Assessing non-uniqueness: An algebraic approach

D. W. Vasco*

(September 16, 2002)

ABSTRACT

Geophysical inverse problems are endowed with a rich mathematical structure. When discretized, most differential and integral equations of interest are algebraic (polynomial) in form. Techniques from algebraic geometry and computational algebra provide a means to address questions of existence and uniqueness for both linear and non-linear inverse problem. In a sense, the methods extend ideas which have proven fruitful in treating linear inverse problems.

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

INTRODUCTION

From a mathematical standpoint, geophysical inverse problems are highly structured. The rich algebraic structure of linear inverse problems is well know and extensively exploited. For example, ideas such as Hilbert spaces, linear functionals, and spectral expansion have proven fruitful in treating linear problems (Halmos 1957, Luenberger 1969, Dorny 1980). For non-linear inverse problems the associated mathematical structure has not been extensively utilized. For the most part, non-linear problems are simply linearized and methods from linear inverse theory employed (Parker 1994). Alternatively, purely stochastic techniques are frequently used in treating the inverse problem, and the inherent mathematical structure is effectively ignored (Tarantola 1987). In truth, most non-linear geophysical inverse problems possess a high degree of mathematical structure which may be used to some advantage in addressing important questions related to existence and uniqueness. In this paper I discuss an algebraic framework in which both linear and non-linear geophysical inverse problems may be treated. I point to developments in computational algebra and numerical algebraic geometry which are useful in attacking geophysical inverse problems. The techniques outlined in this paper are given a more extensive treatment in Vasco (1999, 2000). It is still early in the development and implementation of these methods. Thus, we might expect substantial improvements in the current algorithms over the coming years.

Inverse problems as algebraic equations:

As pointed out by Everett (1996), many of the differential equations governing geophysical phenomena take the form of algebraic equations, that is polynomial equations, when discretized. Specifically, the equations governing geophysical phenomena are partial differential equations with spatially varying coefficients (Menke and Abbott 1990). The spatially varying coefficients signify variations in physical properties within the Earth. Often the equations governing the field variations are linear partial differential equations. Another large class of equations govern geophysical flows in the Earth, oceans, and atmosphere. Such equations describe the evolution, in space and time, of quantities such as the magnetic and electric fields within the Earth, the displacement fields associated with earthquakes, fluid pressure variations in the subsurface, the phase field of a wave traveling through the Earth. A simple example is the Helmholtz equation

$$\nabla^2 u + i\omega\sigma(\mathbf{x})u = 0, \tag{1}$$

where $u(\mathbf{x}, \omega)$ is a component of the electric or magnetic field and $\sigma(\mathbf{x})$ is the electrical conductivity. Another well known equation is the Eikonal equation,

$$\nabla T \cdot \nabla T - \sigma(\mathbf{x}) = 0 \tag{2}$$

where $T(\mathbf{x})$ is the spatially varying phase field of a wave propagating through the Earth and $\sigma(\mathbf{x})$ is the square of the slowness (Aki and Richards 1980). The equation governing the evolution of hydraulic head $H(\mathbf{x}, \omega)$ in an aquifer is a bit more involved. In that case there are two properties governing the head distribution, the storage coefficient and the hydraulic conductivity (de Marsily 1986). In the frequency domain, the governing equation is

$$\omega S(\mathbf{x})H - K(\mathbf{x})\nabla^2 H - \nabla K(\mathbf{x}) \cdot \nabla H = 0, \tag{3}$$

where $S(\mathbf{x})$ denotes the storage coefficient and $K(\mathbf{x})$ represents the hydraulic conductivity (Vasco and Karasaki 2001). All of the above equations, are composed of terms in which a field quantity, or its derivative, multiplies a material property coefficient or its own spatial derivative.

