I. THREE DIMENSIONAL RAY-TRACING AND RAY-INVERSION IN LAYERED MEDIA.

II. INVERSE SCATTERING AND CURVED RAY TOMOGRAPHY WITH APPLICATIONS TO SEISMOLOGY.

Thesis by

John Alan Fawcett

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1983.

(Submitted May 17, 1983)

ACKNOWLEDGEMENTS

I would like to thank my advisor Professor Herbert B. Keller for his help, guidance and patience during my graduate career at Caltech. I am very grateful for the time and consideration he has given me throughout my graduate tenure here. I have also had many invaluable discussions on various topics in geophysics with Professor Robert W. Clayton of the Department of Geological and Planetary Sciences. The last chapter of this thesis on tomography benefitted greatly from Professor Clayton's advice and suggestions. I am also grateful for several discussions on inverse scattering and the theory of integral equations I had with Professor Tim Minzoni of the National University of Mexico, while he was visiting Caltech.

Financial support for the last three years was provided by Post-Graduate Scholarships of the National Science and Engineering Research Council of Canada. I was also supported by Institute tuition stipends and research assistantships, under United States Department of Energy contract number, DE-AS03-76SF-00767. The Departments of Computing Science and Mathematics at the University of Victoria provided me with office space and computing facilities for a three month stay there during 1982. I would especially like to thank Professor Ian Barrodale for his interest and support, while I was there.

I have made many friends at Caltech, and I would like to thank all the faculty, staff, and students of the Department of Applied Mathematics for the help and friendship shown to my family and me.

Finally, I would like to dedicate this thesis to my wife Anne, and

daughter Helen. They have led the hard life of a graduate student's family with constant cheeriness and support, making my graduate life a very happy one.

ABSTRACT

In seismology, the basic problem is that of deducing some knowledge of the geological structure of portions of the Earth from observed seismic signals. This leads to the concepts of seismic interpretation, or more mathematically, the formulation of inverse problems.

Some aspects of seismic wave propagation can be interpreted in terms of asymptotic ray theory. In Chapter 1 of Part I, we describe the numerical ray tracing algorithm we developed for layered media with interfaces that can vary in three dimensions. We describe in Chapter 2, how this ray tracing method is implemented in an inversion procedure. This method is based on the theory of non-linear least-squares inversion.

In Part II of the thesis, we discuss two formulations of seismic inverse problems, which are more analytical in nature. Chapter 1 deals with the use of inverse scattering theory for the Schroedinger operator in the seismological problem. In chapter 2 of Part II, we develop the theory of the tomographical inversion of travel time anomalies to determine velocity anomalies within the Earth. Here, we have extended, in an approximate sense, the Inverse Radon Transform to situations where the "background" velocity field varies with depth.

TABLE OF CONTENTS

I.0 Introduction Part I.

In Part I of the thesis, we examine the use of seismic ray theory, in direct and inverse seismic ray problems. In chapter 1 the direct problem considered, is that of finding the seismic signal, within the geometric optics approximation, at a specified receiver location. The source is known, and the medium consists of layers of homogeneous elastic material, that are separated by known, non-intersecting, smooth interfaces that can vary in three dimensions. We develop efficient and accurate methods for the determination of seismic rays and we calculate exactly the rays' travel times, amplitudes and phases.

In chapter 2, we consider the inverse problem of estimating elastic and interface parameters of a layered elastic medium, from a knowledge of the travel times and amplitudes of rays between various sources and receivers. We utilize the efficient ray generation methods of chapter 1, in a non-linear least squares algorithm for parameter estimation.

Chapter 1. Direct Problem

Introduction. 1.0

In our ray tracing methods, we use the first order geometric optics approximation to the solution of the linearized equations of elasticity. The particle displacement, $\underline{U}$, is approximated for high frequencies ω by:

$$\underline{U} = e^{i\omega(t-\sigma(\underline{x}))}\,\underline{U}_0(\underline{X}) \qquad (0.1)$$

Here $\sigma(\underline{X})$ is the "phase function", or travel time to a point $\underline{X}$, and $\underline{U}_0(\underline{X})$ is the amplitude coefficent. Using this approximation it is found that two forms of uncoupled waves are possible for $\underline{U}$. The wave fronts travel with velocities $Vp=((\lambda+2\mu)/\rho)^{1/2}$, which is called the compressional velocity, and $Vs=(\mu/\rho)^{1/2}$, which is called the shear velocity. Here ρ is the density and μ and λ are Lame's elastic parameters. For a compressional (P) wave, $\underline{U}$ will represent a longitudinal particle motion, and for the shear (S) wave, $\underline{U}$ will represent a transverse particle motion. The orthogonal trajectories to the wavefronts are called the "rays", and within homogeneous layers the rays are straight line segments. The first order ray approximation to the elastic wave equation solution is invalid for many interesting and important seismic phenomena, such as surface waves, head waves, diffraction, and shadow zone behaviour. However, even with these limitations, this simple ray theory approach is very useful in seismic interpretation.

Our problem in chapter 1 consists of 2 parts: i) determining a ray

which intersects a given sequence of interfaces, starting and ending at specified points, and ii) determining for this ray the amplitude coefficent and phase shift of the signal at the receiver.

Section 1. The Kinematic Problem.

We seek to find all physical rays which join given initial and final points, and which intersect the specified sequence of interfaces. The compressional and shear velocities of each layer are assumed known. In general, there may be several physical rays which satisfy the above constraints. The existence and computation of all the multiple solutions is an important but difficult problem. We will concentrate on the problem of finding one of possibly many solutions. However, we will discuss the problem of multiple solutions in a later section.

The seismic ray tracing problem has often been posed as a "shooting" problem (eg., Shah[16]). The shooting angles at the source are varied until the ray ends sufficiently close to the receiver. For a sufficiently good initial guess of the starting angles, a Newton iteration scheme can be used to refine the answer. However, a good initial estimate of the starting angles may be hard to find and the endpoint of the ray may depend unstably upon these initial angles.

Our approach follows that of Keller and Perozzi[9],[15] for the two dimensional problem. We formulate the problem as a system of non-linear equations with the source and receiver positions $\underline{X}_S$ and $\underline{X}_R$ as known. An Euler-Newton continuation method is used to solve the system.

There are several important differences between the two dimensional and three dimensional ray tracing problems. There are twice as many equations and unknowns in the three dimensional problem. In the two dimensional case the ray must satisfy the scalar Snell's law at each interface. In the three dimensional case, the ray must also satisfy a

scalar Snell's law at each interface, but the plane in which it is satisfied must also be determined. This is done by requiring that the incident ray, the emergent ray, and the normal to the interface at the ray contact point, be coplanar. These conditions can be formulated compactly as a vector Snell's law. Hence our system of equations is larger and different from the corresponding two dimensional system. However for a two dimensional problem it is easy to show that our system can be reduced to the smaller two dimensional system of Keller and Perozzi [9].

