SEISMIC CROSSHOLE TOMOGRAPHY AND NONLINEAR CONSTRAINED OPTIMIZATION

James G. Berryman[*]

Abstract. When an inverse problem can be formulated so the data are minima of one of the variational problems of mathematical physics, feasibility constraints can be found for the nonlinear inversion problem. These constraints guarantee that optimal solutions of the inverse problem lie in the convex feasible region of the model space. Furthermore, points on the boundary of this convex region can be found in a constructive fashion. Finally, for any convex function over the model space, it is shown that a local minimum of the function is also a global minimum. The proofs in the paper are formulated for definiteness in terms of first arrival traveltime inversion, but apply to a wide class of inverse problems including electrical impedance tomography.

1. Introduction

In medical x-ray tomography [Cormack 1963; Hounsfield 1973], measurements of x-ray attenuation for straight rays that pass through the body provide a set of line integrals for analysis. Then, the classical methods of tomography [Gordon, Bender, and Herman 1970; Gilbert 1972; Herman 1980; Natterer 1986] may be applied to reconstruct a two-dimensional function from the set of line integrals. Backprojection along the known straight ray paths in these problems is a typical method used for such reconstructions. The attenuation data are then used to generate a picture of a slice through the inhomogeneous density distribution of the body. This picture may subsequently be used for purposes of diagnosis.

---

*University of California, Lawrence Livermore National Laboratory, P. O. Box 808 L-156, Livermore, CA 94550

Tomography also has many uses outside the field of medicine. Some of these uses include electron microscopy [Gordon, Bender, and Herman 1970; Gilbert 1972; Herman 1980; Natterer 1986], acoustical and optical tomography [Schomberg 1978; Kak 1984; Natterer 1986], radio astronomy [Bracewell and Riddle 1967], and seismic tomography [Nolet 1987]. In geophysical applications to the whole earth [Aki, Christoffersson, and Husebye 1976; Aki and Richards 1980] or to smaller scale reconstruction problems such as borehole-to-borehole scanning with either seismic or electromagnetic probes [Bois, LaPorte, Lavergne, and Thomas 1972; Lager and Lytle 1977], the assumption of straight ray paths, although very common, is often a poor approximation [Dines and Lytle 1979; Lytle and Dines 1980; Justice, Vassiliou, Singh, Logel, Hansen, Hall, Hutt, and Solanki 1989a,b]. Using both wave amplitude and phase in the reconstruction can improve the results dramatically [Wolf 1969; Mueller, Kaveh, and Wade 1979; Devaney 1984] if a starting model of sufficient accuracy is known. Thus, progress on the reconstruction problem with ray bending might be made if some method for finding good starting models could be found. A new iterative algorithm for computing such a starting model from traveltime data is therefore highly desirable.

In a series of papers [Berryman 1989a; Berryman 1989b; Berryman 1990], the author has developed a stable iterative reconstruction method for first arrival traveltime inversion. The general theory behind this new approach and its extensions will be described in the present paper. The main idea behind the new approach may be summarized as follows: When an inverse problem can be formulated so the data are minima of one of the variational problems of mathematical physics, feasibility constraints can be found for the nonlinear inversion problem. These constraints guarantee that optimal solutions of the inverse problem lie in a convex feasible region of the model space. Furthermore, points on the boundary of this convex region can be found in a constructive fashion. For any convex function over the model space, a local minimum of the function is also a global minimum. In light of the structure induced on the model space by the feasibility constraints, we can also obtain a series of results about the structure of the solution set that would not be possible to establish otherwise. Although these ideas were first presented for traveltime inversion problems, the methods apply to a wide class of inverse problems including electrical impedance tomography.

The four main goals for the paper are these: (a) to establish before a mathematical audience that the idea of using variational/feasibility constraints for inversion is both rigorous and applicable to a wide class of physical problems, (b) to provide elementary proofs that will be accessible to a broad audience (including physicists, geophysicists, engineers, etc., as well as mathematicians) of the consequences of this idea, (c) to present the proofs in an abstract setting so as to be independent of the particular choice of discretization made in practical algorithms for solving the inverse problem, and (d) to summarize one new reconstruction algorithm based on these ideas.

For definiteness, we use first arrival traveltime inversion as our primary example. However, it will become clear that the methods developed here apply to a wide class of inverse problems. Whenever the data can be chosen to be the minima of one of the variational problems of mathematical physics, variational constraints can be introduced and the concepts of feasible set and feasibility boundary follow

Figure 1. Diagram to illustrate the computation of the traveltime when the model is composed of cells of constant slowness.

immediately [Berryman 1989b; Berryman 1990].

2. Convexity Properties of Seismic Inversion

Our principal example will be first arrival traveltime inversion. The problem is this: Given the locations of sources and receivers of some type of exciting wave (*e.g.*, acoustic, seismic, or electromagnetic) and the first arrival traveltimes T_i for waves propagated between the m pairs of sources and receivers (labelled by $i = 1, \ldots, m$), deduce the wave speed $v(\vec{x})$ in the region probed by these waves. For iterative methods involving the use of least-squares methods, it is common to solve for the reciprocal of the wave speed $s(\vec{x}) = 1/v(\vec{x})$ — which is the wave slowness. For practical applications, a rectangular grid is generally chosen and the wave slowness is discretized either by treating it as constant in the cells determined by the grid or by specifying values of slowness at the nodal points and choosing some interpolation scheme (such as bilinear) between the nodes. For either choice of discretization, the model slowness is determined by a vector $s^T = (s_1, \ldots, s_n)$, where n is respectively either the number of cells or the number of nodes. Thus, in terms of the practical applications, we can speak of a particular model slowness s as a *point*, *i.e.*, a point in the model vector space.

Fermat's principle [Born and Wolf 1959] says that the first arrival traveltime for the i-th ray path is given by

$$t_i(s) = \min_{\{paths\}} \int s \, dl_i^{(path)} \equiv \int s \, dl_i^*[s] \tag{1}$$

where $l_i^{(path)}$ is the arc length along any connected path between the source and

receiver and where $l_i^*[s]$ is the arc length along a ray path that minimizes the integral of the traveltime for the i-th path and wave slowness s. Figure 1 illustrates the computation associated with (1.1) for a particular choice of model discretization, i.e., for cells of constant slowness. If more than one path minimizes the traveltime, then $l_i^*[s]$ is any particular choice among those minimizing the traveltime. Since the slowness is a positive quantity, the traveltime can never vanish unless the source and receiver are located at the same point. We exclude this case, so the traveltime is also a strictly positive quantity.

