Variational Structure of Inverse Problems in Wave Propagation and Vibration

James G. Berryman
Lawrence Livermore National Laboratory
P. O. Box 808 L-202
Livermore, CA 94551-9900

Version date: July 27, 1995

Abstract

Algorithms for solving inverse problems arising in practice 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 compact or (ideally) convex, 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 of the constraints provided by the data. One example of an inverse problem in wave propagation (traveltime tomography) and two examples in vibration (the plucked string and free oscillations of the Earth) are presented to illustrate how the variational structure of these problems may (or may not) be used to create nonlinear programs using implicit variational constraints. A detailed analysis of the string density inversion problem shows that the feasibility set constructed with data consisting of two or more eigenfrequencies is nonconvex, but the solutions (which are nonunique) nevertheless lie on the feasibility boundary.

1 Introduction

Predicting the dynamical behavior of physical systems has become an elegant part of physics and applied mathematics through the efforts of many researchers trying to find descriptions of such systems that identify constants of motion for each system and means for altering the values of those constants when the system is undergoing some change. (See for example Lanczos [1970] and/or Goldstein [1980].) Elementary systems such as pairs of billiard balls have correspondingly elementary constants of motion such as momentum, angular momentum, and energy, which get altered by common means such as forces, torques, and friction (causing dissipation of energy into heat). More complex systems require more abstract notions of what is meant by "a constant of the motion."

For vibrating systems, the apparently very complex behavior of an excited string, drumhead, or the Earth can begin to be understood by first conceiving the notion of a mode of vibration, for which the total energy is conserved during a cycle — even though that energy changes from kinetic to potential within a cycle and back again from one cycle to the next. The first concept of the mode is that it stands alone and does not interact with other modes. Therefore, each mode is itself a constant of the motion — analogous to a fixed trajectory of a particle, or to a simple closed orbit of a satellite. The time period of a cycle is an important measurable characteristic for each mode of vibration. A plucked string may exhibit an extremely complex oscillatory behavior involving a large (or even infinite) number of modes excited simultaneously, but by clever devices (such as the use of strobe lights tuned to flash at one of the characteristic frequencies of the string vibration) it is possible to isolate individual modes and see that an analysis of the motion into Fourier components is not just mathematical fantasy but actually observable in a laboratory setting.

Nevertheless, these modes are in fact abstract features of the system that are virtually never (quite) observed in practice, because real systems interact with their surroundings — vibrating strings and drumheads lose energy each cycle, both to the air and to the tuning screws providing the tension that both enables the vibration and determines the pitches or frequencies of oscillation. To see the modes of oscillation in a strobe light experiment, we normally cheat a little and continue to drive the vibrating system, so that energy lost to dissipation is replaced by mechanical means. The abstract but robust notion of the existence of constants of the motion and/or the existence of modes of vibration proves to be much more important to our understanding of such systems than are the small errors introduced by neglecting energy loss mechanisms or (in this and many other problems of physical interest) the fact that the linear equations we use to describe the motion are only approximations to more correct nonlinear ones.

And what (you may ask) does this have to do with either inverse problems or variational principles? The answer is: Everything! We will expand on this theme, but briefly: Parameters of system behavior often uniquely determine the constants of motion (read: forward problem/variational solution). Conversely, observations of constants of the motion are expected to constrain the range of parameters of the system (read: inverse problem/unknown solution).

What is a variational principle? A variational principle is a mathematical method used to determine the expected state of some physical system when that system is subjected to a known set of boundary conditions and when that system is also known to obey a system of equations (often characterized by a functional of system descriptors) for which the behavior may be characterized as a minimum, maximum, or stationary point. For example, the frequencies of vibration of a string may be characterized as the minima of the Rayleigh-Ritz quotient, higher frequencies being subject to progressively more constraints on the eigenfunctions. Thus, in applications of variational methods, we know the form of the equations of motion, and even know (or think we know) the coefficients for those equations, but we need to determine the constants of the motion. Variational principles have traditionally been used to solve forward problems to predict system behavior using known system parameters.

What is an inverse problem? An inverse problem is usually associated with a physical system in which measurements have been made of system constants (such as the frequencies of vibration for a string oscillation problem) or possibly complex combinations of such constants and it is desired to deduce more precisely the nature of the physical system (*e.g.*, the density distribution along the string). Thus, for inverse problems, we often know (or think we know) the form of the equations and some of the constants of the motion, but not the coefficients of the equations. The solution to an inverse problem is generally obtained by deducing system parameters from observations of system behavior.

Why should variational principles and inverse problems be related? Just as any other forward method of solution is expected to be related to its inverse, variational principles should have much to tell us about inverse problems — especially so, since the constants of motion are often the extrema of the variational functionals. In forward problems predicting system behavior, variational methods often lead to robust methods for numerical solution. In inverse problems deducing system parameters, variational methods may therefore be expected to lead to robust methods as well.

In this paper, we want to take a broad view, considering several different types of inverse problems and the relationships between these and the corresponding variational principles. We consider three classes of problems: (1) vibration of a one-dimensional string, (2) sound wave propagation through a two- or three-dimensional Earth, and (3) free oscillations of a spherically symmetric, but otherwise three-dimensional, Earth. These three examples appear to cover the full range of possible relations between inverse problems and variational methods of analysis

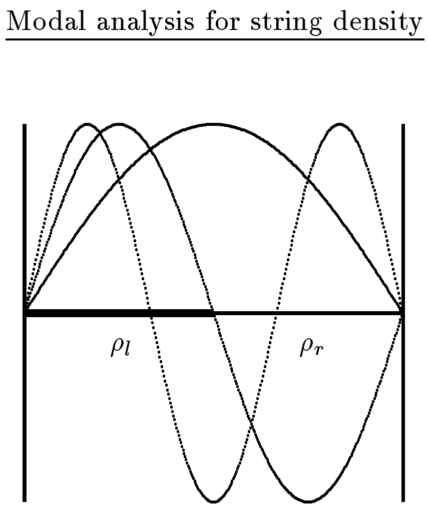

Figure 1: Density distribution of the string may be determined by analyzing the eigenfrequencies associated with its modes of vibration. The left half of the string has density ρ_l, and the right half ρ_r.

for the solution of those problems.

2 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 $u(0,t) = 0 = u(1,t)$. We assume that the string's temporal motion may be decomposed into its Fourier components so that we may study standing waves on the plucked string. The time dependence is assumed to separate into 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, equation (1) reduces to

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

where we have 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.

We 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 only knowledge of the frequencies of vibration of the string. In particular, we 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 $T = 1$, we 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), we 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, for $n = 1$, 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 $v_1(0) = 0 = v_1(1)$ and 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 is often 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 we impose constraints on the trial function such as requiring $v_n(x)$ to have $n-1$ interior nodes [Coddington and Levinson, 1955] plus orthogonality to lower order eigenfunctions, 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 (*i.e.*, the ω_ns) are those defined in (3) for the higher order modes. We refer to the equations (4) and (5) as the "forward problem," since in these equations we 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 $\rho(x)$.