The velocity continuation method to generate the initial, purely compressional ray discussed in [9], [15] for the two dimensional problem, is no longer directly applicable to the three dimensional problem. We have devised a new continuation scheme, where the continuation parameter is the deviation of the interfaces from parallel planar interfaces. The details of these methods are discussed below.

1.1a Notation Used.

Since a ray segment in every layer is a straight line segment, any ray is determined by specifying its initial point $\underline{X}_S$ and its final point $\underline{X}_R$, and in order, the N points at which it intersects the interfaces, say $\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_N$. We adopt the notation that the k'th node, $\underline{X}_K$ lies on the interface z_{i_k}, k=1,2,...,N. In each specific case, we must specify the integers i1,i2,...,iN. The source and/or receiver points will often lie on the earth's surface which we denote as the interface z_0. The normal to the interface at the k'th node, is:

$$\underset{\sim}{N}_k \equiv \left(-\frac{\partial z_{i_k}}{\partial x}, -\frac{\partial z_{i_k}}{\partial y}, 1 \right) \tag{1.1}$$

The vector $\underline{X}_1 - \underline{X}_s$, denotes the ray segment from the source to the first node point, and similarly for the other segments. We also use the notation:

$$\underline{X}_k = (x_k, y_k, z_{i_k}(x_k, y_k)) \quad k = 1, \ldots, N$$

$$\underline{X}_0 \equiv \underline{X}_s \qquad \underline{X}_{N+1} \equiv \underline{X}_R \tag{1.2}$$

$$d_k \equiv \sqrt{(\Delta x_k)^2 + (\Delta y_k)^2 + (\Delta z_k)^2} \quad k = 1, \ldots, N+1$$

$$\Delta x_k \equiv x_k - x_{k-1}, \quad \Delta y_k \equiv y_k - y_{k-1}, \quad \Delta z_k \equiv z_{i_k} - z_{i_{k-1}}$$

The velocity on the k'th ray segment is V_k (k=1,...,N+1) and it is also to be specified as either Vp or Vs for the layer containing that segment. The above notation is illustrated in figure 1.1.

1.1b Formulation of the Ray Tracing Problem.

In each constant velocity layer the ray is a straight line segment, but at the k'th node (the i_k interface) the two ray segments (the incident and the emergent) must satisfy constraints. For the two dimensional case (see Keller,Perozzi[9],[15]) this is simply the scalar Snell's law

$$V_{k+1}\left(\frac{\hat{\tau}_k \cdot (\underline{X}_k - \underline{X}_{k-1})}{d_k}\right) = V_k\left(\frac{\hat{\tau}_k \cdot (\underline{X}_{k+1} - \underline{X}_k)}{d_{k+1}}\right) \tag{1.3}$$

where $\hat{\tau}_k$ is a tangent to the interface at point $\underline{X}_k = (x_k, z_{i_k}(x_k))$ and V_k, V_{k+1} the 2 layer velocities. For the three dimensional case we require

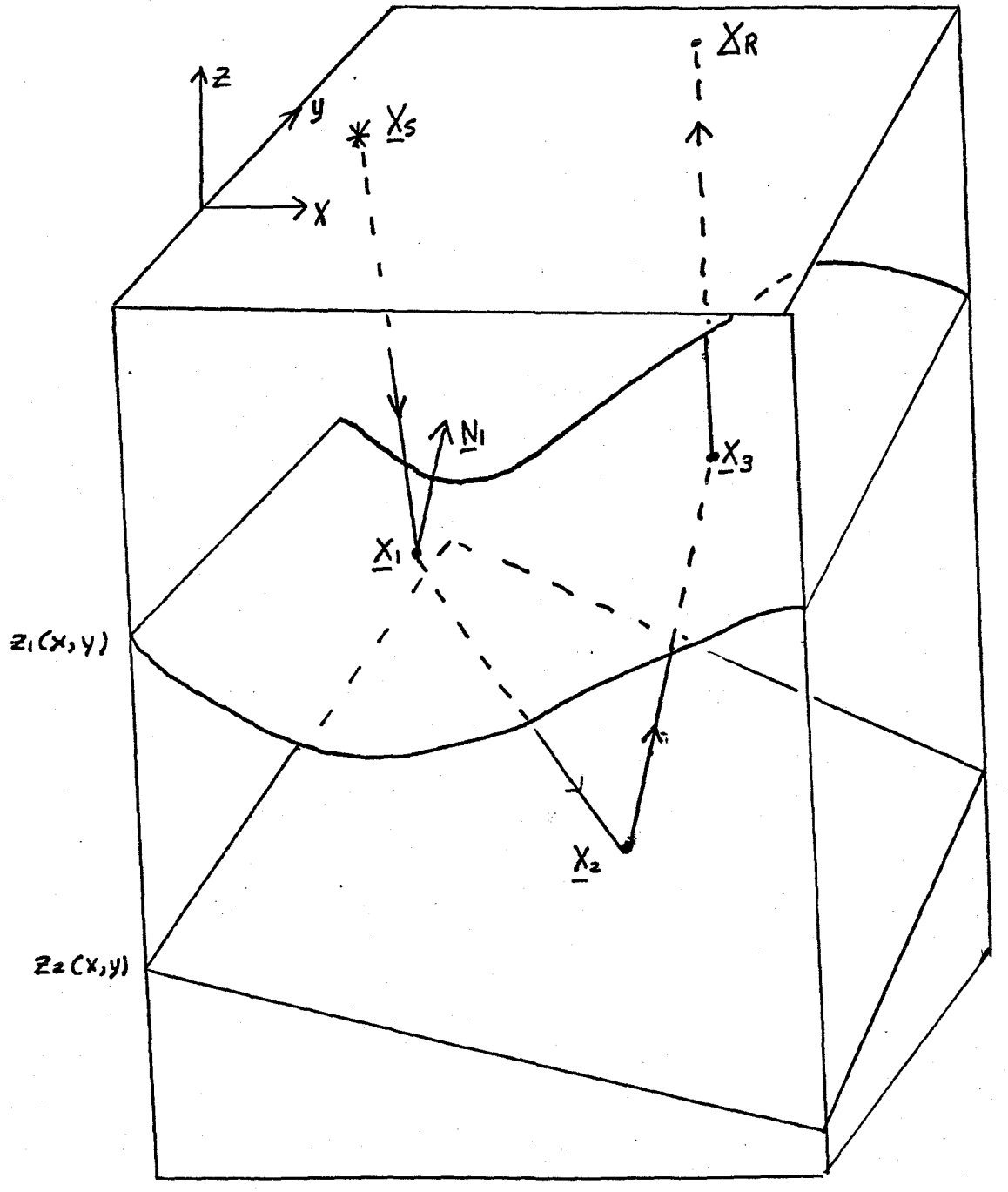


Figure 1.1 Illustration of Notation Used

that the incident ray, $\underline{X}_k - \underline{X}_{k-1}$, the emerging ray, $\underline{X}_{k+1} - \underline{X}_k$, and the normal N_k, at the ray contact point be coplanar. Their plane is called the plane of incidence. On this plane we require that (1.3) hold. The coplanarity condition can be written as:

