LBNL-42397

ERNEST ORLANDO LAWRENCE
BERKELEY NATIONAL LABORATORY

Intersection, Ideals, and Inversion

D.W. Vasco

Earth Sciences Division

October 1998

REFERENCE COPY
Does Not
Circulate
Bldg. 50 Library - Ref.
Lawrence Berkeley National Laboratory
Copy 1
LBNL-42397

LBNL-42397

INTERSECTION, IDEALS, AND INVERSION

D. W. Vasco

Earth Sciences Division
Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720

October 1998

This work was supported by a Laboratory Directed Research and Development grant, Office of Energy Research, Division of Basic Energy Sciences, Engineering, and Geosciences, of the U.S. Department of Energy under contract No. DE-AC03-76SF00098.

Intersections, ideals, and inversion

D. W. Vasco*

(October 7, 1998)

ABSTRACT

Techniques from computational algebra provide a framework for treating large classes of inverse problems. In particular, the discretization of many types of integral equations and of partial differential equations with undetermined coefficients lead to systems of polynomial equations. The structure of the solution set of such equations may be examined using algebraic techniques. For example, the existence and dimensionality of the solution set may be determined. Furthermore, it is possible to bound the total number of solutions. The approach is illustrated by a numerical application to the inverse problem associated with the Helmholtz equation. The algebraic methods are used in the inversion of a set of transverse electric (TE) mode magnetotelluric data from Antarctica. The existence of solutions is demonstrated and the number of solutions is found to be finite, bounded from above at 50. The best fitting structure is dominantly one-dimensional with a low crustal resistivity of about 2 ohm-m. Such a low value is compatible with studies suggesting lower surface wave velocities than found in typical stable cratons.

*Center for Computational Seismology, Berkeley Laboratory, University of California, Berkeley, CA 94720

INTRODUCTION

The complete solution of an inverse problem requires addressing questions concerning the existence and uniqueness of an Earth model compatible with a given set of observations. In this regard linear inverse problems have been treated rather successfully. There are now well established methods for calculating model parameter resolution and uncertainty. Much of the power of these techniques derives from formalisms developed in linear algebra and functional analysis (Halmos 1957, Luenberger 1969, Dorny 1980). Mathematical structures in these areas, in particular vector and Hilbert spaces, provide a framework for answering the questions posed when solving inverse problems (Parker 1994). To date, we have no direct link between non-linear inverse problems and any such algebraic formalism. To be sure, there are extensive numerical techniques for finding solutions as well as methods for linearization. But such methods are usually local in nature or do not fully utilize the mathematical structure of the equations defining the inverse problem.

What I wish to demonstrate in this paper is that there is a body of mathematical work providing a framework for treating many non-linear inverse problems. Specifically, techniques from algebraic geometry and computational algebra can answer questions concerning the existence and uniqueness of solutions to non-linear inverse problems. Algebraic structures such as polynomial ideals and their geometric counterparts, algebraic varieties, are useful for describing the solution set of an inverse problem. To a significant degree these methods are generalizations of techniques used in linear algebra. For example, there are procedures akin to Gaussian elimination applicable to polynomial systems of equations. It is possible to construct an algebraic basis, similar to the basis of a vector space, associated with the solution set of a non-linear inverse problem. As shown below, the methods are global and do not depend on linearization about some initial model. Furthermore, they are enumerative, providing information about the dimension of the solution set. In the case of a zero-dimensional solution set consisting of a finite number of distinct solutions, one may bound the number of solutions to the inverse problem. Finally, the approach outlined here may be used in the presence of noisy data and in conjunction with regularization penalty terms. This work was motivated by Everett's (1996) treatment of inverse problems in the form of polynomial equations and by a desire to gain a greater understanding of the solution sets of such equations. To this end I consider the under-determined case in which there are an infinite number of solutions as well as the situation in which the solution set is empty.

METHODOLOGY

Many important inverse problems may be formulated in terms of functionals or as partial differential equations with undetermined coefficients. In the course of the numerical solution of these kinds of inverse problems one is usually compelled to discretize the defining equations. The methods described here assume that such a discretization has been performed. As pointed out by Everett (1996), when considered as inverse problems, the discretized differential equations are often in the form of systems of polynomial equations. Under certain assumptions discretized functionals also may be put in the form of polynomial equations. Such equations are the domain of algebraic geometry, commutative algebra, and computational algebra. Much insight into inverse problems may be gained from techniques developed in these areas. Furthermore, explicit computational procedures have been developed to address the existence, enumeration, and dimensionality of solutions to systems of polynomial equations. In this section we review the concepts that are key to understanding these useful algorithms. Algebraic geometry and commutative algebra are subjects of great depth, and it is not feasible nor desirable to provide an extensive review here. I merely wish to highlight the most important concepts in the context of solving inverse problems.

Governing Equations and Discretization

Functionals and Integral Equations First, consider an inverse problem in which the observed data d_i are functionals of the Earth's structure

$$d_i = \int_\Omega f(\mathbf{x}_i{}^s, \mathbf{x}, \sigma(\mathbf{x}))d\mathbf{x}, \tag{1}$$

where $\mathbf{x}_i{}^s$ denotes the location of the *ith* station, the vector $\mathbf{x}$ denotes position within the Earth, $\sigma(\mathbf{x})$ is the unknown function describing the spatial variation in a material property such as density, and the integration is over some volume of interest Ω. In general, the integrand will be a non-linear function of the unknown σ. A direct discretization of the integral (1) comes from evaluating it as a sum over a set of lattice points. In fact, there are accurate lattice summation techniques for the numerical evaluation of integrals (Sloan and Joe 1994). In its basic form the integral appears as a sum over the function value at N points of an integration lattice,

$$d_i = \frac{1}{N}\sum_{j=0}^{N-1} f(\mathbf{x}_i{}^s, \mathbf{x}_j, \sigma_j), \tag{2}$$

where $\sigma_j = \sigma(\mathbf{x}_j)$ is the material property at lattice point j.

Many functionals and integral equations occurring as inverse problems reduce to systems of polynomial equations when they are discretized. In particular, many integral equations which are derived from differential equations of the form

$$L[u(\mathbf{x},\omega)] + \omega\sigma(\mathbf{x})u(\mathbf{x},\omega) = -s(\mathbf{x},\omega)$$

where L signifies a self-adjoint, second-order, linear operator, $u(\mathbf{x},\omega)$ denotes the wavefield, $\sigma(\mathbf{x})$ signifies the material property variation in the subsurface, ω is the frequency, and $s(\mathbf{x},\omega)$ is the source term. Using a Green's function and the representation theorem, it is possible to derive an equivalent integral equation (Stakgold 1979). Discretization of the equation produces a system of polynomial equations in the unknown internal field and material property variables.

As an example consider the Helmholtz equation, which represents the wave-equation in the frequency domain,

$$\nabla^2 u + \omega^2\sigma(\mathbf{x})u = -s(\mathbf{x},\omega). \tag{3}$$

This equation occurs frequently in acoustic (Miller et al. 1987) and electro-magnetic wave propagation (Hohmann 1987). Consider the material property variation as a sum of some specified prior structure $\sigma_0(\mathbf{x})$ and a deviation (perhaps a large deviation) from the prior structure $\sigma'(\mathbf{x})$

$$\sigma(\mathbf{x}) = \sigma_0(\mathbf{x}) + \sigma'(\mathbf{x}). \tag{4}$$

Often we may choose the prior structure such that it is possible to construct a Green's function $G(\mathbf{x},\mathbf{y},\omega)$, which satisfies

$$\nabla^2 G + \omega^2\sigma_0(\mathbf{x})G = -\delta(\mathbf{x}-\mathbf{y}). \tag{5}$$

Substituting equation (4) into equation (3) and making use of equation (5) we may construct an integral equation for $u(\mathbf{x},\omega)$

$$u(\mathbf{x},\omega) = \int G(\mathbf{x},\mathbf{y},\omega)\left[s(\mathbf{y},\omega) + \omega^2\sigma'(\mathbf{y})u(\mathbf{y},\omega)\right] d\mathbf{y}.$$

Noting that the first term on the right is simply the wavefield predicted by the prior model, the reference wavefield which is denoted $u_0(\mathbf{x},\omega)$, we may write

$$u(\mathbf{x}_s,\omega) = u_0(\mathbf{x}_s,\omega) + \omega^2\int G(\mathbf{x}_s,\mathbf{y},\omega)\sigma'(\mathbf{y})u(\mathbf{y},\omega)d\mathbf{y}$$

where we have taken $\mathbf{x}$ to be the location of a station $\mathbf{x}_s$. When this equation is discretized at the N points of an integration lattice it becomes a quadratic polynomial in u and σ' at the lattice points. Given a set of M observations we arrive at M polynomial equations in approximately $2N$ unknowns (boundary conditions will reduce the number of unknowns somewhat). Regularization, as described below, will introduce additional polynomial constraint equations. An alternative discretization involves expanding σ' in a series of rectangular basis functions. If the blocks are taken small enough we may assume that u is essentially constant in each cell. We could then solve for the average value of u in each cell as well as the expansion coefficients.

For more complicated non-polynomial integrands we might resort to a representation of the function f in a power series expansion in σ_j, truncated to order l, we have

$$d_i = \frac{1}{N}\sum_{j=0}^{N-1}\sum_{0}^{l} c_{ijl}{\sigma_j}^l \tag{6}$$

where c_{ijl} are the expansion coefficients. The result is a set of polynomial equations for the unknown material properties σ_j at each lattice point. The power series expansion introduces an additional approximation which limits the range of applicability of this particular technique. Because of this limitation, the range of applicability should be determined for each specific inverse problem.

Differential Equations Next, consider an inverse problem in the form of a partial differential equation with variable but unknown coefficients. For the most part, the governing equations we employ are linear partial differential equations (Menke and Abbott 1990), though this is not absolutely necessary. In the case of linear partial differential equations we have the form

$$\frac{\partial^k u}{\partial t^k} = L(u,\nabla u), \tag{7}$$

where $L(u,\nabla u)$ is a linear function that depends on some combination of u and its spatial derivatives as well as on a set of unknown coefficients. The kth time derivative on the left may be removed by a Fourier or Laplace transform, thus introducing a transform variable such as ω or s.

As an example of an equation in the transformed domain, consider the two-dimensional modified Helmholtz equation associated with the interpretation of magnetotelluric (MT) data, for the transverse electric (TE) mode,

$$\nabla^2 u + i\omega\mu_0\sigma(x,y)u = 0, \tag{8}$$

where μ_0 is the magnetic permeability of free space, ω is the angular frequency, $\sigma(x,y)$ is the conductivity distribution in space, and u is the electric field associated with a steady state monochromatic ($e^{i\omega t}$) excitation (Vozoff 1985). The existence question for electromagnetic induction has been answered for the one-dimensional inverse problem (Parker and Whaler 1981), but is still an open question for full three-dimensional conductivity variations. We shall treat this equation in some detail is this paper, using it to illustrate the concepts we introduce. The forward problem associated with this partial differential equation is to find the electric field u at a set of 'observation' points given the distribution of material

properties, in this case the conductivity $\sigma(x, y)$ and the necessary boundary values. The inverse problem entails finding the $\sigma(x, y)$ given a set of observations, usually at discrete points on the surface.

The approach we shall take is to discretize the governing differential equation and consider the forward and inverse problem for a lattice of points. There are a number of ways to discretize continuous equations ranging from finite difference, finite element (Mikhlin 1979), and lattice gas methods (Doolen et al. 1990), to direct formulation on a point lattice (Toda 1981). We follow the approach of Everett (1996) based upon finite differences, in which the spatial derivatives are approximated as numerical differences. Furthermore, we shall assume that the time derivative has been removed either by transformation or by considering a field u with the time variations $e^{i\omega t}$. As described in Everett (1996) our finite difference representation of the linear partial differential equation takes the form of a linear system of equations

$$A\mathbf{u} = \mathbf{b}. \tag{9}$$

As an example, consider the magnetotelluric equation (8), where A is an $N \times N$ block tridiagonal matrix of the form

$$A = \begin{pmatrix} \mathbf{A}_1 & \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{I} & \mathbf{A}_2 & \mathbf{I} & \cdots & \mathbf{0} \\ \cdot & \cdot & & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & & \cdot & \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I} & \mathbf{A}_N \end{pmatrix} \tag{10}$$

$\mathbf{I}$ and $\mathbf{0}$ are $N \times N$ identity and zero submatrices, respectively. There are N rows and N columns in the finite difference grid. The submatrices $\mathbf{A}_i, i = 1, ..., N$ are tridiagonal matrices of the form

$$\begin{pmatrix} -4 + i\omega\mu_0 h^2\sigma_{1i} & 1 & 0 \\ 1 & -4 + i\omega\mu_0 h^2\sigma_{2i} & 0 \\ \cdot & & \cdot \\ \cdot & \cdot & \\ \cdot & & 1 \\ 0 & \cdot & -4 + i\omega\mu_0 h^2\sigma_{Ni} \end{pmatrix}$$

where the σ_{jk} denote nodal conductivities $\sigma(x, y)$ for row j and column k, and h is the distance between finite difference nodes. As noted by Everett (1996) the right hand side of equation (9) also may be written in a partitioned form $\mathbf{b} = (\mathbf{b}_1, ..., \mathbf{b}_N)^T$ where the subvectors $\mathbf{b}_i$ are composed of combinations of force terms (if present) and Dirichlet and, possibly, Neumann boundary values, including the surface measurements. We also may write the electric field vectors $\mathbf{u}$ as $(\mathbf{u}_1, ..., \mathbf{u}_N)^T$ where each subvector $\mathbf{u}_i = (u_{1i}, ..., u_{Ni})^T$ contains the field components for the ith layer of the mesh. As in Everett (1996) the $0th$ layer of the mesh denotes the lower boundary and the Nth layer signifies the surface. The constraints on the interior values of both $\sigma(x, y)$ and $u(x, y)$ are provided by the boundary conditions, such as far-field conditions, and measured surface values in row N. For example, we might have observations for all surface nodes, all elements of $\mathbf{u}_N$, for several frequencies $\omega = \omega_1, ..., \omega_m$.

To solve the inverse problem we may work downwards, solving for $\mathbf{u}_{N-1}$ in terms of the boundary and measured values. Then, using back-substitution in equation (9), we solve successively for $\mathbf{u}_{N-2}, \mathbf{u}_{N-3}...$ until $\mathbf{u}_1$. The final equation is a set of N polynomial equations in terms of the data, boundary values, and unknown nodal conductivities.

As an illustration and to motivate the ideas presented in this paper we shall consider the special case $N = 2$ in some detail. This is the same situation treated in Everett (1996). The field vector is $\mathbf{u} = (u_1, u_2, u_3, u_4)^T$ where u_1, u_2 are unknown interior field values while u_3, u_4 are fixed by surface Dirichlet and Neumann observations d_1 and d_2. The entire system consists of four polynomial equations in six unknowns $u_1, u_2, \sigma_1, \sigma_2, \sigma_3$, and σ_4 (Everett 1996):

$$\begin{aligned} i\omega\mu_0 h^2 u_1\sigma_1 - 4u_1 + u_2 + (d_1 - b_1) &= 0 \\ i\omega\mu_0 h^2 u_2\sigma_2 + u_1 - 4u_2 + (d_2 - b_2) &= 0 \\ i\omega\mu_0 h^2 d_1\sigma_3 + u_1 + (-4d_1 + d_2 - b_3) &= 0 \\ i\omega\mu_0 h^2 d_2\sigma_4 + u_2 + (-4d_2 + d_1 - b_4) &= 0. \end{aligned}$$

If these equation were linear we would have techniques from linear algebra and linear inverse theory at our disposal to determine the existence and the uniqueness of the solutions. Furthermore, we could introduce constraints or penalty terms and construct solutions to the regularized inverse problem. However, the equations are polynomial due to the presence of the product terms $u_i\sigma_i$. Therefore, we must find other means to address the existence and uniqueness issues.

Numerical Lattices and Affine Varieties

As shown above, starting from functionals, integral equations, or differential equations we often arrive at a system of polynomial equations in the field variables and material properties. Such equations are highly structured and have been intensively studied (Kendig 1977, Dieudonne 1985, Mishra 1993). In this paper I shall examine some of the techniques which are applicable to inverse problems. To state the ideas and results efficiently we shall need a uniform and compact notation for polynomials in several variables $z = (z_1, ..., z_n)$. A polynomial consists of a sum of terms

$$f = \sum_\alpha a_\alpha z^\alpha, \qquad a_\alpha \in k,$$

where k is a number field, α is a vector of integer indices, $\alpha = (\alpha_1, ..., \alpha_n)$, and the monomial term z^α is given by

$$z^\alpha = {z_1}^{\alpha_1} \cdot {z_2}^{\alpha_2} \cdots {z_n}^{\alpha_n}.$$

The degree of the monomial term is denoted $|\alpha| = \alpha_1 + \cdots + \alpha_n$ and the total degree of the polynomial f, signified by $\deg(f)$, is the maximum $|\alpha|$ such that the coefficient, a_α, is nonzero. The set of all polynomials in the variables $z_1, ..., z_n$ with coefficients in the number field k is symbolized by $k[z_1, ..., z_n]$. In modern treatments the field k is usually unspecified. Here we shall always work with either the real ($\mathbf{R}$) or complex ($\mathbf{C}$) number fields. The advantage of working with $\mathbf{C}$ is that it is a closed field, any polynomial equation with complex coefficient has a solution in the field of complex numbers (Gallian 1990, p. 305). In all that follows we shall assume that k is algebraically closed, we shall take $\mathbf{C}$ as the number field. As an example, let us re-write the unknowns in the system of equations as $z_1 = u_1, z_2 = u_2, z_3 = \sigma_1, z_4 = \sigma_2, z_5 = \sigma_3$, and $z_6 = \sigma_4$ and scale the conductivities σ_1, σ_2 by $i\omega\mu_0 h^2$, σ_3 by $i\omega\mu_0 h^2 d_1$, and σ_4 by $i\omega\mu_0 h^2 d_2$, so that the equations become

$$\begin{aligned} z_1 z_3 - 4z_1 + z_2 + c_1 &= 0 \\ z_2 z_4 + z_1 - 4z_2 + c_2 &= 0 \\ z_5 + z_1 + c_3 &= 0 \\ z_6 + z_2 + c_4 &= 0 \end{aligned}$$

where $c_1 = (d_1 - b_1)$, $c_2 = (d_2 - b_2)$, $c_3 = (-4d_1 + d_2 - b_3)$, and $c_4 = (-4d_2 + d_1 - b_4)$.

A central concern in any inverse problem is the structure of the solution set. Is there a unique solution, a finite set of solutions, or an infinity of solutions? Alternatively, if there is no solution, can this be determined from the set of equations? For systems of polynomial equations, the primary object of study is the geometric quantity defined by the zero set in the affine space k^n,

Definition An *affine variety* associated with the set of p polynomials $f_1, ..., f_p$ in n variables, denoted $\mathbf{V}(f_1, ..., f_p)$, is the set of points $(\zeta_1, ..., \zeta_n) \in k^n$ for which $f_i(\zeta_1, ..., \zeta_n) = 0$ for all $1 \geq i \geq p$.

(Kendig 1977, p. 1). Intuitively, affine varieties are defined by the intersection of a number of hyper-surfaces defined by polynomial equations. In our example there are $p = 4$ equations in $n = 6$ unknowns, with the polynomials given by

$$\begin{aligned} f_1 &= z_1 z_3 - 4z_1 + z_2 + c_1 \\ f_2 &= z_2 z_4 + z_1 - 4z_2 + c_2 \\ f_3 &= z_5 + z_1 + c_3 \\ f_4 &= z_6 + z_2 + c_4. \end{aligned}$$

Affine varieties consist of collections of points in affine space k^n, which may define individual points, lines, surfaces, or hypersurfaces, or some such combination. In general, affine varieties may kink and self-intersect and hence do not constitute differentiable manifolds because the tangent vector at a point may not be unique (Auslander and MacKenzie 1963). However, as we shall see, because they are specified by polynomials these sets of points have a rich algebraic structure. For simplicity we shall only treat the case of affine space, which may be thought of as an n-dimensional normed space such as $\mathbf{R}^n$ or $\mathbf{C}^n$. There are some advantages working in projective space (Kendig 1977), where it is easier to handle points at infinity, but such an added complication will not be considered here.

