Stanford Exploration Project, Report 84, May 9, 2001, pages 1–375

Variational structure of inverse problems in wave propagation and vibration

James G. Berryman[1]

ABSTRACT

Practical algorithms for solving realistic inverse problems may often be viewed as problems in nonlinear programming with the data serving as constraints. Such problems are most easily analyzed when it is possible to segment the solution space into regions that are feasible (satisfying all the known constraints) and infeasible (violating some of the constraints). Then, if the feasible set is convex or at least compact, the solution to the problem will normally lie on the boundary of the feasible set. A nonlinear program may seek the solution by systematically exploring the boundary while satisfying progressively more constraints. Examples of inverse problems in wave propagation (traveltime tomography) and vibration (modal analysis) are presented to illustrate how the variational structure of these problems may be used to create nonlinear programs using implicit variational constraints.

INTRODUCTION

Although the most common method used to analyze inverse problems is perturbation theory and/or linearization, it is generally recognized that most inverse problems of practical interest are mathematically nonlinear. The inherent nonlinearities lead to the necessity of distinguishing local minima from global minima. Optimization methods based on objective functionals typically produce "solutions" that minimize the functional "locally" — *i.e.*, in the vicinity of the starting point of the minimization process. The methods discussed in the present paper have been developed specifically to improve our insight into the structure of certain classes of nonlinear inverse problems. Furthermore, once this structure is understood, methods of solution of the inverse problems using nonlinear programming methods are immediately suggested by the particular variational structure observed. The first example presented uses vibration frequencies of a taut string to deduce its mass distribution. The pertinent variational principle for this problem is Rayleigh-Ritz. The second example uses traveltime data to deduce wave velocity using Fermat's principle of least-time. The remaining examples use frequencies of free oscillation of the Earth to deduce density and elastic parameters. The pertinent variational principle is again Rayleigh-Ritz, but significant differences in the variational structure arise between these examples and the string density vibration analysis problem. The structures possessed by members of this set of problems appears to span the range of possible variational

[1]**email:** berryman@sep.stanford.edu

Modal analysis for string density

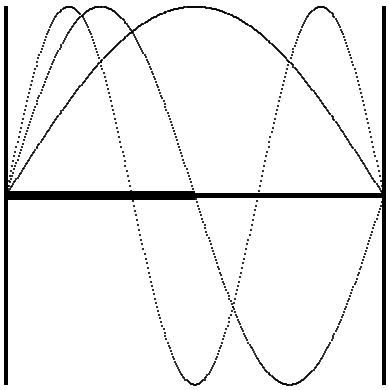

Figure 1: Density distribution of the string may be determined by analyzing the eigenfrequencies associated with its modes of vibration.

structures and provides some new insights into potential inversion algorithms for some of the problems.

STRING VIBRATION ANALYSIS

The equation of vibration for a simple string is

$$\rho \frac{\partial^2 u}{\partial t^2} = T \frac{\partial^2 u}{\partial x^2}, \tag{1}$$

where t is the time, x is the spatial coordinate of the string along its length, $\rho(x)$ is the density distribution of the string, T is the tension in the string (assumed constant), and $u(x,t)$ is the normal displacement of the string from its equilibrium position as a function of both position and time. The ends of the string are fixed, so the boundary conditions are $u(0,t) = 0 = u(1,t)$. I assume that the string's temporal motion may be decomposed into its Fourier components so that I may study standing waves on the plucked string. The time dependence is assumed to separate out in one of the forms $\sin\omega t$, $\cos\omega t$, or $\exp i\omega t$, where $\omega = 2\pi f$ is angular frequency with dimensions of radians/sec, while f is frequency in Hz. Then, the equation (1) reduces to

$$-\omega^2 \rho u = T u_{xx}, \tag{2}$$

where I have now introduced the subscript notation for derivatives such that $u_x \equiv \partial u/\partial x$, $u_{xx} \equiv \partial^2 u/\partial x^2$, etc. I assume that the tension T in the string is known, but the density distribution $\rho(x)$ of the string is unknown. Our goal is to determine to what extent $\rho(x)$ can be determined using knowledge of modes of vibration of the string. In particular, I assume that

some measurements of the vibration frequencies for standing waves have been made. Let ω_n be the eigenfrequency and $u_n(x)$ the eigenfunction of the n-th Fourier component for $n \geq 1$. Higher n corresponds to eigenfunctions with more internal nodes, *e.g.*, $n = 1$ has no internal nodes, $n = 2$ has one internal node, etc. Then, after choosing unit tension $T = 1$, I multiply (2) by $u_n(x)$ and integrate along the length of the string to show that

$$\omega_n^2(\rho) = \frac{\int_0^1 u_{n,x}^2(x)\,dx}{\int_0^1 \rho(x) u_n^2(x)\,dx}. \tag{3}$$

To arrive at (3), I integrated once by parts ($\int u_n u_{n,xx}\,dx = -\int u_{n,x}^2\,dx$) using the fact that $u_n(0) = 0 = u_n(1)$ to eliminate the boundary contribution. Equation (3) is an identity satisfied by the n-th eigenfrequency and relating it to the integrated properties of the n-th eigenfunction. Now the Rayleigh-Ritz method for characterizing eigenvalues [Courant and Hilbert, 1953] may be applied to the string problem and it shows that

$$\omega_1^2(\rho) \leq \mathcal{R}[\rho, v_1] \equiv \frac{\int_0^1 v_{1,x}^2(x)\,dx}{\int_0^1 \rho(x) v_1^2(x)\,dx}, \tag{4}$$

where $\mathcal{R}$ is the Rayleigh-Ritz quotient with $v_1(x)$ being any trial function satisfying the boundary conditions $v_1(0) = 0 = v_1(1)$ whose first derivative with respect to x is finite everywhere along the interval containing the string. When the string density distribution $\rho(x)$ is known, the Rayleigh-Ritz method may be used as a numerical method for finding estimates of both the eigenfrequency ω_1 and the eigenfunction $u_1(x)$; the theory shows that $\mathcal{R}[\rho, v_1] \equiv \omega_1^2(\rho)$ if and only if $v_1(x) = u_1(x)$ almost everywhere. Similarly, if I impose constraints on the trial function such as requiring $v_n(x)$ to have $n - 1$ interior nodes [Coddington and Levinson, 1955], then (4) generalizes to

$$\omega_n^2(\rho) \leq \mathcal{R}[\rho, v_n] \equiv \frac{\int_0^1 v_{n,x}^2(x)\,dx}{\int_0^1 \rho(x) v_n^2(x)\,dx}. \tag{5}$$

The characteristic frequencies (the ωs) are those defined in (3) for the higher order modes. I will sometimes refer to the equations (4) and (5) as the "forward problem," since in these equations I assume that the density distribution is a known quantity. The "inverse problem" is the harder problem of taking measured values of the eigenfrequencies ω_n and attempting to solve for the unknown density distribution $\rho(x)$.

Feasibility analysis

There are two tricks that make it possible in some cases to use the variational functionals to analyze inversion problems. The first trick is a result of what I call the "scale invariance" property of the eigenfunctions in certain problems. For a given density distribution ρ and its corresponding eigenfunction $u(x)$, the only effect of multiplying ρ by a constant γ is to change the eigenfrequency by the factor $1/\sqrt{\gamma}$. This result follows since

$$\mathcal{R}[\gamma\rho, u] = \mathcal{R}[\rho, u]/\gamma, \tag{6}$$

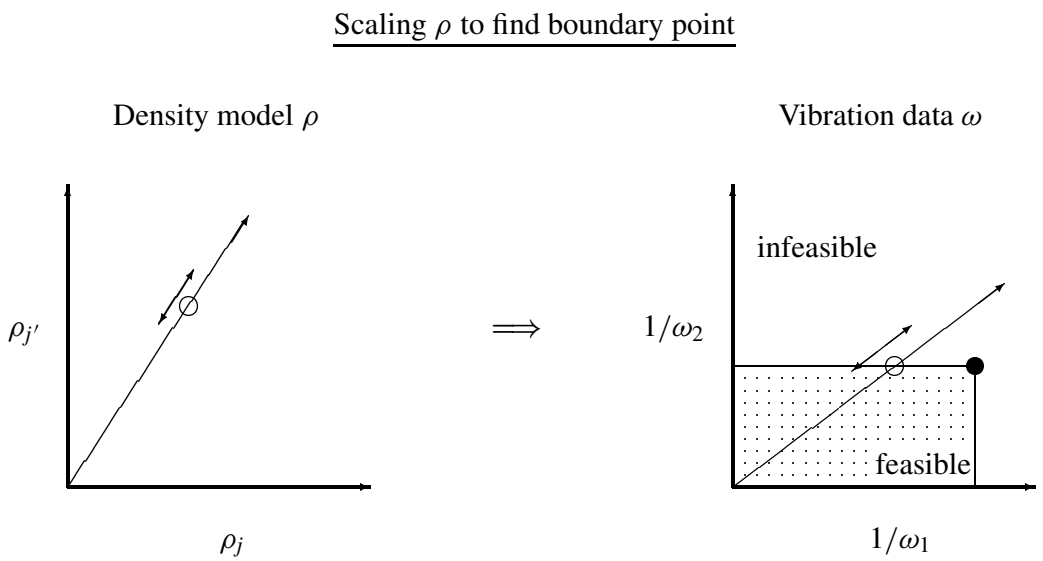

Figure 2: Feasible part of the density space is determined implicitly by the explicit boundaries defined by the frequency data.