$$\frac{(\underline{X}_{k+1} - \underline{X}_k)}{d_{k+1}} \times \underline{N}_k = \alpha \frac{(\underline{X}_k - \underline{X}_{k-1}) \times \underline{N}_k}{d_k} \tag{1.4}$$

where α is an arbitrary scalar. When this holds we can write:

$$\frac{(\underline{X}_k - \underline{X}_{k-1})}{d_k} = \gamma_1 \underline{e}_n + \beta_1 \underline{e}_\tau \qquad \gamma_1^2 + \beta_1^2 = 1$$

$$\frac{(\underline{X}_{k+1} - \underline{X}_k)}{d_{k+1}} = \gamma_2 \underline{e}_n + \beta_2 \underline{e}_\tau \qquad \gamma_2^2 + \beta_2^2 = 1 \tag{1.5}$$

where $\underline{e}_n$ is a unit vector in the normal direction, $\underline{N}_k$, and $\underline{e}_\tau$ is a unit vector orthogonal to $\underline{e}_n$, and in the plane of incidence (i.e., it is the unit vector parallel to $\underline{\tau}$ of (1.3)). We will choose $\underline{e}_\tau$ so that $\beta_1 > 0$. Then the scalar Snell's law (1.3) reduces to:

$$V_{k+1} \beta_1 = V_k \beta_2 \tag{1.6}$$

since $\underline{\tau}_k \cdot \underline{e}_n = 0$. Using expressions (1.5) in (1.4), and using $\underline{e}_n \times \underline{N}_k = 0$, we obtain:

$$\beta_2 \underline{e}_\tau \times \underline{N}_k = \alpha \beta_1 \underline{e}_\tau \times \underline{N}_k \tag{1.7}$$

From (1.6) we obtain $\alpha = V_{k+1} / V_k$. Thus we obtain Snell's vector law

$$\frac{V_{k+1}}{d_k}(\underline{X}_k - \underline{X}_{k-1}) \times \underline{N}_k = \frac{V_k}{d_{k+1}}(\underline{X}_{k+1} - \underline{X}_k) \times \underline{N}_k \qquad (1.8)$$

From (1.1) and (1.8) we get:

$$\frac{V_{k+1}\left(\Delta y_k + \Delta z_k \frac{\partial z_{ik}}{\partial y}\right)}{d_k} = \frac{V_k\left(\Delta y_{k+1} + \Delta z_{k+1} \frac{\partial z_{ik}}{\partial y}\right)}{d_{k+1}} \qquad (1.9a)$$

$$\frac{V_{k+1}\left(\Delta x_k + \Delta z_k \frac{\partial z_{ik}}{\partial x}\right)}{d_k} = \frac{V_k\left(\Delta x_{k+1} + \Delta z_{k+1} \frac{\partial z_{ik}}{\partial x}\right)}{d_{k+1}} \qquad (1.9b)$$

$$\frac{V_{k+1}\left(\frac{\partial z_{ik}}{\partial y}\Delta x_k - \frac{\partial z_{ik}}{\partial x}\Delta y_k\right)}{d_k} = \frac{V_k\left(\frac{\partial z_{ik}}{\partial y}\Delta x_{k+1} - \frac{\partial z_{ik}}{\partial x}\Delta y_{k+1}\right)}{d_{k+1}} \qquad (1.9c)$$

Only two of these three conditions are independent. For example if (1.9a) and (1.9b) hold then we can multiply (1.9a) by $\partial z_{ik}/\partial x$ and (1.9b) by $\partial z_{ik}/\partial y$, subtract and obtain (1.9c). Geometrically, this is to be expected. The two vectors of (1.8) lie in the same plane; thus equating two components of the vectors automatically implies that the third component of these coplanar vectors will also be equal. There are several equivalent sets of equations which can be obtained from (1.9) From (1.9a) and (1.9b) we have:

$$\frac{\left(\Delta y_{k+1} + \Delta z_{k+1} \frac{\partial z_{ik}}{\partial y}\right)}{\left(\Delta y_k + \Delta z_k \frac{\partial z_{ik}}{\partial y}\right)} = \frac{\left(\Delta x_{k+1} + \Delta z_{k+1} \frac{\partial z_{ik}}{\partial x}\right)}{\left(\Delta x_k + \Delta z_k \frac{\partial z_{ik}}{\partial x}\right)} = \frac{V_k d_{k+1}}{V_{k+1} d_k} \qquad (1.10)$$

Multiplying these ratios out we obtain a new equation:

$$\frac{\partial z_{ik}}{\partial x}(\Delta y_k \Delta z_{k+1} - \Delta z_k \Delta y_{k+1}) - \frac{\partial z_{ik}}{\partial y}(\Delta x_k \Delta z_{k+1} - \Delta z_k \Delta x_{k+1}) - (\Delta x_k \Delta y_{k+1} - \Delta y_k \Delta x_{k+1}) = 0 \qquad (1.11)$$

This is simply the vector equation:

$$((\underline{X}_K - \underline{X}_{K-1}) \times N_K) \cdot (\underline{X}_{K+1} - \underline{X}_K) = 0 \qquad (1.12)$$

which is the coplanarity condition.

Another interesting derivation of (1.9) is to consider Fermat's principle; that is, the travel time on the ray is extremal with respect to coordinate perturbations. Using our notation we can write for the travel time on a ray with N intersections:

$$tr(\underline{X}_S, \underline{X}_R) = \sum_{K=1}^{N+1} \frac{d_K}{V_K} \qquad (1.13)$$

From Fermat's principle, varying the j'th intersection point:

$$\frac{\partial tr}{\partial x_j} = \frac{1}{d_j V_j}(\Delta x_j + \frac{\partial z_{ij}}{\partial x_j}\Delta z_j) - \frac{1}{d_{j+1} V_{j+1}}(\Delta x_{j+1} + \frac{\partial z_{ij}}{\partial x_j}\Delta z_{j+1}) = 0 \qquad (1.14a)$$

$$\frac{\partial tr}{\partial y_j} = \frac{1}{d_j V_j}(\Delta y_j + \frac{\partial z_{ij}}{\partial y_j}\Delta z_j) - \frac{1}{d_{j+1} V_{j+1}}(\Delta y_{j+1} + \frac{\partial z_{ij}}{\partial y_j}\Delta z_{j+1}) = 0 \qquad (1.14b)$$

These are the same as (1.9a) and (1.9b).