Some easy but important facts follow from the variational definition (1) of the first arrival traveltime. They are given in these lemmas:

Lemma 1. *(Concavity and Homogeneity)* The traveltime $t_i(s)$ is a concave and homogeneous function of the model slowness s.

Proof: First, note that, for $s_1 > 0$, $s_2 > 0$, and $0 \le \lambda \le 1$,

$$t_i(s_1) = \int s_1 \, dl_i^*[s_1] \le \int s_1 \, dl_i^*[\lambda s_1 + (1-\lambda)s_2] \tag{2}$$

and

$$t_i(s_2) = \int s_2 \, dl_i^*[s_2] \le \int s_2 \, dl_i^*[\lambda s_1 + (1-\lambda)s_2] \tag{3}$$

both follow immediately from the definitions in (1). Then, taking the appropriate linear combination of the inequalities in (2) and (1.3) with $0 \le \lambda \le 1$, we have

$$\lambda t_i(s_1) + (1-\lambda)t_i(s_2) \le \int [\lambda s_1 + (1-\lambda)s_2] \, dl_i^*[\lambda s_1 + (1-\lambda)s_2]$$
$$\equiv t_i(\lambda s_1 + (1-\lambda)s_2) \tag{4}$$

thus completing the proof that $t_i(s)$ is a concave function. That the traveltime is homogeneous in slowness so $t_i(\gamma s) = \gamma t_i(s)$ follows easily from the statement (1) of Fermat's principle.

Lemma 2. *(Scale Invariance of Ray Paths)* A ray path with arc length $l_i^*[s]$ that minimizes the traveltime for s also minimizes the traveltime for γs where γ is any positive scalar.

Proof: Lemma 2 follows immediately from the homogeneity property of $t_i(s)$.

Next we need to introduce the notion of a *feasible set* of model slownesses and the associated *feasibility boundary*. The introduction of these physically-based feasibility conditions is the principle new contribution from which the rest of the present results follow.

The concept of feasibility sets arises commonly in the study of nonlinear programming techniques [Fiacco and McCormick 1968]. Algorithms for practical solution of the inverse problems discussed here fall into this class — although inverse problems tend to be substantially more difficult than the optimization problems typically considered (since the constraints are implicit rather than explicit). Feasibility constraints for inverse problems have been introduced elsewhere [Berryman 1989b; Berryman 1990]. We will present only a brief outline of the motivation here.

The inverse problem for first arrival traveltime is to determine a slowness model s given a set of measured traveltimes T_i between pairs (labelled by index i) of

sources and receivers whose locations are known. One method for solving the inverse problem is to guess a model s_g that might have given rise to the measured data, compute the set of traveltimes $t_i(s_g)$ for the trial model, and then use some method (often based on least-squares fitting) to update the model and obtain a better fit to the data. However, such programming methods are generally limited by the fact that it may be computationally difficult to find the exact $l^*[s_g]$ associated with the trial slowness. For this reason, we define a trial traveltime

$$\tau_i^{(p_i)}(s) = \int s \, dl_i^{(p_i)} \tag{5}$$

for arc length $l_i^{(p_i)}$ associated with the trial ray path p_i. Then, neglecting experimental error in the T_is and defining s_0 as an exact solution of the inverse problem, then s_0 clearly satisfies

$$T_i = t_i(s_0) \leq \tau_i^{(p_i)}(s_0) \tag{6}$$

for all source-receiver pairs i and any trial ray path p_i between them. Thus, in trying to formulate a constructive method for locating an s_0, it is useful to consider splitting the model space into two parts: (a) a feasible part whose members s are like s_0 in that they satisfy the constraints

$$T_i \leq \tau_i^{(p_i)}(s) \tag{7}$$

for all source-receiver pairs i and all ray paths in the trial set and (b) a nonfeasible part whose members s violate at least one of the inequalities (7) for some ray and, hence, are unlike s_0.

With this motivation, we can now distinguish between local (path dependent) feasibility sets $F^{\{p\}}$ and the global (path independent) feasibility set F as in other recent publications [Berryman 1989b; Berryman 1990]. We will also be able to establish a definite relationship between these two types of feasibility set.

Definition: The (local) *feasible set* $F^{\{p\}}$ of slownesses for the nonlinear traveltime inversion problem is given by $F^{\{p\}} = \{s \,|\, \tau_i^{(p_i)}(s) \geq T_i, i = 1, \ldots, m\}$ where the T_is are the measured traveltime data and $\{p\}$ is a particular set of trial ray paths.

Definition: The (global) *feasible set* F of slownesses for the nonlinear traveltime inversion problem is given by $F = \{s \,|\, t_i(s) \geq T_i, i = 1, \ldots, m\}$ where the T_is are the measured traveltime data and all the T_is are finite.

Definition: Let $\Sigma = \{s \,|\, s > 0\}$ be the physical set of slownesses. Then, we define the *absolute* or *physical* feasible sets as $A = F \cap \Sigma$ and $A^{\{p\}} = F^{\{p\}} \cap \Sigma$. Note that Σ is a convex set.

Theorem 1. The (global) feasible set F is a nonempty convex set. All points on the boundary of F are determined by finding, for each model $s > 0$, the smallest value of the scalar γ such that $t_i(\gamma s) \geq T_i$ for all $i = 1, \ldots, m$.

Theorem 2. The (local) feasible set $F^{\{p\}}$ is a nonempty convex set. All points on the boundary of $F^{\{p\}}$ are determined by finding, for each model s, the smallest value of the scalar γ such that $\tau_i^{(p_i)}(\gamma s) \geq T_i$ for all $i = 1, \ldots, m$.

Proof: If s_1 and s_2 satisfy the feasibility constraints $t_i(s) \geq T_i$ for all i and if $0 \leq \lambda \leq 1$, then

$$t_i(\lambda s_1 + (1 - \lambda)s_2) \geq \lambda t_i(s_1) + (1 - \lambda)t_i(s_2) \geq T_i \tag{8}$$

follows from the concave property of the traveltime function. Thus, $\lambda s_1 + (1 - \lambda)s_2$ also satisfies the feasibility constraints and the feasible set F is therefore convex.

