{e}_k^* + \mathbf{e}_k\boldsymbol{\alpha}_k^*\mathbf{U}^* + \mathbf{e}_k\mathbf{e}_k^* \tag{4.1.6}$$

Set

$$\mathbf{Z}_k = \mathbf{e}_k\mathbf{e}_k^* - \sigma^2\mathbf{I}_m + \mathbf{U}\boldsymbol{\alpha}_k\mathbf{e}_k^* + \mathbf{e}_k\boldsymbol{\alpha}_k^*\mathbf{U}^* \tag{4.1.7}$$

Then, substituting (4.1.7) into (4.1.6), and noting that $\boldsymbol{E}(\mathbf{Z}_k) = 0$ we have

$$\mathbf{X}_k\mathbf{X}_k^* = [\ \mathbf{U}\boldsymbol{\alpha}_k\boldsymbol{\alpha}_k^*\mathbf{U}^* + \sigma^2\mathbf{I}_m\ ] + \mathbf{Z}_k = \boldsymbol{E}(\mathbf{X}_k\mathbf{X}_k^*) + \mathbf{Z}_k \tag{4.1.8}$$

We may thus write for the SDM

$$\mathbf{S} = N^{-1}\sum_{k=1}^{N}\mathbf{X}_k\mathbf{X}_k^* = \mathbf{U}[N^{-1}\boldsymbol{\alpha}\boldsymbol{\alpha}^*]\mathbf{U}^* + \sigma^2\mathbf{I}_m + N^{-1}\sum_{k=1}^{N}\mathbf{Z}_k \tag{4.1.9}$$

$$= \boldsymbol{E}(\mathbf{S}) + \overline{\mathbf{Z}} = \boldsymbol{\Sigma} + \overline{\mathbf{Z}}$$

As we have discussed in Chapter 2, the p dominant eigenvectors of $\boldsymbol{\Sigma} = \boldsymbol{E}(\mathbf{S})$ determine the response space (and hence $\mathbf{T}$) exactly. The SDM $\mathbf{S}$ is perturbed from its expected value $\boldsymbol{\Sigma}$ by the random matrix $\overline{\mathbf{Z}}$. As a result, the eigenvectors of $\mathbf{S}$, and thus the estimate $\hat{\mathbf{T}}$, will be randomly perturbed from the true parameters $\mathbf{T}$. The magnitude of these perturbations determines the estimation errors. Equation (4.1.9) gives an exact expression for $\mathbf{S}$ in terms of model parameters and the error vectors $\mathbf{e}_k$. The relation between the error vectors $\mathbf{e}_k$ and the perturbations to $\boldsymbol{\Sigma}$ is very simple and, at least in some cases (e.g. Gaussian errors) it will be possible to give an explicit (albeit complicated) expression for the exact distribution of $\mathbf{S}$ (for Gaussian errors $\mathbf{S}$ will have a non-central complex Wishart distribution (e.g. Brillinger, 1981)). The estimates $\hat{\mathbf{T}}$, on the other hand, are non-linear functions of $\mathbf{S}$ so the computation of the exact distribution of $\hat{\mathbf{T}}$ will be impossible in practice. We must then turn to approximate methods.

The standard statistical approach to this approximation is to compute the asymptotic properties of $\hat{\mathbf{T}}$ in the limit as the sample size $N \to \infty$. We give a qualitative outline of the idea behind this approach. In the case at hand we note that $\overline{\mathbf{Z}}$, the perturbation to $\boldsymbol{\Sigma}$, is the average of N independent (but not identically distributed) random matrices, $\mathbf{Z}_k$. Since $\boldsymbol{E}(\mathbf{Z}_k) \equiv 0$ we would expect that as N gets large and we average over more and more terms that $\overline{\mathbf{Z}}$ should converge to zero. More importantly, we may invoke (some version of) the central limit theorem (CLT) to demonstrate that $N^{1/2}\overline{\mathbf{Z}}$ converges to a random matrix $\mathbf{G}$ whose elements are jointly Gaussian with zero means. The best known standard CLT requires independent identically distributed random variables (e.g. Lamperitti, 1966) but there are many variants of the CLT which relax some or all of these requirements (Gnedenko and Kolmogorov, 1967). In particular with sufficient regularity it is shown in Gleser (1981) that for the situation considered here, for N large we have

$$\mathbf{S} = \boldsymbol{\Sigma} + N^{-1/2}\mathbf{G} + O(N^{-1})$$

where the elements of $\mathbf{G}$ are jointly Gaussian.

For the second step of this asymptotic approach we note that since $\mathbf{S}$ differs from $\boldsymbol{\Sigma}$ by a small perturbation, we may linearize the relationship between $\hat{\mathbf{T}}$ and $\mathbf{S}$ so that we can write $N^{1/2}(\hat{\mathbf{T}} - \mathbf{T})$ (approximately) as a linear function of the perturbation $\mathbf{G}$. It then follows easily that the elements of $N^{1/2}(\mathbf{T} - \hat{\mathbf{T}})$ are approximately jointly Gaussian with zero mean and covariances which can be easily computed from the linearization in terms of the covariances of the elements of $\mathbf{G}$.

Note that other approximate approaches are possible. Backus *et al.* (1981) and Park and Chave (1984) also used a perturbation approach to compute approximate estimation errors in their applications of this model. In these papers the authors expanded $\mathbf{S}$ as in (4.1.9)

$$\mathbf{S} = N^{-1}[\ \mathbf{U}\boldsymbol{\alpha}\boldsymbol{\alpha}^*\mathbf{U}^* + (\mathbf{U}\boldsymbol{\alpha}\mathbf{E}^* + \mathbf{E}\boldsymbol{\alpha}^*\mathbf{U}^*) + \mathbf{E}\mathbf{E}^*\ ] = N^{-1}[\ \mathbf{U}\boldsymbol{\alpha}\boldsymbol{\alpha}^*\mathbf{U}^* + \sigma\mathbf{G}_1 + \sigma^2\mathbf{G}_2\ ]$$

where σ^2 is here the noise power normalized by the signal power. Assuming $\sigma \ll 1$, they then used a perturbation expansion in the small parameter σ to write the perturbations to the eigenvectors of $\mathbf{\Sigma}$ approximately as linear combinations of the elements of the error vectors. In a sense the main difference between this approach and the standard asymptotic approach used in statistics lies in the choice of the small parameter used for the perturbation expansion. Although the standard asymptotic statistical theory can be more rigorously justified, there are certainly circumstance (i.e. when $N^{-1/2}$ is not small but σ is) where this alternative approach will be preferable. This is certainly the case for the problem considered in Backus *et al.* (1981), but probably not for the MT impedance tensor estimation problem considered by Park and Chave (1984).

Gleser (1981) discusses the large sample properties (asymptotics) of the MEV estimate, using the approach outlined above. To prove any results about the large sample properties of the estimates $\hat{\mathbf{T}}$ we must assume that the nuisance parameters α_{jk} are well behaved as $N \rightarrow \infty$. Specifically we require the existence of the limit

$$\lim_{N\rightarrow\infty} N^{-1}\boldsymbol{\alpha}\boldsymbol{\alpha}^* = \mathbf{\Sigma}_\alpha$$

and we require that $\mathbf{\Sigma}_\alpha$ be non-singular. The matrix $\mathbf{\Sigma}_\alpha$ gives the limiting second moments of the source polarization parameters - i.e. it gives the long term average distribution of power in the different possible source polarizations. The assumptions required for 'nice' large sample properties are thus quite reasonable - as more and more data are collected the average properties of the sources must settle down to a long term average, and all possible source polarizations which we have included in the model must have finite power (on average).

With these assumptions Gleser proves that the estimates of $\mathbf{T}$ given in (4.1.5) are strongly consistent, satisfying $\hat{\mathbf{T}}^{(N)} \rightarrow \mathbf{T}$ with probability one; i.e. as N gets large the estimate for sample size N, $\hat{\mathbf{T}}^{(N)}$, always gets closer and closer to the true

parameter $\mathbf{T}$. Following the procedure outlined above, Gleser also shows that provided the fourth moments of the error vectors $\mathbf{e}_k$ exist, the estimates $\hat{\mathbf{T}}$ are asymptotically normal (i.e. the joint distribution of the elements of $N^{1/2}(\hat{\mathbf{T}} - \mathbf{T})$ converges to a multivariate Gaussian distribution as $N \rightarrow \infty$.) There are some differences between the complex case and the real case treated by Gleser, but the final result is the same. Because of the differences for the complex case we outline Gleser's proof and indicate the modifications needed in Appendix B.

The exact form of the asymptotic covariance depends on third and fourth moments of the error distribution. These are not known and they would be difficult to estimate. To obtain useful expressions for the asymptotic covariance we must therefore assume that the errors are Gaussian (more precisely we must assume that all moments up to the fourth moments are the same as those for some multivariate Gaussian distribution). Note that if the *noise* process is stationary then the frequency domain noise vectors $\mathbf{e}_k$ will be (approximately) Gaussian. This will not always be true, particularly if deviations from a plane wave source model are included in the noise (Egbert and Booker, 1986). It is probable that the error estimates given here will be optimistic unless a robust variant of the basic MEV estimate described here is used. With this caveat, the asymptotic covariances of the ij and $i'j'$ elements of $\mathbf{T}$ are given by

$$\mathbf{Cov}(\hat{T}_{ij}, \hat{T}_{i'j'}) = \tag{4.1.10}$$
$$N^{-1}\sigma^2 \left[\sigma^2\mathbf{\Sigma}_\alpha^{-1}(\mathbf{I}_p + \mathbf{T}^*\mathbf{T})\mathbf{\Sigma}_\alpha^{-1} + \mathbf{\Sigma}_\alpha^{-1}\right]_{jj'} \left[\mathbf{I}_{m-p} + \mathbf{T}\mathbf{T}^*\right]_{ii'}$$

To compute these estimation errors the error variance σ^2 and the matrix $\mathbf{\Sigma}_\alpha$ must be estimated. Consistent estimates of these quantities are given in Gleser; these results remain applicable to the complex case considered here. As in (4.1.5), partition the matrix eigenvectors $\mathbf{V}$ to obtain the $p \times p$ and $(m-p) \times p$ matrices $\mathbf{V}_{11}$ and $\mathbf{V}_{21}$. Let γ_i, $i=1,m$ be the ordered eigenvalues of $\mathbf{S}$ and set

$\mathbf{D} = \mathbf{diag}(\gamma_1, \cdots, \gamma_p)$. Then we estimate the desired quantities by

$$\hat{\Sigma}_\alpha = N^{-1}\left[\mathbf{V}_{11}\mathbf{D}\mathbf{V}_{11}^* - \hat{\sigma}^2\mathbf{V}_{11}\mathbf{V}_{11}^* \right] \tag{4.1.11a}$$

$$\hat{\sigma}^2 = \frac{1}{m-p} \sum_{i=p+1}^{m} \gamma_i \tag{4.1.11b}$$

Together, the results given in (4.1.10) and (4.1.11) allow us to compute estimation errors of any linear combination of the parameters **T**.

4.2: Estimation with a General, Known Error Covariance Matrix

We now consider the estimation of the response space $\boldsymbol{R}$ for the general case where the noise covariance Σ_N is not isotropic. In contrast to the previous section where we assumed the MEV model to uniquely identify the parameters which represent the response space, we here represent the response space as the projection matrix $\mathbf{P}_R$ which projects $\mathbf{C}^m$ onto $\boldsymbol{R}$. As we have already indicated in section 2.4, the eigenvector estimate of the response space $\hat{\boldsymbol{R}}$ of (2.4.7) will generally be biased, yielding incorrect answers even for an infinite amount of data. We first give an explicit form for this bias.

To do this we express the total expected covariance as the sum of signal and noise parts $\boldsymbol{E}(\mathbf{S}) = \Sigma = \Sigma_S + \Sigma_N$. We will assume that the signal matrix Σ_S is in some sense large relative to the noise matrix Σ_N and use a perturbation approach. This perturbation technique can also be used for the direct computation of estimation errors as we will discuss later. Note that this perturbation technique is similar to that used by Backus *et al.* (1981) and Park and Chave (1984) to calculate estimation errors for the model of (4.0.1).

Watson (1983) used results from Kato (1966) to give a treatment for this problem which is at once elegant, mathematically rigorous and clear. We will base our discussion here on a generalization to the case of complex matrices of the formulation and results given in Watson. We give a general statement of the setup and the key result since we will use this result several times in the sequel.

Let $\mathbf{T}_0$ and $\mathbf{T}_1$ be $m \times m$ complex Hermetian matrices and set

$$\mathbf{T}(\varepsilon) = \mathbf{T}_0 + \varepsilon\mathbf{T}_1 \tag{4.2.1}$$

The spectral representation of $\mathbf{T}_0$ in terms of its eigenvalues and eigenvectors is

$$\mathbf{T}_0 = \sum_{i=1}^{m} \gamma_i \mathbf{v}_i \mathbf{v}_i^* \tag{4.2.2}$$

We would like to consider the perturbations of the eigenvectors and eigenvalues of $\mathbf{T}(\varepsilon)$ from those of $\mathbf{T}_0$ for small values of ε. To do this in the general case where degenerate eigenvalues are allowed requires a bit of care. If a group of k eigenvalues are equal, say $\gamma_1 = \gamma_2 = \cdots = \gamma_k$ the corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ are not uniquely defined, and a discussion of perturbations becomes rather ambiguous. Instead, the set of vectors satisfying

$$\mathbf{T}_0\mathbf{v} = \gamma_1\mathbf{v}$$

form a k dimensional subspace of $\mathbf{C}^m$, the eigenspace $\boldsymbol{V}_1$. If $\mathbf{v}_i$, $i = 1, k$ is any orthonormal basis for $\boldsymbol{V}_1$ we may write the matrix which projects onto $\boldsymbol{V}_1$ as $\mathbf{P}_1 = \sum_{i=1}^{k} \mathbf{v}_i\mathbf{v}_i^*$. Although the vectors $\mathbf{v}_i$, $i = 1, k$ are not uniquely determined, the matrix $\mathbf{P}_1$ is. We modify our notation slightly and let $\gamma_1, \cdots, \gamma_q$ denote the q ($q \le m$) distinct eigenvalues of $\mathbf{T}_0$ and we let $\mathbf{P}_i$, $i = 1, q$ be the projection matrices for the corresponding eigenspaces $\boldsymbol{V}_i$. Note that the projections satisfy

$$\mathbf{P}_i^* = \mathbf{P}_i \qquad\qquad \mathbf{P}_i\mathbf{P}_j = \delta_{ij}\mathbf{P}_j$$

Then (4.2.2) can be written

$$\mathbf{T}_0 = \sum_{i=1}^{q} \gamma_i \mathbf{P}_i$$

We now wish to characterize the perturbations to the projection matrices $\mathbf{P}_i$, $i = 1, q$ caused by the addition of the small matrix $\varepsilon\mathbf{T}_1$ to $\mathbf{T}_0$.

In general we would expect that $\mathbf{T}(\varepsilon)$ will have m distinct eigenvalues but that these will condense to q clusters centered at the $\gamma_i's$ as $\varepsilon \to 0$. The eigenvectors corresponding to the k_i eigenvalues of the i^{th} cluster define a projection matrix which we would expect to converge to $\mathbf{P}_i$ as $\varepsilon \to 0$. It is the form of this perturbed projection matrix which we seek. The proof of this convergence and the derivation of the perturbed projection matrix is accomplished in Watson for the case of real symmetric matrices $\mathbf{T}_0$, $\mathbf{T}_1$ via a fairly simple application of Cauchy's theorem from complex analysis. the complex Hermetian case considered here is virtually identical so we simply state the results. Let C_j denote a contour in the complex plane which contains only the eigenvalue γ_j (which is of multiplicity k_j). Let $\gamma_i(\varepsilon)$, $\mathbf{v}_i(\varepsilon)$ denote the eigenvalues and eigenvectors of $\mathbf{T}(\varepsilon)$. Then

$$\#\gamma_i(\varepsilon) \text{ inside } C_j = k_j + O(\varepsilon^3) \tag{4.2.3a}$$

and, if we define

$$\hat{\mathbf{P}}_j(\varepsilon) = \sum_{i:\gamma_i(\varepsilon)\ C_j} \mathbf{v}_i \mathbf{v}_i^*$$

Then

$$\hat{\mathbf{P}}_j(\varepsilon) = \mathbf{P}_j + \varepsilon \sum_{k \neq j} \frac{\mathbf{P}_k\mathbf{T}_1\mathbf{P}_j + \mathbf{P}_j\mathbf{T}_1\mathbf{P}_k}{\gamma_j - \gamma_k} + O(\varepsilon^2) \tag{4.2.3b}$$

Equation (4.2.3), then, gives a general expression, correct to first order in ε, for the perturbed projections and is the key result for our applications. We can use this to approximate the asymptotic bias in the eigenvector estimate of the response space.

For simplicity we initially assume that the eigenvalues of $\mathbf{\Sigma}_S$ are all equal to γ, so that $\mathbf{\Sigma}_S = \gamma\mathbf{P}_R$ where $\mathbf{P}_R$ is the rank p projection onto the p dimensional space $\boldsymbol{R}$. Let $\sigma_N^2 = \mathbf{Tr}[\mathbf{\Sigma}_N]$ be the total noise power, set $\mathbf{\Sigma}_N' = (1/\sigma_N^2)\mathbf{\Sigma}_N$ so that $\mathbf{\Sigma}_N'$ is non-dimensional and of order one, and let $\varepsilon = \sigma_N^2/\gamma$ be the noise-to-signal power which we assume small ($\varepsilon \ll 1$). The eigenvector estimate of the response space $\hat{\boldsymbol{R}}$ is formed from the p dominant eigenvectors of

$$\mathbf{\Sigma} = \gamma\mathbf{P}_R + \sigma_N^2\mathbf{\Sigma}_N' = \gamma(\mathbf{P}_R + \varepsilon\mathbf{\Sigma}_N')$$

These eigenvectors are identical to those of $\mathbf{P}_R + \varepsilon\mathbf{\Sigma}_N'$, so we may use (4.2.3) to obtain an expression for the projection onto the estimated response space $\hat{\mathbf{P}}_R$. In this case

$$\mathbf{T}_0 = \mathbf{P}_R \quad q = 2 \quad \gamma_1 = 1 \quad \gamma_2 = 0 \quad \mathbf{P}_1 = \mathbf{P}_R \quad \mathbf{P}_2 = \mathbf{I}_m - \mathbf{P}_R$$

$$\hat{\mathbf{P}}_R = \mathbf{P}_R + \varepsilon\left[\mathbf{P}_2\mathbf{\Sigma}_N'\mathbf{P}_1 + \mathbf{P}_1\mathbf{\Sigma}_N'\mathbf{P}_2\right] + O(\varepsilon^2)$$

$$= \mathbf{P}_R + (1/\gamma)\left[(\mathbf{I}_m - \mathbf{P}_R)\mathbf{\Sigma}_N\mathbf{P}_R + \mathbf{P}_R\mathbf{\Sigma}_N(\mathbf{I}_m - \mathbf{P}_R)\right] + O(\varepsilon^2) \tag{4.2.4}$$

Equation (4.2.4) demonstrates that the eigenvector estimate of $\boldsymbol{R}$ is in general asymptotically biased and, since $\left[\mathbf{P}_2\mathbf{\Sigma}_N'\mathbf{P}_1 + \mathbf{P}_1\mathbf{\Sigma}_N'\mathbf{P}_2\right]$, the fractional bias is of order ε (the noise-to-signal ratio).

Equation (4.2.4) generalizes readily to the case where the eigenvalues of $\mathbf{\Sigma}_S$ are not equal. Suppose that

$$\mathbf{\Sigma}_S = \sum_{i=1}^{q} \gamma_i\mathbf{P}_i \qquad \text{so that} \quad \mathbf{P}_R = \sum_{i=1}^{q} \mathbf{P}_i$$

Then, if we estimate $\mathbf{P}_R$ using the p dominant eigenvectors of $\mathbf{\Sigma}$, an application of (4.2.3) yields

$$\hat{\mathbf{P}}_R = \mathbf{P}_R + \sum_{i=1}^{q} \left[\frac{(\mathbf{I}_m - \mathbf{P}_R)\mathbf{\Sigma}_N\mathbf{P}_i + \mathbf{P}_i\mathbf{\Sigma}_N(\mathbf{I}_m - \mathbf{P}_R)}{\gamma_i} \right] \tag{4.2.5}$$

Note that if $\Sigma_N = \sigma^2 \mathbf{I}_m$ we have

$$(\mathbf{I}_m - \mathbf{P}_R)\Sigma_N \mathbf{P}_R = \sigma^2(\mathbf{I}_m - \mathbf{P}_R)\mathbf{P}_R = 0$$

so the bias is zero as expected. More generally, it is not hard to show that the bias term will vanish if and only if the matrices $\mathbf{P}_R$ and Σ_N commute and that this will occur if and only if the response space is an invariant subspace of the matrix Σ_N (i.e. $\mathbf{v} \in \boldsymbol{R} \rightarrow \Sigma_N \mathbf{v} \in \boldsymbol{R}$). We will see that in some cases of interest this property will hold at least approximately.

An asymptotically unbiased estimate of the response space can be computed if the the structure of the noise covariance matrix Σ_N is known or can be consistently estimated up to a multiplicative constant (e.g. Gleser, 1981; Park and Chave, 1984). This can be easily accomplished by using the matrix Σ_N to transform the sample SDM $\mathbf{S}$ so that the isotropic error assumption holds. Specifically, let $\Sigma_N^{1/2}$ be any square root of Σ_N - i.e any matrix satisfying $\Sigma_N^{1/2}(\Sigma_N^{1/2})^* = \Sigma_N$. Note that the existence of such a square root is guaranteed by the Hermetian property of the error covariance matrix Σ_N, and that the construction of a (lower triangular) square root can easily be accomplished using a complex version of the Cholesky decomposition. We will assume that the error covariance matrix is positive definite (this will always be true if there is any instrumental noise) so that the square root matrix is non-singular with inverse Σ_N^{-half}.

Then transforming the data vectors, the response space vectors and the error vectors

$$\mathbf{X}'_k = \Sigma_N^{-1/2}\mathbf{X}_k = \sum_{j=1}^{p} \alpha_{jk}(\Sigma_N^{-1/2}\,\mathbf{u}_j) + \Sigma_N^{-1/2}\,\mathbf{e}_k = \sum_{j=1}^{p} \alpha_{jk}\,\mathbf{u}'_j + \mathbf{e}'_k \qquad (4.2.6)$$

we find that we now have a problem with isotropic errors since

$$\mathbf{E}(\mathbf{e}'\mathbf{e}'^*) = \Sigma_N^{-1/2}\mathbf{E}(\mathbf{e}\mathbf{e}^*)\Sigma_N^{-1/2^*} = \mathbf{I}_m$$

If

$$\mathbf{S}' = N^{-1} \sum_{k=1}^{N} \mathbf{X}'_k \mathbf{X}'^*_k = \Sigma_N^{-1/2}\mathbf{S}\Sigma_N^{-1/2^*}$$

and $\mathbf{v}'_i$, $i = 1, m$ are the eigenvectors of $\mathbf{S}'$ ordered by the magnitude of the corresponding eigenvalues, then as above,

$$\hat{\boldsymbol{R}}' = \mathbf{Sp}\left\{\mathbf{v}'_i : i = 1, p\right\}$$

is an asymptotically unbiased estimate of

$$\boldsymbol{R}' = \mathbf{Sp}\left\{\mathbf{u}'_i : i = 1, p\right\}$$

Consequently,

$$\hat{\boldsymbol{R}} = \mathbf{Sp}\left\{\Sigma_N^{1/2}\mathbf{v}'_i : i = 1, p\right\} \qquad (4.2.7)$$

will be an asymptotically unbiased estimate of $\boldsymbol{R}$.

If we again adopt the MEV model to obtain an identifiable parametrization we can use the transformation of the data described above to convert to the standard MEV problem so that all of the results described in section 4.1 can be used. Specifically, choose the transformation matrix $\Sigma_N^{-1/2}$ so that it can be partitioned

$$\Sigma_N^{-1/2} = \begin{bmatrix} \mathbf{A}_{11} & 0 \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \quad \text{where} \quad \mathbf{A}_{11} \text{ is } p \times p$$

(e.g. form the inverse of Σ_N and then compute the Cholesky decomposition so that $\Sigma_N^{-1/2}$ is lower triangular). Then transforming both sides of (4.1.1) we find

$$\Xi' = \Sigma_N^{-1/2}\Xi = \begin{bmatrix} \mathbf{A}_{11} & 0 \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I}_p \\ \mathbf{T} \end{bmatrix} \boldsymbol{\alpha} + \mathbf{E}'$$

$$= \begin{bmatrix} \mathbf{I}_p \\ \mathbf{A}_{21}\mathbf{A}_{11}^{-1} + \mathbf{A}_{22}\mathbf{T}\mathbf{A}_{11}^{-1} \end{bmatrix} \mathbf{A}_{11}\,\boldsymbol{\alpha} + \mathbf{E}' = \begin{bmatrix} \mathbf{I}_p \\ \mathbf{T}' \end{bmatrix} \boldsymbol{\alpha}' + \mathbf{E}' \qquad (4.2.8)$$

where the columns $\mathbf{e}_i'$ of the error matrix $\mathbf{E}'$ now satisfy the isotropic error condition so that the model takes the standard MEV form of (4.1.1).

The estimates $\hat{\mathbf{T}}'$ computed from the eigenvectors of $\mathbf{S}'$ have all of the properties given in section 4.1. Estimates of the original parameters can be easily obtained in terms of $\hat{\mathbf{T}}'$

$$\hat{\mathbf{T}} = \mathbf{A}_{22}^{-1}\hat{\mathbf{T}}'\mathbf{A}_{11} - \mathbf{A}_{22}^{-1}\mathbf{A}_{21} \tag{4.2.9}$$

Since the elements of $\hat{\mathbf{T}}$ are linear combinations of the elements of $\hat{\mathbf{T}}'$, the asymptotic covariance of these parameters can be easily computed in terms of the asymptotic covariances of the elements $\hat{\mathbf{T}}'$. These are obtained using (4.1.10) and (4.1.11) of section 4.1. Note that this result for estimation errors assumes that $\mathbf{\Sigma}_N$ is known. If $\mathbf{\Sigma}_N$ must be estimated, the computation of estimation errors will have to be modified to take this into account.

4.3 : Some Comments on the Computation of Errors

In Chapter 2 we stressed that the response space provides a more fundamental and natural description of the interelationships of all field components at all stations than the usual transfer functions which assume a special 'normal' reference station. On the other hand, in the discussion of error estimates in the last two sections, we have assumed a MEV model which is essentially equivalent to the usual transfer function parametrization with a special reference station. (Note however, that the MEV estimates take account of the noise in all channels and thus differ from the usual transfer function estimates). This is because we must describe (parametrize) the response space precisely in terms of a set of numbers (parameters) in order to define and compute estimation errors. The parameterization may be specified in many different ways, although all parametrizations describe the same object (R). All parameter estimates will be computed in terms

of the eigenvectors of the SDM $\mathbf{S}$, but the computation of estimation errors will in general depend on the parametrization. In this section we discuss how the results given above for estimation error covariances for the MEV parametrization can be used to compute error covariances for other possible parametrizations. We also discuss briefly how the perturbation methods described above can be used to derive estimation errors.

In Chapter 2 we suggested imposing constraints of the form $\mathbf{W}_1\mathbf{U} = \mathbf{C}$ to uniquely define the vectors $\mathbf{u}_i$ which span R. The MEV parametrization is a special case of this. Conversely, all parametrizations specified by constraints of this form can be transformed to the MEV form with a general error covariance. To demonstrate this we first note that we can rewrite the constraint of (2.3.7) as $(\mathbf{C}^{-1}\mathbf{W}_1)\mathbf{U} = \mathbf{I}_p$ so, without loss of generality, we may assume that $\mathbf{C} = \mathbf{I}_p$. Let $\mathbf{W}_1^T = (\mathbf{w}_1 \ \cdots \ \mathbf{w}_p)$ and pick vectors $\mathbf{w}_i$, $i = p{+}1, m$ such that $\{ \mathbf{w}_i : i = 1,m \}$ is a basis for $\mathbf{C}^m$. Setting $\mathbf{W}_2^T = (\mathbf{w}_{p+1} \ \cdots \ \mathbf{w}_m)$, we have

$$\mathbf{W}\mathbf{U} = \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \end{bmatrix}\mathbf{U} = \begin{bmatrix} \mathbf{I}_p \\ \mathbf{W}_2\mathbf{U} \end{bmatrix}$$

Transforming the data matrix $\mathbf{\Xi}$ by pre-multiplying by $\mathbf{W}$ thus transforms (4.0.1b) to

$$\mathbf{\Xi}' = \mathbf{W}\mathbf{\Xi} = \begin{bmatrix} \mathbf{I}_p \\ \mathbf{W}_2\mathbf{U} \end{bmatrix}\alpha + \mathbf{E}' \tag{4.3.1}$$

where the error vectors $\mathbf{e}_i'$ now have error covariance $\mathbf{\Sigma}'_N = \mathbf{W}\mathbf{\Sigma}_N\mathbf{W}^*$. The model of (4.3.1) is in the standard MEV form (with general error covariance, even if the original error covariance is isotropic). Letting

$$(\mathbf{\Sigma}'_N)^{-1/2} = \begin{bmatrix} \mathbf{A}_{11} & 0 \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

we may transform the data as in (4.2.7) to reduce to a MEV model of the standard

form (4.1.1) with isotropic errors. If $\hat{\mathbf{T}}'$ is the estimate of $\mathbf{T}$ for this transformed model (computed via (4.1.5) from the eigenvectors of $\boldsymbol{\Sigma}_N'^{-\frac{1}{2}}\mathbf{WSW}^*\boldsymbol{\Sigma}_N'^{-\frac{1}{2}\,*}$) then we find that the estimate of $\mathbf{U}$ satisfying the constraint (2.3.7) is given by

$$\hat{\mathbf{U}} = \mathbf{W}^{-1}\begin{pmatrix} \mathbf{I}_p \\ \mathbf{A}_{22}^{-1}\hat{\mathbf{T}}'\mathbf{A}_{11} - \mathbf{A}_{22}^{-1}\mathbf{A}_{11} \end{pmatrix} \tag{4.3.2}$$

With the regularity assumptions of section 4.1 the parameter estimates $\hat{\mathbf{T}}'$ are asymptotically normal with covariances which may be computed using (4.1.10). Again, the elements of $\hat{\mathbf{U}}$ are linear combinations of the $\hat{T}_{ij}'s$ so the asymptotic covariances of the $\hat{U}_{ij}'s$ are easily computed.

The argument given above demonstrates that for any linear constraint of the form (2.3.7) we may transform to a standard MEV problem, use the methods of section 4.1 to compute estimation errors for the transformed problem, and then transform the results back to the desired parametrization. An alternative, more direct, approach is possible. Suppose we have an estimate $\hat{\mathbf{U}}'$ of the parameters $\mathbf{U}'$ with $\mathbf{U}'$ identified by the constraint $\mathbf{W}'\mathbf{U}' = \mathbf{I}_p$ and we wish to find estimates for the parametrization satisfying the constraints $\mathbf{WU} = \mathbf{I}_p$. We can easily compute the estimates of $\mathbf{U}$ directly in terms of $\hat{\mathbf{U}}'$ - i.e. $\hat{\mathbf{U}} = \hat{\mathbf{U}}'(\mathbf{W}\hat{\mathbf{U}}')^{-1}$. In a similar fashion, if we know the asymptotic distribution of $\hat{\mathbf{U}}'$ we can find the asymptotic distribution of $\hat{\mathbf{U}}$.

If $\hat{\mathbf{U}}'$ is asymptotically normal then for the sequence of estimates $\hat{\mathbf{U}}_N'$ based on samples of size N, we have, as $N \to \infty$

$$N^{\frac{1}{2}}\,(\hat{\mathbf{U}}'_N - \mathbf{U}) \to \mathbf{G}$$

where $\mathbf{G}$ is a matrix whose elements are jointly Gaussian. We may thus write

$$\hat{\mathbf{U}}'_N = \mathbf{U} + N^{-\frac{1}{2}}\,\mathbf{G} + O(N^{-1})$$

so to first order in $N^{-\frac{1}{2}}$

$$\hat{\mathbf{U}}_N = (\mathbf{U}' + N^{-\frac{1}{2}}\mathbf{G})\Big[\mathbf{W}(\mathbf{U}' + N^{-\frac{1}{2}}\mathbf{G})\Big]^{-1} + O(N^{-1})$$

For the second term we have

$$\begin{aligned}[\mathbf{W}(\mathbf{U}' + N^{-\frac{1}{2}}\mathbf{G})]^{-1} &= [\mathbf{I}_m + N^{-\frac{1}{2}}(\mathbf{WU}')^{-1}\mathbf{WG}]^{-1}(\mathbf{WU}')^{-1} \\ &= [\mathbf{I}_m - N^{-\frac{1}{2}}(\mathbf{WU}')^{-1}\mathbf{WG} + O(N^{-1})\,](\mathbf{WU}')^{-1} \\ &= (\mathbf{WU}')^{-1} - N^{-\frac{1}{2}}(\mathbf{WU}')^{-1}\mathbf{WG}(\mathbf{WU}')^{-1} + O(N^{-1})\end{aligned}$$

where we have used $(\mathbf{I} + \varepsilon\mathbf{A})^{-1} = \mathbf{I} - \varepsilon\mathbf{A} + O(\varepsilon^2)$ in the second step. Multiplying the two terms out and simplifying we find

$$\hat{\mathbf{U}}_N = \mathbf{U} + N^{-\frac{1}{2}}[\mathbf{I}_m - \mathbf{U}'(\mathbf{WU}')^{-1}\mathbf{W}]\mathbf{G}(\mathbf{WU}')^{-1} + O(N^{-1})$$

or as $N \to \infty$

$$N^{\frac{1}{2}}(\hat{\mathbf{U}}_N - \mathbf{U}) \to [\mathbf{I}_m - \mathbf{U}'(\mathbf{WU}')^{-1}\mathbf{W}]\mathbf{G}(\mathbf{WU}')^{-1} = \mathbf{FGH} \tag{4.3.3}$$

Equation (4.3.3) gives a linear expression for the asymptotic distribution of $\hat{\mathbf{U}} - \mathbf{U}$ in terms of the Gaussian random matrix $\mathbf{G}$ whose covariance is assumed know, so asymptotic estimation errors can be easily computed. Specifically, if the covariances of $\mathbf{G}$ are given by $\sigma_{kk'l'} = \mathbf{Cov}(G_{kl},G_{k'l'})$, then the asymptotic covariances of the elements of $\hat{\mathbf{U}}$ are given by

$$\mathbf{Cov}(\hat{U}_{ij},\hat{U}_{i'j'}) = N^{-1}\sum_{kk'l'} F_{ik}H_{lj}F_{i'k'}H_{l'j'}\sigma_{kk'l'} \tag{4.3.4}$$

where the form of $\mathbf{F}$ and $\mathbf{G}$ can be inferred from (4.3.3). If $\mathbf{U}'$ represents the MEV parametrization so that

$$\hat{\mathbf{U}}' = \begin{pmatrix} \mathbf{I}_p \\ \mathbf{T} \end{pmatrix}$$

then the covariances $\sigma_{kk'l'}$ needed in (4.3.4) are obtained from the results given in (4.1.10) and (4.1.11).

There is yet another approach to the computation of estimation errors. This is based upon the perturbation methods described in (4.2.3)-(4.2.4). Briefly, assuming isotropic errors with variance σ^2 and using (4.1.9) we have $\mathbf{S} = \mathbf{\Sigma} + N^{-1/2}\mathbf{G}$ where the asymptotic distribution of the elements of the matrix $\mathbf{G}$ is complex multivariate Gaussian. The covariance of the elements of $\mathbf{G}$ are given for the Gaussian errors case in Appendix B. We may then use (4.2.3) and (4.2.5) with $\mathbf{T}_0 = \mathbf{\Sigma}$, $\mathbf{T}_1 = \mathbf{G}$ and $\varepsilon = N^{-1/2}$ to find

$$\hat{\mathbf{P}}_R - \mathbf{P}_R = N^{-1/2}\sum_{i=1}^{q}\left[\frac{(\mathbf{I}_m - \mathbf{P}_R)\mathbf{G}\mathbf{P}_i + \mathbf{P}_i\mathbf{G}(\mathbf{I}_m - \mathbf{P}_R)}{\gamma_i - \sigma^2}\right] + O(N^{-1}) \qquad (4.3.5)$$

where the $\gamma_i's$ are the eigenvalues of $\mathbf{\Sigma}$ and the $\mathbf{P}_i's$ are the projections onto the corresponding eigenspaces. Equation (4.3.5) gives an asymptotically correct expression for the random perturbations of the projection $\mathbf{P}_R$ as a linear function of $\mathbf{G}$. The asymptotic covariance of $N^{1/2}(\hat{\mathbf{P}}_R - \mathbf{P}_R)$ can thus easily be calculated as in (4.3.4) using the covariances of the elements of $\mathbf{G}$ given in Appendix B.

4.4:Estimation of the Noise Covariance I: Parameter Identification

To obtain efficient, unbiased estimates of the response plane $\boldsymbol{R}$, together with meaningful estimation errors we must know, or be able to estimate, the noise covariance matrix $\mathbf{\Sigma}_N$ (up to a multiplicative constant). In all of the above discussion we have assumed that $\mathbf{\Sigma}_N$ was known. We now turn our attention to the more realistic situation where $\mathbf{\Sigma}_N$ must be estimated from the data.

Our general model for the population SDM is

$$\mathbf{\Sigma} = E(\mathbf{S}) = \mathbf{\Sigma}_S + \mathbf{\Sigma}_N = \mathbf{U}\mathbf{\Sigma}_\alpha\mathbf{U}^* + \mathbf{\Sigma}_N(\boldsymbol{\theta})$$

where $\mathbf{U}$ and $\mathbf{\Sigma}_\alpha$ are, as above, the parameters which characterize the signal matrix $\mathbf{\Sigma}_S$. The noise covariance $\mathbf{\Sigma}_N$ depends on the vector of l unknown parameters

$\boldsymbol{\theta} = (\theta_1, \cdots \theta_l)$ which we must estimate. As we have discussed above, the unique identification of the parameters $\mathbf{U}$ and $\mathbf{\Sigma}_\alpha$ requires some constraints on $\mathbf{U}$. For the geomagnetic array models which we are considering, these constraints are equivalent to making a 'normal' reference field assumption. When the error covariance also depends on the unknown parameters $\boldsymbol{\theta}$ there is a further, and more difficult parameter identification problem.