We assume that the source position $\underline{X}_S$ and the receiver position $\underline{X}_R$ are known. Thus we wish to find the unknown coordinates (x_K, y_K) k=1,N of the N contact points of the ray with the interfaces. The depth (z-coordinate) of the ray's k'th contact point is implicitly known from $z=z_{i_K}(x_K, y_K)$. Hence we have 2N unknowns, but at each contact point we have

two conditions from Snell's law (1.9). Thus we can form a system of 2N equations in 2N unknowns. As mentioned above, there are several equivalent ways to formulate the problem. We chose two such formulations to study numerically. However, the concepts of the methods which we derive and test for these formulations, apply for any of the other formulations. We denote our system of equations in the vector form:

$$\underline{F}(\underline{X}\,;\,\underline{V},\underline{X}_S,\underline{X}_R) = \underline{O} \tag{1.15}$$

Here $\underline{X}$ denotes the vector of unknown contact points:

$$\underline{X} \equiv (x_1, y_1, x_2, y_2, \ldots, x_N, y_N) \tag{1.16}$$

$\underline{V}$ is the known vector of encountered velocities. We have explicitly stressed in (1.15) that $\underline{X}_S$ and $\underline{X}_R$ are important parameters in the problem. Using (1.9a) and (1.11) at each of the N contact points, we define the components:

$$F_{2k-1} \equiv \frac{V_{k+1}}{d_k}\left(\Delta y_k + \Delta z_k \frac{\partial z_{ik}}{\partial y_k}\right) - \frac{V_k}{d_{k+1}}\left(\Delta y_{k+1} + \Delta z_{k+1}\frac{\partial z_{ik}}{\partial y_k}\right) \tag{1.17a}$$

$$F_{2k} \equiv \frac{\partial z_{ik}}{\partial x}(\Delta y_k \Delta z_{k+1} - \Delta z_k \Delta y_{k+1}) - \frac{\partial z_{ik}}{\partial y}(\Delta x_k \Delta z_{k+1} - \Delta z_k \Delta x_{k+1}) - (\Delta x_k \Delta y_{k+1} - \Delta y_k \Delta x_{k+1}) \qquad k = 1, \ldots, N \tag{1.17b}$$

or from (1.9a) and (1.9b) we could use:

$$F_{2k-1} \equiv \frac{V_{k+1}}{d_k}\left(\Delta y_k + \Delta z_k \frac{\partial z_{ik}}{\partial y_k}\right) - \frac{V_k}{d_{k+1}}\left(\Delta y_{k+1} + \Delta z_{k+1}\frac{\partial z_{ik}}{\partial y_k}\right) \tag{1.18a}$$

$$F_{2K} = \frac{V_{K+1}}{d_K}\left(\Delta X_K + \Delta Z_K \frac{\partial Z_{i_K}}{\partial X_K}\right) - \frac{V_K}{d_{K+1}}\left(\Delta X_{K+1} + \Delta Z_{K+1} \frac{\partial Z_{i_K}}{\partial X_K}\right) \qquad (1.18b)$$

The first two components F_1 and F_2 involve $\underline{X}_S$ explicitly. Similarly the last two components F_{2N-1}, F_{2N} involve $\underline{X}_R$ explicitly.

Initially, we employed (1.17) before we considered (1.18). However, we note that for situations where the $\underline{e}_x$ component of $(\underline{X}_K - \underline{X}_{K-1}) \times \underline{N}_K$ is zero at some node the system (1.17) is no longer valid. Geometrically, this is because (1.17) forces $(\underline{X}_K - \underline{X}_{K-1}) \times \underline{N}_K$ and $(\underline{X}_{K+1} - \underline{X}_K) \times \underline{N}_K$ to be colinear. When the "$\underline{e}_x$" components of these vectors are non-zero, their equality implies the equality of the two vectors, and hence Snell's law is satisfied. However, this argument is not correct when $((\underline{X}_K - \underline{X}_{K-1}) \times \underline{N}_K) \cdot \underline{e}_x = 0$. This problem can often be circumvented by simply redefining the coordinates for the problem. However, for the case of zero-offset rays where $\underline{X}_K - \underline{X}_{K-1}$, $\underline{X}_{K+1} - \underline{X}_K$, and $\underline{N}_K$ may be colinear, (1.17) is not appropriate. Formulation (1.18) is valid for these situations. Thus from this standpoint, formulation (1.18) is preferable.

1.2 Solution of the System of Non-Linear Equations.

We wish to solve (1.16) using either definition (1.17) or (1.18) of $\underline{F}$, for the ray $\underline{X}$. One commonly used method to solve such a system is Newton's method. If $\underline{X}$ is "sufficiently" close to the root of $\underline{F}$, then an improved value $\underline{X}$ is given by:

$$\underline{X}^{(n+1)} = \underline{X}^{(n)} + \delta\underline{X}^{(n)}$$

$$\delta\underline{X}^{(n)} = -\underline{\underline{J}}^{(n)^{-1}} \underline{F}(\underline{X}^{(n)}; \underline{V}, X_s, X_R) \qquad (1.19)$$

Here, $\underline{\underline{J}}$ is the jacobian of the system: $\underline{\underline{J}} \equiv \partial\underline{F}/\partial\underline{X}$. When the iterates in Newton's method are "close" to a simple root, the iterates converge quadratically. However, in complicated problems, it may be difficult to supply a good initial guess $\underline{X}^{(0)}$.

1.2a Generation of Initial Purely Compressional Ray.

