# RAY TRACING IN COMPLEX THREE-DIMENSIONAL EARTH MODELS

Thesis by

Tomas Paul Girnius

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1986

(Submitted 9/26/1985 )

# ACKNOWLEDGEMENTS

I thank Caltech for providing financial support in the form of Graduate Teaching Assistantships. The Dept. of Energy graciously provided additional support under Contract DE-AA03-76SF0076.

I am grateful to Professor H. B. Keller for suggesting topics of research and for having the patience and wisdom to allow me to explore them at my own phlegmatic pace. Those to whom my completion of this thesis mattered are my friends. To them I dedicate it.

# ABSTRACT

The problem of tracing seismic rays between specified source and receiver is discussed for Earth models consisting of layers, in which velocity varies linearly, that are separated by material interfaces of arbitrary shape. The calculation of travel times, amplitudes, and phase shifts is considered. Fast and efficient numerical algorithms are developed. Computed examples are presented.

# TABLE OF CONTENTS


| | page |
|---|---|
| Acknowledgements | ii |
| Abstract | iii |
| Table of Contents | iv |
| List of Figures | v |
| I. Introduction | 1 |
| II. The Forward Problem | 4 |
| II.0 Introduction to the Forward Problem | 4 |
| II.1 Description of Problem and Notational Conventions | 4 |
| II.2 Calculation of Rays and Travel-Times | 7 |
| 2.1 The Ray Equations and Jump Conditions (Snell's Law) | 7 |
| 2.2 Rays in a Linearly Stratified Medium | 8 |
| 2.3 Computation of Travel-Times | 11 |
| 2.4 Some Limits and Rescaling for Computation | 12 |
| II.3 Calculation of Amplitudes and Phase Shifts, and Detection of Caustics | 14 |
| 3.1 Wavefront Curvatures Before Encounter With an Interface | 14 |
| 3.2 Wavefront Curvatures After Encounter With an Interface and Geometric Spreading Factors | 16 |
| 3.3 Effect on Wavefront Curvature by an Interface | 20 |
| 3.4 Reflection, Transmission, and Conversion Coefficients | 24 |
| 3.5 Detection of Caustics | 27 |

3.6 Phase Shifts 28

II.4 Numerical Methods for the Forward Problem 30

4.1 Nonlinear Algebraic System for the Nodal Points 30

4.2 Method of Solution - Newton's Method 33

4.3 Selection of Initial Guess $\mathbf{X}^{(0)}$ - General Considerations 35

4.4 Initial Guess for Rays of Type I 37

4.5 Initial Guess for Rays of Type II 38

4.6 Some Combinatorics 41

4.7 Continuation Along a Gather of Receivers 42

4.8 Normal Incidence Rays 44

II.5 Numerical Examples 50

5.1 Phantom Interfaces and Regions of Convergence 50

5.2 Continuation in Receiver Location and Class Transitions 53

5.3 A Search for All Rays 56

III. Some Comments On the Inverse Problem 70

III.1 An Inverse Problem and Some Heuristic Considerations 70

References 74

# LIST OF FIGURES


| | |
|---|---|
| Figure 1 | 17 |
| Figure 2a | 36 |
| Figure 2b | 36 |
| Figure 3a | 40 |
| Figure 3b | 40 |
| Figure 3c | 40 |
| Figure 3d | 41 |
| Figure 3e | 41 |
| Figure 4a | 54 |
| Figure 4b | 55 |
| Figure 5a | 58 |
| Figure 5b | 59 |
| Figure 5c | 60 |
| Figure 5d | 61 |
| Figure 5e | 62 |
| Figure 5f | 63 |
| Figure 5g | 64 |
| Figure 5h | 65 |
| Figure 5i | 66 |
| Figure 5j | 67 |
| Figure 5k | 68 |

# I. INTRODUCTION

The forward problem in seismology is to describe completely the wave propagation resulting from a known disturbance in a general medium with arbitrary, but specified, variations in elastic properties. In general, only numerical methods are available to analyze this problem for even relatively simple heterogeneities in elastic properties. Numerical methods for the full elastodynamic wave equations are limited by excessive requirements both for storage and for computational time.

The ray theoretic development provides an attractive alternative. It is computationally simpler. Moreover, it provides information precisely about those aspects of seismic wave propagation which can be observed: the arrival times of disturbances at a known receiver location, as well as their relative amplitudes and phase shifts.

Because ray theory is based on an asymptotic series expansion in inverse powers of frequency, a major drawback is that it is useful only in the high frequency limit. However, when the assumption of a high frequency seismic signal is valid, even the leading order term of the asymptotic ray series can be used to interpret many features of seismogram recordings.

Because of the usefulness of ray theory in interpreting seismic records, seismologists have developed several methods for computing the 'rays' ( the normal trajectories to the wavefronts ) joining a known source and receiver. *

The equations defining these rays are nonlinear, and can, in general, only be

---

* Strictly speaking, the rays are normal to the wavefronts only in isotropic media. Because anisotropy in the Earth is thought to be small ($\leq$5%), we will adhere to this assumption throughout.

integrated numerically. Essentially, there are two approaches to solving these ray equations for trajectories between a given source and receiver pair. In the first of these, commonly known as 'shooting,' a sequence of initial value problems for rays emanating from the source with different initial tangents is solved until several rays are found which emerge close to the receiver. By interpolating the initial conditions for these nearby rays, a ray which emerges arbitrarily close to the receiver is found. This approach has been explored, for example, by Shah (1973).

Alternatively, one tries to solve the two-point boundary value problem directly. This approach has been suggested by Julian and Gubbins (1977). Pereyra, Lee, and Keller (1980) developed efficient algorithms capable of solving the two-point ray problem in very general heterogeneous media. Their approach involved using finite difference approximations to the coupled, first order, nonlinear differential ray equations, and then solving the resulting system of nonlinear algebraic equations with an iterative scheme.

An approach similar in spirit to that of Peyrera, Lee, and Keller has been developed by Keller and his students Perozzi (1980) and Fawcett (1983). Working from the premise that in many regions of the Earth the velocity distribution for elastic waves may be satisfactorily modeled as piecewise homogeneous with discontinuities across possibly complicated interfaces, they developed fast and accurate numerical algorithms for the calculation of rays in both two-dimensional (Perozzi) and three-dimensional (Fawcett) cases.

Because geometric rays in a homogeneous medium are straight lines, they were able to forego the numerical integration of the differential ray equations, solving instead a system of nonlinear algebraic equations corresponding to Snell's law of refraction across each interface.

The number of equations in this method is directly proportional to the number of velocity discontinuities which the ray encounters. Consequently, although media

with sizeable velocity gradients can, in principle, be modeled as a dense cascade of homogeneous layers, it becomes impractical to compute rays in such media with the method of Keller, Perozzi, and Fawcett without further modification.

It has been observed that continuously varying velocities in the Earth (even in certain regions of the Earth's crust) can often be approximated accurately by linear distributions with non-negligible constant gradients. As an example, for typical undisturbed tertiary sand-shale clastics, the empirical relationship

$$V(z) = 1600(m/s) + 0.55z$$

($z$ depth in meters) has been found to hold down to depths of about 3,000 meters, representing a velocity doubling in just several kilometers (Kley, A.M., 1983). When such regions have been disturbed, for example, by the intrusion of diapiric salt domes, it can be expected that the magnitude of the velocity gradients will still be comparable. Their directions, however, may be significantly displaced from the vertical.

The oil potential beneath the lobes of such diapiric salt intrusions motivated our study of fast and accurate ray calculations in complex three-dimensional regions containing such linearly plane-stratified regions.

# II. THE FORWARD PROBLEM

## II.0 INTRODUCTION TO THE FORWARD PROBLEM

In this chapter, we discuss the forward ray problem for piecewise linearly stratified media. In II.1, we provide a description of the problem and introduce some notational conventions to which we shall adhere. In II.2, we discuss the calculation of the actual rays and travel times. The calculation of wave amplitudes and phases, as well as the detection of caustics, is discussed in II.3. Our numerical treatment of the problem is detailed in II.4. Finally, in II.5, we provide and discuss numerical results for several examples.

## II.1 DESCRIPTION OF THE PROBLEM AND SOME NOTATIONAL CONVENTIONS

We adopt a right-handed Cartesian coordinate system $\mathbf{x} = (x, y, z)$, with $z$ increasing with depth in the Earth.

The Earth is assumed to be isotropic, linearly elastic, with piecewise continuous variations in the elastic parameters $\rho(\mathbf{x})$, $\lambda(\mathbf{x})$, and $\mu(\mathbf{x})$. Here $\mathbf{x}$ denotes position, $\rho(\mathbf{x})$ the density, and $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$ the Lamé parameters.

Discontinuities in the elastic parameters or their spatial derivatives will be assumed to occur across prescribed, non-intersecting interfaces

$$z = f_i(x, y), \quad i = 0, 1, \ldots, i_{max}, \tag{1.1}$$

where $f_i(x, y)$, together with its first and second derivatives, is continuous.

The free surface of the Earth will be defined by

$$z = f_0(x, y). \tag{1.2}$$

Because our main concern is with problems in exploration geophysics, for which length scales are of the order of only several kilometers, we will usually adopt a flat free surface with $f_0(x,y) = 0$.

For the same values of their arguments, $f_j(x,y) < f_k(x,y)$ if $j < k$. We will denote by the $k^{th}$ layer that region of the Earth between the interfaces defined by $z = f_{k-1}(x,y)$ and $z = f_k(x,y)$. The $k^{th}$ layer can sustain wave motions with two distinct wave speeds, (compressional $(p)$, and shear $(s)$), given, respectively, by

$$V_k^p(\mathbf{x}) = \left( \frac{\lambda_k(\mathbf{x}) + 2\mu_k(\mathbf{x})}{\rho_k(\mathbf{x})} \right)^{1/2} \tag{1.3a}$$

and

$$V_k^s(\mathbf{x}) = \left( \frac{\mu_k(\mathbf{x})}{\rho_k(\mathbf{x})} \right)^{1/2}. \tag{1.3b}$$

We constrain the variations in the elastic parameters within the $k^{th}$ layer to be such that the compressional and shear wave velocities satisfy

$$V_k^p(\mathbf{x}) = D_k^p + \alpha_k^p(\hat{\mathbf{A}}_k \cdot \mathbf{x}) \tag{1.4a}$$

and

$$V_k^s(\mathbf{x}) = D_k^s + \alpha_k^s(\hat{\mathbf{A}}_k \cdot \mathbf{x}), \tag{1.4b}$$

respectively. Here $\alpha_k^s$, $\alpha_k^p$, $D_k^s$, and $D_k^p$ are prescribed constants, while $\hat{\mathbf{A}}_k$ is a prescribed unit vector.

Strictly speaking, vertically polarized shear waves $(sv)$ do not decouple from compressional waves $(p)$ in such a medium even within a layer of continuous variation in the elastic parameters. However, we confine our attention to waves which travel within any continuous layer as well-defined $(p)$ or $(s)$ waves, with conversions from one type to another occuring only at an interface $z = f_i(x,y)$.

Given the assumed velocity distribution described above, we seek to determine all physical rays joining known source and receiver locations and passing through

a prescribed sequence of interfaces, as well as the travel-times, relative amplitudes, and phases of disturbances propagating along these rays.

Within each layer, the ray segments are circular arcs (or, in the limit of homogeneous velocity, straight line segments) which, as will be shown, are uniquely determined by their endpoints on successive interfaces along with the velocity distribution.

We refer to those points at which a ray intersects an interface as *nodes*, with the coordinates of the $k^{th}$ node given by

$$\mathbf{x}_k \equiv (x_k, y_k, z_k) = (x_k, y_k, f_{i_k}(x_k, y_k)), \tag{1.5}$$

where it is assumed that the $k^{th}$ node lies on the $i_k^{th}$ interface.

For convenience, we augment this sequence of *a priori* unknown nodes with

$$\mathbf{x}_0 \equiv \text{source location} \tag{1.6a}$$

and

$$\mathbf{x}_{N+1} \equiv \text{receiver location}, \tag{1.6b}$$

either of which may or may not lie on an interface.

At each of the nodes $\mathbf{x}_k$, $(k = 1, \ldots, N)$, the ray must satisfy Snell's law in its most general three-dimensional form. This requirement leads to a system, which we shall subsequently formulate precisely, of $2N$ nonlinear equations for the $2N$ unknowns $x_k, y_k$, $k$ $1, \ldots, N$. This system is analogous to that for the piecewise homogeneous velocity case (studied by Fawcett, 1983), and, in fact, reduces to it in the limiting case of $\alpha_k^p = 0$ and $\alpha_k^s = 0$ for all $k$.

In subsequent sections, we shall adopt the following notational conventions. Either superscript or subscript, or both, may be omitted from the velocities defined by the above equations when no confusion will result.

By $\mathbf{T}_k^-$, we denote the unit tangent vector to the incident ray segment at the $k^{th}$ node. Similarly, we denote by $\mathbf{T}_k^+$ the unit tangent vector to the emergent ray segment.

Note that the functional dependence of $\mathbf{T}_k^-$ is on $\mathbf{x}_{k-1}$, $\mathbf{x}_k$, $D_k$, $\alpha_k$ and $\hat{\mathbf{A}}_k$, whereas that of $\mathbf{T}_k^+$ is on $\mathbf{x}_k$, $\mathbf{x}_{k+1}$, $D_{k+1}$, $\alpha_{k+1}$ and $\hat{\mathbf{A}}_{k+1}$.

When we wish to refer to the unit tangent at an arbitrary point $\mathbf{x}$ along the ray, we write simply $\mathbf{T}[\mathbf{x}]$, suppressing the implicit dependence on ray segment endpoints and velocity of the appropriate layer.

## II.2 CALCULATION OF RAYS AND TRAVEL-TIMES

### II.2.1 The Ray Equations and Jump Conditions (Snell's Law)

In the first order geometric optics approximation, the 'rays' (i.e., the orthogonal trajectories to the wavefronts) may be derived from Fermat's Principle.

Let $\mathbf{x} = (x, y, z)$, and let $U(\mathbf{x}) = 1/V(\mathbf{x})$, where $V(\mathbf{x})$ is a smoothly-varying velocity field of a heterogeneous, but isotropic medium.

Let '$s$' denote arclength along a ray, and let '$'$' denote differentiation with respect to $s$.

The variational principle, which a ray joining endpoints $\mathbf{P}_0$ and $\mathbf{P}_1$ must satisfy, is given by

$$\delta T = 0 \tag{2.1}$$

where

$$T = \int_{s_0}^{s_1} U(\mathbf{x}(s))\, ds \tag{2.2}$$

subject to the non-holonomic constraint

$$g \equiv \mathbf{x}' \cdot \mathbf{x}' - 1 = 0, \tag{2.3}$$

as well as the boundary conditions

$$\mathbf{x}(s_0) = \mathbf{P}_0 \tag{2.4a}$$

and

$$\mathbf{x}(s_1) = \mathbf{P}_1. \tag{2.4b}$$

Here $\delta$ denotes the first variation. $T$ is the time for the wavefront to travel from $\mathbf{P}_0$ to $\mathbf{P}_1$, and the solution $\mathbf{x}(s)$ is the path of stationary ("least") time.

The Euler equations corresponding to the variational principle above are

$$\nabla\big(U(\mathbf{x})\big) - \frac{d}{ds}\big(U(\mathbf{x})\frac{d\mathbf{x}}{ds}\big) = 0. \tag{2.5}$$

These are the ray equations in their most general three-dimensional form.

Equations (2.5) must hold along the ray wherever $U(\mathbf{x})$ is continuous. At a surface of discontinuity in $U(\mathbf{x})$, the tangent to the ray is discontinuous, and (2.5) must be replaced by the three-dimensional Snell's law

$$V^+(\mathbf{x})\big(\mathbf{N}(\mathbf{x}) \times \mathbf{T}^-(\mathbf{x})\big) = V^-(\mathbf{x})\big(\mathbf{N}(\mathbf{x}) \times \mathbf{T}^+(\mathbf{x})\big) \tag{2.6}$$

Here $\mathbf{N}(\mathbf{x})$ denotes any normal to the surface of discontinuity at the node $\mathbf{x}$. $V^-(\mathbf{x})$ and $\mathbf{T}^-(\mathbf{x})$ denote, respectively, the velocity and unit tangent vector to the ray on the incident side of the discontinuity, while $V^+(\mathbf{x})$ and $\mathbf{T}^+(\mathbf{x})$ are similarly defined for the emergent side.

### II.2.2 Description of Rays in a Linearly Stratified Medium

In this section, we develop formulae that conveniently describe the ray joining given endpoints in a linearly stratified three-dimensional medium. We discuss first the case of two dimensions with the velocity field given by

$$V(x, y) = D + \alpha(\hat{A}_1 x + \hat{A}_2 y) \tag{2.7}$$

We will generalize to three-dimensional case by way of the following two observations. First, all solutions to (2.5) for a plane stratified velocity field are plane curves. Moreover, the tangent vectors at any two points along a given ray are coplanar with the direction of stratification (i.e., with the normal vector to planes of constant velocity).

It is well-known that the rays in the velocity distribution given by (2.7) are arcs of circles. We pin down the precise solution for the ray with endpoints at $(x_0, y_0)$ and $(x_1, y_1)$.

The Euler equations (2.5) may, in this case, be rewritten as

$$(D + \alpha(\hat{A}_1 x + \hat{A}_2 y))y'' - \alpha(\hat{A}_1 y' - \hat{A}_2)(1 + y'^2) = 0, \tag{2.8}$$

with prime now denoting differentiation with respect to $x$.

Inserting into (2.8) the *ansatz*

$$(x - x_c)^2 + (y - y_c)^2 = R^2 \tag{2.9}$$

for the circle of, as yet, undetermined radius $R$ and center coordinates $(x_c, y_c)$, we obtain immediately

$$D + \alpha(\hat{A}_1 x_c + \hat{A}_2 y_c) = 0. \tag{2.10}$$

Hence, the center of any ray lies along that line on which the velocity field vanishes.

Substituting into (2.9) each of the endpoints $(x_0, y_0)$ and $(x_1, y_1)$ in turn, and subtracting the two resulting equations, we obtain a second linear equation for the center coordinates:

$$(x_1 - x_0)x_c + (y_1 - y_0)y_c = \frac{1}{2}(x_1^2 + y_1^2 - x_0^2 - y_0^2). \tag{2.11}$$

The minor arc of the circle determined uniquely by the solution of (2.10) and (2.11) is then the physical ray joining $(x_0, y_0)$ and $(x_1, y_1)$.

We can now readily extend this result to the three-dimensional case of velocity distribution

$$V(x,y,z) = D + \alpha(\hat{A}_1 x + \hat{A}_2 y + \hat{A}_3 z) \tag{2.12}$$

and ray endpoints at $\mathbf{x}_0 = (x_0, y_0, z_0)$ and $\mathbf{x}_1 = (x_1, y_1, z_1)$. The previous result motivates the *ansatz* that the ray joining $\mathbf{x}_0$ and $\mathbf{x}_1$ is a minor geodesic on the sphere

$$(x - x_c)^2 + (y - y_c)^2 + (z - z_c)^2 = R^2. \tag{2.12}$$

Two conditions on the location of the center $\mathbf{x}_c = (x_c, y_c, z_c)$ are obtained immediately from (2.10) and (2.11). These are

$$\alpha\hat{\mathbf{A}}\cdot\mathbf{x}_c = -D \tag{2.13}$$

and

$$(\mathbf{x}_1 - \mathbf{x}_0)\cdot\mathbf{x}_c = \frac{1}{2}(\mathbf{x}_1\cdot\mathbf{x}_1 - \mathbf{x}_0\cdot\mathbf{x}_0). \tag{2.14}$$

A third linear constraint on $\mathbf{x}_c$ is obtained from the requirement that the tangent vectors at any two points along the ray be coplanar with the direction of stratification $\hat{\mathbf{A}}$. One expression for this constraint is

$$\Big((\mathbf{x}_0 - \mathbf{x}_c) \times (\mathbf{x}_1 - \mathbf{x}_c)\Big)\cdot\hat{\mathbf{A}} = \mathbf{0} \tag{2.15}$$

or, more conveniently,

$$\Big(\hat{\mathbf{A}} \times (\mathbf{x}_1 - \mathbf{x}_0)\Big)\cdot\mathbf{x}_c = -Det\Big\{\hat{\mathbf{A}}, \mathbf{x}_0, \mathbf{x}_1\Big\}, \tag{2.16}$$

where '*Det*' denotes the determinant of the column-vector matrix.

The simple linear system for $\mathbf{x}_c$ which we have derived may become singular or very ill-conditioned in either of two cases. In the first of these, the velocity is homogeneous or very nearly so, and $\alpha \to 0$. This case may be dealt with computationally by an appropriate rescaling, as will be shown in a later section. In the second case, $(\mathbf{x}_1 - \mathbf{x}_0)$ is parallel to $\hat{\mathbf{A}}$, and the physical ray is the straight line segment joining $\mathbf{x}_0$ and $\mathbf{x}_1$, even if $\alpha$ is large. This case must be dealt with separately in subsequent computations.

### II.2.3 Computation of Travel-Times

Assume the velocity dependence

$$V(\mathbf{x}) = D + \alpha \hat{\mathbf{A}} \cdot \mathbf{x}, \tag{2.17}$$

and consider the ray segment with endpoints $\mathbf{x}_0$ and $\mathbf{x}_1$.