2.1 Feasibility analysis

There are two tricks that make it possible in some cases to use the variational functionals to analyze inversion problems. In the context of the string problem, the first trick is a result of what we call the "scale invariance" property of the eigenfunctions. 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 directly from (2) or from (5) since

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

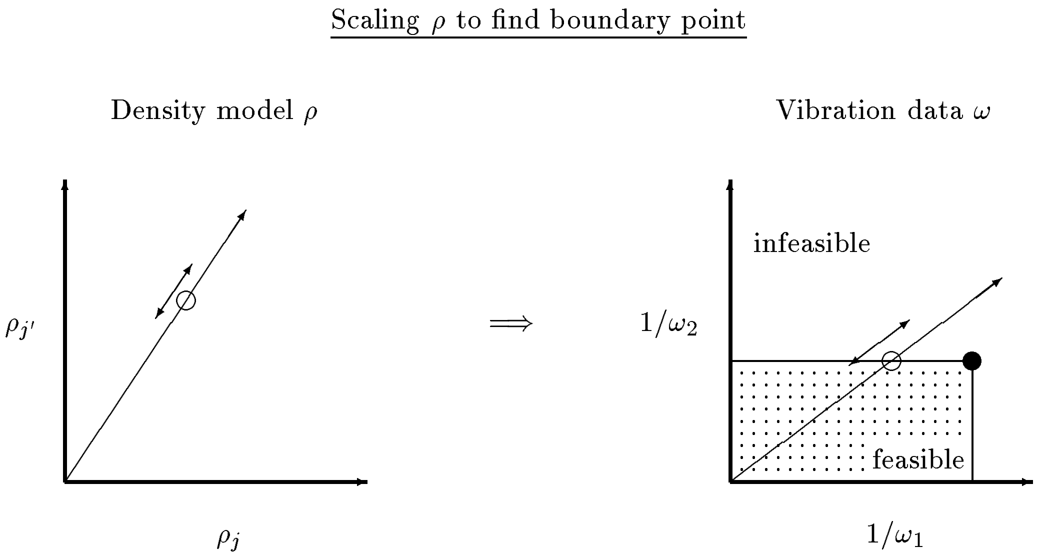

Figure 2: Feasible part of the density space is determined implicitly by the explicit boundaries defined by the frequency data.

which implies that $\omega(\gamma\rho) = \omega(\rho)/\sqrt{\gamma}$. Equation (6) is a statement of the homogeneity property of the Rayleight-Ritz functional in the density scale factor. We can 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. We can take advantage of this linearity in ρ with greater ease by working with the reciprocal of $\mathcal{R}$, so 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 we have made measurements of some of the eigenfrequencies of the string, we can ask the following question: Are there density distributions that violate the inequality in (7)? That is, if we try to compute the left hand side of (7) using an arbitrary trial density distribution $\rho(x)$, when will the inequality be satisfied and when will it be violated? We use the answers to these questions to define two distinct classes of trial model density distributions $\rho(x)$:

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

However,

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

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 in some problems [Berryman, 1991] that universal (global) feasible and infeasible sets exist that are independent of the particular trial functions chosen (dependence on the particular subset of 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 we 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, we derive a very useful property of the class of feasible density distributions. For this derivation, we neglect the correlations that arise between trial eigenfunctions for eigenfrequencies higher than the first, induced by required orthogonality conditions. Thus, the following argument is rigorous only for the first (lowest) eigenfrequency. Consider two feasible trial densities $\rho_a(x)$ an $\rho_b(x)$, which therefore (by definition) satisfy

$$\frac{1}{\mathcal{R}[\rho_a, v]} \le \omega^{-2}(\rho) \quad \text{and} \quad \frac{1}{\mathcal{R}[\rho_b, v]} \le \omega^{-2}(\rho), \tag{10}$$

for the same trial eigenfunction v, with ρ being the actual density of the string when $\omega(\rho)$ was measured. Let λ be a number in the range $0 \le \lambda \le 1$. Then, taking a linear combination of the two expressions in (10) gives

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

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

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

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

$$\rho_\lambda(x) \equiv \lambda\rho_a(x) + (1-\lambda)\rho_b(x). \tag{13}$$

Combining (11) and (12) shows that

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

But any trial function that satisfies (14) for all measured frequencies is by definition a member of the feasible set. So, if $\rho_a(x)$ and $\rho_b(x)$ are both feasible, their convex combination $\rho_\lambda(x)$ is also feasible.

The definition of a convex set is this: a set such that the convex combination — *i.e.*, $\rho_\lambda = \lambda\rho_a + (1-\lambda)\rho_b$ — of any two members is also a member of the set. This fact implies that *the set of all feasible density distributions based on the lowest eigenfrequency data is a convex set* (see Fig. 3). 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 we have any trial density distribution ρ_w and we have found the eigenfunctions $w_n(x)$ and eigenvalues $\omega_n(\rho_w)$ associated with this distribution. Then, the eigenfunctions $w_n(x)$ are also the eigenfunctions for all densities of the form $\gamma\rho_w(x)$, where γ is an arbitrary positive scalar. It is not hard to show [Berryman, 1991] that there always exists a choice of $\gamma = \gamma_w$ such that $\gamma_w\rho_w(x)$ lies exactly on the boundary of the convex feasible set for the inversion problem. It follows then that, if $\gamma \leq \gamma_w$, the density $\gamma\rho_w$ lies in the feasible part of the model space; while, if $\gamma > \gamma_w$, the density $\gamma\rho_w$ lies in the infeasible part. This characteristic of the feasible set allows us to produce a simple geometrical interpretation of the feasible set as we will show.

In particular, for the vibrating string inversion problem, *the feasible set of model densities is compact and occupies a region in the neighborhood of the origin of the model space* [*i.e.*, near $\rho(x) \equiv 0$]. We can easily prove this statement. First note a physical density must be nonnegative, so the general model space is the convex set of all possible nonnegative density distributions. Then, by taking the trial eigenfunction to be $v_1(x) = \sin \pi x$ with the trial density $\rho(x) = \rho_0$ (constant), and noting that

$$\frac{1}{\mathcal{R}[\rho_0, v_1]} = \frac{\rho_0 \int \sin^2 \pi x \, dx}{\pi^2 \int \cos^2 \pi x \, dx} = \frac{\rho_0}{\pi^2} \tag{15}$$

satisfies the feasibility constraint if

$$\rho_0 \leq \pi^2 \omega_1^{-2}(\rho). \tag{16}$$

The first measured eigenfrequency ω_1 then determines an upper bound on the value of ρ_0 for constant density models. For nonconstant models such as the two component string with $\rho(x) = \rho_l$ for $0 \leq x < \frac{1}{2}$ and $\rho(x) = \rho_r$ for $\frac{1}{2} < x \leq 1$, the same argument shows that

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

Thus, (17) shows that ρ_r is bounded above by the straight line

$$\rho_r \leq 2\pi^2 \omega^{-2}(\rho) - \rho_l, \tag{18}$$

producing tighter bounds on ρ_r as ρ_l increases. When $\rho_l = \rho_r$, this problem reduces to the constant model result found earlier.

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, admissible trial functions 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. The consequences of these correlations will be emphasized when the general analysis of the string density is presented for a particular example later in this section. 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.

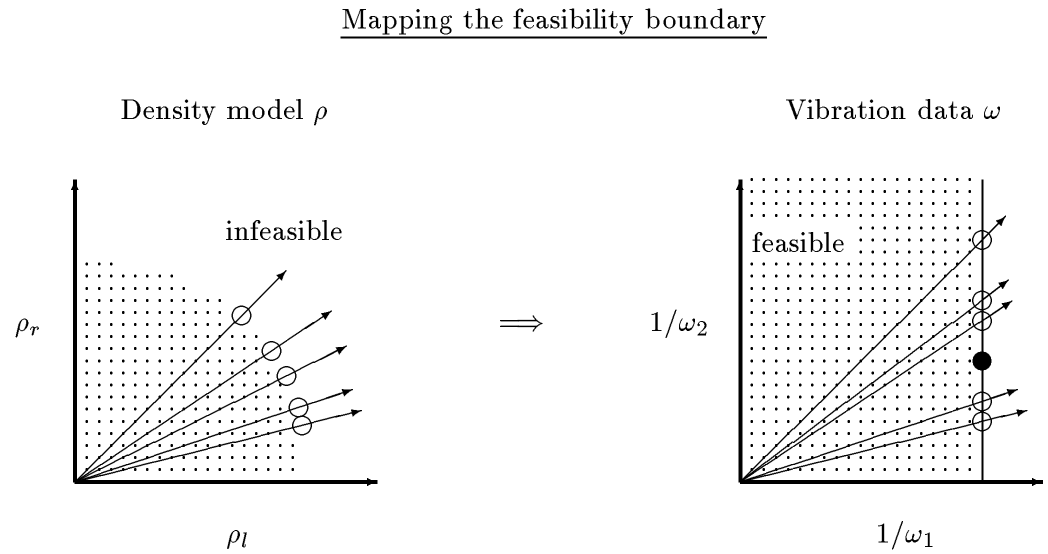

Figure 3: By scaling many density distributions, the location of the feasibility boundary can be mapped. If the only information available is the lowest eigenfrequency ω_1, then (and only then) the feasible region for string vibration is a convex set near the origin, as shown on the left.