To be definite, we identify the signal covariance parameters by making the MEV assumptions so that we have

$$\mathbf{\Sigma} = \begin{bmatrix}\mathbf{I}_p \\ \mathbf{T}\end{bmatrix}\mathbf{\Sigma}_\alpha\begin{bmatrix}\mathbf{I}_p & \mathbf{T}^*\end{bmatrix} + \mathbf{\Sigma}_N(\boldsymbol{\theta}) \qquad (4.4.1)$$

where $\mathbf{\Sigma}_\alpha = N^{-1}\boldsymbol{\alpha}\boldsymbol{\alpha}^*$. The matrix $\mathbf{\Sigma}$ depends upon the signal parameters $\boldsymbol{\tau} = (\mathbf{T}, \mathbf{\Sigma}_\alpha)$ and the noise parameters $\boldsymbol{\theta}$. Suppose there is at least one set of parameters $(\boldsymbol{\tau}, \boldsymbol{\theta})$ such that (4.4.1) is satisfied. Are there $\boldsymbol{\theta}' \neq \boldsymbol{\theta}$, $\boldsymbol{\tau}' \neq \boldsymbol{\tau}$ such that $\mathbf{\Sigma} = \mathbf{\Sigma}_S(\boldsymbol{\tau}') + \mathbf{\Sigma}_N(\boldsymbol{\theta}')$? If so then the parameters $\boldsymbol{\theta}$ and $\boldsymbol{\tau}$ can't be determined from even perfect knowledge of $\mathbf{\Sigma}$ and the separation of the data covariance into uniquely determined noise and signal parts will be impossible. Note that we have already made an assumption about the parametrization of $\mathbf{\Sigma}_S(\boldsymbol{\tau})$ to guarantee the unique identification of $\boldsymbol{\tau}$ from knowledge of $\mathbf{\Sigma}_S$. We will make a similar assumption about the the parametrization of $\mathbf{\Sigma}_N$: we assume that if $\mathbf{\Sigma}_N(\boldsymbol{\theta}_1) = \mathbf{\Sigma}_N(\boldsymbol{\theta}_2)$ then $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$. This eliminates any trivial identifiability problems which can be solved by changing the way we parametrize the noise covariance without changing the class of possible noise models. The question that we address here then, is how general a class of noise models can we assume and still be able to uniquely separate signal and noise. It is clear that $\mathbf{\Sigma}_N$ can't be arbitrary, because if we allow a completely arbitrary noise covariance, one way of fitting the data is to say that everything is noise. We must have *a priori* knowledge of the parametric form of the error covariance in order to have any hope of separating signal from

noise.

Note that the notion of identifiability in statistics can be somewhat more general than this (e.g. Bickel and Doksum, 1977). One can ask if the model parameters can be uniquely recovered from the full distribution of the data. Here we consider only the possibility of recovering the parameters from the second moments. Note that for the model of (4.4.1) with Gaussian errors these will be equivalent, since in this case the second moments completely determine the distribution. We should also note that identifiability is often defined as a property of the entire parameter space. In this case the parametrization is said to be identifiable only if the mapping between the parameter space and data distribution is one-to-one. With the definition used here, identifiability is a property of individual points within the parameter space. If a point $\boldsymbol{\theta}$ in the parameter space has the property that no other point in the parameter space produces the same data distribution then the model is identifiable at $\boldsymbol{\theta}$. The approach used here is similar to that used by Anderson and Rubin (1956) in their discussion of parameter identification for factor analysis models.

The problem of second moment identifiability reduces to a question about the uniqueness of the solution to the non-linear system of equations (4.4.1). We may get a very rough feel for the restrictions imposed by the requirement of identifiability by counting parameters. The signal covariance depends on $2mp-p^2$ real parameters and (4.4.1) is equivalent to a system of m^2 real equations. Consequently we would expect that $\boldsymbol{\Sigma}_N$ can depend on at most $(m^2-2mp+p^2)$ real parameters.

For a single 3 component MV station, which we fit a plane wave response to, we have p=2, m=3 and $(m^2-2mp+p^2)$ = 1 so $\boldsymbol{\Sigma}_N$ (a 3×3 matrix) can depend on only a single parameter (the scale of the total misfit of the plane wave model). On the other hand, for a five station MV array with p=2, m=15 and

$(m^2-2mp+p^2)$ = 169. Even for small arrays, this constraint is very weak. Note however that its satisfaction does not guarantee identifiability; in general further conditions will be required.

The problem of parameter identifiability has been studied extensively for the case of a diagonal error covariance $\boldsymbol{\Sigma}_N = \mathbf{diag}(\sigma_1^2, \cdots, \sigma_m^2)$. A classic result in this case (e.g. Kendall and Stewart, 1979) is that when m=2, p=1 and $\boldsymbol{\Sigma}_N = \mathbf{diag}(\sigma_1^2, \sigma_2^2)$ the parameters are *not* identified. (In fact, this can easily be verified by the simple parameter counting approach.) A large number of identifiability results for the case of a diagonal $\boldsymbol{\Sigma}_N$ are given in Anderson and Rubin (1956). A generalization of one of these will be useful to us.

For geomagnetic array data it may be necessary to allow for correlated local noise. Cultural noise and some forms of instrument noise (section 2.4) may be coherent between channels at a fixed station, but not coherent with noise at other stations. A straightforward generalization of the diagonal covariance matrix model which allows for this is the block diagonal covariance model

$$\boldsymbol{\Sigma}_N = \mathbf{diag}(\boldsymbol{\Sigma}_i) \tag{4.4.2}$$

where $\boldsymbol{\Sigma}_i$ is the (k×k for k component data) local noise covariance for the i^{th} station. Following theorem 5.1 in Anderson and Rubin we have the following sufficient condition for parameter identifiability for this error model.

Theorem 1: Suppose $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_S(\boldsymbol{\tau}) + \boldsymbol{\Sigma}_N(\boldsymbol{\theta})$ where $\boldsymbol{\tau} = (\mathbf{T}, \boldsymbol{\Sigma}_\alpha)$ as in (4.4.1) and where $\boldsymbol{\Sigma}_N$ has the block diagonal form of (4.4.2) with k×k matrices on the diagonal. Suppose that $\boldsymbol{\Sigma}_\alpha$ is non-singular (of rank p) and that whenever the k rows of

$$\mathbf{U} = \begin{bmatrix} \mathbf{I}_p \\ \mathbf{T} \end{bmatrix}$$

corresponding to any block on the diagonal are deleted from **U**, the remaining rows of **U** can be rearranged to form two disjoint matrices of rank p.

Then the parameters of Σ_S and Σ_N are uniquely determined.

The proof of this is a straightforward extension of the proof give in Anderson and Rubin (whose proof, in turn, follows a result give by Albert (1944)). Since there are several significant differences with our setup (block diagonal error covariances, complex matrices, different notation, and a different parametrization of the signal matrix Σ_S) we offer a proof of the theorem in Appendix C.

As an application of this theorem, consider the case were we assume plane wave sources $p = 2$. Partition $\mathbf{U}$ by stations $\mathbf{U}^T = (\mathbf{U}_1^T \mid \cdots \mid \mathbf{U}_n^T)$ where n is the number of stations. The two columns of $\mathbf{U}_i$, $i = 1, n$ are the fields at station i which are coherent with two orthogonal polarizations of unit magnitude at station 1. We thus expect the two column vectors for the i^{th} station $\mathbf{U}_{i1}$, $\mathbf{U}_{i2}$ to be roughly orthogonal (and certainly linearly independent) so that the matrices $\mathbf{U}_i$ should all be of rank 2. The theorem implies then, that whenever the number of stations is at least 3 the block diagonal covariance parametrization should be identified. Thus even for small arrays, arbitrary local noise covariances can be assumed. This is a significant advantage to multiple station data analysis - no assumptions about local noise are required.

If the total noise is small compared to the signal it is possible to establish more general parameter identification results which allow for models with noise which is coherent between stations. We now state a general theorem of this nature. Our theorem, which we prove in Appendix C, will justify the identifiability of the parameters for a model for the noise covariance which we describe in the next section. We will apply this model to MT and MV data in sections 4.8 - 4.10.

Theorem 2: Suppose $\Sigma = \Sigma_S + \Sigma_N(\boldsymbol{\theta})$ where Σ_S is fixed and of rank p with $||\Sigma_S|| = 1$. Let $\Sigma_S = \sum_{i=1}^{p} \gamma_i \mathbf{v}_i \mathbf{v}_i^*$, let $\mathbf{P}_R = \sum_{i=1}^{p} \mathbf{v}_i \mathbf{v}_i^*$ and $\mathbf{Q}_R = \mathbf{I}_m - \mathbf{P}_R$ be the projections on to the response space $\boldsymbol{R}$ and its orthogonal complement.

Suppose that there are δ, M such that whenever

$$|| \mathbf{Q}_R[\Sigma_N(\boldsymbol{\theta}_1) - \Sigma_N(\boldsymbol{\theta}_2)]\mathbf{Q}_R || < \delta$$

then

$$|| \Sigma_N(\boldsymbol{\theta}_1) - \Sigma_N(\boldsymbol{\theta}_2) || < M || \mathbf{Q}_R[\Sigma_N(\boldsymbol{\theta}_1) - \Sigma_N(\boldsymbol{\theta}_2)]\mathbf{Q}_R || \qquad (4.4.3)$$

Then there is $\varepsilon > 0$ such that when $|| \Sigma_N(\boldsymbol{\theta}) || < \varepsilon$ the parameters are uniquely identified - i.e. if $\boldsymbol{\theta}$ is such that $|| \Sigma_N(\boldsymbol{\theta}) || < \varepsilon$ and $\Sigma_S + \Sigma_N = \Sigma_S' + \Sigma_N(\boldsymbol{\theta}')$ then $\Sigma_S = \Sigma_S'$ and $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$.

To explain this formally stated theorem we give a 'heuristic proof' which suggests an iterative method for recovering the parameters of interest, $\boldsymbol{R}$ and $\boldsymbol{\theta}_1$, from Σ. We will use the approach described below as the basis for our parameter estimation scheme which we describe in section 4.6. Suppose we knew $\mathbf{P}_R$, the projection onto $\boldsymbol{R}$ (and hence $\mathbf{Q}_R$). Then we could compute

$$\mathbf{Q}_R \, \Sigma \, \mathbf{Q}_R = \mathbf{Q}_R \, \Sigma_N(\boldsymbol{\theta}) \, \mathbf{Q}_R \qquad (4.4.4)$$

By assumption there is $\boldsymbol{\theta}$ which satisfies (4.4.1); the condition (4.4.3) for theorem 2 guarantees that there is only one such $\boldsymbol{\theta}$. If we knew $\mathbf{P}_R$ we could, at least in theory, find $\boldsymbol{\theta}$. We are not given $\mathbf{P}_R$, but since Σ_N is small we can get a good approximation $\mathbf{P}_0$ from the p dominant eigenvectors of Σ. By (4.4.2) this will be approximately of the form $\mathbf{P}_R + \varepsilon \mathbf{A}_0$, where $||\mathbf{A}_0||$ is of order one. Then $\mathbf{Q}_0 = \mathbf{I}_m - \mathbf{P}_0 = \mathbf{Q}_R - \varepsilon \mathbf{A}_0$ and, since $\mathbf{Q}_R \Sigma_S = \Sigma_S \mathbf{Q}_R = 0$, we have

$$\mathbf{Q}_0 \Sigma \mathbf{Q}_0 = \mathbf{Q}_0 \Sigma_N(\boldsymbol{\theta}) \mathbf{Q}_0 + \varepsilon^2 \mathbf{A}_0 \Sigma_S \mathbf{A}_0 \qquad (4.4.5)$$

Since $||\Sigma_N|| = \varepsilon$ and $||\Sigma_S|| = 1$, the second term is small compared to the first. Thus

$$\mathbf{J}_0(\boldsymbol{\theta}) = \mathbf{Q}_0\, \boldsymbol{\Sigma}\, \mathbf{Q}_0 - \mathbf{Q}_0\, \boldsymbol{\Sigma}_N(\boldsymbol{\theta})\, \mathbf{Q}_0 \;\simeq\; 0$$

so we may approximate $\boldsymbol{\theta}$ by choosing $\boldsymbol{\theta}_0$ to minimize $\| \mathbf{J}_0(\boldsymbol{\theta}) \|$. The condition (4.4.3) guarantees that $\boldsymbol{\Sigma}_N(\boldsymbol{\theta}_0)$ will be a good approximation to $\boldsymbol{\Sigma}_N(\boldsymbol{\theta})$ provided $\| \mathbf{J}_0(\boldsymbol{\theta}_0) \|$ is small enough.

If we now use the approximation to the noise covariance $\boldsymbol{\Sigma}_N(\boldsymbol{\theta}_0)$ to transform $\boldsymbol{\Sigma}$ to

$$\boldsymbol{\Sigma}' = \boldsymbol{\Sigma}_N^{-1/2}(\boldsymbol{\theta}_0)\, \boldsymbol{\Sigma}\, \boldsymbol{\Sigma}_N^{-1/2\,*}(\boldsymbol{\theta}_0)$$

we can get a better approximation $\mathbf{P}_1$ of the projection $\mathbf{P}_R$. This will have the form

$$\mathbf{P}_1 = \mathbf{P}_R + \varepsilon^2 \mathbf{A}_1 \qquad \text{where } \|\mathbf{A}_1\| \simeq 1$$

We may then proceed as above to find an improved estimate $\boldsymbol{\theta}_1$ of $\boldsymbol{\theta}$. This procedure can be iterated to convergence, so the parameters $\boldsymbol{\theta}$ and $\boldsymbol{R}$ can be recovered from $\boldsymbol{\Sigma}$.

4.5: Estimation of the Noise Covariance II: Parametrization

We now turn our attention to the problem of specifying a parametric form for the noise covariance matrix $\boldsymbol{\Sigma}_N$. In general it is useful to write the noise vector $\mathbf{e}_i$ as a sum of two uncorrelated vectors

$$\mathbf{e}_i = \boldsymbol{\xi}_i + \boldsymbol{\eta}_i \qquad \text{with} \quad E(\boldsymbol{\xi}_i \boldsymbol{\eta}_i^*) = 0$$

where $\boldsymbol{\xi}_i$ represents local noise which is incoherent between stations and $\boldsymbol{\eta}_i$ represents non-local noise which occurs coherently at more than one station in the array.

The local noise vector $\boldsymbol{\xi}_i$ will include instrument noise, local cultural noise (whose length scale is small compared to station separation), the effects of wind

and ground vibrations, etc. Letting $\boldsymbol{\xi}_{ji}$ be the noise vector at station j for time segment i, the assumption that this noise is local implies $E(\boldsymbol{\xi}_{ji}\boldsymbol{\xi}_{j'i}) = \delta_{jj'}\boldsymbol{\Sigma}_j$ where $\boldsymbol{\Sigma}_j$ gives the error covariance for the local noise at station j. Our application of theorem 1 of section 4.4 indicates that even for small arrays we can assume arbitrary local noise covariances $\boldsymbol{\Sigma}_j$. We will thus in general assume a block diagonal form (with no other restrictions) for the covariance of the local noise vectors

$$E(\boldsymbol{\xi}_i \boldsymbol{\xi}_i) = \mathbf{diag}(\, \boldsymbol{\Sigma}_1, \; \cdots, \boldsymbol{\Sigma}_n \,) \tag{4.5.1}$$

This model can be simplified if warranted.

The specification of the covariance structure for the non-local errors is much more problematic - there is no general model (whose parameters can be uniquely identified) which is appropriate to all situations. Indeed, modeling the structure of the non-local noise requires some additional information (or assumptions) and must be treated on a case by case basis. We will consider here only the case of fitting a plane wave source model ($p = 2$) to data from a small regional array. We thus assume the response space is of dimension two and treat the gradient terms (and their effect on the vertical magnetic and electric field components) as noise. If we assume that all coherent noise is related to such source effects, we can then use the results from the exploratory data analysis of Chapter 2, and our random source model from Chapter 3 as guidance in formulating a model for the covariance of the coherent noise.

Although we cannot be sure *a priori* that such a model is correct, we can test the fit of the model to the data. If we can obtain a reasonable fit to the data for a range of frequencies and for a number of arrays we can be confident that the model is general enough to describe the true situation adequately. If this is not the case, then we must explore physically reasonable extensions to the model which do fit the data. The model proposed here is thus a starting point - in a sense a hypothesis to be tested - and is subject to modification as we learn what

the data has to tell us.

One could question the treatment of of the horizontal source gradients (and their effects on the vertical magnetic and electric fields) as noise. These gradients clearly represent a signal which contains useful geophysical information. Indeed, in order to adequately model the source-related noise, we will have to treat the gradients as signal to estimate the parameters which describe $\mathbf{\Sigma}_N$. We would further argue that the utility of small geomagnetic arrays will be greatly enhanced when both plane wave and source gradient responses can be reliably estimated. There are, nonetheless, good reasons to treat source gradient effects as noise for the data analyzed in this thesis.

The fact is, the plane wave signal is by far the dominant signal. While power due to gradient effects can be quite large for vertical magnetic fields, the power in the horizontal gradients is often only barely above background (local) noise levels for small arrays. Furthermore, the arrays we consider here were not designed to measure gradients - they often are linear, making a meaningful estimation of gradient responses impossible. As a result we will see that we can generally obtain only qualitative information about gradient effects. On the other hand we can usually obtain high quality quantitative results for the plane wave responses. It is thus reasonable to take the perspective that the plane wave response space is the parameter of primary interest. Finally, with a random source model, fluctuations in the source terms will perturb the dominant eigenvectors, and hence our estimate of the plane wave response space. To minimize these effects and to obtain meaningful estimation errors for the plane wave response space we must include these effects in the noise.

Before proceeding we set some notation. We assume that $l = 3$ or 5 components of the fields are measured at n stations (so that $m = 3n$ or $5n$) and we assume that the components of the data vectors $\mathbf{X}_k$ are ordered so that the first $2n$

components are horizontal magnetic fields

$$\mathbf{X}_k = \begin{bmatrix} \mathbf{X}_{hk} \\ \mathbf{X}_{zk} \end{bmatrix}$$

where $\mathbf{X}_{hk}$ is the $2n$-vector of horizontal field components and $\mathbf{X}_{zk}$ is the n- or $3n$-vector consisting of vertical or vertical plus electric fields for each of the n stations. We will use similar notation for all vectors in the remainder of this section and throughout section 4.9. We will also assume that the station coordinates are $\mathbf{x}_i = (x_i, y_i)$ and that the origin of the coordinate system is chosen so that the average coordinate $\bar{\mathbf{x}} = 0$.

Consider initially only the horizontal magnetic fields. Let

$$\mathbf{\Psi}_{h1} = \begin{bmatrix} x_1 \\ y_1 \\ - \\ \cdot \\ \cdot \\ \cdot \\ x_n \\ y_n \end{bmatrix} \qquad \mathbf{\Psi}_{h2} = \begin{bmatrix} x_1 \\ -y_1 \\ - \\ \cdot \\ \cdot \\ \cdot \\ x_n \\ -y_n \end{bmatrix} \qquad \mathbf{\Psi}_{h3} = \begin{bmatrix} y_1 \\ x_1 \\ - \\ \cdot \\ \cdot \\ \cdot \\ y_n \\ x_n \end{bmatrix} \tag{4.5.2}$$

so that $\mathbf{\Psi}_{h1}, \mathbf{\Psi}_{h2}, \mathbf{\Psi}_{h3}$ are just the gradient functions $\mathbf{g}_3, \mathbf{g}_4, \mathbf{g}_5$, respectively, evaluated at the station coordinates. With our assumption that all coherent noise is source related, the results of Chapter 3 imply that we can approximate the coherent noise vector for the horizontal fields as

$$\mathbf{\eta}_{hk} = \mathbf{\Psi}_h \mathbf{\beta}_k \tag{4.5.3}$$

where $\mathbf{\Psi}_h = (\mathbf{\Psi}_1 \; \mathbf{\Psi}_2 \; \mathbf{\Psi}_3)$ and $\mathbf{\beta}_k$ is the three vector of coefficients of the gradient terms for the k^{th} data segment. This will only be strictly correct for a 1-d earth. Just as lateral variations in conductivity perturb the uniform fields of the plane wave source, they will perturb the total fields associated with gradients. We will assume (4.5.3) initially in the hope that the effect will be small enough to neglect, at least for a first try. We may thus write the covariance matrix for the noise in

the horizontal field components as

$$\Sigma_{hN} = E(\mathbf{e}_{hk}\mathbf{e}_{hk}^*) = \Psi_h \Sigma_\beta \Psi_h^* + \mathbf{diag}(\Sigma_{h1}, \cdots, \Sigma_{hn}) \qquad (4.5.4)$$

where $\Sigma_\beta = E(\beta\beta^*)$ and Σ_{hi} is the local noise covariance for the horizontal components at station i. The noise covariance of the horizontal field components thus depends upon the nine real parameters which determine the unique real and imaginary parts of the 3×3 Hermetian matrix Σ_β and the $4n$ real parameters which determine the 2×2 Hermetian matrices Σ_{hi}, $i = 1, n$. Note that Σ_{hN} is linear in these parameters. As a result parameter estimation can be accomplished by relatively simple means. We shall discuss this in more detail in section 4.6.

If vertical fields are considered as well, the situation becomes more complicated. For a 1-d earth we can use the results of Chapter 3 to write the three gradient vectors as

$$\psi_i = \begin{bmatrix} \psi_{hi} \\ \psi_{zi} \end{bmatrix} \quad i = 1, 3 \quad \text{where} \quad \psi_{z_1} = 2C(\omega)\begin{bmatrix} 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix} \quad \psi_{z_2} = \psi_{z_3} = 0 \qquad (4.5.5)$$

Letting Ψ be the matrix whose i^{th} column is ψ_i we can then write the noise covariance for the three component data as

$$\Sigma_N = E(\mathbf{e}_k\mathbf{e}_k^*) = \Psi\Sigma_\beta\Psi^* + \mathbf{diag}(\Sigma_1, \cdots, \Sigma_n) \qquad (4.5.6)$$

The only significant difference now is that the noise covariance depends in addition on $C(\omega)$, (Schmuker's inductive length scale), and the dependence of Σ_N on this parameter is not linear (some elements of Σ_N involve products of $C(\omega)$ and elements of Σ_β).

The form for the noise covariance for the case where electric fields are also included can also be derived for a 1-d earth using the results from Chapter 3. In fact, as we have discussed in section 3.7, in 2- and 3-dimensional environments, the results from the simple 1-d model are not likely to be particularly relevant for

the vertical magnetic and (especially) the electric fields. For exploratory purposes we suggest that it may be more reasonable to allow all of the elements of the vectors ψ_{zi}, $i = 1, 3$ to be free parameters which must be estimated. That is, we assume that the gradients are perturbed only slightly from what they would be for the 1-d case and then attempt to estimate the form of the vertical magnetic and electric fields which are associated with each gradient term. We will discuss this more fully in sections 4.9 and 4.10. For now we note that if we can estimate the vectors ψ_{zi}, $i = 1, 3$, (or any parameters such as $C(\omega)$ which determine them) by some means, the model for the noise covariance has the same form as for the case where only horizontal fields are considered. For the remainder of this section we consider this simpler case only.

Often the array is too small (relative to source length scales) in one or more directions so that one or more of the three gradient terms will not be measurable. As we have pointed out in Chapter 3, there will still be vertical magnetic and (possibly) electric fields associated with this unobservable gradient. In particular, for small MV arrays where the actual gradients may often be too small to be measured reliably, the vertical fields associated with these gradients can still be quite large. We have seen this both for synthetic arrays (figure 3.11) and for real arrays (figures 2.8 and 2.9). In this case the model for the 'gradient noise' of (4.5.3) has to be modified. In the simplest case, where the array is so small that no gradients can be measured, the vertical fields associated with the gradients will be approximately constant at all stations in the array -i.e. the fields associated with this vertical field 'noise' will be proportional to the vector $\mathbf{u}_z$ given in (2.5.2). We may thus model the noise covariance matrix in this case as

$$\Sigma_N = \beta_z^2\mathbf{u}_z\mathbf{u}_z^* + \mathbf{diag}(\Sigma_1, \cdots, \Sigma_n) \qquad (4.5.7)$$

which is of the same general form as (4.5.6). More generally, when one or two (but not three) linearly independent linear combinations of the gradient vectors ψ_{hi}

can be measured, we may consider a combination of (4.5.6) and (4.5.7). This situation will arise, for instance, for a linear array in which the gradients of the two horizontal field components in the direction of the array strike can be measured. These two gradients can be represented as two linear combinations of the vectors ψ_{hi}. A third linear combination, consisting of gradients perpendicular to the array strike is not measured.

The models for the noise covariance matrix discussed above are linear in the unknown parameters. In general we can write these models in the form

$$\Sigma_N = \sum_{i=1}^{r} \theta_i \mathbf{Y}_i \tag{4.5.8}$$

where the $\mathbf{Y}_i$, $i = 1, r$ are known $m \times m$ matrices and the θ_i, $i = 1, r$ are real parameters. We give the explicit relationship between the parametrizations of (4.5.6) and (4.5.8) in Appendix D. Parameter estimation for this sort of linear model for real symmetric covariance matrices is treated by Anderson (1969; 1970; 1973). We have independently developed an equivalent estimation scheme for complex Hermetian matrices which we discuss in the next section.

We close this section by stating a theorem, proved in Appendix D, which can be used to establish condition (4.4.3) of theorem 2 for linear covariance models of the form (4.5.8).

Theorem 3: Assume $\Sigma_N(\theta) = \sum_{i=1}^{r} \theta_i \mathbf{Y}_i$ where the $\mathbf{Y}_i$ are known matrices. Let $\mathbf{Q}_R$ be the projection onto the orthogonal complement of the response space R. Then, if the r matrices $\mathbf{Y}_i' = \mathbf{Q}_R \mathbf{Y}_i \mathbf{Q}_R$ are linearly independent, there is M such that for all θ_1, θ_2

$$\| \Sigma_N(\theta_1) - \Sigma_N(\theta_2) \| < M \| \mathbf{Q}_R [\Sigma_N(\theta_1) - \Sigma_N(\theta_2)] \mathbf{Q}_R \|$$

The linear independence of the matrices $\mathbf{Y}_i'$ is easily checked during the estimation procedure described in the next section. We find that for the specific models discussed here this condition always holds. Theorem 2 thus implies that, for these models, the noise and signal parameters can always be uniquely separated provided signal to noise ratios are large enough.

4.6: Estimation of the Noise Covariance III: Estimation

We have already outlined, with our heuristic proof of theorem 2 at the end of section 4.4, our general approach to the simultaneous estimation of the noise covariance Σ_N and response space R. The approach is iterative, and is based on the observation that if we knew Σ_N we could use the results from section 4.2 to obtain an unbiased estimate of R, while if we knew R we could eliminate the signal portion of the SDM and estimate the parameters θ which determine Σ_N. We can thus begin with an estimate of R, use this to obtain an initial estimate of θ, use this to improve our estimate of R etc. When the signal to noise ratio is large, and condition (4.4.3) of theorem 2 holds, the argument given at the end of section 4.4 suggests that this approach ought to work well and converge in a few iterations. In this section we consider the second step in this procedure - estimating the noise covariance parameters θ - for the case discussed in the previous section where Σ_N is linear in the unknown parameters θ. For this discussion then, we assume that we know, or have an estimate of the projection matrix $\mathbf{P}_R$, and that

$$\Sigma_N(\theta) = \sum_{i=1}^{r} \theta_i \mathbf{Y}_i \tag{4.6.1}$$

If $\mathbf{P}_R$ is known exactly and $\mathbf{Q}_R = \mathbf{I}_m - \mathbf{P}_R$ then

$$\mathbf{S}_N' = \mathbf{Q}_R \mathbf{S} \mathbf{Q}_R = \mathbf{Q}_R \mathbf{E} \mathbf{E}^* \mathbf{Q}_R \tag{4.6.2}$$

depends only on the errors $\mathbf{e}_i$ (the columns of $\mathbf{E}$) and not on any of the signal

parameters. We will base our estimate of $\boldsymbol{\theta}$ on $\mathbf{S}_N'$.

Any matrix of the form $\mathbf{Q}_R\mathbf{A}\mathbf{Q}_R$ has rank at most $m-p$. Before proceeding it will be useful to reformulate the problem slightly so that we may use matrices which are generally of full rank. We may write

$$\mathbf{P}_R = \sum_{i=1}^{p} \mathbf{v}_i\mathbf{v}_i^* \qquad\qquad \mathbf{Q}_R = \sum_{i=p+1}^{m} \mathbf{v}_i\mathbf{v}_i^*$$

where the vectors $\mathbf{v}_i$, $i = 1, m$ are an orthonormal basis for C^m. Let

$$\mathbf{V} = (\,\mathbf{V}_1\ \mathbf{V}_2\,) = (\,\mathbf{v}_1\ \cdots\ \mathbf{v}_p\ \mid\ \mathbf{v}_{p+1}\ \cdots\ \mathbf{v}_m\,)$$

Then transforming matrices $\mathbf{A} \rightarrow \mathbf{V}^*\mathbf{A}\mathbf{V}$ corresponds to a rotation of the coordinate system of $\mathbf{C}^m$. A simple calculation shows that for any $m\times m$ matrix $\mathbf{A}$

$$\mathbf{V}^*\mathbf{Q}_R\mathbf{A}\mathbf{Q}_R\mathbf{V} = \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{V}_2^*\mathbf{A}\mathbf{V}_2 \end{bmatrix} \tag{4.6.3}$$

where the matrix $\mathbf{V}_2^*\mathbf{A}\mathbf{V}_2$ is $(m-p)\times(m-p)$ and will be of full rank if $\mathbf{A}$ is. This demonstrates that we may consider the $(m-p)\times(m-p)$ matrices of the form $\mathbf{V}_2^*\mathbf{A}\mathbf{V}_2$ instead of the $m\times m$ matrices $\mathbf{Q}_R\mathbf{A}\mathbf{Q}_R$ without any loss of information. We thus rewrite (4.6.2) as

$$\mathbf{S}_N' = \mathbf{V}_2^*\mathbf{S}\mathbf{V}_2 = \mathbf{V}_2^*\mathbf{E}\mathbf{E}^*\mathbf{V}_2 \tag{4.6.4}$$

$\mathbf{S}_N'$ is the sample covariance matrix of the $(m-p)$ dimensional error vectors $\mathbf{V}_2^*\mathbf{e}_i$. Its expectation is

$$E(\mathbf{S}_N') = \mathbf{V}_2^*\boldsymbol{\Sigma}_N(\boldsymbol{\theta})\mathbf{V}_2 = \sum_{i=1}^{r} \theta_i\mathbf{V}_2^*\mathbf{Y}_i\mathbf{V}_2 = \sum_{i=1}^{r} \theta_i\mathbf{Y}_i' \tag{4.6.5}$$

The expectation of the 'projected' SDM $\mathbf{S}_N'$ (now of dimension $m-p \times m-p$) thus is a linear combination of the projected matrices $\mathbf{Y}_i'$. Note that the parameters are not effected (i.e. the $\theta_i's$ are the same in (4.6.1) and (4.6.5)). The problem of estimating the noise covariance parameters $\boldsymbol{\theta}$ from the full sample covariance

matrix $\mathbf{S}$ with $\boldsymbol{R}$ known can thus be reduced to that of estimating the same parameters for a virtually identical linear model from the smaller SDM $\mathbf{S}_N'$.

We thus consider the general problem of estimation of $\boldsymbol{\theta}$ for complex covariance matrices with the linear structure of (4.6.1). To simplify notation for this discussion we let $M = m-p$, $\mathbf{C} = \mathbf{S}_N'$ with $E\mathbf{C} = \boldsymbol{\Sigma}_C$, we drop the primes on the matrices $\mathbf{Y}_i$, and we consider the problem in general with the understanding that for our application we will use the projected matrices.

One obvious approach to estimation of θ for the linear model (4.6.1) would be to minimize

$$\| \mathbf{C} - \sum_{i=1}^{r} \theta_i\mathbf{Y}_i \|^2 = \sum_{k,l=1}^{M} (C_{kl} - \sum_{i=1}^{r} \theta_i Y_{ikl})^2$$

Note that this is a standard least squares (LS) problem which can be easily solved. The estimate obtained from solving this LS problem is unbiased, and with sufficient regularity for the error distribution will be consistent.

This procedure can be refined. Let

$$\mathbf{C}' = \boldsymbol{\Sigma}_C^{-½}\,\mathbf{C}\,\boldsymbol{\Sigma}_C^{-½} \qquad\qquad \mathbf{Y}_i' = \boldsymbol{\Sigma}_C^{-½}\,\mathbf{Y}_i\,\boldsymbol{\Sigma}_C^{-½},\ i = 1, r \tag{4.6.6}$$

Then

$$E\mathbf{C}' = \mathbf{I} = \sum_{i=1}^{r} \theta_i\mathbf{Y}_i' \tag{4.6.7}$$

We can form the M^2 dimensional real vectors from the elements on and below the diagonals of the matrices $\mathbf{C}'$, $\mathbf{Y}_i'$, $i = 1, r$

$$\mathbf{c} = \begin{bmatrix} 1/\sqrt{2}C'_{11} \\ \cdot \\ \cdot \\ \cdot \\ 1/\sqrt{2}C'_{MM} \\ \mathrm{Re}C'_{12} \\ \mathrm{Im}C'_{12} \\ \cdot \\ \cdot \\ \cdot \\ \mathrm{Re}C'_{M,M-1} \\ \mathrm{Im}C'_{M,M-1} \end{bmatrix} \qquad \mathbf{y}_i = \begin{bmatrix} 1/\sqrt{2}Y'_{i11} \\ \cdot \\ \cdot \\ \cdot \\ 1/\sqrt{2}Y'_{iMM} \\ \mathrm{Re}Y'_{i12} \\ \mathrm{Im}Y'_{i12} \\ \cdot \\ \cdot \\ \cdot \\ \mathrm{Re}Y'_{iM,M-1} \\ \mathrm{Im}Y'_{iM,M-1} \end{bmatrix} \qquad i = 1, r$$

Then (4.6.6) implies

$$E\mathbf{c} = \sum_{i=1}^{r} \theta_i \mathbf{y}_i \qquad \text{or} \qquad \mathbf{c} = \sum_{i=1}^{r} \theta_i \mathbf{y}_i + \boldsymbol{\upsilon} \qquad \text{where } E\boldsymbol{\upsilon} = 0 \qquad (4.6.8)$$

For the Gaussian case $\mathbf{C}'$ has a standard complex Wishart distribution and the covariances of the elements of $\mathbf{C}'$ (equivalently, the elements of the vector $\boldsymbol{\upsilon}$) can easily be computed (see Appendix B). We can apply these results to find that the covariance matrix of the real random vector $\boldsymbol{\upsilon}$ is

$$E(\boldsymbol{\upsilon}\boldsymbol{\upsilon}^T) = \tfrac{1}{2}\mathbf{I}_{M^2} \qquad (4.6.9)$$

By the Gauss-Markov theorem (cf. Graybill, 1976) the best linear unbiased estimate of $\boldsymbol{\theta}$ is given by the standard LS estimate

$$\hat{\boldsymbol{\theta}} = (\Omega^T\Omega)^{-1}\Omega^T\mathbf{c} \qquad \text{where} \quad \Omega = (\,\mathbf{y}_1 \;\cdots\; \mathbf{y}_r\,) \qquad (4.6.10)$$

We cannot exactly compute the vectors $\mathbf{c}$, $\mathbf{y}_i$ $i = 1, r$ needed to compute $\hat{\boldsymbol{\theta}}$ in (4.6.10) since these depend on the unknown covariance matrix $\boldsymbol{\Sigma}_C$, but this general argument suggests an iterative procedure. Specifically, we start with an estimate of $\hat{\boldsymbol{\Sigma}}_C^{(0)}$ ($= \mathbf{I}$, say) and use this in place of $\boldsymbol{\Sigma}_C$ in (4.6.6) to compute approximate vectors $\mathbf{c}^{(0)}$, $\mathbf{y}_i^{(0)}$ $i = 1, r$. These are then used in (4.6.10) to compute an estimate

$\hat{\boldsymbol{\theta}}^{(1)}$, which is used to form an improved estimate $\boldsymbol{\Sigma}_C^{(1)}$ of $\boldsymbol{\Sigma}_C$, which is in turn used to compute an improved estimate of the parameters of interest, $\boldsymbol{\theta}^{(2)}$. This procedure may then be iterated to convergence.

An equivalent iterative scheme for estimating the model parameters for (real) covariance matrices with linear structure has been described by Anderson (1969, 1970, 1973). He proves that if the error vectors are Gaussian this procedure is asymptotically equivalent to maximum likelihood. As a result, the procedure is asymptotically efficient (for the Gaussian case). In fact, this holds even if the procedure is stopped after two iterations : the estimate $\hat{\boldsymbol{\theta}}^{(2)}$ is asymptotically efficient.

There are some complications with this procedure in practice. To be a valid covariance matrix for which we can compute the inverse square root $\hat{\boldsymbol{\Sigma}}_C^{-1/2}$ needed to transform the matrices in (4.6.6), the estimated covariance must be positive definite. Using the unconstrained (iteratively weighted) LS estimates of $\boldsymbol{\theta}$ we cannot guarantee that

$$\hat{\boldsymbol{\Sigma}}_C = \sum_{i=1}^{r} \hat{\theta}_i \mathbf{Y}_i$$

satisfies this condition. For our specific parametrization the parameter estimates $\hat{\boldsymbol{\theta}}$ determine estimates to a series of smaller covariance matrices $\hat{\boldsymbol{\Sigma}}_\beta$, $\hat{\boldsymbol{\Sigma}}_i$, $i = 1, r$ (see (4.5.6)). The elements of these matrices are essentially the parameters $\boldsymbol{\theta}$; the connection is made explicit in Appendix C. If the first of these matrices, which represents the non-local (source related) noise is positive semi-definite and the others, which represent local noise, are all positive definite then $\hat{\boldsymbol{\Sigma}}_C$ will be positive definite. It is fairly easy to ensure that the estimates of these smaller matrices satisfy the necessary conditions. To do this we essentially diagonalize each of these estimated matrices and eliminate negative eigenvalues. The local and non-local noise matrices are treated slightly differently.