In the forward problem we fix the material properties [such as $\sigma(\mathbf{x})$, $S(\mathbf{x})$, and $K(\mathbf{x})$ in the equations given above] and solve for the field variables [$u(\mathbf{x}, \omega)$, $T(\mathbf{x})$, and $H(\mathbf{x}, \omega)$]. This renders many equations, such as (1) and (3), linear in the unknown field quantities, u and H respectively. However, in the inverse problem we seek both the material properties as well as the field variables within the Earth. In this setting, equations (1) and (3) are non-linear because they contain product terms in which field variables (or their derivatives) are multiplied by material property coefficients (or their derivatives) [for example, the product term $\nabla K(\mathbf{x}) \cdot \nabla H$ in equation (3)]. Typically the field variables are consider to be implicit functions of the material property distribution. In reality, the situation is more symmetric, because if the field is known throughout the Earth one may solve for the material property distribution. When differential equations such as these are discretized, the result is a system of sparse algebraic equations. The study of such equations is the relm of algebraic geometry, a very deep branch of mathematics (Kendig 1977, Eisenbud 1995).

As an illustration, consider the discretization of the Helmholtz equation (1). In two-dimensions we discretized equation (1) over a finite difference grid. The result is the matrix equation

$$\mathbf{Au} = \mathbf{b}. \tag{4}$$

where $\mathbf{A}$ is an $N \times N$ block tridiagonal matrix of the form

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{I} & \mathbf{A}_2 & \mathbf{I} & \cdots & \mathbf{0} \\ \cdot & \cdot & & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & & \cdot & \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I} & \mathbf{A}_N \end{pmatrix} \tag{5}$$

and the block matrices $\mathbf{A}_i, i = 1, ..., N$ is given by

$$\begin{pmatrix} -4 + h^2\sigma_{1i} & 1 & 0 \\ 1 & -4 + h^2\sigma_{2i} & 0 \\ \cdot & & \cdot \\ \cdot & \cdot & \\ \cdot & & 1 \\ 0 & \cdot & -4 + h^2\sigma_{Ni} \end{pmatrix} \tag{6}$$

and h is the distance between consecutive nodes (Everett 1996). The vector $\mathbf{b}$ contains the boundary values and data, that is the observations on the boundary nodes. Equation (4) represents a large, sparse system of algebraic equations in the unknown nodal field values, u_{ni}, and the nodal conductivities σ_{ni}.

As pointed out in Vasco (2000), discrete representations of functionals and integral equations, derived from the governing differential equations, are also algebraic in nature. Furthermore, squaring and differentiating algebraic equations results in algebraic equations. Thus, least squares formulations and most regularization penalty terms, which are quadratic forms, do not alter the algebraic nature of the equations. Thus, large classes of geophysical inverse problems may be represented by algebraic equations. As we shall see, the algebraic approach provides a unified framework for treating both linear and non-linear geophysical inverse problems. For linear problems the calculations reduce to computational linear algebra and matrix operations (Parker 1994). For non-linear algebraic equations there are methods from computational algebra which can be invoked (Cox et al. 1998, Sturmfels 2002).

The importance of being algebraic

What is the significance of the fact that many, if not most, geophysical inverse problems take the form of a system of algebraic equations? One consequence is that the solution set of a system of algebraic equations, known as an affine variety, can be quite different from the solution set of a linear inverse problem. For example, the solution set may intersect itself and the tangent space to the solution set may not be defined at points. Furthermore, the solution set may consist of pieces of differing dimension, such as a line intersecting a plane. Thus, the solution set does not necessarily define a differentiable manifold (Boothby 1975). Hence, concepts such as the dimension of an affine variety require careful consideration.

More importantly, polynomial equations possess a rich algebraic structure, similar in many respects to that associated with linear equations, which may be used to examine fundamental properties of the solution set. Questions related to the existence and uniqueness of a solution may be addressed using methods from computational algebra (Vasco 2000). Furthermore, techniques for finding all solutions to the inverse problem follow from the definite algebraic structure (Sturmfels 2002). In the Methodology section that follows I discuss the fundamental algebraic structure, the polynomial ideal, associated with a set of algebraic equations. I also explore the analogy between the treatment of linear equations and the treatment of algebraic equations.

METHODOLOGY

Many of the questions we wish to answer are geometric in nature. Does a solution to the inverse problem even exist? What is the dimensionality of the solution set? If the solution set is zero-dimensional, a collection of points, how many solutions are there? In order to answer these geometrical questions using a given set of algebraic equations one must relate the geometrical quantity, the affine variety, to an algebraic structure. The relevant questions are then framed in the language of algebra and computations are performed on the algebraic structure. This approach is identical to that adopted in solving linear inverse problems. In fact, linear equations are a special case of algebraic equations, in which the degree of each term in the equation is one. Thus, the concepts employed in solving linear problems have counterparts in the treatment of arbitrary algebraic equations. Thus, I will motivate the algebraic approach by first examining a linear system of equations.