2.2 Algorithm

Now we can 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 (ρ_1, ρ_2) solves the problem then so does (ρ_2, ρ_1). (We are presently considering only the frequencies as data. This nonuniqueness of the solution can be easily removed by considering the eigenfunctions as data as well.) Thus, we only need to consider half of the positive quadrant, say for models such that $\rho_1 \geq \rho_2$. Then, picking models evenly spaced in angle for (say) 10 angles up to $\theta = \pi/4$, so $(\rho_1, \rho_2) = (\rho_0 \cos\theta, \rho_0 \sin\theta)$ with $\theta = \pi/40, \pi/20, \ldots, \pi/4$. (The precise value of ρ_0 does not matter, since these points will eventually be scaled onto the feasibility boundary.) Having chosen a set of initial points, we now do forward modeling on each of these string models. After scaling to the feasibility boundary, we check satisfaction of the data and pick the two adjacent points that best satisfy the data. Then, we divide the angular region bracketed by these two angles into 10 smaller angles, and repeat this process until we have a satisfactory solution. The algorithm just described is basically what is known as 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 apparent when we consider complex models involving three or more

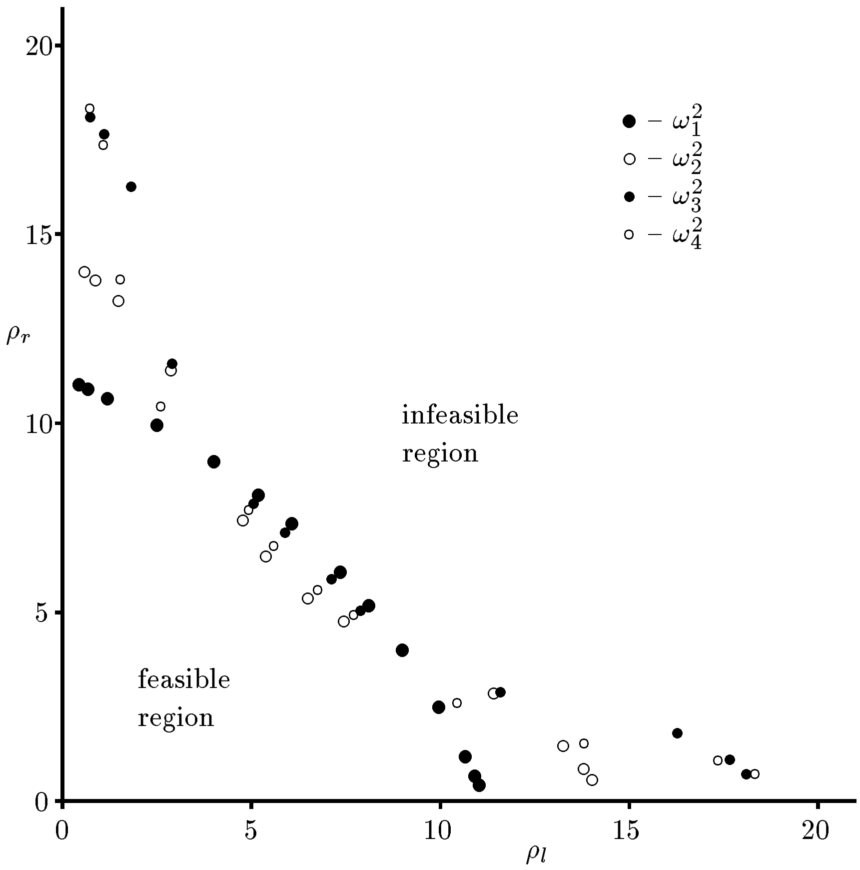

Figure 4: Using the four measurements of eigenfrequencies $\omega_1^2, \ldots, \omega_4^2$ and exact calculations of the eigenfrequencies for various density ratios in TABLE 1, we obtain this plot through scaling of the densities for fixed density ratios. The region near the origin is feasible, while the exterior region is infeasible.

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.

2.3 String example

To provide a definite example of the general ideas presented so far, consider the string density inversion problem and suppose that we know that the string has two segments of equal length and of unequal but constant mass. A measurement of the first four frequencies of oscillation is made and the results are found to be:

$$\omega_1^2 = 1.467$$

$$\omega_2^2 = 6.712$$
$$\omega_3^2 = 13.634$$
$$\omega_4^2 = 25.727$$

To use these data to produce a set of feasibility constraints, we need information about the forward problem, *i.e.*, relationships between densities and frequencies. To obtain these, we can solve this nonuniform string problem exactly for a few particular cases. First, we use a standard transformation (see Coddington and Levinson [1955]) of this one-dimensional problem:

$$u(x) = A(x)\sin\theta(x) \quad \text{and} \quad u_x(x) = \omega A(x)\cos\theta(x), \tag{19}$$

with the conditions that $\theta(0) = 0$ and that $\theta(1) = n\pi$, so that $u(0) = u(1) = 0$ as required. Substituting (19) into (2), we find that

$$\theta_x(x) = \omega\left[\cos^2\theta + \rho(x)\sin^2\theta\right], \tag{20}$$

which is a nonlinear equation for θ that can be integrated easily when the density is piecewise constant.

We use the scale invariance property to simplify the analysis of the eigenvalues. We need to compute the eigenvalues for only one point along a given ray in the model space (ρ_l,ρ_r). The analysis is greatly simplified by choosing models such that

$$\begin{aligned} \rho(x) &= \rho_l = 1 && \text{for} \quad 0 \le x < \tfrac{1}{2} \\ \rho(x) &= \rho_r = 1/s && \text{for} \quad \tfrac{1}{2} < x \le 1. \end{aligned} \tag{21}$$

Then, (20) shows that the general expression for $\theta_n(x)$ for the nth mode is

$$\int_0^{\theta_n(x)} \frac{d\theta}{\cos^2\theta + \rho\sin^2\theta} = \omega_n x. \tag{22}$$

The indefinite integral over an interval in which the density is constant is given by

$$\int \frac{d\theta}{\cos^2\theta + \rho\sin^2\theta} = \frac{1}{\sqrt{\rho}}\tan^{-1}(\sqrt{\rho}\tan\theta). \tag{23}$$

Because of our choice of scaling, we find easily that the integral over the first half of the string gives

$$\theta_n(\tfrac{1}{2}) = \frac{\omega_n}{2}, \tag{24}$$

which provides a relationship between θ_n and the characteristic frequency ω_n. The integral over the second half of the string gives the condition

$$\sqrt{s}\tan^{-1}\left(\frac{\tan\theta_n(\frac{1}{2})}{\sqrt{s}}\right) = \frac{\omega_n}{2}, \tag{25}$$

or equivalently [using(24)] that

$$\tan\left(\frac{\theta_n(\frac{1}{2})}{\sqrt{s}}\right) + \frac{1}{\sqrt{s}}\tan\theta_n(\tfrac{1}{2}) = 0. \tag{26}$$

Equation (26) determines the eigenfrequencies, since for each value of n we have

$$\omega_n^2 = 4\theta_n^2(\tfrac{1}{2}). \tag{27}$$

We use (26) to compute various examples of density/eigenfrequency data. A selection of these is displayed in TABLE 1.