For the non-local noise we don't know *a priori* what the rank of Σ_β should be since some or all of the gradient terms may be too small to detect. For this matrix we thus set all eigenvalues which are too small (including all negative ones) to zero. Given a unit eigenvector $\mathbf{v}$ of $\hat{\Sigma}_\beta$ the associated eigenvalue is just $\hat{\gamma} = \mathbf{v}^* \hat{\Sigma}_\beta \mathbf{v}$. This is a linear combination of the elements of the estimated covariance, and hence of the parameter estimates $\hat{\boldsymbol{\theta}}$. Ignoring the fact that $\mathbf{v}$ is itself a random quantity, which depends on the estimated parameters $\hat{\boldsymbol{\theta}}$ and using the estimation error covariance for the parameter estimates $\hat{\boldsymbol{\theta}}$ (computed using the standard LS approach), we may compute an approximate estimation error for $\hat{\gamma}$. We thus set to zero all eigenvalues which are less than some constant c times their estimation error. Values of c between 1.0 and 2.0 seem to work fairly well; for the results discussed in section 4.8 we have used a value of $c = 1.5$. Using the remaining positive eigenvalues, and their associated eigenvectors, we then reconstruct our estimated covariance matrix $\hat{\Sigma}_\beta$ which is now guaranteed to be positive semi-definite.

We assume that the true local noise covariance matrices are all positive definite, since with instrument noise that is incoherent between channels this will always be true. To ensure that the estimated local noise matrices are positive definite we diagonalize the matrices $\hat{\Sigma}_i$ and replace all eigenvalues which are too small (including all non-positive eigenvalues) with a minimum value. This minimum value is determined as a constant (we have used 0.2) times the median of all eigenvalues from all of the estimated noise covariance matrices. These procedures are admittedly quite *ad hoc* but they seem to work fairly well.

We remind the reader that the full estimation scheme that we have described here involves a number of steps - we alternately improve estimates of $\boldsymbol{R}$ and $\Sigma_N(\boldsymbol{\theta})$. In this section we have emphasized estimation of $\Sigma_N(\boldsymbol{\theta})$. As outlined at the beginning of this section and in section 4.4, this procedure is combined with

the estimation scheme for $\boldsymbol{R}$ with a general noise covariance matrix described in section 4.2. While the individual steps can be rigorously justified, the total algorithm is rather *ad hoc*. There is an alternative way to put the two estimation steps together which may be preferable to our approach. In order to estimate the noise covariance we have eliminated the signal part of $\mathbf{S}$ by projecting into the orthogonal complement of the estimated response space. Instead of doing this we could have modeled the full covariance matrix

$$\boldsymbol{\Sigma} = E\mathbf{S} = \mathbf{U}\,\boldsymbol{\Sigma}_\alpha\,\mathbf{U}^* + \psi\Sigma_\beta\psi^* + \mathbf{diag}(\,\boldsymbol{\Sigma}_1,\ \cdots\ ,\boldsymbol{\Sigma}_n\,) \qquad (4.6.11)$$

as a covariance matrix with linear structure. We now must have estimates of $\mathbf{U}$ for (4.6.11) to have the linear form discussed in this section. Again, an iterative approach is possible. We can estimate $\mathbf{U}$ using the methods of section 4.2, estimate the parameters of (4.6.11), use these to refine the estimates of $\mathbf{U}$ etc. We have not developed this scheme, but we conjecture that, at least in some circumstances, this approach will be asymptotically equivalent to maximum likelihood for the case where everything (including the α_{jk}'s) are jointly Gaussian (see (Joreskog and Goldberger (1972) and Browne (1974) for suggestive examples in the context of factor analysis models).

The difference between this second scheme and the one which we have developed and applied to real data should not be significant when the signal to noise ratio is large. With our approach we completely ignore all information about the error covariance parameters in the subspace $\boldsymbol{R}$. On the other hand, the weighted LS approach to estimating the noise covariance parameters $\boldsymbol{\theta}$ down-weights information about the parameters in directions which have large variances. Since, with a large signal to noise ratio, the directions which lie in $\boldsymbol{R}$ have very large variances (relative to all other directions), the inclusion of this extra information will have a minimal effect on the noise covariance parameter estimates. The development of this alternative approach, and a comparison to the approach

suggested here would be a worthwhile exercise.

4.7: Testing Model Fit

We have proposed a general statistical model for geomagnetic array data which implies that the expectation of the SDM takes the general form

$$E\mathbf{S} = \mathbf{\Sigma} = \mathbf{U}\mathbf{\Sigma}_\alpha\mathbf{U}^* + \mathbf{\Sigma}_N(\boldsymbol{\theta}) \qquad \text{where } \mathbf{U}^T = (\,\mathbf{I}_p\ \mathbf{T}^T\,) \tag{4.7.1}$$

and we have suggested approaches to the estimation of the parameters $\mathbf{T}$, $\mathbf{\Sigma}_\alpha$, and $\boldsymbol{\theta}$. We now consider the problem of testing the adequacy or 'goodness of fit' of the proposed model (and of simpler possible models). If the model is adequate to describe the true situation, the matrix $\hat{\mathbf{\Sigma}}$ formed from substituting the parameter estimates in (4.7.1) should be very similar to the full sample SDM $\mathbf{S}$ when N gets large. If these two matrices differ by 'to much', then a more general model will be required to fit the data. What we need is a statistic which allows us to quantify how much of a difference is 'to much'.

A general approach to this problem is possible if maximum likelihood estimation is used. We outline this approach, which is discussed in many texts on mathematical statistics (e.g. Bickel and Doksum, 1977), and then discuss its application to our situation. The general set up is as follows. We assume that the distribution of our data $\mathbf{x}$ is a member of a family of models which depend on a finite dimensional parameter $\boldsymbol{\theta}$. We write for the density function $p(\mathbf{x};\boldsymbol{\theta})$. $\Theta \subseteq \mathbf{R}^M$ is the full parameter space (of dimension M), and $\Theta_0 \subseteq \Theta$ consists of a subset of these parameters of dimension $P < M$. We want to test the null hypothesis that the unknown parameter vector $\boldsymbol{\theta}$ is an element of the lower dimensional parameter space Θ_0 against the alternative that it is not. If we cannot reject the null hypothesis, then we can say that the data can be adequately described by the simpler model; there is no need to invoke the more complicated full parametrization to model the data.

Let $\hat{\boldsymbol{\theta}}_0$ be the parameter maximizing the likelihood $p(\mathbf{x};\boldsymbol{\theta})$ over the restricted parameter space Θ_0 and let $\hat{\boldsymbol{\theta}}$ be the maximum likelihood estimate (MLE) over the full parameter space Θ. The generalized log-likelihood ratio statistic is defined as

$$\boldsymbol{\lambda} = -2\ln\left[\frac{p(\mathbf{x};\hat{\boldsymbol{\theta}}_0)}{p(\mathbf{x};\hat{\boldsymbol{\theta}})}\right] \tag{4.7.2}$$

Under the null hypothesis ($\boldsymbol{\theta} \in \Theta_0$), and with sufficient regularity, (e.g. Lehman, 1983) the asymptotic distribution of the statistic $\boldsymbol{\lambda}$ is chi-square with M–P degrees of freedom.

To apply this result to our problem we must completely specify the distribution of the data. For this purpose, then, we assume that the vectors $\mathbf{X}_k$ (the columns of the data matrix Ξ) are N independent, identically distributed complex m–variate Gaussian random vectors with covariance matrix $\mathbf{\Sigma}$. Then the joint density of the data can be written (Giri, 1977)

$$p(\Xi;\mathbf{\Sigma}) = \pi^{-Nm}|\mathbf{\Sigma}|^{-1}\exp(\,-N\mathbf{Tr}[\mathbf{\Sigma}^{-1}\,\mathbf{S}]\,) \qquad \text{where } \mathbf{S} = N^{-1}\Xi\Xi^* \tag{4.7.3}$$

For our application of the likelihood ratio test the full parameter space Θ is the set of all $m\times m$ positive definite Hermetian matrices (which can be parametrized by a set of m^2 real parameters). and the smaller space Θ_0 is the subset of matrices in Θ of the form (4.7.1).

The MLE for $\mathbf{\Sigma}$ for the full parameter space is (Goodman, 1962) $\mathbf{S}$. If we let $\hat{\mathbf{\Sigma}}_0$ be the MLE over the restricted parameter space we can then write the likelihood ratio statistic, after a bit of simplification, as

$$\boldsymbol{\lambda} = 2N(\,\ln|\hat{\mathbf{\Sigma}}_0| + \mathbf{Tr}[\hat{\mathbf{\Sigma}}_0^{-1}\mathbf{S}] - \ln|\mathbf{S}| - m\,) \tag{4.7.4}$$

For our purposes this can be simplified further. With Θ_0 defined as the set of positive definite Hermetian matrices satisfying (4.7.1), where $\mathbf{\Sigma}_N(\boldsymbol{\theta})$ has the linear form of (4.6.1), we see that if $\mathbf{\Sigma}_0 \in \Theta_0$ so is $c\mathbf{\Sigma}_0$ where c is any positive constant. As a result, we can reparametrize the model so that $c = \|\mathbf{\Sigma}_0\|$ is one of the model

parameters. Let $\Sigma'_0 = \Sigma_0/c$, where Σ'_0 depends on the parameters θ'. Then we can write the log-likelihood as

$$L(c, \Sigma_0';\Xi) = \ln[p(c, \Sigma_0';\Xi)] = Nm\ln\pi - Nm\ln c - N\ln|\Sigma_0'| - Nc^{-1}\mathbf{Tr}[\Sigma_0'^{-1}\mathbf{S}]$$

At the maximum of the likelihood the derivatives of L with respect to all parameters must vanish. Setting the derivative of L with respect to c to zero and simplifying we find that the MLE's $\hat{c}$ and $\hat{\Sigma}_0'$ must satisfy

$$m = \mathbf{Tr}[(\hat{c}\hat{\Sigma}_0')^{-1}\mathbf{S}] = \mathbf{Tr}[(\hat{\Sigma}_0)^{-1}\mathbf{S}]$$

since $\hat{c}\hat{\Sigma}'_0 = \hat{\Sigma}_0$ where $\hat{\Sigma}_0$ is the MLE expressed in the original parametrization. Thus the likelihood ratio statistic simplifies to

$$\lambda = 2N(\ln|\hat{\Sigma}_0| - \ln|\mathbf{S}|) \tag{4.7.5}$$

The simplest application of this test is to the case where Θ_0 is the one dimensional parameter space consisting of all isotropic covariance matrices $\Sigma_0 = \sigma^2\mathbf{I}_m$. The MLE of Σ_0 is then

$$\hat{\sigma}^2\mathbf{I}_m \qquad \text{where } \hat{\sigma}^2 = m^{-1}\mathbf{Tr}[\mathbf{S}] = m^{-1}\sum_{i=1}^{m}\gamma_i = \bar{\gamma}$$

where γ_i, $i = 1, m$ are the eigenvalues of $\mathbf{S}$. The determinant of $\mathbf{S}$ is the product of these same eigenvalues so after some simplification we may write the likelihood ratio statistic for this test as

$$\lambda_0 = 2N\sum_{i=1}^{m}\ln(\gamma_i/\bar{\gamma})$$

Under the null hypothesis λ is distributed as a χ^2 random variable with $l = m^2-1$ degrees of freedom. This statistic would be used to test the null hypothesis that all structure in the spectral density matrix could have arisen from isotropic noise - i.e. that there is nothing of interest in the data. If λ exceeds the appropriate (say 95%) critical point for the chi-square distribution with the appropriate number of

degrees of freedom, we would reject the null hypothesis. In fact, for all SDM's we have looked at, it is not necessary to do a formal hypothesis test to be convinced that the data cannot be modeled as isotropic noise.

A more useful null hypothesis is that the data can be modeled as signal (from a p dimensional response space) plus isotropic noise - i.e. (4.7.1) holds with $\Sigma_N = \sigma^2\mathbf{I}_m$. Now the null hypothesis parameter space Θ_0 is of dimension $2mp-p^2+1$ and is equivalent to the space of all complex Hermetian $m\times m$ matrices whose eigenvalues satisfy $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{p+1} = \cdots = \lambda_m$. The MLE $\hat{\Sigma}_0$ can be written in terms of the eigenvalues γ_i and eigenvectors $\mathbf{v}_i$ of $\mathbf{S}$ (Morrison, 1967)

$$\hat{\Sigma}_0 = \sum_{i=1}^{p}\gamma_i\mathbf{v}\mathbf{v}^* + \sum_{i=p+1}^{m}\bar{\gamma}\mathbf{v}\mathbf{v}^* \qquad \text{where now } \bar{\gamma} = (m-p)^{-1}\sum_{i=p+1}^{m}\gamma_i$$

We thus have

$$\ln|\mathbf{S}| = \sum_{i=1}^{m}\ln\gamma_i \qquad \ln|\Sigma_0| = \sum_{i=1}^{p}\ln\gamma_i + (m-p)\ln\bar{\gamma}$$

so that the test statistic is now

$$\lambda_p = 2N\sum_{i=p+1}^{m}\ln(\gamma_i/\bar{\gamma}) \tag{4.7.6}$$

In this case λ will be asymptotically distributed χ^2_l where $l = m^2-2mp+p^2-1$. Note that λ_0 can be thought of as a special case of λ_p. These tests were first proposed (for the case of real Gaussian random vectors) by Mauchly (1946). Refinements of the likelihood ratio statistic (to improve its performance for small samples) have been considered for the real versions of the cases discussed above by Bartlett (1954), Lawely (1956) and James (1969). These corrections are fairly small and will not be considered here.

Finally, we consider testing the goodness of fit of the model (4.7.1) when Σ_N has the linear structure of (4.6.1). Although the estimation scheme which we have

proposed in sections (4.2) - (4.6) is not exactly maximum likelihood we can still use our estimates to form $\hat{\mathbf{\Sigma}}_0$, substitute this into (4.7.4), and compute the statistic $\boldsymbol{\lambda}$. While the asymptotic distribution of this statistic will generally be different from that which would be obtained if the true MLE were used, this is still a useful index of model fit. We believe that our estimates do not differ greatly from the MLE's, and we have conjectured that with a minor modification of the scheme (which should have a very small effect on parameter estimates) the estimates will be asymptotically equivalent to maximum likelihood. We are furthermore of the opinion that even if the MLE of $\mathbf{\Sigma}$ is used in (4.7.4) the exact asymptotic distribution of $\boldsymbol{\lambda}$ should not be taken too seriously since the distribution of the data vectors is almost certainly not Gaussian. Finally note that using the true MLE to compute $\boldsymbol{\lambda}$ will always make the statistic smaller. Thus if we conclude that the model fit is adequate with our parameter estimates, we will conclude this more strongly with the MLE's, but if we conclude that the model fit is inadequate it is possible that we may reverse our decision with the true MLE's.

We close this section by giving a slightly more convenient form for the computation of $\boldsymbol{\lambda}$ for this last case. First, let $\mathbf{A}$ be any non-singular $m \times m$ matrix. Then for the determinant of the product of two $m \times m$ matrices $\mathbf{A}$ and $\mathbf{B}$ we have $|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}|$, so that $\ln|\mathbf{AS}| - \ln|\mathbf{A}\mathbf{\Sigma}_0| = \ln|\mathbf{S}| - \ln|\mathbf{\Sigma}_0|$. A similar result holds for post-multiplication of the covariance matrices by $\mathbf{A}$. Let $\mathbf{\Sigma}_N(\hat{\boldsymbol{\theta}})$ be the estimated noise covariance. The signal part of the estimated covariance matrix $\hat{\mathbf{\Sigma}}_0$ is calculated from the p dominant eigenvectors and eigenvalues of the transformed sample SDM $\mathbf{S}'$. Thus if

$$\mathbf{S}' = \mathbf{\Sigma}_N(\hat{\boldsymbol{\theta}})^{-½}\,\mathbf{S}\,\mathbf{\Sigma}_N(\hat{\boldsymbol{\theta}})^{-½*} = \sum_{i=1}^{p} \gamma_i' \mathbf{v}_i' \mathbf{v}_i'^{*} + \sum_{i=p+1}^{m} \gamma_i' \mathbf{v}_i' \mathbf{v}_i'^{*}$$

we will have for the estimated covariance matrix for the model

$$\hat{\mathbf{\Sigma}}_0' = \mathbf{\Sigma}_N(\hat{\boldsymbol{\theta}})^{-½}\,\hat{\mathbf{\Sigma}}_0\,\mathbf{\Sigma}_N(\hat{\boldsymbol{\theta}})^{-½*} = \mathbf{\Sigma}_N(\hat{\boldsymbol{\theta}})^{-½}\,[\hat{\mathbf{\Sigma}}_S + \mathbf{\Sigma}_N(\hat{\boldsymbol{\theta}})]\,\mathbf{\Sigma}_N(\hat{\boldsymbol{\theta}})^{-½*}$$

$$= \sum_{i=1}^{p} \gamma_i' \mathbf{v}_i' \mathbf{v}_i'^{*} + \sum_{i=p+1}^{m} \mathbf{v}_i' \mathbf{v}_i'^{*}$$

Here γ_i' and $\mathbf{v}_i'$ are the eigenvalues and eigenvectors of the transformed SDM $\mathbf{S}'$. Then we have,

$$\boldsymbol{\lambda} = 2N(\ln|\mathbf{S}| - \ln|\hat{\mathbf{\Sigma}}_0|) = 2N(\ln|\mathbf{S}'| - \ln|\hat{\mathbf{\Sigma}}_0'|) = 2N \sum_{i=p+1}^{m} \ln(\gamma_i')$$

so that the test statistic can be easily computed in terms of the m–p smallest eigenvalues of $\mathbf{S}'$. Note that these are found as a byproduct of the estimation scheme.

4.8: Examples of Noise Covariance Estimation

In this section we consider some example applications of the noise covariance estimation methods described in sections 4.4 - 4.7. We apply these methods to four small geomagnetic arrays - the five station EMSLAB long period MT array (figure 2.1) and three MV arrays from Western Washington (figure 4.1). Two of the MV arrays (numbered 1 and 8; solid dots in figure 4.1) are small five station arrays having a maximum station separation of approximately 50 kilometers. The third MV array (number 13; open circles in figure 4.1) consists of four stations with a maximum station separation of approximately 150 kilometers (comparable to the length of the EMSLAB line). We will concentrate initially on application of the model to horizontal magnetic field data. We will see that the model for the horizontal field noise covariance of (4.5.4) can fit the major features in the data, although the χ^2 goodness of fit statistics described in section 4.7 are often large enough to reject the null hypothesis that the model fit is adequate. This indicates that, at least in some cases, a more complicated model may be necessary. The estimates of the signal and noise covariance parameters can be used to compute estimates of the power spectra of the plane wave source, gradient and local noise

terms. We will see that for all four arrays considered here, the power in the gradients is small relative to local noise power for periods shorter than about 1000 seconds. For the smaller arrays the gradient effects on the horizontal fields can be virtually negligible for all periods. Based on this observation we consider application of the simple model of (4.5.7) to the three component MV data. This model neglects horizontal gradients but includes a term for vertical fields, constant at all stations in the array. Again, this model provides a qualitatively reasonable fit to the data, but, based on the χ^2 statistics the model fit is generally not completely adequate. In sections 4.9 and 4.10 we will extend these models to allow for a better treatment of vertical magnetic and electric field data.

We consider first the horizontal magnetic fields for the EMSLAB array. In figure 4.2 we plot the eigenvalues of the transformed horizontal field SDM

$$\mathbf{S'_H} = \mathbf{\Sigma_N^{-1/2}}(\hat{\boldsymbol{\theta}})\mathbf{S_H}\mathbf{\Sigma_N^{-1/2*}}(\hat{\boldsymbol{\theta}})$$

for a period of 1000 seconds. The eigenvalues of the untransformed matrix have been plotted previously in figure 2.11a. The power in the smallest eight eigenvalues of $\mathbf{S'_H}$ is virtually flat (compare to figure 2.11a where the smallest eight eigenvalues have a one-and-a-half order of magnitude range). This demonstrates, in a qualitative fashion, that the noise in the transformed data is reasonably isotropic, so that the gradients plus local noise model provides a reasonable fit to the data. The χ^2 statistics of section 4.7 allow for a more rigorous assessment of model fit. In figure 4.3a we plot the χ^2 statistics obtained from fitting three models to the EMSLAB horizontal field data for 16 periods ranging from 10^2 to 10^4 seconds. Model one assumes that the error covariance is isotropic $\mathbf{\Sigma_{hN}} = \sigma^2\mathbf{I}$. Model two assumes that all noise is local (i.e. incoherent between stations) so that $\mathbf{\Sigma_{hN}} = \mathbf{diag}(\mathbf{\Sigma_{h1}}, \cdots, \mathbf{\Sigma_{hn}})$, and model three is the full local noise plus gradients model given in (4.5.4). Note that because the stations in the array are all in a line so only two gradient vectors are included in this model.

The solid, dotted and dashed lines give the χ^2 statistics for models one, two and three, respectively. The degrees of freedom for the appropriate χ^2 distributions for the three models are, respectively, 63, 44 and 40, with 95% critical points of approximately 83, 60, and 56. These are indicated by horizontal lines on figure 4.3(a). The isotropic noise model fits the data very poorly at all frequencies with values of the χ^2 statistics more than an order of magnitude larger than expected. The full model, on the other hand, fits the data fairly well for the full range of frequencies. Note, however, that even for this model the χ^2 statistic exceeds the 95% levels for 5 out of the 16 periods. This result, taken at face value indicates that the simple local noise plus gradients model for the noise covariance matrix is not exactly correct. However other interpretations are possible. It could be that the discrepancy arises because our parameter estimates are not exactly maximum likelihood, because the assumption of a Gaussian data distribution is not valid, or because sample sizes are to small for the asymptotic χ^2 approximation to be valid. Given these uncertainties, and given the large improvement in model fit (figure 4.3(b), where the χ^2 is plotted on a linear scale makes this point much more dramatically) we believe it is fair to say that the model of (4.5.4) reasonably describes the major features of the noise covariances.

For model two, we omit the gradient terms. While this model is adequate at the shortest periods (below 400 seconds) it provides a poor fit to the data at longer periods. The results of figure 4.3 indicate that beyond about 1000 seconds the gradient terms become the most important source of 'noise'. The relative power in the plane wave, gradient and local noise terms can be computed from the model parameter estimates. Specifically, we have estimated the power in the gradients by $\hat{\sigma}_g^2 = \mathbf{Tr}[\hat{\mathbf{\Sigma}}_\beta]$, and estimated total noise power $\hat{\sigma}_N^2$ (i.e. total local noise in all 10 components from the diagonal elements of the matrices $\hat{\mathbf{\Sigma}}_{Ni}$, $i = 1, 5$. Plane wave signal power σ_p^2 is very well approximated as the sum of the two largest

eigenvalues of $\mathbf{S_H}$, $\hat{\sigma}_p^2 = \gamma_1 + \gamma_2$ (this can be refined to account for the effects of noise on the largest eigenvalues). We plot estimates of $\hat{\sigma}_p^2(\omega)$, $\hat{\sigma}_g^2(\omega)$ and $\hat{\sigma}_N^2(\omega)$ in figure 4.4. This figure is consistent with the χ^2 plots of figure 4.3. For very short periods, $\hat{\sigma}_g^2(\omega)$ is much smaller than $\hat{\sigma}_N^2(\omega)$, but for longer periods. $\hat{\sigma}_g^2(\omega)$ is larger than $\hat{\sigma}_N^2(\omega)$. At the longest periods the gradient power exceeds the local noise power by a factor of 3 or 4, so that the gradients become the dominant source of misfit to the plane wave model.

The local noise power spectrum is substantially flatter than the signal power spectrum. Roughly, we have for these estimated spectra

$$\hat{\sigma}_p^2(\omega) = c_p\omega^{-2} = c_pT^2 \qquad (4.8.1a)$$

$$\hat{\sigma}_g^2(\omega) = c_g\omega^{-2} = c_gT^2 \qquad (4.8.1b)$$

$$\hat{\sigma}_N^2(\omega) = c_N\omega^{-1} = c_NT \qquad (4.8.1c)$$

The increased importance of gradient terms in the noise covariance models for longer periods is due to the relative flatness of the local noise spectrum. There is no evidence from these results that the ratio $\hat{\sigma}_g^2(\omega)/\hat{\sigma}_p^2(\omega)$ increases with period (if anything, this ratio decreases with increasing period). This suggests that typical spatial length scales for sources in this frequency range are relatively independent of frequency (although we may interpret the apparent decrease of $\hat{\sigma}_g^2(\omega)/\hat{\sigma}_p^2(\omega)$ with frequency as weak evidence that longer period variations are characterized by slightly longer spatial length scales). We can use the results given in figure 4.4, together with our simple model from Chapter 3, to get a very rough feel for effective source length scales (but note that for this array, we have information about source variations in the east-west direction only). From figure 4.4 we find $\hat{\sigma}_g^2(\omega)/\hat{\sigma}_p^2(\omega) \approx 10^{-3}$. This can be compared to the results of figure 3.11 where we plot γ_3/γ_1 computed from the synthetic SDM $\mathbf{\Sigma}$ with $\delta_0 = 0.0$ (curve (d)). For the synthetic model $\gamma_3/\gamma_1 = 10^{-3}$ when $x_0 = x/r_0 \approx .03$. Since for the EMSLAB line

$x \approx 150km$, this yields source length scales $r_0 \approx 5000\ km$. This is a very crude approximation of the spatial characteristics of the sources; nonetheless the value is consistent with that obtained from larger array studies (e.g. Porath *et al.*, 1971). Based on the results of Chapter 3 where we obtained good results with a relatively larger non-dimensional array size ($x_0 = .1$), these results provide a strong justification for the plane wave source assumption for arrays the size of the EMSLAB line.

We have obtained similar results for the 3-component MV arrays. We discuss these briefly. In figure 4.5 - 4.7 we plot the χ^2 statistics for the three models described above, along with the estimated signal and noise power spectra, for the three MV arrays of figure 4.1. Degrees of freedom for all models are given in the figure headings. For all arrays the fit of the isotropic noise model (model 1) is poor. The success of the other two models is somewhat variable. For the small array 8 (figure 4.5) both model 2 and model 3 provide a reasonable fit to the data. Model 3, of course, provides the best fit, but the improvement in model fit due to the addition of the gradient terms is small. For array 13 (figure 4.6), which is comparable in size to the EMSLAB MT line, the inclusion of gradient terms does improve the fit substantially. For array 13, (as for array 8) the full model fits the horizontal field data quite well. The difference in the importance of the gradient terms can in part be ascribed to the difference in array size. The results from array 1 (figure 4.7), however, demonstrate that this may be only part of the story. The χ^2 plots for this array, which is comparable in size to array 8, are somewhat different than those for array 8. Here the gradient terms provide a substantial improvement in fit at longer periods. This difference reflects in part the variations of the source characteristics with time. It is also quite possible that a part of the difference is due to the strike of the arrays : Array 1 is essentially a north-south line, while array 8 is elongated more in an east-west direction, so the arrays detect

gradients in different directions. It is very likely that source gradients are larger in the north-south direction than in the east-west direction. Note also that neither model two or model three fits the data for array 1 particularly well; the fit at shorter periods is substantially worse than for the other two arrays.

In figures 4.5(b), 4.6(b), and 4.7(b) we plot the signal and noise spectra $\hat{\sigma}_p^2(\omega)$, $\hat{\sigma}_g^2(\omega)$ and $\hat{\sigma}_N^2(\omega)$ for the three MV arrays. The results are consistent with the χ^2 plots - $\hat{\sigma}_g^2(\omega)$ is relatively large at longer periods for arrays 1 and 13, but remains smaller than the local noise at all periods for array 8. The results for all parameters are also quite consistent both among themselves and with those discussed above for the EMSLAB array. The power law formulas of (4.8.1) for the spectra of these terms are approximately valid for all three arrays. Applying our approximate estimate of source length scales we find for arrays 1 and 13 $r_0 \approx 3000 - 4000\ km$. It is possible that the difference between the EMSLAB line results and these results reflects a real difference in typical north-south and east-west source length scales. At any rate, these results again offer strong justification for the uniform plane wave source model for arrays of this size.

We now consider the application of the simple linear model of (4.5.7) for the noise covariance matrix for three component data. For this model we neglect horizontal gradients and allow only local noise plus vertical fields which are coherent and in phase at all stations in the array. In figure 4.8(a) we plot the ordered eigenvalues of the 15×15 SDM **S** for array 8 at a period of 1600 seconds; in figure 4.8(b) we plot the eigenvalues computed from the transformed matrix

$$\mathbf{S}' = \boldsymbol{\Sigma}_N^{-\frac{1}{2}}(\hat{\boldsymbol{\theta}})\mathbf{S}\boldsymbol{\Sigma}_N^{-\frac{1}{2}*}(\hat{\boldsymbol{\theta}})$$

where $\hat{\boldsymbol{\theta}}$ is the parameter estimate for the model of (4.5.7). Again, the smaller eigenvalues for the transformed matrix are nearly constant, indicating qualitatively that the model captures at least the major features of the noise structure. For a more quantitative assessment of model fit we again turn to plots of the χ^2 statistic.

In figure 4.9(a) and (b) we plot these χ^2 statistics for fits to an isotropic noise model, and for the fit to the model of (4.5.7), for arrays 1 and 8. The asymptotic distributions of these statistic have, respectively, 169 and 149 degrees of freedom. For both arrays, the fit of the isotropic noise model is substantially worse for the full SDM **S** than for the horizontal field sub matrix $\mathbf{S_H}$ (compare to figures 4.5 and 4.6). As we have discussed in section 2.5, (see figures 2.8 - 2.10) the eigenvector associated with the third largest eigenvalue of the SDM can be well approximated by the vector $\mathbf{u_z}$ which consists of vertical fields which are equal for all stations. These vertical fields represent the major deviation of the data from the perfect plane wave source model. It is not surprisingly, then, that the vertical field plus local noise model of (4.5.7) provides a reasonable fit to the data. For both arrays 1 and 8, however, the misfit for this model is still substantially larger than the 95% critical values for the appropriate χ^2 distributions. This simple model thus does not provide a statistically adequate adequate fit to the noise covariance matrix. Not surprisingly, the failure of the model is more serious for array 1, where the gradients (which have been neglected in this model) are larger (figures 4.5 and 4.7).

To fit the noise covariance for three (or five) component data we will have to consider more complicated models. We consider these in the next two sections. We should emphasize that there is a substantial difference between statistical significance and physical significance. The simple model described here *does* capture the main features in the data. It is likely that this model will in practice prove quite adequate for routine processing of 3 component MV data from small arrays. The additional complications introduced in the next sections may not prove to be worth the trouble in many circumstances. The properties of the actual plane wave response parameters (and their associated error estimates) for the different noise parametrizations should be carefully examined. If the results for

the simpler model are essentially identical to those from the more complicated models considered below, then the simpler models will be preferable.

We close this section with a look at some estimates of the local noise covariance estimates. In figures 4.10(a)-(c) we plot the noise covariance parameters estimated for each station in the EMSLAB array - i.e. the noise power in the two horizontal field measurement channels, (geomagnetic north (H) 4.10(a); and geomagnetic east (D) 4.10(b)), and the magnitude of the correlation between the noise in the two channels (4.10c). The most striking feature occurs in the D noise components for station JEF. The noise power at this station is roughly a factor of 5 greater than at any other station. This station also has the largest noise power in H. Furthermore, note that, at least at shorter periods, noise in the H and D channels at JEF is correlated. There is also a strong peak in the correlation of noise between channels at a period of 2000 seconds for all stations. It is not clear at this point what the significance of this is. It is quite possibly a statistical artifact of some sort. Note that in general the noise curves for the stations in the EMSLAB profile generally remain parallel - sites (or instruments) tend to be quiet or noisy over a range of frequencies. Note also that the range of noise powers can be quite large - roughly an order of magnitude. This is consistent with the universally poor fit of the isotropic noise models which fail even when gradients can be neglected.

The results discussed here are just a sample of the sorts of analysis of noise structure that is possible with a multiple station interpretation of geomagnetic array data. More careful studies of a larger number of arrays should be conducted to help clarify the relative contributions of cultural, system and source related noise to the total model misfit. We have only given a very brief sketch of a few examples here.

4.9: Treatment of Electric and Vertical Magnetic Field Noise

In section 4.5 we developed a model for the noise covariance matrix which included local noise and the effects of gradients in the source fields. The model, and the methods of estimation developed in section 4.6, were completely satisfactory for horizontal magnetic field data. We now develop some techniques which allow a more satisfactory treatment of the more general case where the vertical magnetic and/or horizontal electric fields are included in the analysis. Again, we consider the special case of fitting a plane wave source model to data from a small array.

The methods described here are useful for both an exploratory analysis of the data and for more formal model fitting. We consider both sorts of applications in the next section, although we will emphasize the exploratory aspects. It is necessary to have a fairly clear picture of what sorts of effects must be included in a model before a more formal model fitting procedure can be applied. We do not consider the results presented here as final by any means. The results and methods discussed in this and the next section must be considered as somewhat tentative at this point. Further developments, in both model parametrization and estimation are needed.

As in section 4.5 we assume that the noise can be divided into local and gradient components, and we assume that the covariance matrix of the local noise is block diagonal. In addition we assume that the local noise vectors for the magnetic and electric field components at each station are uncorrelated. This will usually be justified for true system noise because magnetic and electric fields are measured by separate instruments. This should also be nearly true for local cultural noise. The E-fields are determined by the magnetic fields at the earth's surface in a region roughly one skin depth in radius; hence very localized disturbances of the magnetic fields have very little effect on the electric fields. With a

station spacing substantially less than the skin depth, truly local cultural noise in the magnetic fields has little effect on (and hence will be very weekly correlated with) the electric fields.

Our model and notation for the non-local gradient related 'noise' is as in section 4.5. Specifically, the three gradient vectors are written

$$\boldsymbol{\psi}_i = \begin{bmatrix} \boldsymbol{\psi}_{\mathbf{h}i} \\ \boldsymbol{\psi}_{\mathbf{z}i} \end{bmatrix} \qquad i = 1, 3$$

where the vectors $\boldsymbol{\psi}_{\mathbf{h}i}$ give the horizontal gradients (as in (4.5.2)), and the vectors $\boldsymbol{\psi}_{\mathbf{z}i}$ give the vertical magnetic and/or horizontal electric components which are associated with these gradient terms. We consider here the problem of estimating the vectors $\boldsymbol{\psi}_{\mathbf{z}i}$, $i = 1, 3$. The model results for the 1-d case considered in Chapter 3 suggest a simple form for the vectors $\boldsymbol{\psi}_{\mathbf{z}i}$, particularly for the case of three component MV data (see (4.5.5)). However, as discussed in section 4.5, it is not clear that these results will be relevant for the case of 2- or 3-dimensional conductivity distributions. In this section we will thus consider the general case where the form of the $\boldsymbol{\psi}_{\mathbf{z}i}$'s is unrestricted.

Our full model for the data is

$$\mathbf{X}_k = \mathbf{U}\boldsymbol{\alpha}_k + \boldsymbol{\Psi}\boldsymbol{\beta}_k + \boldsymbol{\xi}_k \tag{4.9.1}$$

where $\mathbf{U}$ and $\boldsymbol{\Psi}$ are the matrices whose columns are the plane wave and gradient vectors, respectively. We change notation slightly. Let

$$\boldsymbol{\eta}_k = \begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \end{bmatrix} \qquad \mathbf{W} = \begin{bmatrix} \mathbf{U} & \boldsymbol{\Psi} \end{bmatrix} = \begin{bmatrix} \mathbf{U}_\mathbf{h} & \boldsymbol{\Psi}_\mathbf{h} \\ \mathbf{U}_\mathbf{z} & \boldsymbol{\Psi}_\mathbf{z} \end{bmatrix} = \begin{bmatrix} \mathbf{W}_\mathbf{h} \\ \mathbf{W}_\mathbf{z} \end{bmatrix} \tag{4.9.2}$$

where $\mathbf{U}$ and $\boldsymbol{\Psi}$ are partitioned, as above, into horizontal magnetic ($\mathbf{h}$ subscript) and electric/vertical magnetic ($\mathbf{z}$ subscript) components. Note also that the vector $\boldsymbol{\eta}_k$ has a different meaning here and in section 4.5. With this new notation (4.9.1) is

$$\mathbf{X}_k = \mathbf{W}\boldsymbol{\eta}_k + \boldsymbol{\xi}_k \tag{4.9.3}$$

The covariance matrices for the random vectors $\boldsymbol{\eta}_k$ and $\boldsymbol{\xi}_k$ are

$$E(\boldsymbol{\xi}_k\boldsymbol{\xi}_k^*) = \boldsymbol{\Sigma}_\xi \qquad E(\boldsymbol{\eta}_k\boldsymbol{\eta}_k^*) = \boldsymbol{\Sigma}_\eta = \begin{bmatrix} \boldsymbol{\Sigma}_\alpha & 0 \\ 0 & \boldsymbol{\Sigma}_\beta \end{bmatrix} \tag{4.9.4}$$

so we may write for the expectation of the SDM:

$$\boldsymbol{\Sigma} = E(\mathbf{S}) = \mathbf{W}\boldsymbol{\Sigma}_\eta\mathbf{W}^* + \boldsymbol{\Sigma}_\xi = \begin{bmatrix} \mathbf{W}_\mathbf{h}\boldsymbol{\Sigma}_\eta\mathbf{W}_\mathbf{h}^* & \mathbf{W}_\mathbf{h}\boldsymbol{\Sigma}_\eta\mathbf{W}_\mathbf{z}^* \\ \mathbf{W}_\mathbf{z}\boldsymbol{\Sigma}_\eta\mathbf{W}_\mathbf{h}^* & \mathbf{W}_\mathbf{z}\boldsymbol{\Sigma}_\eta\mathbf{W}_\mathbf{z}^* \end{bmatrix} + \boldsymbol{\Sigma}_\xi \tag{4.9.5}$$

Partitioning $\mathbf{S}$ in the same way,

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{\mathbf{hh}} & \mathbf{S}_{\mathbf{zh}} \\ \mathbf{S}_{\mathbf{zh}} & \mathbf{S}_{\mathbf{zz}} \end{bmatrix}$$

we find, if the local noise covariance matrix $\boldsymbol{\Sigma}_\xi$ is diagonal,

$$E(\mathbf{S}_{\mathbf{zh}}) = \mathbf{W}_\mathbf{z}\boldsymbol{\Sigma}_\eta\mathbf{W}_\mathbf{h}^*$$

so that

$$\mathbf{W}_\mathbf{z} = E(\mathbf{S}_{\mathbf{zh}})[\mathbf{W}_\mathbf{h}(\mathbf{W}_\mathbf{h}^*\mathbf{W}_\mathbf{h})^{-1}\boldsymbol{\Sigma}_\eta^{-1}]$$

We don't know $\boldsymbol{\Sigma}\eta$, but we can estimate it from an analysis of horizontal magnetic fields using the model and methods developed in section 4.6. Using this estimate $\hat{\boldsymbol{\Sigma}}_\eta$, we may estimate $\mathbf{W}_\mathbf{z}$ by

$$\hat{\mathbf{W}}_\mathbf{z} = \mathbf{S}_{\mathbf{zh}}[\mathbf{W}_\mathbf{h}(\mathbf{W}_\mathbf{h}^*\mathbf{W}_\mathbf{h})^{-1}\hat{\boldsymbol{\Sigma}}_\eta^{-1}] \tag{4.9.6}$$

The last three columns of the estimated matrix $\hat{\mathbf{W}}_\mathbf{z}$ are the desired estimates of the vectors $\boldsymbol{\psi}_{\mathbf{z}i}$, $i = 1, 3$. With these parameters estimated, we may treat the vectors $\boldsymbol{\psi}_i$ as known and proceed as outlined in section 4.6 to estimate the remaining noise covariance parameters.