The set F is nonempty because for any slowness s we can always find a finite value of the positive scalar γ such that $t_i(\gamma s) = \gamma t_i(s) \geq T_i$ for all $i = 1, \ldots, m$. This result follows from the positivity of t_i and the finiteness and positivity of the measured traveltimes. Furthermore, the smallest such scalar is given by

$$\gamma_{min}(s) = \max_i T_i/t_i(s), \tag{9}$$

so we can locate the feasibility boundary in the direction of s which is then given by the point $\gamma_{min}(s)s$. This completes the proof of Theorem

The proof of Theorem 2 is completely analogous — just replace $t_i(s)$ everywhere by $\tau_i^{(p_i)}(s)$ and note that

$$\tau_i^{(p_i)}(\lambda s_1 + (1 - \lambda)s_2) = \lambda \tau_i^{(p_i)}(s_1) + (1 - \lambda)\tau_i^{(p_i)}(s_2), \tag{10}$$

i.e., that $\tau_i^{(p_i)}$ is a linear function of s.

Corollary 1. The physical feasible sets A and $A^{\{p\}}$ are nonempty convex sets.

Proof: The sets A and $A^{\{p\}}$ are both the intersections of convex sets and therefore convex. That the intersection of these convex sets is nonempty is clear from the proofs of Theorems 1 and 2.

Remark: We would like to avoid cluttering the remainder of the paper with constant reminders that the physical space is limited to the set Σ. All our results on convexity should be followed by a step finding the intersection of the set under study and Σ. We assume that, from now on, the reader will supply this step in each case.

Corollary 2. The global feasible set F is the intersection of the local feasible sets $F^{\{p\}}$ for all possible sets of ray paths.

Proof: First, the intersection of all local feasible sets is a convex set. Let I be the intersection. Then it is convex if $\lambda s_1 + (1 - \lambda)s_2$ lies in I for every s_1 and s_2 in I and $0 \leq \lambda \leq 1$. But if s_1 and s_2 are in I, then they lie in each individual convex region and, hence, so does the convex combination.

Second, since it follows from Fermat's principle for first arrivals that

$$t_i(s) \leq \tau_i^{(p_i)}(s), \tag{11}$$

we want to consider the values of the scalar γ such that

$$T_i \leq t_i(\gamma s) \leq \tau_i^{(p_i)}(\gamma s). \tag{12}$$

Let $\gamma_{min}^{\{p\}}$ be the minimum such γ for $\tau_i^{(p_i)}$ defined by

$$\gamma_{min}^{\{p\}}(s) = \max_i T_i/\tau_i^{(p_i)}(s). \tag{13}$$

Then, it follows immediately from (12) that

$$\gamma_{min}^{\{p\}}(s) \leq \gamma_{min}(s) \tag{14}$$

Figure 2. Illustration of the distinction between local and global feasibility boundaries. Local feasibility is determined by the traveltime data and one set of ray paths. Global feasibility is determined by the traveltime data and by all possible sets of ray paths (Corollary 2) or by the optimum ray paths for a given direction in the model space (Theorem 1).

for every set of ray paths and every s. Thus, the global feasibility boundary is bounded below by the local feasibility boundaries for all sets of ray paths $\{p\}$. Furthermore, equality in (14) is achieved when the set of ray paths $\{p\}$ is one that minimizes the traveltime. So the intersection of all the local feasibility sets is convex and has the same extreme points as F. It follows that the intersection set I and the global feasibility set F must be the same. This completes the proof of the Corollary. Figure 2 illustrates the distinction between the local and global feasibility boundaries.

Next we introduce the concept of feasibility violation number and determine the convexity properties (or lack thereof) to be asssociated with this function. The feasibility violation number is another useful concept that arises in the practical applications to be described later. The feasibility constraints bound the region where the feasiblity violation number is zero everywhere. By counting the number of violations present as we move away from the feasibility boundary into the infeasible region, we obtain a useful measure of how far a model point is from the feasibility boundary without doing a prohibitively expensive computation [Berryman 1989b;

Berryman 1990].

Definition: The feasibility violation number at s for a given set of ray paths $\{p\}$ is defined to be

$$N^{\{p\}}(s) \equiv \sum_i \theta(T_i - \tau_i^{(p_i)}(s)) \tag{15}$$

where the step function $\theta(x)$ is defined by

$$\theta(x) = \begin{cases} 0, & \text{for } x \le 0; \\ 1, & \text{for } x > 0. \end{cases} \tag{16}$$

Definition: The set $V^{\{p\}}(n)$ is given by $V^{\{p\}}(n) = \{s \mid n \ge N^{\{p\}}(s)\}$, where n is a non-negative integer and $\{p\}$ is some set of ray paths. The set so defined is the set of all slowness models s that violate n or fewer than n of the feasibility constraints.

To see that the set $V^{\{p\}}(n)$ is generally a nonconvex set, suppose s_1 violates constraint i but not j, while s_2 violates j but not i. Then, for some choices of s_1, s_2, and λ, it is possible for the convex combination to violate both constraints i and j. Unless this point has also simultaneously ceased to violate some other constraint (which cannot happen for example if $n = 1$), the point lies outside the set and $V^{\{p\}}(n)$ is therefore nonconvex.

On the other hand, if s_1 and s_2 satisfy the feasibility constraint $\tau_i^{(p_i)}(s) \ge T_i$ for some i, then $\lambda s_1 + (1 - \lambda)s_2$ also satisfies it from the property (10) of the traveltime trial function. Thus, $V^{\{p\}}(n)$ is nonempty and includes both the global and local feasible sets which are also nonempty

$$F \subset F^{\{p\}} \subseteq V^{\{p\}}(n). \tag{17}$$

Furthermore, if $m \ge n_2 \ge n_1 + 1$, then it is clear that

$$V^{\{p\}}(n_1) \subset V^{\{p\}}(n_2). \tag{18}$$

3. Convex Programming for Seismic Tomography

We will first define convex programming for first arrival traveltime inversion. Then we present some basic Theorems about convex programming in this context. Finally, we give a discussion of implications for practical implementations.

Definition: Let $\varphi(s)$ be any convex function of s. Then the *convex nonlinear programming problem* associated with φ is to minimize $\varphi(s)$ subject to the global feasibility constraints $t_i(s) \ge T_i$ for $i = 1, \ldots, m$.

Definition: Let $\psi_0^{\{p\}}(s) = \sum_i w_i (\tau_i^{(p_i)}(s) - T_i)^2$ for some set of ray paths $\{p\}$ where the w_is are some positive weights. Then the *convex linear programming problem* associated with $\psi_0^{\{p\}}$ is to minimize $\psi_0^{\{p\}}(s)$ subject to the local feasibility constraints $\tau_i^{(p_i)}(s) \ge T_i$ for $i = 1, \ldots, m$.

Theorem 3. Every local minimum s^* of the convex nonlinear programming problem associated with $\varphi(s)$ is a global minimum.