which implies that $\omega(\gamma\rho) = \omega(\rho)/\sqrt{\gamma}$. I conclude that each eigenfunction $u(x)$ is determined only by the relative variations in the density, not by the absolute scale. The second trick is a result of the linear dependence of the denominator of the Rayleigh-Ritz functional on the density distribution ρ. I take advantage of this linearity in ρ with greater ease by working with the reciprocal of $\mathcal{R}$, so I first note that

$$\frac{1}{\mathcal{R}[\rho, v]} = \frac{\int \rho(x) v^2(x)\, dx}{\int v_x^2(x)\, dx} \leq \omega^{-2}(\rho). \tag{7}$$

Now, if I have made measurements of some of the eigenfrequencies of the string, I can ask the following question: Are there density distributions that violate the inequality in (7)? That is, if I try to compute the left hand side of (7) using an arbitrary trial density distribution $\tau(x)$, when will the inequality be satisfied and when will it be violated? I use the answers to these questions to define two distinct classes of trial model density distributions $\tau(x)$:

$$\text{If } \frac{1}{\mathcal{R}[\tau, v_n]} \leq \omega_n^2(\rho) \text{ for each measured frequency } \omega_n, \text{ then } \tau(x) \text{ is } \textit{feasible}. \tag{8}$$

However,

$$\text{if } \frac{1}{\mathcal{R}[\tau, v_n]} > \omega_n^2(\rho) \text{ for any measured frequency } \omega_n, \text{ then } \tau(x) \text{ is } \textit{infeasible}. \tag{9}$$

As defined here, the concepts of *feasible* and *infeasible* trial density distributions depend explicitly on the available data ω_n and implicitly on the trial functions $v_n(x)$. However, it is also possible to show [Berryman, 1991] that universal (global) feasible and infeasible sets exist that are independent of the particular trial functions chosen (dependence on the particular choice

of the subset of those measured frequencies which are used in the analysis still remains). The concepts of *feasible* and *infeasible* sets are commonly found in texts on nonlinear programming methods [Fiacco and McCormick, 1990]. Since the method I develop for solving the inverse problem is a type of nonlinear programming method, it is not surprising that these concepts also arise in the context of numerical methods for solving inverse problems. Now, I derive a very useful property of the class of feasible density distributions. Consider some trial eigenfunction $v(x)$ and two *feasible* trial densities $\tau_a(x)$ and $\tau_b(x)$, which therefore (by assumption) satisfy

$$\frac{1}{\mathcal{R}[\tau_a, v]} \leq \omega^{-2}(\rho) \quad \text{and} \quad \frac{1}{\mathcal{R}[\tau_b, v]} \leq \omega^{-2}(\rho). \tag{10}$$

Let λ be a number in the range $0 \leq \lambda \leq 1$. Then, taking a linear combination of the two expressions in (10) gives

$$\frac{\lambda}{\mathcal{R}[\tau_a, v]} + \frac{1-\lambda}{\mathcal{R}[\tau_b, v]} \leq \omega^{-2}(\rho). \tag{11}$$

But, the left hand side of (11) may be rewritten as

$$\frac{\int [\lambda \tau_a(x) + (1-\lambda)\tau_b(x)]\, v^2(x)\, dx}{\int v_x^2(x)\, dx} = \frac{1}{\mathcal{R}[\tau_\lambda, v]}, \tag{12}$$

where I have defined the convex combination of the trial densities to be

$$\tau_\lambda(x) \equiv \lambda \tau_a(x) + (1-\lambda)\tau_b(x). \tag{13}$$

Combining (11) and (12) shows that

$$\frac{1}{\mathcal{R}[\tau_\lambda, v]} \leq \omega^{-2}(\rho). \tag{14}$$

But any trial function that satisfies the analog of (14) for all frequencies considered from the measurement set is by definition feasible. So, if $\tau_a(x)$ and $\tau_b(x)$ are both feasible, their convex combination $\tau_\lambda(x)$ is also feasible. The definition of a convex set is this: a set such that the convex combination — *e.g.*, $\tau_\lambda = \lambda\tau_a + (1-\lambda)\tau_b$ — of any two members is also a member of the set. So, this fact implies that *the set of all feasible density distributions for a given set of vibration data (and a given set of trial eigenfunctions) is a convex set*. In general, convex sets are very useful in computations because they are often compact and always have smooth boundaries. Now, the first trick (scale invariance) plays an interesting and important role in the analysis. Suppose I have any trial density distribution τ_u and I have found the eigenfunctions $u_n(x)$ and eigenvalues $\omega_n(\tau_u)$ associated with this distribution. Then, the eigenfunctions $u_n(x)$ are also the eigenfunctions for all densities of the form $\gamma\tau_u(x)$, where γ is an arbitrary positive scalar. It is not hard to show [Berryman, 1991] that there always exists a choice of $\gamma = \gamma_u$ such that $\gamma_u\tau_u(x)$ lies exactly on the boundary of the feasible set for the inversion problem. It follows then that, if $\gamma \leq \gamma_u$, the density $\gamma\tau_u$ lies in the feasible part of the model space; while, if $\gamma > \gamma_u$, the density $\gamma\tau_u$ lies in the infeasible part. This characteristic of the feasible set allows us to produce a simple geometrical interpretation of the feasible set. In particular,

for the vibrating string inversion problem, *the feasible set of model densities is compact and occupies a convex region in the neighborhood of the origin of the model space* [*i.e.,* near $\rho(x) \equiv 0$]. I easily prove this statement. First note that a mass density must be nonnegative, so the physical model space is the convex set (actually a cone) of all possible nonnegative density distributions. Then, by taking the trial eigenfunction to be $v_1(x) = \sin \pi x$ with the trial density $\tau(x) = \tau_0$ (constant), and noting that

$$\frac{1}{\mathcal{R}[\tau_0, v_1]} = \frac{\tau_0 \int \sin^2 \pi x \, dx}{\pi^2 \int \cos^2 \pi x \, dx} = \frac{\tau_0}{\pi^2} \tag{15}$$

satisfies the feasibility constraint if

$$\tau_0 \leq \pi^2 \omega_1^{-2}(\rho). \tag{16}$$

The first measured eigenfrequency ω_1 then determines an upper bound on the value of τ_0 for constant density models. For nonconstant models such as the two component string with $\tau(x) = \tau_a$ for $0 \leq x < \frac{1}{2}$ and $\tau(x) = \tau_b$ for $\frac{1}{2} < x \leq 1$, the same argument shows that

$$\frac{1}{\mathcal{R}[\tau, v_1]} = \frac{\tau_a \int_0^{\frac{1}{2}} \sin^2 \pi x \, dx + \tau_b \int_{\frac{1}{2}}^1 \sin^2 \pi x \, dx}{\pi^2 \int_0^1 \cos^2 \pi x \, dx} = \frac{\tau_a + \tau_b}{2\pi^2} \leq \omega_1^{-2}(\rho). \tag{17}$$

Thus, (17) show that τ_b is bounded above by the straight line

$$\tau_b \leq 2\pi^2 \omega_1^{-2}(\rho) - \tau_a, \tag{18}$$

producing tighter bounds on τ_b as τ_a increases. When $\tau_a = \tau_b$, this problem reduces to the constant model result found earlier. These simple examples can be used to produce further restrictions on model densities. Although the analysis presented is fine for the lowest eigenfrequency, complications arise for the higher frequencies ω_n, for all $n \geq 2$. For these higher eigenfrequencies, the Rayleigh-Ritz functional requires some constraints on the trial eigenfunctions (*i.e.*, for ω_n, the trial function v_n should be orthogonal to the eigenfunctions u_m for $m < n$). These restrictions on the eigenfunctions guarantee the existence of correlations between the model and the trial eigenfunctions that would be neglected if we tried to make direct use of the preceding analysis for the higher order eigenfunctions. I will show what the consequences of these correlations are when I present the general analysis of variational structure for inverse problems later in the paper. The main result will be that a feasible set for the Rayleigh-Ritz variational problem still exists, but it can be nonconvex when the variational functionals become nonlinear as they do when such correlations arise.