To start the calculation of rays for a given sequence of interfaces, we first calculate the purely compressional ray (i.e., each ray segment has the layer's compressional velocity) between a specified source and receiver pair. To provide a good initial estimate for this ray, we use a continuation method. We write each interface $z_{i_k}(x,y)$ in the form

$$z_{i_k}(x_k, y_k) = f_{i_k}(x_k, y_k) + c_{i_k} \qquad (1.20)$$

The constant c_{i_k} represents a mean depth of the interface. Instead of (1.20) we can introduce the one parameter family of interfaces:

$$z_k = \lambda f_{i_k}(x_k, y_k) + c_{i_k} \qquad 0 \leq \lambda \leq 1 \qquad (1.21)$$
$$k = 1, \ldots, N$$

We will solve the ray problem for plane parallel interfaces (λ=0) and then "deform" the plane interfaces into the desired (curved) interfaces (λ=1), as λ goes from zero to one. To determine the ray for a parallel plane

geometry is easily done, and the details of solution are given in Appendix A. We can write the non-linear system of equations for the ray, as λ varies, as:

$$\underline{F}\left(\underline{X}(\lambda), \underline{Z}(\lambda)\right) = \underline{0} \tag{1.22}$$

where $\underline{z}(\lambda)$ is the vector defined in (1.21). If a ray solution to (1.22), $\underline{X}(\lambda)$, is known for some value of λ, then as an initial ray estimate for the solution of (1.22) at $\lambda + \Delta\lambda$, we take:

$$\underline{X}^{(0)}(\lambda + \Delta\lambda) = \underline{X}(\lambda) + (\Delta\lambda)\,\dot{\underline{X}}(\lambda) \tag{1.23}$$

This estimate for $X(\lambda + \Delta\lambda)$ is accurate to order $(\Delta\lambda)$**2. We show later how to calculate $d\underline{X}/d\lambda$. Thus, between $\lambda = 0$ and $\lambda = 1$, we can consider a sequence of problems at $\lambda = \{\lambda_\nu\}, (\nu = 1, 2, \ldots, J)$ $\Delta\lambda_\nu = (\lambda_{\nu+1} - \lambda_\nu)$, and at each successive problem, we start the Newton iterations with the estimate:

$$\underline{X}^{(0)}(\lambda_{\nu+1}) = \underline{X}(\lambda_\nu) + \Delta\lambda_\nu\,\dot{\underline{X}}(\lambda_\nu) \qquad \lambda_1 = 0 \qquad \lambda_J = 1 \tag{1.24}$$

We will call the step (1.24), from the solution at $\lambda = \lambda_\nu$ to the initial estimate at $\lambda = \lambda_{\nu+1}$, an Euler step. The use of Newton's method to solve (1.22) at each λ_ν in conjunction with (1.24) is called Euler- Newton continuation.

In practice, to minimize the amount of computation, we use (1.23) with $\lambda = 0$, and $\Delta\lambda = 1$. In most of our numerical experiments, this value of $\underline{X}$ $(\lambda = 1)$ led to a convergent Newton's sequence. If however, Newton's method

does not converge quickly (say 7 iterations), we return to (1.24) with $\lambda_1 = 0$, $\lambda_2 = 1/2$, and $\lambda_3 = 1$. In general, we can increase the total number, J, of Euler steps to some Jmax, in order that the sequence (1.24) eventually lead to a convergent Newton's sequence at $\lambda = 1$.

For the two dimensional problem (i.e., interfaces can vary in only two dimensions: z=f(x)) Keller, Perozzi [9],[15] used a different continuation method to generate the initial compressional ray. They used a continuation in the vector of velocities $\underline{V}(\lambda) = (Vp,1(\lambda), Vp,2(\lambda), \ldots, VN+1(\lambda))^T$. At $\lambda = 0$, this vector is generated so that a simply chosen ray $\underline{X}$ is an exact solution to the system $\underline{F}(\underline{X}, \underline{V}(0)) = \underline{0}$. The velocity sequence $\underline{V}(0)$ is, in general, unphysical; some velocities may be negative and velocities corresponding to the same layer and type (P or S) may be multi-valued. For the three-dimensional problem, there are 2N unknown nodes, and so we could not arbitrarily choose a ray vector $\underline{X}$, and generate a N+1 component vector $\underline{V}$, such that $\underline{F}(\underline{X}, \underline{V}) = \underline{0}$. Thus we used the idea of continuation in interfaces, as outlined above. The advantages of this method are that: 1) it is a continuation in a physical parameter; that is, for each λ the resulting raytracing problem is a physical problem and it is even of interest for some situations to know the path of solutions; 2) often problems arise in seismology where the interfaces are only slightly different from parallel planes, and hence the ray solution for $\lambda = 0$ is an excellent approximation; 3) as we shall see in the numerical examples, this seems to be a robust method in conjunction with (1.17) and (1.18). That is, even for interfaces with appreciable slopes, $\underline{X}(\lambda = 0)$ is often a sufficiently good estimate for $\underline{X}(\lambda = 1)$.

1.2b Continuation in Velocities.

Associated with a given sequence of N interfaces (i1,i2,...,iN) there is a N+1 component velocity vector, containing the velocities of each of the N+1 ray segments. Each segment can have either a compressional velocity Vp or shear velocity Vs. Thus allowing for every possible sequence of shear and compressional velocities, there are 2^{N+1} possible velocity vectors, associated with a sequence of N interfaces. We denote the set of all such possible velocity vectors as $\{\underline{V}_k\}_{k=1}^{2^{N+1}}$. If we know the ray solution, $\underline{X}_k$, for the velocity sequence $\underline{V}_k$, then to find the ray solution $\underline{X}_{k+1}$ for the sequence $\underline{V}_{k+1}$ it is natural to consider the continuation problem:

$$\underline{F}(\underline{X}(\lambda);\underline{V}(\lambda)) = \underline{0} \qquad 0 \le \lambda \le 1$$
$$\underline{V}(\lambda) = (1-\lambda)\underline{V}_k + \lambda \underline{V}_{k+1} \;;\; \underline{X}(0) = \underline{X}_k \qquad (1.25)$$

Equations (1.23) and (1.24) (Euler-Newton continuation equations) are also applicable to this continuation (1.25). Starting with the purely compressional ray, it is easy to generate the rays for all 2^{N+1} velocity sequences, in such a way that successive velocity sequences differ by only one component.

1.2c Continuation in Receiver Position.

Often, one wishes to calculate rays to many different receivers $\underline{X}_{R,k} = (x_R, y_R, z_R)$. Assuming we know the ray solution $\underline{X}_i$ for the receiver at $\underline{X}_{R,i}$, then a natural continuation method, to generate the ray for the receiver at

$\underline{X}_{R,i+1}$ is to consider the parameter dependant problem:

$$\underline{F}(\underline{X}(\lambda); \underline{X}_R(\lambda)) = \underline{0} \qquad 0 \leq \lambda \leq 1$$
$$\underline{X}_R(\lambda) = (1-\lambda)\,\underline{X}_{R,i} + \lambda\,\underline{X}_{R,i+1} \qquad (1.26)$$
$$\underline{X}(0) = \underline{X}_i$$

Once again equations (1.23) and (1.24) (Euler-Newton continuation) are used with (1.26).

1.3 Calculation of Variational Derivatives.

We use the term, variational derivative, to refer to the derivative or variation of the ray solution $\underline{X}$ with respect to some parameter or parameter vector, $\underline{p}$, in the problem. We have already alluded above, to the calculation of $\dot{\underline{X}}(\lambda)$, where λ is the continuation parameter for the deviation of the interfaces from parallel planes, or for the different velocity sequences, or receiver/source positions. It is the algebraic formulation of the problem, that makes possible the easy calculation of these quantities. We will present the general formula, for arbitrary parameter variation and then give the resulting formulae for some specific cases.