There is one serious problem with this approach. The estimates of the matrices $\boldsymbol{\Sigma}_\alpha$ and $\boldsymbol{\Sigma}_\beta$ must be reliable and non-singular. If array dimensions in one

or more directions are too small, it will not be possible to obtain a reliable estimate of the full gradient covariance matrix Σ_β. For the small arrays considered in this dissertation this is the typical case. We consider here one partial solution. To be explicit we consider the case of a linear array (in the x-direction, say) so that only two linearly independent gradient vectors are observable. More general cases are treated in an analogous fashion. We rewrite the horizontal gradient vectors as

$$\Psi_{h1} = \begin{bmatrix} x_1 \\ 0 \\ - \\ \cdot \\ \cdot \\ \cdot \\ - \\ x_n \\ 0 \end{bmatrix} \qquad \Psi_{h2} = \begin{bmatrix} 0 \\ x_1 \\ - \\ \cdot \\ \cdot \\ \cdot \\ - \\ 0 \\ x_n \end{bmatrix} \qquad \Psi_{h3} = 0 \tag{4.9.7}$$

and let Ψ_{zi}, $i = 1, 3$ be the corresponding vertical magnetic/electric vectors. Note that the vectors Ψ_{h1} and Ψ_{h2} of (4.9.7) are just linear combinations of the canonical gradient vectors of (4.5.2). To distinguish the last column of $\mathbf{W} = (\mathbf{U}\ \Psi)$, partition the matrices $\mathbf{W}$ and Σ_η of (4.9.2) and (4.9.4) as

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_h \\ \mathbf{W}_z \end{bmatrix} = \left[\begin{array}{c|c} \mathbf{W}_{h1} & 0 \\ - & - \\ \mathbf{W}_{z1} & \mathbf{W}_{z2} \end{array}\right]$$

$$\Sigma_\eta = \left[\begin{array}{c|c} \Sigma_{\eta 11} & \Sigma_{\eta 12} \\ - & - \\ \Sigma_{\eta 21} & \sigma_{\eta 22} \end{array}\right] = \left[\begin{array}{cc|c} \Sigma_\alpha & 0 & 0 \\ 0 & \Sigma_{\beta 11} & \Sigma_{\beta 12} \\ - & - & - \\ 0 & \Sigma_{\beta_{12}} & \sigma_{\beta 22} \end{array}\right]$$

Note that the vector $\mathbf{W}_{z2} = \Psi_{z3}$ gives the vertical magnetic/electric fields associated with the (unobservable) gradients perpendicular to the array strike.

For this case where we can estimate $\Sigma_{\eta 11}$, but not the full matrix Σ_η, there is a fundamental non-uniqueness in the model parametrization. To see this let

$$\mathbf{W}'_{z1} = \mathbf{W}_{z1} + \mathbf{W}_{z2}\Sigma_{\eta 21}\Sigma_{\eta 11}^{-1} \qquad \sigma'_{\eta 22} = \sigma_{\eta 22} - \Sigma_{\eta 21}\Sigma_{\eta 11}^{-1}\Sigma_{\eta 12} \tag{4.9.8}$$

$$\mathbf{W}' = \begin{bmatrix} \mathbf{W}_{h1} & 0 \\ \mathbf{W}'_{z1} & \mathbf{W}_{z2} \end{bmatrix} \qquad \Sigma'_\eta = \begin{bmatrix} \Sigma_{\eta 11} & 0 \\ 0 & \sigma'_{\eta 22} \end{bmatrix} \tag{4.9.9}$$

Then it is easily verified that

$$\mathbf{W}'\Sigma'_\eta\mathbf{W}'^* = \mathbf{W}\Sigma_\eta\mathbf{W}^* \tag{4.9.10}$$

The vertical magnetic/electric field vectors, represented by the matrix $\mathbf{W}_z$, thus cannot be recovered uniquely. For the case of a linear array considered here we can, however, estimate $\Sigma_{\eta 11}$. If we substitute $\hat{\Sigma}_{\eta 11}$ and $\mathbf{W}_{h1}$ for $\hat{\Sigma}_\eta$ and $\mathbf{W}_h$ in (4.9.6), we can form

$$\tilde{\mathbf{W}}_{z1} = \mathbf{S}_{zh}[\mathbf{W}_{h1}(\mathbf{W}_{h1}^*\mathbf{W}_{h1})^{-1}\Sigma_{\eta 11}^{-1}] \tag{4.9.11}$$

A simple calculation shows that this is generally biased as an estimate of $\mathbf{W}_{z1}$, but is (asymptotically) unbiased as an estimate of the matrix $\mathbf{W}'_{z1}$ defined in (4.9.8); i.e.

$$E(\tilde{\mathbf{W}}_{z1}) \rightarrow \mathbf{W}'_{z1} = \mathbf{W}_{z1} + \mathbf{W}_{z2}\Sigma_{\eta 21}\Sigma_{\eta 11}^{-1} \tag{4.9.12}$$

In fact, it is easily verified that with our assumption that gradient and plane wave terms are uncorrelated, the first two columns of $\mathbf{W}_{z1}$ and $\mathbf{W}'_{z1}$ (i.e. the two columns of $\mathbf{U}$; see (4.9.2)) are identical - only the estimates of the fields associated with the gradient terms Ψ_{z1} and Ψ_{z2} are biased. The same holds for the non-uniqueness of the covariance parametrization noted above - only the gradient part of covariance is over-parametrized. Thus for $\mathbf{W}'$ and Σ'_η defined in (4.9.8) and (4.9.9) we have

$$\mathbf{W} = (\mathbf{U}\ \Psi) \qquad \mathbf{W}' = (\mathbf{U}\ \Psi') \qquad \Psi' = \begin{bmatrix} \Psi_{h1} & \Psi_{h2} & 0 \\ \Psi'_{z1} & \Psi'_{z2} & \Psi_{z3} \end{bmatrix} = \left[\begin{array}{c|c} \Psi'_1 & \Psi_3 \end{array}\right]$$

$$\Sigma'_\eta = \begin{bmatrix} \Sigma_\alpha & 0 \\ 0 & \Sigma'_\beta \end{bmatrix} \qquad \text{where} \quad \Sigma'_\beta = \begin{bmatrix} \Sigma_{\beta 11} & 0 \\ 0 & \sigma'_{\beta 33} \end{bmatrix} \tag{4.9.13}$$

We can thus rewrite the parametrization of the noise covariance matrix as

$$\Sigma_N = \Psi\Sigma_\beta\Psi^* + \Sigma_\xi = \Psi'\Sigma'_\beta\Psi'^* + \Sigma_\xi$$
$$= \Psi'_1\Sigma_{\beta 11}\Psi'^*_1 + \sigma'_{\beta 33}\psi_{z3}\psi^*_{z3} + \Sigma_\xi \qquad (4.9.14)$$

We can estimate the parameters which determine Ψ'_1 from the columns of $\tilde{W}_{z1}$ of (4.9.11). If we know, or can estimate, ψ_{z3} (up to a multiplicative constant) then, treating the estimated vectors $\hat{\psi}'_{z1}$, $\hat{\psi}'_{z2}$ (and ψ_{z3}) as known, (4.9.14) is a linear covariance matrix model of the form discused in section 4.5. As an example of a situation where ψ_{z3} is (approximately) known we may cite the case of three component MV data. Here we expect that ψ_{z3} should be approximately proportional to the vector $\mathbf{u}_z$ given in (2.5.2). We summarize the suggested procedure here. First analyze the horizontal field data (as in sections 4.5,4.6 and 4.8), estimating Σ_α and $\Sigma_{\beta 11}$. Use these in the estimator of (4.9.11) to find estimates $\hat{\psi}'_{z1}$ and $\hat{\psi}'_{z2}$. Taking these, together with the assumed form of ψ_{z3} ($= \mathbf{u}_z$ for MV data) as known, the model for the noise covariance matrix of (4.9.14) is linear in the parameters Σ_ξ, $\sigma'_{\beta 33}$, $\Sigma_{\beta 11}$. These can be estimated (along with the response space parameters) using the generalized least squares estimate of section 4.6.

Note that, if the model assumptions hold, this procedure yields a consistent estimate of the noise covariance matrix Σ_N even though the vectors ψ_{z1} and ψ_{z2} cannot be estimated. From the standpoint of estimating Σ_N, then, the problem of unobservable gradients need not be significant. On the other hand, the vectors ψ_{zi} have a physical interpretation, and they are of interest in their own right. These vectors, which give the vertical fields associated with the gradients represented by ψ_{hi} (i.e. the vertical fields which would be seen if the horizontal magnetic fields were given exactly by the pure horizontal gradient represented by ψ_{hi}), can not be estimated. The vectors which we can estimate, ψ'_{z1} and ψ'_{z2}, do not have such a nice physical interpretation. These vectors do represent vertical fields which are

correlated with the gradients, but if the coefficient β_3 of the unobservable gradient vector ψ_{h3} is correlated with the coefficients β_1 and β_2 of the observable gradient vectors ψ_{h1} and ψ_{h2}, these vertical fields will, in part, be due to the unobserved gradient. It will not be possible in this case to separate the fields due to the observable gradient from those due to the unobservable gradient. Note that even for the idealized simple case of isotropic, spatially homogeneous sources with an evenly spaced linear array, β_3 is correlated with β_1, so this sort of correlation should be expected (see discussion in section 3.6). In summary, for three component MV data, we can always obtain approximate estimates of the noise covariance matrix, but we can only estimate the vertical fields associated with specific gradients when all three gradient terms are observable.

There are other approaches to the problem of estimating/exploring the form of the electric/vertical magnetic field components associated with the gradient terms. The most obvious approach is to treat these as signal so that the response space $\boldsymbol{R}$ is now of dimension $p = 5$. By including the gradient terms as signal we can (at least for the small arrays considered here) assume that the remaining noise is all local so that the noise covariance matrix Σ_N is block diagonal. With this simple (linear) form for Σ_N we may use the iterative methods of estimation for noise and signal parameters outlined in sections 4.4 - 4.6. In fact, in this case our approach to the simultaneous estimation of the signal and noise parameters is somewhat less than ideal. The justification for our approach is based upon the assumption that the signal power is large compared to the noise power. In fact, the power in gradient and local noise terms does not always satisfy this condition (see figures 4.4 - 4.7). If the gradients are treated as signal, then, the condition of a large signal-to-noise ratio is usually violated. In our experience the iterative scheme of sections 4.4 - 4.6 usually works fairly well even in this case; the method does however break down at times, particularly when local noise is very

large in one or more channels.

If we take the point of view that the response space is of dimension 5, we can estimate the response space $\boldsymbol{R}$ (along with the local noise parameters) as in sections 4.4 - 4.6. For interpretation purposes we would like to identify vectors in $\boldsymbol{R}$ with plane wave and gradient source terms. We propose one simple approach to this problem. Let $\{ \mathbf{v}_i, i = 1, 5 \}$ span $\boldsymbol{R}$, and let

$$\mathbf{V} = \begin{bmatrix} \mathbf{V_h} \\ \mathbf{V_z} \end{bmatrix}$$

where the columns of the $m \times 5$ matrix $\mathbf{V}$ are the vectors $\mathbf{v}_i$, and where, as above, the matrix has been partitioned into horizontal magnetic components ($\mathbf{V_h}$) and vertical magnetic/electric components ($\mathbf{V_z}$). If we analyze the horizontal magnetic fields separately, we can, as in sections 4.4 - 4.6, fit a plane wave response space ($p = 2$) model, obtaining an estimate of the horizontal field components for the plane wave sources - in the notation of (4.9.2) these are given by the columns of the matrix $\mathbf{U_h}$. We expect the horizontal components of the gradient vectors to be approximately equal to the three horizontal gradient vectors $\boldsymbol{\Psi}_{\mathbf{h}i}$. We thus expect that the horizontal components of the response space vectors will approximately be linear combinations of these five (two plane wave and three gradient) vectors. If this expectation is satisfied exactly we will have, in the notation of (4.9.2),

$$\mathbf{V} = \begin{bmatrix} \mathbf{V_h} \\ \mathbf{V_z} \end{bmatrix} = \begin{bmatrix} \mathbf{W_h} \\ \mathbf{W_z} \end{bmatrix} \mathbf{A} \qquad (4.9.15)$$

for some matrix $\mathbf{A}$. With these assumptions we know (or at least have estimates of) $\mathbf{W_h} = (\mathbf{U_h} \ \boldsymbol{\Psi}_{\mathbf{h}})$, so if (4.9.15) holds exactly we could solve the overdetermined system of equations $\mathbf{V_h} = \mathbf{W_h A}$ for the matrix $\mathbf{A}$. In fact, this system of equations will never be satisfied exactly for any matrix $\mathbf{A}$, so we must seek an approximate solution using least squares. Specifically, we can find $\hat{\mathbf{A}}$ which

minimizes $\| \mathbf{V_h} - \mathbf{W_h}\hat{\mathbf{A}} \|$. As usual we find

$$\hat{\mathbf{A}} = (\mathbf{W_h^* W_h})^{-1} \mathbf{W_h^* V_h} \qquad (4.9.16)$$

so we can then estimate $\mathbf{W_z}$ by

$$\hat{\mathbf{W}}_\mathbf{z} = \mathbf{V_z}\hat{\mathbf{A}}^{-1} = \mathbf{V_z}(\mathbf{W_h V_h})^{-1} \mathbf{W_h^* W_h} \qquad (4.9.17)$$

Note that this method of estimating $\mathbf{W_z}$ suffers from the same problems as the estimator of (4.9.6) - the estimates of (4.9.16) and (4.9.17) can only be computed if all three gradient terms are observable (otherwise $\mathbf{W_h^* W_h}$ will not be invertible). Of course, it is always possible to fit a 5-dimensional response plane plus local noise model to the data, but identification of vertical magnetic and electric field components with specific gradient terms (i.e. estimating $\mathbf{W_z}$) requires that all three gradient terms be observable. Note that the approximate 'fix' of reparametrizing the noise covariance as in (4.9.14) can be adapted to this method of estimating $\mathbf{W_z}$ as well.

One advantage to an approach which treats the gradient terms as signal is that it allows a check on the validity of the assumption that the horizontal magnetic fields can be well approximated as the sum of a pair of arbitrary plane wave source vectors and three fixed gradient vectors. If this assumption holds, then the residual $\mathbf{V_h} - \mathbf{W_h}\hat{\mathbf{A}}$ will be small. We will show in the next section that the residual vectors associated with the gradient terms are not always small, and that for a fixed array they can have a similar form for a range of frequencies. These results suggest that there can be systematic deviations in the horizontal magnetic fields from the simple gradient model which was suggested by the results from the 1-d models of Chapter 3. We will explore these effects briefly in the next section where we consider applications of the methods described here to real data.

4.10: Some Applications to Vertical Magnetic and Electric Field Data

We now consider application of the methods of section 4.9 to some actual arrays. The results described here are of a somewhat exploratory nature; we will concentrate on outlining the types of gradient related effects on electric and vertical magnetic field data which are revealed by application of these methods. We will see in general that there are coherent electric and vertical magnetic fields which are correlated with gradients in the horizontal magnetic fields. The effects on the vertical magnetic fields are qualitatively similar to those predicted for the simple 1-d model, although there are some significant differences. The effects on the electric fields are substantially different than predicted by the 1-d model. Our results strongly suggest that in 2- and 3-dimensional situations non-zero wavenumber sources have a first order effect on the impedances. This contrasts with the 1-d case where such source effects are second order effects (Dmitriev and Berdichevsky, 1979; Mareschal, 1986). Note, however, that none of the arrays that we have analyzed are of sufficient size in all directions to produce consistent estimates of the power in all three gradient terms. As a consequence, our conclusions are somewhat tentative. More definitive statements about some of the points raised here will require the analysis of data from larger two dimensional arrays.

We begin with an example application of the estimate of the vectors Ψ_{zi} given by (4.9.6) for three component (MV) data (so that the vectors Ψ_{zi} give the vertical fields associated with the corresponding gradient terms). In figure 4.11 we give the results of this exercise for array 1 (see figure 4.1 for array locations). This array is very narrow in the east west direction and, as a result, we find that obtaining consistent results for all three gradients is impossible. In general two gradient vectors are observable and these are linear combinations of the three fundamental gradient vectors given in (4.5.2). Our method for automatically deciding which two linear combinations to include in the model is given at the end of

section 4.6. In this case, the resulting vectors correspond approximately (but not exactly) to north-south gradients in B_x and B_y. The actual gradient vectors are plotted at the station locations in figures 4.11(a) and (b) with the estimated value of the corresponding (complex) vertical field printed beside the station. Note that the horizontal field vectors are all real (zero phase) and, as discussed in Chapter 3, they, and the corresponding vertical components, have been scaled in kilometers (the units of $C(\omega)$). Note also that at the longest period plotted (3600 seconds) only a single gradient vector was resolvable (by the criterion of section 4.6). As discussed in the previous section, the estimated vertical field components should be interpreted as estimates of the vertical fields *correlated* with the observable gradients, not the vertical fields *caused* by these gradients.

At all periods B_z is very roughly equal and in phase at all stations in the array. This is what we would expect from the 1-d model of Chapter 3. The deviations from this simple form are greatest at short periods. We believe that this mostly reflects the effects of noise on the estimated parameters since the gradient related signals are smaller (relative to local system noise) at shorter periods. Larger amplitudes of B_z are associated with the north-south gradients in B_x, and amplitudes increase with increasing periods. Both of these results are expected from the analysis of the 1-d model. In fact, however, for a 1-d earth there should be no vertical fields associated with north-south gradients in the east-west field component B_y. This is clearly not the case here. Again, however, we must regard the significance of this observation with some caution since our estimates cannot be interpreted as the vertical fields caused by the observable gradient. Note also that the phase of the vertical fields appears to be somewhat unstable between periods. We believe that this also is due to the effect of the unobservable gradient. In summary, while the general results given here (and these are typical of those obtained for other small arrays) are reasonably consistent with the 1-d

model, there is some indication of deviations (which may be significant) from the simple model. The effect of unobservable gradients on the vertical fields renders any attempt at quantitative physical interpretation of these results hopeless.

Even though the physical interpretation of the parameter estimates is questionable we can still estimate the noise covariance matrix and assess the adequacy of the model. For three component data we can estimate the vertical fields which are correlated with the resolvable gradients and adopt the model of (4.9.14) for the noise covariance matrix with $\Psi_{z3} = \mathbf{u}_z$. We can then use the methods of section 4.6 to estimate the noise covariance matrix, and then use this estimate of the noise covariance in our iterative procedure to improve estimates of the plane wave response. As with the horizontal magnetic field models, we can assess the goodness of fit of the overall model using the χ^2 statistics of section 4.7. In figure 4.12(a)-(c) we plot these χ^2 statistics versus frequency for the three arrays 1, 8 and 13. In each figure we plot results for three models - the solid lines give the χ^2 for model one which assumes isotropic noise, the dashed line gives the χ^2 for the model of (4.9.14) (model two) and the dotted line is for the model discussed at the end of the previous section which assumes that the response space is of dimension $p = 5$ and that all noise is local (model three).

The number of degrees of freedom for the χ^2 statistics are given in the figure captions. The isotropic model does not come close to providing an adequate fit to the data. The goodness of fit for model two is substantially better, but not statistically adequate. The results for model three, on the other hand, indicate that this model generally provides an adequate fit to the data (although the fit for the smaller array 13 is not so good). This last result indicates that the dimension of the response space which can be resolved above the noise is at most five (as expected). We have tried fitting lower dimensional response spaces and have generally found the fit to be inadequate. An example of this is given in figure 4.13

where χ^2 statistics for fits with $p = 2$, 3, 4 and 5 are plotted for array 1 along with the 95% confidence levels for the χ^2 statistics. The results in this figure are typical of those obtained for a number of arrays.

We can apply the methods of section 4.9 to estimate the effects of gradients on the electric fields. In figures 4.14 and 4.15 we plot the electric field vectors which are correlated with the two observable horizontal field gradients for the five stations in the EMSLAB line at periods of 2700 and 900 seconds, respectively. Note that the assumed horizontal gradients together with the correlated vertical components are plotted together with the corresponding electric fields. The results for 2700 seconds are clearest (figure 4.14). For both of the observable gradients the correlated electric field vectors point in roughly the same direction (east-west) and have approximately the same phase at all stations in the array. This demonstrates that there is a component of the electric fields, coherent across the array, which is correlated with gradients in the horizontal fields. There is considerable variation in the magnitude of the correlated electric fields between stations. These variations in magnitude are approximately the same as those noted for the electric fields associated with the plane wave sources - the same stations have large (or small) electric fields for both cases (see figure 2.6). At 900 seconds there is some evidence of a similar pattern for the correlated electric fields, but there is considerably more scatter in both the direction and phase of the electric fields. The vertical fields given for these two periods are roughly equal across the array although, at this period, there are some significant deviations from uniformity, particularly at JEF (the eastern-most station).

We have also applied the second estimation approach suggested in section 4.9 : we have treated the gradients (and the associated electric/vertical magnetic fields) as signal and fit a response space of dimension $p = 4$. We then used the approach suggested in (4.9.15) - (4.9.17) to find linear combinations of the

response space vectors whose horizontal components are most nearly equal to the two observable gradient vectors. The results of this exercise are presented in figures 4.16 and 4.17 again for periods of 2700 and 900 seconds. The results are very similar for the two periods. For both periods the horizontal fields are very nearly equal to the two observable gradient vectors; the fit is slightly better at 2700 seconds, but is still quite good at 900 seconds. The electric fields correlated with these gradients all point in the same direction (roughly east-west) and they are in phase across the array. The vertical fields are also roughly equal across the array. These results are similar to those obtained with the other estimation approach, particularly at 2700 seconds. At 900 seconds however, the results from the second approach are much cleaner - they appear to be less noisy, and the aberrations at JEF are not evident.

With both methods of estimating the electric fields which are correlated with horizontal magnetic field gradients, we find significant interstation variations in the amplitudes of the electric fields. Since the pattern of these variations is similar to that observed for the dominant plane wave related electric fields, it seems likely that these variations are at least in part due to local distortion of the electric fields (Larsen, 1977; Zhang *et al.*, 1987). To demonstrate this we have applied an extremely crude correction to the estimates of gradient related electric fields. Assuming that the inter-station variations of the amplitudes of the plane wave caused electric fields are due only to local distortions (this is clearly only a very crude first approximation), we may adjust the gradient caused electric fields by dividing by the plane wave electric field amplitudes. The results of doing this are presented in figure 4.18 for three periods - 900, 2700 and 8200 seconds. Indeed, we now see that the electric field vectors at all sites are of comparable amplitude; clearly the local distortions of the electric fields are similar for the plane wave and gradient terms.

Without the effect of local surface distortion, then, the electric fields correlated with the horizontal field gradients are approximately constant across the array. This result is not predicted by a 1-d model. In Chapter 3 we showed that for a 1-d earth there were gradients in the electric fields associated with the gradients in the magnetic fields. In fact, for a 1-d earth, the electric fields at any station are related to the magnetic fields at that station via the standard plane wave impedance

$$E_x = \zeta(\omega)B_y \qquad E_y = -\zeta(\omega)B_x$$

For the EMSLAB data the electric fields correlated with the gradients are much larger than predicted by the 1-d model and they are, allowing for local surface distortion, roughly constant across the array.

As we have discused in Chapter 3 (see also Dmitriev and Berdichevsky, 1979) the electric fields over a 1-d earth are related via a convolution operator to B_x and B_y on the earth's surface (see 3.2.2 - 3.2.3). Since the kernel of this operator is rotationally symmetric, the electric fields are determined by the local value of B_x and B_y. The electric fields are completely insensitive to local gradients in the magnetic fields while the vertical magnetic fields B_z, are essentially determined by the local gradients. Thus, for a uniform magnetic field gradient over a 1-d earth, B_z is constant while the electric fields have uniform gradients. Put another way, finite wavenumber source fields have a first order effect on vertical magnetic fields but only a second order effect on the electric fields (Mareschal, 1986). Our results show that the results of the 1-d analysis are not always relevant in 2- and 3-dimensional environments. The results for the electric fields strongly suggest that finite wavenumber fields can have a first order effect on the electric fields in some circumstances.

It is noteworthy that the electric fields associated with the gradients are much larger in the east-west direction. Since the dominant conductivity gradient is

east-west, this corresponds to the B-polarization or TM mode. We suggest a simple physical explanation for our results. The east-west electric fields in the region of EMSLAB array are in part due to the accumulation of charge in regions with large conductivity gradients - in particular at the continental margin where the conductive ocean abuts the relatively resistive continental crust. The magnitude of this charge will depend on the electric fields in the ocean; these fields will in turn depend mostly on the remote (i.e. over the ocean) magnetic sources. As a consequence, the electric fields at a station on land will depend relatively more on the magnetic fields to the west than to the east of the station. This asymmetry is precisely what is required to produce the observed relationship between the gradients and the correlated electric fields seen in figures 4.14-4.18.

The results here are suggestive, but as we have stressed above, the fact that gradients can be measured in only one direction makes a quantitative interpretation questionable at best. In fact, the EMSLAB experiment consists of a large array of instruments, both on land and on the ocean bottom (The EMSLAB Group, 1987). When this data becomes available a more careful study of gradient effects on the electric (and vertical magnetic) fields will be conducted.

We have seen that at least some of the 'noise' observed in the electric fields can be ascribed to source effects. To assess the relative importance of this we must consider the nature of local noise in the electric fields. In figures 4.19(a) and (b) we plot the estimated local noise powers for the north and east components of the electric fields respectively. We see that the noise power is quite variable from channel to channel, particularly at short periods (< 300s) where variations in noise power between different stations exceed two orders of magnitude. Note that the spread in noise power decreases with period, but that stations which are quiet at shorter periods tend to remain so at longer periods. Note also that at longer periods the local noise tends on average to be larger in the east-west component of

the fields. This is probably an artifact of the parameter fitting procedure. We have fit a model which allows two gradient terms plus local noise. The third, unobserved gradient has some effect on the electric fields, and this is undoubtedly greater for the east-west component. Some of this source effect is probably included in the local noise parameter estimates.

We can use the estimates of power in the plane wave and gradient magnetic fields, together with the estimates of the amplitudes of the associated electric fields, to estimate the power in the plane wave and gradient components for each electric field component. In figure 4.20 and 4.21 we plot these two quantities, (solid lines are plane wave; dashed are gradients) together with the local noise estimates (dotted lines) for stations VAL (figure 4.20) and AME (figure 4.21) on the EMSLAB profile. In general, we see that there are significant differences between components and between stations. This reflects both the variability in the local noise power and the variations of power in gradient related effects caused by surface distortion. VAL represents a 'resistive' site with large electric fields, and power in the gradient terms is large relative to the local noise. This is particularly true for the east-west component where the gradient noise exceeds the local noise by more than an order of magnitude. At this site, where the plane wave signal is large, the local noise is comparatively negligible at all periods. AME represents the opposite situation - this is a 'conductive' site with small electric fields. Here gradient related effects are not to dominant. For the east-west component, gradient and local noise are of comparable amplitude while for the north-south component, the local noise is larger than the gradient effects. These results show that deviations of electric field data from the simplifying plane wave assumption can have very significant source gradient related components, even at periods of 100 seconds or less. The relative importance of local and source related noise is, however, highly site, and polarization, dependent.

In our analysis of horizontal magnetic field data we assumed that the total fields due to gradients in the sources had the same simple forms as those obtained for the 1-d model. In fact, with a 2-d or 3-d conductivity distribution, the total horizontal fields associated with source gradients will be perturbed from the simple 1-d form. If we treat the gradient terms as signal we can, as outlined in section 4.9, estimate the form of these perturbations (along with the vertical fields associated with the source gradients). Recall that this is accomplished by treating additional eigenvectors as signal and fitting a noise model which allows only for local noise. We find (as illustrated for the EMSLAB line in figures 4.16 and 4.17 above) that these additional eigenvectors typically do resemble the simple form predicted for the 1-d conductivity case (gradients in the horizontal fields, nearly uniform vertical fields) plus a perturbation. To get a better picture of the form of these deviations from the 1-d gradient model for a fixed eigenvector, it is useful to subtract the best fitting linear combination of the canonical gradient vectors given in (4.5.2) from the horizontal fields and to subtract the average vertical field from the estimated vertical field components so that the perturbations can be separated from the gradients. In figure 4.22 we plot the deviations from the 1-d gradient model computed in this manner at three periods for the third eigenvector from MV array 8. Note that this 'eigenvector' is computed from the model which assumes a response space of dimension $p = 5$ and fits the local noise structure (so it does not really represent an eigenvector of the raw SDM). The power in this eigenvector, relative to the dominant plane wave eigenvectors, is on the order or 10^{-2}, and most of this power is in the relatively uniform vertical fields. The relative power in the perturbations from the expected gradient vector form is on the order of 10^{-3} or less. These perturbations represent a small amount of the power in observed fields.

For reference in figure 4.22(a), we plot the anomalous horizontal and vertical fields computed for a plane wave source, linearly polarized in the north south direction. For this plot we have assumed that the average field is normal so these plots represent the deviations of the total magnetic fields from a uniform north-south field. The plot of anomalous horizontal fields shows that the horizontal fields are larger in the center of the array and smaller at the stations on the northeastern and southwestern edges, while the vertical fields reverse sign between the northern and southern stations. This pattern of anomalous fields indicates that there is a region of enhanced conductivity passing through the center of the array. This conductive anomaly has been more fully mapped, and will be described in slightly more detail in a regional context in the next chapter. The three sets of perturbation fields plotted in 4.22(b)-(d) are all very similar in form and, at least qualitatively, they are quite similar to the horizontal and vertical anomalous fields plotted in figure 4.22(a). It is thus reasonable to ascribe these fields to current flowing in the central region of anomalous conductivity. These currents are correlated with the gradients but are not correlated with the two dominant (plane wave) eigenvectors. Note that the fields of 4.22(a) and 4.22(b)-(c) are scaled in an entirely different manner; the pattern of variations is all that can be compared directly.

This is a graphic demonstration of the phenomena of current channeling (Babour and Mosnier, 1979; Woods and Lilley, 1980; Summers, 1982; Jones, 1983; Hart *et al.*, 1983; Mosnier, 1985) - a portion of the current flowing in the anomaly is not coherent with the local horizontal source fields and hence must be non-locally induced. It is worth noting here that with perfect plane wave sources all current flowing in the anomaly (indeed all current everywhere) will be coherent with the local fields, although a substantial fraction of the current may be induced in a region which is much larger than the array. The effects seen here are due to

the interaction of the finite spatial scale of the sources with the 3-dimensional conductivity distribution. The magnitude of this effect thus depends on properties of both the conductivity distribution and the sources. The fact that this effect is as small as it is does not imply that most of the current flowing in the anomaly is locally induced. It does imply that most, but not all, of the current is induced in a region where the fields are highly correlated with the local fields.

The current channeling seen for array 8, and the source field effects on the electric fields seen for the EMSLAB line, are almost certainly manifestations of the same physical phenomenon. In both cases the electric fields on the land side of the coast are influenced by the magnetic fields over the ocean, and, because of the finite spatial scales of the sources, these magnetic fields are not perfectly correlated with the local inducing fields. As a result, there is a portion of the electric fields which is coherent over a large area on land, but which is not coherent with the local inducing fields. This effect is strongest in the east-west component of the electric fields, as we have seen for the EMSLAB data. In the case of MV array 8, which is about 50 kilometers from the coast and straddles an east west trending conductivity anomaly (which probably is reasonably well connected electrically to the ocean), these electric fields lead to anomalous current flow which is not coherent with the local inducing fields.

In this section we have concentrated on a more or less exploratory analysis of some of the effects of finite spatial scale sources on magnetic and electric fields observed in arrays. Our motivation for this analysis derived from our need to estimate the form of the noise covariance matrix in order to carry out the estimation scheme developed in previous sections of this chapter. In many ways, however, the results obtained are more interesting for their own sake, at least to anyone with an interest in the complications of electromagnetic induction by natural sources in a 3-dimensional earth. It is important at this point to reiterate that the multivariate

methods developed in this dissertation can be useful both for exploratory data analysis, and for the estimation of parameters within a formal statistical model. Although philosophically there is a wide chasm between these approaches to data analysis, in practice the distinction often becomes a difficult one to make. The results presented in this section represent only a very tentative exploratory analysis of the effect of finite wavenumber sources on vertical magnetic and electric fields. Substantial additional analysis of data and of the physics of electromagnetic induction will be required to develop an adequate model for these features in the data. Such a model will be required to carry out the statistical parameter estimation described in this chapter.

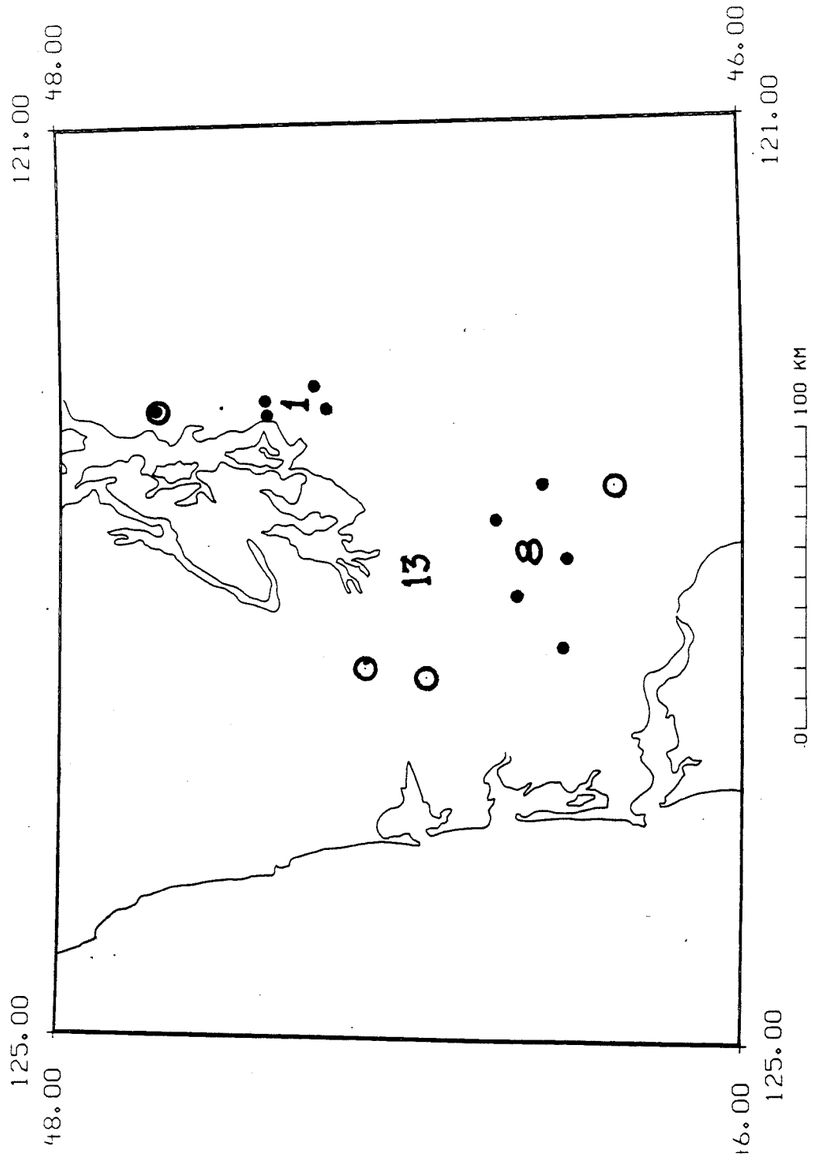


Figure 4.1 Maps of three Western Washington MV arrays discussed in sections 4.8 and 4.10.

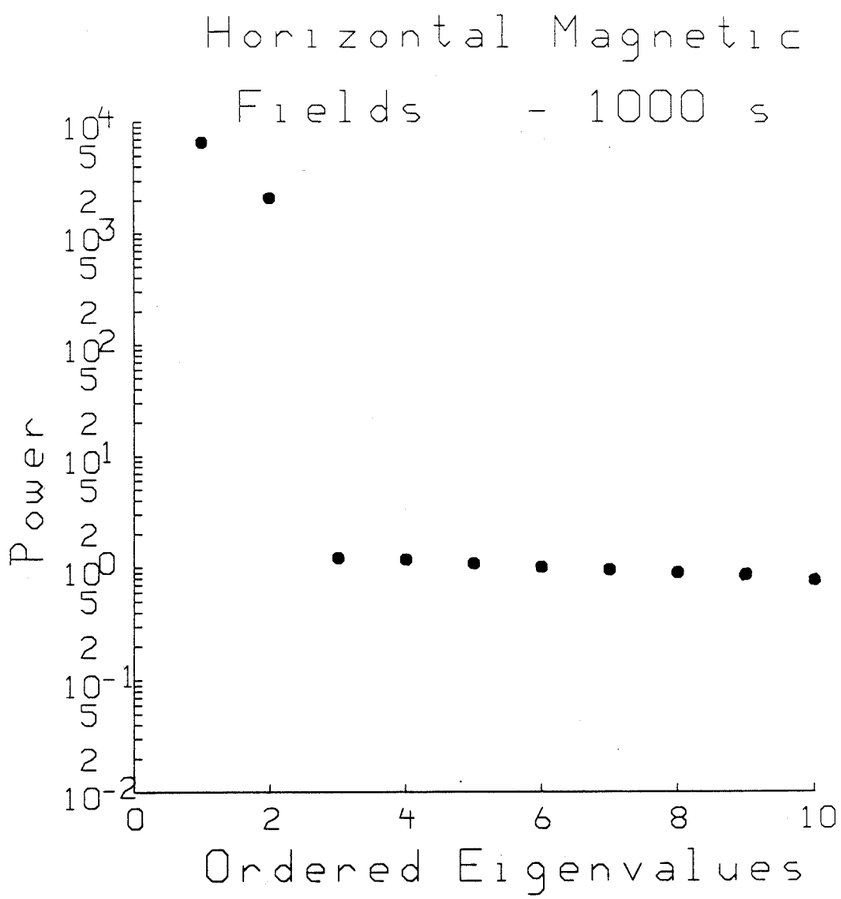


Figure 4.2 Eigenvalues of $S'_H = \Sigma_N^{-1/2} S_H \Sigma_N^{-1/2*}$ for EMSLAB array; period is 1000 seconds (compare to figure 2.11a).