Ideals and the Ideal-Variety Correspondence

The most important algebraic structure associated with solution sets of polynomial equations are ideals (Gallian 1990, p. 210). In terms of polynomials an ideal has three distinguishing properties (Cox et al. 1997, p. 29) as indicated by the following definition.

Definition A subset of all polynomials, denoted by $I \subset k[z_1, ..., z_n]$, is an *ideal* if $0 \in I$, and for $f, g \in I$ and $h \in k[z_1, ..., z_n]$ both $f + g$ and $h \cdot f$ are in I.

The property that the product of an element in the ideal I with any member of the space $k[z_1, ..., z_n]$ (i.e. any polynomial in the variables $z_1, ..., z_n$) results in an element in the ideal I may be new to those familiar with vector spaces. For our purposes one of the most useful realizations of this abstract definition is the ideal generated by a set of polynomials, that is, the sum of all possible products of the defining polynomials $f_1, ..., f_p$ with all other polynomials $h_i \in k[z_1, ..., z_n]$

$$\langle f_1, ..., f_p \rangle \equiv \left\{ \sum_{i=1}^{p} h_i f_i \right\}.$$

Note that such sums vanish on the variety defined by $f_1 = 0, ..., f_n = 0$ as well as at the zeros of the polynomials h_i. Thus, the ideal may be thought of as the consequence of the equations $f_i = 0$, $i = 1, ..., p$ on the space of all possible polynomials. An ideal is said to be finitely generated if $f_1, ..., f_p$ is a finite set; these polynomials are the *generators* of the ideal. In many respects an ideal is similar to a subspace in linear algebra. For example, an ideal is closed under addition and multiplication by polynomials (which take the place of scalars), and finitely generated ideals are similar to the span of a finite set of vectors. Carrying the analogy further, it may be shown that a variety depends only on the ideal generated by its defining equations, not on the particular equations. This is akin to the fact that a subspace in linear algebra does not depend on the particular set of basis vectors used for its representation. The formal theorem for ideals is

Theorem *If* $\langle f_1, ..., f_p \rangle = \langle g_1, ..., g_q \rangle$, *then* $\mathbf{V}(f_1, ..., f_p) = \mathbf{V}(g_1, ..., g_q)$

Thus, as in linear algebra, we may change the basis set in certain ways without affecting the variety, the zero set.

This fact is key to our ability to manipulate the defining equations into a desired form. In our example, the set

$$\begin{aligned} -z_3z_5 + z_2 - c_3z_3 + 4z_5 + c_1' &= 0 \\ z_4z_6 + z_1 - c_4z_4 + 4z_6 + c_2' &= 0 \\ z_5 + z_1 + c_3 &= 0 \\ z_6 + z_2 + c_4 &= 0, \end{aligned}$$

where $c_1' = c_1 + 4c_3$ and $c_2' = c_2 + 4c_4$, defines the same variety as the original system of equations.

An ideal that is central to affine varieties and hence the solution set of a system of polynomial equations is

Definition The *ideal of a variety* $\mathbf{V} \subset k^n$ is the set

$$I(\mathbf{V}) = \{f \in k[z_1, ..., z_n]\}$$

such that $f(\zeta_1, ..., \zeta_n) = 0$ for all $(\zeta_1, ..., \zeta_n) \in \mathbf{V}$.

Simply stated, the ideal of a variety $\mathbf{V}$ is the set of polynomials that vanish at every point of $\mathbf{V}$. A proof that $I(\mathbf{V})$ is in fact an ideal is given in Cox et al. (1997, p. 32). The most relevant properties of the ideal of a variety are that $\langle f_1, ..., f_p\rangle \subset I(\mathbf{V})$, but equality, $\langle f_1, ..., f_p\rangle = I(\mathbf{V})$, does not necessarily hold. Consider our example: the variety defined by $I = \langle f_1, f_2, f_3, f_4\rangle$, the set of zeros is identical to that of $I^2 = \langle {f_1}^2, {f_2}^2, {f_3}^2, {f_4}^2\rangle$, while the ideals differ (I^2 may not contain f_1 for example). To recover equality we must consider the possibility of such 'roots' of our defining polynomials, which leads to the notion of the radical of an ideal (Cox et al. 1997, p. 173).

Definition An ideal I is *radical* if $f^m \in I$ for any integer $m \geq 1$ implies that $f \in I$.

Given an arbitrary ideal I we may introduce notation symbolizing the derivation of a radical ideal from I, the ideal containing the 'roots' of any polynomials of the form f^m,

Definition For an ideal $I \subset k[z_1, ..., z_p]$ the *radical* of I, denoted by $\sqrt{I}$, is the set

$$\{f : f^m \in I\}$$

for some integer $m \geq 1$.

Algorithms are available for extracting the radical of a given ideal (Gianni et al. 1989, Eisenbud et al. 1992, Becker and Weispfenning 1993) and determining if a polynomial is a member of the radical of a particular ideal (Cox et al. 1997, p. 176). Radical ideals are all that is required to produce a one-to-one correspondence between affine varieties (zero sets) and an ideal generated by the defining polynomial equations. That is, for any ideal we have $\mathbf{V}(I(\mathbf{V})) = \mathbf{V}$, but for radical ideals we also have that $I(\mathbf{V}(I)) = I$. This is formalized by a landmark theorem of Hilbert's, a vast generalization of the fundamental theorem of algebra to multi-dimensional spaces (Kendig 1977,p. 124).

Theorem (Hilbert's Nullstellensatz) *For $f, f_1, ..., f_p \in k[z_1, ..., z_n]$, we have $f \in I(\mathbf{V}(f_1, ..., f_p))$, if and only if there exists an integer $m \geq 1$ such that $f^m \in \langle f_1, ..., f_p\rangle$.*

In light of this theorem the investigation of varieties, solution sets of polynomial equations, may be framed in terms of the algebra of ideals. As shown in the next three subsections, this shift of emphasis generates techniques of considerable power.

Ideal Bases and Gröbner Bases

The importance of a basis for a polynomial ideal is tied to our ability to change such bases and carry out calculations using the elements of the basis generating set. As with vector spaces, the question of the size of the basis set determines the practically of the computations. If we restrict our attention to polynomials in a finite number of variables the following theorem by Hilbert (Kendig 1977, p. 119, Cox et al. 1997, p. 74) assures us that basis set calculations are at least finite in length.

Theorem (Hilbert's Basis Theorem) *Every ideal $I \subset k[z_1, ..., z_n]$ has a finite generating set.*

Thus, in describing our solution set we only need a finite set of generators. As shown in the previous section, equivalent sets of generators (generators producing identical ideals) produce the same variety (zero set). So there is a certain freedom in adopting a particular generating set of polynomials. Are some generating sets more desirable than others? As we shall soon see, the Gröbner bases (Buchberger 1985) have many useful properties when it comes to questions of the existence of solutions to polynomial systems of equations and addressing the issue of non-uniqueness.

In simplest terms, a Gröbner basis provides an algorithmic means to determine if a system of polynomial equations is solvable and to estimate the dimension of the solution set. If the solution set is 0-dimensional and consists of a finite number of points, a Gröbner basis may be used to define the total degree of the polynomial system. This information may be used to bound the number of solutions to the inverse problem. Finally, using this special basis we may perform something akin to Gaussian elimination upon polynomial systems of equations. That is, using this basis set we may derive a 'triangular' set of polynomial equations in which the final equations are in the fewest number of variables (Mishra 1993). In the optimal case the final equation is a polynomial equation in a single variable, say z_n. This equation is solved for all values of z_n and these values are successively back-substituted into the previous equation, which is solved for z_{n-1}, and so on.

An essential component of the reduction or systematic manipulation of the defining equations of a variety is an ordering of the monomial terms in the set of polynomials.

For example, if we wish to perform some type of elimination of variables we must work systematically from z_1 to z_2 and so on, until the final variable z_n. Such an ordering is standard in algorithms in linear algebra and necessary for systematic computation. When considering systems of polynomial equations attention must also be paid to the total degree of each term in the polynomial as well as to the variable ordering. In our illustrative example, z_1 might be considered 'larger' than z_2. However, is $z_2 z_4$ 'larger' than z_5? The idea of monomial orderings is to construct a systematic ordering of the terms in multivariate polynomials. That is, given the vector of integer indices, $\alpha = (\alpha_1, ..., \alpha_n)$ introduced above, we must be able to order any two such vectors as greater than, less than, or equal to. In Appendix A two such monomial orderings are described. The first, lexicographic ordering is a dictionary type ordering in which the degree of the left-most variable, such as z_1 takes precedence over all others. In graded lexicographic ordering, the other type considered here, total degree dominates. In the case of monomials of equal total degree we use lexicographic ordering.

The final ingredient needed to define a Gröbner basis is the notion of an ideal of leading terms. For a particular polynomial $f = \sum_\alpha a_\alpha z^\alpha$ the leading term, denoted $\mathrm{LT}(f)$ is that monomial $a_\alpha z^\alpha$ with the greatest index vector α (with respect to the monomial ordering, such as lexicographic or graded lexicographic). In other words, $\mathrm{LT}(f)$ is the 'largest' or 'leading' term of the polynomial. For the illustrative system of polynomials, the leading terms of f_1, f_2, f_3, f_4 are $z_1 z_3, z_2 z_4, z_1$, and z_2 with respect to graded lexicographic ordering. For any ideal I we can construct the ideal of leading terms which is defined as follows.

Definition The *ideal of leading terms*, denoted by $\langle \mathrm{LT}(I) \rangle$, is the ideal generated by the set of leading terms of the elements of I (symbolized by $\mathrm{LT}(I)$).

One fundamental property of the ideal of leading terms $\langle \mathrm{LT}(I) \rangle$ is that it is a monomial ideal, an ideal generated entirely by monomials rather than by the more complicated polynomials. Another critical fact is that there are elements $g_1, ..., g_l \in I$ such that $\langle \mathrm{LT}(I) \rangle = \langle \mathrm{LT}(g_1), ..., \mathrm{LT}(g_l) \rangle$ (Cox et al. 1997, p. 73). Thus, the criterion for an ideal to be a Gröbner basis is based upon the following property of its leading terms:

Definition (Gröbner Basis) A finite subset $G = \{g_1, ..., g_l\}$ of an ideal I is a *Gröbner basis* if

$$\langle \mathrm{LT}(g_1), ..., \mathrm{LT}(g_l) \rangle = \langle \mathrm{LT}(I) \rangle.$$

Note that this is not a trivial statement. Due to possible cancellation, the ideal $\langle \mathrm{LT}(I) \rangle$ may contain generators contributed by lower degree terms from $g_1, ..., g_l$, monomials other than the those in $\langle \mathrm{LT}(g_1), ..., \mathrm{LT}(g_l) \rangle$. The S-polynomials, used to quantify such cancellations in a basis set, are introduced in Appendix B.

It has been established that every non-zero ideal I has a Gröbner basis and that any Gröbner basis is a valid basis for I (generates I) (Mishra 1993, p. 49). Appendix B outlines how a Gröbner basis is derived from a defining set of polynomial equations. A specific procedure, Buchberger's algorithm, is given for calculating the basis elements. There are now numerous commercial (axiom, maple, macsyma, mathematica, reduce, etc.) and public domain (cocoa, macaulay, mas, magma, singular etc.) software packages, primarily computer algebra software, that compute Gröbner bases and perform other tasks in computational ideal theory. For descriptions and information on accessing these routines refer to Adams and Loustaunau (1994), Eisenbud (1995), and Cox et al. (1997).

Unlike linear systems, the size and complexity of an ideal basis depends on the monomial ordering adopted. For example, lexicographic Gröbner bases are often much more difficult to compute than others such as graded orders. If the ideals are 0-dimensional, corresponding to varieties defined by a finite set of points, there is an efficient technique for converting a Gröbner basis with respect to some order into a lexicographic order (Faugere et al. 1993). An advantage of the approach is that it only requires methods from linear algebra. The next two sections discuss two important uses of ideal bases. As we shall see, different monomial orderings may be used to examine various aspects of a solution set.

Existence: Elimination Ideals and Solvability

In this subsection ideal bases are used to address the existence question. That is, does the defining system of polynomial equations have any solutions at all? We also explore how the Gröbner basis algorithm in some sense generalizes Gaussian elimination. Finally, we describe a more numerical approach to solving polynomial systems of equations.

First consider the question of the existence of a solution. It turns out that the existence question is simply stated in terms of the Gröbner basis G of the ideal I:

Theorem *Given polynomials $f_1, ..., f_p$ there are no solutions to the system $f_1 = 0, ..., f_p = 0$ in k^n if and only if $1 \in G$*

(Adams and Loustaunau 1994). The proof of this follows directly from a weak form of Hilbert's Nullstellensatz presented earlier. This says that if we compute the Gröbner basis, with respect to any monomial ordering and find that it contains the constant 1, no solution exists.

In order to actually calculate solutions we must examine the notion of eliminating a variable in more detail. There are two processes involved in solving a system of equations in many variables. First, one must isolate a particular variable, say z_n, in a single equation. Next, having found a value of z_n that solves the single equation

we must extend the solution to the full system. That is, the value is back-substituted into the full system of equations and the reduced set is used to find z_{n-1}. The procedure is repeated until all values of $z_1, ..., z_n$ are found. In what follows only elimination will be examined. For details on solution extension the reader may consult more complete sources. A criterion for the extension of a solution is given in Mishra (1993) for example.

Given a polynomial ideal I, the elimination of a variable z_i consists of finding the intersection of I with the space of all possible polynomials in z_i. It may be proven that this intersection is in fact an ideal (Mishra 1993, p. 137) known as the *lth* elimination ideal.

Definition Given an ideal $I \subset k[z_1, ..., z_n]$ the *lth elimination ideal*, denoted I_l is the ideal $k[z_{l+1}, ..., z_n]$ defined by

$$I_l = I \cap k[z_{l+1}, ..., z_n].$$

The elimination ideal I_l consists of all consequences of $f_1 = 0, ..., f_p = 0$ that eliminate the variables $z_1, ..., z_l$. Translated into the algebra of ideals we find that eliminating $z_1, ..., z_l$ means finding nonzero polynomials in the *lth* elimination ideal I_l.

The primary computational question is: how do we find polynomials in the elimination ideal of I? It turns out that Gröbner bases, coupled with lexicographic ordering, automatically provide the needed polynomials (Adams and Loustaunau 1994, p. 65). In particular, such an ordering produces a 'triangular' system of polynomial equations from which the we may immediately determine if elements of the elimination ideal exist and may be extended.

Theorem *Let I be a 0-dimensional ideal and $G = \langle g_1, ..., g_m \rangle$ its reduced Gröbner basis with respect to lexicographic ordering with $z_1 < z_2 < ... < z_n$. Then we can order $g_1, ..., g_m$ such that g_1 contains only the variable z_1, g_2 contains only the variables z_1 and z_2, g_3 contains only the variables z_1, z_2, and z_3, and so forth until g_m.*

Thus, when a finite solution set exists, a Gröbner basis with respect to lexicographic ordering produces an ordered set of equations in which the first equation contains only z_1, the second equation contains only z_1 and z_2, and so on. As noted above, if there is no solution set the Gröbner basis will contain the constant $\{1\}$ as well. If the solution set is not finite, the Gröbner basis with respect to lexicographic ordering will not contain an equation solely in z_1. Rather, the first equation will contain z_1 and a subset of $z_2, ..., z_n$. We can illustrate these ideas by considering our example problem again. A Gröbner basis with respect to pure lexicographic ordering is identical to the basis G derived in Appendix B using graded lexicographic ordering:

$$g_1 = z_1 + z_5 + c_1{}'$$
$$g_2 = z_2 + z_6 + c_2{}'$$
$$g_3 = z_4 z_6 + c_4 z_4 + z_5 - 4z_6 + c_3{}'$$
$$g_4 = z_3 z_5 + c_3 z_3 - 4z_5 + z_6 + c_4{}'.$$

Note that the final two equations do not contain the variables z_1 and z_2 which are the electric field variables u_1 and u_2. Thus, the final two equations are relationships that contain only the conductivities $\sigma_1, \sigma_2, \sigma_3$, and σ_4. These equations are similar to those obtained by Everett (1996) upon explicit elimination of u_1 and u_2 from the original system. Also note that the Gröbner basis does not contain a constant polynomial, indicating that the system is solvable. However, the system is not finitely solvable; the variety is not 0-dimensional, because it cannot be put in a triangular form, that is, there is no one equation in a single variable. There are an infinite number of solutions which form a variety of some dimension. Characterizing this dimension is the subject of the next section.

The above theorem is satisfying from theoretical point of view, but as a way to find solutions it runs into computational problems (Adams and Loustaunau 1994). At present, Gröbner bases constructed using lexicographic ordering are only computationally feasible for small problems. There is some evidence that parallel computation and approaches taking account of system sparsity could extend the range of the calculations, but this remains to be seen. The source of some of this difficulty stems from trying to derive a single polynomial equation in one unknown. The total degree of such an equation will be quite high and the coefficients of the equation will vary by many orders of magnitude. Consider the factored form of such an m-degree polynomial equation for z_l:

$$(z_l - r_1)(z_l - r_2) \cdots (z_l - r_m) = 0$$

where the r_i denote the roots. The constant term will consist of the product of the roots $r_1 \cdot r_2 \cdots r_m$ while the coefficient of the term ${z_l}^m$ will be 1. Similarly, the other coefficients will consist of sums of various products of the roots (Tignol 1988). Depending on the size of the roots, the terms likely to vary greatly in size for m large. In order that the roots may be determined accurately, extremely high precision is required leading to storage problems and excessive computation manipulating coefficients. It should be emphasized that such problems are not associated with graded Gröbner bases, those which order by monomial degree first, because such orderings do not produce triangular systems. Graded orderings are used in the next section to extract information concerning polynomial varieties. Finally, Gröbner basis calculations are efficient when a finite number field, modulo a large prime number for example, is used. There are techniques for relating calculations made using several such number fields to a Gröbner basis with respect to infinite fields such as the rational numbers $\mathbf{Q}$ or $\mathbf{R}$ (Manocha 1994).

Feasible approaches to solving systems of polynomial equations are largely numerical and work directly with

the defining equations. For example, there are purely numerical approaches based upon Newton iteration and deflation (Dobbs and Hanks 1992). A Newton-Raphson algorithm is used to find a root $\zeta = (\zeta_1, ..., \zeta_n)$, and then the linear factors $z_1 - \zeta_1, ..., z_n - \zeta_n$ are successively divided out of the equations using polynomial division. A more sophisticated approach uses a combination of multi-polynomial resultants and matrix computations (Manocha 1994). Such an approach follows upon early work in algebraic geometry (Macaulay 1902). The underlying idea is to take the polynomial system of equations, $f_1 = 0, ..., f_p = 0$, say of maximum order d, in the variables $z_1, ..., z_n$ and write it as a matrix equation

$$\mathbf{M}(z_1)\mathbf{z} = 0$$

where $\mathbf{M}(z_1)$ is a square matrix of polynomials in z_1 and $\mathbf{z}$ is a vector containing monomial multipliers in the other variables $z_2, ..., z_n$

$$\mathbf{z} = \begin{pmatrix} 1 \\ z_2 \\ \vdots \\ z_n \\ \vdots \\ {z_2}^d {z_3}^d \\ \vdots \\ {z_n}^d \end{pmatrix}.$$

The condition for the existence of a non-trivial solution of the above system is that the determinant of $\mathbf{M}(z_1)$ vanish. The matrix $\mathbf{M}(z_1)$ may be decomposed into terms of various orders in z_1

$$\mathbf{M}(z_1) = \mathbf{M}_0 + \mathbf{M}_1 z_1 + ...\mathbf{M}_l {z_1}^l$$

where $\mathbf{M}_i$ are matrices with numerical entries and l is the maximum degree of z_1 in $\mathbf{M}(z_1)$. The equation may be normalized if we multiply by ${\mathbf{M}_l}^{-1}$, assuming that it is non-singular. The singular case is treated in Manocha (1994). Denote the transformed matrices by

$$\overline{\mathbf{M}}_i = {\mathbf{M}_l}^{-1}\mathbf{M}_i$$

similarly for $\overline{\mathbf{M}}(z_1)$. For a non-trivial solution to exist the determinant of the matrix $\overline{\mathbf{M}}(z_1)$ must vanish. It turns out that the determinant of $\overline{\mathbf{M}}(z_1)$ is a polynomial whose roots correspond to the eigenvalues of the matrix

$$\mathbf{P} = \begin{pmatrix} 0 & \mathbf{I}_m & 0 & ... & 0 \\ 0 & 0 & \mathbf{I}_m & ... & 0 \\ \vdots & \vdots & \vdots & ... & \vdots \\ -\overline{\mathbf{M}}_0 & -\overline{\mathbf{M}}_1 & -\overline{\mathbf{M}}_2 & ... & -\overline{\mathbf{M}}_{l-1} \end{pmatrix}$$

(Manocha 1994). One advantage of this approach is that we never derive a single polynomial equation for z_1. Furthermore, we may make use of techniques from linear algebra. The computation requirements and numerical characteristics of such techniques are well understood. One difficulty of this approach is that the order of the matrix $\mathbf{P}$ grows exponentially with the order and number of equations. This fact is offset by the sparseness of the matrix that often increases with the size of the problem. Sparse matrix eigen-decomposition techniques can now treat very large problems (Sehmi 1989).