Introduce projected distances $\xi$ and $\eta$, where $\xi$ is the distance traversed by the ray parallel to the direction of stratification, and $\eta$ is the resolvent normal distance. That is,

$$\xi = \hat{\mathbf{A}} \cdot (\mathbf{x}_1 - \mathbf{x}_0) \tag{2.18}$$

and

$$\eta = \left( \|\mathbf{x}_1 - \mathbf{x}_0\|^2 - \xi^2 \right)^{\frac{1}{2}}. \tag{2.19}$$

The travel-time $T_{01}$ of the ray from $\mathbf{x}_0$ to $\mathbf{x}_1$ is equivalent to the travel-time of a ray from $(\xi_0, 0)$ to $(\xi_0 + \xi, \eta)$ in the $\xi - \eta$ plane with a velocity distribution

$$V(\xi) = D + \alpha \xi \tag{2.20}$$

and where

$$\xi_0 = \hat{\mathbf{A}} \cdot \mathbf{x}_0. \tag{2.21}$$

Let $(\xi_c, \eta_c)$ denote the center of the ray segment, and $R$ its radius. Parametrize the ray segment by

$$\xi = \xi_c + R \sin\theta \tag{2.22}$$

$$\eta = \eta_c - R \cos\theta, \tag{2.23}$$

with $\theta = \theta_0$ and $\theta = \theta_1$ corresponding, respectively, to $(\xi_0, 0)$ and $(\xi_0 + \xi, \eta)$.

Using (2.2), we have

$$T_{01} = \int_{\theta_0}^{\theta_1} \frac{R}{\alpha\xi_c + \alpha R \sin\theta + D} d\theta. \tag{2.24}$$

From (2.10), this simplifies to

$$\begin{aligned} T_{01} &= \frac{1}{\alpha} \int_{\theta_0}^{\theta_1} \frac{d\theta}{\sin\theta} \\ &= \frac{1}{2\alpha} \log \left[ \frac{(1+\cos\theta_0)(1-\cos\theta_1)}{(1-\cos\theta_0)(1+\cos\theta_1)} \right], \end{aligned} \tag{2.25}$$

where

$$\begin{aligned} \cos\theta_0 &= \eta_c/R, & 0 \le \theta_0 \le \pi/2 \\ \cos\theta_1 &= (\eta_c - \eta)/R, & 0 \le \theta_1 \le \pi, \end{aligned} \tag{2.26}$$

and where $\eta_c$ and $R$ are easily found to be

$$\eta_c = \frac{1}{2\eta}\left(\xi^2 + \eta^2 + 2\xi(\xi_0 + D/\alpha)\right) \tag{2.27}$$

$$R = \left\{ (\xi + D/\alpha)^2 + \frac{1}{4\eta^2}\left[(\xi^2+\eta^2) + 2\xi(\xi_0 + D/\alpha)\right]^2 \right\}^{1/2} \tag{2.28}$$

II.2.4 Some Limits as $\alpha \to 0$ and Rescaling for Computations

Equation (2.25), together with the definitions (2.26)-(2.28), provides a closed solution for $T_{01}$, but it is not very useful for computations when $|\frac{\alpha}{D}| << 1$. Anticipating our subsequent discussions of continuation from media that are piecewise homogeneous to those that are piecewise plane-stratified, we provide the following recursive expansion, which is convenient when $|\frac{\alpha}{D}| << 1$.

We define

$$r = 2\xi_0 + {\xi_0}^2 \frac{\alpha}{D} \tag{2.29}$$

$$s = \left\{ (\xi^2+\eta^2) + (\xi^2+\eta^2)(\xi + 2\xi_0)\frac{\alpha}{D} + \left[\frac{1}{4}(\xi^2+\eta^2)^2 + (\xi^2+\eta^2)(\xi\xi_0 + {\xi_0}^2)\frac{\alpha^2}{D^2}\right] \right\}^{1/2}$$

where, evidently, both $r$ and $s$ are $O(1)$ as $\alpha \to 0$. Expanding the logarithm in (2.25) we obtain

$$T_{01} = \frac{1}{2D}\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}\left(\frac{\alpha}{D}\right)^{n-1}\left[(r+s)^n - (r-s)^n\right]. \tag{2.31}$$

From (2.31), we readily obtain

$$\lim_{\alpha\to 0} T_{01} = \frac{(\xi^2+\eta^2)^{1/2}}{D}, \tag{2.32}$$

which recovers the travel-time for the straight-line ray in the homogeneous velocity $V(\mathbf{x}) = D$.

As pointed out in the section II.2.2, as $\alpha \to 0$, the medium tends to a homogeneous one and the system (2.13), (2.14) and (2.16) becomes singular. For homogeneous media, the geometric rays are straight lines, and $\|\mathbf{x}_c\|$ and $R \to \infty$.

It is easy to show, e.g., by solving the two-dimensional case (2.10) and (2.11), that $\|\mathbf{x}_c\|$ and $R$ are $O\left(\frac{1}{\alpha}\right)$, and hence, $\alpha\|\mathbf{x}_c\|$ and $\alpha R$ remain bounded as $\alpha \to 0$. ( The single exception occurs when $(\mathbf{x}_1 - \mathbf{x}_0)$ is parallel to $\hat{\mathbf{A}}$. ) This implies that, for numerical computations, the proper rescaling of (2.13),(2.14) and (2.16) is given by

$$(\mathbf{x}_1 - \mathbf{x}_0)\cdot\xi_c = \frac{1}{2}\alpha(\mathbf{x}_1\cdot\mathbf{x}_1 - \mathbf{x}_0\cdot\mathbf{x}_0) \tag{2.33a}$$

$$\hat{\mathbf{A}}\cdot\xi_c = -D \tag{2.33b}$$

$$\left(\hat{\mathbf{A}}\times(\mathbf{x}_1 - \mathbf{x}_0)\right)\cdot\xi_c = -\alpha Det\left\{\hat{\mathbf{A}}, \mathbf{x}_0, \mathbf{x}_1\right\}, \tag{2.33c}$$

where

$$\xi_c \equiv \alpha\mathbf{x}_c. \tag{2.34}$$

Solving the system (2.33) for $\mathbf{x}_c$, and using

$$\|\mathbf{x} - \mathbf{x}_c\| = \|\mathbf{x}_0 - \mathbf{x}_c\| = \|\mathbf{x}_1 - \mathbf{x}_c\| \tag{2.35}$$

for any point x along the ray, we obtain, after isolating leading order terms in $\alpha$, the straight-line limiting ray

$$\frac{(x - x_0)}{(x_1 - x_0)} = \frac{(y - y_0)}{(y_1 - y_0)} = \frac{(z - z_0)}{(z_1 - z_0)} \tag{2.36}$$

as $\alpha \rightarrow 0$.

Finally, we observe that the quantity $p = (\alpha R)^{-1}$ is, in fact, the usual seismic ray parameter, provided that angles of incidence of the ray are taken with respect to $\hat{\mathbf{A}}$.

## II.3 CALCULATION OF AMPLITUDES AND PHASE SHIFTS, AND DETECTION OF CAUSTICS

### II.3.1 Wavefront Curvatures Before Encounter With an Interface

For subsequent discussions of amplitude variations along a ray, we will need to know how the wavefront curvatures evolve. For simplicity, we consider first the two-dimensional problem of a symmetric point source at $\mathbf{x}_0 = (0, z_0)$ in the vertically stratified velocity field $V(z) = D + \alpha z$. The results will be extended to the more general case by appropriate translations and rotations of the coordinate system.

Let

$$\mathbf{x}_1 = (x_1, z_1) = (\epsilon \cos\psi, z_0 + \epsilon \sin\psi) \tag{3.1}$$

be a point on a given ray emerging from $\mathbf{x}_0$. By varying $\psi$ continuously over $(0, 2\pi)$, and choosing $\epsilon$ sufficiently small, we obtain the complete family of rays emanating from $\mathbf{x}_0$.

Solving (2.10) and (2.11), inserting the result in (2.9) (with $z$ replacing $y$), and, for convenience, passing to the limit $\epsilon \rightarrow 0$, we obtain, for $\mathbf{x} = (x, z)$ along a ray,

$$xF(\psi) = x^2 + z^2 + z_0^2 + 2\frac{D}{\alpha}(z + z_0), \tag{3.2}$$

where

$$F(\psi) = 2\left(z_0 + \frac{D}{\alpha}\right)\tan\psi. \tag{3.3}$$

Equations (3.2) and (3.3) define implicitly $\psi = \psi(x,z)$, the level curves of which are the rays. Because the wavefronts are orthogonal to the rays, we must have

$$\frac{dx}{dz} = \frac{\partial\psi/\partial x}{\partial\psi/\partial z} \tag{3.4}$$

along a wavefront.

As a consequence of (3.2), this condition reduces to the Bernoulli equation

$$\frac{dx}{dz} = \frac{1}{2}\left(z + \frac{D}{\alpha}\right)^{-1} x - \frac{1}{2}\left(z^2 - z_0^2 + 2\frac{D}{\alpha}(z - z_0)\right)x^{-1}, \tag{3.5}$$

which can be solved exactly to yield

$$x^2 + (z - z_0 - \tau)^2 = \tau^2 + 2\tau\left(z_0 + D/\alpha\right), \tag{3.6}$$

where $\tau \geq 0$ is a constant of integration. We have chosen to write the solution in this form because the integration constant $\tau$ has the appealing physical interpretation of arc-length along the vertical straight-line ray. The wavefronts are circles with increasing radii, whose centers propagate along the direction of stratification.

The result (3.6) can be extended to the three-dimensional linear velocity distribution $V(\mathbf{x}) = D + \alpha\hat{\mathbf{A}}\cdot\mathbf{x}$. For a source at $\mathbf{x}_0$, we obtain the radius of curvature, R of the spherical wavefront passing through $\mathbf{x}_1$ from

$$\left\|\alpha(\mathbf{x}_1 - \mathbf{x}_0) + (D - \alpha\tau')\hat{\mathbf{A}}\right\|^2 = \alpha^2\tau'^2 - V^2(\mathbf{x}_0) = \alpha^2\mathrm{R}^2. \tag{3.7}$$

### II.3.2 Wavefront Curvatures After Encounter With an Interface and Geometric Spreading Factors

Referring to Figure 1, we tentatively postpone a detailed discussion of how the wavefront curvatures are locally deformed by transmission through (or reflection from) an interface. Instead, we assume knowledge of $R_1'$ and $R_2'$, the principal radii of curvature of the wavefront immediately below the interface, and focus our attention on the subsequent evolution of the wavefront shape along ray segment II.

For this purpose, it is convenient to introduce the concept of 'equivalent point sources,' which we define in the following way. We conceptually remove the velocity discontinuity that separates regions I and II. Moreover, we continue the rule defining the velocity $V_2(\mathbf{x})$ back into region I.

We define the equivalent point source $\mathbf{p}_1$, situated on the extension of ray segment II, as that point from which an emitted wave (in this new, continuous medium) would arrive at the point $\mathbf{o}$ as a (spherical) wavefront of radius $R_1'$. Similarly, we define the equivalent point source $\mathbf{p}_2$, corresponding to wavefront radius $R_2'$ at $\mathbf{o}$.

Because the ray segment II is torsion-free, those cross-sectional curves, on successive wavefronts, which are of principal curvature, remain in the same two planes, $E_1$ and $E_2$, at least until the next discontinuity is encountered. It is this fact which allows us to reduce a discussion of the evolution of the local wavefront shape along ray segment II to two, conceptually decoupled, spherical wavefronts emanating from the equivalent point sources $\mathbf{p}_1$ and $\mathbf{p}_2$.

We now discuss briefly how the locations of $\mathbf{p}_1$ and $\mathbf{p}_2$ may be obtained. Because, in practice, we are able to confine our attention to the two planes $E_1$ and $E_2$ defined above, we limit this discussion to the two-dimensional case of velocity $V(z) = D + \alpha z$.

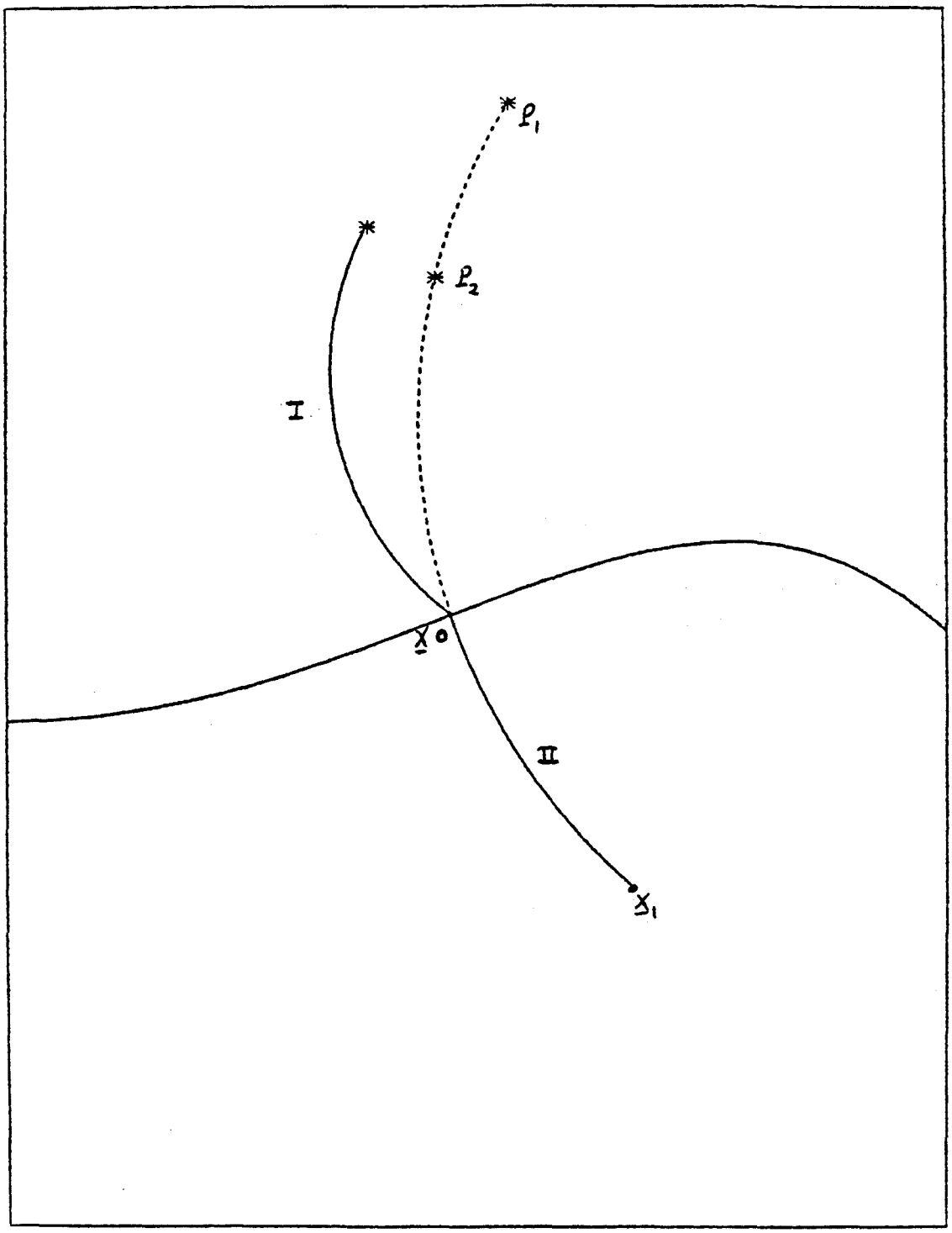


Figure 1

Consider the circular ray segment joining $\mathbf{x}_0 = (x_0, z_0)$ and $\mathbf{x}_1 = (x_1, z_1)$. We parametrize this ray, together with its extension, by

$$x(\theta) = x_c - R\cos\theta \tag{3.8a}$$

and

$$z(\theta) = z_c + R\sin\theta \tag{3.8b}$$

where, in view of our previous results, $x_c, z_c$, and $R$ may be regarded as known, and where $\theta_0$ and $\theta_1$ correspond, respectively, to $\mathbf{x}_0$ and $\mathbf{x}_1$.

We now determine the location $\mathbf{x}_p = \big(x(\theta_p), z(\theta_p)\big)$ of that point source from which such a ray segment must have emanated, subject to the requirement that the radius of curvature of the wavefront at $\mathbf{x}_0$ be R.

From the results of section II.3.1, the center of this wavefront will be at $\big(x(\theta_p), z(\theta_p) + \tau\big)$, for some value of $\tau \geq 0$, such that

$$\mathrm{R}^2 = \tau^2 + 2\tau\big(z(\theta_p) + D/\alpha\big). \tag{3.9}$$

An additional constraint on the two unknowns $\theta_p$ and $\tau$ is provided by

$$\mathrm{R}^2 = \big(x(\theta_0) - x(\theta_p)\big)^2 + \big(z(\theta_0) - z(\theta_p) - \tau\big)^2. \tag{3.10}$$

Upon substituting (3.8) into (3.9) and (3.10), and recalling

$$D + \alpha z_c = 0, \tag{3.11}$$

we obtain, after some manipulation,

$$\theta_p = \cos^{-1}\left(\cos\theta_0 - \frac{\mathrm{R}}{R}\sin\theta_0\right) \tag{3.12}$$

and

$$\tau = \frac{R}{\sin\theta_0}\big(1 - \cos(\theta_0 - \theta_p)\big). \tag{3.13}$$

Replacing R in (3.12) by $R_1'$ and $R_2'$ in turn, and using the results ($\theta_{p1}$ and $\theta_{p2}$, say), in (3.8), we obtain the locations of the equivalent point sources $p_1$ and $p_2$.

Moreover, once $\theta_{p1}$ and $\theta_{p2}$ have been evaluated, we may invert (3.12) to obtain the principal radii of curvature at any point along ray segment II, simply by effecting the appropriate substitution for $\theta_0$.

For example, at the endpoint $x_1$, we have

$$R_1'(x_1) = \frac{R}{\sin\theta_1}(\cos\theta_1 - \cos\theta_{p1}). \tag{3.14}$$

We next address the question of obtaining tractable expressions for the geometric spreading factor along ray segment II. Because our emphasis throughout is on ray tracing as a two-point boundary value problem, the expressions for a linearly stratified medium, which we obtain, are equivalent to, but differ considerably in appearance from, those formulae, in terms of initial take-off angles, that are usually given in the literature (see, e.g., Ben-Menahem and Singh, 1981).

Let $\delta\sigma_0$ be an infinitesimal area cross-section of the wavefront at o on the emergent side of the interface (see Fig. 1). Also, let $\delta\sigma_1$ be the corresponding infinitesimal cross-section of the wavefront at $x_1$, in the sense that those rays, nearby to ray segment II, which define the boundary of $\delta\sigma_0$, determine, as well, the boundary of $\delta\sigma_1$.

In the usual way, we define the geometric spreading factor $G$ along ray segment II by

$$G = \left(\frac{\delta\sigma_0}{\delta\sigma_1}\right)^{1/2}. \tag{3.15}$$

The geometric spreading $G$ may be expressed in terms of the principal radii of curvature of the wavefront along the ray by

$$G = \exp\left\{-\frac{1}{2}\int_{s_0}^{s_1}\left(\frac{1}{R_1'(s)} + \frac{1}{R_2'(s)}\right)ds\right\}, \tag{3.16}$$

where $s$ is arclength along the ray. (See, e.g., Červeny and Ravindra, 1971, p.32; also, Kline and Kay, 1965, pp. 184-186, for a more detailed derivation. )

Using (3.14) and the observation immediately preceding it, we can carry out exactly the integration in (3.16) to obtain, finally,

$$G = \left( \frac{\cos\theta_1 - \cos\theta_{p1}}{\cos\theta_0 - \cos\theta_{p1}} \right)^{1/2} \left( \frac{cos\theta_1 - cos\theta_{p2}}{cos\theta_0 - cos\theta_{p2}} \right)^{1/2}. \tag{3.17}$$

Each term in (3.17) represents the two-dimensional geometric spreading factor in the planes $E_1$ and $E_2$, respectively.

II.3.3 Effect On Wavefront Curvature By An Interface

Corresponding to the incident wavefront, the interface of discontinuity, and the emergent wavefront at a node $x_k$, we define the orthogonal triads of vectors $(\mathbf{T}^-, \mathbf{P}, \mathbf{Q}^-)$, $(\mathbf{N}, \mathbf{P}, \mathbf{Q}^0)$, and $(\mathbf{T}^+, \mathbf{P}, \mathbf{Q}^+)$. As before, $\mathbf{T}^-$ and $\mathbf{T}^+$ are unit tangents to the incident and emergent ray segments, while $\mathbf{N}$ is a unit normal to the interface.

Here we define

$$\mathbf{P} \equiv \mathbf{T}^- \times \mathbf{N} \tag{3.18a}$$

$$\mathbf{Q}^- \equiv \mathbf{T}^- \times \mathbf{P} \tag{3.18b}$$

$$\mathbf{Q}^0 \equiv \mathbf{N} \times \mathbf{P} \tag{3.18c}$$

$$\mathbf{Q}^+ \equiv \mathbf{T}^+ \times \mathbf{P}. \tag{3.18d}$$

The angles of incidence and refraction (or reflection) are given, respectively, by

$$\cos\theta_1 = \mathbf{T}^- \cdot \mathbf{N} \tag{3.19a}$$

$$\cos\theta_2 = \mathbf{T}^+ \cdot \mathbf{N}. \tag{3.19b}$$

We define the scalars

$$\mu \equiv \frac{V^+}{V^-} \quad \text{(ratio of velocities)} \tag{3.20}$$

$$\gamma \equiv \cos\theta_2 - \mu \cos\theta_1. \tag{3.21}$$

With the superscript $s \in (-, 0, +)$ denoting, respectively, the incident wavefront, the interface, and the emergent wavefront, we define:

$\mathbf{T}^s_\xi, \mathbf{T}^s_\eta \equiv$ principal unit direction vectors

$\mathrm{K}^s_\xi, \mathrm{K}^s_\eta \equiv$ principal surface curvatures

$\mathrm{K}^s_p, \mathrm{K}^s_q \equiv$ surface curvatures in the $\mathbf{P}$ and $\mathbf{Q}^s$ directions

$\sigma^s \equiv$ torsion of geodesic in $\mathbf{Q}^s$ direction

and

$$\theta^s \equiv \cos^{-1}(\mathbf{P}\cdot\mathbf{T}^s_\xi). \quad (3.22)$$

For the time being, we regard $K_\xi^{(-,0)}$, $K_\eta^{(-,0)}$ ,$\mathbf{T}_\xi^{(-,0)}$, and $\mathbf{T}_\eta^{(-,0)}$ as known. We wish to then determine $K_\xi^+$, $K_\eta^+$ ,$\mathbf{T}_\xi^+$, and $\mathbf{T}_\eta^+$.