Suppose we have some parameter vector $\underline{p}$ of dimension M and we wish to calculate the 2NxM matrix $d\underline{x}/d\underline{p}$. We indicate the parameter dependance of the kinematic problem by writing (1.17) or (1.18) as:

$$\underline{F}(\underline{X}(\underline{p}); \underline{p}) = \underline{0} \qquad (1.27)$$

Now this relation holds for all $\underline{p}$, and we assume that the solution $\underline{X}=\underline{X}(\underline{p})$ depends smoothly on $\underline{p}$. Then

$$\frac{d\underline{F}}{d\underline{p}} = \frac{\partial \underline{F}}{\partial \underline{X}} \frac{d\underline{X}}{d\underline{p}} + \frac{\partial \underline{F}}{\partial \underline{p}} = \underline{0} \tag{1.28}$$

or

$$\frac{d\underline{X}}{d\underline{p}} = -\underline{\underline{J}}^{-1} \frac{d\underline{F}}{d\underline{p}} \tag{1.29}$$

where $\underline{\underline{J}}$ is the jacobian (1.19) of the non-linear system of equations (formed from the ray solution at the next to last Newton iteration). To find the ray solution $\underline{X}$, (see above) we use the already calculated $\underline{\underline{J}}^{-1}$ (numerically, we know its L-U decomposition). Thus calculation (1.29) is "cheap" numerically. If p is a scalar, λ , then clearly (1.29) yields:

$$\frac{d\underline{X}}{d\lambda} = -\underline{\underline{J}}^{-1} \frac{\partial \underline{F}}{\partial \lambda} \tag{1.30}$$

We note that if we use different formulations for the problem as outlined in 1.2, then the definitions of $\underline{F}$, $\underline{\underline{J}}$, and $\partial \underline{F} / \partial \lambda$ are different for each formulation. However, the result $d\underline{X}/d\lambda$ from (1.29) must be identical (assuming we are at the same $\underline{X}$) for all cases. We used this idea to check our algebra for (1.17) and (1.18).

We now give two particular examples. First, let us consider $d\underline{X}/d\underline{X}_R$, where $\underline{X}_R = (X_R, y_R)$ is the receiver position. We will take $z_{i_{NH}}(X_R, y_R)=0$. From (1.29) we must calculate $\partial \underline{F} / \partial \underline{X}_R$. We will use the formulation (1.18). The quantities X_R , y_R only enter into the vector components F_{2N-1}, and F_{2N}. The derivatives $\partial F_{2N-1}/\partial X_R$, $\partial F_{2N-1}/\partial y_R$, $\partial F_{2N}/\partial X_R$, and $\partial F_{2N}/\partial y_R$ are

respectively:

$$\frac{V_N\,(x_R-x_N)((y_R-y_N)-z_N\,\frac{\partial f_{iN}}{\partial y_N})}{d_{N+1}^3} \equiv b_{11} \qquad (1.31a)$$

$$\frac{-V_N\,((x_R-x_N)^2+z_N^2+z_N\,\frac{\partial f_{iN}}{\partial y_N}(y_R-y_N))}{d_{N+1}^3} \equiv b_{12} \qquad (1.31b)$$

$$\frac{V_N\,(y_R-y_N)((x_R-x_N)-z_N\,\frac{\partial f_{iN}}{\partial x_N})}{d_{N+1}^3} \equiv b_{22} \qquad (1.31c)$$

$$\frac{-V_N\,((y_R-y_N)^2+z_N^2+z_N\,\frac{\partial f_{iN}}{\partial y_N}(x_R-x_N))}{d_{N+1}^3} \equiv b_{21} \qquad (1.31d)$$

Now, we can write

$$\frac{d\underline{X}}{d\underline{X}_R} = -J^{-1}\begin{bmatrix} 0 & 0 \\ \vdots & \vdots \\ b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}\begin{matrix} \\ \\ \leftarrow 2N-1 \\ \leftarrow 2N \end{matrix} \qquad (1.32)$$

We now outline the calculation of the derivative $\dot{X}(\lambda)$, where λ is the continuation parameter in (1.21). We rewrite (1.17) as:

$$F_{2k-1} \equiv \frac{V_{k+1}\,(\Delta y_k+(\lambda(f_{i_k}-f_{i_{k-1}})+(c_{i_k}-c_{i_{k-1}}))\lambda\,\frac{\partial f_{i_k}}{\partial y_k}}{\sqrt{(\Delta x_k)^2+(\Delta y_k)^2+(\lambda(f_{i_k}-f_{i_{k-1}})+(c_{i_k}-c_{i_{k-1}}))^2}} \qquad (1.33a)$$

$$-\frac{V_k\,(\Delta y_{k+1}+(\lambda(f_{i_{k+1}}-f_{i_k})+(c_{i_{k+1}}-c_{i_k})\lambda\,\frac{\partial f_{i_k}}{\partial y_k}}{\sqrt{(\Delta x_{k+1})^2+(\Delta y_{k+1})^2+(\lambda(f_{i_{k+1}}-f_{i_k})+(c_{i_{k+1}}-c_{i_k}))^2}}$$

$$F_{2K} \equiv \lambda \frac{\partial f_{iK}}{\partial x_K} \left(\Delta y_K \left(\lambda \Delta f_{K+1} + \Delta c_{K+1}\right) - \Delta y_{K+1} \left(\lambda \Delta f_K + \Delta c_K\right)\right)$$

$$- \lambda \frac{\partial f_{iK}}{\partial y_K} \left(\Delta x_K \left(\lambda \Delta f_{K+1} + \Delta c_{K+1}\right) - \Delta x_{K+1} \left(\lambda \Delta f_K + \Delta c_K\right)\right) \qquad (1.33b)$$

$$- \left(\Delta x_K \Delta y_{K+1} - \Delta y_K \Delta x_{K+1}\right)$$

Now the analytic calculation of $\partial \underline{F} / \partial \lambda$ is straightforward (although messy); similarly for F as in (1.18). From (1.30) we can now find $\dot{\underline{X}}$ (λ). For the case of velocity continuation, we have

$$W(\lambda) = \lambda \underline{V}_{K+1} + (1-\lambda) \underline{V}_K \qquad (1.34)$$

where here $\underline{V}_K$ is the sequence for which we have the ray solution, and we wish the ray solution for $\underline{V}_{K+1}$. Now from (1.17) we have