Theorem 4. Every local minimum s^* of the convex linear programming problem associated with $\psi_0^{\{p\}}(s)$ is a global minimum.

Proof: (Also see [Fiacco and McCormick 1968].) Let s^* be a local minimum. Then, by definition, there is a compact set C such that s^* is in the interior of $C \cap F$ and

$$\varphi(s^*) = \min_{C \cap F} \varphi(s). \tag{19}$$

If s is any point in the feasible set F and $0 \leq \lambda \leq 1$ such that $\lambda s^* + (1 - \lambda)s$ is in $C \cap F$, then

$$\varphi(s) \geq \frac{\varphi(\lambda s^* + (1 - \lambda)s) - \lambda \varphi(s^*)}{1 - \lambda} \geq \frac{\varphi(s^*) - \lambda \varphi(s^*)}{1 - \lambda} = \varphi(s^*). \tag{20}$$

The first step of (20) follows from the convexity of φ and the second from the fact that s^* is a minimum in $C \cap F$. Convexity of F guarantees that $\lambda s^* + (1 - \lambda)s$ is in the feasible set. This completes the proof of Theorem 3.

To prove Theorem 4, we first need to establish that $\psi_0^{\{p\}}(s)$ is a convex function of s. This fact follows from identity (10), the convexity of the function $f(x) = x^2$, and the convexity of a sum of convex functions. We see that

$$\begin{aligned}
[\tau_i(\lambda s_1 + (1 - \lambda)s_2) - T_i]^2 &= [\lambda \tau_i(s_1) + (1 - \lambda)\tau_i(s_2) - T_i]^2 \\
&= \lambda[\tau_i(s_1) - T_i]^2 + (1 - \lambda)[\tau_i(s_2) - T_i]^2 \\
&\quad - \lambda(1 - \lambda)(\tau_i(s_1) - \tau_i(s_2))^2
\end{aligned} \tag{21}$$

so $\psi_0^{\{p\}}(s)$ is a convex function. The remainder of the proof of Theorem 4 follows that of Theorem 3 with $\psi_0^{\{p\}}$ replacing φ.

Finding useful convex functions for nonlinear programming is a challenging and unsolved problem. For example, if $\varphi(s)$ is chosen to be the convex function $-\sum_i t_i(s)$ (because $t_i(s)$ is concave), or $\sum_i 1/(t_i(s) - T_i)$ (because the harmonic mean is bounded above by the arithmetic mean), or $-\sum_i log(t_i(s) - T_i)$ (because the geometric mean is bounded above by the arithmetic mean), each choice is a convex function of s, but the minima occur for large (actually infinite) values of s and have nothing to do with the inversion problem. However, many useful convex functions (including $\psi_0^{\{p\}}(s)$) are available for the linear programming problem, and any convex function available for that problem may also be used in the nonlinear programming problem. We will now consider some convex and some nonconvex functions that are important in programming for inversion.

One important nonconvex function of s is the weighted squared error in the traveltime given by

$$\psi_0(s) = \sum_i w_i[t_i(s) - T_i]^2, \tag{22}$$

where the w_is are some positive weights [Berryman 1989a]. Techniques that seek to minimize the squared error are probably the most abundant in the literature, and therefore any new insight into the behavior of this function relative to the feasibility boundary is of great practical significance. We can determine the smallest value of $\psi_0(\gamma s)$ for all models with the same relative distribution of slowness but differing scales by varying the scalar γ to find the minimum. Elementary analysis shows that

$$\gamma_{ls}(s) = \frac{\sum_i w_i t_i(s) T_i}{\sum_i w_i t_i^2(s)}. \tag{23}$$

Note that (23) was obtained using only the homogeneity property of $t_i(s)$. Substituting (23) for γ in $\psi_0(\gamma s)$, we obtain

$$\psi_0(\sigma(s)) \equiv \psi_0(\gamma_{ls}(s)s) = \sum_i w_i T_i^2 - \frac{[\sum_i w_i t_i(s)T_i]^2}{\sum_i w_i t_i^2(s)}. \tag{24}$$

Note that $\psi_0(\sigma(s))$ is a function of s independent of scale. Thus, (24) provides a scalar that is characteristic of slowness distributions of the form γs.

Lemma 3. *(Infeasibility of Scaled Least-Squares Points)* The scaled least-squares point $\sigma(s) = \gamma_{ls}(s)s$ either solves the inversion problem or lies outside the feasible set F.

Proof: It follows from (23) that

$$0 = \sum_i w_i t_i [\gamma_{ls}(s)t_i(s) - T_i] = \sum_i w_i t_i(s)[t_i(\gamma_{ls}(s)s) - T_i]. \tag{25}$$

Since the t_is and w_is are all positive, it follows either that

$$t_i(\sigma(s)) - T_i \equiv 0 \text{ for all } i \tag{26}$$

or that

$$t_i(\sigma(s)) - T_i < 0 \text{ for some } i. \tag{27}$$

If (27) is true, then one of the feasibility constraints is violated at the point $\sigma(s)$ so this point lies outside F. If (26) is true, then $\sigma(s)$ solves the inversion problem.

Homogeneity is the only property of $t_i(s)$ used in (22)-(1.27). Therefore, we can immediately prove the corresponding statement for the path dependent problem. In the equations (22)-(1.27), let $t_i(s) \to \tau_i^{(p_i)}(s)$, $\gamma_{ls}(s) \to \gamma_{ls}^{\{p\}}(s)$, $\psi_0(s) \to \psi_0^{\{p\}}(s)$, and $\sigma(s) \to \sigma^{\{p\}}(s)$. Then it is clear that the following statement is true.

Lemma 4. *(Infeasibility of Scaled Least-Squares Points)* The scaled least-squares point $\sigma^{\{p\}}(s) = \gamma_{sl}^{\{p\}}(s)s$ for a particular set of ray paths $\{p\}$ either solves the inversion problem or lies outside the feasible set $F^{\{p\}}$.

Then, the next important Theorem is an easy consequence of Lemmas 3 and 4.

Theorem 5. *(Infeasibility of Least-Squares Points)* If s^* is a global minimum of either $\psi_0(s)$ or $\psi_0^{\{p\}}(s)$ for some choice of ray paths $\{p\}$, then s^* either solves the inversion problem or lies outside the feasible set F.

Proof: If s^* is a global least-squares point, then it is also a scaled least-squares point for distributions of the form γs^*. Thus, Lemmas 3 and 4 apply to s^*. Lemma 3 implies that the Theorem is true for $\psi_0(s)$. When Lemma 4 is pertinent, the point s^* lies outside of $F^{\{p\}}$ and, therefore, outside of F. Thus, the Theorem is true.