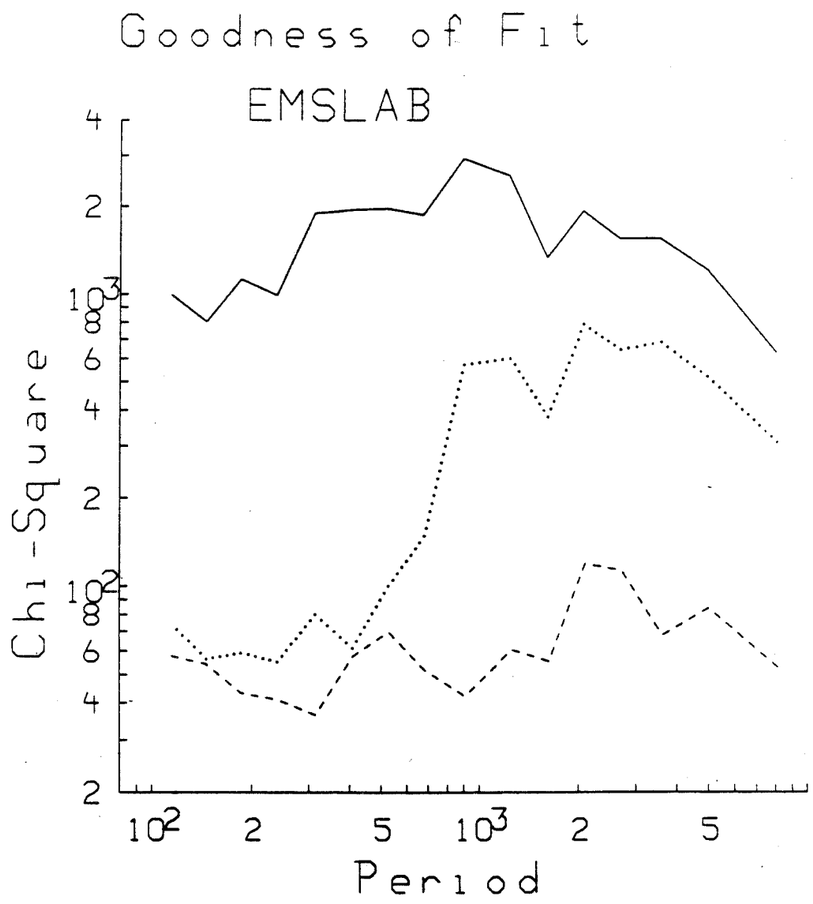


Figure 4.3(a) χ^2 goodness of fit statistics plotted vs. period: horizontal field models for EMSLAB array. 95% critical points (cp) for each model number is indicated on left edge of figure. (1) Solid line : isotropic errors model (df = 63; cp = 83). (2) Dotted line : model with local noise only (no gradients) (df = 44; cp = 60). (3) Dashed line : full model - local noise plus gradients (df = 40; cp = 56). Isotropic noise model is not adequate at any period. For periods greater than about 1000 seconds model fit is improved substantially by inclusion of gradient terms.

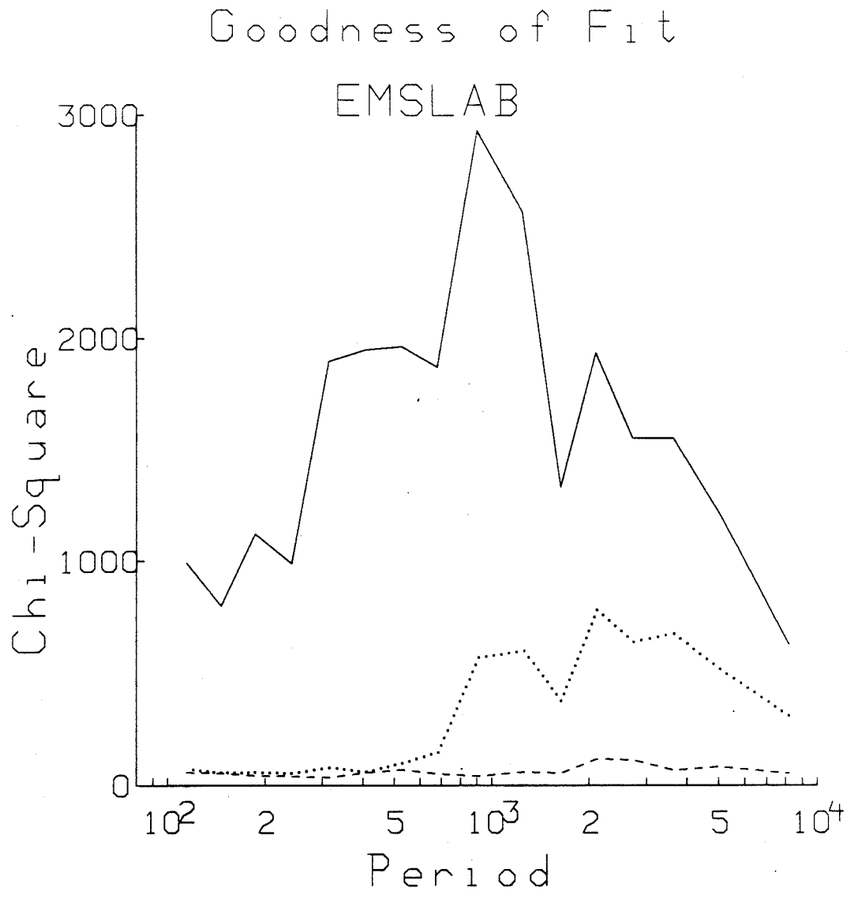


Figure 4.3(b) χ^2 goodness of fit statistics plotted vs. period, as in figure 4.3(a) but plotted on a linear scale. The improvement in fit offered by the full model is quite dramatic.

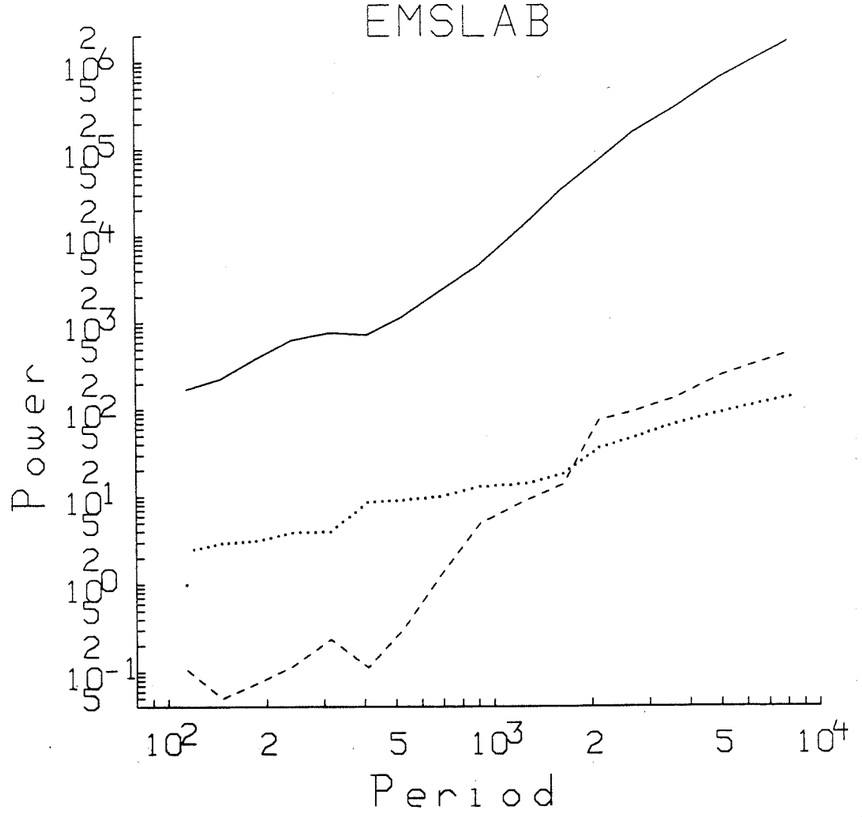


Figure 4.4: Signal and noise power spectra inferred from signal and noise parameter estimates for EMSLAB array. (1) Solid line: plane wave sources - $\delta_p^2(\omega)$; (2) Dashed line: gradients - $\delta_g^2(\omega)$; (3) Dotted line: local noise - $\delta_N^2(\omega)$. For periods above roughly 1000 seconds the power in the gradient vectors exceeds total local noise power. Note also that the local noise power spectrum is flatter than the plane wave and gradient signal spectra (which have roughly equal slopes).

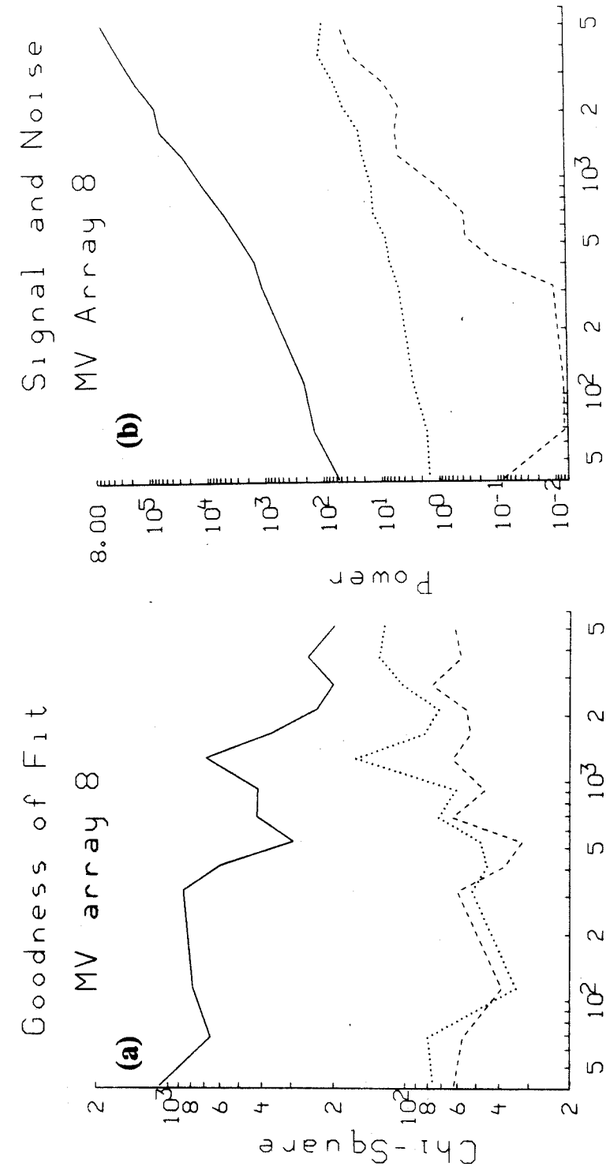


Figure 4.5(a) χ^2 goodness of fit statistics plotted vs. period: horizontal field models for array 8 (df is degrees of freedom, cp is 95% critical point). (1) Solid line : for isotropic errors model (df = 63, cp = 83); (2) Dotted line : model with local noise only (no gradients) (df = 44, cp = 60); (3) Dashed line : full model - local noise plus gradients (df = 43, cp = 59). The addition of the gradient terms does not substantially improve the model fit. Figure 4.5(b): Signal and noise power spectra inferred from signal and noise parameter estimates for array 8. (1) Solid line: plane wave sources - $\delta_p^2(\omega)$; (2) Dashed line: gradients - $\delta_g^2(\omega)$; (3) Dotted line: local noise - $\delta_N^2(\omega)$.

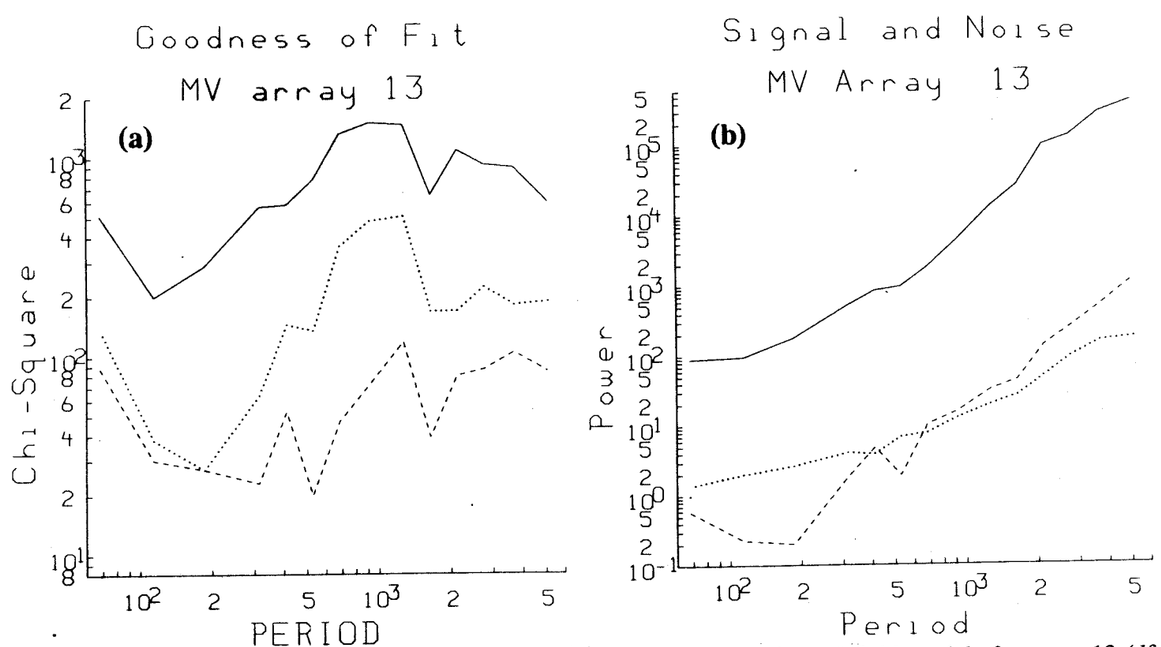


Figure 4.6(a): χ^2 goodness of fit statistics plotted vs. period: horizontal field models for array 13 (df is degrees of freedom, cp is 95% critical point). (1) Solid line : for isotropic errors model (df = 35, cp = 50); (2) Dotted line : model with local noise only (no gradients) (df = 20, cp = 31); (3) Dashed line : full model - local noise plus gradients (df = 16, cp = 20). Results are similar to those for the EMSLAB array: beyond about 1000 seconds gradient terms substantially improves the model fit. Figure 4.6(b): Signal and noise power spectra inferred from signal and noise parameter estimates for array 13. (1) Solid line: plane wave sources - $\hat{\sigma}_p^2(\omega)$; (2) Dashed line: gradients - $\hat{\sigma}_g^2(\omega)$; (3) Dotted line: local noise - $\hat{\sigma}_N^2(\omega)$.

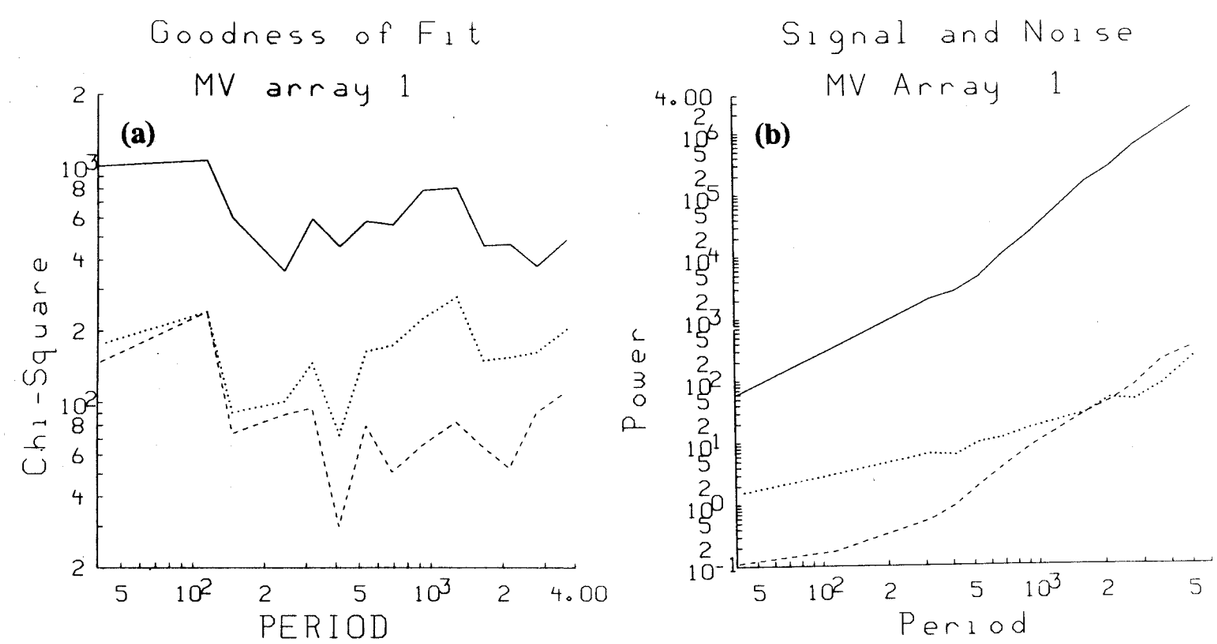


Figure 4.7(a) χ^2 goodness of fit statistics plotted vs. period: horizontal field models for array 1 (df is degrees of freedom, cp is 95% critical point). (1) Solid line : for isotropic errors model (df = 63, cp = 83); (2) Dotted line : model with local noise only (no gradients) (df = 44, cp = 60); (3) Dashed line : full model - local noise plus gradients (df = 40, cp = 56). Although the array size is similar to array 8, the gradient terms are substantially more important Figure 4.7(b): Signal and noise power spectra inferred from signal and noise parameter estimates for array 1. (1) Solid line: plane wave sources - $\hat{\sigma}_p^2(\omega)$; (2) Dashed line: gradients - $\hat{\sigma}_g^2(\omega)$; (3) Dotted line: local noise - $\hat{\sigma}_N^2(\omega)$.

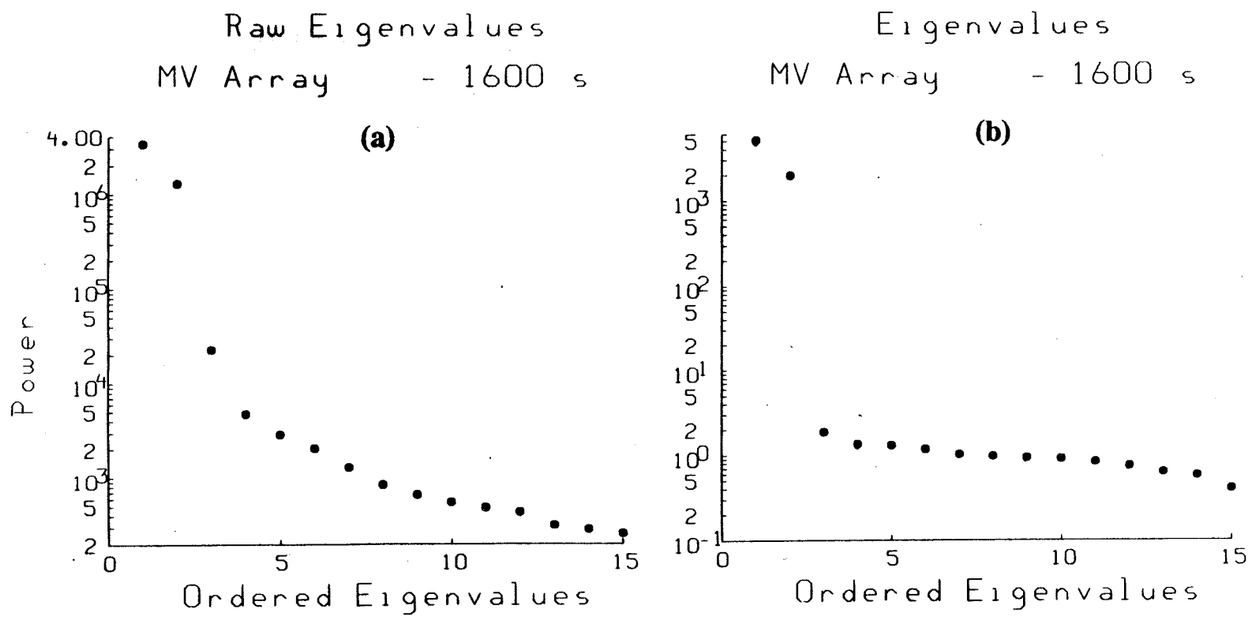


Figure 4.8: Ordered eigenvalues for 3 component data fit to local plus vertical field noise model (no gradients) for array 8, 1600 seconds. (a) Untransformed matrix **S** (b) Transformed matrix $\mathbf{S}' = \Sigma_N^{-1/2}\mathbf{S}\Sigma_N^{-1/2*}$. The eigenvalues of the transformed matrix are fairly constant indicating a qualitatively reasonable fit to the data.

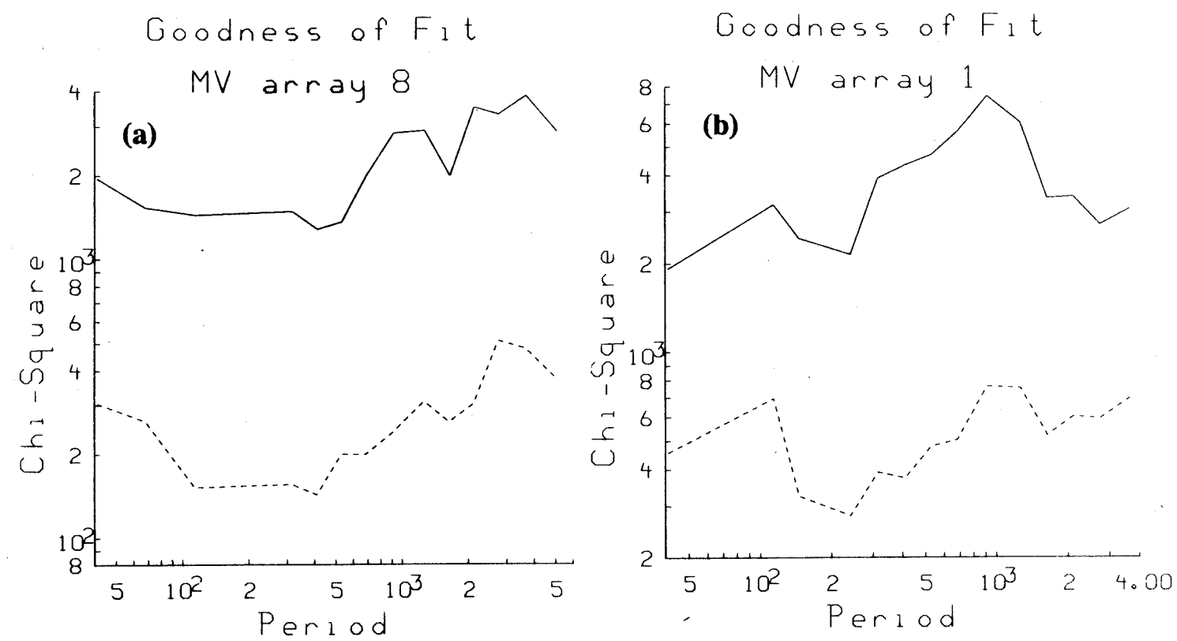


Figure 4.9: χ^2 goodness of fit statistics plotted vs. period: (a) Three component models for array 8. (b) Three component models for array 1. (df is degrees of freedom, cp is 95% critical point).
(1) solid line : isotropic errors model (df = 168, cp 199)
(2) dashed line : local noise plus vertical field source noise (no gradients) model (df = 132, cp = 160).

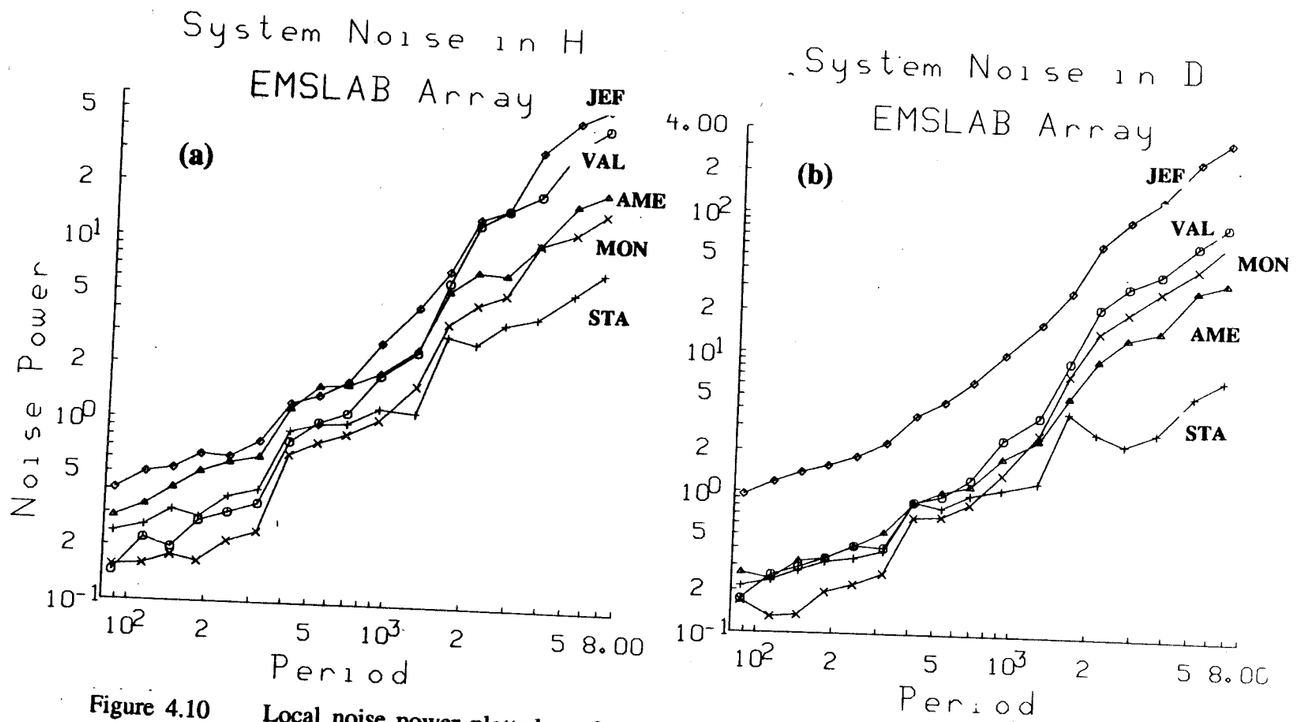


Figure 4.10 Local noise power plotted vs. frequency for each station in the EMSLAB array. (a) noise power in *H* component (geomagnetic north) (b) noise power in *D* component (geomagnetic east) The local noise power is quite variable with roughly an order of magnitude difference between the quietest and noisiest sites. Note that stations tend to be quiet or noisy over most of the frequency band analyzed here. Station JEF is by far the noisiest station in both components and at all frequencies.

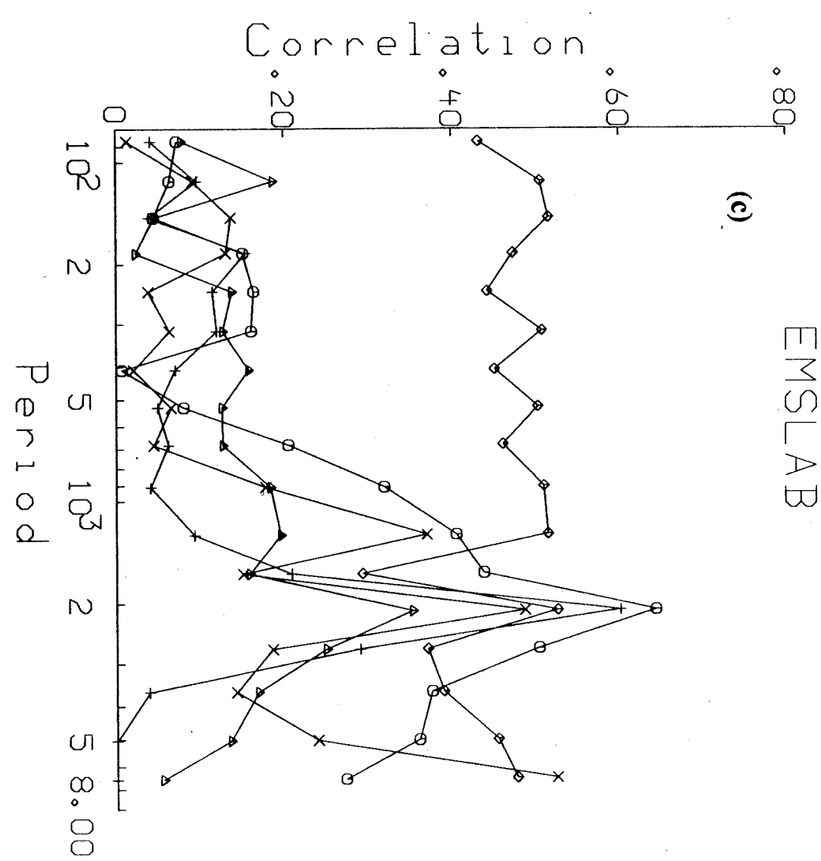


Figure 4.10(c) Magnitude of correlation between *H* and *D* components of local horizontal magnetic field noise for EMSLAB array. Note that the noisy station (JEF) has a large correlation between components.

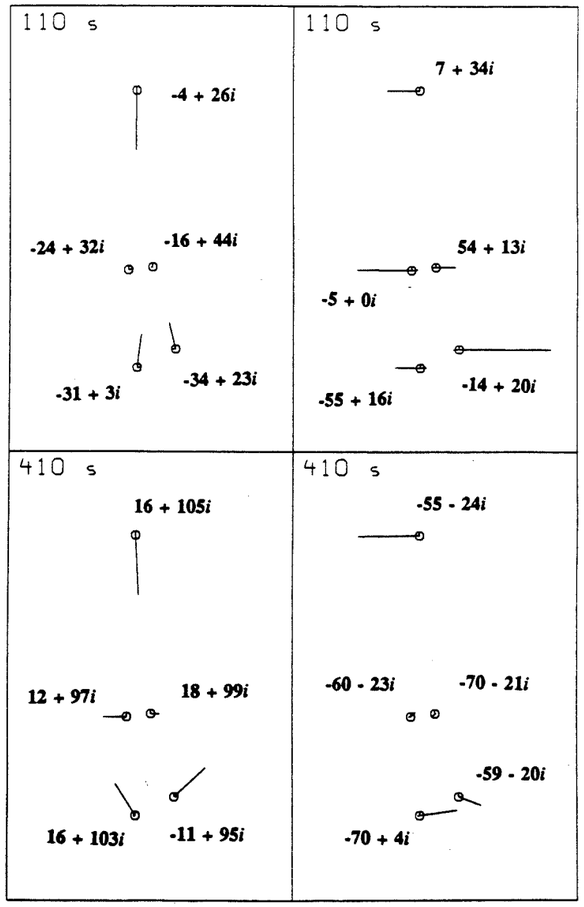


Figure 4.11(a) Vertical fields B_z correlated with observable horizontal gradients for array 1. Observable gradients are plotted as horizontal field vectors at each station; correlated vertical fields at each station (real and imaginary parts) are printed under each station. The array is essentially linear so only two gradient terms can be reliably estimated for each period.

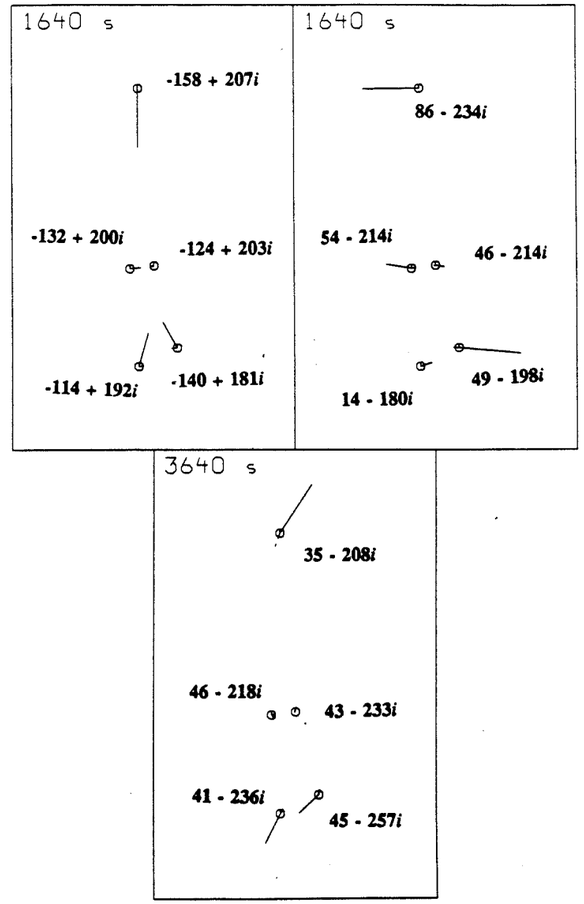


Figure 4.11(b) Vertical fields B_z correlated with observable horizontal gradients for array 1. Only a single gradient term is included for the longest period. Note that the vertical fields become more uniform at longer periods.

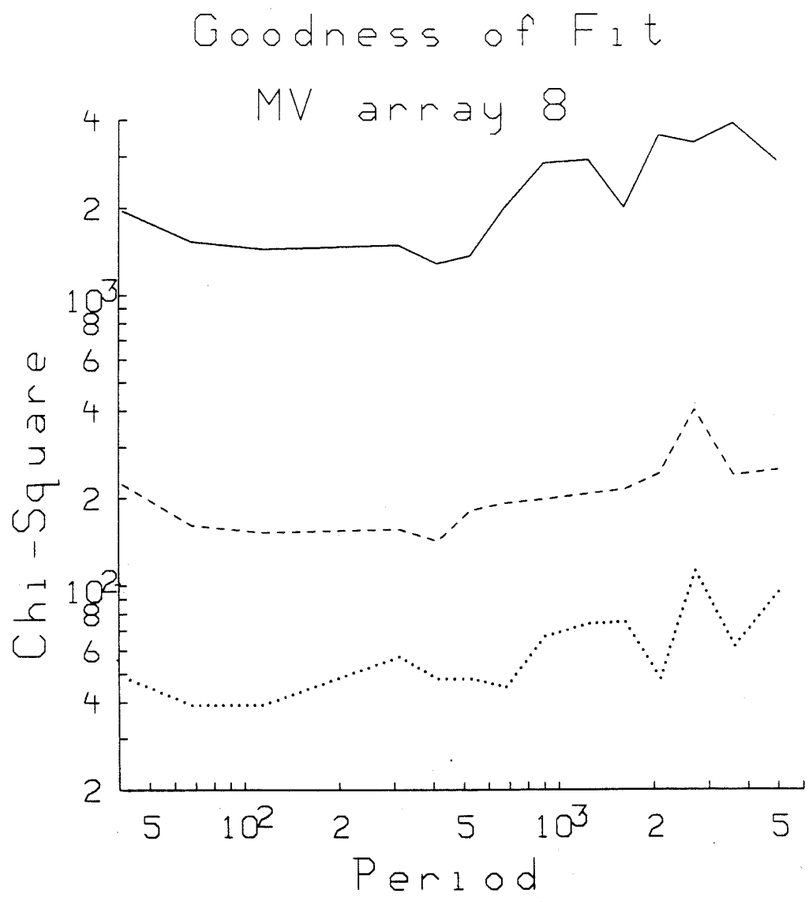


Figure 4.12(a) χ^2 goodness of fit statistics plotted vs. period for three component models for MV array 8. (df is degrees of freedom, cp is 95% critical point). (1) Solid line : Model 1 - isotropic errors model (df = 168, cp = 199); (2) Dotted line : Model 2 (see text) (df = 113, cp = 139); (3) Dashed line : Model 3 : 5-dimensional response space plus local noise (df = 55, cp = 73).

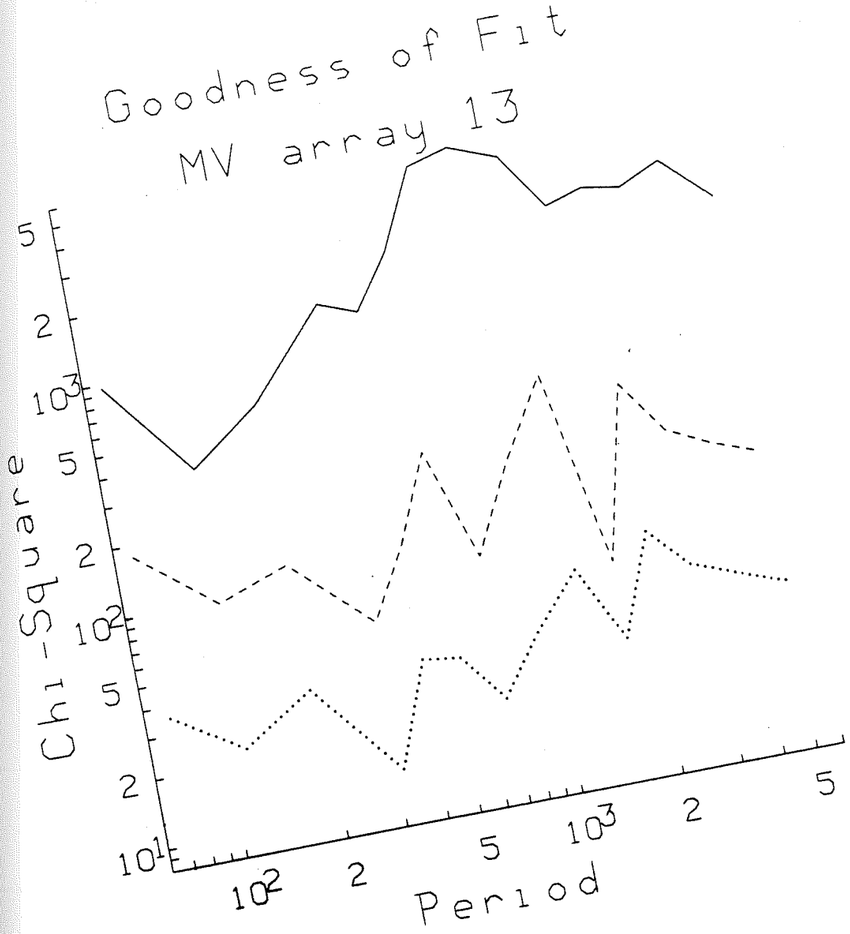


Figure 4.12(b) χ^2 goodness of fit statistics plotted vs. period for three component models for MV array 13. (df is degrees of freedom, cp is 95% critical point). (1) Solid line : Model 1 - isotropic errors model (df = 99, cp = 123); (2) Dotted line : Model 2 (see text) (df = 43, cp = 58); (3) Dashed line : Model 3 : 5-dimensional response space plus local noise (df = 13, cp = 22).