Non-uniqueness: The Dimension and Degree of a Variety

The case of finite solvability corresponds to an affine variety of 0-dimension, consisting of a finite collection of points. The more general case of an infinite solution set of some positive dimension might be encountered in practice. As in the previous subsections, the geometric question of the dimension of a variety may be expressed in terms of the algebra of ideals.

First consider the dimension of a variety $\mathbf{V}$ defined by the vanishing of monomials. This case provides some insight into the calculation of the dimension of a general polynomial variety. Furthermore, the dimension of an arbitrary variety can be related to a specific monomial ideal: the ideal of leading terms $\langle \mathrm{LT}(I) \rangle$. Now the variety defined by a single monomial in a single variable, say ${z_k}^{\alpha_k} = 0$, is a coordinate hyperplane, denoted by $\mathbf{H}_{z_k}$. The variety defined by the vanishing of a single monomial in several variables,

$${z_{i_1}}^{\alpha_1} \cdots {z_{i_r}}^{\alpha_r} = 0,$$

consists of the union of such coordinate hyperplanes. A variety defined by a set of multivariable monomials is the intersection of such unions. Because intersections are distributive over unions and the intersection of coordinate hyperplanes is a coordinate hyperplane we may write $\mathbf{V}(I)$ as a union of coordinate subspaces

$$\mathbf{V}(I) = \mathbf{V}_1 \cup \cdot\cdot \cup \mathbf{V}_p.$$

The dimension of the variety is defined as the largest of the dimensions of the subspaces $\mathbf{V}_k$.

Definition If $\mathbf{V}(I)$ is the union of a finite number of linear subspaces of k^n, the dimension of the variety, $\dim\mathbf{V}$, is the largest of the dimension of these subspaces.

The definition of the dimension of a monomial variety, while simple to compute, does not carry over to general varieties characterized by sets of polynomials. In order to arrive at a workable algorithm we must pursue an alternative strategy. The key observation, first made by Hilbert, is that the dimension of a variety $\mathbf{V}$ can also be characterized by the growth, as a function of total degree, of the number of monomials not contained in the ideal $I(\mathbf{V})$. In fact, the dimension is given by the degree of a polynomial characterizing the number of monomials not in $I(\mathbf{V})$. First we introduce a concise notation for

the set of exponents of monomials not in I, known as the complement of the monomial ideal,

$$C(I) = \{\alpha \in \mathbf{Z}^n : z^\alpha \notin I\}$$

where $\mathbf{Z}$ denotes the integers and $\mathbf{Z}^n$ signifies an n-dimensional vector of integers. The relationship between the complement of the monomial ideal and the dimension of the variety is given as follows (Cox et al. 1997, p. 436).

Theorem *The dimension of* $\mathbf{V}(I)$ *is the dimension of the largest coordinate subspace in* $C(I)$.

The importance of this theorem is that there are explicit formulas for the total number of monomials in $C(I)$. For example, the number of monomials of a total degree $\leq s$ in $k[z_1, ..., z_m]$ is given by the binomial coefficient, and is a polynomial of degree m in s

$$\binom{m+s}{s} = \binom{m+s}{m} = \frac{1}{m!}(s+m)(s+m-1)\cdots(s+1).$$

The intuitive idea for calculating $C(I)$ is that the elements that are not in $C(I)$, those monomials in I, are contained in a finite number of translates of the octant $\mathbf{Z}^n$.

As an example, consider the ideal generated by the monomials ${z_1}^5{z_2}^2$ and ${z_1}^3{z_2}^4$ (Figure 1). The monomials contained in $I = \langle {z_1}^5{z_2}^2, {z_1}^3{z_2}^4\rangle$ are denoted by the filled squares and the shaded region in Figure 1. The monomials in $C(I)$ are those found in hyperplanes between the boundaries of the octant $\mathbf{Z}^n$ and the boundaries of the set of octants (quadrants in the example) translated to (5,2) and (3,4). In Figure 1 these missing monomials, up to degree 6, are indicated by crosses. For a particular total degree, the number of such trapped monomials is given by the following theorem (Cox et al. 1997, p. 442).

Theorem *If* $I \subset k[z_1, ..., z_n]$ *is a monomial ideal with* $\dim \mathbf{V}(I) = d$, *then for all* s *sufficiently large, the number of points in* $C(I)$ *of total degree* $\leq s$ *is a polynomial of degree* d *in* s *that can be written in the form*

$$\sum_{i=0}^{d} a_i \binom{s}{d-i},$$

where $a_i \in \mathbf{Z}$ *for* $0 \leq i \leq d$ *and* $a_0 > 0$

This polynomial function is known as the affine Hilbert polynomial of I and is denoted by $HP_I(s)$.

In order to move on to general varieties, defined by systems of polynomial equations, there is a critical result that depends on how the monomials are ordered. In particular, a graded ordering is necessary because under this ordering we can directly relate the Hilbert polynomial of I to a monomial ideal (Cox et al. 1997, p. 448).

Theorem *Let* $I \subset k[z_1, ..., z_n]$ *be an ideal. The monomial ideal* $\langle \mathrm{LT}(I)\rangle$, *constructed using a graded order on* $k[z_1, ..., z_n]$, *has the same affine Hilbert polynomial as* I.

It also may be shown that both the ideal I and its radical $\sqrt{I}$ have affine Hilbert polynomials of the same degree (Cox et al. 1997, p. 449). Gathering facts, we find a suggestion that the degree of the affine Hilbert polynomial contains geometric information about the variety $\mathbf{V}$. First, it is an integer that is invariant for a large collection of ideals defining the variety, namely $\mathbf{V}(I)$, $\mathbf{V}(\sqrt{I})$ and all varieties in between. Furthermore, the degree of the affine Hilbert polynomial is the same for the monomial ideal $\langle \mathrm{LT}(I)\rangle$ and I. Finally, we know that for monomial ideals the affine Hilbert polynomial corresponds to our intuitive notion of the dimension of the variety. For these reasons the dimension of a general affine variety is often defined in terms of the affine Hilbert polynomial.

Definition The *dimension* of an affine variety $\mathbf{V} \subset k^n$, denoted $\dim \mathbf{V}$, is the degree of the affine Hilbert polynomial of the ideal $I = I(\mathbf{V}) \subset k[z_1, ..., z_n]$.

How do we go about computing the dimension of a variety $\mathbf{V}(I)$? First, we compute a Gröbner basis for I using a graded lexicographic ordering. From the definition of a Gröbner basis we know that if $\{g_1, ..., g_p\}$ generate a Gröbner basis of I then $\langle \mathrm{LT}(g_1), ..., \mathrm{LT}(g_p)\rangle = \langle \mathrm{LT}(I)\rangle$. We then compute the maximum dimension d of the coordinate subspaces contained in $\mathbf{V}(\langle \mathrm{LT}(I)\rangle)$. Because this is a monomial ideal this is easily done, as indicated at the beginning of this subsection. Consider our example problem and the Gröbner basis with respect to graded lexicographic ordering. The ideal of leading terms is generated by $\langle z_1, z_2, z_4z_6, z_3z_5\rangle$ with the corresponding coordinate subspaces

$$\mathbf{H}_{z_1} \cap \mathbf{H}_{z_2} \cap (\mathbf{H}_{z_4} \cup \mathbf{H}_{z_6}) \cap (\mathbf{H}_{z_3} \cup \mathbf{H}_{z_5}).$$

The dimension of the zero set of the system of polynomial equations is given by the dimension of this quantity.

One may devise a systematic algorithm for computing the dimension of a variety given the monomial generating set of its ideal I (Cox et al. 1997, p. 431). The algorithm is outlined here and illustrated by the example problem. For an ideal generated by monomials $m_1, ..., m_p$, $I = \langle m_1, ..., m_p\rangle$, the variety consists of the common zero set of the defining monomials, that is, the intersection of the zero sets of all the components

$$\mathbf{V}(I) = \bigcap_{j=1}^{p} \mathbf{V}(m_j).$$

From the above definition, the dimension of $\mathbf{V}(I)$ is the component that has the largest dimension. The basic idea is to find a collection of variables $z_{i_1}, ..., z_{i_r}$ such that each monomial contains at least one variable in the set. Then the coordinate subspace specified by the vanishing of these variables is contained in $\mathbf{V}(I)$. The largest such subspace is the one defined by the vanishing of the fewest variables. Therefore we should try to find the smallest of all such sets. Specifically, denote the set of subscripts

occurring in monomial m_j

$$M_j = \{i \in (1, ..., n) : z_i \text{ divides the monomial } m_j\}$$

For our example problem $m_1 = z_1, m_2 = z_2, m_3 = z_4 z_6$, and $m_4 = z_3 z_5$ so that

$$\begin{aligned} M_1 &= \{1\} \\ M_2 &= \{2\} \\ M_3 &= \{4, 6\} \\ M_4 &= \{3, 5\} \end{aligned}$$

Now consider all subsets of $\{1, ..., n\}$ that have elements in common with every set M_j

$$M_{common} = \{J \subset (1, ..., n) : J \cap M_j \neq \emptyset\}$$

Let $|J|$ denote the number of elements in any set J from M_{common}. Then it may be shown (Cox et al. 1997, p 431) that $\dim \mathbf{V}(I) = n - \min(|J|)$.

For our small example the elements of M_{common} are sets with 4 members: $(1, 2, 4, 3), (1, 2, 6, 3), (1, 2, 4, 5)$, and $(1, 2, 6, 5)$. Hence, the variety determined by our small system of polynomial equations has $\dim \mathbf{V}(I) = 6 - 4 = 2$. In such a simple case we could compute the dimension almost by inspection. However, in actual applications there will be a large number of equations and variables and we must rely on symbolic computational algorithms such as those mentioned above.

As might be expected, such a general concept as dimension may be approached from a number of directions. This is true in algebraic geometry and there are a number of other definitions for the dimension of an affine variety (Eisenbud 1995). For example, there is the algebraic Krull dimension (Kunz 1985, Eisenbud 1995, p. 227). There are also more geometric approaches related to nonsingularity and the idea of a manifold (Kendig 1977, p. 163). Our approach falls more into the class of computational techniques based upon the codimension of a variety and the number of independent variables with respect to the ideal I (Mishra 1993, p. 141).

The Hilbert polynomial $HP_I(s)$ also may be used to calculate the degree of the ideal I (Giusti 1989). In particular, the coefficient a_0 in the polynomial

$$HP_I(s) = \sum_{i=0}^{d} a_i \binom{s}{d-i}$$

is equal to the degree of I. The degree of I is a generalization of the total degree of the defining equation of a hypersurface. For the case of an ideal I generated by a set of homogeneous polynomials $f_1, ..., f_p$ the degree of I is the product $\deg f_1 \cdots \deg f_p$. Homogeneous polynomials may be thought of as defining varieties in a projective space. In projective space solutions or roots at infinity are well defined (Kendig 1977) and the degree of the defining polynomials may be directly related to the zero set. For general 0-dimensional ideals the degree is an upper bound on the number of solutions to the set of defining equations. That is, there may be fewer solutions because we do not consider roots at infinity.

Regularization Polynomials

It is quite common to treat the issue of non-uniqueness by introducing some form of regularization (Menke 1989, Parker 1994). Regularizing non-linear inverse problems is complicated by the presence of local minima (Vasco 1998). The usual approach is to introduce a function that measures some quantifiable attribute of the model such as model norm or model smoothness. Often these attributes are described by quadratic functionals defined over the model space. For example, we may wish to obtain a model that lies close to a prior structure $\mathbf{z}^0$ as measured by the sum of the squares of the deviations

$$|\mathbf{z} - \mathbf{z}^0| = \sum_{i=1}^{n} (z_i - z^0{}_i)^2.$$

Another common form of regularization is model roughness that is also typically a quadratic form such as

$$|\nabla^2 \mathbf{z}| = \sum_{i=1}^{n} (\Delta^2 z_i)^2,$$

where Δ^2 is the second spatial differencing operator. For an $n = l \times l$ 2D grid of cells $\Delta^2 z_i$ is given by

$$\Delta^2 z_i = 4z_i - z_{i-1} - z_{i+1} - z_{i-l} - z_{i+l}$$

which is a measure of the difference in parameter values between block i and blocks $i-1$, $i+1$, $i-l$, and $i+l$, where l is the number of blocks in a row of cells. There are a number of ways to combine such penalty terms and the data equations. For example, we could minimize the model norm and/or the model roughness subject to the satisfaction of the data equations. Using the method of Lagrange multipliers (Menke 1989, Parker 1994) produces a system of polynomial equations defining an affine variety (Watson and Morgan 1992). Alternatively, we could simultaneously minimize both the quadratic regularization measures and the deviation of our data equations from zero (in a squared-error sense), again arriving at a system of polynomial equations. The point is that the algebraic techniques presented above are still applicable when considering the regularized inverse problem, for most commonly used regularization terms, or when minimizing data misfit.

Consider a specific application of the Lagrange multiplier approach to the system of equations (9). For this example let's seek the model closest to the prior structure $\sigma^0(x, y)$ that fits the data. We shall assume that the constraint equations (9) have been modified to account for the inclusion of the boundary data. For example, in

our small problem we have substituted observed values d_1 and d_2 for nodal electric field variables u_3 and u_4. There are two consequences of this: we have fewer electric field variables and also there will be conductivities in the new 'right-hand-side', denoted by $\mathbf{b}'$. In order to find the model nearest to our prior model that satisfies the modified constraint equations we minimize

$$W(\sigma) = \sum_{i=1}^{M} (\sigma_i - \sigma_i{}^0)^2$$

subject to the M constraints,

$$U_i(\sigma, \mathbf{u}) = \sum_{j=1} A_{ij}{}'(\sigma) u_j - b_i{}'(\sigma) = 0,$$

where M signifies the total number of grid points (N^2 for an $N \times N$ grid). The primes denote the fact that we are working with the modified constraint equations rather than with equation (9). The grid values are numbered sequentially using a single index rather than two indices as before. The notation $A_{ij}{}'(\sigma)$ and $b_i{}'(\sigma)$ is used to specify that the coefficient matrix and the modified right-hand-side depend on the model parameters. Using Lagrange multipliers λ_i we may write the necessary equations for a constrained extrema of $W(\sigma)$,

$$\nabla W(\sigma) + \sum_{i=1}^{M} \lambda_i \nabla U_i(\sigma, \mathbf{u}) = 0$$
$$U_i(\sigma, \mathbf{u}) = 0, \quad i = 1, ..., M$$

where the gradient contains partial derivatives with respect to both σ and $\mathbf{u}$ variables. In total, there are $3M$ equations in $3M$ variables $(\sigma_i, u_i, \lambda_i)$, $i = 1, ...M$.

NUMERICAL ILLUSTRATION

Penalized Helmholtz Equation

As a prelude to the analysis of a set of magnetotelluric field data, consider a numerical test case. As in Everett (1996), consider the Helmholtz equation defined over a region of the x, y plane

$$\nabla^2 u + q\sigma(x, y) u = 0$$

where q is a real scalar. Both Dirichlet (u) and Neumann ($\partial u/\partial y$) data are available on the surface ($y = 0$) and Dirichlet data are given on the remaining three boundaries. Assume that we have written out the discrete equations $\mathbf{A}(\sigma)\mathbf{u} = \mathbf{b}$ for N_R rows of nodes in the y (depth) direction and N_C columns of nodes in the x (horizontal) direction, a total of $N = N_R * N_C$ unknown conductivities and some subset of unknown field components $\mathbf{u}$. As described in Everett (1996) and outlined earlier, given enough data, we may write the field components in the top layer in terms of the boundary values and the data. Using a vector form of back substitution, the entire system of equations may be converted to a system of N-linear polynomial equations in the conductivities, data, and boundary values. Explicitly, the equations for a 2×2 grid are

$$A\sigma_1 + B\sigma_3 + C\sigma_4 + D\sigma_1\sigma_3 + E = 0$$
$$F\sigma_2 + G\sigma_3 + H\sigma_4 + I\sigma_2\sigma_4 + J = 0$$

where the coefficients A through J depend on the boundary values (including the surface data), the mesh interval h, and the scalar q. The system of polynomial equations is written symbolically as $\mathbf{F}(\sigma) = 0$ as in Everett (1996). Everett (1996) worked directly with the constraint equations, adding more frequencies until the problem was over-determined. As with linear problems, the constraint equations may be nearly singular and hence unstable with respect to small perturbations in the coefficients. A penalized least squares approach is somewhat more robust and less influenced by random errors in the data. Specifically, we include a norm penalty and minimize the composite quantity

$$W(\sigma) = \nu \sum_{i=1}^{N} (\sigma_i - \sigma_i{}^0)^2 + \sum_{i=1}^{M} \sum_{j=1}^{N_R} F_{ij}{}^2(\sigma) \qquad (11)$$

where M denotes the number of data and ν is the trade-off parameter controlling the importance of remaining close to the prior model σ^0 relative to fitting the data. It would be no more work to consider a roughness penalty term such as the first or second derivative operators. The necessary equations for an extremum are given by the vanishing of

$$\nabla W(\sigma) = \nu \sum_{i=1}^{N} (\sigma_i - \sigma_i{}^0) \cdot \nabla\sigma + \sum_{i=1}^{M} \sum_{j=1}^{N_R} F_{ij}(\sigma) \cdot \nabla F_{ij}(\sigma)$$

where the gradient is with respect to the conductivity variables. This is a system of N N_R-linear polynomial equations in N unknowns.

For the sake of illustration, we consider the inversion of a collection of synthetic values. In particular, consider the 2×2 difference grid and assume that we have both Dirichlet and Neumann data at the surface nodes for $M = 10$ frequencies q_i. The frequencies start at 1 Hz and vary geometrically by factors of 1.5. The situation is somewhat similar to the inversion of the Helmholtz equation using noisy data described by Everett (1996) except that our penalized least squares approach is followed. Note that adding more data does not change either the degree or the number of defining equations. The additional data are present in the coefficients of the polynomial equations as sums of M products of the individual coefficients (A through J). The model parameters used to generate the data set are $\sigma_1 = 2, \sigma_2 = 2, \sigma_3 = 4, \sigma_4 = 4$

and the prior model σ^0 consists of a uniform background with $({\sigma^0}_1, {\sigma^0}_2, {\sigma^0}_3, {\sigma^0}_4) = (1.5, 1.5, 3.5, 3.5)$. We use techniques from the previous subsections to determine the dimension of the solution set and, if possible, estimate a finite set of solutions.