Following Stavroudis (1976), we have, for $s \in (-,0)$, the auxiliary quantities

$$K_p^s = K_\xi^s \cos^2\theta^s + K_\eta^s \sin^2\theta^s \quad (3.23a)$$

$$K_q^s = K_\xi^s \sin^2\theta^s + K_\eta^s \cos^2\theta^s \quad (3.23b)$$

and

$$\sigma^s = \frac{1}{2}\left(K_\xi^s - K_\eta^s\right) \sin 2\theta^s. \quad (3.23c)$$

We transfer to the emergent ray ($s =$'+') through

$$K_p^+ = \mu K_p^- + \gamma K_p^0 \quad (3.24a)$$

$$K_q^+(\cos^2\theta_2) = \mu(\cos^2\theta_1)K_q^- + \gamma K_q^0 \quad (3.24b)$$

and

$$\sigma^+(\cos^2\theta_2) = \mu(\cos\theta_1)\sigma^- + \gamma\sigma^0. \quad (3.24c)$$

From Euler's Theorem it follows that

$$\tan 2\theta^+ = 2\sigma^+\left(K_q^+ - K_p^+\right). \quad (3.25)$$

Finally, we have

$$\mathbf{T}_\xi^+ = (\cos\theta^+)\mathbf{P} + (\sin\theta^+)\mathbf{Q}^+ \quad (3.26a)$$

$$\mathbf{T}_\eta^+ = (\sin\theta^+)\mathbf{P} - (\cos\theta^+)\mathbf{Q}^+ \quad (3.26b)$$

$$K_\xi^+ = (\cos^2\theta^+)K_p^+ + (\sin^2\theta^+)K_q^+ + (\sin 2\theta^+)\sigma^+ \quad (3.26c)$$

$$K_\eta^+ = (\sin^2\theta^+)K_p^+ + (\cos^2\theta^+)K_q^+ - (\sin 2\theta^+)\sigma^+, \quad (3.26d)$$

which were to be found.

From the results of the previous section, we know how $\mathrm{K}_\xi^+$ and $\mathrm{K}_\eta^+$, which are the reciprocals of the principal radii of wavefront curvature, evolve along the emergent ray segment. Consequently, we are able to evaluate the incident curvatures at the next interface (i.e., the $(k+1)^{st}$).

To evaluate $\mathbf{T}_\xi^-$ and $\mathbf{T}_\eta^-$ at the next interface, we use

$$\mathbf{T}_\xi^-(\mathrm{x}_{k+1}) = \mathbf{T}^-(\mathrm{x}_{k+1}) \times \left(\mathbf{T}_\xi^+(\mathrm{x}_k) \times \mathbf{T}^+(\mathrm{x}_k)\right) \tag{3.27a}$$

$$\mathbf{T}_\eta^-(\mathrm{x}_{k+1}) = \mathbf{T}^-(\mathrm{x}_{k+1}) \times \left(\mathbf{T}_\eta^+(\mathrm{x}_k) \times \mathbf{T}^+(\mathrm{x}_k)\right). \tag{3.27b}$$

These simple relations follow from the fact that an individual ray segment is torsion-free.

To evaluate the interface quantities $\mathrm{K}_\xi^0$, $\mathrm{K}_\eta^0$, $\mathbf{T}_\xi^0$ and $\mathbf{T}_\eta^0$, we make use of some well-known results from differential geometry. Following Stoker (1969), let the $k^{th}$ interface be given by $z = f(x,y)$. We define

$$E = 1 + \left(\partial f/\partial x\right)^2 \tag{3.28a}$$

$$F = \left(\partial f/\partial x\right)\left(\partial f/\partial y\right) \tag{3.28b}$$

$$G = 1 + \left(\partial f/\partial y\right)^2 \tag{3.28c}$$

$$L = \beta\left(\partial^2 f/\partial x^2\right) \tag{3.28d}$$

$$M = \beta\left(\partial^2 f/\partial x \partial y\right) \tag{3.28e}$$

$$N = \beta\left(\partial^2 f/\partial y^2\right) \tag{3.28f}$$

where

$$\beta = \left[1 + \left(\partial f/\partial x\right)^2 + \left(\partial f/\partial y\right)^2\right]^{-1/2}, \tag{3.29}$$

and where each partial derivative is understood to be evaluated at the node $\mathbf{x}_k = (x_k, y_k, f(x_k, y_k))$.

Let $\mathbf{P}_\xi = (p_\xi, q_\xi)^T$ and $\mathbf{P}_\eta = (p_\eta, q_\eta)^T$ be eigenvectors of the system

$$(\underline{\mathrm{A}} - \mathrm{K}\underline{\mathrm{B}})\mathbf{P} = \mathbf{0} \tag{3.30}$$

where

$$\underline{\mathrm{A}} \equiv \begin{pmatrix} L & M \\ M & N \end{pmatrix}, \tag{3.31a}$$

and

$$\underline{\mathrm{B}} \equiv \begin{pmatrix} E & F \\ F & G \end{pmatrix}. \tag{3.31b}$$

The corresponding eigenvalues $\mathrm{K}^0_\xi$ and $\mathrm{K}^0_\eta$ are the principal curvatures of the interface at $\mathbf{x}_k$, while the corresponding principal direction vectors are given by

$$\mathbf{T}^0_{\xi(\eta)} = \frac{p_{\xi(\eta)}\left(1, 0, \partial f/\partial x\right)^T + q_{\xi(\eta)}\left(0, 1, \partial f/\partial y\right)^T}{\left[p^2_{\xi(\eta)} + q^2_{\xi(\eta)} + \left(p^2_{\xi(\eta)}\partial f/\partial x + q^2_{\xi(\eta)}\partial f/\partial y\right)^2\right]^{1/2}}. \tag{3.32}$$

Given our restriction to interfaces that are twice continuously differentiable functions, the only pathological case occurs if $\mathrm{K}^0_\xi = \mathrm{K}^0_\eta$. In that case, every vector tangent to the interface at $\mathbf{x}_k$ is a principal direction vector, and we must choose two orthonormal ones. Otherwise, $\mathbf{T}^0_\xi$ and $\mathbf{T}^0_\eta$, as given by (3.32), are orthonormal (Stoker, p. 92,ff.).

### II.3.4 Reflection, Transmission and Conversion Coefficients

At an interface of discontinuity in elastic parameters, the incident wavefront splits into transmitted and reflected components. The dynamic boundary conditions (i.e., continuity in both particle displacement and traction across an interface) are sufficient to determine uniquely the amplitudes, relative to the amplitude of the incident wave, of each transmitted and reflected wave.

For completeness, we include explicit formulae, which are by now standard, for the calculation of these amplitude coefficients. (See, e.g., Keller, 1964, for a detailed derivation.)

Let $z = f_{i_k}(x, y)$ denote that interface on which the node $\mathbf{x}_k$ is located. We define the following unit vectors, each of which is tangent at $\mathbf{x}_k$ to the ray segment with which it is associated:

$$\mathbf{T}_k^{(0)} \equiv \text{incident ray (either P or S)}$$

$$\mathbf{T}_k^{(1)} \equiv \text{reflected P}$$

$$\mathbf{T}_k^{(2)} \equiv \text{reflected S}$$

$$\mathbf{T}_k^{(3)} \equiv \text{transmitted P}$$

$$\mathbf{T}_k^{(4)} \equiv \text{transmitted S.}$$

In each case, $\mathbf{T}_k^{(j)}$ has the sense defined by the direction of propagation of the wavefront with which it is associated. That is, $\mathbf{T}_k^{(0)}$, $\mathbf{T}_k^{(3)}$, and $\mathbf{T}_k^{(4)}$ point from the $k^{th}$ into the $(k+1)^{st}$ layer, whereas $\mathbf{T}_k^{(1)}$ and $\mathbf{T}_k^{(2)}$ have the opposite sense.

Let $\mathbf{N}$ denote a unit vector normal to the interface at $\mathbf{x}_k$, $T$ a unit tangent to the interface at $\mathbf{x}_k$ and in the plane determined by $\mathbf{N}$ and $\mathbf{T}_k^{(0)}$, and $\mathbf{B}$ a binormal unit vector orthogonal to both $\mathbf{N}$ and $T$. These are determined uniquely by the further requirements

$$\mathbf{N}\cdot\mathbf{T}_k^{(0)} > 0 \tag{3.33a}$$

$$T\cdot\mathbf{T}_k^{(0)} > 0 \tag{3.33b}$$

$$\mathbf{B} = \mathbf{N}\times T. \tag{3.33c}$$

Let $\mathbf{R}^{(j)}$ denote displacement vectors at $\mathbf{x}_k$ for each incident, transmitted, and reflected wave component. The identification of superscripts is identical here to that for $\mathbf{T}_k^{(j)}$. The $\mathbf{R}^{(j)}$ may be written as

$$\mathbf{R}^{(0)} = \begin{cases} a_0 \cos\theta_0\,\mathbf{N} + a_0 \sin\theta_0\,T & \text{incident P} \\ a_0 \sin\theta_0\,\mathbf{N} - a_0 \cos\theta_0\,T + b_0\,\mathbf{B} & \text{incident S} \end{cases} \tag{3.34a}$$

$$\mathbf{R}^{(1)} = a_1 \cos\theta_1\,\mathbf{N} + a_1 \sin\theta_1\,T \tag{3.34b}$$

$$\mathbf{R}^{(2)} = a_2 \sin\theta_2\,\mathbf{N} - a_2 \cos\theta_2\,T + b_2\,\mathbf{B} \tag{3.34c}$$

$$\mathbf{R}^{(3)} = a_3 \cos\theta_3\,\mathbf{N} + a_3 \sin\theta_3\,T \tag{3.34d}$$

$$\mathbf{R}^{(4)} = a_4 \sin\theta_4\,\mathbf{N} - a_4 \cos\theta_4\,T + b_4\,\mathbf{B}. \tag{3.34e}$$

Let $V^{(j)}$ denote the appropriate P or S wave speed, corresponding to each ray segment, evaluated at $\mathbf{x}_k$. These depend, through (1.3) of section II.1, on the densities and Lamé parameters $\rho_{k(k+1)}$, $\lambda_{k(k+1)}$, and $\mu_{k(k+1)}$ of the $k^{th}$ or $(k+1)^{st}$ layer. The angles $\theta_j$ are obtained from Snell's Law in the plane of incidence (defined by $\mathbf{N}$ and $T$ ) as follows:

$$V^{(0)} \sin\theta_j = V^{(j)} \sin\theta_0 \tag{3.35}$$

$$\cos\theta_j = \begin{cases} \pm\sqrt{1 - \sin^2\theta_j}\,, & -1 \le \sin\theta_j \le 1 \\ \mp i\sqrt{\sin^2\theta_j - 1}\,, & \sin\theta_j > 1 \end{cases}. \tag{3.36}$$

Here the first sign is chosen for $j = 1$ or $j = 2$, and the second for $j = 3$ or $j = 4$.

Following Keller (1964), the boundary conditions at the interface imply that

$$b_2 = -b_0\left(\frac{\mu_{k+1}}{V^{(4)}} \cos\theta_4 + \frac{\mu_k}{V^{(0)}} \cos\theta_0\right)\Delta^{-1} \tag{3.37a}$$

$$b_4 = b_0\left(\frac{\mu_k}{V^{(2)}} \cos\theta_2 + \frac{\mu_k}{V^{(0)}} \cos\theta_0\right)\Delta^{-1} \tag{3.37b}$$

where

$$\Delta = \frac{\mu_{k+1}}{V^{(4)}} \cos\theta_4 - \frac{\mu_k}{V^{(2)}} \cos\theta_2. \tag{3.38}$$

Also,

$$\begin{pmatrix} \cos\theta_1 & \sin\theta_2 & -\cos\theta_3 & -\sin\theta_4 \\ \sin\theta_1 & -\cos\theta_2 & -\sin\theta_3 & \cos\theta_4 \\ -\rho_k V^{(1)} \cos 2\theta_2 & -\rho_k V^{(2)} \sin 2\theta_2 & \rho_{k+1} V^{(3)} \cos 2\theta_4 & \rho_{k+1} V^{(4)} \sin 2\theta_4 \\ -\frac{\mu_k}{V^{(1)}} \sin 2\theta_2 & \rho_k V^{(2)} \cos 2\theta_2 & \frac{\mu_{k+1}}{V^{(3)}} \sin 2\theta_3 & -\rho_{k+1} V^{(4)} \cos 2\theta_4 \end{pmatrix}$$

$$\times \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = a_0 \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{pmatrix} \tag{3.39}$$

Here

$$\begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{pmatrix} = \begin{cases} \begin{cases} \left(\cos\theta_1, -\sin\theta_1, \rho_k V^{(1)} \cos 2\theta_2, -\frac{\mu_k}{V^{(1)}} \sin 2\theta_1\right)^T \\ \text{for incident P,} \end{cases} \\ \begin{cases} \left(-\sin\theta_2, -\cos\theta_2, \rho_k V^{(2)} \sin 2\theta_2, -\rho_k V^{(2)} \cos 2\theta_1\right)^T \\ \text{for incident S.} \end{cases} \end{cases} \tag{3.40}$$

It is worth noting that the formulae in this section represent the only occurrence in our description of the forward problem in which the density and Lamé parameters enter in combinations other than those defining the wave velocities. Consequently, any attempt to use our formulation of the forward problem to invert for these parameters must incorporate amplitude observations, and cannot rely entirely on travel-time observations.

### II.3.5 Detection of Caustics

A *caustic* is an envelope at which neighboring rays touch each other. For a single ray, we may determine whether or not a caustic has been encountered simply by monitoring the sense of the principal wavefront curvatures at the endpoints of each ray segment. For a given ray segment in a linearly stratified medium, there are three possibilities. If the sense of wavefront curvature is the same at both endpoints, then no caustic has been encountered by that ray segment. If one principal curvature changes sense, but the other does not, then exactly one caustic has been encountered. Finally, if both principal curvatures change sense, then two caustics have been encountered.

Because the passage of a wavefront through a caustic corresponds to a vanishing radius of wavefront curvature, the location of a caustic along a ray segment can be calculated by using the same geometric construction of an equivalent point source discussed in section II.3.2. Our program systematically computes these equivalent point sources for each ray segment as part of our algorithm for evaluating the geometric spreading factor. Consequently, to detect whether a caustic has occurred on a particular segment, we need only test whether the corresponding equivalent point sources lie on that segment, rather than on its artificial extension. No further computation is necessary to pin down the location of the caustic if one is detected.

Geometric ray theory is not valid in the immediate neighborhood of a caustic, where that theory predicts an infinite wave amplitude. Consequently, ray amplitudes at the receiver, which we calculate by the method outlined in the previous sections, will be incorrect if a caustic occurs within several wavelengths of either the receiver or an interface. Because, for a high- frequency signal, either occurrence is unlikely, our program detects those rays for which amplitude calculations are suspect, but does not attempt to improve upon them. Consequently, in subsequent

inversion algorithms, only the travel-times and phases of such rays should be used.

### II.3.6 Phase Shifts

Following Perozzi (1980), the total phase shift $\tilde{\phi}$ of the signal at the receiver, relative to the phase at the source, may be expressed as

$$\tilde{\phi} = \frac{\pi}{2} N_{caustic} + \pi N_{reflect} + \sum_{j=1}^{N} Im(\theta_j). \tag{3.41}$$

Here, $N_{caustic}$ is the number of caustics that the ray has encountered, $N_{reflect}$ is the number of nodes at which the ray has been reflected, and $Im(\theta_j)$ denotes the imaginary part of the splitting angle for the ray at the $j^{th}$ node, which are given in section II.3.4. These angles may be complex in the case of ray segments that are supercritically refracted.

## II.4. NUMERICAL METHODS FOR THE FORWARD PROBLEM

For the type of elastic media with which we are concerned, the forward ray problem is essentially solved when for each ray joining the source and receiver the nodal points $\mathbf{x}_k$, at which the ray hits interfaces of discontinuity in elastic parameters, are determined. Given all the $\mathbf{x}_k$ for one ray, we may determine the exact trajectory of the ray, the travel-time along it, the wave amplitude along the ray relative to the strength of the source, and the phase shift relative to the source signal. These quantities are computed in a straightforward way using the results of sections II.2 and II.3, and present no noteworthy numerical difficulties.

In this chapter, we present, in some detail, our numerical methods for determining the *a priori* unknown nodal points $\mathbf{x}_k$ for each ray.

### II.4.1 Nonlinear Algebraic System for the Nodal Points

At each nodal point $\mathbf{x}_k$ on surfaces of discontinuous velocity, a ray must satisfy Snell's vector law of refraction

$$V_k^+(\mathbf{N}_k \times \mathbf{T}_k^-) = V_k^-(\mathbf{N}_k \times \mathbf{T}_k^+). \tag{4.1}$$

Here $V_k^-$ denotes the velocity of the incident ray segment at $\mathbf{x}_k$, and $V_k^+$ that of the emergent ray segment. $N_k$ denotes any normal to the interface $z = f_{i_k}(x,y)$ evaluated at $\mathbf{x}_k$, and, for definiteness, we choose

$$\mathbf{N}_k = \left( -\frac{\partial f_{i_k}}{\partial x}, -\frac{\partial f_{i_k}}{\partial y}, 1 \right)^T. \tag{4.2}$$

$\mathbf{T}_k^- = (T_{k1}^-, T_{k2}^-, T_{k3}^-)^T$ denotes the unit tangent to the $k^{th}$ ray segment at $\mathbf{x}_k$, and can be obtained from

$$(\alpha_k \mathbf{x}_k - \xi_{c_k}) \cdot \mathbf{T}_k^- = 0 \tag{4.3a}$$

$$\hat{\mathbf{A}}_k \times (\alpha_k \mathbf{x}_k - \xi_{c_k}) \cdot \mathbf{T}_k^- = 0 \tag{4.3b}$$

and

$$\mathbf{T}_k^- \cdot \mathbf{T}_k^- = 1. \tag{4.3b}$$

Here $\xi_{c_k}$ is the normalized center of the $k^{th}$ ray segment, which can be obtained by the method of II.2.4. $\hat{\mathbf{A}}_k$ and $\alpha_k$ are the stratification direction and velocity gradient magnitude for the layer (either the $i_k^{th}$ or $(i_k+1)^{st}$) traversed by the $k^{th}$ ray segment. Equation (4.3a) is a statement that the tangent must be orthogonal to the radial vector, while (4.3b) ensures that the tangent vector is coplanar with the $k^{th}$ ray segment. One virtue of using our normalized centers $\xi_{c_k}$ is that the system (4.3) remains correct for computing $\mathbf{T}_k^-$ even when $\alpha_k \to 0$. (It is worth recalling here that $\xi_{c_k}$ depends implicitly on $\mathbf{x}_{k-1}$, $\mathbf{x}_k$, $\hat{\mathbf{A}}_k$, $\alpha_k$ and $D_k$.)

$\mathbf{T}_k^+$, the unit tangent at $\mathbf{x}_k$ to the $(k+1)^{st}$ ray segment, is the solution of (4.3) if $\xi_{c_k}$, $\hat{\mathbf{A}}_k$, and $\alpha_k$ are replaced by $\xi_{c_{k+1}}$, $\hat{\mathbf{A}}_{k+1}$, and $\alpha_{k+1}$, with evident definitions.

The ambiguity in sign intrinsic to the solution of (4.3) is resolved by requiring that both $\mathbf{T}_k^-$ and $\mathbf{T}_k^+$ point from the $k^{th}$ layer into the $(k+1)^{st}$. An exception to this convention occurs when the ray is reflected at the $k^{th}$ node, in which instance $\mathbf{T}_k^-$ and $\mathbf{T}_k^+$ are chosen with opposing sense.

Equating components in the vector equation (4.1) leads to a system of three scalar equations, of which any two represent independent constraints, but together imply the third, which may be discarded. Explicitly, for the first two components, we have

$$V_k^+\left(T_{k2}^- + T_{k3}^- \frac{\partial f_{i_k}}{\partial y}\right) = V_k^-\left(T_{k2}^+ + T_{k3}^+ \frac{\partial f_{i_k}}{\partial y}\right) \tag{4.4a}$$

and

$$V_k^+\left(T_{k1}^- + T_{k3}^- \frac{\partial f_{i_k}}{\partial x}\right) = V_k^-\left(T_{k1}^+ + T_{k3}^+ \frac{\partial f_{i_k}}{\partial x}\right). \tag{4.4b}$$

Equations (4.4) must hold at each nodal point $\mathbf{x}_k$, $k = 1, \ldots, N$, forming, collectively, a system of $2N$ nonlinear equations for the $2N$ unknown nodal coordinates $x_k$ and $y_k$, $k = 1, \ldots, N$. We denote this collective system as

$$\mathbf{F}(\mathbf{X}; \underline{\mathbf{M}}, \mathbf{x}_0, \mathbf{x}_{N+1}) = \mathbf{0}. \tag{4.5}$$

Here $\mathbf{X} = (x_1, y_1, \ldots, x_N, y_N)^T$ is the vector of unknown nodal point coordinates. $\underline{\mathbf{M}}$ is the collection of all parameters and functions which, together, describe the model geometry. Specifically, $\underline{\mathbf{M}}$ contains the stratification vector $\hat{\mathbf{A}}_k$ for the velocity associated with the $k^{th}$ ray segment, the background velocity $D_k^{(p,s)}$, the gradient magnitude $\alpha_k^{(p,s)}$, as well as the functions $z = f_{i_k}(x, y)$ for the interfaces. For reasons that will be made clear later, we associate $\hat{\mathbf{A}}_k$, $D_k^{(p,s)}$, and $\alpha_k^{(p,s)}$ with each ray segment, rather than with each layer.