$$\frac{\partial F_{2j-1}}{\partial V_k} = -\frac{1}{d_{K+1}} \left(\Delta y_{K+1} + \Delta z_{K+1} \frac{\partial f_{iK}}{\partial y_K}\right) \quad j = k$$

$$= \frac{1}{d_K} \left(\Delta y_K + \Delta z_K \frac{\partial f_{iK}}{\partial y_K}\right) \quad j+1 = k$$

$$= 0 \qquad \text{otherwise} \qquad (1.35)$$

$$\frac{\partial F_{2j}}{\partial V_k} = 0 \qquad j = 1, \ldots, N$$

Thus, the matrix $\partial \underline{F} / \partial \underline{V}$ is a 2Nx(N+1) matrix of the form

$$\frac{\partial \underline{F}}{\partial \underline{V}} = \quad \overset{N+1 \longrightarrow}{\underset{2N \downarrow}{\left[\begin{array}{cccc} X & X & O & \cdots \\ O & O & O & \cdots \\ O & X & X & \cdots \\ O & O & O & O \\ O & O & X & X \quad O \quad \cdots \\ & & & \\ & & & \\ O & & & \\ O & & & \end{array}\right.}} \quad \left.\begin{array}{c} \\ \\ \\ \\ \\ \\ \\ X \\ O \end{array}\right] \begin{array}{c} \\ \\ \\ \\ \\ \\ \\ X \\ O \end{array} \tag{1.36}$$

Now $\dot{\underline{W}}(\lambda) = (\underline{V}_{N+1} - \underline{V}_k) \equiv \Delta \underline{V}_k$ or

$$\frac{\partial \underline{F}}{\partial \lambda} = \frac{\partial \underline{F}}{\partial \underline{W}} \dot{\underline{W}}(\lambda) = - \left[\begin{array}{c} \dfrac{\Delta V_1 (\Delta y_2 + \Delta z_2 \frac{\partial f_1}{\partial y_1})}{d_2} - \dfrac{\Delta V_2 (\Delta y_1 + \Delta z_1 \frac{\partial f_1}{\partial y_1})}{d_1} \\ 0 \\ \vdots \\ \dfrac{\Delta V_N (\Delta y_{N+1} + \Delta z_{N+1} \frac{\partial f_N}{\partial y_N})}{d_{N+1}} - \dfrac{\Delta V_{N+1} (\Delta y_N + \Delta z_N \frac{\partial f_N}{\partial y_N})}{d_N} \\ 0 \end{array}\right] \begin{array}{c} \uparrow \\ \\ \\ \\ \\ \\ \\ \\ \downarrow \end{array} 2N \tag{1.37}$$

For our other formulation (1.18) we obtain

$$\frac{\partial \underline{F}}{\partial \lambda} = - \left[\begin{array}{c} \dfrac{\Delta V_1 (\Delta y_2 + \Delta z_2 \frac{\partial f_1}{\partial y_1})}{d_2} - \dfrac{\Delta V_2 (\Delta y_1 + \Delta z_1 \frac{\partial f_1}{\partial y_1})}{d_1} \\ \dfrac{\Delta V_1 (\Delta x_2 + \Delta z_2 \frac{\partial f_1}{\partial x_1})}{d_2} - \dfrac{\Delta V_2 (\Delta x_1 + \Delta z_1 \frac{\partial f_1}{\partial x_1})}{d_1} \\ \vdots \\ \dfrac{\Delta V_N (\Delta y_{N+1} + \Delta z_{N+1} \frac{\partial f_N}{\partial y_N})}{d_{N+1}} - \dfrac{\Delta V_{N+1} (\Delta y_N + \Delta z_N \frac{\partial f_N}{\partial y_N})}{d_N} \\ \dfrac{\Delta V_N (\Delta x_{N+1} + \Delta z_{N+1} \frac{\partial f_N}{\partial x_N})}{d_{N+1}} - \dfrac{\Delta V_{N+1} (\Delta x_N + \Delta z_N \frac{\partial f_N}{\partial x_N})}{d_N} \end{array}\right] \begin{array}{c} \uparrow \\ \\ \\ \\ \\ \\ \\ \\ \downarrow \end{array} 2N \tag{1.38}$$

From (1.30) we can now calculate the variational derivative of the ray solution with respect to the velocity sequence. Thus in general, with an algebraic formulation of the problem, it is easy to calculate the variation of the ray solution with respect to medium paramaters.

1.4 Existence and Uniqueness of Ray Solutions; Path Following.

Thus far in the theory, we have discussed finding a particular ray solution. However, it is possible that for a given source/receiver positioning there is no ray solution. For this situation, of course, our methods will fail to converge. Alternatively, many solutions may exist for this non-linear problem. We now examine the path of solutions which can exist for the path of continuation as defined in (1.21). In figures 1.2a–1.2c, we show three possible path situations. The vertical axis coordinate is some distinguishing characteristic of the solution, such as its norm, one of the ray components, etc. Figure 1.2a shows an example of non-existence of a ray solution at $\lambda = 1$. The parameter value $\lambda = \lambda_0$ is called a "limit point." At this point, the Jacobian $d\underline{F} / d\underline{X}$ is singular. However, from Keller [10], one can still follow the path of solutions, by making the continuation parameter the arclength along the path. Figure 1.2b is a bifurcation diagram where $(\lambda_0, X_1 = 0)$ is the bifurcation point. A simple example of where this arises in ray theory is shown below in figure 1.3.

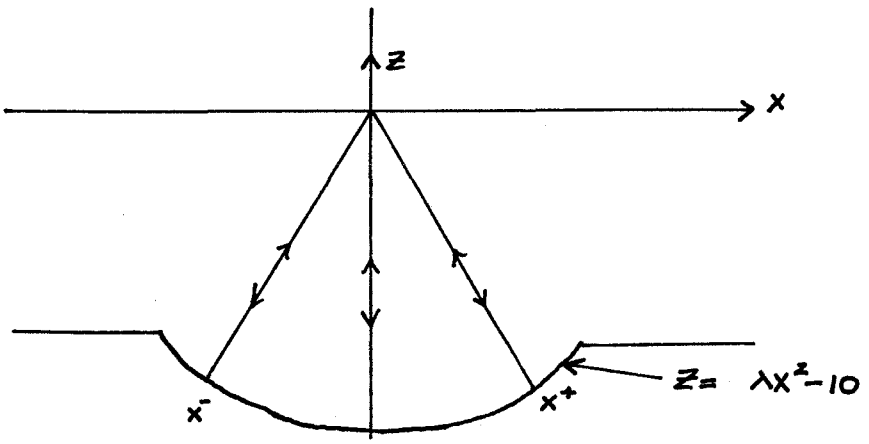


Figure 1.3 A Physical Example of Ray Bifurcation.