First, consider an inversion without regularization, in which a penalty term is not included. Using a graded lexicographic Gröbner basis and methods from the previous subsection, we find that the solution set is two-dimensional. That is, even though we have 20 equations (two for each frequency) constraining four unknowns the problem is still undetermined. Furthermore, a Gröbner basis representation using lexicographic ordering is not in a triangular form. In particular, the basis elements contain the cross product terms $\sigma_1 \cdot \sigma_3$ and $\sigma_2 \cdot \sigma_4$. Both of these calculations were carried out in a fraction (0.63) of a CPU second on a workstation, using the package Singular from the University of Kaiserlautern (Cox et al. 1997). The non-uniqueness is suggested by Figure 2, which displays six two-dimensional sections through the model space. In each section, the sum of the square of the components of $\nabla W(\sigma)$ is shown. The vanishing of these components is a necessary condition for the least squares solution. The sections are all through the solution to the inverse problem $(\sigma_1, \sigma_2, \sigma_3, \sigma_4) = (2, 2, 4, 4)$. There is a trade-off between σ_1 and the conductivity of the overlying node σ_3. In addition an infinite number of combinations of the conductivities σ_2 and σ_4 minimize the misfit to the data. Our set of observations do not allow us to resolve the depth variation in conductivity. Thus, our inverse problem seems to be over-determined, there are more equations than unknowns. However, due to the distribution of frequencies (q_i), trade-offs are present and there is no unique solution.

Next consider the effect of including a regularization penalty term. The weighting associated with the penalty term is taken to be quite small, 10^{-6}, later we shall explore the result of varying this weighting. In Figure 3, the sections through the model space, the trade-offs between σ_1 and σ_3 and between σ_2 and σ_4 are no longer evident. Now the graded lexicographic Gröbner basis calculations indicate that the solution set is 0-dimensional. This means that there is a finite number of solutions to the above system of equations. The degree of the ideal is calculated to be 25, providing an upper bound on the total number of solutions. For such a small problem a lexicographic Gröbner basis can be used to find the solutions. The Gröbner basis consists of four elements in a strongly triangular form. Thus, we have reduced the polynomial system to a sequence of one-dimensional problems that may be solved successively for the conductivities. The Gröbner basis elements are

$$\sum_{i=0}^{25} c_1(i) \cdot {\sigma_4}^i = 0$$

$$\sigma_3 + \sum_{i=0}^{24} c_2(i) \cdot {\sigma_4}^i = 0$$

$$\sigma_2 + \sum_{i=0}^{6} \sum_{j=0}^{6} c_3(i,j) \cdot {\sigma_3}^i {\sigma_4}^j = 0$$

$$\sigma_1 + \sum_{i=0}^{2} \sum_{j=0}^{2} \sum_{k=0}^{1} c_4(i,j,k) \cdot {\sigma_2}^i {\sigma_3}^j {\sigma_4}^k = 0$$

where $c_1(i), c_2(i), c_3(i)$, and $c_4(i)$ are known numerical coefficients produced by Buchberger's algorithm. Of the 25 roots associated with the first equation only three resulted in real conductivities, giving σ_4 equal to either 1.13, 3.88, or 5.06. The second Gröbner basis element is only first degree in the unknown σ_3. Because, the first equation gives three possible values for σ_4 we merely substitute them into this equation and easily solve for σ_3. Possible values for σ_3 are 1.13, 3.87, and 636.14. The third basis element is only first order in σ_2 and sixth order in both σ_3 and σ_4, After substituting the previously determined values for σ_3 and σ_4, we find that only the first two values produce positive estimates for σ_2: 2.78 and 2.07. Lastly, the equation for σ_1 is in the same form as those for σ_2 and σ_3, linear in the unknown conductivity (σ_1) Only the second root gives a positive value of σ_1: 2.06. Thus, we find a single physical solution for the conductivities $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ which is closest to our prior model and which fits the data: 2.06, 2.07, 3.87, and 3.88.

For comparison, a variable metric or quasi-Newton method (Gill et al. 1981) was used to minimize the function $W(\sigma)$ numerically. We are guided by the information provided by the graded Gröbner basis representation. Specifically, the basis calculations show that a finite solution set indeed exists. Furthermore, the degree bound of 25 provides an upper bound on the total number of solutions possible. The approach is to find a solution using the numerical scheme and then to deflate the system by dividing out the root to produce a polynomial system of lower total degree. Note that because we are only interested in real positive solutions we need only deflate the roots that are real and positive. This is important because there are generally many fewer real positive roots. As in the lexicographic Gröbner basis calculations we find a single positive real solution. The estimated values for $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ are 2.09, 2.08, 3.84, and 3.84 respectively.

In order to evaluate the influence of data errors on the parameter estimates, various levels of Gaussian noise are added to the synthetic field values. The normal deviates are zero mean with a standard deviation proportional to a fixed percentage of the synthetic field value. A total of 25,000 noisy field values were generated at each error level. The errors propagate from the field values to the coefficients of the defining polynomial equations. Because the coefficients enter as products, the error propagation will not be linear. Adding noise to the field values perturbs the coefficients resulting in perturbed zeroes of the polynomial equations. In general real roots become complex and roots formerly complex may become real. Because the components of σ are not complex we need a

working definition of a 'real' root. A root is defined to be real if its imaginary component is 1/1000th of its real part and not greater than 0.001 in magnitude. Furthermore, all elements of σ must be positive and only positive roots are retained.

In Figure 4 the 'real' positive roots of σ_1 are shown for four different noise levels. When the data are contaminated by very little noise (Figure 4a) most values are near the actual value of σ_1 used to generate the synthetic field values (2.0). Due to the regularization, there is a bias to values slightly less than 2. Increasing the noise level to just 3% produces a dramatic shift in the configuration of solutions (Figure 4b). Many solutions have σ_1 equal to values near 2 but a larger number have σ_1 around 3. Note how the results of the penalized least squares approach contrast with the distribution of solutions found by Everett (1996). At 1% noise he found a great scatter of values for σ_1 with no significant clustering of solutions. This result illustrates the greater stability of a penalized least squares approach with respect to perturbations in the coefficients A through J. When 10% noise is added to the synthetic field values the number of solutions clustering around $\sigma_1 = 2$ is much less than the number of solutions containing σ_1, slightly less than 3 (Figure 4c). When the noise level is equal to the signal (Figure 4d) most solutions have $\sigma_1 = 1.5$, the value of our prior model $\sigma^0{}_1$. Lastly, the distribution of solutions for all four parameters are shown in Figure 5 for a noise level of 5%. All estimates cluster about one or more distinct values. Note the significant influence of the penalty term in biasing the clusters away from the actual σ values.

The preceding synthetic tests were conducted with a penalty weight of 10^{-6}. Four sets of inversion were also conducted in which the penalty weight varied by ten orders of magnitude (Figure 6). For each of these sets the noise in the field values was 5%. The structure of the solution set varied little for a wide range of penalty weighting, over six orders of magnitude, from 10^{-6} to 1 (Figures 5a and 6a). The penalty weight had to be raised to very large values (Figure 6b) before any significant change in the solution set was noted. At around $\nu = 10^{10}$ the solutions begin to cluster around the prior estimate $\sigma^0{}_1$. As indicated in Figure 4, this behavior may be strongly influenced by the level of noise in the data.

APPLICATION

East Antarctic Magnetotelluric Experiment

Probably the least understood continental structure is that of Antarctica. Up to 95% of the land mass lies enshrouded in ice and the region is relatively inaccessible. Successful geophysical investigations must cope with the thick (1-3 km) layer of ice, which has a fairly high (3.9 km/sec) P-velocity (Bentley 1973, 1991) and is highly resistive (Beblo and Leibig 1990). Crudely, Antarctica is divided into eastern and western regions by the Transantarctic mountains (Craddock 1970). West Antarctica has accreted to East Antarctica from the Lower Paleozoic to the present (Storey and Alabaster 1991). East Antarctica has remained relatively stable through the Jurassic and is overlain by sediments of the Beacon Supergroup that reach thicknesses of up to 2.5 km (Barrett 1991). Seismic surface wave studies indicate continental-type structure in eastern Antarctica (Press and Dewart 1959, Roult et al. 1994) and a differing crustal thickness between east and west Antarctica (Evison et al 1959, Bentley 1991). In particular, there is a 10 km difference in the crustal thickness of east (40 km) and west (30 km) Antarctica (Evison et al. 1959). East Antarctic group velocities fall between those of Eurasia and North America suggesting that the region is not simple crystalline shield but rather a 'typical continental structure' (Dewart and Toksöz 1965, Bentley 1973). There are suggestions that basement elevations of East Antarctica are high, perhaps by as much as 1 km, suggesting some form of dynamic tectonic support (Bentley 1991). The difference in crustal thickness between east and west Antarctica also is supported by seismic refraction work (Bentley 1991) and by the strong Bouguer gravity gradient across the Transantarctic Mountains (Robinson 1964).

In an effort to better constrain east Antarctic crustal structure a group of investigators from the University of Utah and Berkeley Laboratory conducted a South Pole magnetotelluric (MT) survey. This electromagnetic technique is sensitive to conductivity variations of the order of kilometers to tens of kilometers and complements earlier seismic studies of the region. The current experiment was a follow-up to the University of Utah's first successful broadband magnetotelluric survey in Western Antarctica during 1994-1995 (Wannamaker et al. 1996). Using a differential E-field measurement technique this group overcame the difficulty caused by the very high contact impedance of the polar ice. The high impedance prevented high quality broadband electric field measurements. The results of the East and West Antarctic surveys demonstrate the feasibility of conducting MT surveys over the entire continent and the ability of MT to map conductivities to lower mantle depths (1000 km) at the South Pole. Such deep mantle resistivity profiles are lacking in this part of the globe (Everett and Schultz 1996).

The East Antarctic experiment was conducted along a 54 km transect with an origin offset about 6 km from the South Pole and an azimuth of 210° east of the Greenwich meridian (Figure 7). This orientation was approximately perpendicular to the strike of the Transantarctic Mountains. The station spacing was approximately 6 km, with locations provided by hand-held GPS receivers, and equipment was transported between stations using snowmobiles. The station spacing provided adequate spatial sampling because the thick ice sheet eliminated the near-surface 'static' effects observed in many land

surveys (Pellerin and Hohmann 1990). Magnetotelluric observations were recorded over a number of days, on average three days per station. One reason for the fairly long observation times was a wind induced voltage variation in the ice itself which appeared as noise in electric field measurements at periods of 10 s or longer.

The electric and magnetic field measurements are Fourier transformed and averaged over frequency bands to obtain estimates of $E_x(\omega)$, $E_y(\omega)$, $H_x(\omega)$, $H_y(\omega)$, and $H_z(\omega)$ (Stodt 1983). The bandwidth of the measurements spanned a period range of 0.001 to 500 s. Using a cross-power spectral technique, the impedance tensor Z_{ij} is estimated from the field averages (Vozoff 1991). The components of the impedance are in turn used to derive apparent resistivities ρ_{xy} and ρ_{yx} as well as impedance phases ϕ_{xy} and ϕ_{yx} (Figures 8 and 9). There is a notable drop in apparent resistivity with period for short and intermediate periods. Furthermore, the overall similarity of the xy apparent resistivity and phase with the yx components indicates that the structure in this region is not very anisotropic. For example, at periods less than 1 s, the apparent resistivity (Figure 8) does not vary much between stations. At longer periods both ρ_{xy} and ρ_{yx} are somewhat lower, between -35 and -25 km along the survey line. Similarly, the phase (Figure 9) behaves rather uniformly between stations, with a change in phase angle from 80-90° to 40-60° for periods greater than about 10 s but less than 300 s. There is an increase in phase angle between -35 and -25 km in the transect, which correlates with apparent resistivity variations. A rough interpretation of the apparent resistivity and phase pseudosections indicates shallow high resistivity polar ice underlain by lower resistivity sediments, which are on top of a slightly more resistive basement structure.

Integrated one-dimensional modeling, based on the method of Petrick et al. (1977), by the University of Utah group indicates that the apparent resistivity and phase are well matched by three layers over a half-space. The first layer represents highly resistive (250,000 ohm-m) polar ice with a thickness of 2.9 km. The other layers are found at depth intervals of 2.9-3.7 km and 3.7-27.2 km with resistivities of 3.3 and 26 ohm-m's respectively. The layer directly beneath the ice may represent porous sediments, perhaps from the Beacon Supergroup (Barrett 1991). This layered structure serves as our a priori model, σ^0, and the starting point for the inversion. The inversion technique is the penalized least squares algorithm represented by equation (11). However, because the data suggest a dominantly one-dimensional structure, data from each station is inverted for a purely vertical variation. The model is specified by five nodes in depth, centered at the mid-points of each layer (1.45 km, 3.3 km, 15.5 km) and a node just above the ice surface -0.1 km as well as a node at 33.6 km to capture variation at the top of the half-space. There was also a deep boundary node at 60 km used to enforce the bottom boundary condition. The electric ($E_x(\omega)$) and the magnetic fields ($H_y(\omega) = \partial E_x \partial z$) at 34 distinct frequencies provide the basic data at each station.

A block-centered, finite-difference discretization of the TE mode equation provides the relationship between the electric field and subsurface conductivity. A variable node spacing is used to account for the differences in layer thickness. The interior electric field variables are eliminated, resulting in a polynomial equation in the layer conductivities $\sigma_1 - \sigma_4$ (the conductivity of the surface air layer is fixed). The resulting regularized inverse problem consists of four polynomial equations of degree 5 in the four unknown conductivities beneath the station. A graded lexicograpic Gröbner basis indicates that the system of polynomial equations is solvable, that the solution set is 0-dimensional, and the total degree of the system is 50. This provides an upper bound on the total number of solutions to the inverse problem. The graded lexicographic Gröbner basis consists of 35 elements of degrees 4 and 5. An iterative numerical solver is used to find the solutions with which to deflate the defining polynomial equations. Again, only the real positive roots were used to deflate the equations. Of all real positive solutions the one producing the minimum misfit is taken as the 'best' solution. In Figure 10 the resulting solutions for all ten stations are shown. The solutions are plotted beneath each station, at approximately 6 km spacing. Note the largely one-dimensional structure and the lateral continuity of the solutions, even though no lateral smoothing was enforced in the inversion. There would appear to be a trade-off between the resistivity of the top two layers and the resistivity of the 3.7-27.2 km layer. One way to control the trade-off is to fix the conductivity of the ice, assuming it does not vary laterally. However, an inversion in which the resistivity of the top layer was fixed at 250,000 ohm-m produced essentially the same variation in the 3.7 to 27.2 km depth interval.

In general the squared error, representing the sum of the squares of the deviations or the right-hand-side of the equations from zero, decreased from around 10^{-7} to roughly 10^{-10} for each station. The reduction in misfit is shown in Figure 11 for each frequency (34) and each station (10). In this figure the percent ratios (final misfit)/(prior misfit) are shown for all residuals. The misfit is measured by the absolute value of the right-hand-side of the polynomial equation. For a perfect fit and no modeling error, the right-hand-side would vanish. In Figure 11 we see that much of the misfit has been reduced to less than 1% of the initial misfit. With the exception of 26 values, the data misfit has been reduced to under 5% of the initial misfit.

DISCUSSION AND CONCLUSIONS

The solution of non-linear inverse problems is a difficult task. Current approaches generally rely on either a local linearization or some stochastic sampling of the model space. Both of these methodologies suffer from limita-

tions, and neither makes full use of the analytic structure of the equations defining the inverse problem. In specific cases it is possible to transform a non-linear problem to a linear one (Vasco 1997) and then apply conventional techniques, but this in not possible in general. For a large class of inverse problems, differential equations with undetermined coefficients and many integral equations, the discrete form of the equations are polynomial (Everett 1994). Thus, the equations have a well defined structure and the solution sets are restricted in form. That is, the solution sets constitute geometrical entities known in algebraic geometry as analytic varieties. The solution of polynomial equations is a topic of considerable history and efforts at solving such equations have led to many advances in mathematics (Tignol 1988). However, many notable advancements such as the Gröbner basis algorithm (Buchberger 1985) are quite recent. Furthermore, the numerical implementation of these ideas is still at an early stage of development (Fleischer et al. 1995). Compared to linear algebra for example, the development of numerical and symbolic algorithms for polynomial systems is a young field and considerable improvements in efficiency are likely. In order to limit the scope of this paper many algorithms have not been touched upon. For example, techniques for counting real roots (Pederson et al. 1993) have not been described. Combination of the concepts discussed here with new numerical algorithms, such as intermediate factorization of polynomials (Melenk et al. 1989) or algorithms for polynomial programming (Floudas and Visweswaran 1990, Sherali and Tuncbilek 1992), could provide both insight and computability.

In this paper I have tried to illustrate the utility of ideas and algorithms from computational algebra in the solution of inverse problems. The applications are intended to demonstrate that the algorithms and concepts apply to practical problems. By this I mean one can account for data noise and also incorporate regularization into the formulation. However, the robustness of a least squares approach comes at the cost of additional computation. For example, consider the penalized misfit functional in equation (11). If the defining equations are of order d in the variables σ, the necessary equations are of order order $2d-1$ because we are squaring and then differentiating with respect to the variables. Thus, we must solve equations of higher degree than the original set. Even so, it is possible to use this approach in the analysis of actual observations. The inversion of the East Antarctia magnetotelluric observations is a first step in that direction. Working directly with the discrete difference equations we are able to improve the match to the field data. The low crustal resistivity, 2 ohm-m or less, appears to support earlier evidence of low seismic surface wave phase velocities (Bentley 1973, 1991). This, coupled with the anomalously high basement elevation in the region, suggests active tectonic support.

Several aspects of this work could be explored in more detail. For example, it is well established that the number of roots of a polynomial system of equations grows exponentially with the number of variables and the degree of the equations. However, often one is only interested in the real positive roots of the equations. The number of real roots is generally a much smaller subset of the set of all roots. By restricting the number field we are working in one may substantially limit the amount of computation required. An even more restrictive approach is to employ a type of interval constraint on each parameter σ_k. The roots of the equations are restricted to lie between particular upper and lower values: ${\sigma_k}^u > \sigma_k > {\sigma_k}^l$. This would both avoid extreme values and help in culling the roots, thus preventing possible exponential growth in the number of solutions. The degree to which parallel computation could help in solving larger problems also needs to be examined. Parallel computation has proven helpful in examining the set of solutions to the travel time tomographic inverse problem (Vasco et al. 1996) using a purely numerical approach. As shown above, solving multi-variable systems of polynomial equations leads to a large sparse eigenvalue problem, which might be solved efficiently in parallel (Sehmi 1989). The techniques presented here only apply to the discretized inverse problem. It is possible to treat the continuous non-linear inverse problems (Snieder 1991) but the approach is computationally intensive and based upon a perturbative analysis. There is a formalism for polynomial equations in Banach space (Argyros 1998) which may extend the discrete approach described here, but this is a topic for future research. Finally, the algebraic approach presented here needs to be compared to conventional inversion techniques. Only in applications will the strengths and limitations of the algorithms be revealed.

ACKNOWLEDGMENTS

I would like to thank both Louise Pellerin and Phil Wannamaker for supplying the high quality magnetotelluric data from the South Pole experiment. This work was supported by a Laboratory Directed Research and Development grant, Office of Energy Research, Division of Basic Energy Sciences, Engineering, and Geosciences, of the U.S. Department of Energy under contract DE-AC03-76SF00098. All computations were carried out at the Center for Computational Seismology and the National Energy Research Scientific Computing (NERSC) Center at Berkeley Laboratory.

REFERENCES

Adams, W., and P. Loustaunau, 1994. *An Introduction to Gröbner Bases*, American Mathematical Society, Providence.

Argyros, I. K., 1998. *Polynomial Operator Equations in Abstract Spaces and Applications*, CRC Press, Boca Raton.

Auslander, L. and R. E. MacKenzie, 1963. *Introduction to Differentiable Manifolds*, McGraw-Hill, New York.

Barrett, P. J., 1991. The Devonian to Jurassic Beacon Supergroup of the Transantarctic Mountain and correlatives in other parts of Antarctica, in *The Geology of Antarctica*, R. J. Tingey (ed.), Clarendon Press, Oxford.