The scalar entries of the vector equation (4.5) are given by

$$F_{2k-1} = V_k^+ \left(T_{k1}^- + T_{k3}^- \frac{\partial f_{i_k}}{\partial x}\right) - V_k^- \left(T_{k1}^+ + T_{k3}^+ \frac{\partial f_{i_k}}{\partial x}\right) \tag{4.6a}$$

and

$$F_{2k} = V_k^+ \left(T_{k2}^- + T_{k3}^- \frac{\partial f_{i_k}}{\partial y}\right) - V_k^- \left(T_{k2}^+ + T_{k3}^+ \frac{\partial f_{i_k}}{\partial y}\right) \tag{4.6b}$$

for $k = 1, \ldots, N$. These are analogous to equations 1.18b and 1.18a of Fawcett and, in fact, reduce to them in the case of a piecewise homogeneous medium (i.e., $\alpha_k^{(p,s)}$ all identically zero).

II.4.2 Method of Solution – Newton's Method

The nonlinear character of the system (4.5) precludes a direct solution algorithm. Instead, we solve this system by Newton's iterative method. Given a sufficiently close approximation $\mathbf{X}^{(0)}$ to the solution $\mathbf{X}$ of (4.5), the iteration

$$\mathbf{X}^{(n+1)} = \mathbf{X}^{(n)} + \delta\mathbf{X}^{(n)}, \tag{4.7}$$

where

$$\delta\mathbf{X}^{(n)} = -\underline{\mathbf{J}}^{(n)^{-1}}\,\mathbf{F}(\mathbf{X}^{(n)}; \underline{\mathbf{M}}, \mathbf{x}_0, \mathbf{x}_{N+1}), \tag{4.8}$$

converges quadratically to $\mathbf{X}$. Here

$$\underline{\mathbf{J}}^{(n)} \equiv \left.\frac{\partial \mathbf{F}}{\partial \mathbf{X}}\right|_{\mathbf{X}^{(n)}} \tag{4.9}$$

denotes the Jacobian of $\mathbf{F}$ with respect to $\mathbf{X}$, evaluated at $\mathbf{X}^{(n)}$.

If $\mathbf{X}^{(0)}$ is not sufficiently close to $\mathbf{X}$, Newton's method may converge less rapidly, or even fail to converge at all. In subsequent sections, we describe how to obtain, in a systematic way, a sufficiently good estimate for $\mathbf{X}^{(0)}$ to ensure convergence to a solution in many geometries. In this section, however, we assume that such an estimate is already available.

It is easy to show that the Jacobian matrix $\underline{\mathbf{J}}$, of the system (4.5) is banded, with bandwidth 7. This is so because the $(2k-1)^{st}$ and $2k^{th}$ entries of (4.5) depend only on the six unknowns $x_{k-1}$, $y_{k-1}$, $x_k$, $y_k$, $x_{k+1}$ and $y_{k+1}$.

Because we do not display the explicit form of this dependence, a few comments about the algorithm used to compute $\underline{\mathbf{J}}$ are in order.

The spatial derivatives of $V_k^-$, $V_k^+$, and $N_k$ are readily obtained from (1.4) and (4.2), and present no difficulty.

The variations $\partial\mathbf{T}^-/\partial\mathbf{x}_{k-1}$, $\partial\mathbf{T}^-/\partial\mathbf{x}_k$, $\partial\mathbf{T}^+/\partial\mathbf{x}_k$, and $\partial\mathbf{T}^+/\partial\mathbf{x}_{k+1}$ can be evaluated in the following way. From equation (4.3), it is clear that these depend on

the variations of $\xi_{c_k}$ and $\xi_{c_{k+1}}$. Identifying $\mathbf{x}_0$, $\mathbf{x}_1$, $\hat{\mathbf{A}}$, and $\alpha$ of equations (2.3) with $\mathbf{x}_{k-1}$, $\mathbf{x}_k$, $\hat{\mathbf{A}}_k$, and $\alpha_k$, respectively, we obtain a linear system for $\xi_{c_k}$. Differentiating these equations in turn with respect to each of the components of $\mathbf{x}_{k-1}$ and $\mathbf{x}_k$, we obtain six linear systems for each of the columns of $\partial\xi_{c_k}/\partial\mathbf{x}_{k-1}$ and $\partial\xi_{c_k}/\partial\mathbf{x}_k$ . Moreover, the matrix of coefficients for each of these systems is identical to that used to compute $\xi_{c_k}$ in the first place, so that the variations with respect to $\mathbf{x}_{k-1}$ and $\mathbf{x}_k$ can be conveniently computed at the same time as $\xi_{c_k}$, with only one inversion of a $3 \times 3$ matrix for all seven vectors. We observe here that the third columns of $\partial\xi_{c_k}/\partial\mathbf{x}_{k-1}$ and $\partial\xi_{c_k}/\partial\mathbf{x}_k$ , i.e., those corresponding to derivatives with respect to $z_{k-1}$ and $z_k$, are needed in constructing $\underline{\mathbf{J}}$ because the variations, with respect to $x_{k-1}$, $y_{k-1}$, $x_k$, and $y_k$, are not 'free,' but are subject to the subsidiary constraints that $\mathbf{x}_{k-1}$ and $\mathbf{x}_k$ remain on their respective interfaces. The variations $\partial\xi_{c_{k+1}}/\partial\mathbf{x}_k$ and $\partial\xi_{c_{k+1}}/\partial\mathbf{x}_{k+1}$ can be similarly evaluated.

Formally differentiating the system (4.3) with respect to each of $\mathbf{x}_{k-1}$ and $\mathbf{x}_k$, we obtain linear systems for each of the columns of $\partial\mathbf{T}_k^-/\partial\mathbf{x}_{k-1}$ and $\partial\mathbf{T}_k^-/\partial\mathbf{x}_k$ , and, as in the case of the calculation of the variations of $\xi_{c_k}$, the matrix of coefficients for each of these systems is identical to that used to calculate $\mathbf{T}_k^-$. The variations $\partial\mathbf{T}_k^+/\partial\mathbf{x}_k$ and $\partial\mathbf{T}_k^+/\partial\mathbf{x}_{k+1}$ can be evaluated in the same way, requiring only one additional inversion or factorization of a $3 \times 3$ matrix (which is available from the computation of $\mathbf{T}_k^+$) and one post-multiplication of that matrix by another that is $3 \times 6$.

This essentially completes our discussion of the evaluation of all the subsidiary ingredients needed to construct the Jacobian. There is still a singular case (when $(\mathbf{x}_k - \mathbf{x}_{k-1})\cdot\hat{\mathbf{A}}_k = 0$) which must be dealt with. We postpone a discussion of this case until section II.4.8 on normal incidence rays because it is in the calculation of such rays that this situation is most likely to occur.

II.4.3 Selection of Initial Guess $\mathbf{X}^{(0)}$ – Some General Considerations

Following Perozzi (1980), we say that two rays are of the same *class* if the sequence of interfaces $(i_1, i_2, \ldots, i_k, \ldots, i_N)$, which each of the rays encounters, is identical. Because the wave along each ray segment, in a given class, may be either P or S, there are, in general, $2^{N+1}$ rays in each class for fixed source and receiver.

It is convenient, for our purposes, to formally subdivide all possible classes of rays into two *Types*. We say that a particular class of rays is of Type I, if

$$i_k \neq i_{k+1}, \quad \text{for } \textit{all } k = 0, \ldots, N.$$

On the other hand, we say that a class of rays is of Type II, if

$$i_k = i_{k+1}, \quad \text{for } \textit{some } k = 0, \ldots, N.$$

We make this formal subdivision of ray classes into two Types because our algorithm for obtaining sufficiently good approximations to the nodal points $\mathbf{X} = (x_1, y_1, \ldots, x_N, y_N)$ so that Newton's method will converge to a solution is fundamentally different for each.

Figure 2a illustrates a possible ray, of Type I, which has five internal nodes and is in the class (1,2,1,2,1). Figure 2b shows, for the same model geometry, a ray of Type II, with six internal nodes, in the class (1,2,2,1,1,1). Qualitatively, rays of Type I are those for which each ray segment completely traverses the layer with which it is associated. By contrast, rays of Type II are those for which at least one ray segment is completely refracted within a layer. Such ray segments are said to have a *turning point.*

For both types of ray classes, our solution procedure can be qualitatively described in the following way. We choose a model geometry $\underline{\mathbf{M}}_0$, which has the following properties. First, $\underline{\mathbf{M}}_0$ can be deformed, in a continuous way, to the model $\underline{\mathbf{M}}_1$, for which we ultimately wish to solve the ray problem. Second, the model $\underline{\mathbf{M}}_0$

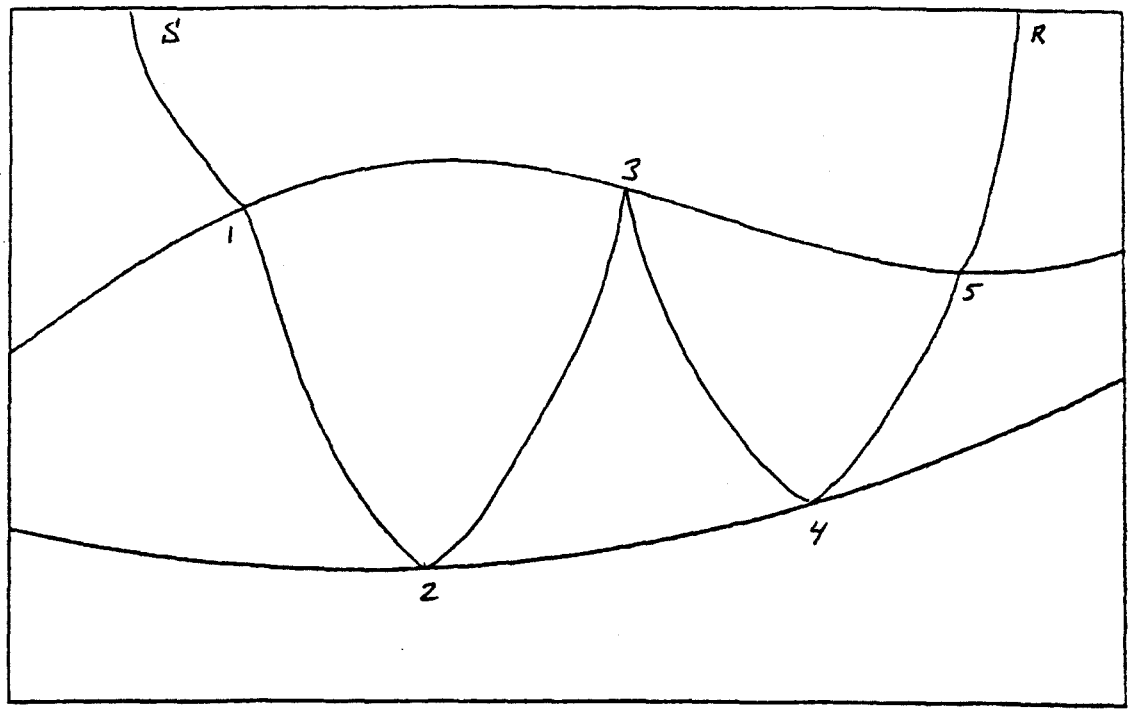


Figure 2a

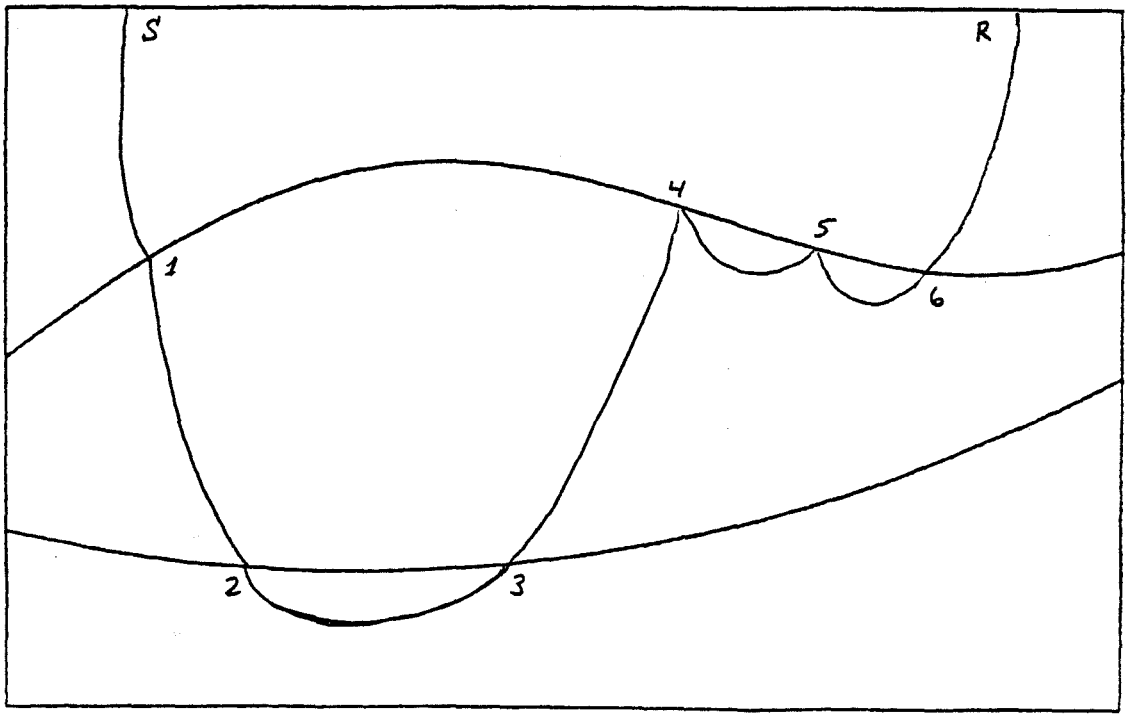


Figure 2b

admits at least one ray in the class of rays we want to compute. Finally, such a ray in $\underline{\mathbf{M}}_0$ can easily be found.

We define the one-parameter family of model geometries

$$\underline{\mathbf{M}}(\lambda) \equiv (1-\lambda)\underline{\mathbf{M}}_0 + \lambda\underline{\mathbf{M}}_1, \quad 0 \leq \lambda \leq 1, \tag{4.10}$$

where, clearly, $\underline{\mathbf{M}}(0) = \underline{\mathbf{M}}_0$ and $\underline{\mathbf{M}}(1) = \underline{\mathbf{M}}_1$. Instead of the problem (4.5), we consider the family of problems

$$\mathbf{F}(\mathbf{X}(\lambda); \underline{\mathbf{M}}(\lambda), \mathbf{x}_0, \mathbf{x}_{N+1}) = \mathbf{0}, \quad 0 \leq \lambda \leq 1, \tag{4.11}$$

where, by assumption, (4.11) is easy to solve for $\lambda = 0$, but it is, in fact, the case $\lambda = 1$ in which we are interested.

It is hoped that, as the model is continuously deformed, the solution vector $\mathbf{X}(\lambda)$ will also trace out a continuous curve in $\mathbf{R}^{2N}$. If this is, in fact, the case, then we may continue in the parameter $\lambda$ from $\mathbf{X}(0)$ to $\mathbf{X}(1)$.

Formally differentiating (4.11) with respect to $\lambda$, we have

$$\frac{d\mathbf{F}}{d\lambda} = \frac{\partial \mathbf{F}}{\partial \mathbf{X}} \frac{d\mathbf{X}}{d\lambda} + \frac{\partial \mathbf{F}}{\partial \underline{\mathbf{M}}} \frac{d\underline{\mathbf{M}}}{d\lambda} = \mathbf{0}, \tag{4.12}$$

and therefore, provided only that $\partial \mathbf{F}/\partial \mathbf{X}$ is not singular,

$$\frac{d\mathbf{X}}{d\lambda} = -\left(\frac{\partial \mathbf{F}}{\partial \mathbf{X}}\right)^{-1} \frac{\partial \mathbf{F}}{\partial \underline{\mathbf{M}}} \frac{d\underline{\mathbf{M}}}{d\lambda}. \tag{4.13}$$

Consequently,

$$\mathbf{X}(\lambda + \delta\lambda) = \mathbf{X}(\lambda) - \left(\frac{\partial \mathbf{F}}{\partial \mathbf{X}}\right)^{-1} \frac{\partial \mathbf{F}}{\partial \underline{\mathbf{M}}} \frac{d\underline{\mathbf{M}}}{d\lambda} \delta\lambda + O(\delta\lambda^2). \tag{4.14}$$

For $\delta\lambda$ sufficiently small, the first two terms of the right hand side provide a reasonable approximation to $\mathbf{X}(\lambda + \delta\lambda)$. In particular, for $\delta\lambda$ small enough, this approximation to $\mathbf{X}(\lambda + \delta\lambda)$ is sufficiently close to provide a first guess for Newton's method to converge to the exact solution of

$$\mathbf{F}(\mathbf{X}(\lambda + \delta\lambda); \underline{\mathbf{M}}(\lambda + \delta\lambda), \mathbf{x}_0, \mathbf{x}_{N+1}) = \mathbf{0}. \tag{4.15}$$

The Newton iterations are performed for the *fixed* model $\underline{\mathbf{M}}(\lambda+\delta\lambda)$. The computed value of the Jacobian **J** for the final iteration can be identified as

$$\mathbf{J} = \left.\frac{\partial \mathbf{F}}{\partial \mathbf{X}}\right|_{\lambda+\delta\lambda}, \tag{4.16}$$

for which the L-U decomposition is available to compute the next continuation step (4.14).

### II.4.4 Initial Guess for Rays of Type I.

For rays of Type I, we choose our starting model $\underline{\mathbf{M}}_0$ in the following way. The interface functions $z = f_i(x,y)$ are replaced by parallel plane interfaces $z = c_i$, where $c_i$ is an approximation, which may be quite coarse, to the mean value of $f_i(x,y)$ over the region of interest. For the velocity within each layer, we choose the homogeneous background velocity $D_i$. Given the ray endpoints $\mathbf{x}_0$ and $\mathbf{x}_{N+1}$, as well as the sequence $(i_1,\ldots,i_N)$ of interfaces that the ray encounters, the problem of finding the ray joining $\mathbf{x}_0$ and $\mathbf{x}_{N+1}$ in this simple model $\underline{\mathbf{M}}_0$ can be reduced to one scalar equation in one unknown:

$$r(\theta_0) \equiv \sum_{k=1}^{N+1} \frac{|c_{i_k} - c_{i_{k-1}}| \frac{D_k}{D_0} \sin\theta_0}{\sqrt{1 - \frac{D_k^2}{D_0^2}\sin^2\theta_0}} = \sqrt{(x_{N+1}-x_0)^2 + (y_{N+1}-y_0)^2}, \tag{4.17}$$

where $c_{i_0} \equiv z_0$, $c_{i_{N+1}} \equiv z_{N+1}$, and $\theta_0$ is the initial take-off angle of the ray measured with respect to the vertical. Since $r(\theta_0)$ is monotonic, there is no difficulty in numerically solving (4.17). Fawcett (1983) provides an efficient and robust algorithm.

For this relatively simple case, the family of model geometries $\underline{\mathbf{M}}(\lambda)$, which we use in our continuation algorithm, is given by

$$f_{i_k}(x,y;\lambda) = (1-\lambda)c_{i_k} + \lambda f_{i_k}(x,y) \tag{4.18}$$

and

$$V_k(\mathbf{x};\lambda) = D_k + \lambda\alpha_k \hat{\mathbf{A}}_k \cdot \mathbf{x}, \tag{4.19}$$

where $0 \leq \lambda \leq 1$ and $k = 1, 2, \ldots, N+1$.

### II.4.5 Initial Guess for Rays of Type II.

With the exception of critically refracted rays, the simple model $\underline{\mathbf{M}}_0$ consisting of homogeneous layers separated by plane interfaces, which we introduced in the previous section, does not admit rays of Type II. Critically refracted rays do not provide a satisfactory first guess for our continuation algorithms because, in the sense of preserving the class of a ray, they are unstable for even infinitesimally small perturbations in the model. We must choose a different starting model $\underline{\mathbf{M}}_0$ to compute rays of Type II. As before, we replace the interface functions $z = f_i(x, y)$ by their mean values, $c_i$, over the domain of interest. We now construct simultaneously a model geometry $\underline{\mathbf{M}}_0$ as well as a ray, which satisfies the ray equations in $\underline{\mathbf{M}}_0$, of any given class of Type II. To illustrate the procedure, we discuss a concrete example that exhibits all the salient features. Suppose that we are solving for rays of the class (1,2,2,2,1,1,), as shown in Figure 3a. First, we coalesce successive nodes that lie on the same interface and, simultaneously, move the receiver closer to the source by some horizontal distance $d$ (Figure 3b). The choice of $d$ is essentially arbitrary, provided only that $0 < d < \|\mathbf{x}_{N+1} - \mathbf{x}_0\|$. In practice, we often make the *ad hoc* selection of $d = \frac{1}{2}\|\mathbf{x}_{N+1} - \mathbf{x}_0\|$. By coalescing nodes in this fashion, we temporarily deflate the class of the ray to (1,2,1), *which is of Type I.* We solve this deflated problem using (4.17), with the right hand side decreased by $d$, using as before, homogeneous background velocities $D_k$ for each ray segment. Next, we split up the $m$ nodes ($m = 3$ in our example), which were coalesced, by equal horizontal distances $d/m$, translating , where necessary, the ray segments that have just been calculated, but otherwise leaving them intact (Figure 3c).