TABLE 1. Eigenvalues ω_n^2 as a function of density ratio s.

n	$s = 1$	$s = \frac{100}{121}$	$s = \frac{16}{25}$	$s = \frac{4}{9}$	$s = \frac{1}{4}$	$s = \frac{1}{9}$
1	39.478	8.912	7.613	5.866	3.651	1.737
2	157.914	35.969	31.913	26.848	19.119	9.870
3	355.306	80.216	68.765	54.537	39.478	24.652
4	631.655	143.846	126.963	102.908	67.139	39.478

The points plotted in Fig. 4 are obtained from TABLE 1 by scaling the density $(1,1/s)$ so that the corresponding eigenfrequency agrees with the measured data for that eigenfrequency. The resulting points determine sets of trial feasibility points along each ray from the origin in the model space (ρ_l,ρ_r). The point nearest the origin along any ray determines the absolute feasibility boundary along that ray for the string density inversion problem. The result of this process is illustrated in Fig. 5. Using the data from Fig. 4 to construct Fig. 5, we find the boundary of the feasible region is determined in this problem by the data for ω_1^2 for large density ratios, and by ω_2^2 for density ratios closer to unity. All four sets of data cross at the two solutions to the inverse problem, which are seen to be two points in the model plane: (4,9) and (9,4). It is not possible to choose between these two solutions using frequency data alone, but knowledge of any one of the eigenfunctions would be sufficient to resolve the ambiguity.

For this relatively simple inverse problem, the solution could have been found by inspection of the feasibility data for any two of the eigenfrequencies. (More precisely, "any two" should be replaced by "almost any two," because in some cases these pairs of curves cross in more than two places.) The corresponding curves all cross at the points (4,9) and (9,4), which gives the pair of solutions immediately. However, with higher dimensional problems (more than two segments along the string), and with modest errors in the frequency data, this type of eyeball analysis would rapidly become difficult or impossible, and some search routine would have to be introduced.

3 Traveltime Inversion Problem

A typical problem arising in seismic traveltime inversion in 2- and 3-dimensional heterogeneous media is this: Infer the (isotropic) compressional-wave slowness (reciprocal of velocity) dis-

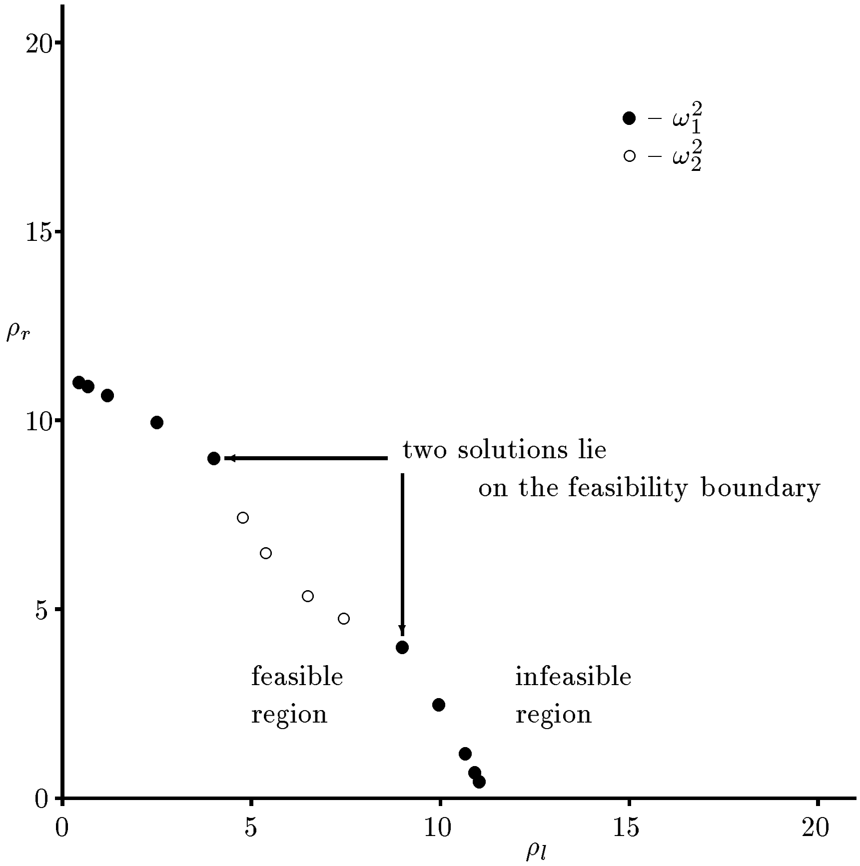

Figure 5: Using the data from Fig. 4, we find the boundary of the feasible region is determined in this problem by the data for ω_1^2 for large density ratios, and by ω_2^2 for density ratios closer to unity. These two sets of data cross at the two solutions to the inverse problem, which are seen to be two points in the model plane: (4,9) and (9,4). It is not possible to choose between these two solutions using frequency data alone.

tribution 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 [Rector, 1995]. We could also consider the problem of inverting for wave slowness when the absolute traveltimes and locations of the sources must also be inferred from the data, as is normally the case in earthquake seismology for whole Earth structure [Aki and Richards, 1980].

3.1 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 have an explicit linear dependence on slowness. The true ray paths are nevertheless implicit functions of slowness, which again makes the full forward modeling problem nonlinear.

We might consider three kinds of slowness models. We could allow the slowness to be a general function $s(\mathbf{r})$ of the position $\mathbf{r}$. However, we 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), we can 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, we refer to the model abstractly as a vector $\mathbf{s}$ in a vector space $\mathcal{S}$. For a block model with n blocks, we 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 we parameterize the model in practice, our models necessarily have far fewer parameters than the actual medium they are intended to represent. Thus, our models are analogous to cartoons, trying to capture the main features with the minimum of detail.

3.2 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, we will define two types of functionals for traveltime.

Let P denote an arbitrary path connecting a given source and receiver in a slowness model. We will refer to P as a *trial ray path*. We define a functional τ^P, which yields the traveltime along path P. Letting the slowness be a continuous distribution $s(\mathbf{r})$, we have

$$\tau^P[s] = \int_P s(\mathbf{r})\, dl^P, \tag{28}$$

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 $\tau^P(s)$ 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

$$\tau^*[s] = \min_{P \in Paths} \tau^P[s], \tag{29}$$

where *Paths* denotes the set of all continuous paths connecting the given source and receiver. The particular path that produces the minimum in (29) 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.