Algorithm

Now I describe a general algorithm for solving the inverse problem for a two-segment string. This algorithm involves a lot of forward modeling, but will always produce a good approximation to the solution. First, note that the symmetry of the problem guarantees that if (τ_1, τ_2) solves the problem then so does (τ_2, τ_1). (I am presently considering only the frequencies as

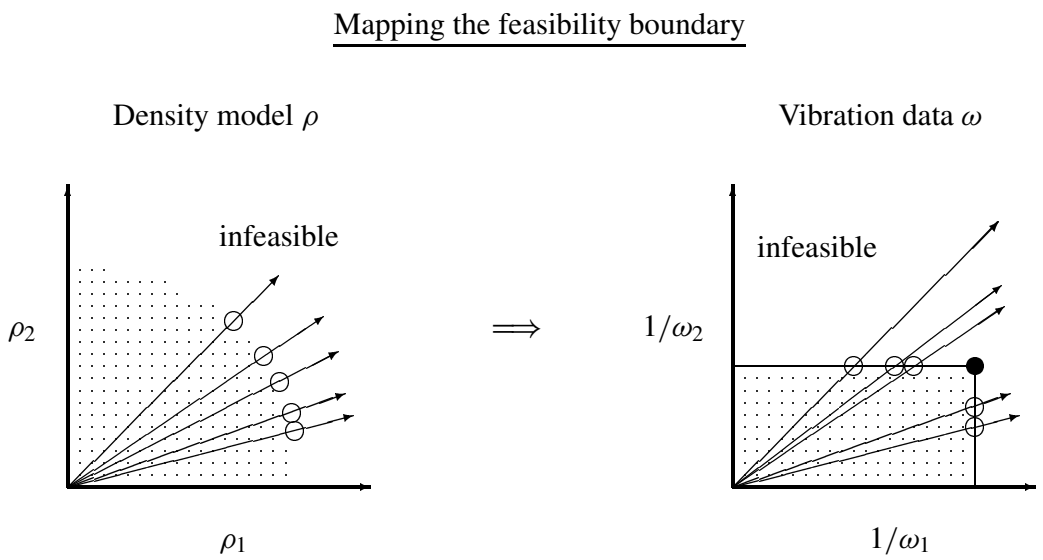

Figure 3: By scaling many density distributions, the location of the feasibility boundary can be mapped.

data. This nonuniqueness of the solution can be easily removed by considering the eigenfunctions as data as well.) Thus, I only need to consider half of the positive quadrant, say for models such that $\tau_1 \geq \tau_2$. Then, picking models evenly spaced in angle for (say) 10 angles up to $\theta = \pi/4$, so $(\tau_1, \tau_2) = (\tau_0 \cos\theta, \tau_0 \sin\theta)$ with $\theta = \pi/40, \pi/20, \ldots, \pi/4$. (The precise value of τ_0 does not matter, since these points will eventually be scaled onto the feasibility boundary.) Having chosen a set of initial points, I now do forward modeling on each of these string models. After scaling to the feasibility boundary, I check satisfaction of the data and pick the two adjacent points that best satisfy the data. Then, I divide the angular region bracketed by these two angles into 10 smaller angles, and repeat this process until I have a satisfactory solution. The algorithm just described is basically a "shooting" algorithm for the inverse problem. The algorithm just described is probably overly complex for the two component string problem. It is not hard to show that the two lowest frequencies are enough to determine the densities of the two segments (although not enough to tell which is which) when it is known that the segments are of equal length. For such a simple problem, other algorithms that are more efficient could be devised to make optimal use of this information. The real advantage of the approach becomes more apparent when we consider more complex models involving three or more string segments. The inversion algorithm just described is easily generalized to many segments of constant density and the resulting algorithm is complicated only by the need to define an efficient means of choosing points in the model space for each forward modeling phase of the calculation. It is also clear that algorithms of this type can easily be parallelized, with an individual processor assigned to do a single forward modeling computation at each step of the algorithm.

TRAVELTIME INVERSION PROBLEM

A typical problem arising in seismic traveltime inversion in 2- and 3-dimensional heterogeneous media is the need to infer the (isotropic) compressional-wave slowness (reciprocal of velocity) distribution of a medium, given a set of observed first-arrival traveltimes between sources and receivers of known location within the medium [Dines and Lytle, 1979; Lytle and Dines, 1980]. This problem is common for crosswell seismic transmission tomography imaging a 2-D region occupying the plane between two vertical boreholes in oil field applications. I could also consider the problem of inverting for wave slowness when the absolute traveltimes are not known perfectly, as is normally the case in earthquake seismology for whole Earth structure [Aki and Richards, 1980].

Wave slowness models

When a sound wave or seismic wave is launched into a medium, it takes time for the influence of the wave to progress from a point close to the source to a more distant point. The time taken by the wave to travel from one point of interest to the next is called the *traveltime*. The *local slowness* is the inverse of the local wave speed. It is most convenient to develop inversion and tomography formulas in terms of wave slowness models, because the pertinent equations are linear in slowness. I might consider three kinds of slowness models. I allow the slowness to be a general function $s(\mathbf{x})$ of the position $\mathbf{x}$. However, I often make one of two more restrictive assumptions that (*i*) the model comprises homogeneous cells (in 2-D), or blocks (in 3-D), with s_j then denoting the slowness value of the jth cell, or block. Or (*ii*) the model is defined in terms of a grid with values of slowness assigned at the grid points together with some interpolation scheme (bilinear, trilinear, spline, etc.) to specify the values between grid points. Of course, as cells/blocks become smaller and smaller (down to infinitesimal), I think of cells/blocks of constant slowness as a special case of continuous models, or of continuous models as a limiting case of cells/blocks. When it is not important which type of slowness model is involved, I refer to the model abstractly as a vector $\mathbf{s}$ in a vector space $\mathcal{S}$. For a block model with n blocks, I have $\mathcal{S} = \mathbf{R}^n$, the n-dimensional Euclidean vector space. A continuous slowness model, on the other hand, is an element of a function space, e.g., the set of continuous functions of three real variables. No matter how I parameterize the model, these models necessarily have far fewer parameters than the actual medium they are intended to represent. Thus, these models are analogous to cartoons, trying to capture the main features with the minimum of detail.

Fermat's principle and traveltime functionals

The traveltime of a seismic wave is the integral of slowness along a ray path connecting the source and receiver. To make this more precise, I define two types of functionals for traveltime. Let P denote an arbitrary path connecting a given source and receiver in a slowness model $\mathbf{s}$. I will refer to P as a *trial ray path*. I define a functional τ^P, which yields the traveltime along

path P. Letting $\mathbf{s}$ be the continuous slowness distribution $s(\mathbf{x})$, I have

$$\tau^P[\mathbf{s}] = \int_P s(\mathbf{x})\, dl^P, \tag{19}$$

where dl^P denotes the infinitesimal distance along the path P. *Fermat's principle* states that the correct ray path between two points is the one of least overall traveltime, *i.e.*, it minimizes $\tau^P(\mathbf{s})$ with respect to path P. Let us define a second traveltime functional τ^* to be the functional that yields the traveltime along the Fermat (least-time) ray path. Fermat's principle then states

$$\tau^*[\mathbf{s}] = \min_{P \in Paths} \tau^P[\mathbf{s}], \tag{20}$$

where *Paths* denotes the set of all continuous paths connecting the given source and receiver. The particular path that produces the minimum in (20) is denoted P^*. If more than one path produces the same minimum traveltime value, then P^* denotes any particular member in this set of minimizing paths. Substituting (19) into (20), I have *Fermat's principle of least time*:

$$\tau^*[\mathbf{s}] = \int_{P^*} s(\mathbf{x})\, dl^{P^*} = \min_P \int_P s(\mathbf{x})\, dl.^P \tag{21}$$

The traveltime functional $\tau^*[\mathbf{s}]$ is well-known to be stationary with respect to small variations in the path $P^*(s)$.