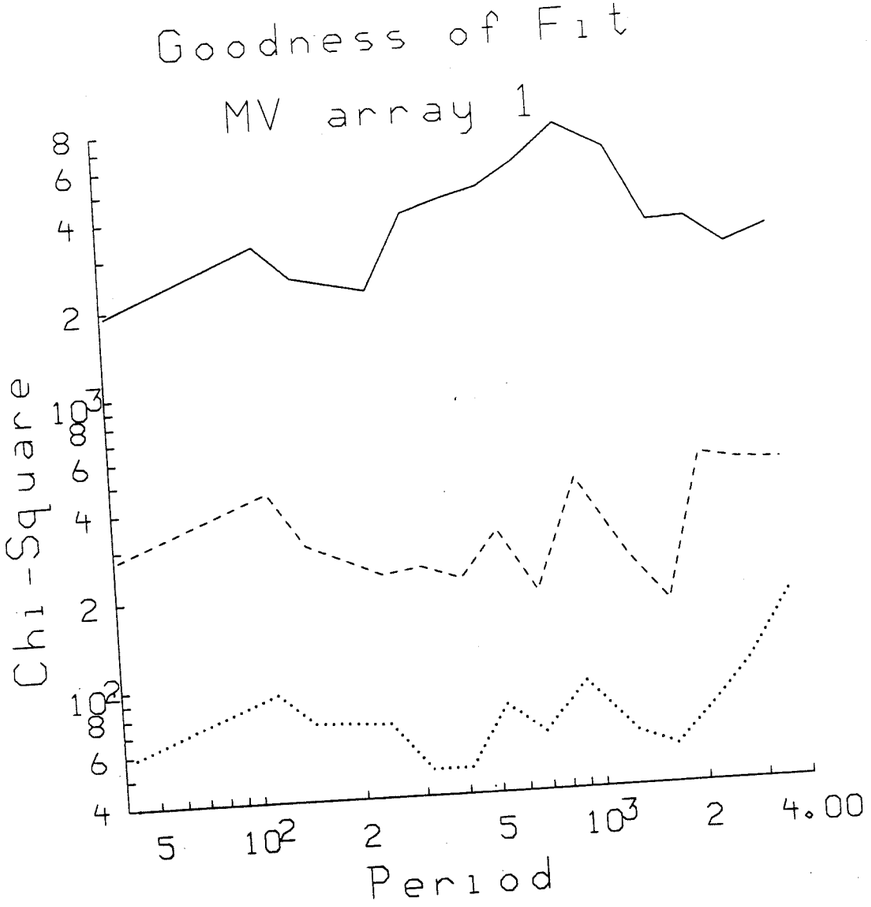


Figure 4.12(c) χ^2 goodness of fit statistics plotted vs. period for three component models for MV array 1. (df is degrees of freedom, cp is 95% critical point). (1) Solid line : Model 1 - isotropic errors model (df = 168, cp = 200); (2) Dotted line : Model 2 (see text) (df = 100, cp = 124); (3) Dashed line : Model 3 : 5-dimensional response space plus local noise (df = 55, cp = 77).

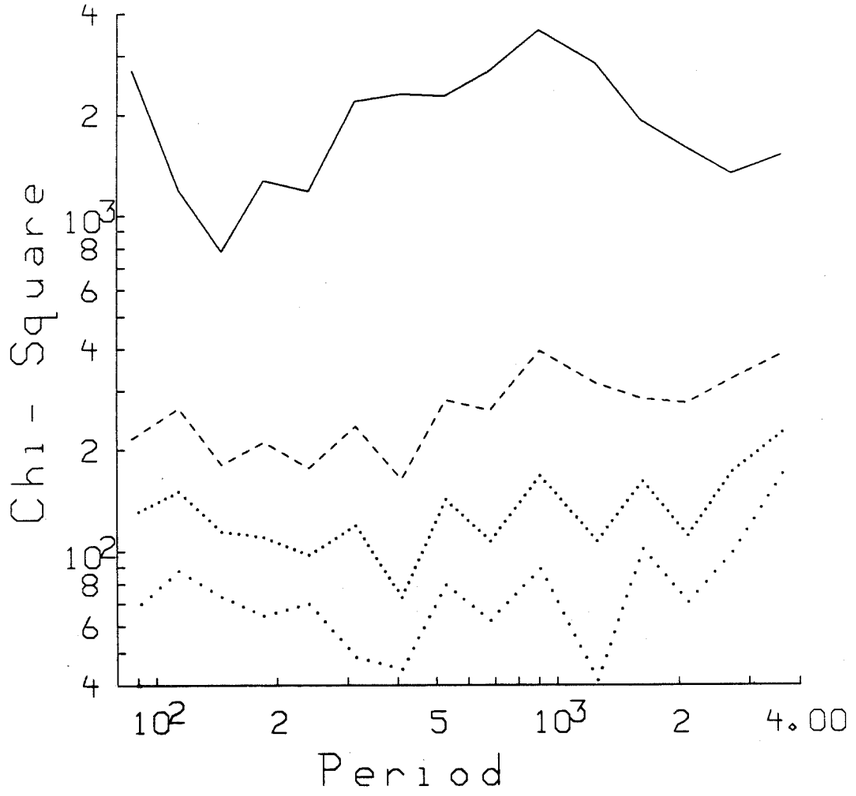


Figure 4.13 χ^2 goodness of fit statistics plotted vs. period for three component models for MV array 1 (df is degrees of freedom, cp is 95% critical point). (1) Solid line : 2-dimensional response space plus local noise; (df = 124, cp = 151); (2) Dashed line : 3-dimensional response space plus local noise; (df = 99, cp = 123); (3) Dashed line : Model 3 : 4-dimensional response space plus local noise (df = 76, cp = 97). (3) Dashed line : Model 3 : 5-dimensional response space plus local noise (df = 55, cp = 73).

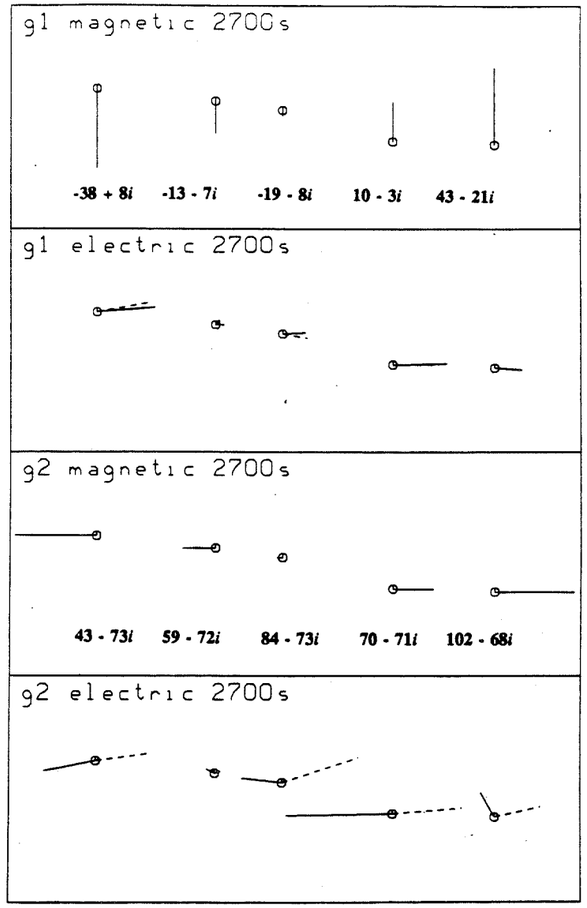


Figure 4.14 Electric field vectors which are correlated with the gradients of the horizontal magnetic fields for period of 2700 seconds. From top: (a) Plot of magnetic field gradient in north-south component. Correlated vertical components for each station is printed under station. (b) Plot of electric fields correlated with the gradient plotted in (a). (c) Plot of magnetic field gradient in east-west component. Correlated vertical components for each station is printed under station. (d) Plot of electric fields correlated with the gradient plotted in (c). The electric field vectors point in the same direction and are in phase across the array; note difference in amplitudes however. Vertical fields are roughly equal and in phase; some deviations from this simple picture are evident.

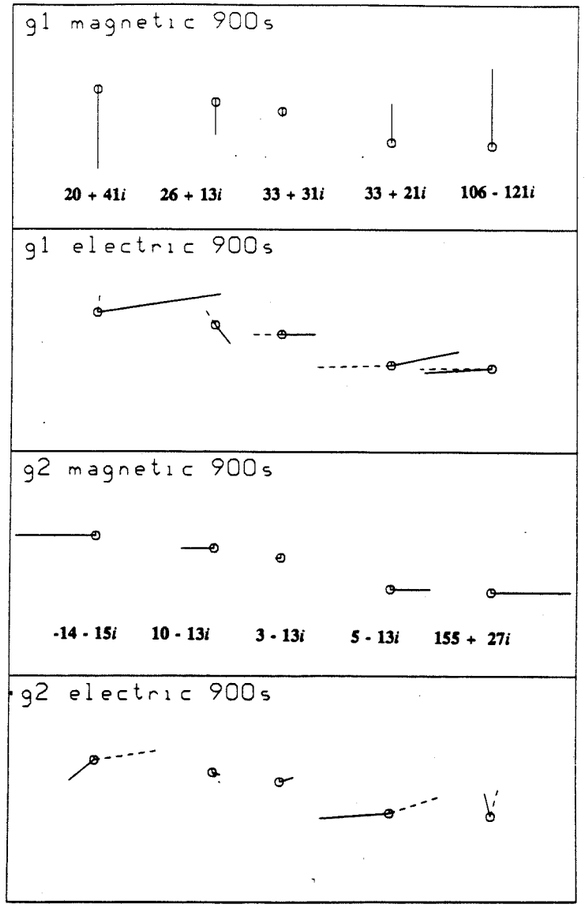


Figure 4.15 Electric field vectors which are correlated with the gradients of the horizontal magnetic fields for period of 900 seconds. From top: (a) Plot of magnetic field gradient in north-south component. Correlated vertical components for each station is printed under station. (b) Plot of electric fields correlated with the gradient plotted in (a). (c) Plot of magnetic field gradient in east-west component. Correlated vertical components for each station is printed under station. (d) Plot of electric fields correlated with the gradient plotted in (c). In contrast to results for 2700 seconds, the electric field vectors exhibit quite a bit of scatter, both in direction and in phase.

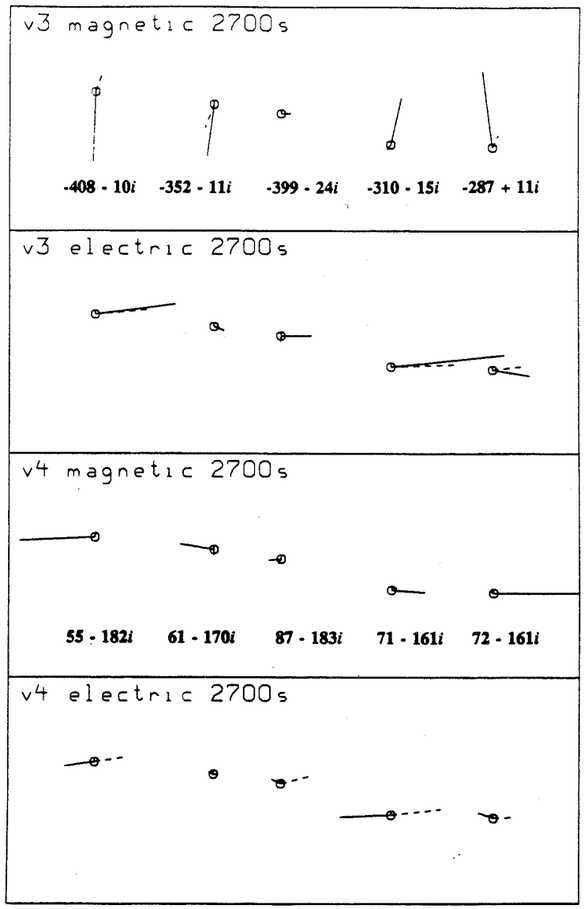


Figure 4.16 Electric field vectors which are correlated with the gradients of the horizontal magnetic fields for period of 2700 seconds. For this figure the gradients and their effects on electric and vertical fields have been treated as signal and a higher dimensional response space has been fit. Linear combinations of the response space vectors whose horizontal components best fit the gradient model are plotted here. From top: (a) Plot of best fit of magnetic field to gradient in north-south component. Correlated vertical components for each station is printed under station. (b) Plot of electric fields correlated with the gradient plotted in (a). (c) Plot of best fit of magnetic field gradient in east-west component. Correlated vertical components for each station is printed under station. (d) Plot of electric fields correlated with the gradient plotted in (c). Results are similar to those presented in figure 4.14. Note that the vertical fields have a much larger amplitude here.

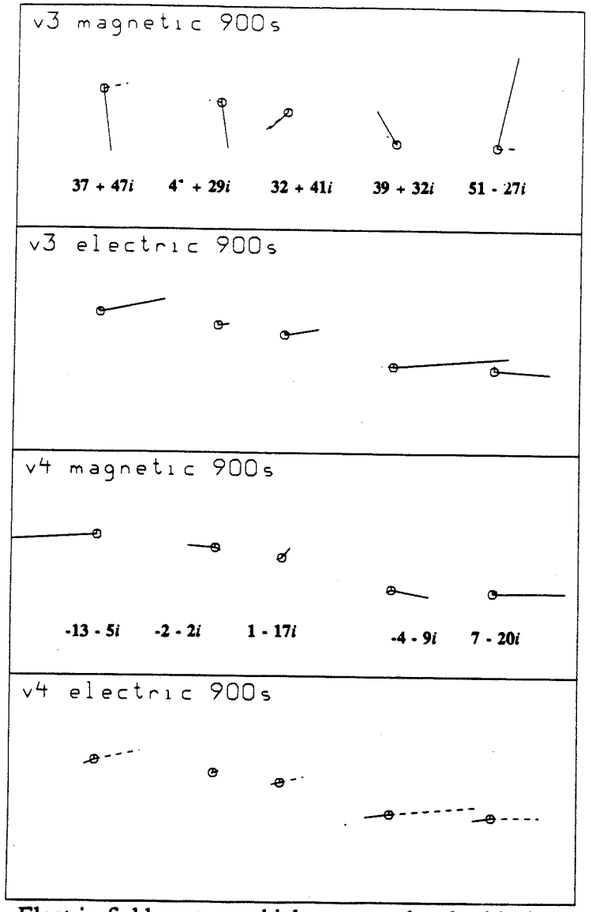


Figure 4.17 Electric field vectors which are correlated with the gradients of the horizontal magnetic fields for period of 900 seconds. For this figure the gradients and their effects on electric and vertical fields have been treated as signal and a higher dimensional response space has been fit. Linear combinations of the response space vectors whose horizontal components best fit the gradient model are plotted here. From top: (a) Plot of best fit to magnetic field gradient in north-south component. Correlated vertical components for each station is printed under station. (b) Plot of electric fields correlated with the gradient plotted in (a). (c) Plot of best fit to magnetic field gradient in east-west component. Correlated vertical components for each station is printed under station. (d) Plot of electric fields correlated with the gradient plotted in (c). In contrast to results presented in figure 4.15, the electric field vectors point in the same direction and are in phase across the array. Results are quite comparable to those obtained for 2700 seconds.

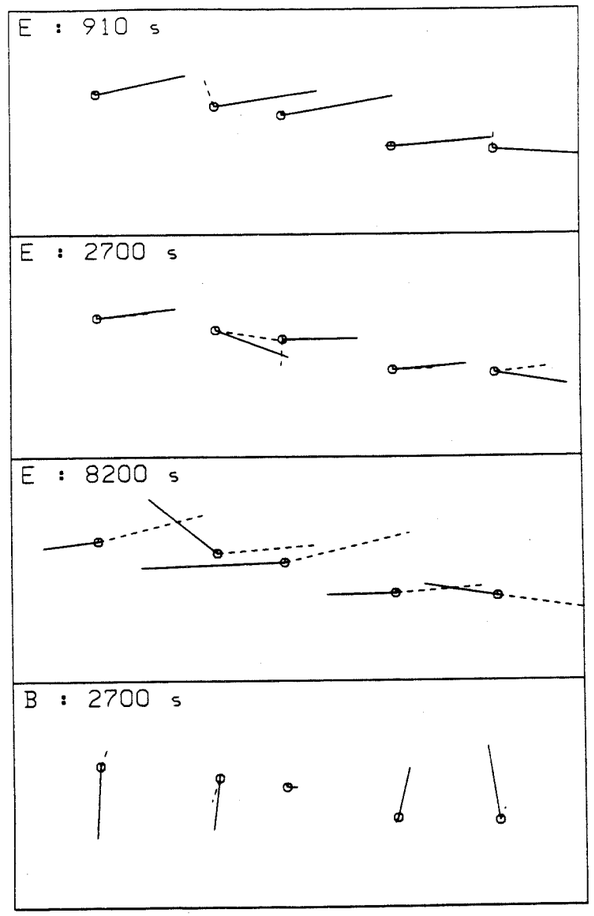


Figure 4.18 Electric fields which are correlated with gradients in magnetic fields. Here the electric fields have been rescaled by dividing by the magnitude of the plane wave impedance for the east-west component of the electric fields. This is a crude correction for surface distortion. Now the electric field vectors correlated with the magnetic gradients are of similar size for all stations in the array. From top: (a) 900 seconds; (b) 2700 seconds; (c) 8200 seconds; (d) corresponding magnetic field gradients for 2700 seconds.

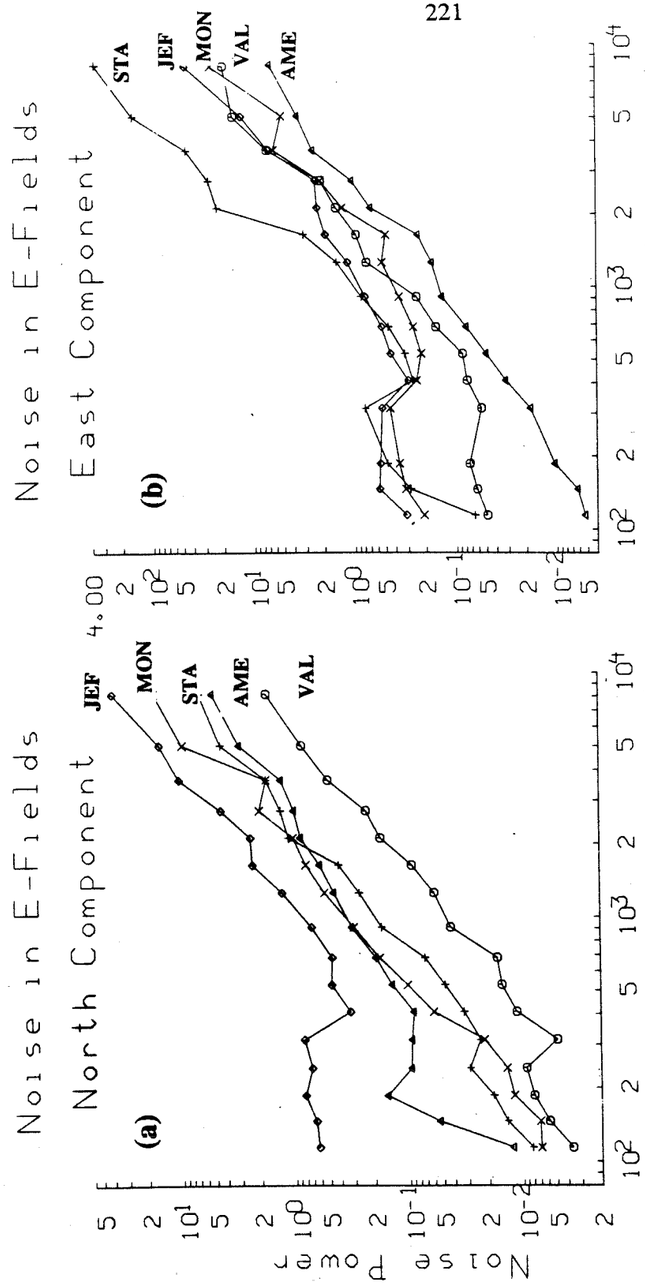


Figure 4.19 Estimated local noise power in electric fields for EMSLAB line. (a) North component; (b) East component. Noise power is quite variable, particularly at shorter periods. As with the magnetic field noise, individual components tend to be quiet or noisy over a wide range of frequencies. Note that the east component is larger at most stations for the longer periods plotted here. This probably reflects an inadequate model for source effects which are stronger in the east-west directions; some of the noise included in the local noise estimates given here is probably coherent between stations.

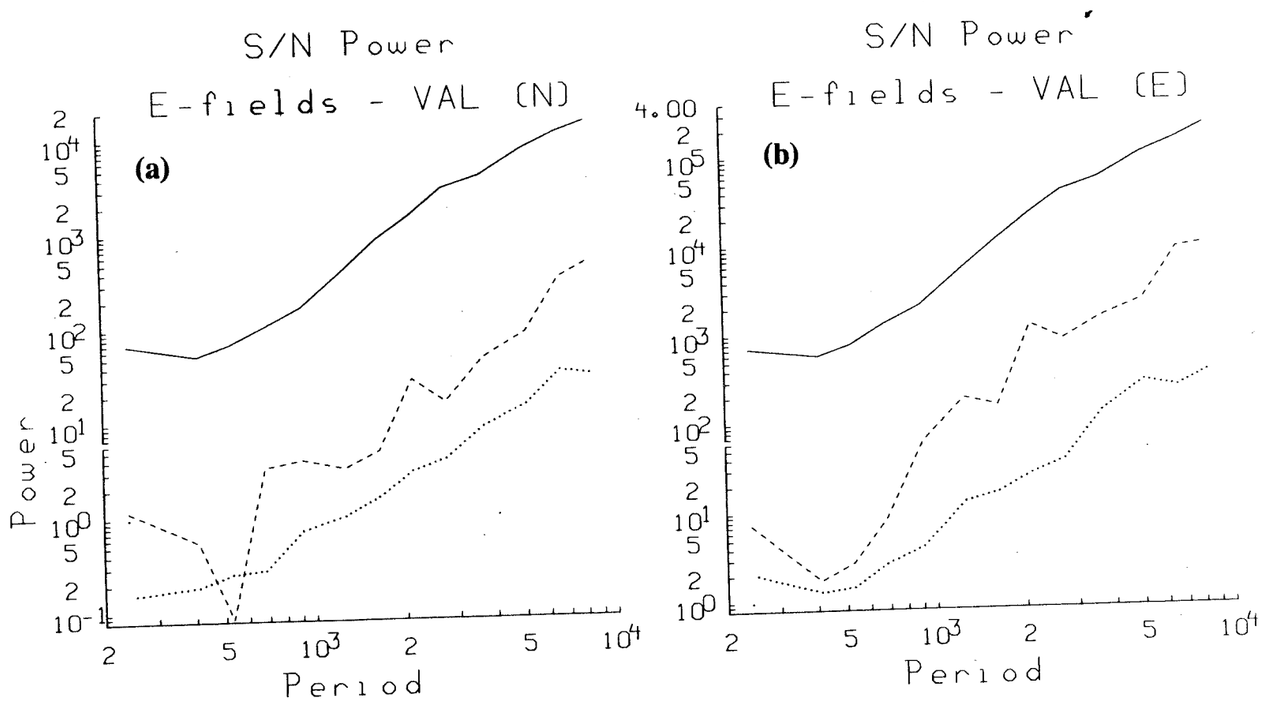


Figure 4.20 Signal and noise power for the electric fields at station VAL in the EMSLAB line for (a) North component; (b) East component. (1) Solid line is estimated plane wave related power; (2) Dashed line is estimated gradient related power; (3) Dotted line is estimated local noise power. At this station gradient related power is generally larger than local noise power, particularly in the north component.

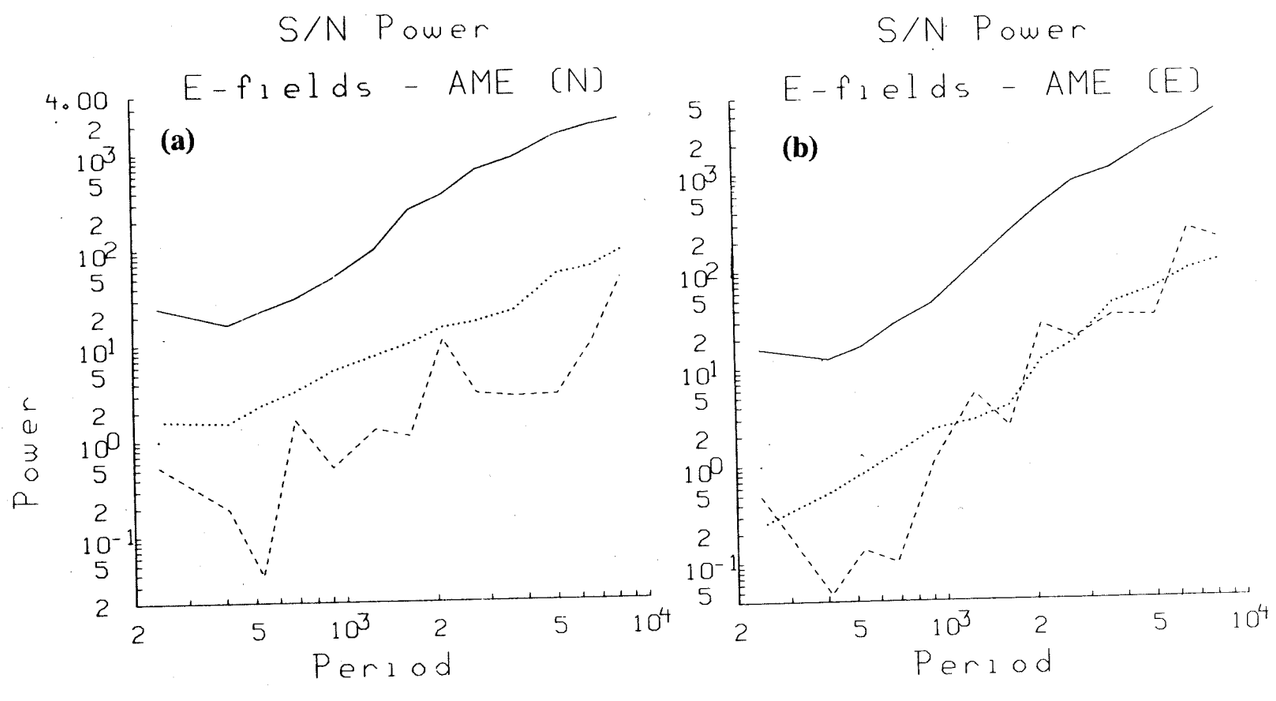


Figure 4.21 Signal and noise power for the electric fields at station AME in the EMSLAB line for (a) North component; (b) East component. (1) Solid line is estimated plane wave related power; (2) Dashed line is estimated gradient related power; (3) Dotted line is estimated local noise power. For this station the local noise is relatively more important than it is at VAL (figure 4.20). In the east component local noise is larger than gradient related noise.

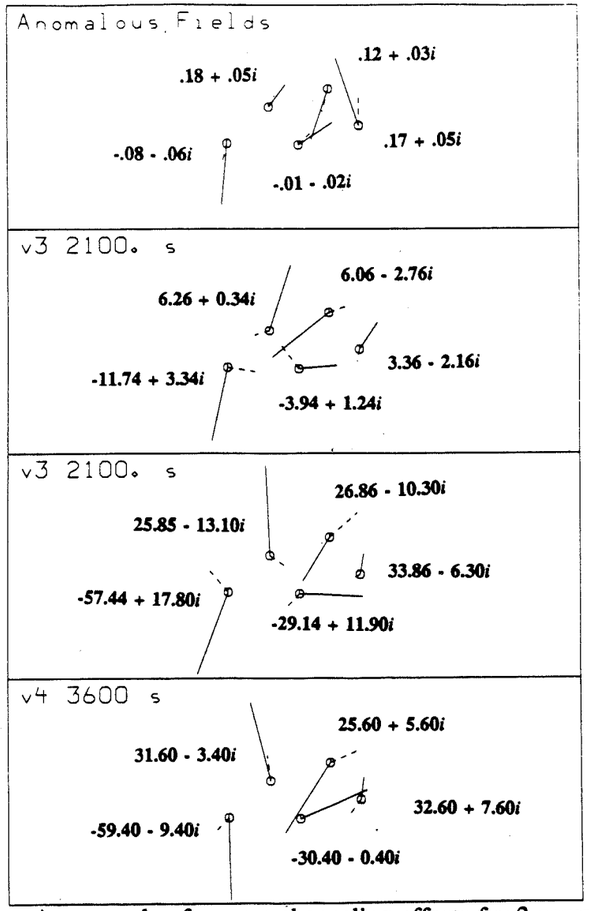


Figure 4.22 An example of current channeling effects for 3-component MV data from array 8. In (a) we plot the anomalous horizontal and vertical magnetic fields for a north-south inducing magnetic field. Normal fields are defined by requiring the average field to be linearly polarized north-south and of unit magnitude. The normal field has been subtracted, so the fields plotted here average to zero. In (b),(c) and (d) we plot deviations of the horizontal fields for the third eigenvector from the best fitting gradient for three periods. Deviations of the vertical components of the third eigenvector from the average vertical field are printed under the station locations. The deviations from the gradient form for the three periods are very similar to each other and to the anomalous horizontal fields plotted in (a). We interpret this as a current channeling effect. Currents, which are induced by non-local sources, and are hence not coherent with the local horizontal fields, flow in a conductive anomaly running through the center of the array.

Chapter 5

Extensions and Applications

In this dissertation we have emphasized the development of, and formal justification for, some new techniques for the analysis of geomagnetic array data. In this final chapter we provide a summary and overview of these techniques, and then discuss very briefly some extensions and applications.

For all of the data processing methods discussed in this dissertation, the fundamental data vectors consist of the complex Fourier coefficients obtained from Fourier transforming short segments of a long multi-channel time series of magnetic and electric fields measured simultaneously at an array of n stations. These data vectors are complex, and of dimension $3n$ (magnetic fields only) or $5n$ (electric plus magnetic fields). The methods developed here can be classified as (complex) multivariate statistical procedures, which are based on the sample spectral density matrix (the Hermetian matrix of second moments of the frequency domain data vectors). As such, these methods are generally complex analogues of classical multivariate procedures. Because multivariate methods are based on a simultaneous analysis of all data channels, considerably more information can be obtained from the data than is possible with univariate (e.g. standard transfer function) methods. As we have seen, the multivariate approach also allows for a reasonable treatment of noise. Our models have included (potentially correlated) noise in all data channels.

The statistical methods considered in this dissertation are of two fairly distinct types - exploratory procedures which are designed to discover what sorts of useful information may be extracted from the data and formal model fitting procedures which assume a specific, parametrized model for the data and then attempt to make inferences about the unknown parameters. Although it is not always

possible to make a clean seperation between procedures of the two types, the distinction is nonetheless important. The primary exploratory method that we have used is essentially the method of principal components - the anlysis of the dominant eigenvalues and eigenvectors of the SDM. We considered a number of applications of this basic method, as well as some variants, in section 2.5. On the other side of the dichotomy lie the more formal model fitting procedures discussed in Chapter 4, where we assumed a specific parametric form for signal and noise components, and developed statistical estimation procedures. Both of these approaches have an important place in geomagnetic array data processing.

The results obtained in section 2.5 amply demonstrated the value of a principal components analysis of the SDM for geomagnetic array data. Consideration of the dominant eigenvalues and eigenvectors of the SDM allowed us to characterize many of the basic properties of both signal and noise processes. The random source model developed in Chapter 3 offered a simple physical interpretation for the general patterns observed in this exploratory analysis of the eigenvalues and eigenvectors of the SDM. More importantly, the qualitative analyisis of general random source models (discussed in section 3.4), showed that we may consider the eigenvalues of the SDM as a sort of discrete spatial power spectrum for the random fields. The corresponding eigenvectors represent realizeable fields, and in general we would expect that the dominant eigenvectors (principal components) will correspond to relativley large scale external source fields with simple morphologies. These eigenvectors will also be the least effected by noise. The extraction of the dominant eigenvectors can thus be thought of as a way to smooth, or average, a number of events to effectively reduce the effects of small scale source field variations and local noise. These conclusions do not depend in any way on a particular assumption about the form of the sources. An eigenvector analysis of the SDM for any array (global or regional) should yield interesting results about both

the large scale source fields and the (usually much smaller spatial scale) variations in electrical conductivity.

The theoretical developments in Chapter 2 showed that interstation and intercomponent transfer functions could be justified with complete rigor if the external source potentials were restricted to a space of finite dimension p. With this finite dimensional source space assumption, we showed that, in the absence of noise, all data vectors would be linear combinations of p fundamental basis vectors which defined the total fields seen at the earths surface for the p possible independent sources. For the specific case of small arrays, where both the array and the skin depth are small compared to typical source scales, the model of Chapter 3, and the exploratory results of sections 2.5, demonstrated that it was reasonable to assume that the external sources were plane waves of infinite horizontal extent ($p = 2$). The finite dimensional source space assumption suggests a formal statistical model - the errors-in-variables model of Chapter 4. To complete the specification of this model it is necessary to parametrize the noise covariance matrix. We discussed an approach to this problem for the special case of small arrays in Chapter 4, and we found that it was necessary to include the effects of violations of our simplifying source assumptions in the noise. Within the assumptions of this model, it is possible to estimate the response of the earth to plane wave sources, and to assess the statistical uncertainty in the estimates, including the effects of deviations from the simplifying source assumptions.

In fact, the exploratory and formal model fitting procedures are not completely seperate. The justification for the plane wave source model depends in part on the exploratory analysis. More importantly, the parametrization of the noise model depends critically on the exploratory phase. On the other hand, exploratory techniques can be enhanced by a formal model fitting procedure. For example, many of the results discussed in Chapter 4 would have to be considered

exploratory. In particular, our findings about gradient related effects in the electric fields are of this nature. Although we were able to demonstrate the presence of coherent noise in the electric fields using the purely exploratory techniques of section 2.5, the results in Chapter 4 are much more satisfactory. Using the model fitting methods of Chapter 4 the nature of the coherent noise was delineated, and the relation of this effect to gradients in the magnetic fields was demonstrated.

While we have tried to keep our discussion of geomagnetic array analysis as general as possible, all of the real arrays we have analyzed are small (both in terms of number of stations and in terms of spatial extent). The case of small arrays is an important one; indeed it was this case which provided the motivation for this study. Both the random source model developed in Chapter 3, and the statistical models for the noise covariance developed in Chapter 4 were geared to this special case. We summarize the results obtained for this special case and give a very brief indication of some holes which need to be filled in.

The synthetic models of Chapter 3, together with the exploratory results of section 2.5 (and Chapter 4), provide a strong justification for the plane wave source approximation. A first order correction to this approximation can be made by including a set of three gradient terms in the model. We have seen that for arrays at geomagnetic mid-lattitudes with spatial extents on the order of 50-200 *Km* that the horizontal magnetic field gradients are small compared to the uniform part of the signal. For EDA fluxgate magnetometers at least, we have also seen that the power in gradient fields is often barely detectable above the local/system noise, particularly at shorter periods (< 1000 seconds). At this point it is not clear how much of the local noise in the magnetic fields is cultural and how much is true system noise, so it is not clear how much this last result depends upon the specific magnetometers used.

For the vertical magnetic and electric fields associated with the gradients the situation is somewhat different. The magnitude of these effects is independent of array size (but dependent on period) and, at least for longer periods, these effects can be substantially larger than the local/system noise levels. If the gradients can be adequately measured (in all directions), the relationship between the gradients and the associated vertical magnetic and/or electirc fields should be considered a potentially very interesting part of the signal. For our situation where gradients are not adequately determined, however, we have argued that it is most reasonable to concentrate on estimation of the plane wave response vectors and treat the horizontal field gradients, as well as the vertical magnetic and electric fields associated with these gradients, as a source of coherent noise. We have developed a parametric model which includes this gradient related noise, and have developed an algorithm for parameter estimation. We have seen that the models generally provide a reasonable fit to the data, at least for 3-component (MV) data.

Our applications of these statistical modeling techniques to small geomagnetic arrays is only a beginning. The modelling procedures we have suggested, while generally succesful, may be unnecessarily complicated. It is entirely possible, for instance, that the noise model of (4.5.7) which includes local (incoherent between stations) noise and a simple form for source gradient related vertical fields, may be completely adequate for MV data in small arrays. We need to carefully compare estimates and estimation errors for various approaches to assess what complications are truly relevant to routine processing of small array data. We have much further to go for the case of MT arrays. We have identified gradient related coherent electric fields in the residuals to the plane wave fit, but our understanding of this effect is, at best, rudimentary. Further study of the EMSLAB array (with the aid of the much larger magnetometer array) will be helpful. Comparison to other small MT arrays in different settings (e.g. not so

close to a coastline) would also be helpful. With a better understanding of these effects, we can attack the problem of specifying the structure of this source related noise in the electric fields.

In addition to the problem of model specification, there are other aspects of the problem which need further work. All of the methods developed here are based on second moments of the data. They are thus appropriate when the data has a Gaussian distribution. We already know (Egbert and Booker, 1986) that the data can not always be treated as Gaussian; for single station transfer function estimation we have found that robust processing techniques are essential. The development of robust variants on the estimation procedures described here will also be necessary. Another area which needs work is in the computation of estimation errors. The estimation errors given in Chapter 4 are only strictly correct for the case where the noise covariance matrix is known (up to a multiplictive constant). When the error covariance must be estimated some modifications will be required.