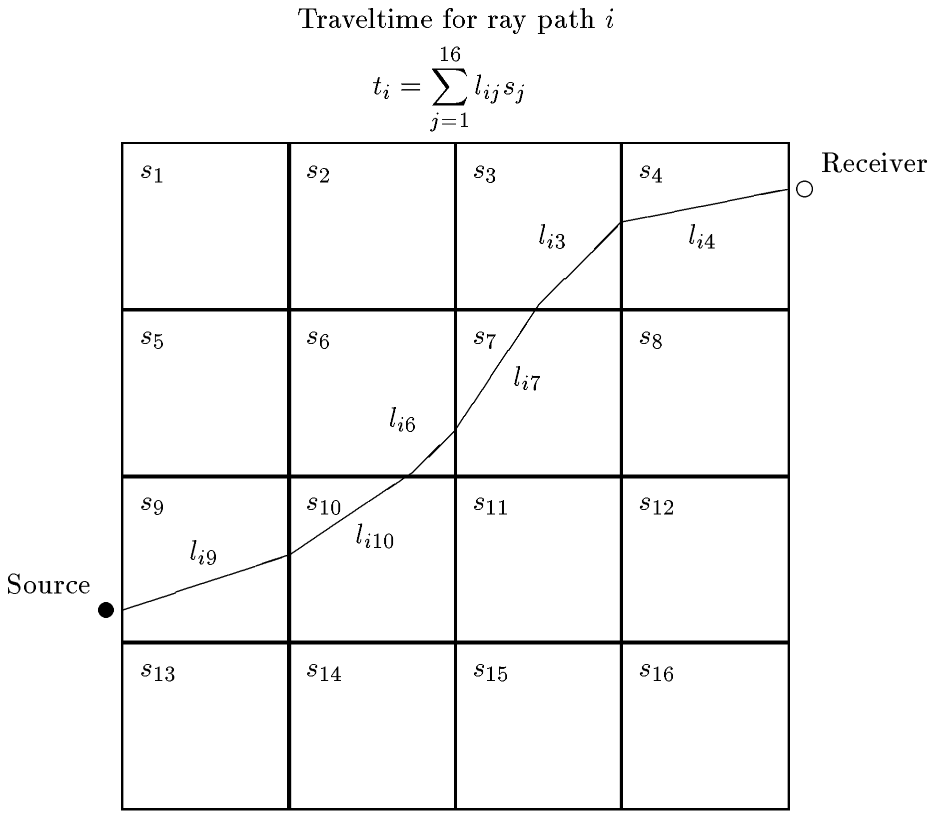

Figure 6: Schematic illustration of ray paths through a cell slowness model.

Substituting (28) into (29), we have *Fermat's principle of least time*:

$$\tau^*[s] = \int_{P^*} s(\mathbf{r})\, dl^{P^*} = \min_P \int_P s(\mathbf{r})\, dl.^P \tag{30}$$

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

3.3 Seismic inversion and tomography

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

$$\int_{P_i} s(\mathbf{r})\, dl^{P_i} = t_i, \quad i = 1, \ldots, m. \tag{31}$$

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

$$l_{ij} = \int_{P_i \cap \text{cell}_j} dl^{P_i}. \tag{32}$$

For n cells, Eq. (31) can then be written as

$$\sum_{j=1}^{n} l_{ij} s_j = t_i, \quad i = 1, \ldots, m. \tag{33}$$

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.

We can rewrite (33) in matrix notation by defining the column vectors $\mathbf{s}$ and $\mathbf{t}$ and the matrix $\mathbf{M}$ as follows:

$$\mathbf{s} = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix}, \qquad \mathbf{t} = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_m \end{pmatrix}, \qquad \mathbf{M} = \begin{pmatrix} l_{11} & l_{12} & \cdots & l_{1n} \\ l_{21} & l_{22} & \cdots & l_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ l_{m1} & l_{m2} & \cdots & l_{mn} \end{pmatrix}. \tag{34}$$

Equation (33) then becomes the basic equation of forward modeling for ray equation analysis:

$$\mathbf{Ms} = \mathbf{t}. \tag{35}$$

Note that equation (35) may be viewed as a numerical approximation to equation (30), *i.e.*, it is just a discretized (or finite element) version of the equation. Equation (35) may be used for any set of ray paths, whether those ray paths minimize (30) or not. If the ray paths used to form the matrix $\mathbf{M}$ are actually minimizing ray paths, $\mathbf{M}$ is then implicitly a function of the slowness model $\mathbf{s}$.

The methods developed apply to both two-dimensional and three-dimensional imaging applications.

3.4 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 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 for seismic inversion) need not be continuous or differentiable and, furthermore, are very often rather difficult to compute. Feasibility constraints for such inversion problems are therefore generally implicit, rather than explicit.

Equation (31) 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

$$\int_{P_i} s(\mathbf{r}) dl^{P_i} \geq t_i, \tag{36}$$

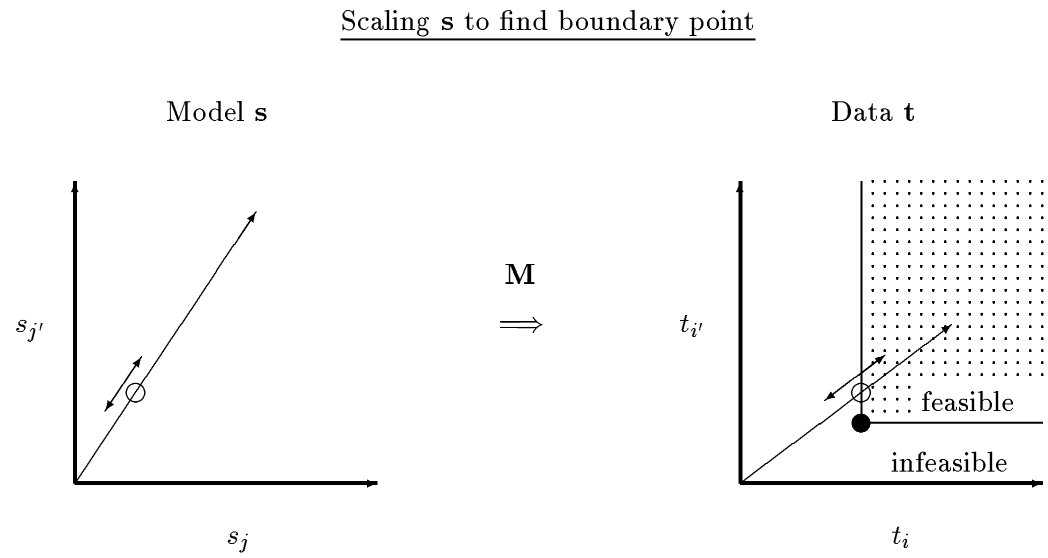

Figure 7: Scaling $\mathbf{s}$ to find feasibility boundary point. Feasibility boundary in data space is determined explicitly by the data. Feasibility boundary in model space is determined implicitly through the traveltime calculation.

where now t_i is the measured traveltime for source-receiver pair i. When we discretize (36) for cell or block models and all ray paths i, the resulting set of m inequalities may be written as

$$\mathbf{Ms} \geq \mathbf{t}. \tag{37}$$

Equations (36) and (37) can be interpreted as a set of inequality constraints on the slowness model $\mathbf{s}$. When $\mathbf{s}$ obeys these m constraints, we say that $\mathbf{s}$ is *feasible.* When any of the constraints is violated, we say $\mathbf{s}$ 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 our inversion problems, an additional computation is required. The feasibility constraints are *explicit* for the traveltime data vector, 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 $s_1(\mathbf{r}), s_2(\mathbf{r})$ are two model slowness functions *in the feasible set* for a given set of ray paths $\{P\}$. Let $s_\lambda(\mathbf{r}) = \lambda s_1(\mathbf{r}) + (1-\lambda)s_2(\mathbf{r})$ be the convex combination of these two slownesses, 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, we have

$$\tau_i^P[s_\lambda] = \lambda\tau_i^P[s_1] + (1-\lambda)\tau_i^P[s_2]. \tag{38}$$

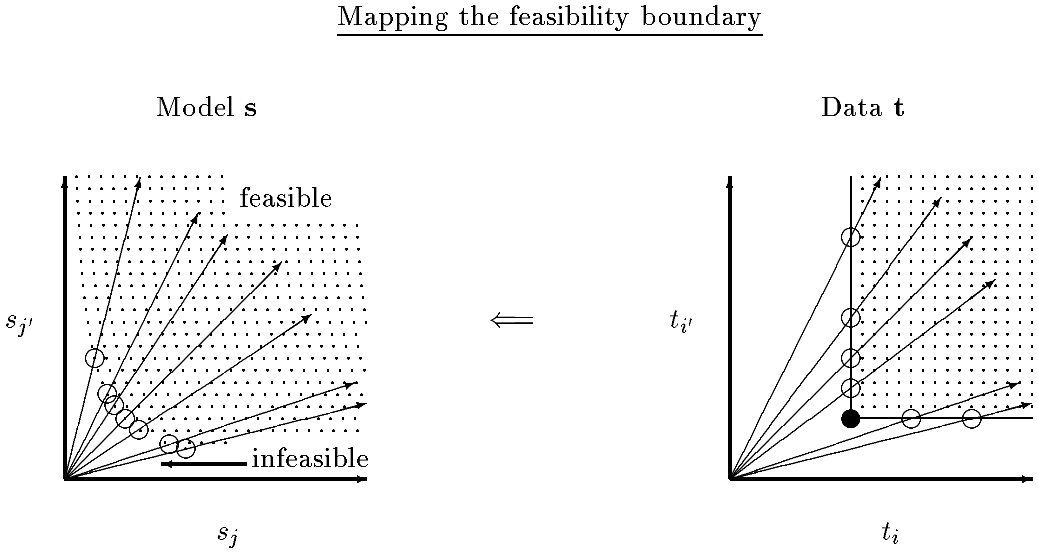

Figure 8: By scaling many slowness vectors $\mathbf{s}$, the location of the feasibility boundary in the model space can be mapped.

