CONVEXITY PROPERTIES OF INVERSE PROBLEMS WITH VARIATIONAL CONSTRAINTS

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

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.

I. Introduction

In a series of papers (**1-3**), 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 principle contribution of this work is the observation that, when an inverse problem can be formulated so the data are minima of one of the variational problems of mathematical physics, rigorous physically-based feasibility (or admissibility) constraints can be found for the corresponding nonlinear inversion problems. These constraints guarantee that any optimal solution of the inverse problem found using convex programming techniques lies in a convex feasible region of the model space. Furthermore, points on the boundary of the feasible set can be found in a constructive fashion. Also, 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 possible to establish otherwise.

We have three main goals for the paper: (i) to establish that the idea of using variational/feasibility constraints for inversion is both rigorous and applicable to a wide class of physical problems, (ii) 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, and (iii) 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.

For definiteness, we use first arrival traveltime inversion as our primary example. However, it will be 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 immediately (**2,3**).

II. Convexity Properties of Inverse Problems

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 $i = 1, \dots, m$), 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 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. 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 (**4**). 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 (**2,3**). 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

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

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: (i) a feasible part whose members s are like s_0 in that they satisfy the constraints

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

for all source-receiver pairs i and all ray paths in the trial set and (ii) 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 (**2,3**). 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

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

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.

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 practical applications. The feasibility boundary bounds 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 (**2,3**).

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 $V^{\{p\}}(n) = \{s \,|\, n \ge N^{\{p\}}(s)\}$, where n is a nonnegative 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

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

III. Convex Programming for Inverse Problems

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 $t_i(s) \geq T_i$ 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 $\tau_i^{(p_i)}(s) \geq T_i$ 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 (**4**).) 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

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

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 the nonlinear programming problem is not easy. 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 (**1**). 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.

Since homogeneity is the only property of $t_i(s)$ used in (22)-(27), we can immediately prove the corresponding statement for the path dependent problem. In the equations (22)-(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 (**4**), 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 (**2,3**).

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 (**1**). 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 (**5-7**).

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

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

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

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

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 $H_i^{\{p\}} = \{s \,|\, \tau_i^{(p_i)}(s) = T_i\}$. 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 $I^{\{p\}} = \{s \,|\, \tau_i^{(p_i)}(s) = T_i$, for $i = 1,\ldots,m$ 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.

IV. Extensions

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 (**8**,**9**). This problem is commonly known as electrical impedance tomography. Then, the set of powers — 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:

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

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

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

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. This fact was pointed out to the author in a private discussion with R. V. Kohn.

V. Discussion

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 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 (**2,3**).

Acknowledgments. I thank A. J. DeGroot, R. V. Kohn, and G. J. F. MacDonald for very helpful conversations. Work performed under the auspices of the U. S. Department of Energy by Lawrence Livermore National Laboratory under contract No. W-7405-ENG-48 and supported specifically by the LLNL Engineering Research Program and by the Geosciences Research Program of the DOE Office of Energy Research within the Office of Basic Energy Sciences, Division of Engineering and Geosciences.

References

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

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

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

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

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

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

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

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

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