There are two further extensions of our work which we wish to consider briefly. The first, which we have already alluded to several times, is to try these techniques on larger arrays - in particular the global array of magnetic observatories and the full EMSLAB array. In both cases many of the specific methods developed in this dissertation will not be particularly relevant. However, as discussed above, the exploratory aspects of our approach can easily be justified, and should prove interesting. In the case of the global array, Schultz and Larsen (1987) have applied transfer function methods to the analysis of observatory daily mean values. This required assuming a specific form for the external sources (the potential was assumed to be the P_1^0 spherical harmonic, see Schultz and Larsen, 1984). If this assumption holds, the response space should be one dimensional, and a multiple station estimation of the P_1^0 response should be straightforward. If

this assumption does not hold exactly, the manner in which it fails should be revealed by a multiple station analysis. If, as seems likely, more complicated models of the source are required our multiple station technique should prove flexible enough to allow for this.

Application of these techniques to large regional arrays (such as the full EMSLAB array) should also prove fruitful. For such a large array (but note: the EMSLAB array is still substantially smaller than typical source length scales) the gradient terms will be quite significant, and they will probably be mixed with the plane wave terms. Additional, source related shorter wavelength features may also appear in the eigenvectors of the SDM. Sorting all of this out may prove to be a significant task, but the dominant few eigenvectors will still be very interesting. With present methods of analysis the large scale overview of a regional array like this is provided by one or maps of specific events (e.g. The EMSLAB Group, 1987). Plots of the dominant eigenvectors of the SDM (or rather of linear combinations chosen to provide simple source polariaztions) would certainly serve this purpose in a far better manner. These will correspond to external sources of the largest possible scale, which represent averages of many events. Such maps will thus combine the advantages of transfer function approaches with the advantages of traditional event maps.

The second extension which we wish to consider is to the case where a number of small overlapping arrays have been run. If two arrays have a station in common the plane wave response spaces estimated for each array seperately can be combined to estimate a plane wave response space for a combined array. This can be accomplished simply by treating the overlapping station as 'normal' and computing the fields in each array relative to the common station. We give an example of this for two overlapping arrays in southwetern Washington in figure 5.1 (array 8, which we have used for some of the examples discussed in Chapter

4, and array 10, a more or less north-south line of five stations; see inset fro array location). Note that the common station (the western station in the two arrays) need be treated as normal only for purposes of forming the combined response space; after the arrays are combined any other definition of normal fields may be used. In figure 5.1 we have chosen the southernmost station as normal for purposes of defining the anomalous fields, which are plotted here for a north-south source polarization at a period of 300 seconds. The anomalous field plots show an enhancement of horizontal fields in the center of the array and a reversal of sign of the vertical fields in the same region. This is consistent with a concentration of anomalous currents in the center of the array. We have already encountered this conductivity anomaly in our discussion of current channelling in Chapter 4. Note that there is a substantial turning of the currents to the northwest in the center of the array. This is presumably due to deflection by a region of high resistivity in the crust in the southwestern part of the region plotted. This region coincides with a high in the Bouger gravity, which is also contoured in figure 5.1.

This simple approach to combining arrays through a common reference station has been described (with transfer function terminology) by Beamish and Banks (1983) who used this technique to form a synthetic large array from a series of smaller arrays run simultaneously with a single fixed reference station. We wish to consider here the more general case where a number of small arrays overlap without any fixed reference station. This is the case for the set of arrays plotted in figure 2.1 - all arrays were overlapped with other arrays and the overlaps are such that all stations can be linked together, but there is no single reference station. The simple approach of using the overlapping station as a reference can also be used for this case. We give examples of the horizontal fields computed using this approach for a 41 station synthetic array in figures 5.2 (east-west polarization) and 5.3 (north-south polarization) for a period of 900 seconds. For

these plots the normal field has been taken as the regional average, using the constraint discussed in section 2.4. Assuming that for the normal fields, half of the fields are internal and half external, we have subtracted half of the assumed normal field. The plotted field vectors thus represent an approximation to the total internal fields. For comparison we also present, in figures 5.4 and 5.5, 'hypothetical events' (Bailey *et al.*, 1974) computed from a series of single station vertical field transfer function estimates for the two polarizations for the same array.

In addition to the coast effect (clearly visible in the vertical field plots of figure 5.3), there are two major anomalous features in this region, both of crustal origin. The most significant is a conductivity anomaly which goes east-southeast out of Puget Sound for roughly 25 kilometers, and then south-southeast under the Cascade range. This anomaly is cleraly evident in the hypothetical event plots (for both polarizations) as a reversal in sign of the fairly large anomalous fields in this region. The anomaly is also quite evident in the horizontal field plots constructed from the multiple station analysis. At some stations in the center of the anomaly the fields are more than double the normal (average) value. Since the anomaly is large in both polarizations, a significant fraction of the anomalous currents must be channeled by the 3-dimensional geometry of these crustal conductivity variations. Note also that the deviations in direction of the horizontal field vectors are all consistent with current being channeled through a narrow region of high conductivity under the Cascades. This anomaly has been reported previously by Law *et al.*, 1980 and Stanley (1984). The second major feature is the one discussed above (figure 5.1) and in Chapter 4 (current channeling in array 8). This feature consists of an essentially east-west conductive anomaly which is clearly evident in the vertical field plots for the north-south polarization (figure 5.5). The fields (particulary the horizontal fields) for this anomaly are considerably smaller than those of the Cascade anomaly to the east. A more

comprehensive analysis of the anomalies in the region covered by this array will, in the near future, be the subject of a more thorough application of the methods discussed in this dissertation.

With a complex series of arrays with no single reference station there may be many ways to connect the small arrays together using the simple overlapping reference station method. In this case the final answer depends upon the choices of reference stations, Furthermore, such simple methods use only a fraction of the available information for combining arrays. We close this dissertation with a sketch of an approach to combining arrays which does not suffer from these defects. The method arises naturally from our response space approach and is based on a direct generalization of the single station estimation criterion given in (4.1.2). We consider K overlapping arrays with array k consisting of m_k total channels of data, sampled in N_k windows. The total number of component-stations is M (i.e. for 3-component data measured at a total of n station distributed over a number of arrays $M = 3n$). Let Ξ_k denote the $m_k \times N_k$ data matrix for the k^{th} array, and let $\hat{\Sigma}_k$ be the estimated error covariance for the k^{th} array. Let $\mathbf{V}$ be the $M \times 2$ matrix whose columns are the plane wave response vectors for the M station synthetic array, and let $\mathbf{P}_k$ be the $m_k \times M$ matrix which projects the response vectors of $\mathbf{V}$ onto the coordinates of the components actually measured in the k^{th} array (so that the columns of $\mathbf{P}_k\mathbf{V}$ define the plane wave response space for the k^{th} array). For each k let Ψ_k be the $2 \times N_k$ matrix giving the polarization parameters for the events of the k^{th} array. We propose minimizing

$$\mathbf{J} = \sum_{k=1}^{K} \| \hat{\Sigma}_k^{-1/2} (\Xi_k - \mathbf{P}_k \mathbf{V} \Psi_k^*) \| \tag{5.1}$$

over $\mathbf{V}$ and the K matrices of nuisance parameters Ψ_k. Note that the multivariate errors-in-variables estimate used for single array estimation minimizes this quantity with K=1.

The function $\mathbf{J}$ can be written in terms of the eigenvalues and eigenvectors of the individual spectral density matrices $\mathbf{S}_m$. Specifically, let $\Xi'_k = \hat{\Sigma}_k^{-1/2}\Xi_k$, $\mathbf{P}'_k = \hat{\Sigma}_k^{-1/2}\mathbf{P}_k$ and form the singular value decomposition of Ξ'_k.

$$\Xi'_k = \mathbf{U}_k\Lambda_k\mathbf{W}_k$$

Then it can be shown that we can rewrite $\mathbf{J}$ as

$$\mathbf{J} = \sum_{k=1}^{K} \|\mathbf{U}_k\Lambda_k\mathbf{W}_k - \mathbf{P}'_k\mathbf{V}\Psi\| = \sum_{k=1}^{K} \|\mathbf{T}_k - \mathbf{P}'_k\mathbf{V}\Psi'_k\| \qquad (5.2)$$

where $\Psi'_k = \Psi_k\mathbf{W}_k^*$ are $2\times m_k$ matrices of nuisance parameters, and $\mathbf{T}_k = \mathbf{U}_k\Lambda_k = \mathbf{S}_k^{-1/2}$ can be formed from the eigenvalues and eigenvectors of the SDM $\mathbf{S}_k$ for the k^{th} array. If $\mathbf{V}$ were known the Ψ'_k's which minimize (5.2) are

$$\hat{\Psi}_k = (\mathbf{V}_k^*\mathbf{V}_k)\mathbf{V}_k^*\mathbf{T}_k \qquad \text{where } \mathbf{V}_k = \mathbf{P}'_k\mathbf{V}$$

On the other hand, with the Ψ_k's known, finding $\mathbf{V}$ to minimize $\mathbf{J}$ can be reduced to a standard least squares problem. These observations suggest an iterative scheme for minimizing $\mathbf{J}$. Using a simple approach for combining arrays, such as one outlined above, we may obtain a starting estimate of the synthetic array response space vectors $\mathbf{V}$. These can be used to estimate the Ψ_k's, which can then be used to improve the estimate of $\mathbf{V}$. This estimation scheme must still be tested with real data, and the derivation of estimation errors has not yet been attempted. Note also that this approach to combining arrays can also be used to treat the case of missing data when estimating the response plane for a single array.

The methods developed in this dissertation have the potential to allow for a substantial improvement in the treatment of source effects in the analysis and interpretation of data from small geomagnetic arrays. With improvements in the basic methods outlined here, with better instruments, and with some care in the layout of stations, it should be possible to minimize the effects of finite spatial

scale sources on plane wave response estimates and to extend the range of useful information to include gradient responses. With the development of optimal methods for combining small arrays, we will have the ability to apply these methods to the detailed mapping of the conductivity variations of the earth over large areas. These techniques thus have the potential to make possible with small arrays many things which have been thought possible only with much larger arrays. Given the current cost of instrumentations for a large array, the advantages of our methods for small arrays could be very significant.

Figure 5.1 Anomalous horizontal and vertical fields for a small synthetic array in Southwestern Washington (inset gives array location); fields are superimposed on a contour plot of Bouger gravity anomaly. The array was constructed from two smaller overlapping 5 station arrays. The period is 300 seconds; source polarization North-South. The southernmost station has been chosen as the normal. Real parts of anomalous horizontal fields are plotted as vectors; real parts of vertical fields are written under the station location. The imaginary parts show a similar pattern and imply that the anomalous fields lead the normal fields by about 20 degrees (so structure is fairly shallow). Horizontal fields are enhanced in the center of the array and there is a reversal in the sign of the vertical fields coincident with the largest horizontal fields. Note the substantial turning of the anomalous fields in the center of the array. The anomalous currents appear to flow out of the east and turn sharply to the north. This is presumably due to deflection by a region of high resistivity in the crust in the lower left hand corner of the figure. Note that this region coincides with a high in the gravity.

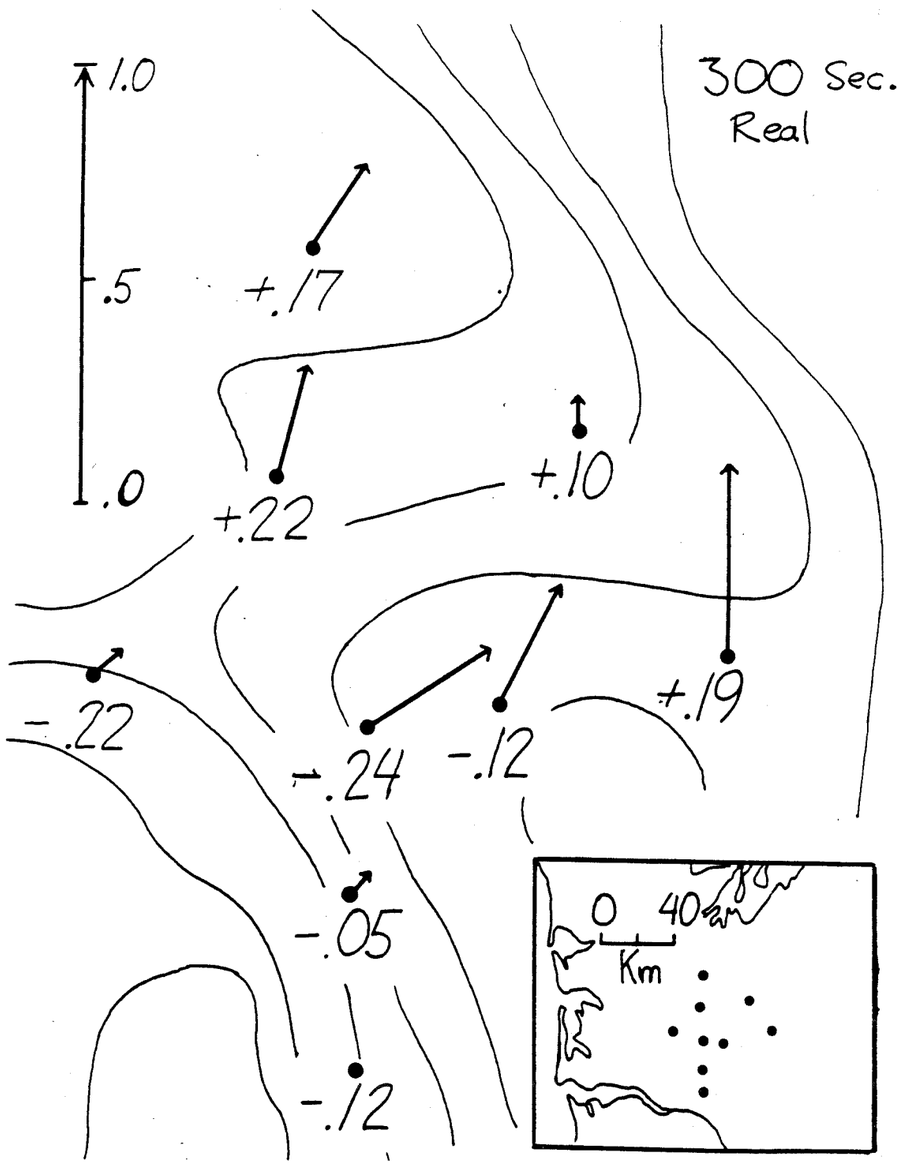


Figure 5.2 Approximate real part of horizontal fields due to internal currents for a synthetic array constructed from a series of small (3-5) station arrays. The external source polarization is defined by assuming that the average field is normal. We then find the vector in the response plane whose average over all stations corresponds to a total field of unit magnitude linearly polarized in the north- south direction. The assumed external field (one half of the normal total field) is subtracted from the field vector at each station before plotting. The variations in horizontal field magnitudes are fairly subtle, but several zones of enhanced current flow can be seen.

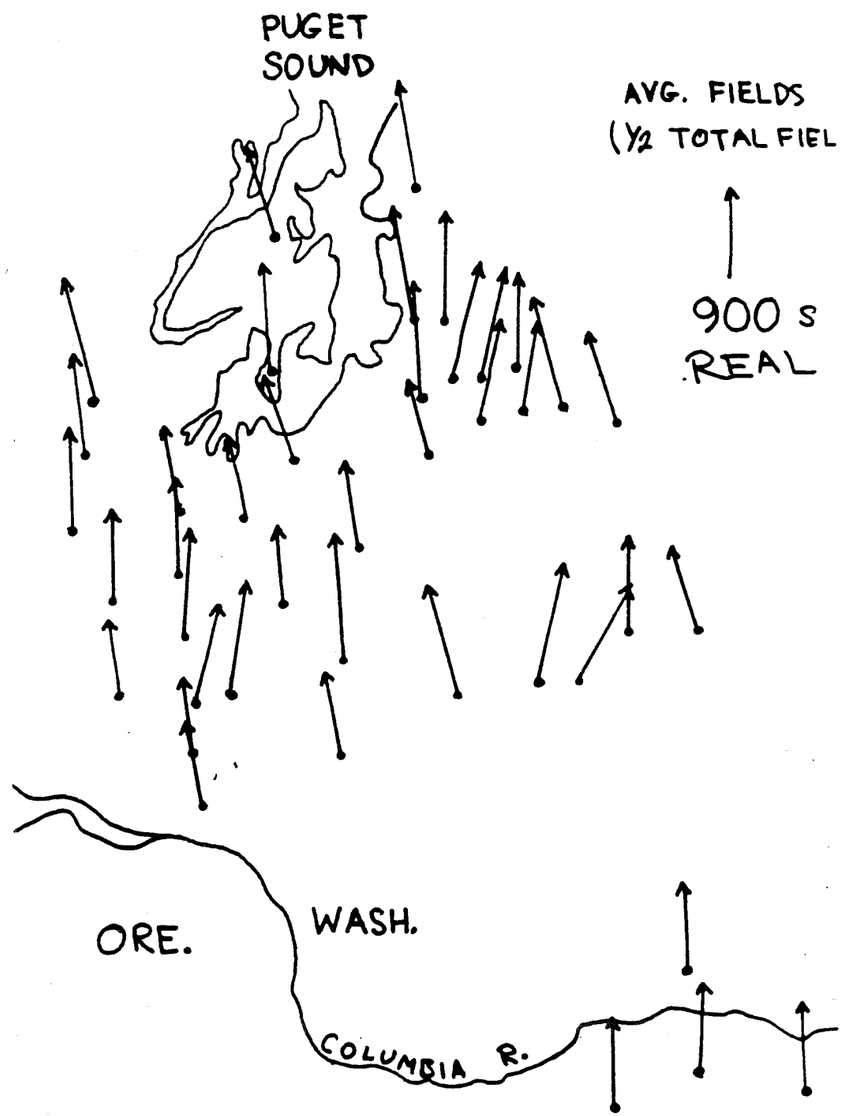


Figure 5.3 Vertical field hypothetical event for array in Southwestern Washington for north-south inducing fields. The large conductive feature under the Cascades evident for this source polarization is larger for the other polarization (figure 5.5). The smaller east-west trending anomaly midway between Puget Sound and the mouth of the Columbia river is shown in more detail in figure 5.1.

Figure 5.4 Approximate real part of horizontal fields due to internal currents for a synthetic array constructed from a series of small (3-5) station arrays. The external source polarization is defined by assuming that the average field is normal. We then find the vector in the response plane whose average over all stations corresponds to a total field of unit magnitude linearly polarized in the east-west direction. The assumed external field (one half of the normal total field) is subtracted from the field vector at each station before plotting. The horizontal fields over the center of the conductivity anomaly are more than twice the normal (average) fields.

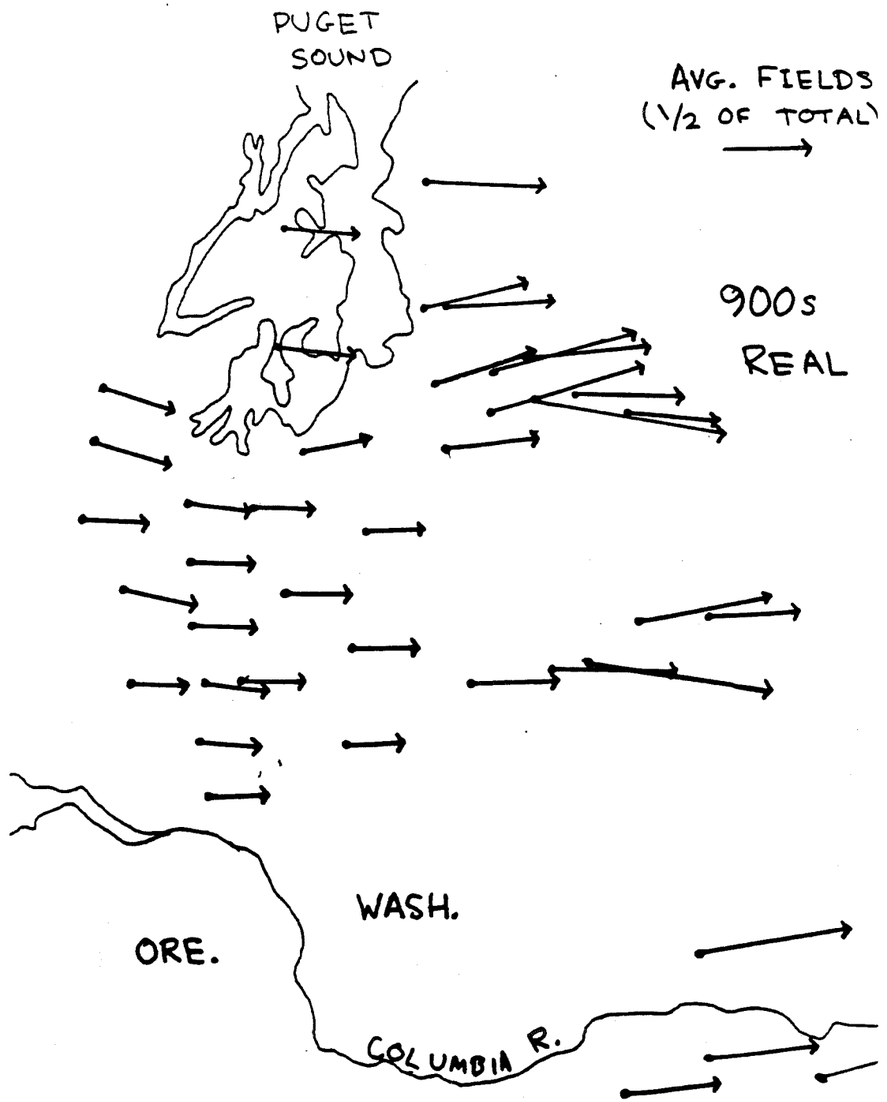


Figure 5.5: Vertical field hypothetical event for array in Southwestern Washington. The real part of the response in Z to linearly polarized east-west inducing fields is contoured (period = 300 seconds). Dashed contours are negative, solid positive. A significant conductive feature under the Cascades is clearly delineated by the sharp reversal of the sign of Z.

List of Refernces

Adler, R.J., *The Geometry of Random Fields*, Wiley, New York, 1981.

Alabi, A.O., Magnetometer array studies, *Geophys. Surv., 6*, 153-172, 1983.

Albert, A.A., The matrices of factor analysis, *Proc. Nat. Acad. Sci., 30*,,90-95, 1944.

Anderson, T.W. Statistical inference for covariance matrices with linear structure, in *Proceedings of the Second International Symposium on Multivariate Analysis*, ed. P.R. Krishnaiah, Academic Press, New York, 55-66, 1969.

Anderson, T.W. Estimation of covariance matrices which are linear combinations or whose inverses are linear combinations of given matrices, in *Essays in Probability and Statistics*, ed. R.C. Bose, I.M. Chakravarti, P.C. Mahlanobis, C.R. Rao, and K.J.C. Smith, University of North Carolina Press, Chapel Hill, 1-24, 1970.

Anderson, T.W., Asymptotically efficient estiamtion of covariance matrices with linear structure, *Ann. Statist., 1*, 135-141, 1973.

Anderson, T.W., Estimating Linear Statistical Relationships, *Ann. Statist., 12*, 1-45, 1984.

Anderson, T.W., and H. Rubin, Statistical inference in factor analysis, *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 5*, 111-150, ed., Jerzy Neyman, University of California Press, 1956.

Babour, K., J. Mosnier, M. Daignieres, G. Vasseur, J.L. Lemouel, and J.C. Rossignol, The geomagnetic variation anomaly in the northern Pyrenees: study of the temporal variation, *Geophys. J. R. astr. Soc, 45*, 593-567, 1976.

Babour, K. and J. Mosnier, Differential geomagnetic sounding in the Rheingraben, *Geophys. J. R. astr. Soc., 58*, 135-144, 1979.

Backus, G.E., J. Park, and D. Garbasz, On the relative importance of the driving forces of plate motion, *Geophys. J. R. astr. Soc., 67, 425-435, 1981.*

Bailey, R.C., R.N. Edwards, G.D. Garland, R.D. Kurtz, and P. Pitcher, Electrical studies over a tectonically active area in Eastern Canada, J. Geomag. Geoelec., 26 125-146.

Bannister, R.J. and D.I. Gough, Development of a polar magnetic storm: a two-dimensional magnetometer array study, *Geophys. J. R. astr. Soc., 51*,75-90, 1977.

Bannister, J.R. and D.I. Gough, A study of two polar magnetic substorms with a two-dimensional magnetometer array, *Geophys. J. R. astr. Soc., 53*, 1-26, 1978.

Banks, R.J., Geomagnetic variations and the electrical conductivity of the upper mantle, *Geophys. J. R. astr. Soc., 17*, 457-487, 1969.

Banks, R.J., and D. Beamish, Local and regional induction in the British Isles, *Geophys. J. R. astr. Soc., 79*, 539-554, 1985.

Banks, R.J. and P. Ottey, Geomagnetic deep sounding in and around the Kenya rift valley, *Geophys. J. R. astr. Soc., 36*, 321-335, 1974.

Bartlett, M.S., A note on multiplying factors for various χ^2 approximations, *J. Roy. Statist. Soc. Ser. B, 16*, 296-298, 1954.

Beamish, D., The mapping of induced currents around the Kenya rift: a comparison of techniques, *Geophys. J. R. astr. Soc., 50*, 311-332, 1977.

Beamish, D., Source field effects on transfer functions at mid-lattitudes, *Geophys. J. R. astr. Soc., 58*, 117-134, 1979.

Beamish, D., A comparison of time and frequency domain geomagnetic sounding, *Geophys. J. R. astr. Soc., 73*, 689-704, 1983.

Beamish, D., and P.M. Johnson, Difficulties in the application of magnetic frequency gradient analysis to induction studies, *Phys. E. Pl. Int., 28*, 1-13, 1982.

Beamish, D., and R.J. Banks, Geomagnetic variation anomalies in northern England: processing and presentation of data from a non-simultaneous array, *Geophys. J. R. astr. Soc., 75*, 513-539, 1983.

Berdichevsky, M.N., and M.S. Zhdanov, *Advanced Theory of Deep Geomagnetic Sounding*, New York: Elsevier, 408 pp., 1984.

Bickel, P.J. and Doksum, K.A. *Mathematical Statistics: Basic Ideas and Selected Topics*, Holden-Day, San Francisco, 1977.

Brillinger, D.R., *Time Series: Data Analysis and Theory*, Holden-Day, San Francisco, 1981.

Browne, M.W., Generalized least squares estimates in the analysis of covariance structures, *South African Statist. J., 8,* 1-24, 1974.

Caignard, L., Basic theory of the magneto-telluric method of geophysical prospecting, *Geophysics, 18*, 605-635.

Camfield, P.A., D.I. Gough, and H. Porath, Magnetometer array studies in the north-western United States and south-western Canada, *Geophys. J. R. astr. Soc., 22*, 201-221, 1971.

Camfield, P.A., and D.I. Gough, A possible Pretorozoic plate boundary in North America, *Can. J. Earth Sci., 14*, 1229-1238, 1977.

Campbell, W.H., Australian region conductivity in the upper mantel determined from *Sq* fields, *Proc. Eighth Workshop on Electromagnetic Induction in the Earth and Moon*, IAGA, Neuchatel, Switzerland, August 24-31, 1986.

Chapman, S. and J. Bartels, Geomagnetism, Oxford University Press, London, 1962.

Chave, A.D., D.J. Thomson, and M.E. Ander, On the robust estimation of power spcetra, coherences, and transfer functions, *J. Geophys. Res., 92*, 633-648, 1987.

Clarke, J., T.D. Gamble, W.M. Goubau, D.A. Koch, and R.F. Miracky, Remote-reference magnetotellurics: equipment and processing, *Geophys. Prosp., 31*, 149-170, 1983.

Dmitriev, V.I., and M.N. Berdichevsky, The fundamental model of magnetotelluric sounding, *Proc. IEEE , 67*, 1033-1044, 1979.

Doob, J.L., *Stochastic Processes,* John Wiley, New york, 1953.

Egbert, G., and J.R. Booker, Robust estimation of geomagnetic transfer functions, *Geophys. J. R. astr. Soc., 87*, 173-194, 1986.

The EMSLAB Group, The EMSLAB electromagnetic sounding experiment,Submitted to *Nature*, 1987.

Everett, J.E. and R.D. Hyndman, Geomagnetic variations and electrical

conductivity structure in south-western Asustralia, *Phys. E. Pl. Int., 1*, 24-34.

Folland, G.B., *Introduction to Partial Differential Equations*, Mathematical Notes, Princeton University Press, Princeton, 1976.

Frazer, M.C., Geomagnetic deep sounding with arrays of magnetometers, *Rev. Geophys. Space, Phys., 12*, 401-420, 1974.

Gamble, T.D., W.M. Goubau, and J. Clarke, Magnetotellurics with a remote reference, *Geophysics, 44*, 53-68, 1979.

Giri, N.C, *Multivariate Statistical Inference* Academic Press, New York, 1977.

Gleser, L.J., Estimation in a multivariate 'errors-in-variables' regression model: large sample results, *Ann. Statist., 9*, 24-44, 1981.

Gleser, L.J., Functional, structural and ultrastructural errors-in-variables models, *ASA Proc. Bus. Econ. Sec.*, 57-66, 1983.

Gleser, L.J., and G.S. Watson, Estimation of a linear transformation, *Biometrika, 60*, 525-534, 1973.

Gnedenko, B.V. and A.N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables*, translated by K.L. Chung, Addison Wesely, Reading, Mass., 1968.

Goodman, N.R., Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction), *Ann. Math Statist., 20*, 225-240, 1963.

Gough, D.I., The interpretation of magnetometer array studies, *Geophys. J. R. astr. Soc., 35*, 83-98, 1973.

Gough, D.I., Electrical conductivity under western North America in relation to heat flow, seismology, and structure, *J. Geomag. Geoelec., 26*, 105-123, 1974.

Gough, D.I., Electromagnetic geophysics and global tectonics, *J. Geophys. Res., 88*, 3367-3377, 1983.

Gough, D.I., and M.R. Ingham, Interpretation methods for magnetometer arrays, *Rev. Geophys., 21*, 805-827, 1983.

Graybill, F.A., *Theory and Applications of the Linear Model*,Duxbury Press, North Scituate, Mass., 1976.

Gregori, G.P., and L.J. Lanzerotti, Geomagnetic depth sounding by induction arrow representation: a review, *Rev. Geophys. Space Phys., 18*, 203-209, 1980.

Hart, A.M., D.J. Krause, and W.G.V. Rosser, Possible contribution of leakage current from the Atlantic Ocean to the magnetic field variations observed at Devon, *Phys. E. Pl. Int., 32*, 107-113, 1983.

Helferty, M.G., J.R. Booker, B.R. Weertman, D.R. Auld, Detection of the south edge of te subducting Gorda Plate using geomagnetic induction, submitted to *Nature*, 1987.

Hermance, J.F., Magnetotelluric and geomagnetic deep sounding studies in rifts and adjacent areas: constraints on phhysical processes in the crust and upper mantle, in *Continental and Oceanic Rifts*, G. Palmason, ed., AGU, Wash. D.C., 169-183, 1982.

Hermance, J.F. Electromagnetic induction studies, *Rev. Geophys., 21*, 652-664, 1983.

Hermance, J.F., Electromagnetic induction by finite wavenumber source fields in a two-dimensional lateral heterogeneity; the transverse electric mode, *Geophys. J. R. astr. Soc., 78*, 159-179, 1984.

Ingham, M.R., D.K. Bingham, and D.I. Gough, A magnetovariational study of a geothermal anomaly, *Geophys. J. R. astr. Soc., 72*, 597-618, 1983.

Jackson, J.D., *Classical Electrodynamics*, second ed., John Wiley and Sons, New York, 1975.

James, A.T., Tests of Latent Roots of the covariance matrix, in *Proceedings of the Second International Symposium on Multivariate Analysis*, ed. P.R. Krishnaiah, Academic Press, New York, 205-218, 1969.

Jenkins, G.M., and D.G. Watts, *Spectral Analysis and its Applications*, Holden-Day, 1968.

Jones, A.G., The problem of current channeling: a critical review, *Geophys. Surv., 6*, 79-122, 1983.

Jones, A.G., Geomagnetic induction studies in Scandanavia I - determination of the inductive response function from magnetometer array data, *J. Geophys., 48*, 181-194, 1980.

Joreskog, K.G., and A.S. Goldberger, Factor analysis by generalized least squares, *Psychometrika, 37*, 183-202, 1972.

Junge, A., Lunar variation of the electric field in Northern Germany induced by tidal currents of the German Sea, *Proc. Eighth Workshop on Electromagnetic Induction in the Earth and Moon*, IAGA, Neuchatel, Switzerland, August 24-31, 1986.

Kato, T., *Perturbation Theory for Linear Operators*, Springer-Verlag, Berlin, 1966.

Kendall, M.G., and A. Stuart, *The Advanced Theory of Statistics, Vol. 2, Inference and Relationship, Fourth Edition*, Griffin, London, 1979.

Kuckes, A.F., Relation between electrical conductivity of a mantle and fluctuating magnetic fields, *Geophys. J. R. astr. Soc., 32*, 119-131, 1973.

Kuckes, A.F., A.G. Nekut, and B.G. Thompson, A geomagnetic scattering theory for evaluation of earth structure, *Geophys. J. R. astr. Soc., 83*, 319-330, 1985.

Kupper, F., J. Untiedt, W. Baumjohann, K. Lange, and A.G. Jones, A two-dimensional magnetometer array for ground-based observations of auroral electric currentsduring the International Magnetospheric Study (IMS), *J. Geophys, 46*, 429-450, 1979.

Lamperitti, J., *Probability*, Benjamin/Cummings, Reading Mass., 1966.

Larsen, J.C., Removal of local surface conductivity effects from low frequency mantle response curves, *Acta Geodaet. Geophys. et Montanist. Acad. Hung., 12*, 183-186, 1977.

Law, L.K., D.R. Auld, and J.R. Booker, A geomagnetic variation anomaly coincident with the Cascade volcanic belt, *J. Geophys. Res., 85*, 5297-5302, 1980.

Lawley, D.N., Tests of significance of the latent roots of covariance and correlation matirces, *Biometrika, 43*, 128-136, 1954.

Lehman, E.L., *Theory of Point Estimation*, Wiley, New York, 1983.

Lilley, F.E.M., and D.J. Bennett, An array experiment with magnetic variometers near the coasts of south-east Australia, *Geophys. J. R. astr. Soc., 32*, 49-64, 1972.

Lilley, F.E.M., and Sloanes, M.N., On estimating electrical conductivity using gradient data from magnetometer arrays, *J. Geomag. Geoelectr., 28*, 321-328, 1976.

Lilley, F.E.M., Woods, D.V. and Sloane, M.N., Electrical conductivity from Australian magnetometer arrays using spatial gradient data, *Phys. E. Pl. Int., 25*, 202-209, 1976.

Lilley, F.E.M., D.V. Woods, and M.N. Sloane, Electrical conductivity from Australian magnetometer arrays using spatial gradient data, *Phys. Earth Planet. Int., 25*, 202-209, 1981.

Mareschal, M., Modelling of natural sources of magnetospheric origin in the interpretation of regional induction studies: a review, *Geophys. Surv., 8*, 261-300, 1986.

Mauchley, J.W., Significance test for sphericity of a normal *n*-variate distribution, *Ann. Math. Statist., 11*, 204-209, 1946.

McMechan, G.A., and I. Barrodale, Processing electromagnetic data in the time domain, *Geophys. J. R. astr. Soc., 81*, 277-294, 1985.

Merrill, R.T., and M.W. McElhinny, *The Earth's Magnetic Field: It's History, Origin, and Planetary Perspective*, Academic Press, London, 1983.

Miller, K.S., *Complex Stochasitic Process*, Addison Weseley, Reading Mass., 1974.

Moran, P.A.P., Estimating structural and functional relationships, *J. Multivariate Anal., 1*, 232-255, 1971.

Morrison, D.F., *Multivariate Statistical Methods*, McGraw-Hill, San Francisco, 1967.

Mosnier, J., A study of the physics of telluric current flow at very low frequencies in the earth's crust, *Geophys. J. R. astr. Soc., 82*, 479-496, 1985.

Park, J., and A.D. Chave, On the estimation of magnetotelluric response functions using the singular value decomposition, *Geophys. J. R. astr. Soc., 77*, 683-709, 1984.

Parkinson, W.D., The influence of continents and oceans on geomagnetic variations, *Geophys. J. R. astr. Soc., 4*, 441-449, 1962.

Pedersen, L.B., and M. Svennekjaer, Extremal bias coupling in magnetotellurics, *Geophysics, 49*, 1968-1978, 1984.

Pedersen, L.B., Some aspects of magnetotelluric field procedures, *Proc. Eighth Workshop on Electromagnetic Induction in the Earth and Moon*, IAGA, Neuchatel, Switzerland, August 24-31, 1986.

Price, A.T. and G.A. Wilkins, New methods of analysis of geomagnetic fields ant their application to the *Sq* field of 1932-1933, *Phil. Trans. R. Soc. London, Ser. A, 256* 31-98, 1963.

Porath, H. and A. Dziewonski, Crustal electrical conductivity anomalies in the Great Plains province of the United States, *Geophys., 36*, 382-395, 1971.

Porath, H. and D.I. Gough, Mantle conductivity structures in the Western United States from magnetometer array studies, *Geophys. J. R. astr. Soc., 22*, 261-275, 1971.

Porath, H., D.I. Gough, and P.A. Camfield, Conductive structures in the Northwestern United States and Southwest Canada, *Geophys. J. R. astr. Soc., 23* 387-398, 1970.

Porath, H., Oldenburg, D.W. and D.I. Gough, Separation of magnetic variation fields and conductive structure in the western United States, *Geophys. J. R. astr. Soc., 19*, 237-260, 1970.

Richmond, A.D., and W. Baumjohann, Three-dimensional analysis of magnetometer array data, *J. Geophys., 54*, 138-156, 1983.

Ripely, B., *Spatial Statistics*, Wiley, New York, 1981.

Rokityansky, I.I., *Geoelectromagnetic Investigation of the Earth's Crust and Mantle*, New York: Springer-Verlag, 381 pp., 1982.

Samson, J.C., Pures states, polarized waves, and principal components in the spectra of multiple, geophysical time-series, *Geophys. J. R. astr. Soc., 72*,647-

664, 1983.