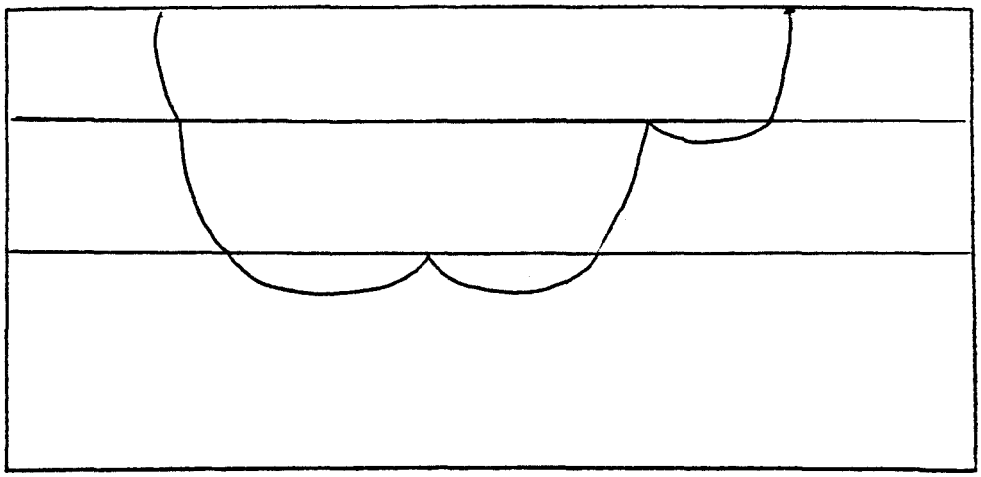

Figure 3a

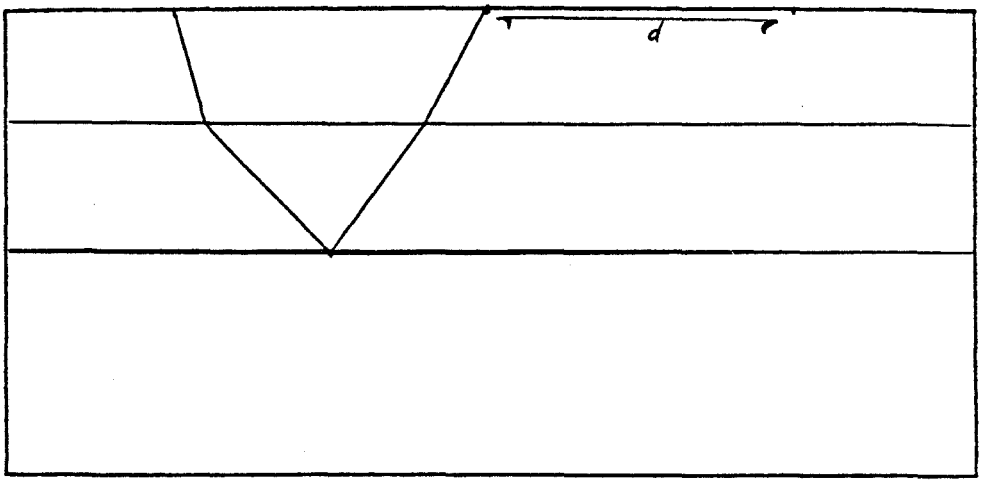


Figure 3b

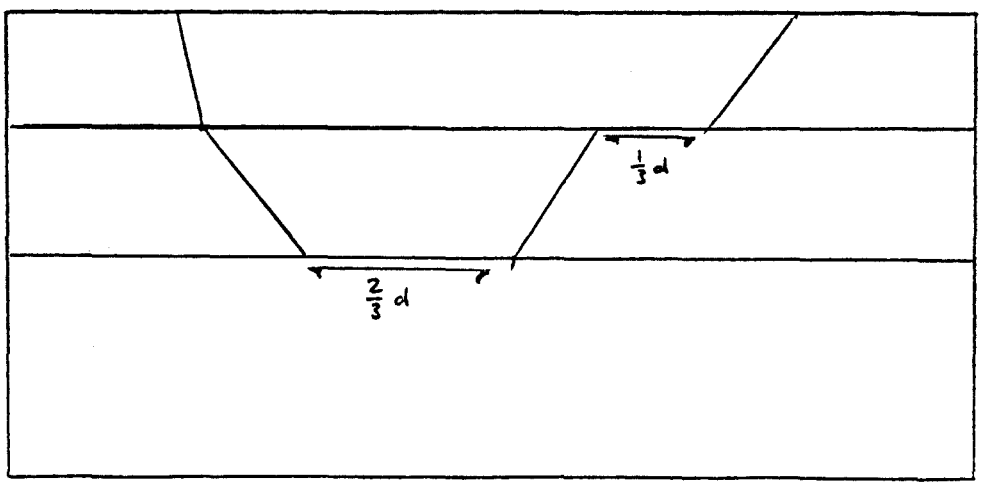


Figure 3c

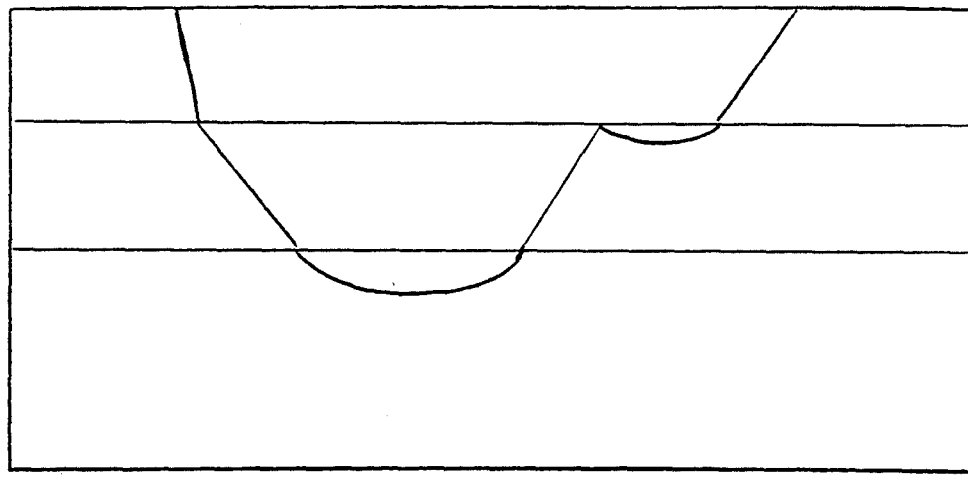

Figure 3d

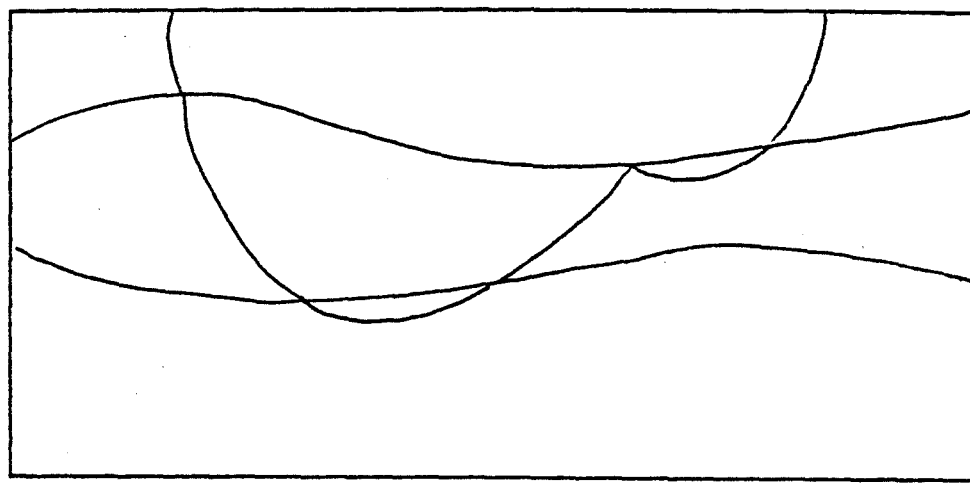

Figure 3e

We must next associate velocities with each of the 'missing' segments ( i.e., those that are to have turning points). This can be done in the following way. If the $k^{th}$ ray segment is one of those with a turning point, we *choose* this segment to be a circular arc of radius $R_k$, where $c_{i_k} < R_k < c_{i_k+1}$. We recall that, for a horizontally stratified linear medium for which $\hat{\mathbf{A}} = (0,0,1)^T$ , the centers of all the rays lie in the plane $z = -D/\alpha$. In order that the ray segment, which we have constructed, satisfies the ray equations for a continuous medium, we must choose for this ratio the value

$$D/\alpha = R_k - c_{i_k}. \tag{4.20}$$

From elementary geometry, we know that the tangent at the node $\mathbf{x}_{k-1}$ (and $\mathbf{x}_k$) forms the angle $\theta_k^+$ with the vertical, where

$$\theta_k^+ = \tan^{-1}\left(\frac{2m}{d}R_k\right). \tag{4.21}$$

Consequently, Snell's law will be satisfied across the node $\mathbf{x}_{k-1}$ provided that we associate with the $k^{th}$ ray segment the velocity

$$V_k(z) = b_k(z + R_k - c_{i_k}) \tag{4.22}$$

where

$$b_k = \frac{2D_{k-1}}{\sin\theta_k^- \sqrt{\left(\frac{d}{m}\right)^2 + 4R_k^2}}. \tag{4.23}$$

We observe that in the above construction, the velocity must be chosen to vary only in $z$, so that Snell's law may be satisfied simultaneously across both $\mathbf{x}_{k-1}$ and $\mathbf{x}_k$. Furthermore, if two or more *successive* segments have turning points (e.g., the third and fourth segments in our example ), then although for the first of these $R_k$ may be chosen arbitrarily the same value must be retained for the subsequent segments in the succession.

A glance at Figure 3d shows why we insist on associating a velocity function with each ray segment rather than with each layer. Although the second, fifth and

sixth ray segments all lie within the same layer, our construction clearly associates a different velocity function with the sixth segment than with the second and fifth in the starting model $\underline{\mathbf{M}}_0$. In fact, in our continuation algorithm, the velocity functions for these three segments would not coincide until the final model $\underline{\mathbf{M}}_1$ has been attained.

Instead of the simple linear variation (4.19) in the continuation parameter $\lambda$, the family of velocities which we use for rays of type II has the more general form

$$V_k(\mathbf{x};\lambda) = D_k(\lambda) + \alpha_k(\lambda)\hat{\mathbf{A}}_k(\lambda)\cdot\mathbf{x} \tag{4.24}$$

where

$$D_k(\lambda) = (1-\lambda)D_k(0) + \lambda D_k(1), \tag{4.25}$$

$$\hat{\mathbf{A}}_k(\lambda) = \frac{(1-\lambda)\alpha_k(0)\hat{\mathbf{A}}_k(0) + \lambda\alpha_k(1)\hat{\mathbf{A}}_k(1)}{\alpha_k(\lambda)}, \tag{4.26}$$

and

$$\alpha_k(\lambda) = [\lambda^2\alpha_k^2(1) + (1-\lambda)^2\alpha_k^2(0) + 2\lambda(1-\lambda)\alpha_k(0)\alpha_k(1)\hat{\mathbf{A}}_k(0)\cdot\hat{\mathbf{A}}_k(1)]^{\frac{1}{2}} \tag{4.27}$$

To complete our discussion of the continuation algorithm used to find a ray of a particular class, we present formulae for the variation of endpoint tangent vectors for a single ray segment with fixed endpoints as we continue in velocity. For the unit tangent vector $\mathbf{T}$ evaluated at $\mathbf{x}$ on the $k^{th}$ ray segment, we have

$$\frac{\partial \mathbf{T}}{\partial \lambda} = -\underline{\mathbf{C}}_1^{-1}\frac{\partial \underline{\mathbf{C}}_1(\lambda)}{\partial \lambda}\tau \tag{4.28}$$

where

$$\mathbf{C}_1(\lambda) \quad \begin{pmatrix} (\alpha_k(\lambda)\mathbf{x} - \xi_{c_k}(\lambda))^T \\ (\hat{\mathbf{A}}_k(\lambda) \times (\alpha_k(\lambda)\mathbf{x} \quad \xi_{c_k}(\lambda)))^T \\ \mathbf{T}^T(\lambda) \end{pmatrix}. \tag{4.29}$$

The derivatives, with respect to $\lambda$ of $\alpha_k(\lambda)$, $\hat{\mathbf{A}}_k(\lambda)$ and $D_k(\lambda)$, are evident from (4.25), (4.26), and (4.27), while those of $\xi_{c_k}(\lambda)$ can be found from

$$\frac{\partial \xi_{c_k}(\lambda)}{\partial \lambda} = \underline{\mathbf{C}}_2^{-1}(\lambda)\left(\frac{\partial \mathbf{f}}{\partial \lambda} - \frac{\partial \underline{\mathbf{C}}_2(\lambda)}{\partial \lambda}\xi_{c_k}(\lambda)\right) \tag{4.30}$$

where

$$\underline{\mathbf{C}}_2(\lambda) = \begin{pmatrix} (\mathbf{x}_k - \mathbf{x}_{k-1})^T \\ (\hat{\mathbf{A}}_k(\lambda))^T \\ (\hat{\mathbf{A}}_k(\lambda) \times (\mathbf{x}_k - \mathbf{x}_{k-1}))^T \end{pmatrix} \tag{4.31}$$

and

$$\mathbf{f}(\lambda) = \begin{pmatrix} \frac{1}{2}(\mathbf{x}_k \cdot \mathbf{x}_k - \mathbf{x}_{k-1} \cdot \mathbf{x}_{k-1}) \\ -D_k(\lambda) \\ \alpha_k(\lambda)\hat{\mathbf{A}}_k(\lambda) \cdot (\mathbf{x}_k \times \mathbf{x}_{k-1}) \end{pmatrix}. \tag{4.32}$$

### II.4.6 Some Combinatorics

The two mutually exclusive types of ray classes which we have introduced together exhaust all possible ray classes in the geometries with which we are concerned. A question that naturally arises is how many ray classes there are for rays with up to some maximum number of nodes and, of this total, how many are of each of the two types.

As the combinatorics are relatively straight forward, we do not belabor the details of the derivation and present only the results.

For a source on the $m^{th}$ interface with the receiver on the free surface ($m = 0$), the number of possible ray classes with $N$ or fewer internal nodes is

$$\sum_{n=1}^{N} \sum_{r=0}^{N-n} \frac{m+1}{n+2} C_r^{n+r+1} C^{n+2}_{\frac{1}{2}(n+m+3)}. \tag{4.33}$$

Here $C_q^p$ is the binomial coefficient, provided $p$ and $q$ are non-negative integers with $p \geq q$. We adopt the convention (not the usual) that $C_q^p \equiv 0$ if $q$ is not an integer or if $q > p$.

For a source between the $m^{th}$ and $(m+1)^{st}$ interfaces, the corresponding formula is

$$\sum_{n=0}^{N} \sum_{r=0}^{N-n} C_r^{n+r} \left( \frac{m+1}{n+1} C^{n+1}_{\frac{1}{2}(n+m+2)} + \frac{m+2}{n+1} C^{n+1}_{\frac{1}{2}(n+m+3)} \right). \tag{4.34}$$

In both cases, the terms with $r = 0$ correspond to rays of type I, and all the rest correspond to rays of type II. If all possible rays in a given class are to be

included (accounting for all P-S conversions), then an additional factor of $2^{n+1}$ is needed for each term in the sum.

The number of ray classes grows very rapidly with $N$. For a surface source and $N = 7$, there are already 530 possible ray classes. Of these, only 22 are of type I, and the rest of type II. Typically, for a specific geometry, only a very small fraction of these possible classes are represented by actual physical rays. Unfortunately, there appears to be no *a priori* way to ascertain which classes will be represented by actual solutions, and an exhaustive search is necessary. Fortunately, our ray tracing algorithm is quite fast requiring, when it does in fact converge, only one or two continuation steps for Newton's method to converge quadratically. In virtually all our test examples, if two continuation steps were not sufficient for convergence, increasing the number of continuation steps still further did not remedy the situation. In such cases, we concluded that a solution of the class and type sought did not exist.

### II.4.7 Continuation Along a Gather of Receivers

In seismic prospecting experiments, the response to a man-made source is typically recorded simultaneously in an array (commonly refered to as a 'gather') of different receivers. We describe in this section an algorithm to quickly obtain rays of the same class at other receivers in a gather after one has been computed by the methods described previously. The method is particularly simple when all the receivers in the gather are located along a line segment on the free surface (a situation that often occurs in exploration practice), but it is applicable to more general gathers as well.

Suppose that a solution $\tilde{\mathbf{x}}$ has been computed for

$$\mathbf{F}(\tilde{\mathbf{x}}; \underline{\mathbf{M}}, \mathbf{x}_S, \tilde{\mathbf{x}}_R) = \mathbf{0} \tag{4.35}$$

for one receiver location $\tilde{\mathbf{x}}_R$ in the gather. We observe that only the entries $F_{2N-1}$ and $F_{2N}$ depend explicitly on $\tilde{\mathbf{x}}_R$ (see equations (4.6)). In particular, only $\mathbf{T}_N^+$, the unit tangent vector to the $(N+1)^{st}$ ray segment evaluated at the $N^{th}$ unknown node, depends on $\tilde{\mathbf{x}}_R$. We wish to find solutions to

$$\mathbf{F}(\mathbf{x}; \underline{\mathbf{M}}, \mathbf{x}_S, \tilde{\mathbf{x}}_R + \beta) = \mathbf{0} \tag{4.36}$$

for some displacement $\beta$ of the receiver position.

We consider the one-parameter family of problems

$$\mathbf{F}(\mathbf{x}(\lambda); \underline{\mathbf{M}}, \mathbf{x}_S, \mathbf{x}_R = \tilde{\mathbf{x}}_R + \lambda\beta) = \mathbf{0} \tag{4.37}$$

for $0 \leq \lambda \leq 1$. By differentiating with respect to $\lambda$ we obtain the continuation step

$$d\mathbf{x} = -d\lambda\, \underline{\mathbf{J}}^{-1} \frac{\partial \mathbf{F}}{\partial \mathbf{x}_R} \beta. \tag{4.38}$$

Here, $\frac{\partial \mathbf{F}}{\partial \mathbf{x}_r}$ is a matrix of dimension $(2N \times 3)$. However, only its bottom two rows are non-zero. Of these non-zero elements, even the calculation of the last column can often be dispensed with because the receivers in the gather will typically be confined to the free surface $z = 0$.

For $\lambda$ small or, alternatively, for $||\beta||$ sufficiently small, $\mathbf{x}^{(0)} = \tilde{\mathbf{x}} + \lambda \frac{d\mathbf{x}}{d\lambda}$ provides a satisfactory guess for Newton's method to converge.

In the special case of a linear gather, appropriate choices of $\lambda$ allow us to compute rays at each receiver location, quite often without updating the value of $\frac{d\mathbf{F}}{d\mathbf{x}_R}$.

II.4.8 Normal Incidence Rays

Besides the two-point problem of finding rays joining a known source and receiver pair, the problem of finding normal incidence rays is also of interest to geophysicists (see, e.g., Kleyn, 1983). These are rays which impinge normally to an interface of discontinuity. Consequently, the reflected ray which is of the same type (P or S) as the incident one tracks back to the source along the same ray path.

The way in which we have organized our calculation for the two-point ray tracing problem, enables us to compute normal incidence rays with relatively little extra work. To adhere most closely to the conventions established in previous sections, we shall refer to the final node at which such a ray hits a normal interface as the 'receiver' $\mathbf{x}_{N+1}$, retaining our convention of N internal nodes between source and receiver.

Instead of 2N unknowns, as in the fixed endpoint problem, there are now 2N+2. The conditions (4.1) still hold for the internal nodes $\mathbf{x}_1, \ldots, \mathbf{x}_N$. However, rather than specifying $\mathbf{x}_{N+1}$, we must solve for it subject to

$$\mathbf{T}^-_{N+1} \times \mathbf{N}_{N+1} = \mathbf{0}. \tag{4.39}$$

This vector equation nominally represents three scalar ones, but, as in the case (4.1), only two of these provide independent constraints. Specifically, we augment the system (4.6) with

$$f_{2N+1} \equiv T^-_{(N+1)1} + T^-_{(N+1)3} \frac{\partial f_{i_{(N+1)}}}{\partial x_{N+1}} = 0 \tag{4.40a}$$

and

$$f_{2N+2} \equiv T^-_{(N+1)2} + T^-_{(N+1)3} \frac{\partial f_{i_{(N+1)}}}{\partial y_{N+1}} = 0. \tag{4.40b}$$

If a sufficiently good guess for the unknowns were available, we would be able to solve this expanded system using Newton's method. Indeed, not only does the

$(2N + 2) \times (2N + 2)$ Jacobian for this problem possess the same banded structure as that for the fixed endpoint problem, but the upper left $2N \times 2N$ block is *exactly* the same.

An obvious model $\underline{\mathbf{M}}_0$ to use to start off the calculation is one that is horizontally stratified with parallel plane interfaces because the solution to the normal incidence problem in such a geometry is trivial, consisting evidently of vertical straight-line ray segments. In principle, the continuation in interface shapes and velocities proceeds exactly as in the two-point problem. However, it is in the very first continuation step that we encounter the difficulty mentioned at the end of II.4.2.

Our representation of the variations with respect to the endpoints of the unit tangent vectors to the ray segments breaks down in the case of the straight-line ray segment when $(\mathbf{x}_1 - \mathbf{x}_0)$ is parallel to $\hat{\mathbf{A}}$. In this case, even the normalized ray segment center $\|\xi_c\| \rightarrow \infty$, and the numerical computation cannot be salvaged by a simple rescaling (as was the case when $\alpha \rightarrow 0$, but $|\hat{\mathbf{A}} \cdot (\mathbf{x}_1 - \mathbf{x}_0)| \geq \epsilon > 0$).

Actually evaluating the endpoint tangents in this special case is trivial; clearly $\mathbf{T}[\mathbf{x}_0] = \mathbf{T}[\mathbf{x}_1] = \pm\hat{\mathbf{A}}$, and for computational purposes this value is almost always sufficiently close for the nearly singular case in which $0 < |(\mathbf{x}_1 - \mathbf{x}_0) \cdot \hat{\mathbf{A}}| << 1$. Hence, the only instance in which the system (4.3) becomes singular is also the only case in which we know the solution *apriori*.