Motivation: The linear problem

Consider the linear system of equations:

$$\begin{aligned} \mathbf{l}_1 &= \mathbf{a}_1 \cdot \mathbf{x} - d_1 = 0 \\ \mathbf{l}_2 &= \mathbf{a}_2 \cdot \mathbf{x} - d_2 = 0 \\ &\vdots \\ \mathbf{l}_N &= \mathbf{a}_N \cdot \mathbf{x} - d_N = 0 \end{aligned} \tag{7}$$

where $\mathbf{x}$ is a vector of unknowns, the model parameters, $\mathbf{a}_i$ are the data kernels or representers (Parker 1994, p. 62), and d_i are the observations. For the linear system (7) there is some freedom in the form of the defining equations. That is, we are not confined to a given set of equations. In particular, it is permissible to multiply any equation in (7) by an arbitrary scalar and add it to any other equation. This freedom is essential in transforming (7) into a useful form. For example, we may employ these operations to change the matrix of data kernels

$$\mathbf{M} = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_N \end{pmatrix} \tag{8}$$

into an upper triangular or diagonal form. From these canonical forms it is possible to compute properties such as the dimensionality of the solution set.

Let us summarize the most important intuitive ideas behind the computations for linear inverse problems. If a solution exists, the solution set forms a geometrical object (a point, line, plane, or hyperplane) defined by the vanishing of the linear forms $\mathbf{l}_1, \mathbf{l}_2, ...\mathbf{l}_N$. Secondly, there is a one-to-one correspondence between the solution set and the algebraic object

$$\langle \mathbf{l}_1, ..., \mathbf{l}_N \rangle \equiv \left\{ \sum^N \beta_j \mathbf{l}_j \right\} \tag{9}$$

where β_j are arbitrary scalar multipliers. This algebraic object consists of all possible linear combinations of the constraints $\mathbf{l}_j$. Stated another way, it is not the particular set of equations that is important, we are free to change the specific set of defining equations. In particular, we can find a simple representation from which we can derive information about the solution set. These ideas form the basis for Gaussian elimination, LU decomposition, and the singular value decomposition (SVD) (Golub and Van Loan 1989).

Algebraic structure: Polynomial ideals

Now consider the situation when the equations are not linear. Rather, we have a system of algebraic equations, polynomial equations in n variables. A polynomial in n variables, $\mathbf{x} = (x_1, ..., x_n)$, is a sum (Cox et al. 1997, p. 2)

$$\mathbf{p}_i = \sum_\alpha a_\alpha{}^i \mathbf{x}^\alpha - d_i \tag{10}$$

where the sum is over a set of integer index vectors $\alpha = (\alpha_1, ..., \alpha_n)$, the coefficients $a_\alpha{}^i$ are real or complex numbers, and

$$\mathbf{x}^\alpha = x_1{}^{\alpha_1} \cdot x_2{}^{\alpha_2} \cdots x_n{}^{\alpha_n}. \tag{11}$$

The set of all polynomials in $x_1, ..., x_n$ with coefficients in the field of complex numbers $\mathbf{C}$ is denoted $\mathbf{C}[x_1, ..., x_n]$. An algebraic equation results when we set the polynomial $\mathbf{p}_i$ equal to zero.

In dealing with algebraic equations we can follow a procedure much like that used in treating a linear system of equations. In order to do this it is necessary to define an appropriate mathematical structure, that is, an algebraic quantity that can be directly related to the zero set of a system of algebraic equations [affine variety]. The particular algebraic structure is similar in form to the linear sum (9).

Definition For a set of polynomials $\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_N$ the *ideal* is defined as

$$\langle \mathbf{p}_1, ..., \mathbf{p}_N \rangle \equiv \left\{ \sum^N \beta_i \mathbf{p}_i \right\} \tag{12}$$

where β_i are arbitrary polynomials.