In nonlinear programming [Fiacco and McCormick 1968], it is common to distinguish between "interior methods" and "exterior methods." As the names imply, interior methods require the trial solutions at each stage to lie inside the feasible set while exterior methods try to approach the optimal solution from outside the feasible set. The significance of Theorem 5 is that the many methods of inversion based on least-squares minimization are all "exterior methods." This fact has apparently escaped notice until recently. The minimum of the least-squares error in

the predicted data will virtually always (unless a solution has been attained) lie outside the feasible set. This fact is important for programming purposes since it suggests steps that can be taken to improve the current trial model slowness. In particular, to gauge the merits of models relative to a given set of trial ray paths, the feasibility violation number (15) may be used effectively to provide a crude but also computationally inexpensive measure of the distance from the trial model to the feasibility boundary. This technique has proven very effective in applications to traveltime inversion [Berryman 1989b; Berryman 1990].

Note that $\psi_0(s)$ is nonconvex, whereas $\psi_0^{\{p\}}(s)$ is convex (see the proof of Theorem 4).

Another important convex function is

$$\psi_\mu^{\{p\}}(s) = \psi_0^{\{p\}}(s) + \mu||s - s_0||^2, \tag{28}$$

where μ is a non-negative scalar and s_0 is any point in the slowness model space. Clearly $\psi_\mu^{\{p\}}(s) = \psi_0^{\{p\}}(s)$ when $\mu = 0$. One possible choice of norm is

$$||s - s_0||^2 = \frac{1}{\Omega} \int_\Omega (s - s_0)^2 c \, d^3x, \tag{29}$$

where Ω is the volume of the region probed by the experiment and $c(\vec{x}) > 0$ is a positive weight that we will call the coverage function [Berryman 1989a]. The function $\psi_\mu^{\{p\}}(s)$ is convex because it is the sum of two convex functions of s.

Lemma 5. Let $R_\mu^{\{p\}}(\bar{s}) \equiv \{s \,|\, \psi_\mu^{\{p\}}(s) \le \psi_\mu^{\{p\}}(\bar{s})\}$. Then, the set $R_\mu^{\{p\}}(\bar{s})$ is convex.

Proof: If $\psi_\mu^{\{p\}}(s_1) \le k$ and $\psi_\mu^{\{p\}}(s_2) \le k$, then from the convexity of this function we have

$$\psi_\mu^{\{p\}}(\lambda s_1 + (1 - \lambda)s_2) \le \lambda \psi_\mu^{\{p\}}(s_1) + (1 - \lambda)\psi_\mu^{\{p\}}(s_2) \le k. \tag{30}$$

Taking $k = \psi_\mu^{\{p\}}(\bar{s})$, it follows that the set $R_\mu^{\{p\}}(\bar{s})$ is convex.

Remark: An immediate consequence of Lemma 5 is this: The global minima of $\psi_\mu^{\{p\}}(s)$ form a convex set. Then we have the following extensions of Theorems 4 and 5.

Corollary 3. The (local-global) minima s^* of the convex linear programming problem associated with $\psi_0^{\{p\}}(s)$ form a convex set.

Proof: Theorem 4, Lemma 5, and the fact that the intersection set $R_\mu^{\{p\}}(s^*) \cap F^{\{p\}}$ is convex.

Corollary 4. The global minima of $\psi_0^{\{p\}}(s)$ for any choice of ray paths $\{p\}$ form a convex set that either contains all solutions of the inversion problem or has no intersection with the feasible set F.

Proof: Theorem 5, Lemma 5, and the preceding Remark.

The convex sets constituting the minima in Corollaries 3 and 4 are the same only if they completely solve the inversion problem. Otherwise they are disjoint and distinct sets. To understand more fully the significance of these sets, we will introduce the concept of ghosts [Smith, Solmon, and Wagner 1977; Grünbaum 1980; Louis 1981].

Definition: A ghost is a nonzero model point g (not in Σ) in the null space of the traveltime operator for a particular set of ray paths, *i.e.*, such that

$$\tau_i^{\{p\}}(g) \equiv 0, \text{ for all } i = 1, \dots, m. \tag{31}$$

Thus, if s_1 and s_2 are two distinct models with the same trial traveltimes for all i so that

$$\tau_i^{\{p\}}(s_1) = \tau_i^{\{p\}}(s_2), \tag{32}$$

then clearly

$$\tau_i^{\{p\}}(s_1 - s_2) = 0 \tag{33}$$

for all i and the difference $g = s_1 - s_2 \not\equiv 0$ is a ghost. The solution of the inversion problem is therefore unique (*i.e.*, the convex solution set contains one and only one member) iff the null space of the traveltime operator corresponding to the solution ray paths is empty. If the convex solution set has two or more distinct members, then there is at least one ghost and an infinite number of solutions in the set.

To finish clarifying the structure of the solution set, we establish two more results.

Lemma 6. Let $H_i^{\{p\}} = \{s \mid \tau_i^{(p_i)}(s) = T_i\}$. Then $H_i^{\{p\}}$ is a convex set.

Proof: If s_1 and s_2 are members of $H_i^{\{p\}}$, then

$$\tau_i^{(p_i)}(\lambda s_i + (1 - \lambda)s_2) = \lambda \tau_i^{(p_i)}(s_1) + (1 - \lambda)\tau_i^{(p_i)}(s_2) = T_i. \tag{34}$$

Remark: If the slowness space has been discretized into cells of constant slowness as in [1-3], then each set $H_i^{\{p\}}$ is a hyperplane in the model vector space [2,3].

Theorem 6. Let $I^{\{p\}} = \{s \mid \tau_i^{(p_i)}(s) = T_i, \text{ for } i = 1, \dots, m \text{ or for some subset of the } is\}$. Then the set $I^{\{p\}}$ is convex.

Proof: The set $I^{\{p\}}$ is the intersection of all the sets $H_i^{\{p\}}$ or the intersection of some subset of this collection of sets. Being the intersection of convex sets, $I^{\{p\}}$ is convex.

Remarks: For inconsistent data (measured traveltimes T_i with errors) or for poor choices of ray paths $\{p\}$, the set $I^{\{p\}}$ may be empty. For cells of constant slowness, each nonempty set $I^{\{p\}}$ consists of overlapping intersections of hyperplanes.