But $\tau_i^P[s_1], \tau_i^P[s_2] \geq t_i$ [by assumption] and λ and $(1-\lambda)$ are non-negative. Therefore, substituting into (38), we find

$$\tau_i^P[s_\lambda] \geq \lambda t_i + (1-\lambda)t_i = t_i. \tag{39}$$

Thus, the convex combination s_λ lies in the feasible set if s_1 and 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 nonlinearity 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 for trial eigenfunctions 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].

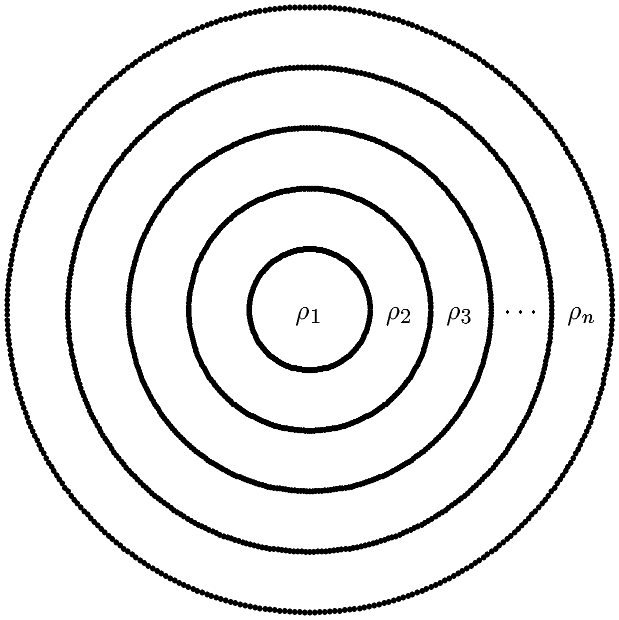

Figure 9: A sphere composed of concentric shells of constant density ρ may be used to model free oscillations of the Earth. Compressional and shear wave velocities are assumed to have been previously determined using traveltime tomography. The inverse problem solves for density structure using measured eigenfrequencies of toroidal oscillation in this example.

3.5 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 later in this paper.

4 Feasibility constraints for density distribution in the Earth

A rather 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]. We concentrate on toroidal modes since they are independent of gravitational effects.

Let B and S be certain functionals of the exciting eigenfunctions. Then, 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. If the compressional and shear wave speeds (v_c and v_s, respectively) have been found using (for example) traveltime tomography, then we can use the identities

$$v_c^2 = (\kappa + 4\mu/3)/\rho \tag{40}$$

and

$$v_s^2 = \mu/\rho, \tag{41}$$

where ρ is again the mass density, to show that the potential energy density may be written in the form

$$\rho U = \rho \left[\left(v_c^2 - 4v_s^2/3 \right) B + v_s^2 S \right]. \tag{42}$$

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 (normally 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. The Fourier transform of this energy density takes the general form $\omega^2 \rho K$, where ω is the angular frequency of oscillation, and K is the square of the trial displacement.

The energy partition in a vibrating medium is known to balance the total kinetic and potential energies in the volume over a cycle. If the volume is Ω, this balance between the two energies may be expressed in the equality

$$\omega^2 \int_\Omega \rho K d^3r = \int_\Omega \rho U d^3r, \tag{43}$$

where 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. Equation (43) can be rewritten in the form of a variational inequality as

$$\omega^2 \le \frac{\int \rho U d^3r}{\int \rho K d^3r}, \tag{44}$$

in which the right hand side is the Rayleigh-Ritz quotient, and the inequality itself follows from the Rayleigh-Ritz variational principle. In (44), K and U are now functionals of *trial* eigenfunctions, and equality is obtained when the minimum of the Rayleigh-Ritz quotient is found. Thus, equality is achieved 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, we 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 important difference between the free oscillation problem and the ones discussed previously (string vibration and traveltime inversion). The

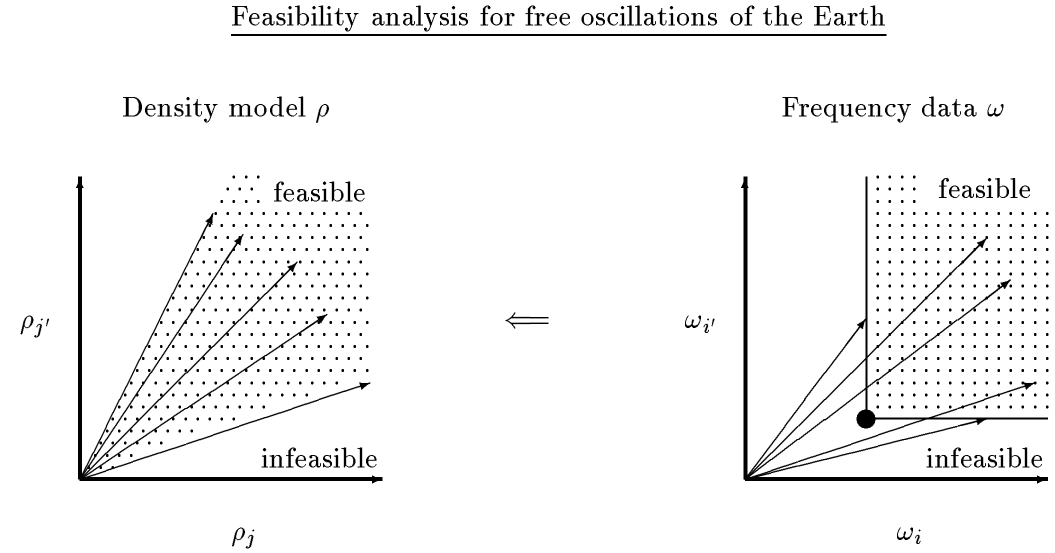

Figure 10: For free oscillations, scaling the density ρ by a positive constant does not change either the trial eigenfunctions or the frequencies, showing that any point along a ray in density space is equally good for satisfying the frequency data. Thus, the feasible region forms a cone in the density space. of the Earth.

boundary of the feasible set cannot be found by scaling the density in this problem. Therefore, the shape of the feasibility boundary is seen to be a cone in the model space.

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

$$\int_\Omega \rho\, d^3r = M. \tag{45}$$

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

$$\int_\Omega \rho r^2 d^3r = I, \tag{46}$$

when it is known.