Unfortunately, this is not the case for the $3 \times 3$ matrices $\frac{\partial \mathbf{T}[\mathbf{x}_0]}{\partial \mathbf{x}_0}$, $\frac{\partial \mathbf{T}[\mathbf{x}_0]}{\partial \mathbf{x}_1}$, etc. Our experience shows that for the two-point ray tracing problem with given source and receiver this is almost never a serious problem as the breakdown of the numerical algorithm occurs only very close to the singular case. However, in the case of computing normal incidence rays, the singular case is the starting point for our continuation algorithm. Evaluating the variations of the tangent vectors in the singular case can be dealt with in a straightforward, although rather cumbersome

way. We focus our discussion on calculating $\frac{\partial \mathbf{T}[\mathbf{x}_0]}{\partial \mathbf{x}_0}$ in the case that $(\mathbf{x}_1 - \mathbf{x}_0)$ is parallel to $\hat{\mathbf{A}}$. The other three differ only in detail, but not in the approach.

The columns of $\frac{\partial \mathbf{T}[\mathbf{x}_0]}{\partial \mathbf{x}_0}$, correspond to the directional derivatives of $\mathbf{T}[\mathbf{x}_0]$ along each of the principal unit vectors. Clearly then, what must be calculated is

$$\lim_{\beta \to 0} \frac{\mathbf{T}[\mathbf{x}_0 + \beta \hat{\mathbf{B}}] - \mathbf{T}[\mathbf{x}_0]}{\beta} \tag{4.41}$$

for $\hat{\mathbf{B}} = \hat{\mathbf{e}}_j, j = 1,2,3$. We adopt the notation $\frac{\partial \mathbf{T}}{\partial \hat{\mathbf{B}}}$ for the above vector limit. So that no confusion will result, we emphasize that $\frac{\partial \mathbf{T}}{\partial \hat{\mathbf{B}}}$ is a 3-vector, whereas $\frac{\partial \mathbf{T}[\mathbf{x}_0]}{\partial \mathbf{x}_0}$ is a $3 \times 3$ matrix. The direct calculation of these limits using the results of previous sections is rather involved. However, the following observations serve to simplify the calculation considerably. First of all, it is clear that $\frac{\partial \mathbf{T}}{\partial \hat{\mathbf{B}}}$ is in the plane of $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ (or $\mathbf{0}$ if $\hat{\mathbf{A}} = \hat{\mathbf{B}}$). Moreover,

$$\mathbf{T}[\mathbf{x}_0 + \beta \hat{\mathbf{B}}] = \hat{\mathbf{A}} + \beta \frac{\partial \mathbf{T}}{\partial \hat{\mathbf{B}}} + O(\beta^2), \tag{4.42}$$

together with the normalization condition $\|\mathbf{T}\| = 1$, implies that $\frac{\partial \mathbf{T}}{\partial \hat{\mathbf{B}}} \cdot \hat{\mathbf{A}} = \mathbf{0}$. Consequently, $\frac{\partial \mathbf{T}}{\partial \hat{\mathbf{B}}} = \gamma \hat{\mathbf{C}}$ where $\hat{\mathbf{C}} = \hat{\mathbf{A}} \times (\hat{\mathbf{A}} \times \hat{\mathbf{B}})$, and only the scalar $\gamma$ has to be found. Finally, if any two columns of $\frac{\partial \mathbf{T}}{\partial \mathbf{x}_0}$ have been found, the third can be evaluated from $\hat{\mathbf{A}} \cdot \nabla T_j = 0, j = 1,2,3$, which follows from the observation that $T$ is stationary to perturbations of $\underline{\mathrm{x}}_0$ in the $\hat{\mathbf{A}}$ direction. This last point has a more fundamental importance than merely that of somewhat reducing the algebra. In the case that $\hat{\mathbf{A}}$ almost coincides with one of the $\mathbf{e}_j$ (as in a medium which is nearly horizontally stratified), then the systems which must be solved for two of the columns of $\frac{\partial \mathbf{T}}{\partial \mathbf{x}_0}$ are extremely well conditioned, whereas the third is nearly singular.

This is most clearly seen if we rewrite the system (2.33) for $\xi_c$ (for the ray segment with endpoints at $\mathbf{x}_0 + \beta \hat{\mathbf{B}}$ and $\mathbf{x}_1$) as

$$\hat{\mathbf{A}} \cdot \xi_c = -D \tag{4.43a}$$

$$\hat{\mathbf{B}} \cdot \xi_c = -\frac{1}{\beta}\left(\frac{\alpha}{2}(\mathbf{x}_1 \cdot \mathbf{x}_1 - \mathbf{x}_0 \cdot \mathbf{x}_0) + \frac{qD}{\beta}\right) + \alpha \mathbf{x}_0 \cdot \hat{\mathbf{B}} + \frac{\alpha\beta}{2} \tag{4.43b}$$

$$(\hat{\mathbf{A}} \times \hat{\mathbf{B}}) \cdot \xi_c = -\frac{1}{\beta}(\alpha \hat{\mathbf{A}} \cdot (\mathbf{x}_1 \times \mathbf{x}_0)) + \alpha \hat{\mathbf{A}} \cdot (\mathbf{x}_1 \times \hat{\mathbf{B}}). \tag{4.43c}$$

Here $q \equiv \|\mathbf{x}_1 - \mathbf{x}_0\|$, while everything else has already been defined.

It follows immediately that $\xi_c$ has the form

$$\xi_c = \frac{1}{\beta}\xi_c^{(-1)} + \xi_c^{(0)} + \beta\xi_c^{(+1)}. \tag{4.44}$$

It has already been established that $\mathbf{T}[\mathbf{x}_0 + \beta\hat{\mathbf{B}}]$ has the form

$$\mathbf{T} = \hat{\mathbf{A}} + \beta\gamma\hat{\mathbf{C}} + O(\beta^2), \tag{4.45}$$

where $\gamma$ is to be found. From the first of ( ) there follows

$$\left(\frac{1}{\beta}\xi_c^{(-1)} + (\xi_c^{(0)} - \alpha\mathbf{x}_0) + \beta(\alpha\hat{\mathbf{B}} - \xi_c^{(+1)})\right) \cdot \left(\hat{\mathbf{A}} + \beta\gamma\hat{\mathbf{C}} + O(\beta^2)\right) = 0. \tag{4.46}$$

Equating coefficients of the different powers of $\beta$, we obtain

$$\gamma = \frac{\hat{\mathbf{C}} \cdot \xi_c^{(-1)}}{\hat{\mathbf{A}} \cdot (\alpha\mathbf{x}_0 - \xi_c^{(0)})}, \tag{4.47}$$

which is needed.

## II.5. NUMERICAL EXAMPLES

In this section we present some computed examples which illustrate both the capabilities as well as some of the limitations of our numerical algorithms.

In the numerical examples presented here, we have assumed that each layer is a Poisson-type solid for which $V_p/V_s = \sqrt{3}$ (see, e.g., Bullen, 1963). This is not a restriction on our code, which admits arbitrary input models. However, besides being a fairly realistic approximation, our choice was motivated by the fact that in such a model, pure shear wave rays and pure compressional wave rays coincide, a circumstance which in itself provides a consistency check for our code.

The two-dimensional examples were generated using our three-dimensional software package, and were chosen as such only because they are easier to illustrate graphically.

### II.5.1 Phantom Interfaces and Regions of Convergence

In our first and simplest example we consider a half-space in which the velocity varies linearly with depth. However, the input model to our ray-tracing program includes several 'phantom' interfaces, which the program treats as actual discontinuities, but across which the velocity is, in fact, continuous.

The purpose of this numerical experiment is two-fold. First, even this simple example (for which we can solve exactly the two-point ray-tracing problem) serves to test substantial parts of our code. Second, we obtain empirical estimates for the radius (in $R^{2N}$) of the ball of convergence (centered at the solution) for Newton's method. We do so in the following way. After a solution has been obtained, we perturb it by vectors whose magnitudes are fixed, but whose directions are generated by a random number simulator. For each fixed magnitude, we use 100 such vector perturbations. If all 100 perturbed 'guesses' yielded convergent Newton iterations,

we assumed that we were still within the ball of convergence, and repeated the experiment with an augmented perturbation magnitude until at least one of the attempts failed to converge.

The estimate for the radius of convergence that we obtain in this way is not rigorous, but serves, nevertheless, as a strict upper bound. For several different velocity profiles and interface shapes, we found that, if a typical scale length of the geometry is on the order of 100, then the radii of convergence are only on the order of 1 to 2 (based on 4 unknown nodes). This result underscores the importance of our continuation algorithms for obtaining sufficiently good initial 'guesses.'

### II.5.2 Continuation in Receiver Location and Class Transitions

Our second example tests our code on a more complex Earth model, implements our algorithm for continuation in receiver location, and illustrates a simple bifurcation phenomenon that may occur under a variety of circumstances. The velocity model used is described by interfaces (including the free surface) given by

$$f_0(x) = 0, \tag{5.1a}$$

$$f_1(x) = 15 + 3\sin\left(\frac{2\pi x}{150}\right), \tag{5.1b}$$

$$f_2(x) = 30 - 3\sin\left(\frac{x}{15}\right), \tag{5.1c}$$

$$f_3(x) = 45 + 4\sin\left(\frac{x}{40} - 1\right), \tag{5.1d}$$

and

$$f_4(x) = 60, \tag{5.1e}$$

as well as velocity functions

$$V_1^p(\mathrm{x}) = -0.2x + z + 100, \tag{5.2a}$$

$$V_2^p(\mathrm{x}) = 0.2x + 3z + 100, \tag{5.2b}$$

$$V_3^P(\mathbf{x}) = 0.2x + 10z - 100, \tag{5.2c}$$

and

$$V_4^P(\mathbf{x}) = 0.5z + 200. \tag{5.2d}$$

Since a Poisson-type solid is assumed, $V_k^S(\mathbf{x}) = V_k^P(\mathbf{x})/\sqrt{3}, \quad k = 1, \ldots, 4.$

We have tested our code on a large number of different velocity models, and a detailed presentation of all the results would be superfluous. We have chosen this particular, perhaps somewhat strange-looking, model as an illustration because it has proved to be one of the most lucrative that we have found in terms of admitting at least one representative ray for each of many different classes of rays, as will be shown in II.5.3.

For the purposes of this section, we limit our attention to the class (1,2,3,2,1), as shown in Figures 4a-c. In each case, the source is located at x=10, while the receiver is at x=60, x=80, and x=100, respectively, for the three cases. All 16 possible rays of the class arrive at the nearest receiver (Figure 4a). As we continue in receiver location ( see section II.4.7 ) to x=80, only four rays of the class persist, while for the receiver at x=100 (Figure 4c), only one such ray remains. It is clear that, for this simple example, the 'disappearance' of rays represents a transition to the deflated class (1,2,2,1). For our continuation algorithms, this type of class deflation to one that is simpler by one level is a canonical bifurcation which can occur not only as we continue in receiver location, but also as we either increase the magnitude of the velocity gradient or deform the interface. In all three cases, the onset of the transition can be detected by monitoring the scalar product of the incident ray tangent and the normal to the interface.

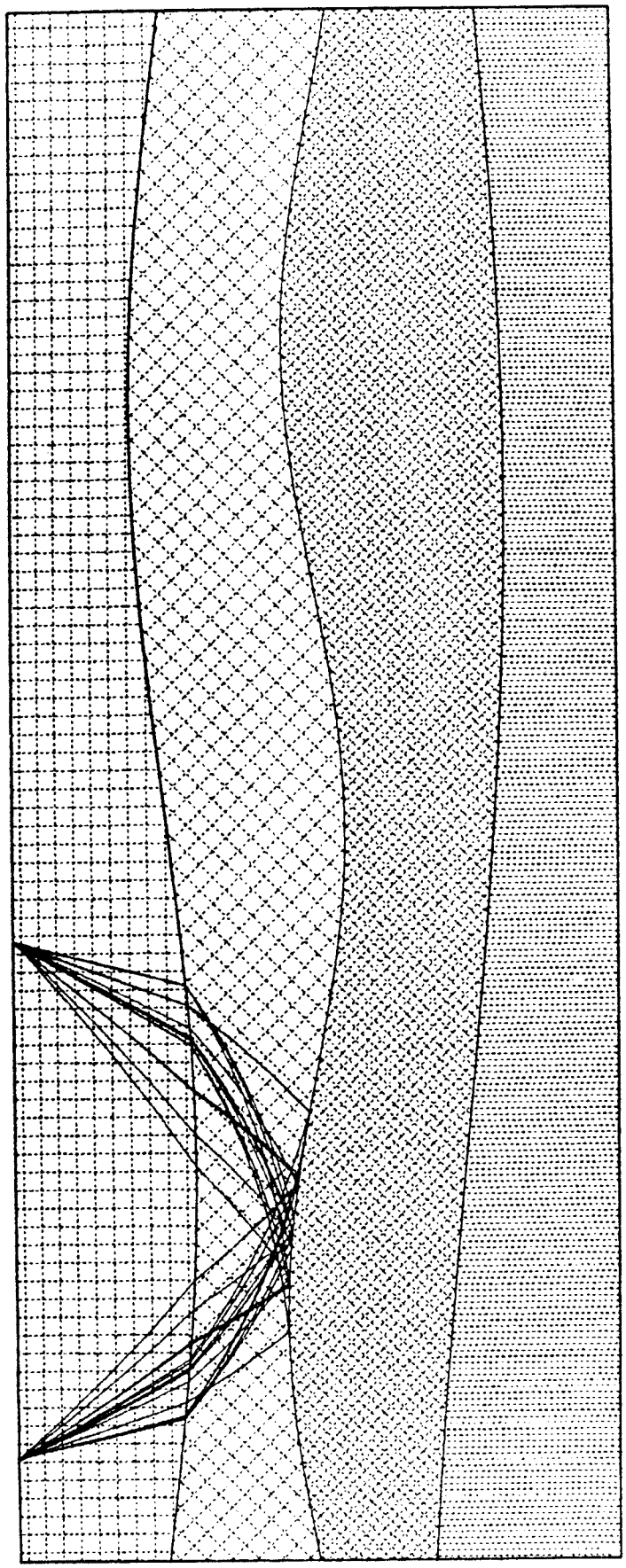

Figure 4a

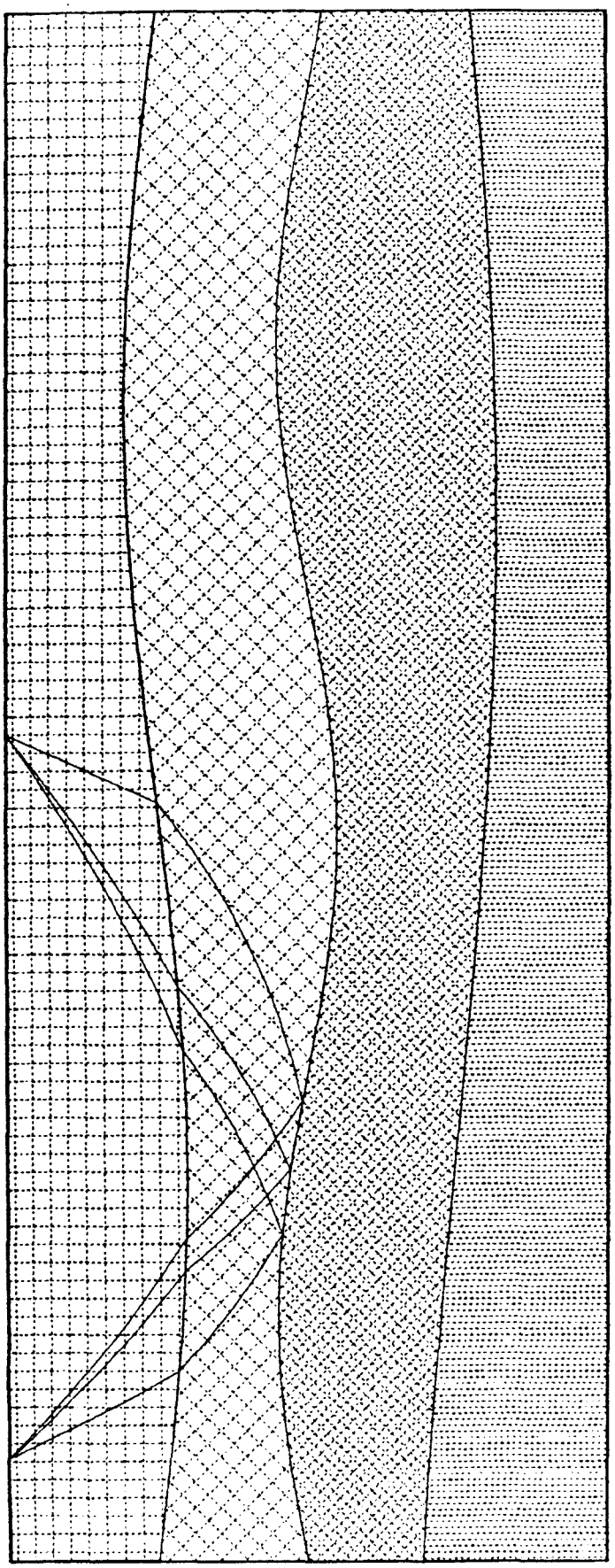

Figure 4b

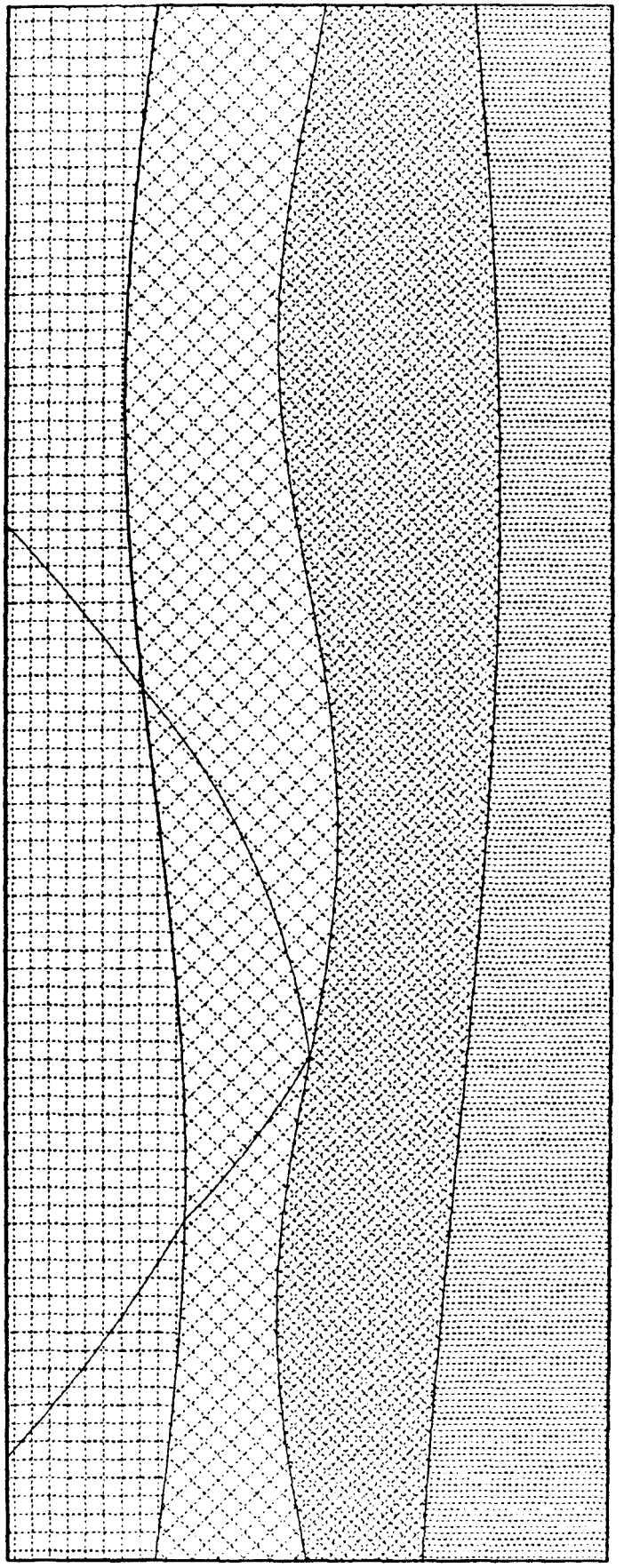

Figure 4c

### III.5.2 A Search For All Rays

In this example, we consider again the velocity model used in the previous example. This time, our purpose is to compute all the rays, with up to four unknown nodes, between a surface source ( at x=10) and a fixed receiver (at x= 80). The results are depicted in Figures 5a-5k, with each figure presenting all the rays for an individual class that were actually found. Our search was exhaustive in that all possible P-S combinations for all possible classes with four or fewer nodes were sought. All the rays depicted were found using at most two continuation steps ($\delta\lambda = 1$ or $\delta\lambda = \frac{1}{2}$). A follow up search using 32 continuation steps successfully reproduced each ray, but failed to reveal any others.