Beblo, M., and V. Liebig, 1990. Magnetotelluric measurements in Antarctica, *Phys. Earth Planet. Inter.*, **60**, 89-99.

Becker, T., and V. Weispfenning, 1993. *Gröbner Bases*, Springer-Verlag, New York.

Bentley, C. R., 1991. Configuration and structure of the subglacial crust, in *The Geology of Antarctica*, R. J. Tingey (ed.), Clarendon Press, Oxford.

Bentley, C. R., 1973. Crustal structure of antarctica, *Tectonophysics*, **20**, 229-240.

Buchberger, B., 1985. Gröbner Bases: An algorithmic method in polynomial ideal theory, in *Multidimensional Systems Theory*, N. K. Bose (ed.), D. Reidel Publishing Co., Dordrecht.

Cox, D., J. Little, and D. O'Shea, 1997. *Ideals, Varieties, and Algorithms*, Springer, New York.

Craddock, C., 1970. Geologic map of antarctica, *Am Geogr. Soc. Antarctic Map Folio*, **12**.

Dewart, G., and M. N. Toksöz, 1965. Crustal structure in East Antarctica from surface wave dispersion. *Geophys. J.*, **10**, 127-139.

Dieudonne, J., 1985. *History of Algebraic Geometry*, Wadsworth Advanced Books, Monterey.

Dobbs, D., and R. Hanks, 1992. *A Modern Course on the Theory of Equations*, Polygonal Publishing House, Washington, New Jersey.

Doolen, G. D., U. Frisch, B. Hasslacher, S. Orszag, S. Wolfram, 1990. *Lattice Gas Methods for Partial Differential Equations*, Addison-Wesley, Redwood City.

Dorny, C. N., 1980. *A Vector Space Approach to Models and Optimization*, Robert E. Krieger Publishing, Malabar, Florida.

Eisenbud, D., 1995. *Commutative Algebra with a View Toward Algebraic Geometry*, Springer, New York.

Eisenbud, D., C. Huneke, and W. Vasconcelos , 1992. Direct methods for primary decomposition, *Invent. Math.*, **110**, 207-235.

Everett, M. E., 1996. Homotopy, polynomial equations, gross boundary data, and small helmholtz systems, *J. Comp. Phys.*, **124**, 431-441.

Everett, M. E., and A. Schultz, 1996. Geomagnetic induction in a heterogeneous sphere: Azimuthally symmetric test computations and the response of an undulating 660-km discontinuity, *J. Geophys. Res.*, **101**, 2765-2783.

Evison, F. F., C. E. Ingham, and R. H. Orr, 1959. Thickness of the earth's crust in Antarctica, *Nature*, **183**, 306-308.

Faugere, J., P. Gianni, D. Lazard, and T. Mora, 1993. Efficient change of ordering for Gröbner bases of zero-dimensional ideals, *J. Symb. Comp.*, **16**, 329-344.

Fleischer, J., J. Grabmeier, F. W. Hehl, and W. Küchlin, 1995. *Computer Algebra in Science and Engineering*, World Scientific, Singapore.

Floudas, C. A., and V. Visweswaran, 1990. A global optimization algorithm (GOP) for certain classes of nonconvex NLPs-I. Theory, *Computers and Chem. Eng.*, **14**, 1397.

Gallian, J. A., 1990. *Contemporary Abstract Algebra*, D. C. Heath, Lexington.

Gianni, P., B. Trager, and G. Zacharias, 1989. Gröbner Bases and primary decomposition of polynomial ideals, in *Computational Aspects of Commutative Algebra*, L. Robbiano (ed.), Academic Press, New York.

Giusti, M., 1989. Combinatorial dimension of algebraic varieties, in *Computational Aspects of Commutative Algebra*, L. Robbiano (ed.), Academic Press, New York.

Gill, P. E., W. Murray, and M. H. Wright, 1981. *Practical Optimization*, Academic Press, London.

Halmos, P. R., 1957. *Introduction to Hilbert Space*, Chelsea, New York.

Hohmann, G., 1987. Numerical modeling for electromagnetic methods of geophysics in *Electromagnetic Methods in Applied Geophysics*, M. N. Nabighian (ed.), Society of Exploration Geophysicists, Tulsa, OK.

Kendig, K., 1977. *Elementary Algebraic Geometry*, Springer-Verlag, New York.

Kunz, E., 1985. *Introduction to Commutative Algebra and Algebraic Geometry*, Birkhäuser, Boston.

Luenberger, D. G., 1969. *Optimization by Vector Space Methods*, John Wiley and Sons, New York.

Macaulay, F. S., 1902. On some formula in elimination, *Proc. London Math. Soc.*, **3**, 3-27.

Manocha, D., 1994. Solving systems of polynomial equations, *IEEE Comp. Graph. Appl.*, March, 46-55.

Melenk, H., H. M. Möller, and W. Neun, 1989. Symbolic solution of large stationary chemical kinetics problems *Impact of Computing*, **1**, 138-167.

Menke, W., 1989. *Geophysical Data Analysis: Discrete Inverse Theory*, Academic Press, San Diego.

Menke, W., and D. Abbott, 1990. *Geophysical Theory*, Columbia University Press, New York.

Mikhlin, S. G., 1979. *Approximation on a Rectangular Grid*, Sijthoff & Noordhoff, Alphen aan den Rijn.

Miller, D., M. Oristaglio, and G. Beylkin, 1987. A new slant on seismic imaging: Migration and integral geometry, *Geophysics*, **52**, 943-964.

Mishra, B., 1993. *Algorithmic Algebra*, Springer-Verlag, New York.

Parker, R. L., 1994. *Geophysical Inverse Theory*, Princeton University Press, Princeton, New Jersey.

Parker, R. L., and K. Whaler, 1981. Numerical methods for establishing solutions to the inverse problem of electromagnetic induction, *J. Geophys. Res.*, **86**, 9574-9584.

Pederson, P., M.-F. Roy, and A. Szpirglas, 1993. Counting real zeros in the multivariate case, in *Computational Algebraic Geometry*, F. Eyssette, and A. Galligo (eds.), Birkhäuser, Boston.

Pellerin, L., and G. W. Hohmann, 1990. Transient electromagnetic inversion: A remedy for magnetotelluric static shifts, *Geophysics*, **55**, 1242-1250.

Petrick, W. R., W. H. Pelton, and S. H. Ward, 1977. Ridge-regression inversion applied to crustal resistivity sounding data from South Africa, *Geophysics*, **42**, 995-1005.

Press, F., and G. Dewart, 1959. Extent of the Antarctic continent, *Science*, **29**, 462-463.

Robinson, E. S., 1964. *Geologic Structure of the Transantarctic Mountains and Adjacent Ice Covered Areas*, Thesis, University of Wisconsin.

Roult, G., D. Rouland, and J. P. Montagner, 1994. Antarctica II: Upper-mantle structure from velocities and anisotropy, *Phys. Earth Planet. Inter.*, **84**, 33-57.

Sehmi, N. S., 1989. *Large Order Structural Eigenanalysis Techniques*, John Wiley and Sons, New York.

Sherali, H. D., and C. H. Tuncbilek, 1992. A global optimization algorithm for polynomial programming problems using a reformulation-linearization technique, *J. Global Optimization*, **2**, 101-112.

Sloan, I. H., and S. Joe, 1994. *Lattice Methods of Multiple Integration*, Clarendon Press, Oxford.

Snieder, R., 1991. An extension of Backus-Gilbert theory to nonlinear inverse problems, *Inverse Problems*, **7**, 409-433.

Stakgold, I., 1979. *Green's Functions and Boundary Value Problems*, John Wiley and Sons, New York.

Stodt, J. A., 1983. *Processing of conventional and remote reference magnetotelluric data*, Ph. D. Thesis, University of Utah.

Storey, B. C., and T. Alabaster, 1991. Tectonomagmatic controls on Gondwana break-up models: Evidence from the proto-Pacific margin of Antarctica, *Tectonics*, **10**, 1274-1288.

Tignol, J-P., 1988. *Galois' Theory of Algebraic Equations*, Longman Scientific & Technical, Essex.

Toda, M., 1981. *Theory of Nonlinear Lattices*, Springer-Verlag, New York.

Vasco, D. W., 1997. Groups, algebras, and the nonlinearity of geophysical inverse problems, *Geophys. J. Int.*, **131**, 9-23.

Vasco, D. W., 1998. Regularization and trade-off associated with nonlinear geophysical inverse problems: penalty homotopies, *Inverse Problems*, **14**, 1033-1052.

Vasco, D. W., Peterson, J. E., and Majer, E. L., 1996. Nonuniqueness in traveltime tomography: Ensemble inference and cluster analysis, *Geophysics*, **61**, 1209-1227.

Vozoff, K., 1991. The magnetotelluric method, in *Electromagnetic Methods in Applied Geophysics, Vol. II*, M. N. Nabighian (ed.), Society of Exploration Geophysicists, Tulsa, OK.

Vozoff, K., 1985. *Magnetotelluric Methods*, Soc. Explor. Geophys., Tulsa.

Wannamaker, P. E., J. A. Stodt, and S. L. Olsen, 1996. Dormant state of rifting in central west Antarctica implied by magnetotelluric profiling, *Geophys. Res. Lett.*, **23**, 2983-2986.

Watson, L. T., and A. P. Morgan, 1992. Polynomial programming using multi-homogeneous polynomial continuation, *J. Comp. Appl. Math.*, **43**, 373-382.

APPENDIX A: MONOMIAL ORDERINGS

Here I introduce two monomial orderings which shall be used in this paper. Each ordering results in an ideal basis with certain desirable properties. One ordering of importance to us is the following.

Definition (Lexicographic Order) For two index vectors $\alpha = (\alpha_1, ..., \alpha_n)$ and $\beta = (\beta_1, ..., \beta_n)$ we say that α is greater than β with respect to a *lexicographic ordering* ($\alpha >_{lex} \beta$) if the left-most non-zero entry of the vector $\alpha - \beta$ is positive.

This ordering is similar to that used in placing entries in a dictionary, we begin with the left-most letter and work to the right. Consider the terms $z^\alpha = z_1 z_3$ and $z^\beta = z_2 z_4$ in our simple example. The respective index vectors are $\alpha = (1,0,1,0,0,0)$ and $\beta = (0,1,0,1,0,0)$ in the $n = 6$ variable space. And the left-most entry of $\alpha - \beta = (1,-1,1,-1,0,0)$ is positive so that with respect to lexicographic ordering $z_1 z_3 >_{lex} z_2 z_4$. The importance of Lexicographic ordering is that when used to construct a Gröbner basis such an ordering produces a 'triangular' form. One disadvantage of lexicograpic ordering is that it can results in a large Gröbner basis (Adams and Loustaunau 1994).

In lexicographic order the variables themselves dominate the total degree of the monomials, but we could let degree take precedence. For example, there is an ordering in which $z^\alpha > z^\beta$ whenever $|\alpha| > |\beta|$, known as a graded order. We can easily define such an ordering based upon the lexicographic ordering.

Definition (Graded Lexicographic Ordering) Let $\alpha, \beta \in \mathbf{Z}^n$ be integer index vectors. We say that α is larger than β under *graded lexicographic* ordering, denoted by $\alpha >_{grlex} \beta$, if

$$|\alpha| = \sum_{i=1}^{n} \alpha_i > \sum_{i=1}^{n} \beta_i = |\beta|$$

or if $|\alpha| = |\beta|$ and $\alpha >_{lex} \beta$.

The monomials are ordered by total degree and we use lexicographic ordering for cases of equal degree. Note that in our example $z_1 z_3 >_{grlex} z_2 z_4$ but that $z_2 z_4 >_{grlex} z_1$, which is not true using purely lexicographic ordering. As we shall see, graded lexicograpic ordering is important when determining the dimension and total degree of a variety. There are other orderings that are useful in accomplishing various tasks. For example, elimination ordering groups collections of variables and produces a type of lexicographic ordering between the groups (Adams and Loustaunau 1994). It is useful because it generates a less complicated basis set while still producing equations in which a set of variables is eliminated.

APPENDIX B: BUCHBERGER'S ALGORITHM

In this appendix the steps in Buchberger's (1985) algorithm to compute a Gröbner basis are outlined.

S-polynomials

A Gröbner basis is a very useful computational tool for determining properties of the zero sets of polynomial equations and their associated algebraic structures polynomial ideals. How does one find a generating set $\{g_1, ...g_l\}$ which is a Gröbner basis? The main impediment to realizing a Gröbner basis is when the leading terms of some combination of polynomials in the generating set cancel, leaving only smaller terms (Cox et al. 1997, p. 81). That is, given two polynomials f_i, f_j in the generating set of the ideal I, and monomials a and b, the leading terms of

$$a \cdot f_i - b \cdot f_j$$

cancel, producing a polynomial of lower total degree. A measure of such cancellations is provided by the S-polynomial

Definition The *S-polynomial* of f and g is given by

$$S(f,g) = \frac{z^\gamma}{\mathrm{LT}(f)} \cdot f - \frac{z^\gamma}{\mathrm{LT}(g)} \cdot g$$

where γ is the index vector with entries $\gamma_i = \max(\alpha_i, \beta_i)$ for each i.

Note that the monomial ordering enters the computation of the S-polynomials in the determination of the leading terms and in calculating γ. For our simple example, considering the first two elements of I, (f_1, f_2): $\gamma = (1,1,1,1,0,0)$, $z^\gamma = z_1 z_2 z_3 z_4$, $\mathrm{LT}(f_1) = z_1 z_3$, $\mathrm{LT}(f_2) = z_2 z_4$, and using graded lexicograpic ordering,

$$S(f_1, f_2) = z_2 z_4 \cdot f_1 - z_1 z_3 \cdot f_2$$

$$= -{z_1}^2 z_3 + 4 z_1 z_2 z_3 - 4 z_1 z_2 z_4 + {z_2}^2 z_4 - c_2 z_1 z_3 + c_1 z_2 z_4$$

Similarly, $S(f_2, f_3) = -z_2 z_4 z_5 + {z_1}^2 - 4 z_1 z_2 - c_3 z_2 z_4 + c_2 z_1$ and $S(f_3, f_4) = c_4 z_1 + c_3 z_2 - z_1 z_6 + z_2 z_5$. By analogy with Gaussian elimination, the S-polynomials, coupled with lexicographic ordering, produce a type of pivoting step in the triangularization algorithm. It is shown in Cox et al. (1997, p. 81) that S-polynomials account for all cancellations within a generating set of an ideal. In fact, there is an equivalent S-polynomial criterion for the basis of an ideal to be a Gröbner basis,

Theorem *A basis $G = \{g_1, ..., g_q\}$ for an ideal I is a Gröbner basis if and only if the remainder on division of $S(g_i, g_j)$ by G is zero for all pairs $i \neq j$.*

Division by a set of polynomials is a generalization of the standard polynomial division algorithm in one variable (Gallian 1990) (f is successively divided by elements of the Gröbner basis), to represent f in terms of the elements of the Gröbner basis $\{g_1, ..., g_l\}$

$$f = h_1 g_1 + \cdots h_l g_l + r$$

where h_i are elements of $k[z_1, ..., z_n]$ and r is the polynomial remainder of total degree less than f. The polynomial f is an element of I if and only if the remainder r is zero. A compact notation for the remainder of f upon division by a set of polynomials $G = (g_1, ..., g_l)$ is $\overline{f}^G$. The above theorem provides an algorithmic criterion which is much more useful than the motivating definition of Gröbner bases. A proof is given in Mishra (1993, p. 57) and Cox et al. (1997, p. 82).

Buchberger's Algorithm

We are now in a position to write down an explicit algorithm for the construction of a Gröbner basis. In essence, the approach is to add more polynomials from I to the original generating set $f_1, ..., f_p$ until the conditions for a Gröbner basis are satisfied. The elements which we add derive from the S-polynomials defined above. In particular, we must keep adding the remainders of the S-polynomials upon division by the current basis set.

Theorem (Buchberger's Algorithm) *Let $I = \langle f_1, ..., f_p \rangle \neq \{0\}$ be a polynomial ideal. A Gröbner basis for I can be constructed in a finite number of steps by the algorithm:*

$$\text{INITIALIZE}: \quad F = (f_1, ..., f_p)$$

$$G \equiv F$$

REPEAT *until $G = G'$:*

$$G' \equiv G$$

FOR *each pair $\{p, q\}, p \neq q$ in G'* DO

$$S \equiv \overline{S(p,q)}^{G'}$$

$$\text{IF } S \neq 0 \text{ THEN } G \equiv G \cup \{S\}$$

The above algorithm is a simplified version of Buchberger's algorithm, detailed enough to understand the underlying principle's. However, many improvements are needed to produce an algorithm efficient enough for applications (Mishra 1993, Adams and Loustaunau 1994, Cox et al. 1997). One of the first improvements involves casting out unneeded generators from the Gröbner basis and writing the basis in a canonical form. The basic idea is encapsulated in the following theorem (Cox et al. 1997, p. 89)

Theorem *If p is a member of G, a Gröbner basis for the ideal I and $\mathrm{LT}(p) \in \langle \mathrm{LT}(G - \{p\}) \rangle$, then $G - \{p\}$ is also a Gröbner basis for I .*

This theorem leads to the idea of a reduced form for the Gröbner basis, the formal definition is

Definition A *reduced Gröbner basis* for an ideal I is a Gröbner basis such that all leading coefficients are unity and for all $p \in G$, no monomial of p lies in $\langle \mathrm{LT}(G - \{p\}) \rangle$.

Let us cycle through Buchberger's algorithm using our illustrative example and graded lexicographic ordering. First, we begin with (f_1, f_2, f_3, f_4) and compute the S-polynomials. From the computations above we know that $S(f_1, f_2) = -{z_1}^2 z_3 + 4 z_1 z_2 z_3 - 4 z_1 z_2 z_4 + {z_2}^2 z_4 - c_2 z_1 z_3 + c_1 z_2 z_4$. Now we compute the remainder upon division by the set of polynomials f_1, f_2, f_3, f_4. After successive polynomial division by the elements $G = (f_1, f_2, f_3, f_4)$ we find that we may write the S-polynomial as the sum

$$S(f_1, f_2) = (4z_2 + z_5 - c_2 + c_3) \cdot f_1$$

$$+(-4z_1 + z_2 + c_1) \cdot f_2 + (-z_1 z_3 + 4z_1 - z_2 - c_1) \cdot f_3$$

with remainder $\overline{S(f_1, f_2)}^G = 0$. Thus, we do not need to add any elements to the initial basis set due to $S(f_1, f_2)$. Similarly, we find that

$$S(f_2, f_3) = z_1 \cdot f_2 - z_2 z_4 \cdot f_3$$

and $\overline{S(f_2, f_3)}^G = 0$. Continuing on, we may write

$$S(f_3, f_4) = -(z_6 + c_4) \cdot f_3 + (z_5 + c_3) \cdot f_4$$

with remainder $\overline{S(f_3, f_4)}^G = -2c_4 z_5 - 2c_3 c_4$. We must now enlarge the basis set to include $f_5 = -2c_4 z_5 - 2c_3 c_4$. We proceed in this manner, considering all remaining $S(f_i, f_j)$ combinations where both i and j now vary from 1 to 5. As more non-zero remainders are encountered we enlarge the set G. Upon completion of the algorithm, which was executed using the public domain software CoCoA (Adams and Loustaunau 1994), we find a reduced Gröbner basis G

$$\begin{aligned}
g_1 &= z_1 + z_5 + {c_1}' \\
g_2 &= z_2 + z_6 + {c_2}' \\
g_3 &= z_4 z_6 + c_4 z_4 + z_5 - 4z_6 + {c_3}' \\
g_5 &= z_3 z_5 + c_3 z_3 - 4z_5 + z_6 + {c_4}'
\end{aligned}$$

for the example ideal I, where ${c_1}' = c_3, {c_2}' = c_4, {c_3}' = -c_2 + c_3 - 4c_4$, and ${c_4}' = -c_1 - 4c_3 + c_4$. Note, the in the reduction process the set of generators was reduced from 5 to 4 .