Seismic inversion and tomography

Suppose I have a set of observed traveltimes, $t_1, \ldots, t_m$, from m source-receiver pairs in a medium of slowness $s(\mathbf{x})$. Let P_i be the Fermat ray path connecting the ith source-receiver pair. Neglecting observational errors, I write

$$\int_{P_i} s(\mathbf{x})\, dl^{P_i} = t_i, \quad i = 1, \ldots, m. \tag{22}$$

Given a block slowness model, let l_{ij} be the length of the ith ray path through the jth cell:

$$l_{ij} = \int_{P_i \cap \text{cell}_j} dl^{P_i}. \tag{23}$$

For n cells, Eq. (22) can then be written

$$\sum_{j=1}^{n} l_{ij} s_j = t_i, \quad i = 1, \ldots, m. \tag{24}$$

Note that for any given i, the ray-path lengths l_{ij} are zero for most cells j, as a given ray path will in general intersect only a few of the cells in the model. I rewrite (24) in matrix notation by defining the column vectors $\mathbf{s}$ and $\mathbf{t}$ and the matrix $\mathbf{M}$ as follows: Equation (24) then becomes the basic equation of forward modeling for ray equation analysis:

$$\mathbf{Ms} = \mathbf{t}. \tag{25}$$

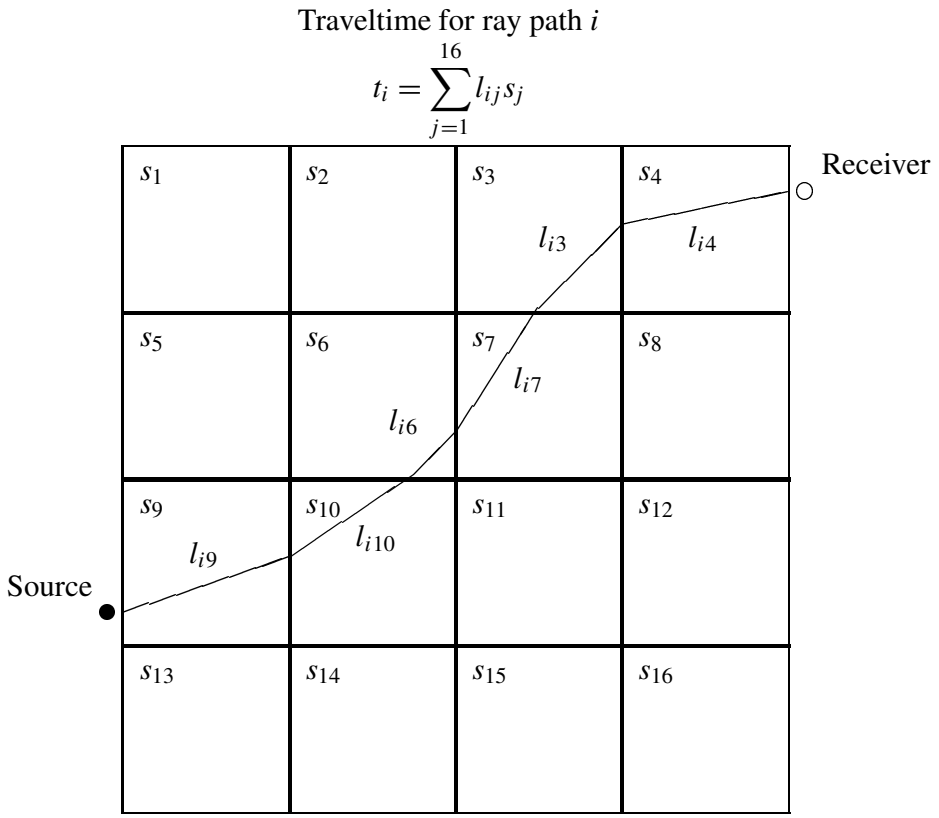

Figure 4: Schematic illustration of ray paths through a cell slowness model.

Note that equation (25) may be viewed as a numerical approximation to equation (21), *i.e.*, it is just a discretized (or finite element) version of the equation. Equation (25) may be used for any set of ray paths, whether those ray paths minimize (21) or not. If the ray paths used to form the matrix **M** are actually minimizing ray paths, **M** is then implicitly a function of the slowness model **s**. The methods developed apply to both two-dimensional and three-dimensional imaging applications.

Feasibility analysis for traveltime inversion

The idea of using feasibility constraints in nonlinear programming problems is well established [Fiacco and McCormick, 1990]. However, comparatively little attention has been given to the fact that physical principles such as Fermat's principle actually lead to rigorous feasibility constraints for nonlinear inversion problems [Berryman, 1991]. The main practical

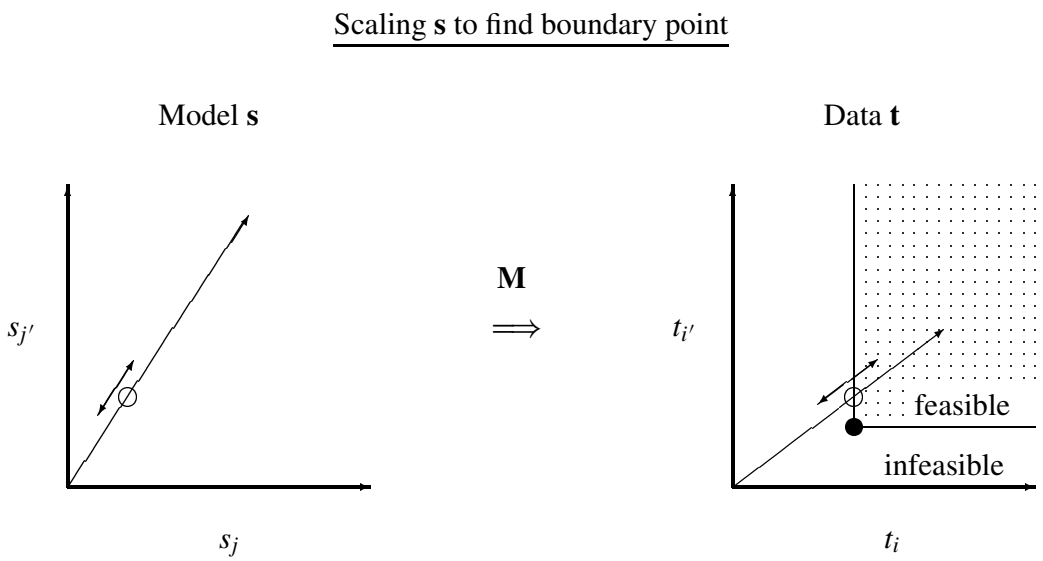

Figure 5: Scaling **s** to find feasibility boundary point. Feasibility boundary in data space is explicitly determined by the data. Feasibility boundary in model space is implicitly determined through the traveltime calculation.

difference between the standard analysis in nonlinear programming and the analysis being developed for nonlinear inversion is that, whereas the functions involved in nonlinear programming are often continuous, differentiable, and relatively easy to compute explicitly, the functionals in nonlinear inversion (*e.g.*, the traveltime functional) need not be continuous or differentiable and, furthermore, are very often rather difficult to compute. Feasibility constraints for such inversion problems are generally implicit, rather than explicit. Equation (22) assumes that P_i is one of Fermat's (least-time) paths and leads to the equalities summarized in the vector-matrix equation $\mathbf{Ms} = \mathbf{t}$. Now suppose instead that P_i is a trial ray path which may or may not be a least-time path. Fermat's principle allows us to write

$$\int_{P_i} s(\mathbf{x}) dl^{P_i} \geq t_i, \tag{26}$$

where now t_i is the measured traveltime for source-receiver pair i. When I discretize (26) for cell or block models and all ray paths i, the resulting set of m inequalities may be written as

$$\mathbf{Ms} \geq \mathbf{t}. \tag{27}$$

Equations (26) and (27) can be interpreted as a set of inequality constraints on the slowness model **s**. When **s** obeys these m constraints, I say that **s** is *feasible*. When any of the constraints is violated, I say **s** is *infeasible*. The set of inequalities collectively are called the *feasibility constraints*. The concept of the feasibility constraint is quite straightforward in nonlinear programming problems [Fiacco and McCormick, 1990] whenever the constraints may be *explicitly* stated for the solution vector. However, in these inversion problems, an additional computation is required. The feasibility constraints are *explicit* for the traveltime data vector,