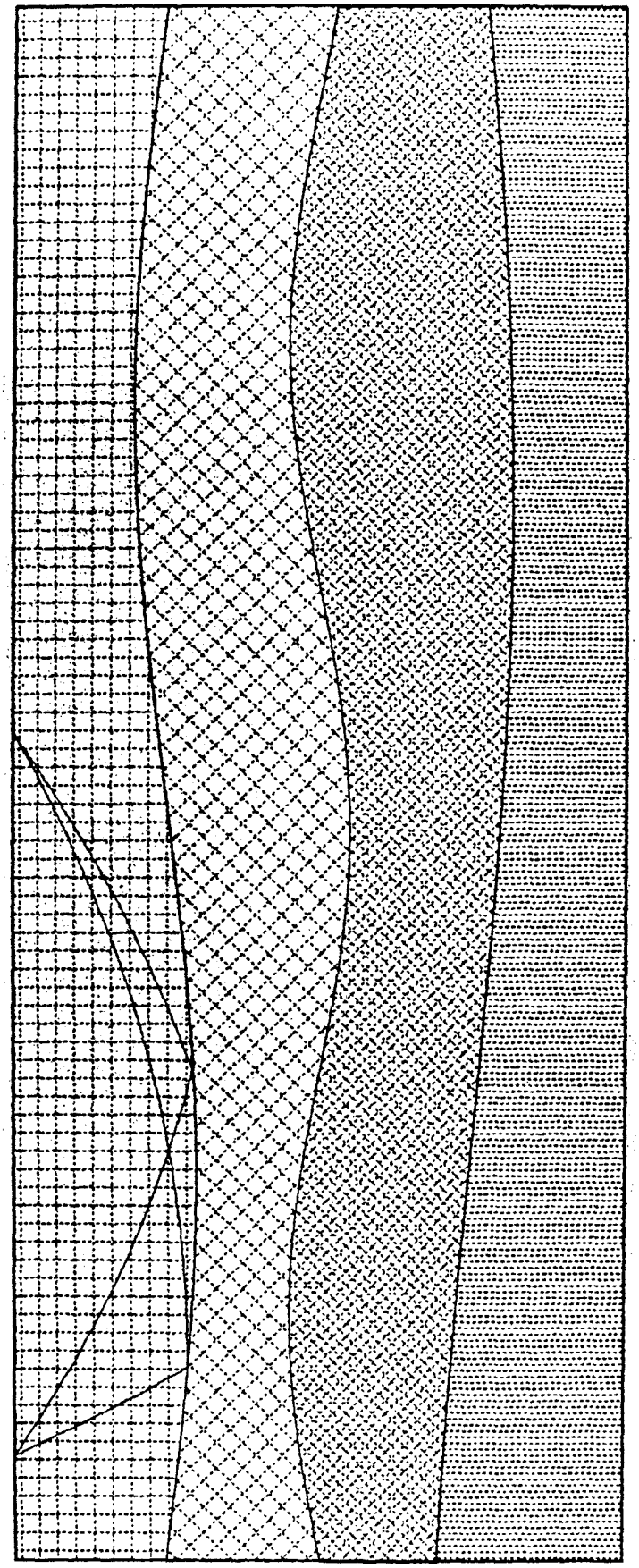

Figure 5a

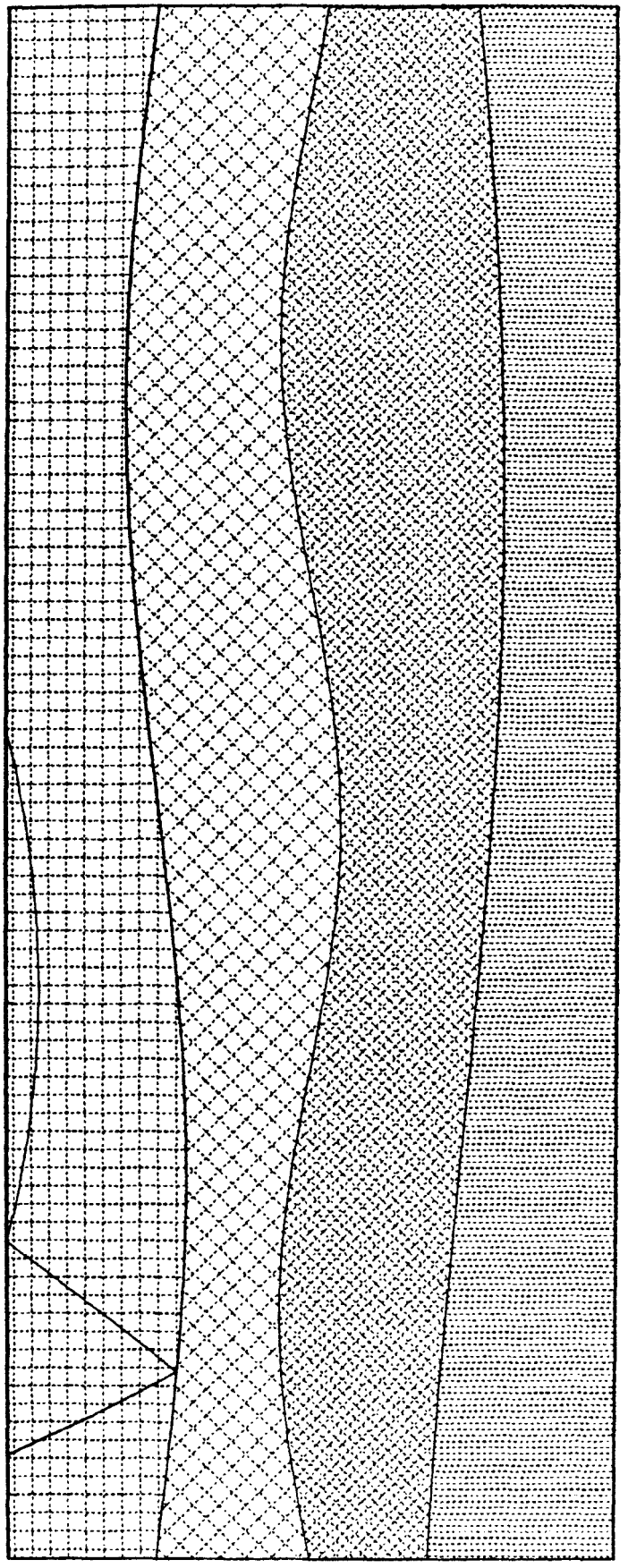

Figure 5b

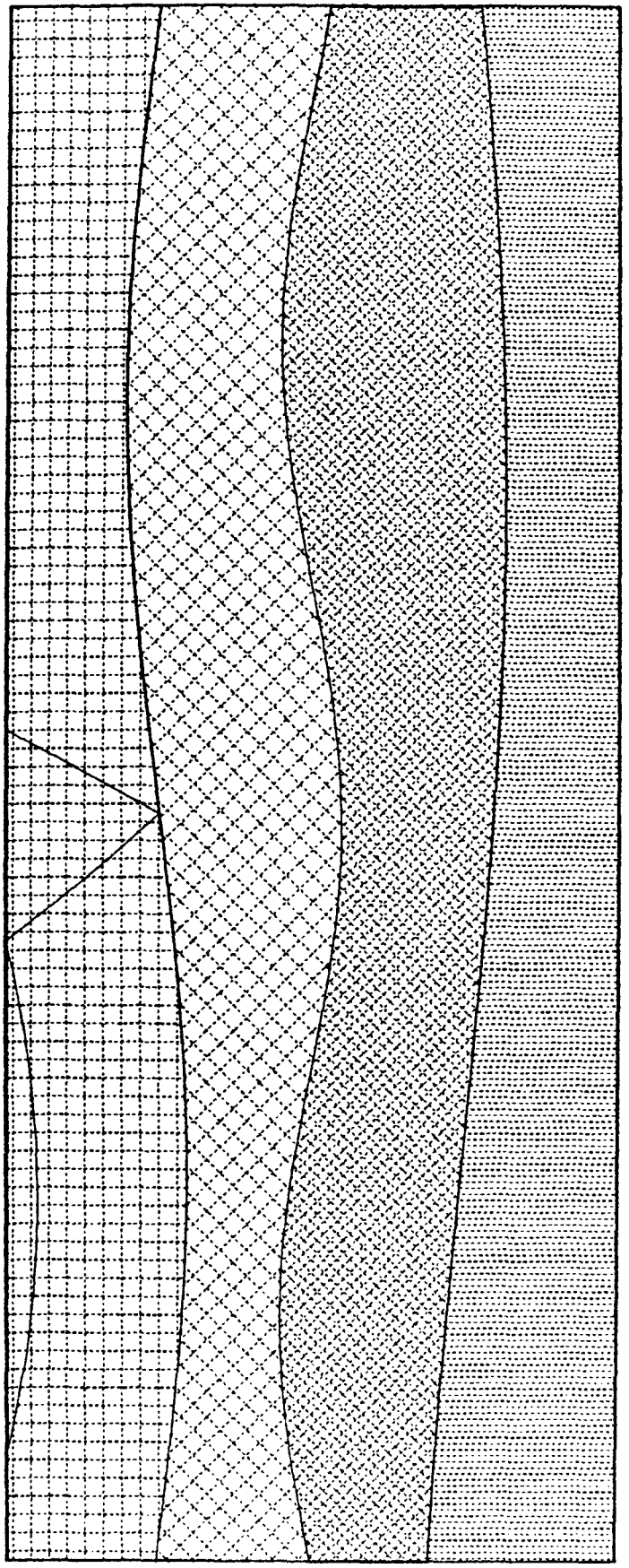

Figure 5c

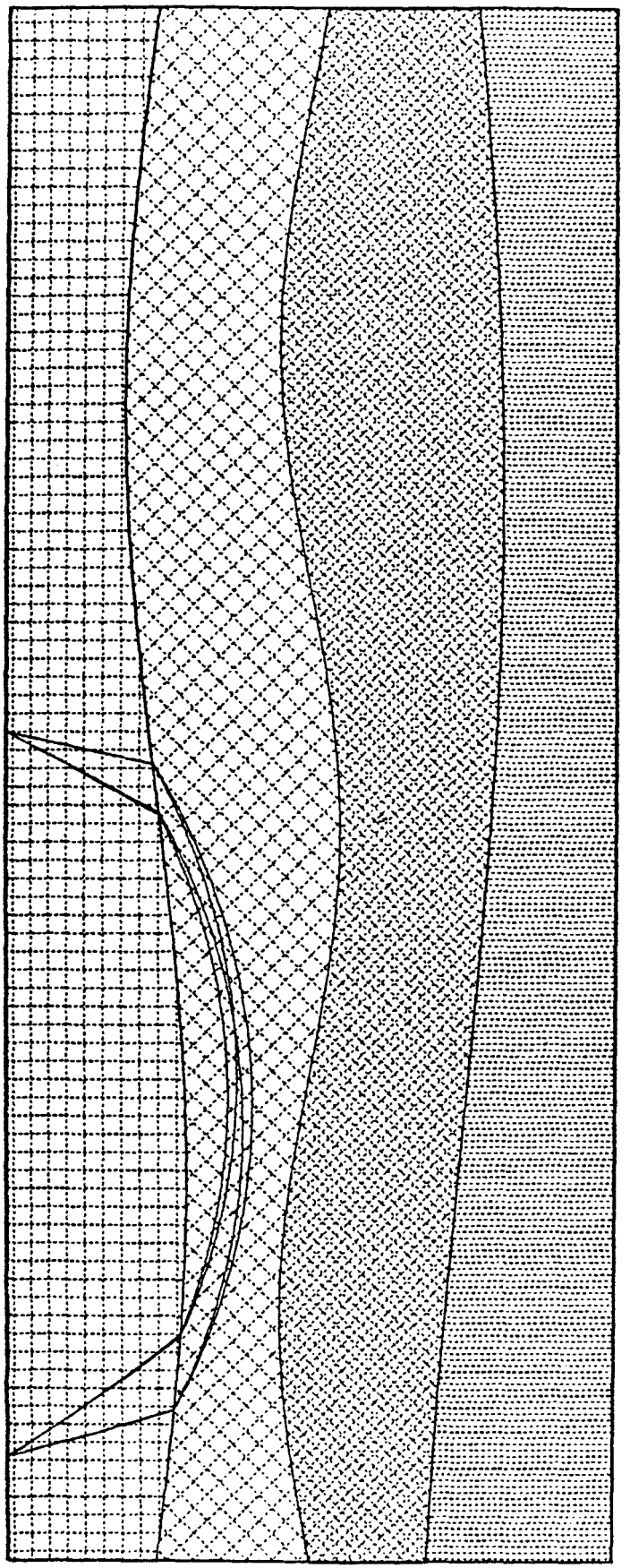

Figure 5d

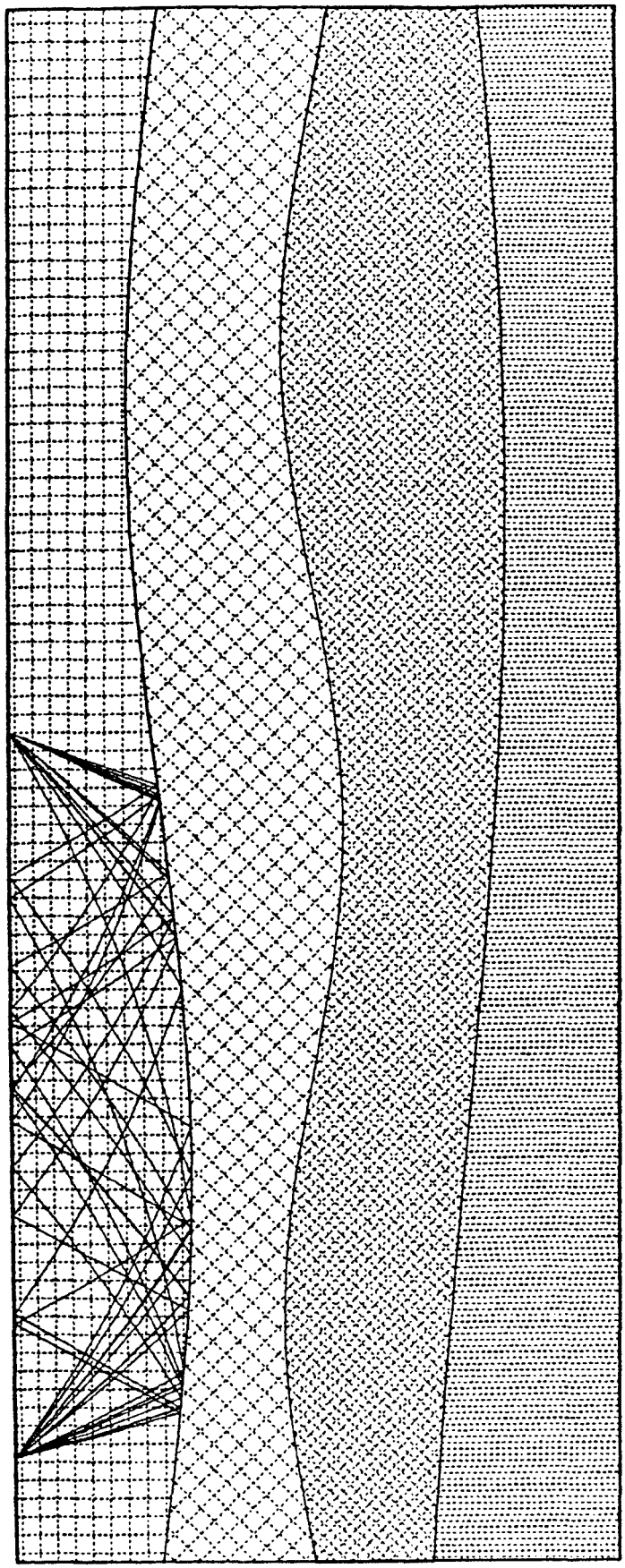

Figure 5e

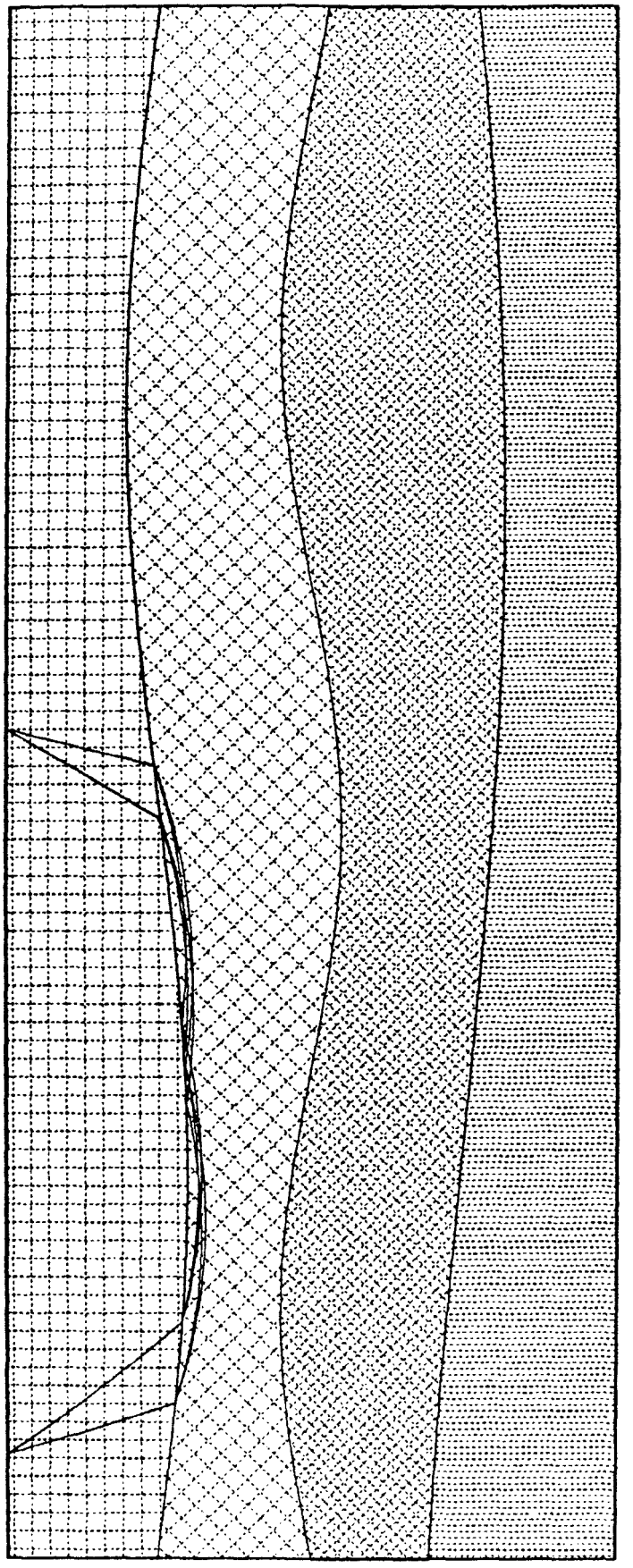

Figure 5f

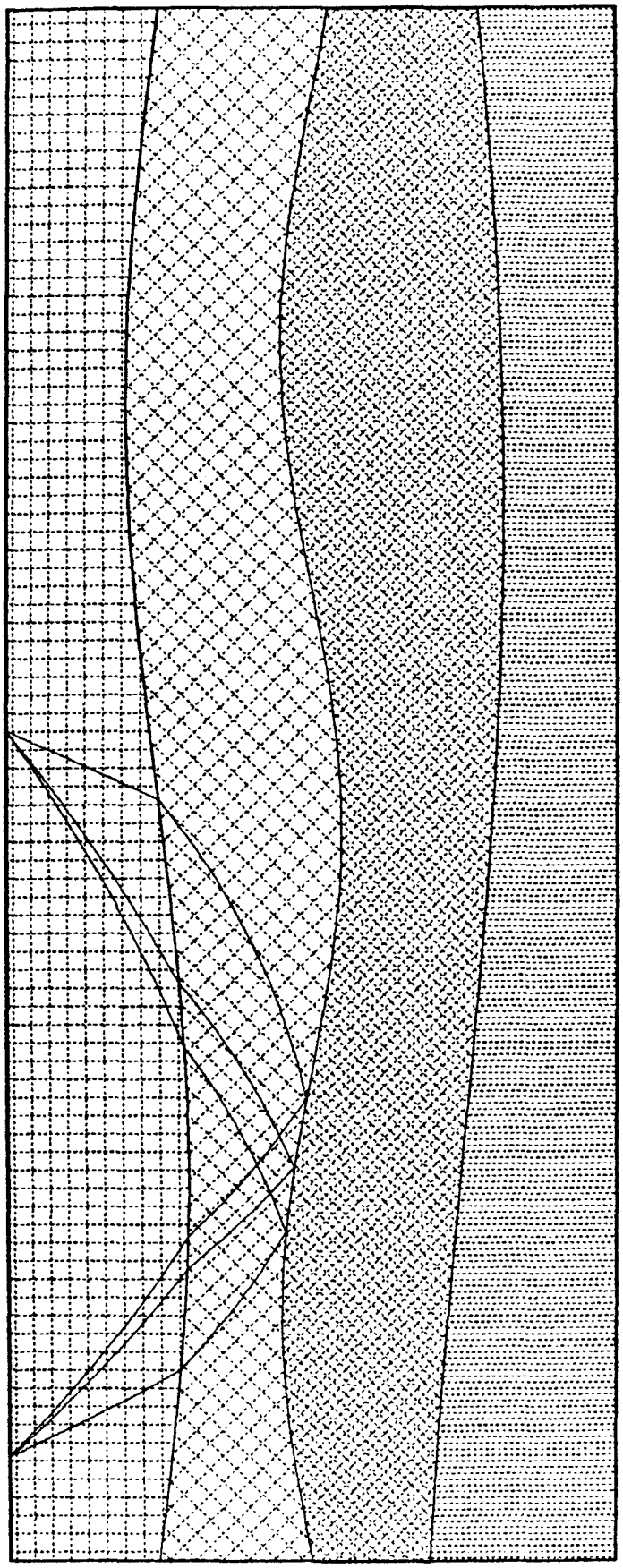

Figure 5g

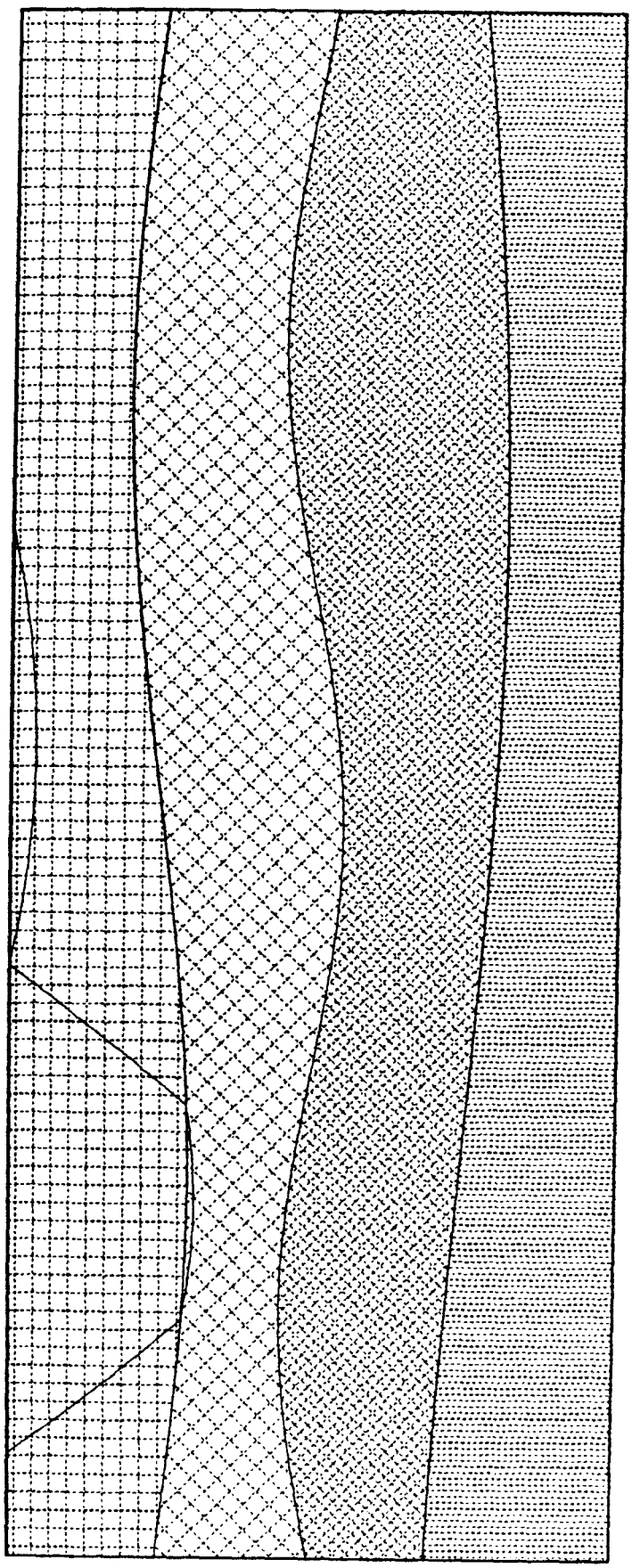

Figure 5h

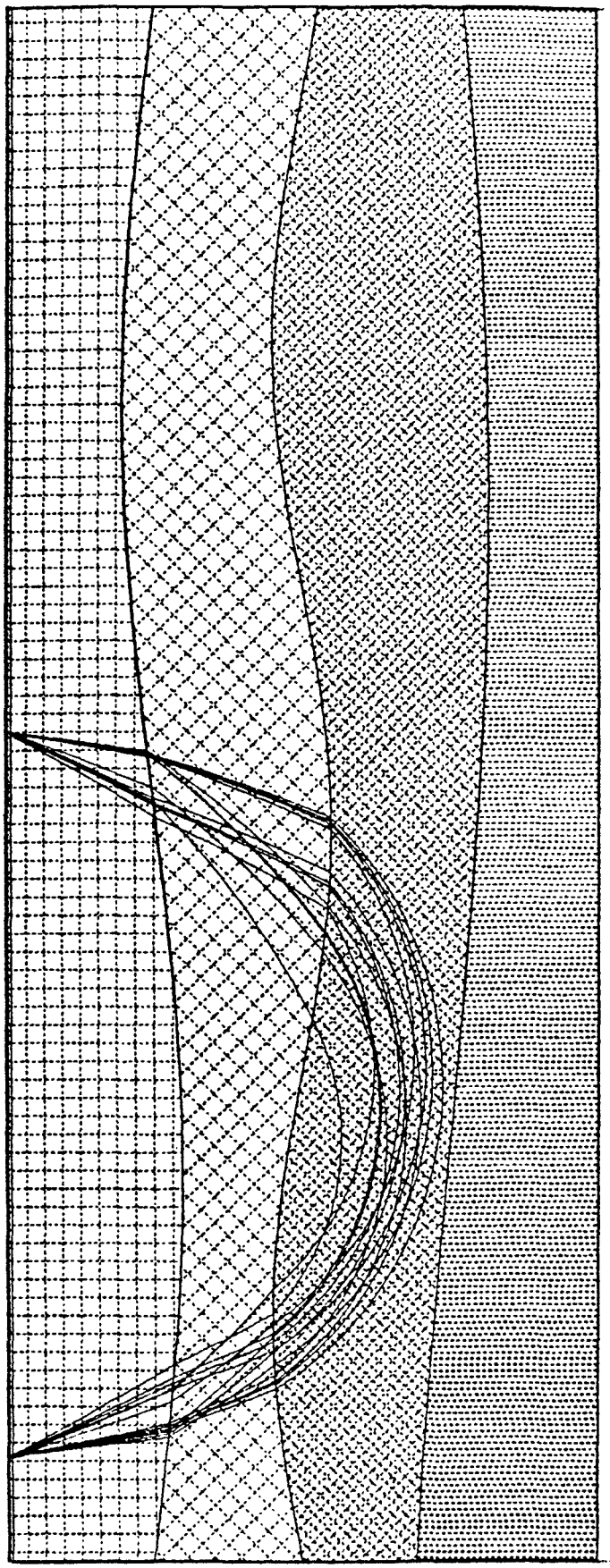

Figure 5i

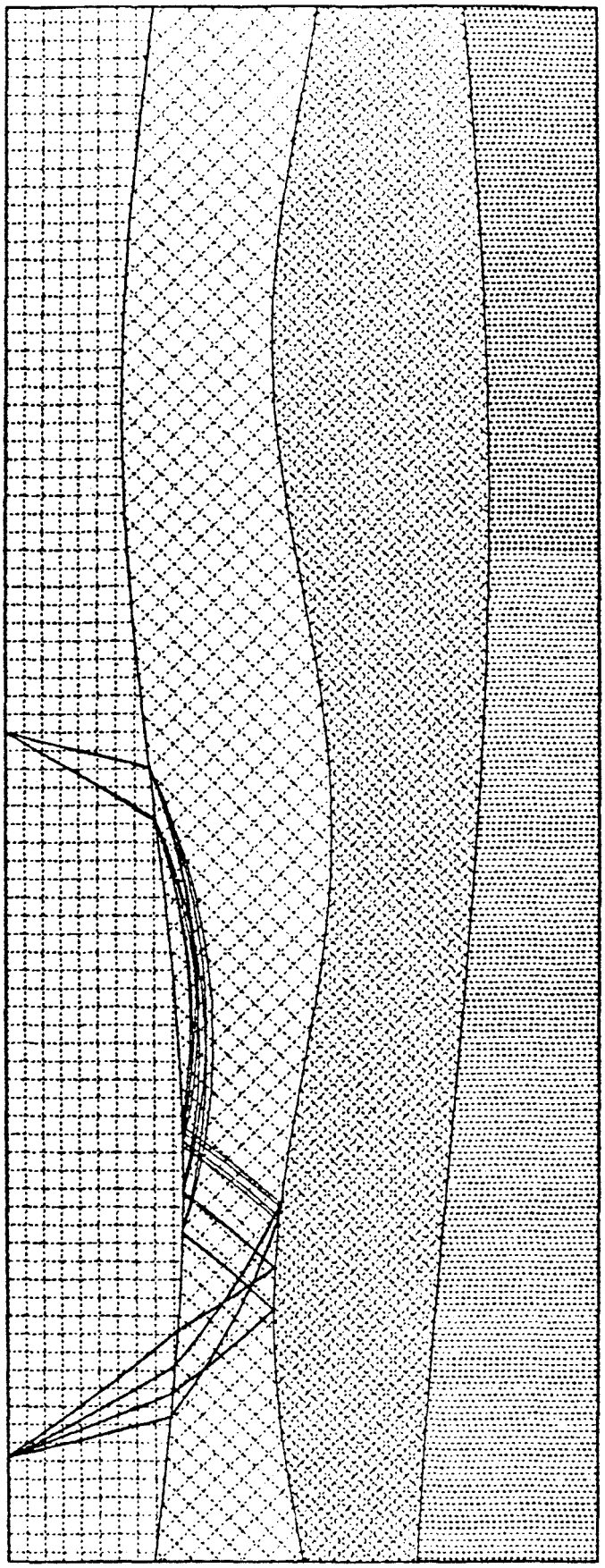

Figure 5j

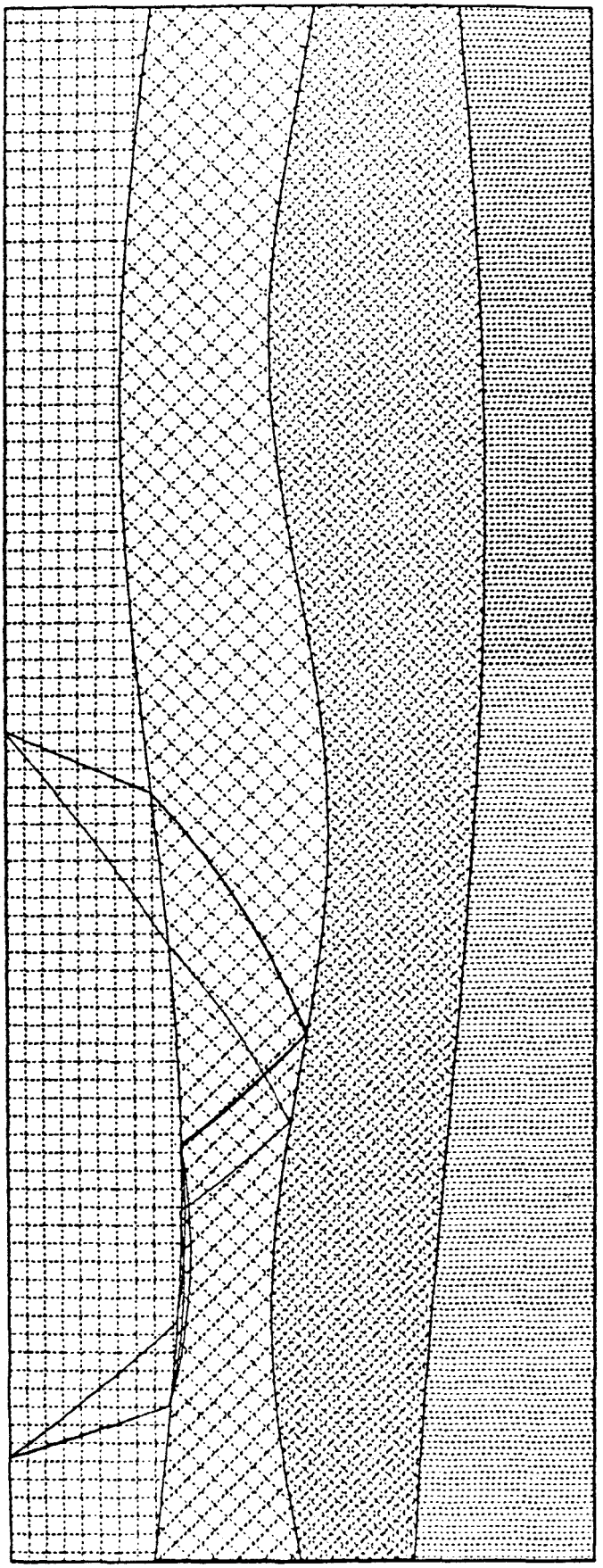

Figure 5k

# III. SOME COMMENTS ON THE INVERSE PROBLEM

Thus far we have discussed in some detail the forward ray tracing problem for a class of fairly general three-dimensional velocity models. However, our ultimate objective is to apply our results toward the solution of inverse problems in exploration seismology. In particular, given data (in the form of seismograms at an array of receivers ) about the arrival times and amplitudes of waves emanating from known sources, we wish to determine properties of the underlying Earth structure. In this section we begin a study of some inverse problems based on the results of previous sections.

## III.1 Solution of an Inverse Problem and Some Heuristic Considerations

For the time being, we limit our discussion to the retrieval of compressional wave velocity parameters based on first arrival times of waves from a point source to a gather of receivers located at the free surface of a linearly stratified half-space.

We assume that in the half-space $z > 0$, the velocity of P-waves satisfies

$$V(\mathbf{x}) = D + \alpha \hat{\mathbf{A}} \cdot \mathbf{x} \tag{1.1}$$

where $\hat{\mathbf{A}} \equiv (A_1, A_2, A_3)$ and $\|\hat{\mathbf{A}}\| = 1$.

The functional form is assumed for $V(\mathbf{x})$, but the values of the four parameters

$$\mathbf{p} \equiv (p_1, p_2, p_3, p_4) \equiv (\alpha, D, A_1, A_2) \tag{1.2}$$

are *a priori* unknown apart from the constraints $D > 0$ and $A_1^2 + A_2^2 \leq 1$.

The source is located at the free surface $z = 0$, at the origin of a cylindrical polar coordinate system $(r, \phi, z)$.

A gather of N receivers are located at

$$\mathbf{x}_j = (r_j \cos \phi_j, r_j \sin \phi_j, 0), \qquad j = 1, \ldots, N. \tag{1.3}$$

We denote by $T_j^{obs}$ the observed first arrival travel time from the source to the $j^{th}$ receiver. If the parameters $\mathbf{p}$ were known, we could compute the arrival times $T(r_j, \phi_j; \mathbf{p})$. In particular, using the results of section II.2.3, we would have

$$T(r_j, \phi_j; \mathbf{p}) = \frac{1}{2\alpha} \ln \left[ \frac{(1 + \cos\theta_{0j})(1 - \cos\theta_{1j})}{(1 - \cos\theta_{0j})(1 + \cos\theta_{1j})} \right] \tag{1.4}$$

where

$$\cos\theta_{0}j = \frac{\eta_{cj}}{R_j} \tag{1.5a}$$

$$\cos\theta_{1}j = \frac{(\eta_{cj} - \eta_j)}{R_j} \tag{1.5b}$$

$$\eta_{cj} = \frac{1}{2\eta_j}(r_j^2 + 2\xi_j D/\alpha) \tag{1.5c}$$

$$R_j = \left[(\xi_j + D/\alpha)^2 + \eta_{cj}^2\right]^{1/2} \tag{1.5d}$$

$$\xi_j = r_j(A_1 \cos\phi_j + A_2 \sin\phi_j) \tag{1.5e}$$

and

$$\eta_j = (r_j^2 - \xi_j^2)^{1/2}. \tag{1.5f}$$

The inverse problem which we consider is that of finding those values of the parameters $\mathbf{p}$ for which computed travel times (based on (1.4)) most closely approximate the $T_j^{obs}$.

We assume that each of the available observations $T_j^{obs}$ is equally reliable. By the principle of least squares we must, therefore, find those values of $\mathbf{p}$ which minimize the function

$$\begin{aligned} F(\mathbf{p}) &= \sum_{j=1}^{N} \left(T_j^{obs} - T(r_j, \phi_j; \mathbf{p})\right)^2 \\ &\equiv \sum_{j=1}^{N} f_j^2, \end{aligned} \tag{1.6}$$

with evident definition of $f_j$.

We discuss several heuristic ideas which will be helpful in solving this optimization problem. Clearly, we expect that irrespective of what algorithm we use, the

parameters p will be better resolved as N increases. However, unlike the problem of velocity inversion for a vertically stratified medium, in which case the receivers may be confined to a radial gather (i.e., one with $\phi_j$ constant for $j = 1, \ldots, N$), it is crucial for a reliable solution that observations be taken over a broad range of azimuth as well. Indeed, it is easy to see that even if $\alpha$ and $D$ were known exactly, there are, in general, two choices of $A_1$ and $A_2$ that will minimize (1.6) if all the observations are taken along a radial gather.

With this caveat in mind, we now describe our algorithm for solving the inverse problem for $\mathbf{p}$. The method we use is essentially the Gauss-Newton method, but with a slight modification. We define the vector $\mathbf{f}(\mathbf{p}) = (f_1, \ldots, f_j, \ldots, f_N)^T$ and observe that (1.6) can be recast in the form

$$\mathbf{f}(\mathbf{p}) \approx \mathbf{0}. \tag{1.7}$$

Denoting by $\mathbf{p}_n$ the $n^{th}$ estimate for the solution, we obtain $\mathbf{p}_{n+1}$ by linearizing about $\mathbf{p}_n$ and solving