FIGURE CAPTIONS

Figure 1. Illustration of the monomials contained in the ideal $I = \langle {z_1}^5 {z_2}^2, {z_1}^3 {z_2}^4 \rangle$. The filled squares and the shaded region denote monomials generated by the elements of I up to degree 6 in each variable. The crosses indicate monomials that are 'missing' from I, up to degree 6 in each variable.

Figure 2. Two-dimensional sections through the four-dimensional model space $(\sigma_1, \sigma_2, \sigma_3, \sigma_4)$. The contours denote the sum of the squares of the gradient components for each parameter in the cross section. The weighting for the penalty term is 0 in this case. The vanishing of the gradient is a necessary condition for an extremum.

Figure 3. As in Figure 2, two-dimensional sections through the four-dimensional model space $(\sigma_1, \sigma_2, \sigma_3, \sigma_4)$. The weighting for the penalty term is 10^{-6} in this case.

Figure 4. Histograms of the number of 'real' solutions of the inverse problem for the penalized Helmholtz equation. These values of σ_1 minimize $W(\sigma)$ in equation (11). Each panel displays the distribution of 'real' solutions for varying levels of Gaussian noise in the field values. A penalty weighting term of 10^{-6} was used.

Figure 5. As in Figure 4, the distribution of 'real' solutions for all four parameters $(\sigma_1, \sigma_2, \sigma_3, \sigma_4)$. In this case the field values are contaminated by 5% Gaussian noise. A penalty weighting term of 10^{-6} was used.

Figure 6. As in Figure 4, the distribution of σ_1 for the 'real' solutions at differing penalty weights ν in equation (11) for a noise level of 5%.

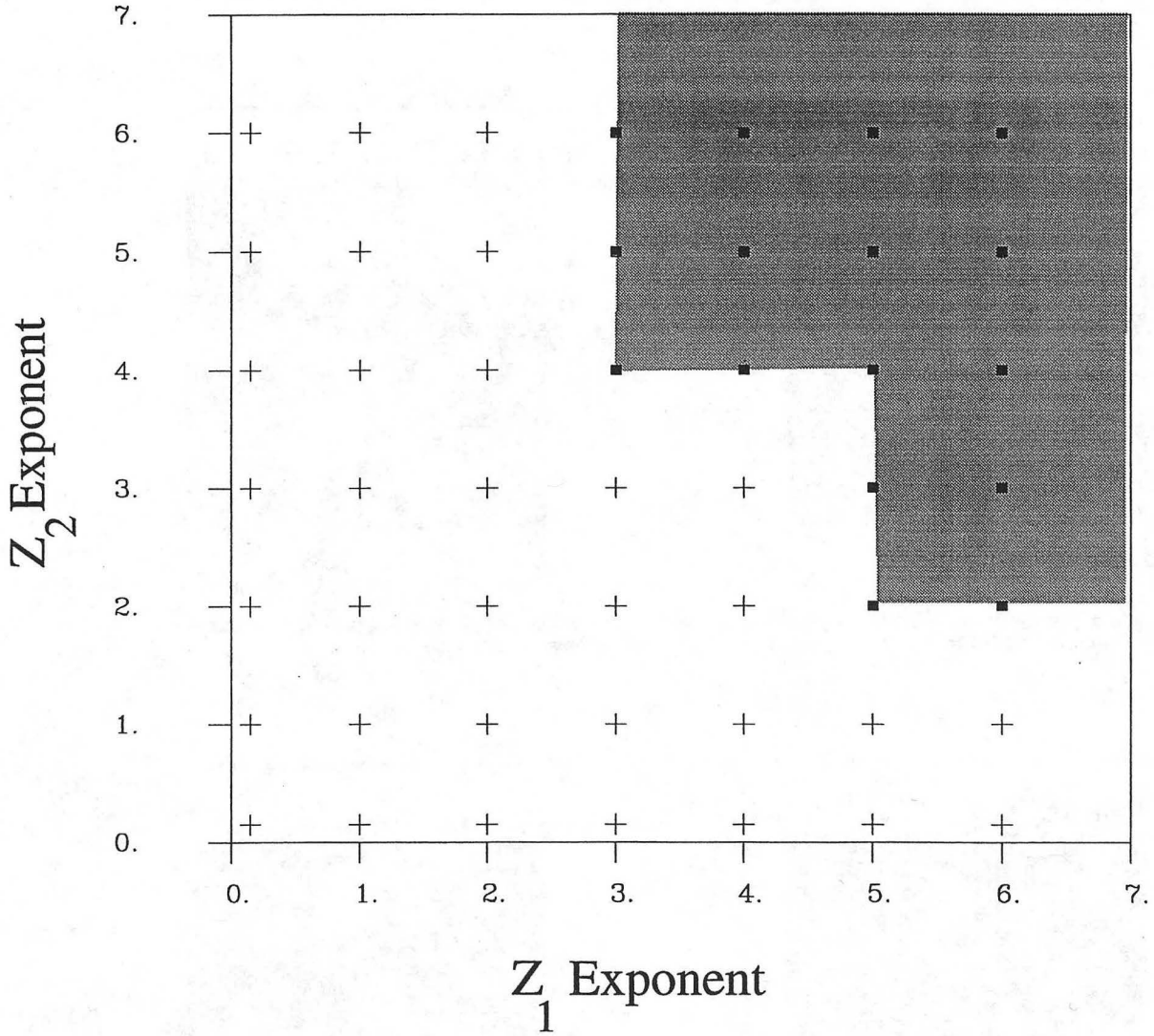
Figure 7. Station distribution for the Antarctic magnetotelluric experiment. The origin of the survey is indicated by the filled square. The filled circle denotes the approximate locate of South Pole Station. The receivers, indicated by the crosses, were laid out along a line of azimuth of 210°, positive values of distance are in that direction.

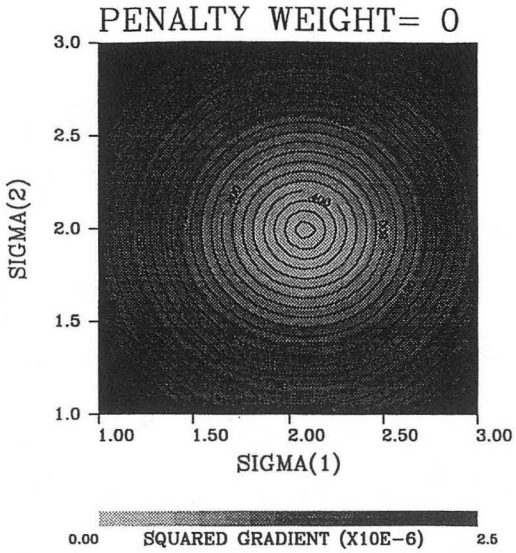
Figure 8. Pseudosections of apparent resistivities ρ_{xy} and ρ_{yx} for the South Pole experiment. The contours are in ohm-m and positive distance along the line is in the direction of azimuth 210°.

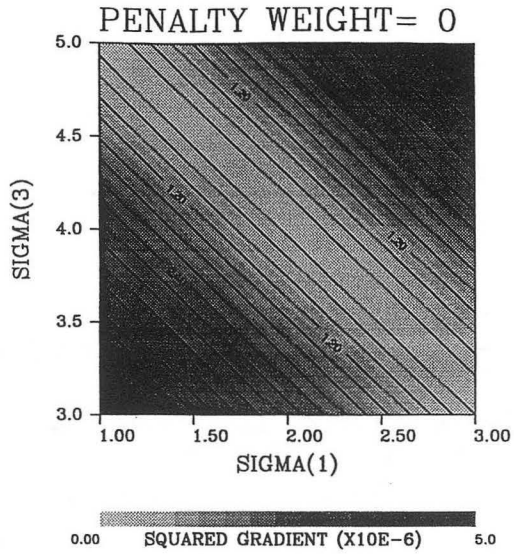
Figure 9. Pseudosections of phase ϕ_{xy} and ϕ_{yx} for the South Pole experiment. The phase angles are a function of the logarithm of the period and the positive distance along the line in the direction of azimuth 210°.

Figure 10. Best fitting solution to the South Pole electric and magnetic observations. The depth scale is logarithmic and the horizontal lines show the location of the three major interfaces in the model. The nodal values have been linearly interpolated onto a finer grid for the purposes of plotting.

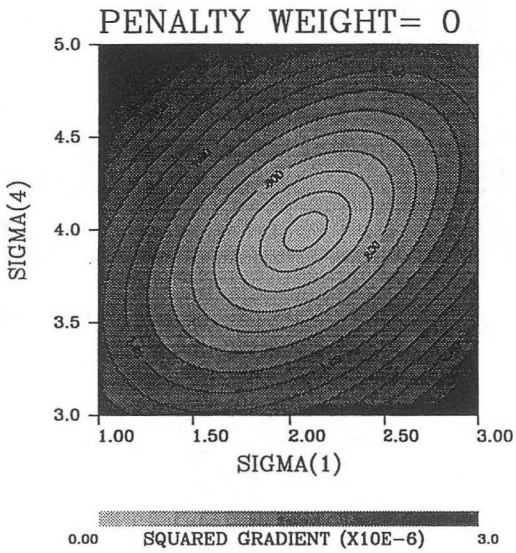
Figure 11. Ratio of final to prior misfit in percent. Misfit is defined as the absolute value of the right-hand-side of the polynomial equations, given by the absolute values of $F_{ij}(\sigma)$ in equation (11). The data set is composed of observations for 34 frequencies at the ten stations.


7.
6.
5.
4.
3.
2.
1.
0.
0.
1.
2.
3.
4.
5.
6.
7.
Z_2 Exponent
Z_1 Exponent

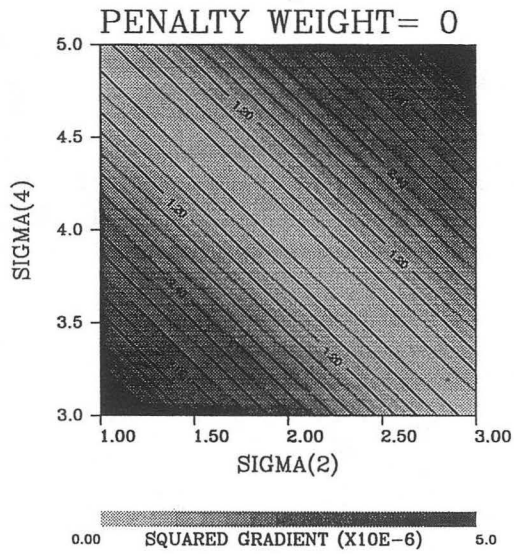




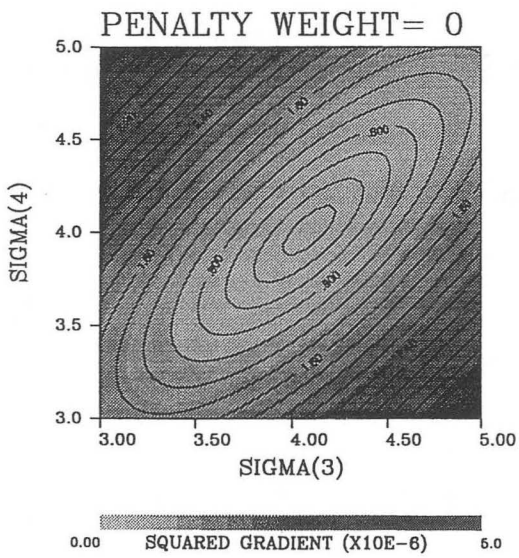
CONTOUR FROM 0. TO 2.7000 CONTOUR INTERVAL OF 0.10000 PT(3,3)= 2.6639



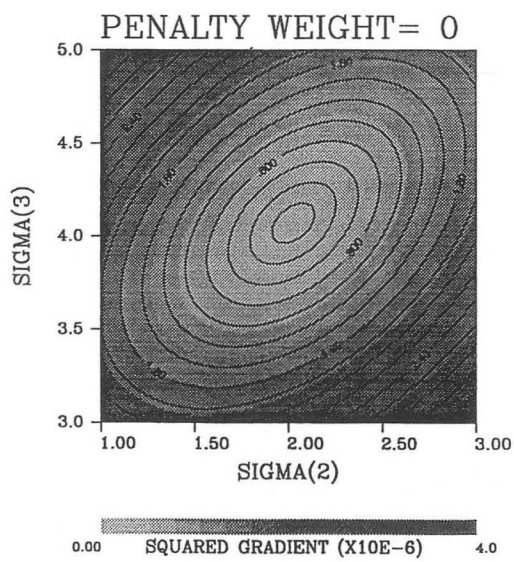
CONTOUR FROM 0. TO 5.4000 CONTOUR INTERVAL OF 0.30000 PT(3,3)= 5.3602

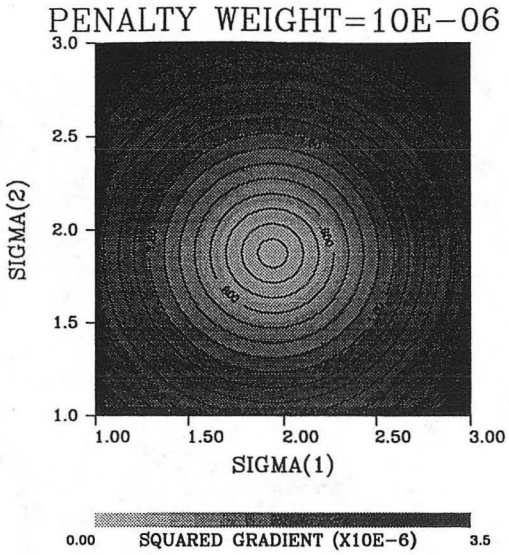


CONTOUR FROM 0. TO 3.4000 CONTOUR INTERVAL OF 0.20000 PT(3,3)= 2.0913

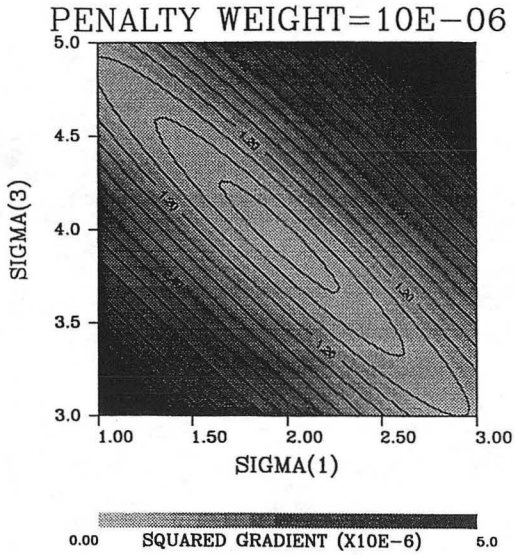


CONTOUR FROM 0. TO 5.4000 CONTOUR INTERVAL OF 0.30000 PT(3,3)= 5.2850

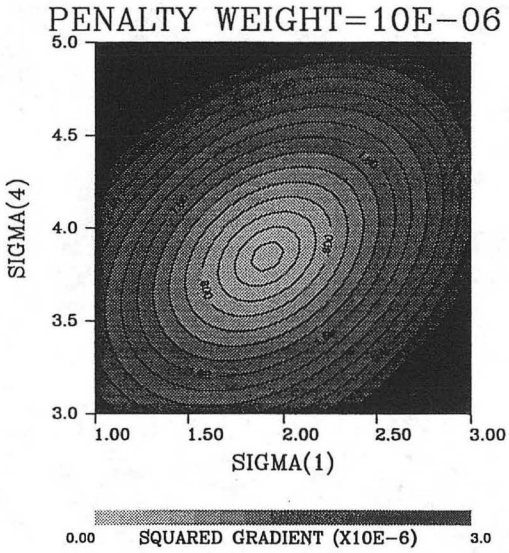




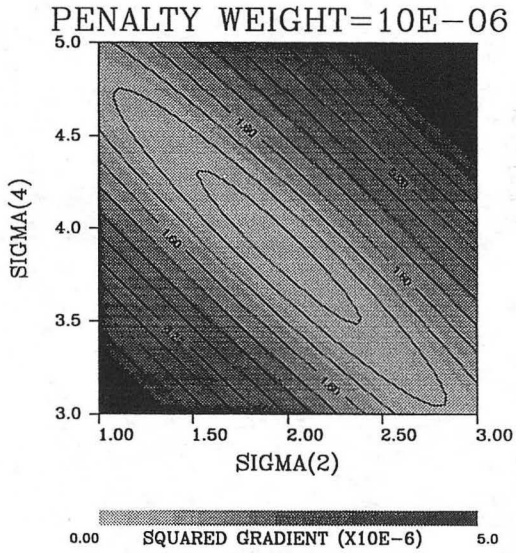
CONTOUR FROM 0. TO 3.6000 CONTOUR INTERVAL OF 0.20000 PT(3,3)= 3.0817



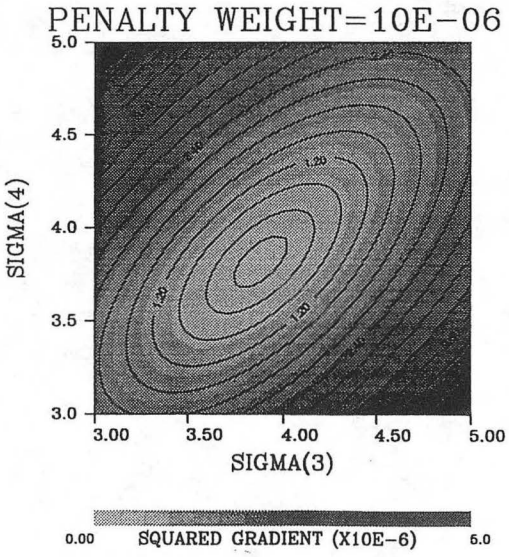
CONTOUR FROM 0. TO 6.3000 CONTOUR INTERVAL OF 0.30000 PT(3,3)= 5.7877

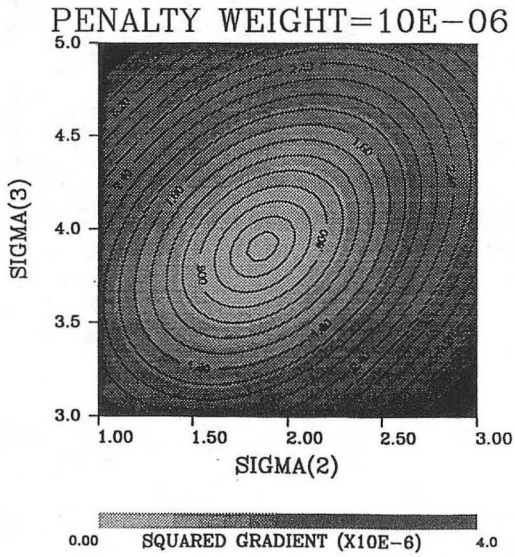


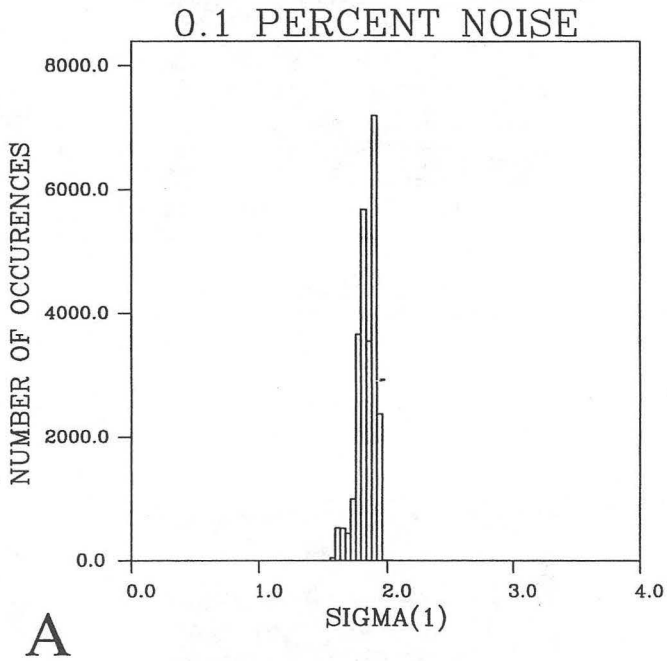
CONTOUR FROM 0. TO 4.4000 CONTOUR INTERVAL OF 0.20000 PT(3,3)= 2.4893