To attack the inverse problem, we again introduce the concept of particular feasible density distributions and collections of these called "feasible sets," as before. We assume that some subset of the modal frequencies has been measured. For simplicity, we display equations for only the first frequency and one trial eigenfunction, but we understand that the same argument applies simultaneously to all trial eigenfunctions, as long as we restrict discussion to the first eigenfrequency. Suppose that $\rho_a(\mathbf{x})$ and $\rho_b(\mathbf{x})$ are two density distributions, both satisfying the Rayleigh-Ritz inequality for the same trial eigenfunction. Thus, we are considering only local or trial feasibility of the densities in this first example. By assumption, we have

$$\omega_1^2 \int \rho_a K d^3r \le \int \rho_a U d^3r \tag{47}$$

Constraints to fix the scale of the density

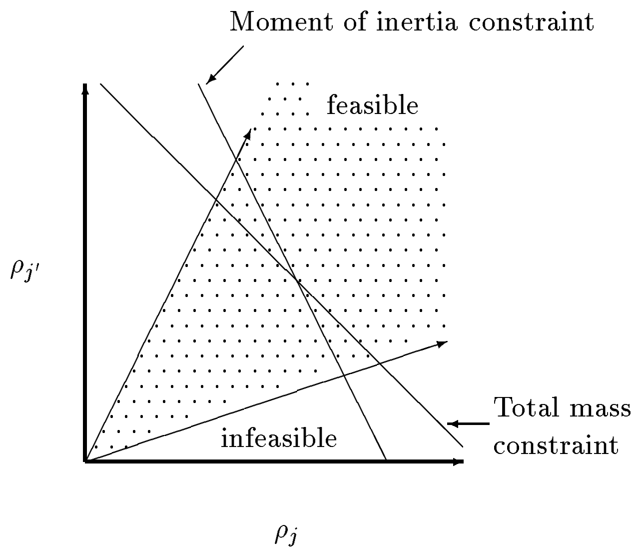

Figure 11: The scale of the density is determined by the total mass and/or the moment of inerita of the Earth.

and

$$\omega_1^2 \int \rho_b K d^3 r x \leq \int \rho_b U d^3 r. \tag{48}$$

If we next multiply (47) by λ and (48) by $(1-\lambda)$ and add the two resulting inequalities, we find

$$\omega_1^2 \int \rho_\lambda K d^3 r \leq \int \rho_\lambda U d^3 r, \tag{49}$$

where $\rho_\lambda(\mathbf{r}) = \lambda\rho_a(\mathbf{r}) + (1-\lambda)\rho_b(\mathbf{r})$ is defined as the convex combination of the two densities. Thus, we 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 the local feasibility set is convex, whenever the set of trial eigenfucntions is uncorrelated with the density.

The proof is similar for global feasibility. Suppose we 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

$$\omega_1^2 \leq \frac{\int \rho_a U_a d^3 r}{\int \rho_a K_a d^3 r} \leq \frac{\int \rho_a U_* d^3 r}{\int \rho_a K_* d^3 r}, \tag{50}$$

and similarly

$$\omega_1^2 \leq \frac{\int \rho_b U_b d^3r}{\int \rho_b K_b d^3r} \leq \frac{\int \rho_b U_* d^3r}{\int \rho_b K_* d^3r}, \tag{51}$$

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, we find that the convex combination of densities ρ_λ must satisfy

$$\omega_1^2 \leq \frac{\int \rho_\lambda U_* d^3r}{\int \rho_\lambda K_* d^3r} \tag{52}$$

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 (52), *i.e.*, when $U_* = U_\lambda$ and $K_* = K_\lambda$. Thus, it follows from (52) that if ρ_a and ρ_b are globally feasible densities, then so is their convex combination ρ_λ since

$$\omega_1^2 \leq \frac{\int \rho_\lambda U_\lambda d^3r}{\int \rho_\lambda K_\lambda d^3r}. \tag{53}$$

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.

When frequencies higher than the first are included in the analysis, the same major problem arises that we found in the string inversion problem. Admissibility constraints on the trial eigenfunctions make it difficult to analyze the feasibility structure, but it seems likely that the feasible set will be nonconvex, as it was for the string problem.

5 Simultaneous feasibility 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

$$\omega^2 \leq \frac{\int (\kappa B + \mu S)\, d^3r}{\int \rho K d^3r}. \tag{54}$$

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

$$\omega_1^2 \int \rho_a K d^3r \leq \int (\kappa_a B + \mu_a S)\, d^3r \tag{55}$$

and

$$\omega_1^2 \int \rho_b K d^3r \leq \int (\kappa_b B + \mu_b S)\, d^3r. \tag{56}$$

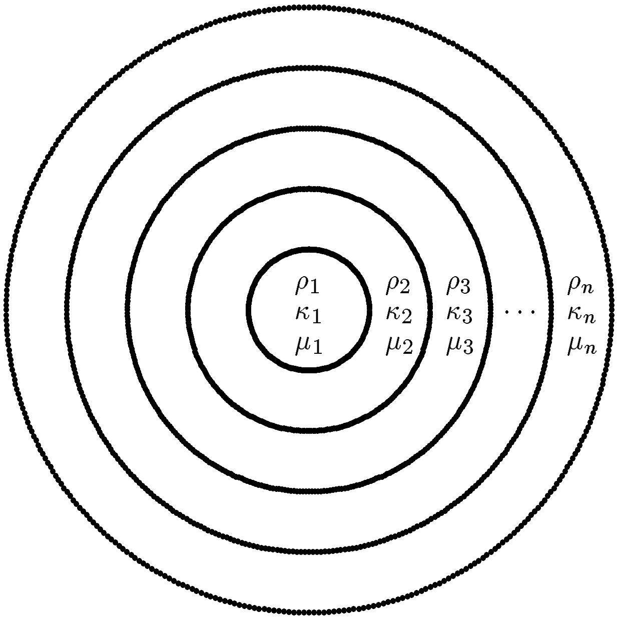

Figure 12: If compressional and shear velocities are not known, then a sphere composed of concentric shells of constant density ρ, bulk modulus κ, and shear modulus μ may be used to model free oscillations of the Earth. As in the last example, model structure is inverted from eigenfrequency data.

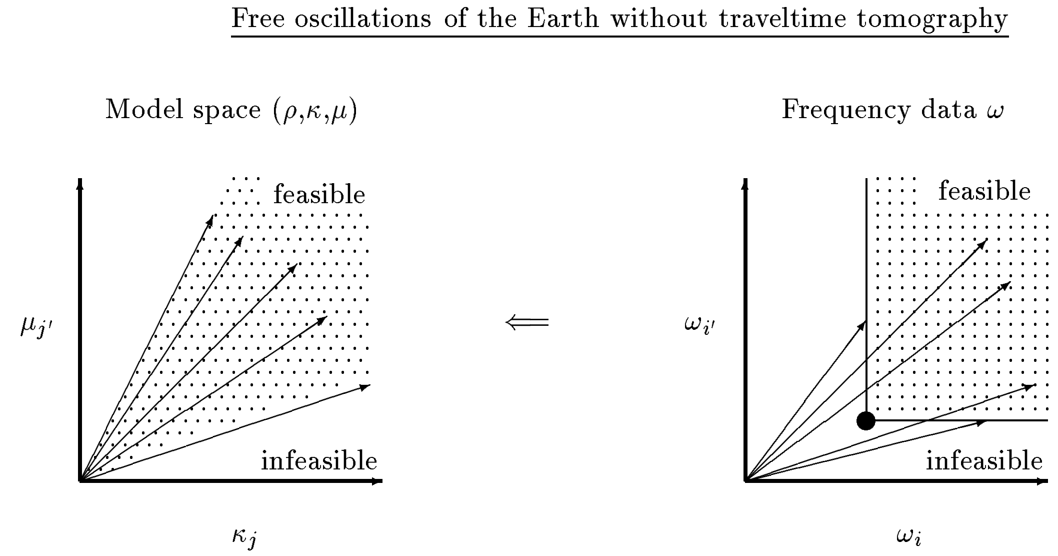

Figure 13: Scaling the triple (ρ,κ,μ) by a positive constant does not change either the trial eigenfunctions or the frequencies, showing again that any point along a ray in the model space is equally good for satisfying the frequency data. The overall scale is again determined by the total mass and/or the moment of inertia of the Earth.