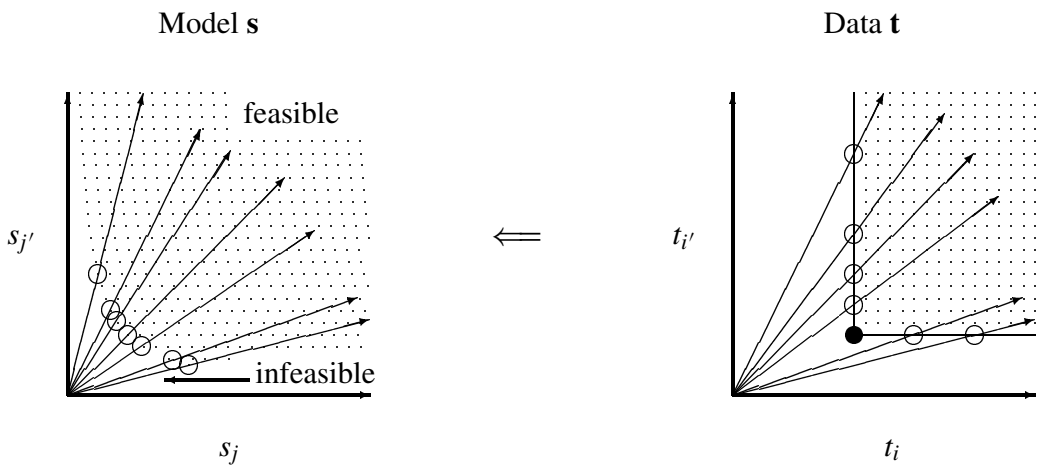

Figure 6: By scaling many slowness vectors **s**, the location of the feasibility boundary in the model space can be mapped.

but they are only *implicit* (*i.e.*, they must be computed) for the slowness vector. This added degree of complication is unavoidable in the inversion problem, but nevertheless it is also very easily handled computationally with only very minor modifications of the usual nonlinear inversion algorithms. Now suppose $\mathbf{s}_1, \mathbf{s}_2$ are two model vectors in the feasible set for a given set of ray paths $\{P\}$. Let $\mathbf{s}_\lambda = \lambda\mathbf{s}_1 + (1-\lambda)\mathbf{s}_2$ be the convex combination of these two vectors, where $0 \leq \lambda \leq 1$. Since, for each fixed ray path i, the traveltime functional τ_i^P is a linear functional of the slowness model, I have

$$\tau_i^P(\mathbf{s}_\lambda) = \lambda\tau_i^P(\mathbf{s}_1) + (1-\lambda)\tau_i^P(\mathbf{s}_2). \tag{28}$$

But $\tau_i^P(\mathbf{s}_1), \tau_i^P(\mathbf{s}_2) \geq t_i$ (by assumption) and λ and $(1-\lambda)$ are non-negative. Therefore,

$$\tau_i^P(\mathbf{s}_\lambda) \geq \lambda t_i + (1-\lambda)t_i = t_i. \tag{29}$$

Thus, the convex combination $\mathbf{s}_\lambda$ lies in the feasible set if $\mathbf{s}_1$ and $\mathbf{s}_2$ do, so it follows again that the feasible set for the traveltime time problem is convex. This shows that for a fixed set of ray paths the local feasibility set is convex. One important difference between the traveltime inversion problem and the string density inversion problem is that whereas the hierarchy of variational functionals used in the Rayleigh-Ritz approach contains inherent nonlinearities due to the necessity of orthogonalizing higher order trial eigenfunctions with respect to lower order ones, no such necessity arises in the traveltime problem. Each traveltime measurement is in a very real sense independent of every other traveltime measurement; all traveltimes are equally important and no hierarchy of traveltimes needs to be established. Although it is certainly possible to use trial ray paths that are highly correlated with the wave speed models under consideration, such correlations are not required by Fermat's principle in the way that they are by the Rayleigh-Ritz method. (For example, we could do the entire tomographic inversion

using straight rays, with some loss of accuracy in the final result.) This fact suggests, but does not prove, that the global feasible set for the traveltime problem is itself convex. Using the fact that the global feasible set must be the intersection of all possible feasible sets, we obtain a proof of convexity [Berryman, 1991].

Algorithms

Stable reconstruction algorithms using feasibility constraints to solve the traveltime tomography problem have been discussed in detail elsewhere [Berryman, 1990]. The key idea is to use the fact that the solution, if one exists, must lie on the feasibility boundary and then to force the solution in an iterative scheme to stay close to this manifold in the model space. An iterative linear least-squares inversion algorithm was easily and cheaply stablized using this approach. Other choices of algorithms based on the feasibility constraints are also possible, as will be discussed at the end of the paper.

CONSTRAINTS FOR EARTH DENSITY DISTRIBUTION

A somewhat different class of problems involves analysis of free oscillations of the Earth in order to deduce its density structure [MacDonald and Ness, 1961; Gilbert, 1971; Jordan and Anderson, 1974; Aki and Richards, 1980; Gilbert, 1980; Ben-Menahem and Singh, 1981; Lapwood and Usami, 1981; Morelli and Dziewonski, 1987; Snieder, 1993]. I concentrate on toroidal modes since they are independent of gravitational effects. For each normal mode of a vibrating elastic medium, the potential energy density is composed of a bulk contribution κB proportional to the bulk modulus κ and a shear contribution μS proportional to the shear modulus. B and S are functionals of the exciting eigenfunctions. If the compressional and shear wave speeds (v_c and v_s, respectively) have been found using (for example) traveltime tomography, then I use the identities

$$v_c^2 = (\kappa + 4\mu/3)/\rho \tag{30}$$

and

$$v_s^2 = \mu/\rho, \tag{31}$$

with ρ being the mass density, to show that the potential energy density may be written in the form

$$\rho U = \rho\left[\left(v_c^2 - 4v_s^2/3\right)B + v_s^2 S\right]. \tag{32}$$

This equation shows explicitly that the potential energy is the sum of terms proportional to the unknown density, the squares of the (assumed) known velocities, and the functionals B and S of the (possibly trial) eigenfunctions. The kinetic energy of a vibrating system is proportional to the density and the square of the time rate of change of the local displacement; so the Fourier transform of this energy density takes the general form $\omega^2 \rho K$, where ω is the angular

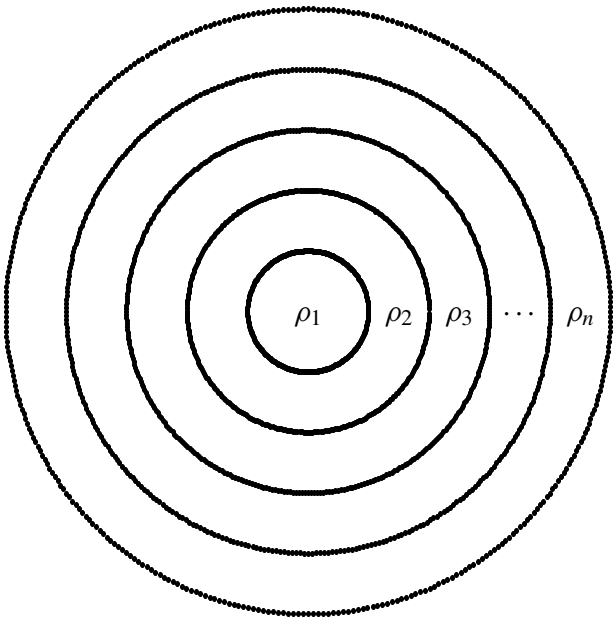

Figure 7: A sphere composed of concentric shells of constant density ρ may be used to model free oscillations of the Earth. Compressional and shear wave velocities are assumed to have been previously determined using traveltime tomography. The inverse problem solves for density structure using measured eigenfrequencies of toroidal oscillation in this example.