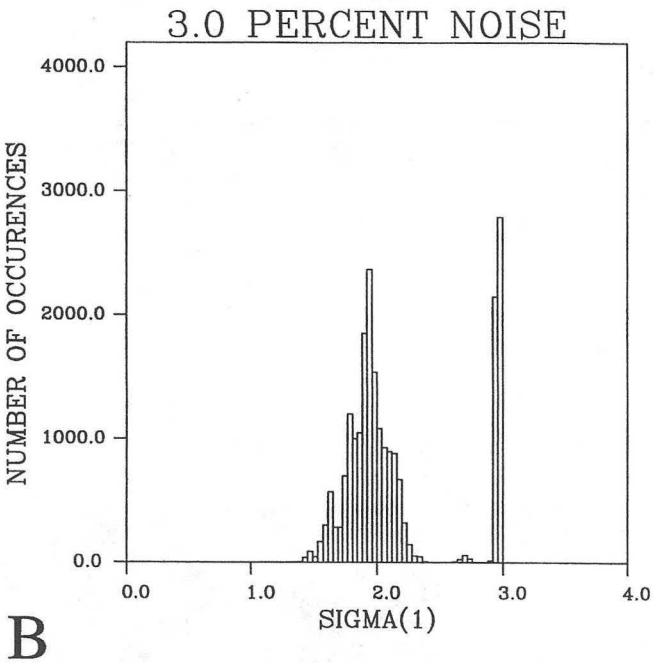
CONTOUR FROM 0. TO 6.4000 CONTOUR INTERVAL OF 0.40000 PT(3,3)= 5.6626



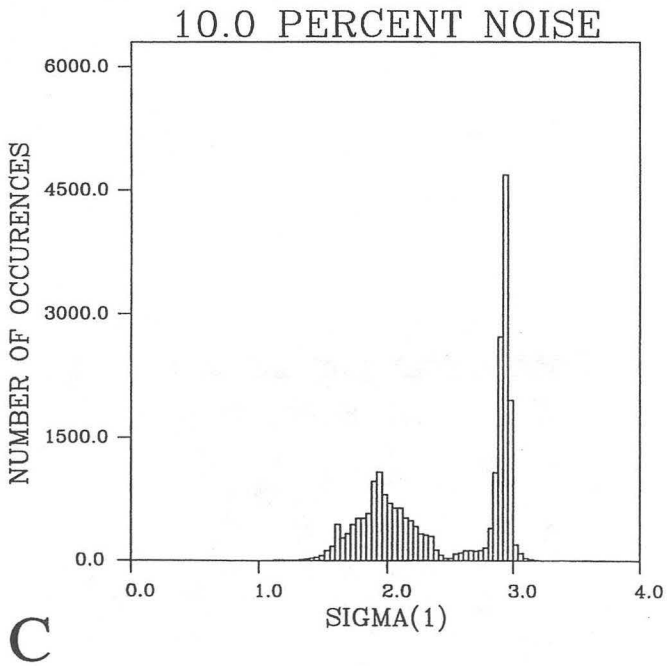



0.1 PERCENT NOISE
NUMBER OF OCCURENCES
8000.0
6000.0
4000.0
2000.0
0.0
0.0
1.0
2.0
3.0
4.0
SIGMA(1)

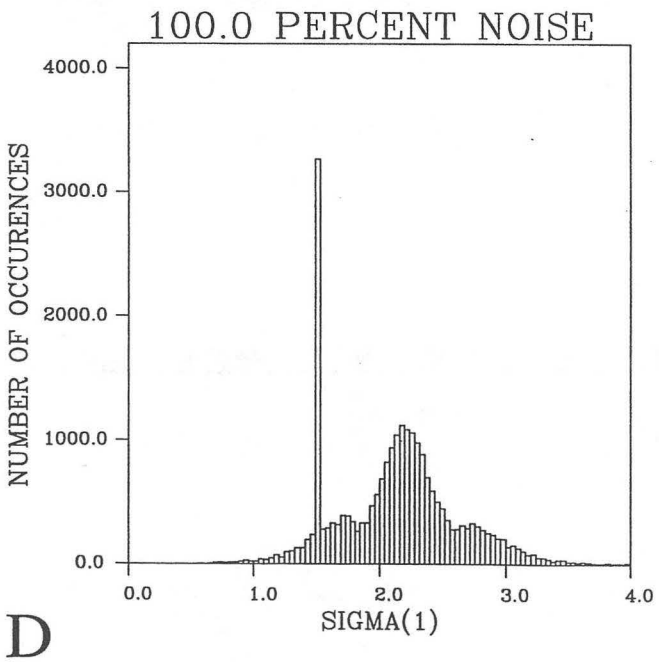
A


3.0 PERCENT NOISE
NUMBER OF OCCURENCES
4000.0
3000.0
2000.0
1000.0
0.0
0.0
1.0
2.0
3.0
4.0
SIGMA(1)

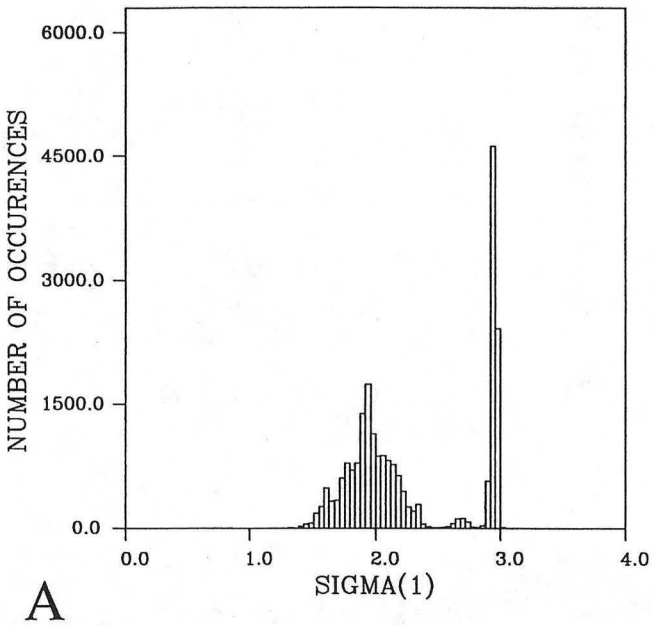
B


10.0 PERCENT NOISE
NUMBER OF OCCURENCES
6000.0
4500.0
3000.0
1500.0
0.0
0.0
1.0
2.0
3.0
4.0
SIGMA(1)

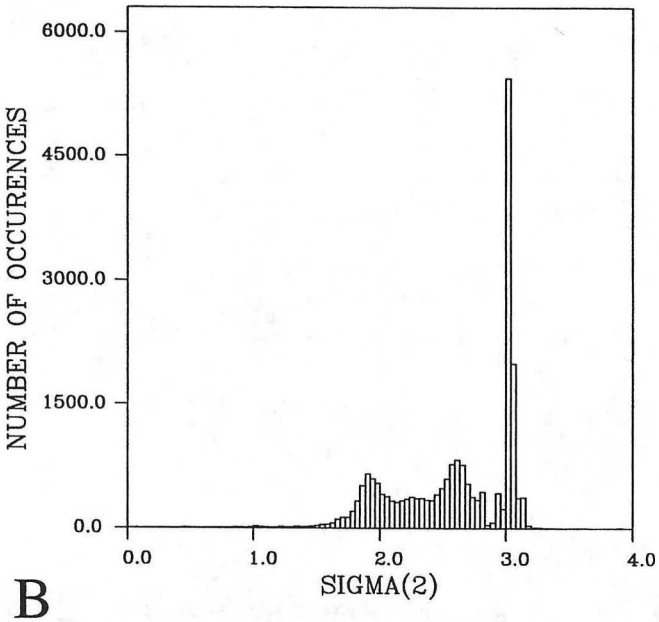
C


100.0 PERCENT NOISE
NUMBER OF OCCURENCES
4000.0
3000.0
2000.0
1000.0
0.0
0.0
1.0
2.0
3.0
4.0
SIGMA(1)

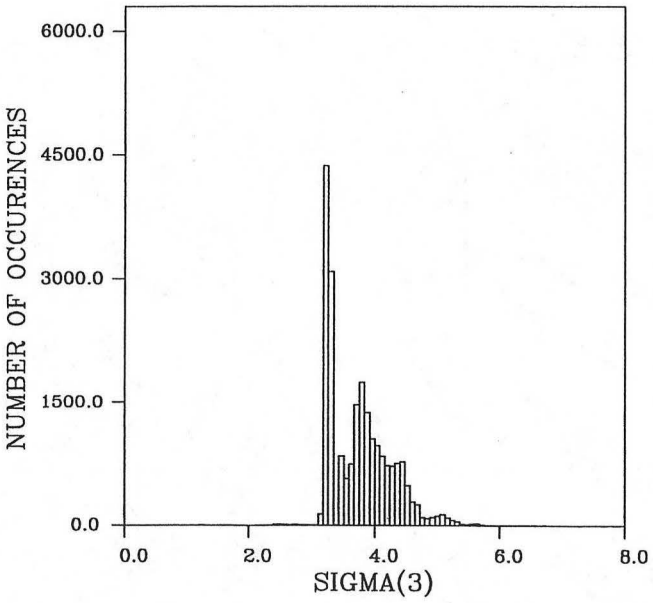
D


NUMBER OF OCCURENCES
6000.0
4500.0
3000.0
1500.0
0.0
0.0
1.0
2.0
3.0
4.0
SIGMA(1)

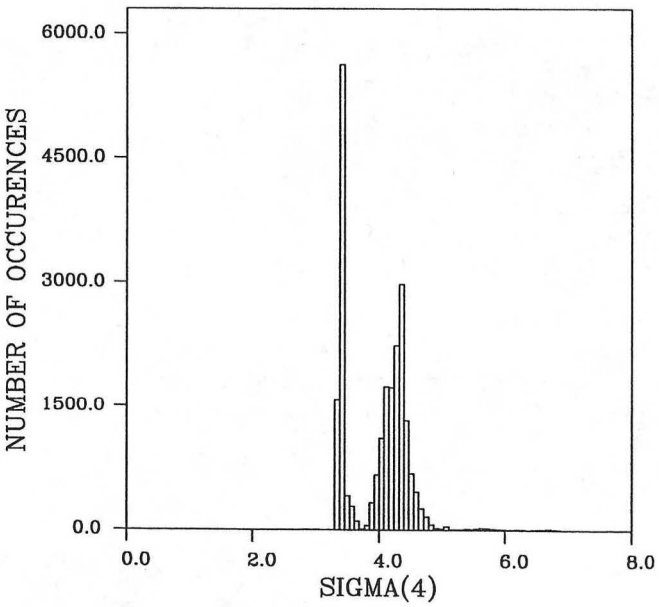
A


NUMBER OF OCCURENCES
6000.0
4500.0
3000.0
1500.0
0.0
0.0
1.0
2.0
3.0
4.0
SIGMA(2)

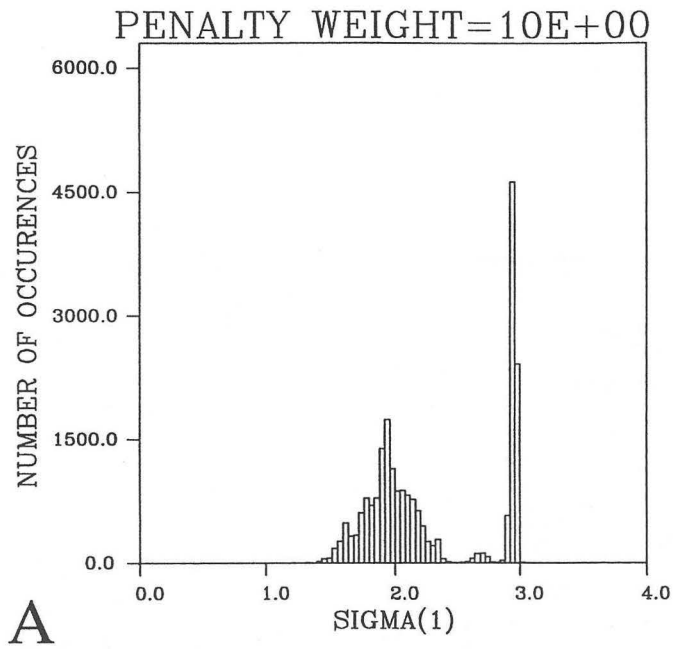
B

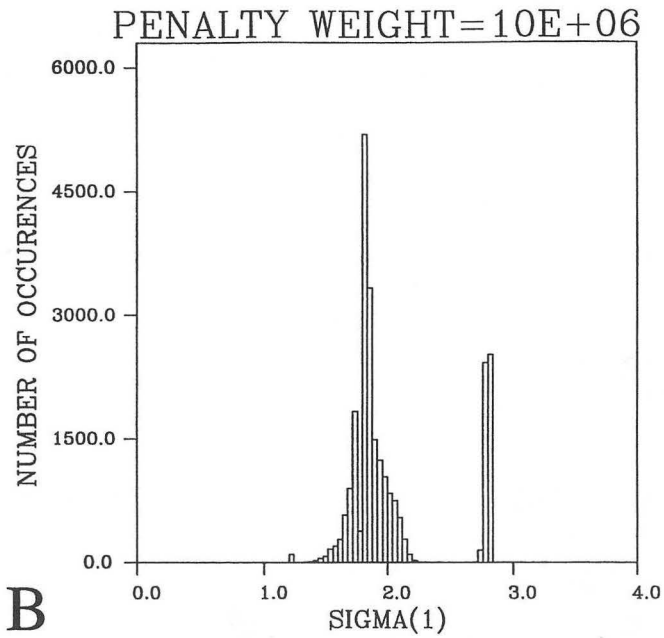

NUMBER OF OCCURENCES
6000.0
4500.0
3000.0
1500.0
0.0
0.0
2.0
4.0
6.0
8.0
SIGMA(3)

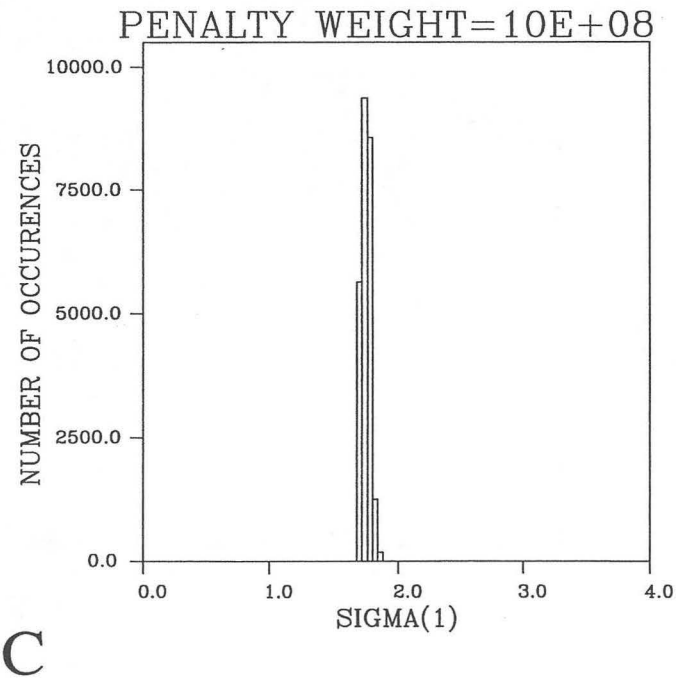
C

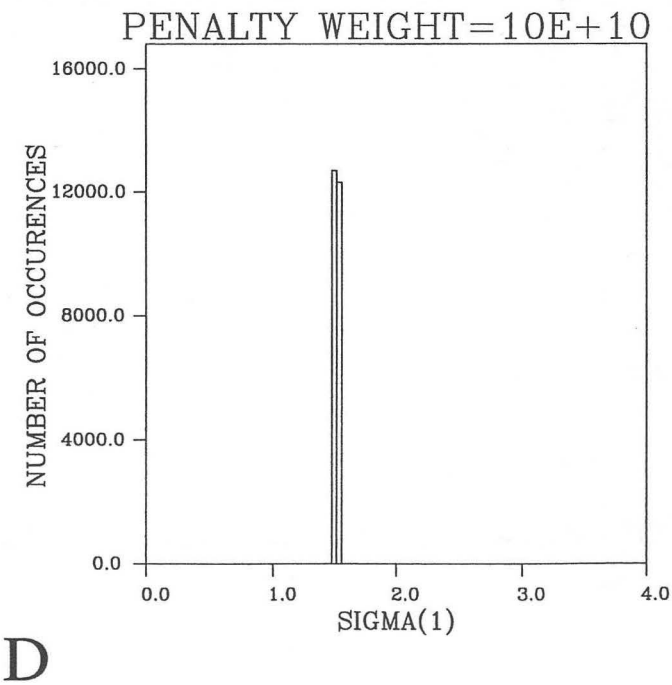

NUMBER OF OCCURENCES
6000.0
4500.0
3000.0
1500.0
0.0
0.0
2.0
4.0
6.0
8.0
SIGMA(4)