And the convex combination

$$(\rho_\lambda, \kappa_\lambda, \mu_\lambda) \equiv (\lambda\rho_a + (1-\lambda)\rho_b, \lambda\kappa_a + (1-\lambda)\kappa_b, \lambda\mu_a + (1-\lambda)\mu_b) \tag{57}$$

clearly satisfies the corresponding condition

$$\omega_1^2 \int \rho_\lambda K d^3r \leq \int (\kappa_\lambda B + \mu_\lambda S)\, d^3r, \tag{58}$$

where the same trial eigenfunctions have been used in (55)–(58). This again establishes convexity of the local feasibility set based on data for the lowest eigenfrequency 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 (unfortunately) scale invariant, so an overall change in scale of the form $(\rho, \kappa, \mu) \to (\gamma\rho, \gamma\kappa, \gamma\mu)$ 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 (45) and/or total moment of inertia (46) 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 the size of the resulting model space. 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.

6 Variational structure for inverse problems

The various inverse problems considered so far may be viewed as special cases of the following general problem.

Suppose we have the set of measured data $\{q_i\}$ and that these data are known — from the mathematical physics of the related forward problem — to be minima of an appropriate variational functional so that

$$q_i \leq Q_i[n_1, \ldots, n_k; d_1, \ldots, d_m], \tag{59}$$

where Q_i is a functional whose $(k+m)$ arguments are parameters of the model (such as density for string vibrations, slowness in traveltime tomography, or bulk and shear moduli in elastic vibration). There are additional dependencies on the trial eigenfuctions contained implicitly in (59), but we suppress these in the present notation. If the functional can be written in the form of a quotient

$$Q_i[n_1, \ldots, n_k; d_1, \ldots, d_m] = \frac{N_i[n_1, \ldots, n_k]}{D_i[d_1, \ldots, d_m]}, \tag{60}$$

where the functionals in the numerator N_i and in the denominator D_i are themselves linear functionals of each of their arguments, then we find that all the examples considered may be written in this form (see Table 2).

It follows easily from this postulated form that, when the trial eigenfunctions can be treated as uncorrelated with the model parameters, 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

$$q_i D_i[\hat{d}_1, \ldots, \hat{d}_m] \leq N_i[\hat{n}_1, \ldots, \hat{n}_k] \quad \text{for all} \quad i \tag{61}$$

and if

$$q_i D_i[\tilde{d}_1, \ldots, \tilde{d}_m] \leq N_i[\tilde{n}_1, \ldots, \tilde{n}_k] \quad \text{for all} \quad i, \tag{62}$$

then, for any λ in the range $0 \leq \lambda \leq 1$, we find

$$q_i \leq \frac{N_i[n_1^{(\lambda)}, \ldots, n_k^{(\lambda)}]}{D_i[d_{(\lambda)}, \ldots, d_m^{(\lambda)}]} \quad \text{for all} \quad i, \tag{63}$$

where

$$n_j^{(\lambda)} = \lambda \hat{n}_j + (1-\lambda)\tilde{n}_j \quad \text{for} \quad 1 \leq j \leq k, \tag{64}$$

and

$$d_l^{(\lambda)} = \lambda \hat{d}_l + (1-\lambda)\tilde{d}_l \quad \text{for} \quad 1 \leq l \leq m, \tag{65}$$

guaranteeing that the convex combinations of the "hat" and "tilde" models parametrized by λ are also feasible if the "hat" and "tilde" models are themselves feasible. The arguments given previously may be repeated to demonstrate the existence of a convex feasible set for all such problems.

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 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 may have feasibility sets that are nonconvex, while traveltime tomography and some other similar inversion problems [Berryman, 1991] in general do have convex feasibility sets.

TABLE 2. Comparison of feasibility sets for three inverse problems.

Inverse 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

Convexity of the feasible set is not really necessary when designing a nonlinear programming method for solving the inverse problem as we have shown in the string density inversion problem, but convexity does help to simplify reconstruction algorithms and is therefore a desirable property worthy of note when present.

7 Discussion

7.1 Finite data sets

Suppose we have a finite set of measurements $\{q_1, \ldots, q_M\}$ and we 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 that in general $\mathcal{F}_{M+1} \subseteq \mathcal{F}_M \subseteq \mathcal{F}_{M-1} \subseteq \cdots \subseteq \mathcal{F}_1$, independent of the convexity or lack of convexity of the feasibility set. For example, we saw that this was so in the string density inversion problem presented earlier in the paper.

7.2 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 place the data in the role of upper or lower bounds on integrals involving the unknown parameters. The data therefore enter 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. We can 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 our 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 our 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.

7.3 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 sometimes convex nature of the feasible set suggests a number of practical algorithms. We 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 difficulty 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 more typical 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 (such as the string density problem), 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 stabilized easily 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). Rather, this approach 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.

8 Conclusions

We have examined the variational structure of the inverse problem for three types of systems: (1) string density inversion, (2) traveltime tomography, and (3) free oscillations of the Earth. Of these three problems, traveltime tomography is the best example of the real power of using variational/feasibility analysis for solving an inverse problem. Every traveltime datum serves as an independent measure of Fermat's minimum traveltime functional. Wave slowness appears linearly in the traveltime integral and this fact together with the homogeneity of the functional in the slowness scale factor provide very strong clues about how to solve the inverse problem when many traveltime data have been measured.

In contrast, the free oscillation problem is the worst example of variational/feasibility analysis. Although a feasibility set exists for this problem, its boundary is not easily mapped because frequencies of oscillation are actually independent of the overall scale for the elastic moduli and the density.

The string density inversion problem proves to be an example of intermediate difficulty. Although the integral appearing in the Rayleigh-Ritz quotient is apparently linear in the density, the overall functional is nonlinear (and in a complex way) because of the constraints on the trial eigenfunctions required for eigenfrequencies of order higher than the first. A feasibility set still exists and scaling can be used to find its boundary, but care must be taken when doing so because this feasible set is not convex.

We conclude that the scaling property of the integrals is more important for solution of an inverse problem than is linearity in the parameters of interest. This suggests that variational/feasibility analysis may prove useful in other problems with more complex structure than the ones considered here, if those problems have useful scaling behavior.

Acknowledgments

Early numerical experiments (not included here) for the string vibration inversion problem were performed by O. S. Tai, a summer research student visiting at LLNL. I thank J. R. McLaughlin for helpful discussions 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. This work was performed under the auspices of the U. S. Department of Energy by the Lawrence Livermore National Laboratory under contract No. W-7405-ENG-48 and supported specifically by the Geosciences Research Program of the DOE Office of Energy Research within the Office of Basic Energy Sciences, Division of Engineering and Geosciences. Part of this work was done while the author was visiting the Geophysics Department at Stanford University, supported in part by sponsors of the the Stanford Exploration Project. The manuscript was completed while the author was on sabbatical at l'Institut de Physique du Globe de Paris, with partial support from the French government. All sources of support for this work are hereby gratefully acknowledged.

References

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

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

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

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

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

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

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

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

F. Gilbert, 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, 1971), pp. 107–121

F. Gilbert, 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, 1980, pp. 41–81.

H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, Massachusetts, 1980), pp. 35–69, 438–498.

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

C. Lanczos, *The Variational Principles of Mechanics* (Dover, New York, 1970), pp. xxvi–xxix, 35–73, 229–290.

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

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

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

A. Morelli and A. M. Dziewonski, 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, 1987, Chapter 11, pp. 251–274.

J. W. Rector, III, Crosswell methods: Where are we, where are we going? *Geophysics* **60**, 629–630 (1995).

R. Snieder, 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.