Schultz, A. and J.C. Larsen, Analysis of zonal field morphology and data quality for a global set of magnetic observatory daily mean values, *J. Geomag. Geoelec., 35*, 835-84, 1983.

Schultz, A., and J.C. Larsen, On the electrical conductivity of the earth's interior I: mid-mantle response function computation, *Geophys. J. Roy. astr. Soc.*, in press, 1987.

Schmuker, U., Anomalies of geomagnetic variations in the sothwestern United States, *J. Geomag. Geoelec., 15*, 193-221, 1964.

Schmuker, U., *Anomalies of geomagnetic variations in the sothwestern United States*, Bull. Scripps Inst. Oceanogr., 13, 1970.

Schnegg, P.A., Fischer, G., Fontes, S., Hutton, V.R.S., Finzi, E., The effect of D.C. railways on MT measurements in Northern Italy, *Proc. Eighth Workshop on Electromagnetic Induction in the Earth and Moon*, IAGA, Neuchatel, Switzerland, August 24-31, 1986.

Sprent, P., A generalized least-squares approach to linear functional relationships (with discussion), *J. Roy. Statist. Soc. Ser. B, 28*, 278-297, 1966.

Stanley, W.D., Tectonic study of Cascade Range and Columbia Plateau in Washington state based upon magnetotelluric soundings, *J. Geophys. Res., 89*, 4447-4460, 1984.

Summers, D.M., On the frequency response of induction anomalies, *Geophys. J. R. astr. Soc., 70*, 487-502, 1982.

Svetov, B.S., Processing of magnetotelluric field variations, *Proc. Eighth Workshop on Electromagnetic Induction in the Earth and Moon*, IAGA, Neuchatel, Switzerland, August 24-31, 1986.

Svetov, B.S. and M.I. Shimelevich, Determination of linear relationships among the magnetotelluric field components: basic principles, *Izv. Akad. Nauk SSSR, ser. fiz. zem.*, 365-372, 1982.

Thomson, D.J., Spectrum estimation and harmonic analysis, *Proc. IEEE, 70*,1055-1096.

Thomson, D.J., Robbins, M.F., MacLennan, C.G. and L.J. Lanzerotti, Spectral and

windowing techniques in power spectral analyses of geomagnetic data, *Phys. E. Pl. Int., 12*, 217-231, 1976.

Tikhonov, A.N., On determination of electric characteristics of deep layers of the earth's crust, *Dokl. Akad. Nauk SSSR, 73*, 295-297, 1950.

Thurstone, L.L., *Multiple-Factor Analysis*, Univ. of Chicago Press, 1947.

Vozoff, K., and C.M. Swift, Magnetotelluric measurements in the North German Basin, *Geophys. Prospecting, 16*, 414-436, 1968.

Watson, G.S., *Statistics on Spheres*, Wiley-Interscience, 1983.

Woods, D.V. and F.E.M. Lilley, Anomalous geomagnetic variations and the concentration of telluric currents in southwest Queensland, Australia, *Gophys. J. R. astr. Soc., 62*, 675-689, 1980.

Zhang, P., Roberts, R.G. and Pedersen, L.B., Magnetotelluric Strike Rules, *Geophysics*, (in press), 1987.

Appendix A

Biases in the Plane Wave Eigenvectors

In chapter three we developed a random source model to study the properties of the SDM for small arrays. The model assumed that conductivity was 1-d and that the source potential could be treated as a spatially homogeneous random field. Under these assumptions we have seen that the dominant two eigenvectors of the SDM correspond very closely to uniform plane wave sources. In this appendix we consider (somewhat schematically) the effects of violations of these assumptions on our conclusions.

The seperation of the plane wave and gradient terms into separate sets of eigenvectors depends on two facts. First, for any plane wave vector $\mathbf{v}_p$ and any gradient vector $\mathbf{v}_g$ we have $\mathbf{v}_p^*\mathbf{v}_g = 0$ - the plane wave and gradient terms are orthogonal. Second, as a consequence of our assumption of spatial homogeneity the plane wave and gradient terms are uncorrelated (as discussed in section 3.7). To understand the consequences of violations of these conditions we consider a simple model. Suppose, for simplicity, that the data vectors are linear combinations of a single plane wave term and a single gradient term

$$\mathbf{X} = \alpha_p \mathbf{v}_p + \alpha_g \mathbf{v}_g \tag{A.1}$$

and let the covariances of the random coefficients in (A.1) be

$$E(\alpha_p \alpha_p^*) = 1 \qquad E(\alpha_g \alpha_g^*) = \varepsilon^2 \qquad E(\alpha_p \alpha_g^*) = \rho\varepsilon$$

so that the ratio of the power in gradient term to the power in the plane wave term is ε^2 and the (complex) correlation of the plane wave and gradient terms is ρ. Then the expectation of the SDM has the form

$$\Sigma = E(\mathbf{S}) = E(\mathbf{X}\mathbf{X}^*) = \mathbf{v}_p\mathbf{v}_p^* + \rho\varepsilon\mathbf{v}_p\mathbf{v}_g^* + \rho^*\varepsilon\mathbf{v}_g\mathbf{v}_p^* + \varepsilon^2\mathbf{v}_g\mathbf{v}_g^*$$

The dominant eigenvector of Σ will not in general be $\mathbf{v}_p$. Thus the dominant eigenvector of $\mathbf{S}$, $\hat{\mathbf{v}}_p$, will in general be a biased estimate of $\mathbf{v}_p$. We calculate the

magnitude of this bias.

We first consider the case where the gradient and plane wave terms are correlated ($\rho \neq 0$). To do this we use the approach of section 4.2. Let $\mathbf{P} = \mathbf{v}_p\mathbf{v}_p^*$ and let $\hat{\mathbf{P}} = \hat{\mathbf{v}}_p\hat{\mathbf{v}}_p^*$. Then using (4.2.3) and simplifing the resulting expression slightly, we have, correct to first order in ε,

$$\hat{\mathbf{P}} = \mathbf{P} + 2\varepsilon\mathrm{Re}[\,\rho\mathbf{v}_p\mathbf{v}_g^*\,] \tag{A.2}$$

For this case (corresponding to the case of non-homogenous sources) the bias is thus of order $\varepsilon|\rho|$.

The case where the gradient and plane wave vectors are not orthogonal (this will occur in the general case where there are anomalous horizontal and vertical fields) is treated in a similar fashion. Assuming now that $\rho = 0$ defining $\mathbf{P}$ and $\hat{\mathbf{P}}$ as above, and letting $\xi = \mathbf{v}_p^*\mathbf{v}_g$ we find, correct to first order in ε^2

$$\hat{\mathbf{P}} = \mathbf{P} + \varepsilon^2\, 2\mathrm{Re}[\xi\mathbf{v}_p\mathbf{v}_g^*] \tag{A.3}$$

For this case, then, the bias is of order $|\xi|\varepsilon^2$. Unless anomalous fields are very large, the perturberd plane wave and gradient terms will still be nearly orthogonal and we will have $|\xi| \ll 1$. Since typically $\varepsilon \ll 1$, this bias is usually very small.

Appendix B

Assymptotic Distribution of the Complex MEV Estimates

In this appendix we sketch the derivation of the assymptotic distribution of the estimates $\hat{\mathbf{T}}$ of section 4.1.

First, a few basic facts about moments of the complex multivariate Gaussian distribution. Suppose that the random complex vector $\mathbf{Z}$ has an m-variate complex Gaussian distribution with mean zero and covariance matrix $\mathbf{I}_m$. Then $\mathbf{Z} = \mathbf{X} + i\mathbf{Y}$ where $\mathbf{X}$ and $\mathbf{Y}$ are independent mean zero m-variate real Gaussian random vectors each with covariance matrix $\frac{1}{2}\mathbf{I}_m$ (e.g. Brillinger, 1981). Using these facts and standard results about the moments of real Gaussian variables the following are easily derived:

(B.1) odd moments are zero; e.g. $E(Z_i Z_j Z_k^*) = 0$

(B.2) $E(Z_i Z_j^*) = \delta_{ij}$, but $E(Z_i Z_j) \equiv 0$

(B.3) $\mathbf{Cov}(Z_i Z_j^*, Z_{i'} Z_{j'}^*) = E(Z_i Z_j^* Z_{i'}^* Z_{j'}) - E(Z_i Z_j^*)\, E(Z_{i'} Z_{j'}^*)^* = \delta_{ii'}\delta_{jj'}$

(B.4) Real and imaginary parts of cross product terms such as $Z_i Z_j^*$ where $i \neq j$ are uncorrelated; each has variance one half.

These results imply that if $\mathbf{Z}_i$, $i = 1, N$ are independent m-variate Gaussian random vectors with covariance matrix $\mathbf{I}_m$, and if $\mathbf{S} = N^{-1} \sum_{i=1}^{N} \mathbf{Z}_i \mathbf{Z}_i^*$ then the elements of $\mathbf{S}$ on and below the diagonal are uncorrelated with each other. Elements on the diagonal are real and have variance N^{-1}, while the real and imaginary parts of the off diagonal terms are uncorrelated and have variance $(2N)^{-1}$.

We now consider assymptotics for the SDM $\mathbf{S}$ when the model of (4.1.1) holds. We follow Gleser (1981) closely. All notation and assumptions are as in section 4.1. In particular the α_{jk}'s are thought of as fixed but unknown parameters (they are not random). Let $\mathbf{Y}_i = \mathbf{U}\boldsymbol{\alpha}_i$ then we may write (4.1.7) as

$$\mathbf{Z}_i = \mathbf{e}_i \mathbf{e}_i^* - \sigma^2 \mathbf{I}_m + \mathbf{Y}_i \mathbf{e}_i^* + \mathbf{e}_i \mathbf{Y}_i^*$$

The random matrices $\mathbf{Z}_i$, $i = 1, N$ are independent, but not identically distributed. A simple calculation and application of (B.1), (B.2), and (B.3) yields

$$E(Z_{ijk} Z_{ij'k'}) = \sigma^4 \delta_{jj'}\delta_{kk'} + \sigma^2 [Y_{ij} Y_{ij'}^* \delta_{kk'} + Y_{ik}^* Y_{ik'} \delta_{jj'}] \qquad \text{(B.5)}$$

Note that this is *not* the same as in the real case treated by Gleser.

Let $\overline{\mathbf{Z}} = N^{-1} \sum_{i=1}^{N} \mathbf{Z}_i$. Lemma 4.4 in Gleser can be adapted to show that provided the limit

$$\lim_{N\to\infty} N^{-1}\, \boldsymbol{\alpha}\boldsymbol{\alpha}^* = \boldsymbol{\Sigma}_\alpha$$

exists the elements of the complex random matrix $N^{1/2}\overline{\mathbf{Z}}$ are assymptotically complex multivariate Gaussian with mean zero and covariance

$$\lim_{N\to\infty} \mathbf{Cov}(N^{1/2}\overline{Z}_{jk}, N^{1/2}\overline{Z}_{j'k'}) = \lim_{N\to\infty} N^{-1} \sum_{i=1}^{N} E(Z_{ijk} Z_{ij'k'})$$

$$\rightarrow \sigma^4 \delta_{jj'}\delta_{kk'} + \sigma^2 [\, (\Sigma_\alpha)_{jj'}\delta_{kk'} + (\Sigma_\alpha)_{kk'}\delta_{jj'}\,] \qquad \text{(B.6)}$$

Since $\mathbf{S} = \boldsymbol{\Sigma} + \overline{\mathbf{Z}}$, this specifies the asymptotic distribution of $\mathbf{S}$.

We next consider the computation of the asymptotic distribution of the estimates $\mathbf{T}$. Let

$$\boldsymbol{\Theta} = -\boldsymbol{\Sigma}_\alpha^{-1} (\mathbf{I}_p + \mathbf{T}^*\mathbf{T})^{-1} (\, \mathbf{I}_p \;\; \mathbf{T}^* \,) \qquad \boldsymbol{\Psi} = \begin{bmatrix} \mathbf{T}^* \\ -\mathbf{I}_{m-p} \end{bmatrix} \qquad \text{(B.7)}$$

and let $\mathbf{G} = N^{1/2}\overline{\mathbf{Z}}$. Then results given in Gleser for the real case can be adapted to show that the asymptotic distributions of $N^{1/2}(\hat{\mathbf{T}} - \mathbf{T})$ and $[\, \boldsymbol{\Theta}\mathbf{G}\boldsymbol{\Psi} \,]$ are identical. We can thus conclude that $N^{1/2}(\hat{\mathbf{T}} - \mathbf{T})$ is asymptotically Gaussian with covariance

$$\mathbf{Cov}(\, (\hat{\mathbf{T}} - \mathbf{T})_{ij}, (\hat{\mathbf{T}} - \mathbf{T})_{i'j'} \,) = \sum_{kk'l'} \theta_{ik}\theta_{i'k'}\psi_{lj}\psi_{l'j'}\, E(G_{kl}G_{k'l'}) \qquad \text{(B.8)}$$

where the covariances of the elements of $\mathbf{G}$ are given in (B.6).

Substituting the covariances from (B.6) into (B.8) and simplifying we find

$$\mathbf{Cov}(\,(\hat{\mathbf{T}})_{ij},\,(\hat{\mathbf{T}})_{i'j'}\,)$$

$$= N^{-1}\left\{\sigma^4[(\Theta\Theta^*)_{ii'}(\Psi^*\Psi)_{jj'}] + \sigma^2[(\Theta\Sigma_\alpha\Theta^*)_{ii'}(\Psi^*\Psi)_{jj'}]\right\}$$

Noting that

$$\Theta\Theta^* = \Sigma_\alpha^{-1}(\mathbf{I}^p + \mathbf{T}^*\mathbf{T})^{-1}\Sigma_\alpha^{-1} \qquad \Psi^*\Psi = (\mathbf{I}_{m-p} + \mathbf{T}\mathbf{T}^*)$$

we find

$$\mathbf{Cov}((\hat{\mathbf{T}})_{ij},\,(\hat{\mathbf{T}})_{i'j'}) \tag{B.9}$$

$$= N^{-1}\sigma^2\left[\sigma^2\Sigma_\alpha^{-1}\,(\,\mathbf{I}_p + \mathbf{T}^*\mathbf{T}\,)\,\Sigma_\alpha^{-1} + \Sigma_\alpha^{-1}\right]_{jj'}\left[\mathbf{I}_{m-p} + \mathbf{T}\mathbf{T}^*\right]_{ii'}$$

Note that this result is identical (after substituting conjugate transpose for all occurences of transpose in matrix operations) to that given by Gleser for the real case although the same substitutions will not yield the correct asymptotic distribution of $\overline{\mathbf{Z}}$.

Appendix C

Parameter Identification Theorems

*Theorem*1 : Suppose $\mathbf{\Sigma} = \mathbf{\Sigma}_S(\tau) + \mathbf{\Sigma}_N(\boldsymbol{\theta})$ where $\tau = (\mathbf{T}, \mathbf{\Sigma}_\alpha)$ as in (4.4.1) and where $\mathbf{\Sigma}_N$ has the block diagonal form of (4.4.2) with $k\times k$ matrices on the diagonal. Suppose that $\mathbf{\Sigma}_\alpha$ is non-singular (of rank p) and that whenever the k rows of

$$\mathbf{U} = \begin{bmatrix}\mathbf{I}_p \\ \mathbf{T}\end{bmatrix}$$

corresponding to any block on the diagonal are deleted from $\mathbf{U}$, the remaining rows of $\mathbf{U}$ can be rearanged to form two disjoint matrices of rank p. Then the parameters of $\mathbf{\Sigma}_S$ and $\mathbf{\Sigma}_N$ are uniquely determined (identified).

Proof: Let

$$\mathbf{\Lambda} = \mathbf{U}\mathbf{\Sigma}_\alpha^{1/2}$$

so that $\mathbf{\Sigma}_S = \mathbf{\Lambda}\mathbf{\Lambda}^*$. Note that since the i^{th} row of $\mathbf{\Lambda}$ is just the i^{th} row of $\mathbf{U}$ post-multiplied by the nonsingular matrix $\mathbf{\Sigma}_\alpha^{1/2}$, the conditions of the theorem hold for $\mathbf{\Lambda}$ whenever they hold for $\mathbf{U}$. Thus it suffices to prove the theorem with $\mathbf{\Lambda}$ in place of $\mathbf{U}$.

Suppose then, that

$$\mathbf{\Lambda}\mathbf{\Lambda}^* + \mathbf{\Sigma}_N = \mathbf{L}\mathbf{L}^* + \mathbf{\Sigma}_N'$$

We wish to show that $\mathbf{\Lambda} = \mathbf{L}$. Note that the elements of $\mathbf{\Lambda}\mathbf{\Lambda}^*$ and $\mathbf{L}\mathbf{L}^*$ that are not in the diagonal blocks must be equal (since the corresponding elements of $\mathbf{\Sigma}_N$ and $\mathbf{\Sigma}'_N$ are zero). We must show that the elements on the diagonal blocks are also equal. The condition of the theorem implies that we may reorganize the rows of $\mathbf{\Lambda}$ (and the

corresponding rows of $\mathbf{L}$) as

$$\Lambda = \begin{bmatrix} \Lambda_1 \\ \Lambda_D \\ \Lambda_2 \\ \Lambda_3 \end{bmatrix} \qquad \mathbf{L} = \begin{bmatrix} \mathbf{L}_1 \\ \mathbf{L}_D \\ \mathbf{L}_2 \\ \mathbf{L}_3 \end{bmatrix}$$

where Λ_D, $\mathbf{L}_D$ are the deleted rows and where Λ_1, Λ_2 are non-singular $p \times p$ matrices. Write out the product $\Lambda\Lambda^*$ as a partitioned matrix

$$\Lambda\Lambda^* = \begin{bmatrix} \Lambda_1\Lambda_1^* & \Lambda_1\Lambda_D^* & \Lambda_1\Lambda_2^* & \Lambda_1\Lambda_3^* \\ \Lambda_D\Lambda_1^* & \Lambda_D\Lambda_D^* & \Lambda_D\Lambda_2^* & \Lambda_D\Lambda_3^* \\ \Lambda_2\Lambda_1^* & \Lambda_2\Lambda_D^* & \Lambda_2\Lambda_2^* & \Lambda_2\Lambda_3^* \\ \Lambda_3\Lambda_1^* & \Lambda_3\Lambda_D^* & \Lambda_3\Lambda_2^* & \Lambda_3\Lambda_3^* \end{bmatrix}$$

and do the same for $\mathbf{L}\mathbf{L}^*$. Since the elements of $\Lambda\Lambda^*$ and $\mathbf{L}\mathbf{L}^*$ which are not in the diagonal blocks are equal we have $\Lambda_1\Lambda_D^* = \mathbf{L}_1\mathbf{L}_D^*$ and $\Lambda_1\Lambda_2^* = \mathbf{L}_1\mathbf{L}_2^*$. We will now show that $\Lambda_D\Lambda_D^* = \mathbf{L}_D\mathbf{L}_D^*$.

Let l_{i1} be the i^{th} column of $\Lambda_1\Lambda_D^* = \mathbf{L}_1\mathbf{L}_D^*$ and let l_{j2}^* be the j^{th} row of $\Lambda_D\Lambda_2^* = \mathbf{L}_D\mathbf{L}_2^*$ and form the $(p+1)\times(p+1)$ matrices

$$\mathbf{A} = \begin{bmatrix} l_{i1} & (\Lambda_1\Lambda_2^*) \\ (\Lambda_D\Lambda_D^*)_{ij} & l_{j2}^* \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} l_{i1} & (\Lambda_1\Lambda_2^*) \\ (\mathbf{L}_D\mathbf{L}_D^*)_{ij} & l_{j2}^* \end{bmatrix}$$

$\mathbf{A}$ and $\mathbf{B}$ are submatrices of $\Lambda\Lambda^*$ and $\mathbf{L}\mathbf{L}^*$ respectively. Since $\Lambda\Lambda^*$ and $\mathbf{L}\mathbf{L}^*$ are of rank p, $\mathbf{A}$ and $\mathbf{B}$ are singular and hence have determinants equal to zero. Expanding the determinants, we may write

$$0 = |\mathbf{A}| = (-1)^p\,(\Lambda_D\Lambda_D^*)_{ij}\,|\Lambda_1\Lambda_2^*| + f(\Lambda) = (-1)^p\,(\mathbf{L}_D\mathbf{L}_D^*)_{ij}\,|\Lambda_1\Lambda_2^*| + f(\Lambda) = |\mathbf{B}|$$

where $f(\Lambda)$ is a function which depends only on the elements of Λ. Since Λ_1 and Λ_2^* are non-singular, we have $|\Lambda_1\Lambda_2^*| \neq 0$. Thus we conclude

$$(\Lambda_D\Lambda_D^*)_{ij} = (\mathbf{L}_D\mathbf{L}_D^*)_{ij}$$

Since this argument holds for any ij and for any block on the diagonal we are

finished.

Theorem 2: Suppose $\Sigma = \Sigma_S + \Sigma_N(\boldsymbol{\theta})$ where Σ_S is fixed and of rank p with $\|\Sigma_S\| = 1$. Let $\Sigma_S = \sum_{i=1}^{p} \gamma_i \mathbf{v}_i\mathbf{v}_i^*$ and let $\mathbf{P}_R = \sum_{i=1}^{p} \mathbf{v}_i\mathbf{v}_i^*$ and $\mathbf{Q}_R = \mathbf{I}_m - \mathbf{P}_R$ be the projections on to the response space $\boldsymbol{R}$ and its orthogonal complement.

Suppose that there are δ, M such that whenever

$$\|\mathbf{Q}_R[\,\Sigma_N(\boldsymbol{\theta}_1) - \Sigma_N(\boldsymbol{\theta}_2)\,]\mathbf{Q}_R\| < \delta$$

then

$$\|\,\Sigma_N(\boldsymbol{\theta}_1) - \Sigma_N(\boldsymbol{\theta}_2)\,\| \;<\; M\,\|\,\mathbf{Q}_R[\,\Sigma_N(\boldsymbol{\theta}_1) - \Sigma_N(\boldsymbol{\theta}_2)\,]\mathbf{Q}_R\,\| \qquad \text{(C.1)}$$

Then there is $\varepsilon > 0$ such that when $\|\,\Sigma_N(\boldsymbol{\theta})\| < \varepsilon$ the parameters are uniquely identified - i.e. if $\boldsymbol{\theta}$ is such that $\|\,\Sigma_N(\boldsymbol{\theta})\,\| < \varepsilon$ and $\Sigma_S + \Sigma_N = \Sigma_S' + \Sigma_N(\boldsymbol{\theta}')$ then $\Sigma_S = \Sigma_S'$ and $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$.

Proof: We begin with a simple lemma:

Lemma: If $\mathbf{A}$, $\mathbf{B}$ are positive semi-definite Hermetian matrices, then $\|\mathbf{A}\|, \|\mathbf{B}\| \leq \|\mathbf{A} + \mathbf{B}\|$.

Proof of Lemma: First note that, for non-negative definite matrices $\mathbf{A}$, $\mathbf{Tr}[\mathbf{A}] \geq 0$. Also, for square matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{Tr}[\mathbf{AB}] = \mathbf{Tr}[\mathbf{BA}]$, so we can write for positive semi-definite Hermetian matrices $\mathbf{A}$, $\mathbf{B}$

$$\mathbf{Tr}[\mathbf{AB}] = \mathbf{Tr}[\mathbf{A}^{1/2}\mathbf{A}^{1/2*}\mathbf{B}] = \mathbf{Tr}[\mathbf{A}^{1/2*}\mathbf{B}\mathbf{A}^{1/2}] \geq 0$$

where the inequality follows from the fact that $\mathbf{A}^{1/2*}\mathbf{B}\mathbf{A}^{1/2}$ is positive semi-definite. Then

$$\|\mathbf{A} + \mathbf{B}\| = (\mathbf{Tr}[(\mathbf{A} + \mathbf{B})(\mathbf{A} + \mathbf{B})^*])^{1/2}$$

$$= (\mathbf{Tr}[\mathbf{A}\mathbf{A}^*] + \mathbf{Tr}[\mathbf{A}\mathbf{B} + \mathbf{B}\mathbf{A}] + \mathbf{Tr}[\mathbf{B}\mathbf{B}^*])^{1/2}$$

$$\geq (\mathbf{Tr}[\mathbf{AA}^*] + \mathbf{Tr}[\mathbf{BB}^*])^{1/2} \geq \|\mathbf{A}\|, \|\mathbf{B}\|$$

Proof of theorem: Fix Σ_S and let $\mathbf{A}$ be any $m \times m$ complex Hermetian matrix. Let $\mathbf{P}_R$, $\mathbf{P}$ denote the projections onto the subspaces spanned by the eigenvectors associated with the largest p eigenvalues of Σ_S and $\mathbf{A}$ respectively. Theorem 1 of chapter 4 implies that the mapping $\mathbf{A} \rightarrow \mathbf{P}$ is continuous (with respect to the norm $\|\cdot\|$). Thus there are δ', $K > 0$ (which depend on Σ_S) such that such that

$$\|\mathbf{P}_R - \mathbf{P}\| < K\delta' \qquad \text{whenever} \qquad \|\Sigma_S - \mathbf{A}\| < \delta' \tag{C.2}$$

Pick

$$\varepsilon_0 < \min[\delta/2,\ \delta'/2M,\ (4M(K^2+K))^{-1}\ ] \tag{C.3}$$

Then suppose

$$\Sigma = \Sigma_S + \Sigma_N(\theta) \qquad\qquad \Sigma = \Sigma_S' + \Sigma_N(\theta') \tag{C.4}$$

and

$$\| \Sigma_N(\theta) \| < \varepsilon \leq \varepsilon_0 \tag{C.5}$$

Since for any projection matrix $\mathbf{P}$ and any matrix $\mathbf{A}$, $\|\mathbf{AP}\|$, $\|\mathbf{PA}\| \leq \|\mathbf{A}\|$, and since

$$\mathbf{Q}_R\Sigma\mathbf{Q}_R = \mathbf{Q}_R\Sigma_N\mathbf{Q}_R = \mathbf{Q}_R(\ \Sigma'_S + \Sigma_N(\theta')\ )\mathbf{Q}_R$$

we have

$$\varepsilon > \| \Sigma_N(\theta) \| > \| \mathbf{Q}_R\ \Sigma_N(\theta)\ \mathbf{Q}_R \| = \| \mathbf{Q}_R\ \Sigma_S'\ \mathbf{Q}_R + \mathbf{Q}_R\ \Sigma_N(\theta')\ \mathbf{Q}_R \|$$

Hence by the lemma

$$\| \mathbf{Q}_R\ \Sigma_N(\theta')\ \mathbf{Q}_R \| < \varepsilon$$

so that

$$\| \mathbf{Q}_R\ [\Sigma_N(\theta) - \Sigma_N(\theta')]\ \mathbf{Q}_R \| < 2\varepsilon$$

and hence,

$$\| \Sigma_S - \Sigma_S' \| = \| \Sigma_N(\theta) - \Sigma_N(\theta') \| = \gamma < 2M\varepsilon < \delta' \tag{C.6}$$

Letting R' be the response space for Σ_S' and applying (C.2), we can write $\mathbf{P}_{R'} = \mathbf{P}_R + \mathbf{A}$ where $\|\mathbf{A}\| < K\gamma$. Let $\mathbf{Q}_{R'} = \mathbf{I}_M - \mathbf{P}_{R'} = \mathbf{Q}_R - \mathbf{A}$. Then, since $\mathbf{Q}_R\ \Sigma_S = \Sigma_S\ \mathbf{Q}_R = \mathbf{Q}_{R'}\ \Sigma_S' = 0$,

$$\mathbf{Q}_{R'}\ \Sigma_N(\theta')\ \mathbf{Q}_{R'} = \mathbf{Q}_{R'}\ \Sigma\ \mathbf{Q}_{R'} = \mathbf{A}\ \Sigma_S\ \mathbf{A} + \mathbf{Q}_{R'}\ \Sigma_N(\theta)\ \mathbf{Q}_{R'}$$

and so, letting $\mathbf{C} = (\Sigma_N(\theta) - \Sigma_N(\theta'))$ we have

$$\mathbf{Q}_{R'}\ \mathbf{C}\ \mathbf{Q}_{R'} = \mathbf{A}\ \Sigma_S\ \mathbf{A} \tag{C.7}$$

Thus

$$\mathbf{Q}_R\ \mathbf{C}\ \mathbf{Q}_R = (\mathbf{Q}_{R'} + \mathbf{A})\ \mathbf{C}\ (\mathbf{Q}_{R'} + \mathbf{A})$$

$$= \mathbf{A}\ \Sigma_S\ \mathbf{A} + \mathbf{Q}_{R'}\ \mathbf{C}\ \mathbf{A} + \mathbf{A}\ \mathbf{C}\ \mathbf{Q}_{R'}\ + \mathbf{A}\ \mathbf{C}\ \mathbf{A} \tag{C.8}$$

Now, for any matrices $\mathbf{A}$, $\mathbf{B}$ the matrix norm $\|\cdot\|$ satisfies $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$ and $\|\mathbf{AB}\| \leq \|\mathbf{A}\|\ \|\mathbf{B}\|$. Hence using (C.3),(C.6) and (C.8) we find

$$\gamma = \|\mathbf{C}\| \leq \|\mathbf{Q}_R\ \mathbf{C}\ \mathbf{Q}_R\| \leq K^2\gamma^2 + 2K\gamma^2 + K^2\gamma^3 \tag{C.9}$$

Suppose $\gamma > 0$. Then, provided we take $\gamma < 1$, (C.9) implies

$$\gamma \geq (2K^2 + 2K)^{-1}$$

But by (C.3) and (C.6)

$$\gamma < 2M\varepsilon < (2K^2 + 2K)^{-1}$$

Hence the only way (C.9) can be satisfied is to have $\gamma = 0$ so that

$$\Sigma_N(\theta) = \Sigma_N(\theta') \qquad\qquad \Sigma_S = \Sigma'_S$$

and we are done.

Appendix D

Covariance Matrices with Linear Structure

The model for the noise covariance matrix used in chapter 4 takes the general form

$$\Sigma_N = \Psi\, \Sigma_\beta\, \Psi^* \;+\; \mathbf{diag}(\, \Sigma_1\, ,\; \cdots\; ,\, \Sigma_n\,) \tag{D.1}$$

where Σ_β is $q \times q$ and the matrices Σ_i, $i = 1, n$ are $l \times l$. In this appendix we show how to write this model as a linear combination of known complex matrices with real coefficients. First, to simplify notation let

$$(\Sigma_\beta)_{jk} = \tau_{jk} \qquad (\Sigma_i)_{jk} = \sigma_{ijk}$$

Let ψ_j be the j^{th} column of Ψ, and let $\mathbf{E}_{ijk}$ be the matrix for which the jk element in the i^{th} diagonal block is 1 and all other elements are 0. Then we can write

$$\Sigma_N = \sum_{jk} \tau_{jk} \psi_j \psi_k^* + \sum_{ijk} \sigma_{ijk} \mathbf{E}_{ijk} \tag{D.2}$$

This gives an expression for the noise covariance matrix as a linear combination of known matrices, but the parameters τ_{jk}, σ_{ijk} are in general complex. Furthermore, because the matrices Σ_β and Σ_i, $i = 1, n$ are Hermetian, these parameters are not all independent (e.g. $\tau_{jk} = \tau_{kj}^*$). We can easily rewrite this so that the parameters are real and independent. If $c = a + ib$ is a complex number and $\mathbf{A} = \mathbf{R} + i\mathbf{P}$ a complex matrix then a simple calculation shows

$$c\mathbf{A} + c^*\mathbf{A}^* = a\,[\,(\mathbf{R} + \mathbf{R}^T) + i(\mathbf{P} - \mathbf{P}^T)\,] + \tag{D.3}$$
$$b\,[\,-(\mathbf{P} + \mathbf{P}^T) + i(\mathbf{R} - \mathbf{R}^T)\,]$$

Terms in the sum of (D.2) with $j \neq k$ occur in conjugate transpose pairs $\tau_{jk}\psi_j\psi_k^* = (\tau_{kj}\psi_k\psi_j^*)^*$. For $j \neq k$ identify $\psi_j\psi_k^*$ with $\mathbf{A}$ in (D.3) and let

$$\mathbf{Y}_{1jk} = [\,(\mathbf{R} + \mathbf{R}^T) + i(\mathbf{P} - \mathbf{P}^T)\,]$$

$$\mathbf{Y}_{2jk} = [\,-(\mathbf{P} + \mathbf{P}^T) + i(\mathbf{R} - \mathbf{R}^T)\,]$$

For $j = k$ let $\mathbf{Y}_{jj} = \psi_j\psi_j^*$. Then the first sum in (D.2) can be written

$$\sum_{jk} \tau_{jk}\psi_j\psi_k^* = \sum_j \tau_{jj}\mathbf{Y}_{jj} + \sum_{j>k} (\,\mathrm{Re}\tau_{jk}\,\mathbf{Y}_{1jk} + \mathrm{Im}\tau_{jk}\,\mathbf{Y}_{2jk}\,)$$

Since the diagonal terms τ_{jj} are real this sum is over the independent q^2 real parameters which make up the elements on and below the diagonal of the parameter matrix Σ_β. The second term in (D.2) is treated in a similar manner. Thus, changing notation slightly, the model for the noise covariance matrix is of the general form

$$\Sigma_N(\boldsymbol{\theta}) = \sum_{i=1}^{r} \theta_i \mathbf{Y}_i \tag{D.4}$$

Theorem 3 shows that if the noise covariance model is a linear model of the form (D.1), then condition (4.4.3) of theorem 2 holds, so that such models have identifiable parametrizations if the signal to noise ratio is large enough. We prove this theorem here.

Theorem 3: Assume $\Sigma_N(\boldsymbol{\theta}) = \sum_{i=1}^{r} \theta_i \mathbf{Y}_i$ where the $\mathbf{Y}_i$ are known matrices. Let $\mathbf{P}$ be any projection matrix. Then, if the r matrices $\mathbf{Y}_i' = \mathbf{P}\mathbf{Y}_i\mathbf{P}$ are linearly independent, there is M such that for all $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$

$$\| \Sigma_N(\boldsymbol{\theta}_1) - \Sigma_N(\boldsymbol{\theta}_2) \| < M \,\| \mathbf{P}[\, \Sigma_N(\boldsymbol{\theta}_1) - \Sigma_N(\boldsymbol{\theta}_2)\,]\mathbf{P} \| \tag{D.5}$$

Proof: We first define two matrix operators which will be useful to us. For any $m \times n$ matrix $\mathbf{A} = (\mathbf{a}_1\ \cdots\ \mathbf{a}_n)$ we will denote the mn dimensional vector consisting of the m columns of $\mathbf{A}$ stacked one on top of the other $\mathbf{vec}(\mathbf{A}) = (\,\mathbf{a}_1^T \mid \cdots \mid \mathbf{a}_n^T\,)^T$. If $\mathbf{A}$ and $\mathbf{B}$ are $m \times n$ and $p \times q$ matrices we can form the $mp \times nq$ matrix

$$\mathbf{A} \times \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix}$$

Then it is easily shown that for any matrices **A**, **B**, **C** of appropriate dimensions

$$\mathbf{vec(ABC)} = [\mathbf{A} \times \mathbf{C}]\ \mathbf{vec(B)} \tag{D.6}$$

Note also that $\|\mathbf{A}\| = \mathbf{Tr}[\mathbf{A}^*\mathbf{A}] = \|\mathbf{vec(A)}\|_2$, where $\|\cdot\|_2$ denotes the usual Euclidean norm.

Then define the m^2-dimensional vectors

$$\mathbf{y}_j = \mathbf{vec}(\mathbf{Y}_j), \quad j = 1, r \qquad\qquad \boldsymbol{\sigma}_i = =\mathbf{vec}(\boldsymbol{\Sigma}_N(\boldsymbol{\theta}_i)) \qquad i = 1, 2$$

and the $m^2 \times m^2$ and $m^2 \times r$ matrices

$$\boldsymbol{\Pi} = \mathbf{P} \times \mathbf{P} \qquad\qquad \boldsymbol{\Gamma} = (\mathbf{y}_1 \cdots \mathbf{y}_r)$$

Since the noise covariance matrices $\boldsymbol{\Sigma}_N(\boldsymbol{\theta}_i)$, $i = 1, 2$ satisfy the linear model (D.4) we have

$$\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2 = \boldsymbol{\Gamma}[\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2] \tag{D.7}$$

Now by (D.6), $\mathbf{vec}(\mathbf{PY}_j\mathbf{P}) = \boldsymbol{\Pi}\mathbf{y}_j$, so the linear independence of the matrices $\mathbf{PY}_j\mathbf{P}$ implies that the $m^2 \times r$ matrix $\boldsymbol{\Pi}\boldsymbol{\Gamma}$ is of full rank r. We may thus solve (D.7) for $(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2)$

$$\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2 = [\boldsymbol{\Gamma}^*\boldsymbol{\Pi}^*\boldsymbol{\Pi}\boldsymbol{\Gamma}]^{-1}\boldsymbol{\Gamma}^*\boldsymbol{\Pi}^*\boldsymbol{\Pi}[\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2]$$

or using (D.7)

$$\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2 = [\ \boldsymbol{\Gamma}(\boldsymbol{\Gamma}^*\boldsymbol{\Pi}^*\boldsymbol{\Pi}\boldsymbol{\Gamma})^{-1}\boldsymbol{\Gamma}^*\boldsymbol{\Pi}^*\ ][\ \boldsymbol{\Pi}(\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2)\ ] \tag{D.8}$$

Letting

$$M = \|\ \boldsymbol{\Gamma}(\boldsymbol{\Gamma}^*\boldsymbol{\Pi}^*\boldsymbol{\Pi}\boldsymbol{\Gamma})^{-1}\boldsymbol{\Gamma}^*\boldsymbol{\Pi}^*\ \|$$

and using the fact that $\|\mathbf{AB}\| \le \|\mathbf{A}\|\ \|\mathbf{B}\|$ together with (D.6) we have

$$\|\ \boldsymbol{\Sigma}_N(\boldsymbol{\theta}_1) - \boldsymbol{\Sigma}_N(\boldsymbol{\theta}_2)\ \| = \|\ \boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2\ \| \le M\ \|\ \boldsymbol{\Pi}(\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2)\ \|$$

$$= M\ \|\ \mathbf{P}\ (\boldsymbol{\Sigma}_N(\boldsymbol{\theta}_1) - \boldsymbol{\Sigma}_N(\boldsymbol{\theta}_2))\ \mathbf{P}\ \|$$

and we are done.