Note that the multipliers β_i are arbitrary polynomials rather than simple scalars. Thus, an ideal consists of all polynomials generated by the basis set $\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_N$. All the elements of the ideal vanish on the zero set of the defining equations.

The importance of the ideal associated with a set of algebraic equations derives from its unique relationship with the zero set of the equations. That is, it we account for multiplicities, there is a one-to-one relationship between the ideal of a set of algebraic equations, $\langle \mathbf{p}_1, ..., \mathbf{p}_N \rangle$, and the zero set of the polynomials (Cox et al. 1997). The multiplicities arise because the zero set for the power of a polynomial is identical to the zero set of the polynomial itself.

The value of Algebraic Structure:

At this juncture I can now outline, in an intuitive fashion, the primary idea underlying an algebraic approach to the study of inverse problems. The solution set forms a geometrical object [Affine Variety] defined by the vanishing of a set of polynomials

$$\begin{aligned}
\mathbf{p}_1 &= \sum_\alpha a_\alpha{}^1 \mathbf{x}^\alpha - d_1 \\
\mathbf{p}_2 &= \sum_\alpha a_\alpha{}^2 \mathbf{x}^\alpha - d_2 \\
&\vdots \\
\mathbf{p}_N &= \sum_\alpha a_\alpha{}^N \mathbf{x}^\alpha - d_N
\end{aligned} \tag{13}$$

To this solution set there corresponds an algebraic object, the associated ideal,

$$\langle \mathbf{p}_1, ..., \mathbf{p}_N \rangle \equiv \left\{ \sum^N \beta_i \mathbf{p}_i \right\}. \tag{14}$$

As in the case of the linear system of equations, we are not constrained to one specific form of defining equations. Rather, we are free to modify the defining equations as long as the ideal is preserved. For example, as pointed out by Morgan (1987), one is free to multiply an equation by a scalear and add it to another equation. Using this generalized form of row reduction, one can examine properties such as the number of solutions at infinity.

In a relatively recent development Buchberger (1985) produced a basis set, the Gröbner basis, with many useful properties. The primary computational advantage of a Gröbner basis is that it allows one to extend polynomial division to division by a set of polynomials. In particular, when the polynomials are elements of a Gröbner basis, the remainder upon division is unique (Cox et al. 1997, p. 79). Though this may seem like a modest accomplishment, the consequences of this are far reaching.

Gröbner basis calculations now form the core of many algorithms in computational algebra (Adams and Loustaunau 1994, Eisenbud 1995, p. 321, Cox et al. 1997, Cox et al. 1998). In particular, they may be used to examine the existence of a solution to a set of algebraic equations. If the unit element, 1, is an element of the Gröbner basis than the equations are inconsistent and no solution exists (Cox et al. 1997). Furthermore, Gröbner basis computations allow one to perform something akin to the triangularization of a set of algebraic equations (Cox et al. 1997). That is, using this basis set I may derive a 'triangular' set of polynomial equations in which the final equations are in the fewest number of variables (Adams and Loustaunau 1994). In the optimal case the final equation is a polynomial equation in a single variable, say x_n. This equation is solved for all values of x_n and these values are successively back-substituted into the previous equation, which is solved for x_{n-1}, and so on. Thus, one may extend the approach used for linear systems of equations to simplify a system of algebraic equations. Gröbner basis calculations may also be used to compute the dimension of the zero set of a system of algebraic equations. In the event that the solution set has a dimension of zero, is a collection of isolated points, one can compute the number of solutions to the equations. Finally, there are Gröbner basis algorithms for finding all of the solutions for a system of algebraic equations (Cox et al. 1997, Cox et al. 1998). Such algorithms complement existing methods based upon polynomial continuation (Morgan 1987, Watson and Morgan 1992, Verschelde et al. 1994) and resultants (Manocha 1998, Emiris and Canny 1995).

CONCLUSIONS

When discretized, many non-linear geophysical inverse problems take the form of algebraic (polynomial) equations. A mathematical framework, based upon polynomial ideals, is available for treating such equations. The approach extends ideas used in solving linear inverse problems, such as Gaussian elimination, to nonlinear inverse problems. Thus, the framework unifies both linear and non-linear inverse problems. As such, it provides a useful vantage point from which to view many geophysical inverse problems. From this vantage point we gain insight into some of the issues pertaining to non-linear inverse problems. That is, we can examine questions of existence, uniqueness, and the computability of solutions to the inverse problems. Furthermore, there are a wealth of methods from computational algebra which may be of value in treating non-linear inverse problems.