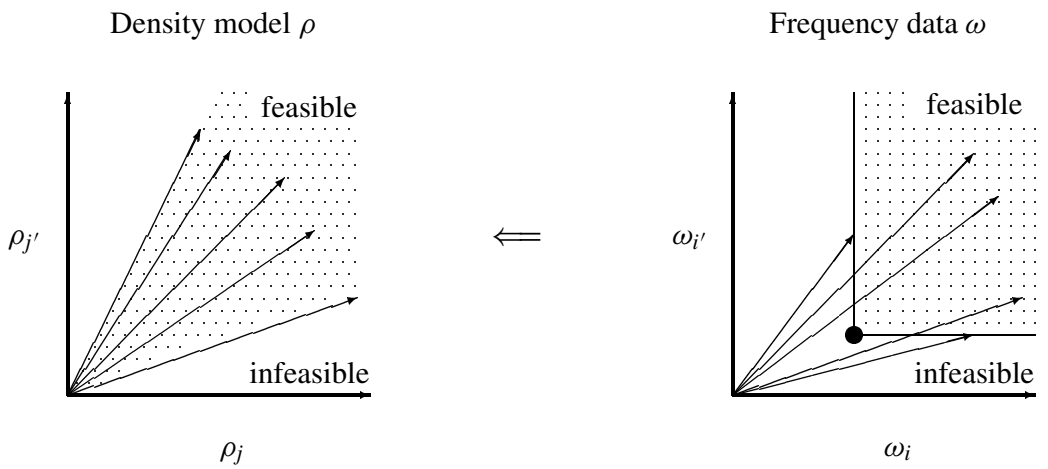

Figure 8: For free oscillations, scaling the density ρ by a positive constant does not change either the trial eigenfunctions or the frequencies, showing that any point along a ray in density space is equally good for satisfying the frequency data. Thus, the feasible region forms a cone in the density space. of the Earth.

frequency of oscillation, and K is the square of the trial displacement. The partitioning of energy in a vibrating medium is known to balance the total kinetic and potential energy in the volume over a cycle. If the volume is Ω, this balance between the two energies may be expressed in the equality

$$\omega^2 \int_\Omega \rho K \, d\Omega = \int_\Omega \rho U \, d\Omega, \tag{33}$$

where again ρ is the local density, K and U are functionals of the eigenfunctions of the normal modes of vibration, and ω is the angular frequency of vibration for a particular mode. It is well-known that equation (33) can be rewritten in the form of a variational inequality as

$$\omega^2 \leq \frac{\int \rho U \, d\Omega}{\int \rho K \, d\Omega}, \tag{34}$$

in which the right hand side is again the Rayleigh-Ritz quotient, and the inequality itself follows from the Rayleigh-Ritz variational principle. In (34), K and U are now functionals of *trial* eigenfunctions, and equality is obtained when the minimum of the Rayleigh-Ritz quotient is achieved, which occurs in variational calculations as the difference between the trial eigenfunction and the true eigenfunction approaches zero. When looking for constraints on the density distribution ρ, I first notice that for this problem the Rayleigh-Ritz quotient *does not place any constraint on the scale* of the density. Multiplying any particular mass distribution ρ by a positive constant γ produces no change in this ratio, since the constant may be moved outside the integrals in both the numerator and the denominator and then cancels. This is an

Constraints to fix the scale of the density

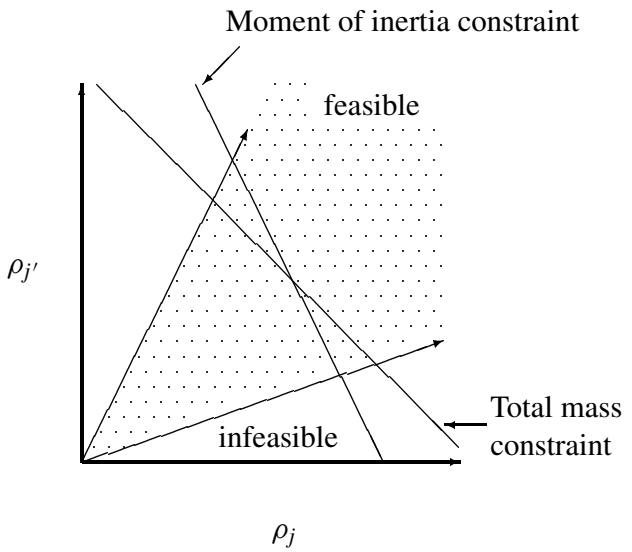

Figure 9: The scale of the density is determined by the total mass and/or the moment of inerita of the Earth.

important difference between the present problem and the ones discussed previously (string vibration and traveltime inversion). The boundary of the feasible set is not found by scaling the density in this problem, and therefore the shape of the feasibility boundary is that of a cone in the model space. However, an independent constraint on the density scale is easily obtained when the total mass M of a body such as the Earth is known, for then

$$\int_\Omega \rho d\Omega = M. \tag{35}$$

Another constraint on the overall scale is provided by the moment of inertia

$$\int_\Omega \rho r^2 d\Omega = I, \tag{36}$$

when it is known. To attack the inverse problem, I again introduce the concept of particular feasible density distributions and collections of these called "feasible sets," as before. I assume that some subset of the modal frequencies has been measured. For simplicity, I display equations for only one frequency and trial eigenfunction, but understand that the same argument applies simultaneously to all eigenfrequencies and trial eigenfunctions. Suppose that ρ_a and ρ_b are two density distributions, both satisfying the Rayleigh-Ritz inequality for the same trial eigenfunction. Thus, I am considering only local or trial feasibility of the densities in this first example. Then, by assumption,

$$\omega_n^2 \int \rho_a K d\Omega \leq \int \rho_a U d\Omega \tag{37}$$

and

$$\omega_n^2 \int \rho_b K \, d\Omega \leq \int \rho_b U \, d\Omega \tag{38}$$

both hold. If I next multiply (37) by λ and (38) by $(1-\lambda)$ and add the two inequalities, I find

$$\omega_n^2 \int \rho_\lambda K \, d\Omega \leq \int \rho_\lambda U \, d\Omega, \tag{39}$$

where $\rho_\lambda = \lambda\rho_a + (1-\lambda)\rho_b$ is defined as the convex combination of the two densities. Thus, by the preceding argument, I have established the feasibility of the convex combination of any two density distributions that are themselves feasible (*i.e.*, ρ_λ also satisfies the Rayleigh-Ritz constraints obtained from the data) for any particular choice of trial eigenfunction. This argument establishes that local feasibility sets are convex. The proof is similar for global feasibility. Suppose I have found the actual eigenfunctions for the densities ρ_a and ρ_b, and therefore know the corresponding potential and kinetic energy densities K_a, U_a, K_b, U_b for these density distributions. These eigenfunctions (by assumption) produce minima in the respective Rayleigh-Ritz quotients; so if the resulting inequalities show that both ρ_a and ρ_b are feasible, then

$$\omega_n^2 \leq \frac{\int \rho_a U_a d\Omega}{\int \rho_a K_a d\Omega} \leq \frac{\int \rho_a U_* d\Omega}{\int \rho_a K_* d\Omega}, \tag{40}$$

and similarly

$$\omega_n^2 \leq \frac{\int \rho_b U_b d\Omega}{\int \rho_b K_b d\Omega} \leq \frac{\int \rho_b U_* d\Omega}{\int \rho_b K_* d\Omega}, \tag{41}$$

where the functionals K_* and U_* have been evaluated using some other convenient trial eigenfunctions (a specific choice will be made later in the discussion). Then, using the same argument as in the last paragraph, I find that the convex combination of densities ρ_λ must satisfy

$$\omega_n^2 \leq \frac{\int \rho_\lambda U_* d\Omega}{\int \rho_\lambda K_* d\Omega} \tag{42}$$

for *any suitably constrained choice of trial eigenfunction*. So, in particular the inequality must hold for the actual eigenfunctions that minimize the Rayleigh-Ritz quotient for (42), *i.e.*, when $U_* = U_\lambda$ and $K_* = K_\lambda$. Thus, it follows from (42) that if ρ_a and ρ_b are globally feasible densities, then so is their convex combination ρ_λ since

$$\omega_n^2 \leq \frac{\int \rho_\lambda U_\lambda d\Omega}{\int \rho_\lambda K_\lambda d\Omega}. \tag{43}$$

This proves that the global feasible set is convex. It is not hard to show that the global feasibility set may also be characterized as the intersection of all local feasibility sets, and is therefore a smaller (but still expected to be nonempty) set than any of the local feasibility sets (for particular trial eigenfunctions). It is harder to prove that the set is nonempty for this problem, than for the previous expamples.