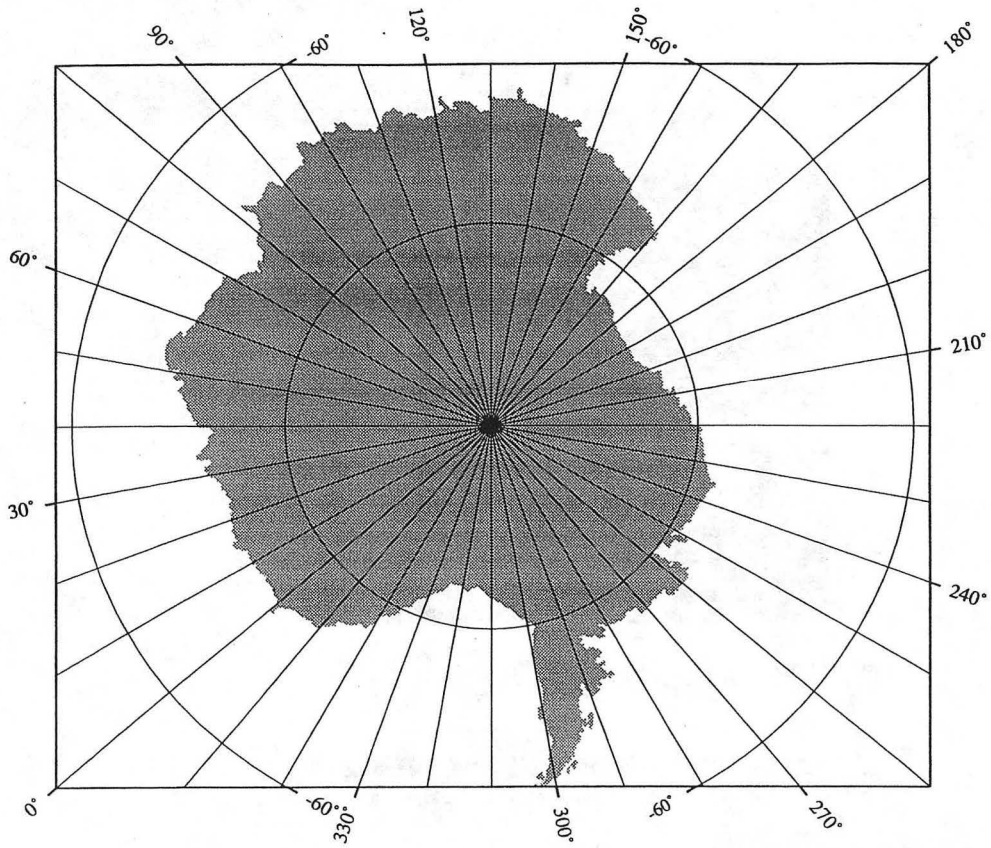
D

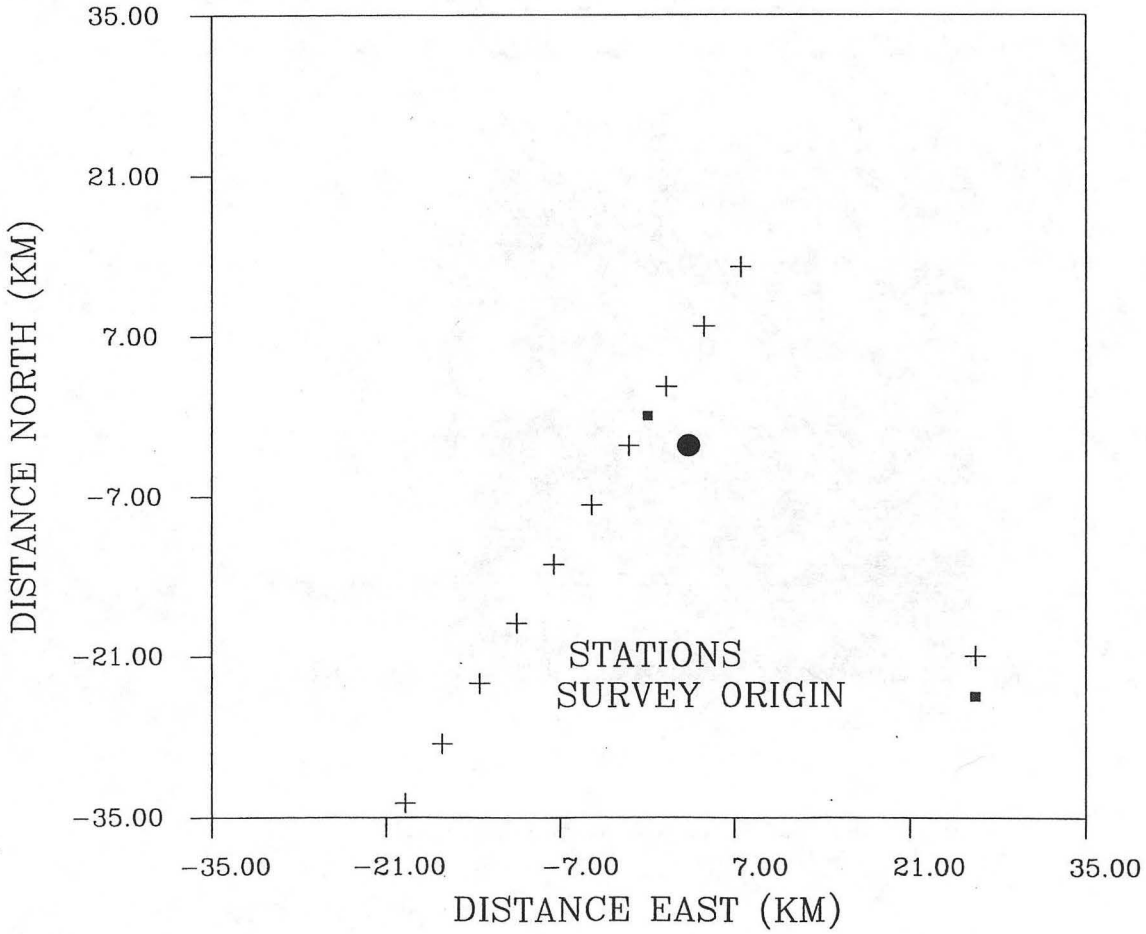

PENALTY WEIGHT=10E+00
NUMBER OF OCCURENCES
6000.0
4500.0
3000.0
1500.0
0.0
0.0
1.0
2.0
3.0
4.0
SIGMA(1)
A


PENALTY WEIGHT=10E+06
NUMBER OF OCCURENCES
6000.0
4500.0
3000.0
1500.0
0.0
0.0
1.0
2.0
3.0
4.0
SIGMA(1)
B

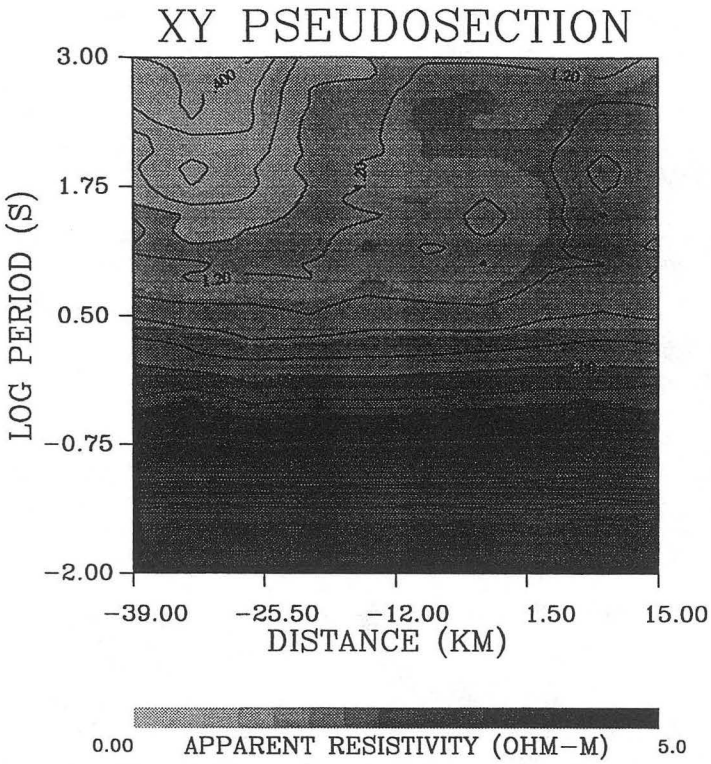

PENALTY WEIGHT=10E+08
NUMBER OF OCCURENCES
10000.0
7500.0
5000.0
2500.0
0.0
0.0
1.0
2.0
3.0
4.0
SIGMA(1)
C


PENALTY WEIGHT=10E+10
NUMBER OF OCCURENCES
16000.0
12000.0
8000.0
4000.0
0.0
0.0
1.0
2.0
3.0
4.0
SIGMA(1)
D


90°
-60°
120°
150°
-60°
180°
60°
210°
30°
240°
0°
-60°
330°
300°
-60°
270°

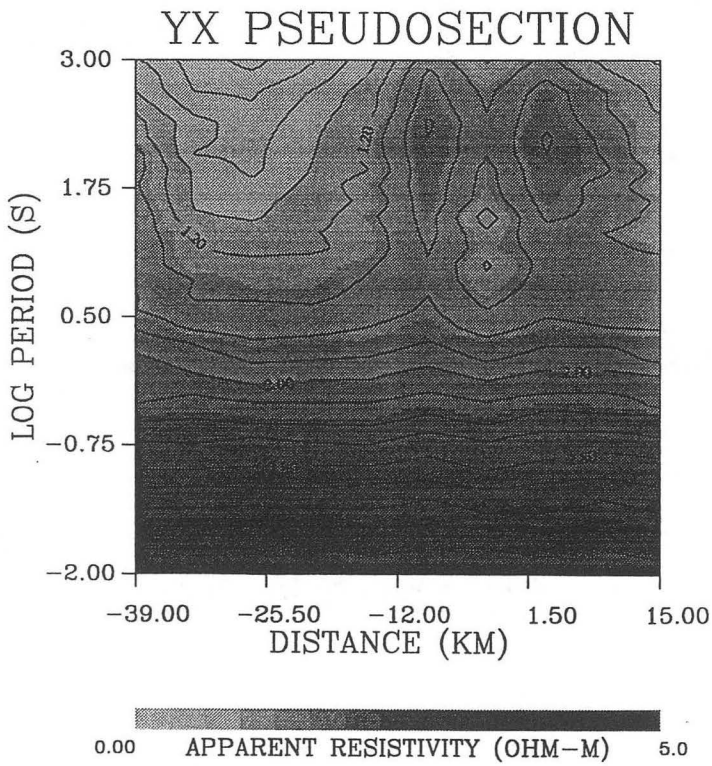

35.00
21.00
7.00
-7.00
-21.00
-35.00
DISTANCE NORTH (KM)
STATIONS
SURVEY ORIGIN
-35.00
-21.00
-7.00
7.00
21.00
35.00
DISTANCE EAST (KM)

Apparent Resistivity



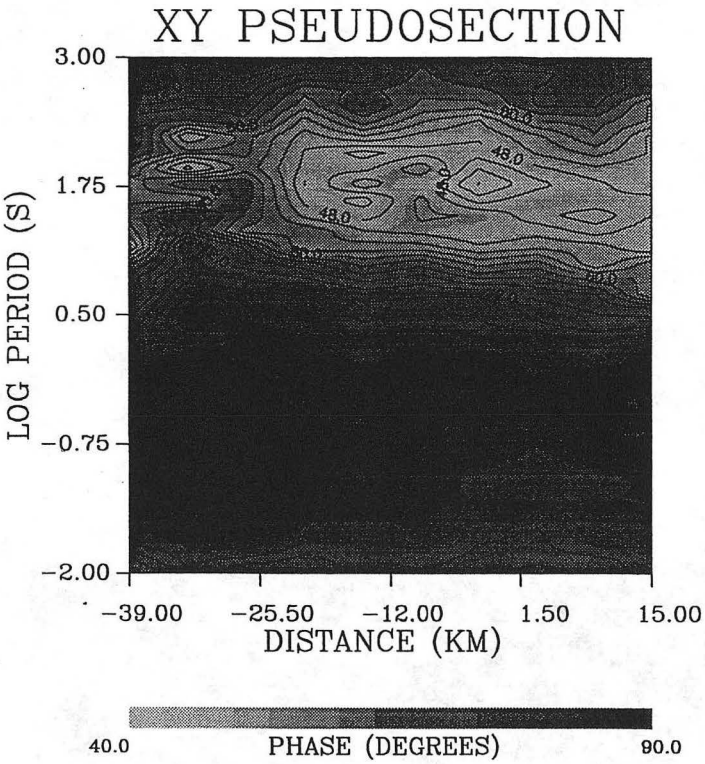
A

CONTOUR FROM 0.20000 TO 3.6000 CONTOUR INTERVAL OF 0.20000 PT(3,3)= 3.7219



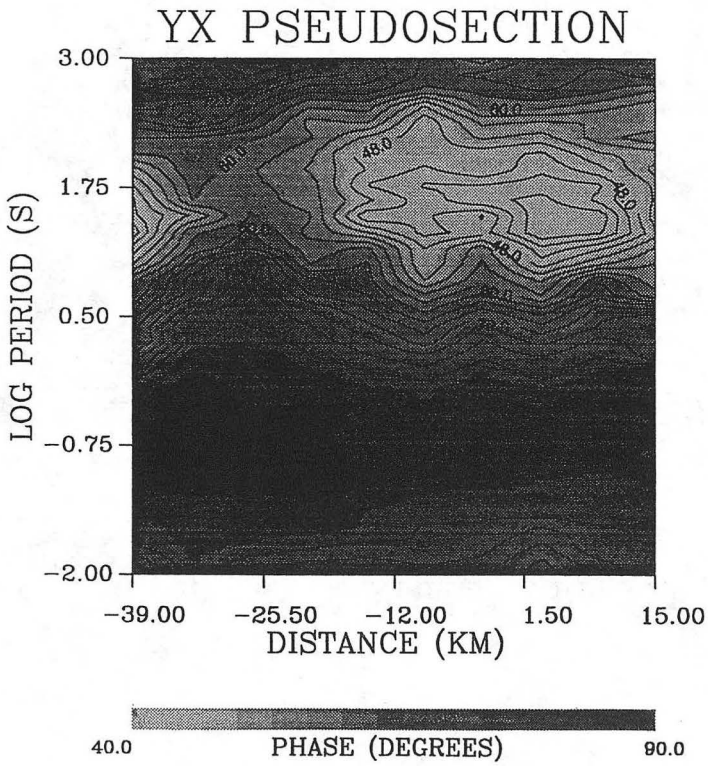
B

Phase

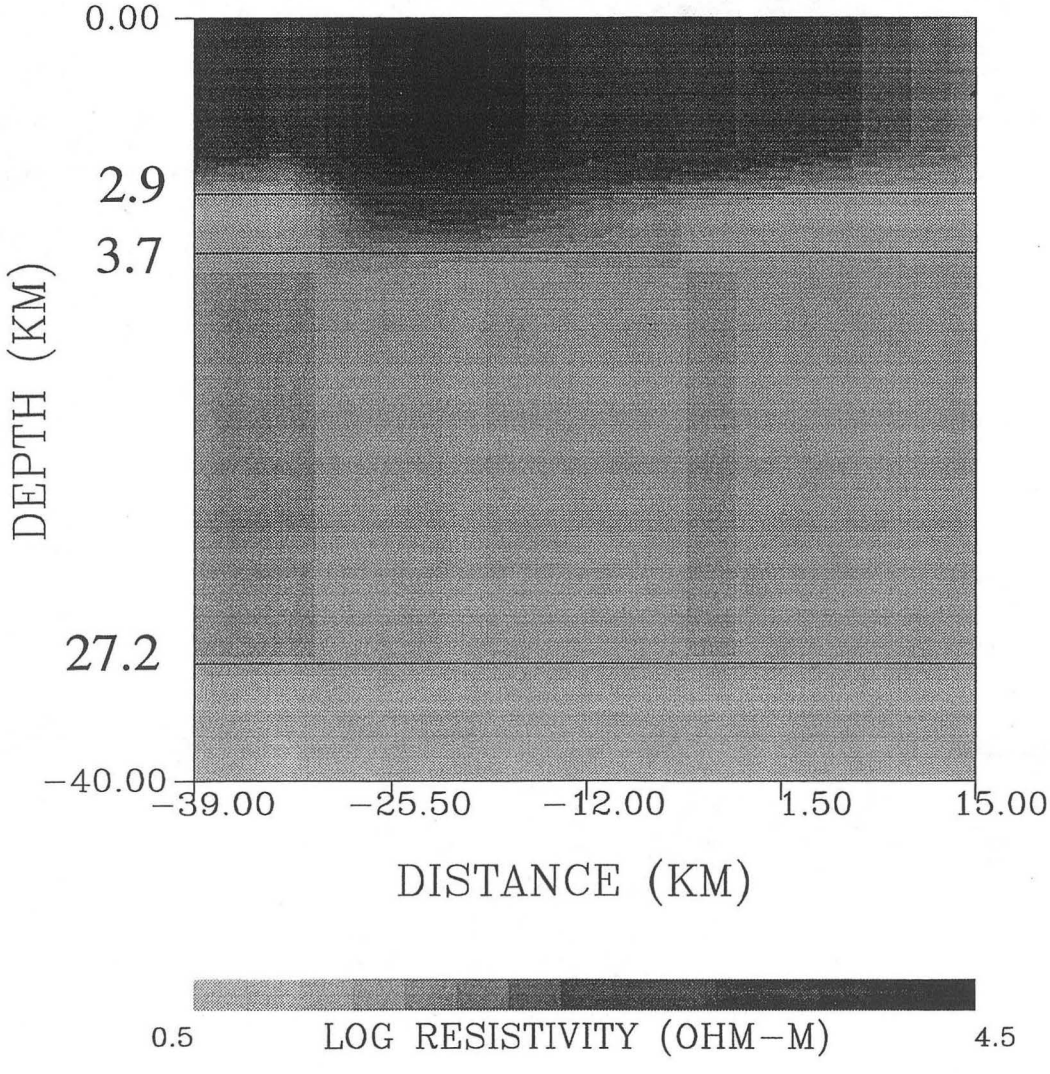


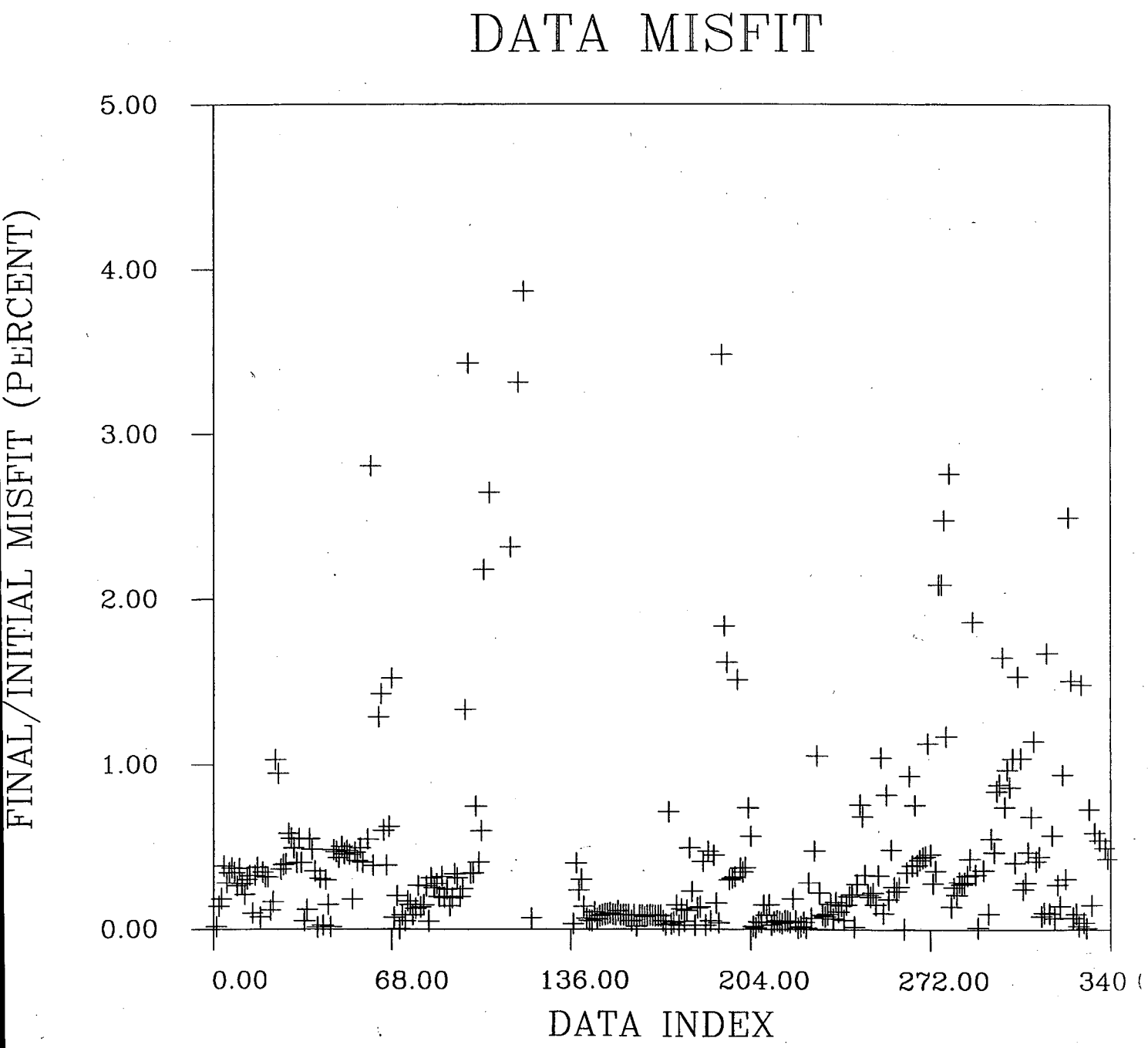
A

CONTOUR FROM 36.000 TO 90.000 CONTOUR INTERVAL OF 3.0000 PT(3,3)= 73.109



B


0.00
2.9
3.7
27.2
−40.00
DEPTH (KM)
−39.00
−25.50
−12.00
1.50
15.00
DISTANCE (KM)
0.5
LOG RESISTIVITY (OHM−M)
4.5


DATA MISFIT
5.00
4.00
3.00
2.00
1.00
0.00
FINAL/INITIAL MISFIT (PERCENT)
0.00
68.00
136.00
204.00
272.00
340
DATA INDEX

Ernest Orlando Lawrence Berkeley National Laboratory
One Cyclotron Road | Berkeley, California 94720

Prepared for the U.S. Department of Energy under Contract No. DE-AC03-76SF00098