As an example of the application of these results to practical problems, suppose s^* is a point where $\psi_0^{\{p\}}(s)$ is minimum. For definiteness, suppose further that $\tau_i^{\{p\}}(s^*) = T_i$ for all i. So the point s^* seems to solve the inversion problem for the set of ray paths $\{p\}$. However, $\{p\}$ may not be the optimal set of ray paths for s^*, *i.e.*, we may find $t_i(s^*) < \tau_i^{\{p\}}(s^*)$ for some or all i. In this situation (which would be fairly typical in practice if we also include some slight errors due to numerical roundoff in these supposed equalities), the apparent agreement between the predicted traveltimes and the data is actually spurious to some degree — often to a large degree.

If there are other models that satisfy the data for the same set of ray paths, Corollary 4 shows that the set of all such models forms a convex set. This convex set either contains all solutions of the inversion problem and therefore lies exactly on the boundary of the global feasible set F, or — the more likely situation — this set does

not solve the problem at all and actually is a convex set of infeasible points. The reason this latter situation is more likely is because in practical implementations we are virtually always working with an approximate set of ray paths. Thus, in programming for inversion, it will often be true that it is counterproductive to try to find models that produce optimum fits to the traveltime data for a given set of (approximate) ray paths. Trying to obtain such optimum fits, as is often done in least-squares or other inversion techniques based only on the magnitudes of the traveltime errors regardless of sign, is programming to fit the noise rather than the signal. This is one of the reasons that such methods so often fail.

4. Implications for Other Inverse Problems

The methods and results presented here apply to a wide class of inverse problems. The proofs were given for the sake of definiteness in terms of the first arrival traveltime inversion problem, but they apply equally to any problem that can be formulated so the data are minima of the pertinent variational problem.

For example, suppose that we wish to invert electrical boundary measurements to obtain the interior conductivity distribution of a body [Kohn and Vogelius 1984; Kohn and Vogelius 1985]. This problem is now commonly known as electrical impedance tomography [Henderson and Webster 1978; Dines and Lytle 1981; Wexler, Fry, and Neuman 1985; Barber and Brown 1986; Daily, Lin, and Buscheck 1987; Yorkey, Webster, and Tompkins 1987; Gisser, Isaacson, and Newell 1988; Lin and Daily 1988]. Then, the set of powers — electrical energy dissipated while current is injected between pairs i of electrodes $\{P_i\}$ — is the pertinent data set. The variational formulation (Dirichlet's principle) states that

$$p_i(\sigma) = \min_{\phi_i^{(trial)}} \int \sigma |\nabla \phi_i^{(trial)}|^2 \, d^3x = \int \sigma |\nabla \phi_i^*[\sigma]|^2 \, d^3x, \tag{35}$$

where $\phi_i^{(trial)}(\vec{x})$ is the trial potential field for the i-th injection pair and $\phi_i^*[\sigma](\vec{x})$ is the potential field distribution that actually minimizes the power dissipation for conductivity distribution $\sigma(\vec{x})$. We define a trial power dissipation by

$$\tilde{p}_i^{(\phi_i)}(\sigma) = \int \sigma |\nabla \phi_i^{(trial)}|^2 \, d^3x. \tag{36}$$

Then, the correspondence between first arrival traveltime inversion and electrical impedance tomography is this:

$$s \rightarrow \sigma,$$
$$t_i(s) \rightarrow p_i(\sigma),$$
$$\tau_i^{(p_i)}(s) \rightarrow \tilde{p}_i^{(\phi_i)}(\sigma),$$
$$dl_i^{(path)} \rightarrow |\nabla \phi_i^{(trial)}|^2 \, d^3x,$$
$$dl_i^*[s] \rightarrow |\nabla \phi_i^*[\sigma]|^2 \, d^3x,$$
$$T_i \rightarrow P_i.$$

All the concepts such as feasibility constraints and feasible sets carry over immediately since

$$P_i = p_i(\sigma_0) \leq \tilde{p}_i(\sigma_0) \tag{37}$$

must again be true if σ_0 is a solution of the inversion problem.

Another remarkable fact is this: For the electrical impedance tomography problem, there are actually two different sets of feasibility constraints. One set is for the variational method (Dirichlet's principle) outlined above. The other is for its dual (Thomson's principle). The existence of dual variational principles will be a general result whenever the variational principles involved are true minimum principles. Fermat's principle is actually not in this class since it is only a stationary principle; but for first arrival traveltime inversion, it is nevertheless valid to treat this principle as a minimum principle, since the data are truly minima. Further discussion of the feasibility constraints for electrical impedance tomography may be found in Berryman and Kohn [1990].

5. Discussion of Convexity Results

If a minimum of an objective function has been attained in programming for nonlinear inversion, several questions often come to mind: "Is this a local or a global minimum?" "If more than one local minimum is found, how are the minima related to global minima?" These issues have been at least partially clarified by Theorems 3-6 and Corollaries 3 and 4.

The observation that variational constraints rigorously imply the existence of feasible sets of trial models for inversion problems has played a crucial role in the analysis presented here. Only the proofs of Lemmas 1, 2, 5, and 6 and Theorem 6 are independent of the definitions of the feasible sets. These constraints are therefore vital in our efforts to elucidate the nature of solution sets for inversion problems. They have also proven to be vital for improving the behavior of iterative numerical techniques for solving the inversion problems [Berryman 1989b; Berryman 1990].

6. Reconstruction Algorithm

Many reconstruction algorithms making use of the feasibility constraints are possible. Only one will be summarized here.

Practical algorithms for reconstruction must be formulated in terms of discrete variables. Therefore, the notation used in this section will differ slightly from that used in the preceeding sections. To set notation, let τ be the computed traveltime vector such that $\tau' = (\tau_1, \ldots, \tau_m)$, where τ_i is the traveltime along the i-th ray path (a superscript ' denotes the transpose), and let T such that $T' = (T_1, \ldots, T_m)$ be the measured traveltime vector along the same set of ray paths. We form our model in two-dimensions by dividing the rectangular region enclosed by our sources and receivers into rectangular cells of constant slowness. Then, s is the model slowness vector $s' = (s_1, \ldots, s_n)$, with s_j being the slowness of the j-th cell. For forward modeling, s and τ are related by the equation

$$Ms = \tau, \tag{38}$$

where M is an $m \times n$ matrix whose matrix elements $l_{i,j}$ are determined by the length of the i-th ray path as it passes through the j-th cell. Eq. (38) simply states that the total traveltime along a ray path is the sum of the traveltimes through each of the cells traversed by the ray. To find M, we use a modified version of the two-point bending algorithm of [Prothero, Taylor, and Eikemeyer 1988].