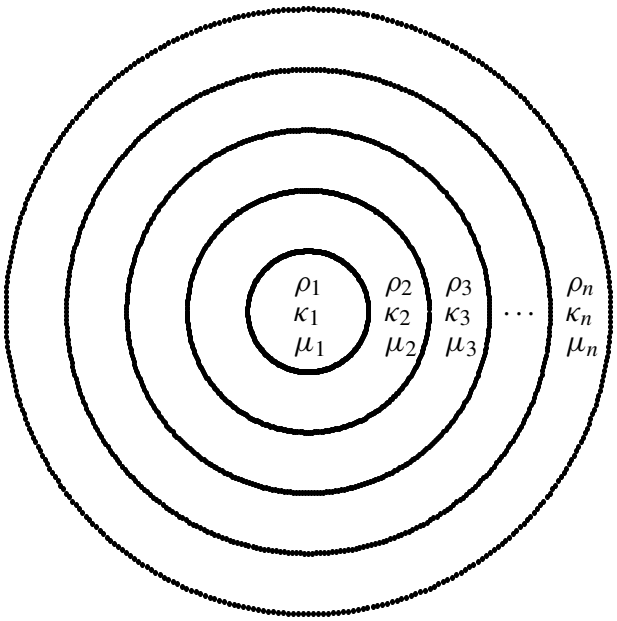

Figure 10: If compressional and shear velocities are not known, then a sphere composed of concentric shells of constant density ρ, bulk modulus κ, and shear modulus μ may be used to model free oscillations of the Earth. As in the last example, model structure is inverted from eigenfrequency data.

CONSTRAINTS FOR ELASTIC MODULI AND DENSITY

Now suppose that the only data available for determining Earth structure is the free oscillation data. What can be said about the corresponding nonlinear inversion problem? It turns out that the analysis goes through essentially as before, but now the model space is larger, involving not only the density distribution ρ but also the bulk modulus κ and the shear modulus μ. The Rayleigh-Ritz inequality is still of the form

$$\omega^2 \leq \frac{\int (\kappa B + \mu S)\, d\Omega}{\int \rho K\, d\Omega}. \tag{44}$$

When two models $(\rho_a, \kappa_a, \mu_a)$ and $(\rho_b, \kappa_b, \mu_b)$ satisfy the feasibility constraints, I have

$$\omega_n^2 \int \rho_a K\, d\Omega \leq \int (\kappa_a B + \mu_a S)\, d\Omega \tag{45}$$

and

$$\omega_n^2 \int \rho_b K\, d\Omega \leq \int (\kappa_b B + \mu_b S)\, d\Omega. \tag{46}$$

And the convex combination

$$(\rho_\lambda, \kappa_\lambda, \mu_\lambda) \equiv (\lambda\rho_a + (1-\lambda)\rho_b, \lambda\kappa_a + (1-\lambda)\kappa_b, \lambda\mu_a + (1-\lambda)\mu_b) \tag{47}$$

clearly satisfies the corresponding condition

$$\omega_n^2 \int \rho_\lambda K\, d\Omega \leq \int (\kappa_\lambda B + \mu_\lambda S)\, d\Omega, \tag{48}$$

where the same trial eigenfunctions have been used in (45)–(48). This again establishes convexity of any local feasibility set and the proof of the convexity of the global feasibility set for (ρ, κ, μ) follows immediately using the same arguments as before. The Rayleigh-Ritz ratio is again scale invariant, so an overall change in scale of the form $(\rho, \kappa, \mu) \rightarrow (\gamma\rho, \gamma\kappa, \gamma\mu)$ does not affect the frequency predictions. But, as long as the scale of any one of the parameters is known, the others follow from it, so knowledge of the total mass (35) and/or total moment of inertia (36) is again sufficient to determine the scale of the model. The main difference between the results for a single parameter ρ and those for a set of parameters such as the triple (ρ, κ, μ) is that, depending on details of how the three parameters are discretized for numerical treatment, the model space will be on the order of three times larger for the case of pure free oscillation data and therefore correspondingly more of this vibration data will be required to obtain enough constraints to produce the same degree of model resolution.

VARIATIONAL STRUCTURE OF INVERSE PROBLEMS

The various inverse problems considered so far may be viewed as special cases of the following general problem. Suppose I have the set of measured data $\{q_i\}$ and that these data are

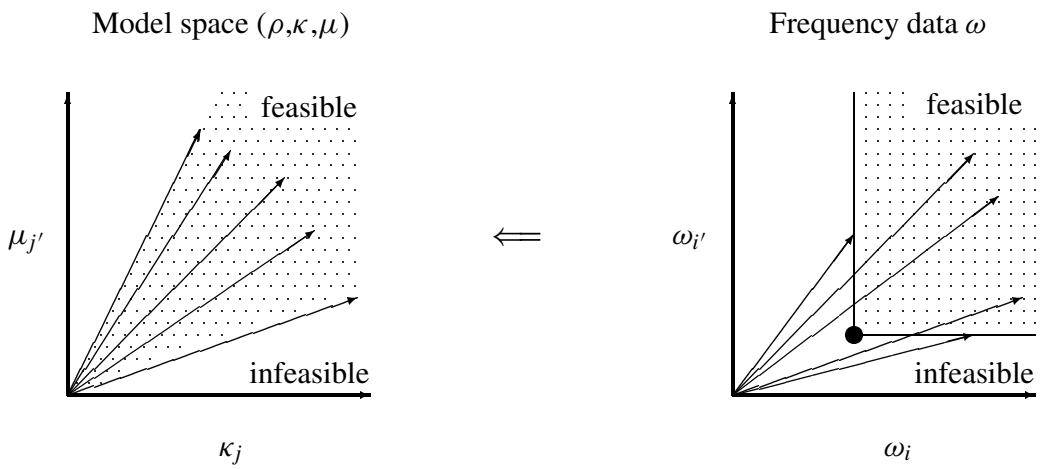

Figure 11: Scaling the triple (ρ,κ,μ) by a positive constant does not change either the trial eigenfunctions or the frequencies, showing again that any point along a ray in the model space is equally good for satisfying the frequency data. The overall scale is again determined by the total mass and/or the moment of inertia of the Earth.

known — from the mathematical physics of the related forward problem — to be minima of an appropriate variational functional so that

$$q_i \le Q_i[n_1,\ldots,n_k,d_1,\ldots,d_m;v_1,\ldots,v_p], \tag{49}$$

where Q_i is a functional whose first $(k+m)$ arguments are parameters of the model (such as density for string vibrations, slowness in traveltime tomography, or bulk and shear moduli in elastic vibration) and whose p remaining arguments are trial eigenfunctions. The dependencies on the trial eigenfuctions will be suppressed in the following discussion. If the functional can be written in the form of a quotient

$$Q_i[n_1,\ldots,n_k,d_1,\ldots,d_m] = \frac{N_i[n_1,\ldots,n_k]}{D_i[d_1,\ldots,d_m]}, \tag{50}$$

where the functionals in the numerator N_i and in the denominator D_i are themselves linear functionals of each of their arguments, then I find that all the examples considered may be written in this form (see Table 1). It follows easily from this postulated form that the inverse problem leads in all cases to a convex feasible region in the multiparameter model space. This follows directly from the observation that, if

$$q_i D_i[\hat{d}_1,\ldots,\hat{d}_m] \le N_i[\hat{n}_1,\ldots,\hat{n}_k] \quad \text{for all} \quad i \tag{51}$$

and if

$$q_i D_i[\tilde{d}_1,\ldots,\tilde{d}_m] \le N_i[\tilde{n}_1,\ldots,\tilde{n}_k] \quad \text{for all} \quad i, \tag{52}$$

then for any λ in the range $0 \le \lambda \le 1$ I find

$$q_i \le \frac{N_i[n_1^{(\lambda)},\ldots,n_k^{(\lambda)}]}{D_i[d_{(\lambda)},\ldots,d_m^{(\lambda)}]} \quad \text{for all} \quad i, \tag{53}$$

where

$$n_j^{(\lambda)} = \lambda\hat{n}_j + (1-\lambda)\tilde{n}_j \quad \text{for} \quad 1 \le j \le k, \tag{54}$$

and

$$d_l^{(\lambda)} = \lambda\hat{d}_l + (1-\lambda)\tilde{d}_l \quad \text{for} \quad 1 \le l \le m, \tag{55}$$

guaranteeing that, if the "hat" and "tilde" models are themselves feasible for the given set of trial eigenfunctions, then the convex combinations of the "hat" and "tilde" models parametrized by λ are also feasible. The argument just given depends strongly on *the implicit assumption that the conditions for admissibility of eigenfunctions are independent of the convex combination parametrized by* λ. Such independence is definitely true for the traveltime inversion problem where the ray paths play the role of eigenfunctions, but also definitely not true for the string density inversion problem or for the free oscillations problem. Thus, whether or not the global feasible set is convex depends on the problem, and in particular on the variational principle being used in the analysis. Rayleigh-Ritz variational problems will not have convex feasibility sets, while traveltime tomography and some other similar inversion problems [Berryman, 1991] in general do have convex feasibility sets.