$$\mathbf{f}(\mathbf{p}) + \frac{\partial \mathbf{f}(\mathbf{p}_n)}{\partial \mathbf{p}} \cdot (\mathbf{p}_{n+1} - \mathbf{p}_n) \approx \mathbf{0}, \tag{1.8}$$

where $\frac{\partial \mathbf{f}(\mathbf{p}_n)}{\partial \mathbf{p}}$ denotes the $N \times 4$ Jacobian of $\mathbf{f}$ with respect to $\mathbf{p}$. The least squares solution to the over-determined system (1.8) can be written as

$$\mathbf{p}_{n+1} = \mathbf{p}_n + \left( \frac{\partial \mathbf{f}(\mathbf{p}_n)}{\partial \mathbf{p}}^T \frac{\partial \mathbf{f}(\mathbf{p}_n)}{\partial \mathbf{p}} \right)^{-1} \frac{\partial \mathbf{f}(\mathbf{p}_n)}{\partial \mathbf{p}}^T \mathbf{f}(\mathbf{p}_n), \tag{1.9}$$

where the matrix product premultiplying $\mathbf{f}(\mathbf{p}_n)$ is the pseudo-inverse of $\frac{\partial \mathbf{f}(\mathbf{p}_n)}{\partial \mathbf{p}}$ .

What we have described thus far is just the Gauss-Newton method. However, because our problem is a constrained minimization, we modify this method slightly by performing a line search in the direction $\mathbf{p}_{n+1} - \mathbf{p}_n$ in the event that a full step kicks $\mathbf{p}$ outside the acceptable bounds. The sample output of our code reproduced below indicates that this method works quite well. The output is based on 21 receivers (seven each at three different $r_j$). The initial guess for $\mathbf{p}$ was (4.,400.,0.,0.).

```
ITERATION  1  STEPS IN LINE SEARCH        2
NEW GUESS FOR PARAMETERS
       1.63200542
     301.27083606
       0.04641738
       0.03094492
ITERATION  2  STEPS IN LINE SEARCH        1
NEW GUESS FOR PARAMETERS
       0.78545221
     225.26051521
       0.13409438
       0.08939625
ITERATION  3  STEPS IN LINE SEARCH        0
NEW GUESS FOR PARAMETERS
       0.96922272
     196.93835755
       0.17754576
       0.11843051
ITERATION  4  STEPS IN LINE SEARCH        0
NEW GUESS FOR PARAMETERS
       0.96752796
     200.05066382
       0.15661578
       0.10441052
ITERATION  5  STEPS IN LINE SEARCH        0
NEW GUESS FOR PARAMETERS
       1.00011785
     200.00153033
       0.14976438
       0.09985625
ITERATION  6  STEPS IN LINE SEARCH        0
NEW GUESS FOR PARAMETERS
       0.99999820
     200.00000437
       0.15000030
       0.10000020
```

# REFERENCES

[1] Ben-Menahem, A. and S. J. Singh, 1981. *Seismic Waves and Sources.* Springer-Verlag, New York.

[2] Bullen, K. E., 1963. *An Introduction to the Theory of Seismology.* Cambridge University Press, Cambridge.

[3] Červény, V., I. Molotkov and I. Psenčik, 1977. *Ray Method in Seismology.* Univerzita Karlova, Prague.

[4] Červény, V. and R. Ravindra, 1971. *Theory of Seismic Head Waves.* University of Toronto Press, Toronto.

[5] Chorlton, F., 1976. *Vector and Tensor Methods.* John Wiley and Sons, New York.

[6] Fawcett, J., 1983. *Three Dimensional Ray-Tracing and Ray Inversion in Layered Media*, Ph.D. Thesis, Part I. California Institute of Technology, Pasadena.

[7] Fawcett, J. and H. B. Keller, 1985. *Three Dimensional Ray Tracing and Geophysical Inversion in Layered Media.* SIAM Journal on Applied Mathematics, **45**, pp. 491-501.

[8] Feller, W., 1950. *An Introduction to Probability Theory and Its Applications.* John Wiley and Sons, New York.

[9] Julian, B. R. and D. Gubbins, 1977. *Three Dimensional Seismic Ray Tracing.* Journal of Geophysics, **43**, pp. 95-114.

[10] Keller, H. B., 1964. *Propagation of Stress Discontinuities in Inhomogeneous Elastic Media.* SIAM Review, **6**, pp.356-382.

[11] Keller, H. B. and D. Perozzi, 1983. *Fast Seismic Ray Tracing*. SIAM Journal on Applied Mathematics, **43**, pp. 981-992.

[12] Kennett, B. L. N., 1983. *Seismic Wave Propagation in Stratified Media*. Cambridge University Press, Cambridge.

[13] Kleyn, A. M., 1983. *Seismic Reflection Interpretation*. Applied Sciences Publishers, London.

[14] Kline, M. and I. W. Kay, 1965. *Electromagnetic Theory and Geometric Optics*. Interscience Publishers, New York.

[15] Pereyra, V., W. H. K. Lee and H. B. Keller, 1980. *Solving Two-Point Seismic-Ray Tracing Problems in a Heterogeneous Medium*. Bulletin of the Seismological Society of America, **70**, pp. 79-99.

[16] Pereyra, V., H. B. Keller and W. H. K. Lee, 1980. *Computational Methods for Inverse Problems in Geophysics: Inversion of Travel Time Observations*. Physics of the Earth and Planetary Interiors, **21**, pp.120-125.

[17] Perozzi, D., 1980. *Seismic Ray Tracing in Piecewise Homogeneous Media*, Ph.D. Thesis, Part I. California Institute of Technology, Pasadena.

[18] Saaty, T. L. and J. Bram, 1964. *Nonlinear Mathematics*. Dover Publications, New York.

[19] Shah, P. M., 1973. *Ray Tracing in Three Dimensions*. Geophysics, **38**, pp. 600-604.

[20] Stavroudis, O. N., 1976. *Simpler Derivation of the Formulas for Generalized Ray Tracing*. Journal of the Optical Society, **66**, pp. 1330-1333.

[21] Stoker, J. J., 1969. *Differential Geometry*. John Wiley and Sons, New York.