To solve the inverse problem, the forward problem (38) is now replaced by the

m feasibility constraints

$$(Ms)_i \geq T_i. \tag{39}$$

This fact follows from Fermat's principle: the first arrival necessarily followed the path of minimum traveltime for the model s. Thus, (39) must be satisfied by any ray path matrix M if s is the true model and therefore any model that violates (39) along any of the ray paths is not a feasible model. An exact solution to the inversion problem is found if and only if all of inequalities in (39) become identities for some choice of model slowness vector s.

The feasibility violation number defined in (15) is now given explicitly for any combination of ray path matrix M, slowness vector s, and measured traveltime times T by

$$N_M(s) = \sum_{i=1}^{m} \theta[T_i - (Ms)_i]. \tag{40}$$

This formula amounts to counting the number of rays that violate the constraints (39). This number equals zero in the feasible region. Furthermore, it is a monotonically increasing function of distance from the feasibility boundary – once one of the hyperplanes of (39) is crossed we never cross it again if we keep moving in the same direction. Thus, this number is cheap to compute and gives us a rough idea of how close we are to the feasibility boundary.

The key ideas behind one new reconstruction algorithm may now be summarized as follows: Given a set of transmitter-receiver pairs and any model slowness s, Fermat's principle may be used to find the ray path matrix M associated both with s and with any slowness γs (where $\gamma > 0$) in the same direction as s. An optimum scale factor γ may be found by doing a weighted least-squares fit to the traveltime data. The best weights to use are described in detail elsewhere [Berryman 89a].

Having found the optimum slowness $s_b = \gamma s$ in the given direction, we next attempt to improve the model by finding another direction in the slowness vector space that gives a still better fit to the traveltime data. As many others have done, we first compute a damped least-squares [Marquardt 1963; Aki and Richards 1980; Nolet 1987; Bording, Gersztenkorn, Lines, Scales, and Treitel 1987; Scales, Gersztenkorn, and Treitel 1988] solution s_μ. Next we note that both of the points found so far are guaranteed to lie in the nonfeasible part of the vector space – at least one and generally about half of the ray paths for both of these models will have traveltimes shorter than that of the measured data. Furthermore, although the point s_μ gives a better fit to the traveltime data, this fit is certainly spurious to some extent because it is based on the wrong ray-path matrix; the ray-path matrix used in the computation of s_μ from s_b is the one that was correct for slownesses along the direction s_b. Thus, both of the points we have found so far are off the feasibility boundary and the second point s_μ is of questionable worth also because its value was obtained in an inconsistent manner.

Recall also that the solution of (38), if one exists, must lie on the feasibility boundary. So we would like to use s_μ and s_b to help us find a point on this boundary that is optimum in the sense that it is as consistent as possible with the ray path matrix M, with the measured traveltimes T, and with the feasibility constraints. The fact that traveltime error may be reduced by moving in the direction of s_μ may still give us an important clue as to the best direction to move in the vector space,

Figure 3. A snapshot of the main features of one new reconstruction algorithm during one iteration step. The axes are the slownesses for any two cells (j and k) in the model space. Point s_b is the initial value for the next step of the iteration sequence. Point s_μ is the damped weighted least-squares solution. Point s_l is a simple linear combination of s_b and s_μ chosen so it has the least feasibility violations. Point s_f is the unique point on the feasibility boundary in the same direction as s_l.

i.e., we may want to move in the direction $s_\mu - s_b$ but perhaps we should stop before we arrive at s_μ. How far then should we move in this direction?

If we consider Figure 3, we are reminded that the feasible region is convex. Therefore, there may exist a point between the points s_μ and s_b that is closer to the feasible region than either of the two end points. If we could find this point s_l and then scale up to the point in the same direction lying on the feasibility boundary, then we have found s_f in the figure. In principle, it is possible to find the point on this line closest to the feasibility boundary. But it is much easier to compute the feasibility violation number $N_M(s)$. As we move along the direction $s_\mu - s_b$ from s_b we generally find that this number achieves a minimum value at some intermediate point. This point of minimum $N_M(s)$ is the point s_l in the figure.

Now it is possible to prove that all three of these points are distinct unless we have found an exact solution to the inversion. So unless we have already solved the problem, these three points form a triangle and the size of the triangle provides an estimate of how far we are from a solution.

These ideas have all been repeatedly confirmed in a large number of reconstructions on synthetic examples and on real seismic and electromagnetic traveltime data. Examples (including color illustrations) may be found in the references [Berryman 89b; Berryman 90]. Whereas the standard damped least-squares method typically becomes unstable after several iterations when the contrasts are high, the new algorithm is completely stable and gives reasonable results.

7. Discussion of Reconstruction Algorithm

For real problems with high contrasts and noisy data, a damped least-squares method generally does not converge and no criterion is has previously been available to determine when to terminate the iteration sequence. However, by modifying the damped least-squares approach using the information available in the the feasibility criteria, the resulting algorithm converges relatively quickly to a solution in its convex soution set. Stable iteration to such a convergence set is the most that could be expected when the traveltime data have errors and are therefore inconsistent.

References

K. Aki, A. Christoffersson, and E. S. Husebye (1976), *Determination of the three-dimensional seismic structure of the lithosphere*, J. Geophys. Res., 82, pp. 277-296.

K. Aki and P. G. Richards (1980), *Quantitative Seismology — Theory and Methods*, Vol. II, Freeman, San Francisco, Section 12.3.

D. C. Barber and B. H. Brown (1986), *Recent developments in applied potential tomography – APT*, in Information Processing in Medical Imaging, S. L. Bacharach, ed., Martinus Nijhoff, Dordrecht, pp. 106-21.

J. G. Berryman (1989a), *Weighted least-squares criteria for seismic traveltime tomography*, IEEE Trans. Geosci. Remote Sensing, 27, pp. 302-309.

J. G. Berryman (1989b), *Fermat's principle and nonlinear traveltime tomography*, Phys. Rev. Lett., 62, pp. 2953-2956.

J. G. Berryman (1990), *Stable iterative reconstruction algorithm for nonlinear traveltime tomography*, Inverse Problems, 6, pp. 21-42.

J. G. Berryman and R. V. Kohn (1990), *Variational constraints for electrical-impedance tomography*, Phys. Rev. Lett., 65, pp. 325-328.

P. Bois, M. LaPorte, M. Lavergne, and G. Thomas (1972), *Well-to-well seismic measurements*, Geophysics, 37, pp. 471-480.