TABLE 1.

Problem	$\{q_i\}$	$\{n_i\}$	$\{d_i\}$	Description of feasible set
Density from string vibration	$\{\omega_i\}$	—	$d_1 = \rho$	Finite, nonconvex, bounded region near the origin
Wave slowness from traveltime tomography	$\{\tau_i\}$	$n_1 = s$	—	Infinite, convex, bounded away from the origin
Elastic constants from free oscillations	$\{\omega_i\}$	$n_1 = \kappa$ $n_2 = \mu$	$d_1 = \rho$	Infinite cone, nonconvex

Although convexity of the feasible set is not really necessary when designing a nonlinear programming method for solving the inverse problem, convexity does help to simplify the reconstruction algorithms when it is a characteristic of the problem of interest.

DISCUSSION

Finite data sets

Suppose I have a finite set of measurements $\{q_1,\ldots,q_M\}$ and I compute the feasible set $\mathcal{F}_M$ associated with that measurement set. Now suppose that an additional measurement is made

so the measurement set is now $\{q_1,\ldots,q_{M+1}\}$ with the associated feasible set $\mathcal{F}_{M+1}$. What is the relationship between $\mathcal{F}_M$ and $\mathcal{F}_{M+1}$? The addition of a constraint can only decrease the size of the feasible set, so it follows easily from the convexity properties that in general $\mathcal{F}_{M+1} \subseteq \mathcal{F}_M \subseteq \mathcal{F}_{M-1} \subseteq \cdots \subseteq \mathcal{F}_1$.

Measurement error

An issue that is often raised about the usefulness of the feasibility constraints concerns the effects of measurement errors on the location of the feasibility boundary. The variational constraint equations always find the data in the role of upper or lower bounds on integrals involving the unknown parameters. The data therefore enters these constraints linearly, so small measurement errors will generally lead to small (or possibly no) changes in the location of the bounding curves, depending on which measurements are in error and which measurements are the most constraining. I think of the feasibility boundary in these circumstances, not as a sharp but rather, as a fuzzy boundary. If the errors are small, then the fuzzy region is also small. It is useful to take this fuzziness into account in practical algorithms that make use of the feasibility boundary to reconstruct the desired model parameters in these inverse problems. This can be accomplished by using either the least (or the most) constraining range of data-minus-error (or data-plus-error) when computing this estimate of the boundary location. Alternatively, the fuzziness of the boundary can be incorporated directly into algorithms that use only the approximate location of the feasibility boundary as a means of constraining a search for the desired model parameters based on other criteria (such as minimizing the least-squares error in predicted versus measured data, or some other choices of objective functional minimization). Experience has shown that practical algorithms based on feasibility constraints are less sensitive (more robust) to the presence of measurement errors than most other algorithms for inversion.

Algorithms

The variational structure of nonlinear inverse problems does not by itself provide an algorithm for reconstructing the desired model parameters from the data. Nevertheless, knowing the existence of the feasibility boundary and the convexity properties of the feasible set suggests a number of possible algorithms that may be useful depending on the particular problem. I may suggest three types of algorithms: (1) The most obvious and probably the least practical algorithm entails a search along the feasibility boundary for the model that best fits the data. This approach has much in common with linear programming methods of the simplex type, but is probably not practical for most of the large dimensional problems that would benefit from the methods discussed in this paper. The reason for the difference is that the feasibility boundary in nonlinear inversion problems is determined only implicitly and therefore requires considerable computation to find each boundary point, whereas in linear and most other nonlinear programming problems the boundary constraints are given explicitly. (2) A Monte Carlo or shooting method that tries to sample a region of the model space and map the feasibility boundary in that local region has been tried on both the string density inversion problem and

on the traveltime inversion problem. This approach has been found to work extremely well in the problems of lower dimensionality, and it also works well in higher dimensional problems that are easily parallelized (such as traveltime tomography) when many processors are available. Finally, (3) virtually any existing inversion algorithm can be easily modified to incorporate the feasibility constraints as a natural means of regularizing, *i.e.*, preventing divergences from occurring. In particular, iterative linear least-squares inversion algorithms that might diverge due to inconsistencies arising from the forward modeling (trial eigenfuctions or trial ray paths) based on a previous guess of the model can be easily stabilized by forcing the stepsize for model updates to remain small enough so that the successive iterates do not wander away from the feasibility boundary. Such a constraint does not tie the result to any particular part of the model space that must be chosen prior to the inversion (*i.e.*, hard constraints on the maximum and minimum values of the parameters in the model are not needed), but rather tethers the final result to a manifold determined strictly by the data and the measurement configuration. Thus, the data itself is used to determine the appropriate means of regularizing the solution to the inverse problem in such algorithms. This approach is probably the one that will find the most use in practical solutions to inverse problems.

ACKNOWLEDGMENTS

Numerical experiments for the string vibration problem were performed by O. S. Tai, a summer research student visiting at LLNL. I thank J. R. McLaughlin for insight into the general solution of the string vibration problem. I also thank G. C. Beroza, J. M. Rice, and G. Zandt for helpful leads into the literature on free oscillations of the Earth.

REFERENCES

Aki, K., and Richards, P. G., 1980, *Quantitative Seismology: Theory and Methods*, Vol. I, Freeman, New York, Chapter 8, pp. 337–381.

Ben-Menahem, A., and Singh, S. J., 1981, *Seismic Waves and Sources*, Springer-Verlag, New York, Chapter 6, pp. 337–419.

Berryman, J. G., 1990, Stable iterative reconstruction algorithm for nonlinear traveltime tomography: *Inverse Problems* **6**, 21–42.

Berryman, J. G., 1991, Convexity properties of inverse problems with variational constraints: *J. Franklin Inst.* **328**, 1–13.

Coddington, E. A., and Levinson, N., 1955, *Theory of Differential Equations*, McGraw-Hill, New York, pp. 208–221.

Courant, R., and Hilbert, D., 1953, *Methods of Mathematical Physics*, Vol. I, Interscience, New York, pp. 132–134.

Dines, K. A., and Lytle, R. J., 1979, Computerized geophysical tomography: *Proc. IEEE* **67**,1065–1073.

Fiacco, A. V., and McCormick, G. P., 1990, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, SIAM, Philadelphia, Chapter 6, pp. 86–112.

Gilbert, F., 1971, Inverse problems for the Earth's normal modes: in *Mathematical Problems in the Geophysical Sciences, 2*, *Lectures in Applied Mathematics*, vol. 14, edited by W. H. Reid (American Mathematical Society), pp. 107–121

Gilbert, F., 1980, An introduction to low-frequency seismology: in *Proceedings of the International School of Physics "Enrico Fermi"*, Course LXXVIII, edited by A. M. Dziewonski and E. Boschi, North-Holland, Amsterdam, pp. 41–81.

Jordan, T. H., and Anderson, D. L., 1974, Earth structure from free oscillations and travel times: *Geophys. J. R. Astr. Soc.* **36**, 411–459.

Lapwood, E. R., and Usami, T., 1981, *Free Oscillations of the Earth*, Cambridge University Press, London.

Lytle, R. J., and Dines, K. A., 1980, Iterative ray tracing between boreholes for underground image reconstruction: *IEEE Trans. Geosci. Remote Sens.* **18**, 234–240.

MacDonald, G. J. F., and Ness, N., 1961, A study of the free oscillations of the earth: *J. Geophys. Res.* **66**, 1865–1911.

Morelli, A., and Dziewonski, A. M., 1987, The harmonic expansion approach to the retrieval of deep Earth structure: in *Seismic Tomography: With Applications in Global Seismology and Exploration Geophysics*, edited by G. Nolet, Reidel, Dordrecht, Chapter 11, pp. 251–274.

Snieder, R., 1993, Global inversions using normal modes and long-period surface waves: in *Seismic Tomography: Theory and Practice*, edited by H. M. Iyer and K. Hirahara, Chapman and Hall, London, 1993, Chapter 3, pp. 23–63.