One recent development in computational algebra, Gröbner basis computations, has been the subject of considerable interest (Adams and Loustaunau 1994, Cox et al. 1997, 1998). At present Gröbner basis calculations are computationally intensive and are practical for rather small systems of equations. Future needs in the area of Gröbner basis algorithms involve the practical implementation and application of these techniques. The software implementation of Gröbner basis algorithms has been primarily linked with mathematical research. Issues such as computational efficiency and stability may garner greater attention in the future.

Computational algebra is a relatively young field and there are many open questions and avenues to explore. Already, there are a wealth of methods for treating systems of algebraic equations (Cox et al. 1998, Sturmfels 2002). These include earlier homotopy algorithms (Morgan 1987) and more recent sparse polynomial variants (Verschelde et al. 1994). In addition, there are algorithms based upon resultants (Manocha 1998) and sparse resultants (Emiris and Canny 1995). The quesion of real solutions to algebraic equations is a topic of interest in solving inverse problems. This is a relatively unexplored area but there are algorithms for counting the number of real solutions (Pederson et al. 1993). The field of semidefinite programming (Stengle 1974, Vandenberghe and Boyd 1996) provides a means to verify the existence of a real solution to a system of algebraic equations (Sturmfels 2002). There is the question of extending these ideas to an infinite dimensional setting which has not yet been explored. Such an extension might involve exploring the properties of multilinear mappings in Banach spaces (Abraham et al. 1983).

REFERENCES

Abraham, R., J. E. Marsden, and T. Ratiu, 1983. *Manifolds, Tensor Analysis, and Applications*, Addison-Wesley Publishing Co., Reading.

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

Aki, K., and P. Richards, 1980. *Quantitative Seismology*, W. H. Freeman, San Francisco.

Boothby, W. M., 1975. *An Introduction to Differentiable Manifolds and Riemannian Geometry*, Academic Press, New York.

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, 1998. *Using Algebraic Geometry*, Springer, New York.

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

de Marsily, G., 1986. *Quantitative Hydrogeology*, Academic Press, San Diego, California.

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.

Emiris, I. Z., and J. F. Canny, 1995. Efficient incremental algorithms for the sparse resultant and the mixed volume, *J. Symbolic Comp.*, **20**, 117-149.

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

Golub, G. H., and C. F. Van Loan, 1989. *Matrix Computations*, The Johns Hopkins University Press, Baltimore.

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

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

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

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

Manocha, D., 1998. Numerical methods for solving polynomial equations, in *Applications of Computational Algebraic Geometry*, D. A. Cox and B. Sturmfels (eds.), American Math. Soc., Providence.

Morgan, A. P., 1987. *Solving Polynomial Systems Using Continuation for Engineering and Scientific Problems*, Prentice-Hall, Englewood Cliffs, New Jersy.

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

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.

Stengle, G., 1974. A nullstellensatz and a positivstellensatz in semialgebraic geometry, *Math. Ann.*, **207**, 87-97.

Sturmfels, B., 2002. *Solving Systems of Polynomial Equations*, Amer. Math. Soc., Providence, RI.

Tarantola, A., 1987. *Inverse Problem Theory*, Elsevier, New York.

Vandenberghe, L., and S. Boyd, 1996. Semidefinite programming, *SIAM Review*, **38**, 49-95.

Vasco, D. W., 2000. An algebraic formulation of geophysical inverse problems, *Geophy. J. Int.*, **142**, 970-990.

Vasco, D. W., 1999. Intersections, ideals, and inversion, *Inverse Problems*, **15**, 1573-1602.

Vasco, D. W., and K., Karasaki, 2001. Inversion of pressure observations: an integral formulation *J. Hydrol.*, **253**, 27-40.

Verschelde, J., P. Verlinden, and R. Cools, 1994. Homotopies exploiting Newton polytopes for solving sparse polynomial systems, *SIAM J. Numer. Anal.*, **31**, 915-930.

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