R. P. Bording, A. Gersztenkorn, L. R. Lines, J. A. Scales, and S. Treitel, *Applica-*

tions of seismic travel-time tomography, Geophys. J. Roy. Astron. Soc., 90, pp. 285-303.

M. Born and E. Wolf (1959), *Principles of Optics*, Pergamon, London, pp. *xix-xx*, 127-129.

R. N. Bracewell and A. C. Riddle (1967), *Inversion of fan-beam scans in radio astronomy*, Astrophys. J., 150, pp. 427-434.

A. M. Cormack (1963), *Representation of a function by its line integrals, with some radiological applications*, J. Appl. Phys., 34, pp. 2722-2727.

W. Daily, W. Lin, and T. Buscheck (1987), *Hydrological properties of Topopah Spring tuff: Laboratory measurements*, J. Geophys. Res., 92, pp. 7854-7864.

A. J. Devaney (1984), *Geophysical diffraction tomography*, IEEE Trans. Geosci. Remote Sensing, GE-22, pp. 3-13.

K. A. Dines and R. J. Lytle (1979), *Computerized geophysical tomography*, Proc. IEEE, 67, pp. 1065-1073.

K. A. Dines and R. J. Lytle (1981), *Analysis of electrical conductivity imaging*, Geophysics, 46, pp. 1025-1036.

A. V. Fiacco and G. P. McCormick (1968), *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, Wiley, New York.

P. Gilbert (1972), *Iterative methods for the three-dimensional reconstruction of an object from projections*, J. Theor. Biol., 36, pp. 105-117.

R. Gordon, R. Bender, and G. T. Herman (1970), *Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography*, J. Theor. Biol., 29, pp. 471-481.

D. G. Gisser, D. Isaacson, and J. C. Newell (1988), *Theory and performance of an adaptive current tomography system*, Clin. Phys. Physil. Meas., 9A, pp. 35-41.

F. A. Grünbaum (1980), *A study of Fourier space methods for "limited angle" image reconstruction*, Numer. Funct. Anal. Optimiz., 2, pp. 31-42.

R. P. Henderson and J. G. Webster (1978), *An impedance camera for spatially specific measurements of the thorax*, IEEE Trans. Biomech. Eng., BME-25, pp. 250-254.

G. T. Herman (1980), *Image Reconstruction from Projections — The Fundamenetals of Computerized Tomography*, Academic, New York, pp. 1-18, 103-106.

G. N. Hounsfield (1973), *Computerized transverse axial scanning (tomography): Part I. Description of system*, Brit. J. Radiol., 46, pp. 1016-1022.

J. H. Justice, A. A. Vassiliou, S. Singh, J. D. Logel, P. A. Hansen, B. R. Hall, and J. J. Solanki (1989a), *Acoustic tomography for monitoring enhanced oil recovery*, Geophysics: The Leading Edge of Exploration, 8, pp. 12-19.

J. H. Justice, A. A. Vassiliou, S. Singh, J. D. Logel, P. A. Hansen, B. R. Hall, and J. J. Solanki (1989b), *Tomographic imaging in hydrocarbon reservoirs*, Intern. J. Imaging Sys. Techn., 1, pp. 62-72.

A. C. Kak (1984), *Image reconstruction from projections*, in Digital Image Processing Techniques, M. P Ekstrom, ed., Academic, Orlando, pp. 111-170.

R. V. Kohn and M. Vogelius (1984), *Determining conductivity by boundary measurements*, Comm. Pure Appl. Math., 37, pp. 289-298.

R. V. Kohn and M. Vogelius (1985), *Determining conductivity by boundary measurements II, Interior results*, Comm. Pure Appl. Math., 38, pp. 643-667.

D. L. Lager and R. J. Lytle (1979), *Determining a subsurface electromagnetic profile from high-frequency measurements by applying reconstruction-technique algorithms*, Radio Sci., 12, pp. 249-260.

W. Lin and W. Daily (1988), *Laboratory-determined transport properties of core from the Salton Sea Scientific Drilling Project*, J. Geophys. Res., 93, pp. 13047-13056.

A. K. Louis (1981), *Ghosts in tomography – The null space of the Radon transform*, Math. Meth. Appl. Sci., 3, pp. 1-10.

R. J. Lytle and K. A. Dines (1980), *Iterative ray tracing between boreholes for underground image reconstruction*, IEEE Trans. Geosci. Remote Sensing, GE-18, pp. 234-240.

D. W. Marquardt (1963), *An algorithm for least-squares estimation of nonlinear parameters*, J. Soc. Indust. Appl. Math., 11, pp. 431-441.

R. K. Mueller, M. Kaveh, and G. Wade (1979), *Reconstructive tomography and applications to ultrasonics*, Proc. IEEE, 67, pp. 567-587.

F. Natterer (1986), *The Mathematics of Computerized Tomography*, Wiley, New York, pp. 1-8, 102-157.

G. Nolet (1987), *Seismic wave propagation and seismic tomography*, in Seismic Tomography With Applications in Global Seismology and Exploration Geophysics, G. Nolet, ed., Reidel, Dordrecht, pp. 1-23. This book also contains fourteen other articles on seismic tomography by various authors summarizing the state of the art as of the date of publication.

W. A. Prothero, W. J. Taylor, and J. A. Eickemeyer (1988), *A fast two-point, three-dimensional raytracing algorithm using a simple step search method*, Bull. Seismol. Soc. Am., 78, pp. 1190-1198.

J. A. Scales, A. Gersztenkorn, and S. Treitel (1988), *Fast l_p solution of large, sparse, linear systems: Application to seismic travel time tomography*, J. Comput. Phys., 73, pp. 314-333.

H. Schomberg (1978), *An improved approach to reconstructive ultrasound tomography*, J. Phys. D: Appl. Phys., 11, pp. L181-L185.

K. T. Smith, D. C. Solmon, and S. L. Wagner (1977), *Practical and mathematical aspects of the problem of reconstructing objects from radiographs*, Bull. Am. Math. Soc., 83, pp. 1227-1270.

A. Wexler, B. Fry, and M. R. Neuman (1985), *Impedance-computed tomography algorithm and system*, Appl. Opt., 24, pp. 3985-92.

E. Wolf (1969), *Three-dimensional structure determination of semi-transparent objects from holographic data*, Optics Commun., 1, pp. 153-156.

T. J. Yorkey, J. G. Webster, and W. J. Tompkins (1987), *Comparing reconstruction algorithms for electrical impedance tomography*, IEEE Trans. Biomed. Eng., BME-34, pp. 843-852.