We consider the zero-offset ray (ie $\underline{X}_s = \underline{X}_g$ for $\underline{X}s=0$). The vertical ray is the only solution until the curvature of the interface at the node (0,-10) exceeds the radii of curvature of the wavefront, or $\lambda > 1/20$. Then, there are three ray solutions (x=0,z=-10) and $(X = \pm\sqrt{\frac{20\lambda - 1}{2\lambda^2}}, -10 + \lambda X^2)$. Finally, in figure 1.2c, we show a situation with multiple solutions, but not with bifurcation points.

We now give a two dimensional example. If we start with the planar ray solution, figure 1.4a, use an Euler step with $\Delta\lambda = 1$, then for $\lambda = 1$ the resulting ray solution is shown in figure 1.4b. However, if we follow the path from $\lambda = 0$ more closely with steps of $\Delta\lambda = 1/32$, we find that the "continuous" path changes rapidly and in fact does not lead to the solution of figure 3b. The first path steps with $\lambda = 1/32$, $\lambda = 1/16$, and $\lambda = 3/32$ are shown in figures 1.5a-1.5c. Thus our continuation method with $\Delta\lambda = 1$ leads us to an initial estimate for which Newton's method converges to a ray solution on a different path of solutions. As will be seen in the numerical examples, our method will often converge to a ray solution with $\Delta\lambda = 1$. However, from above (where multiple solutions exist) this solution may not lie on the path leading from the planar solution to $\lambda = 1$.

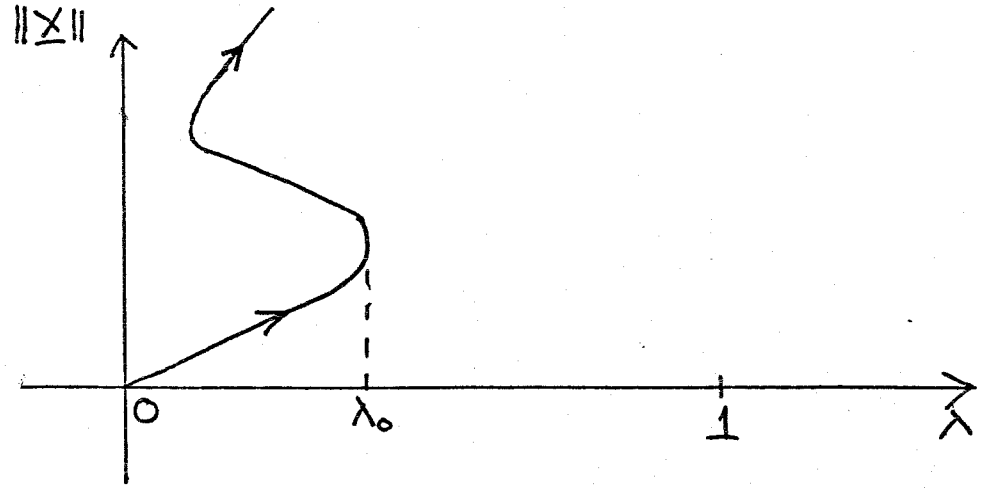


Figure 1.2a Limit Point Behaviour

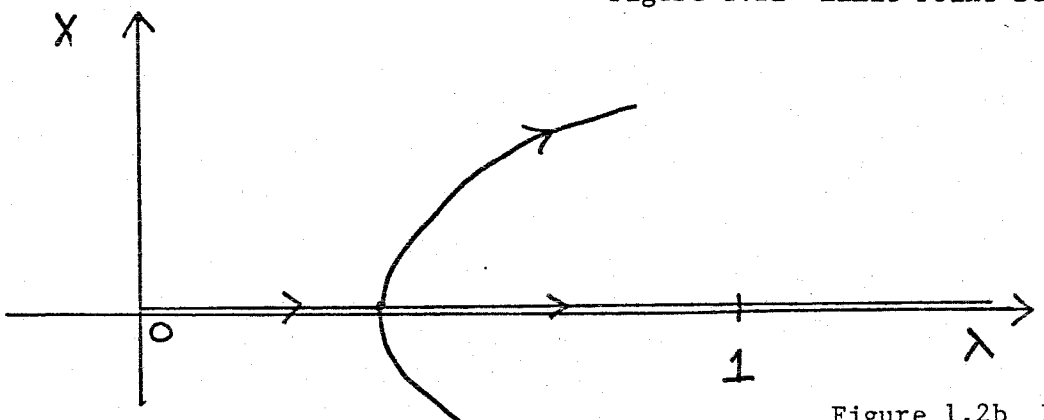


Figure 1.2b Bifurcation

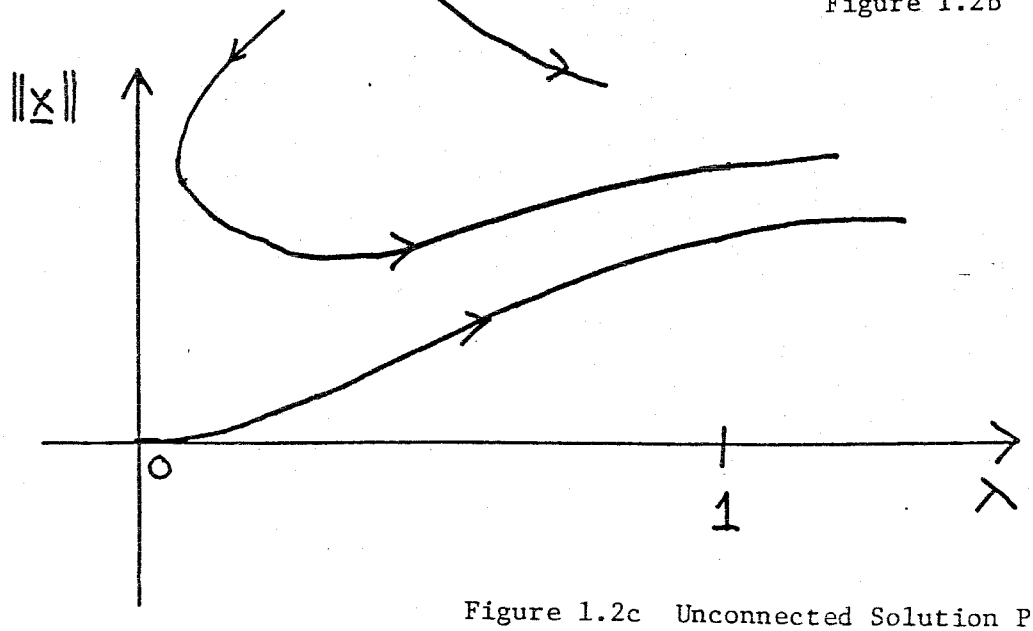


Figure 1.2c Unconnected Solution Paths

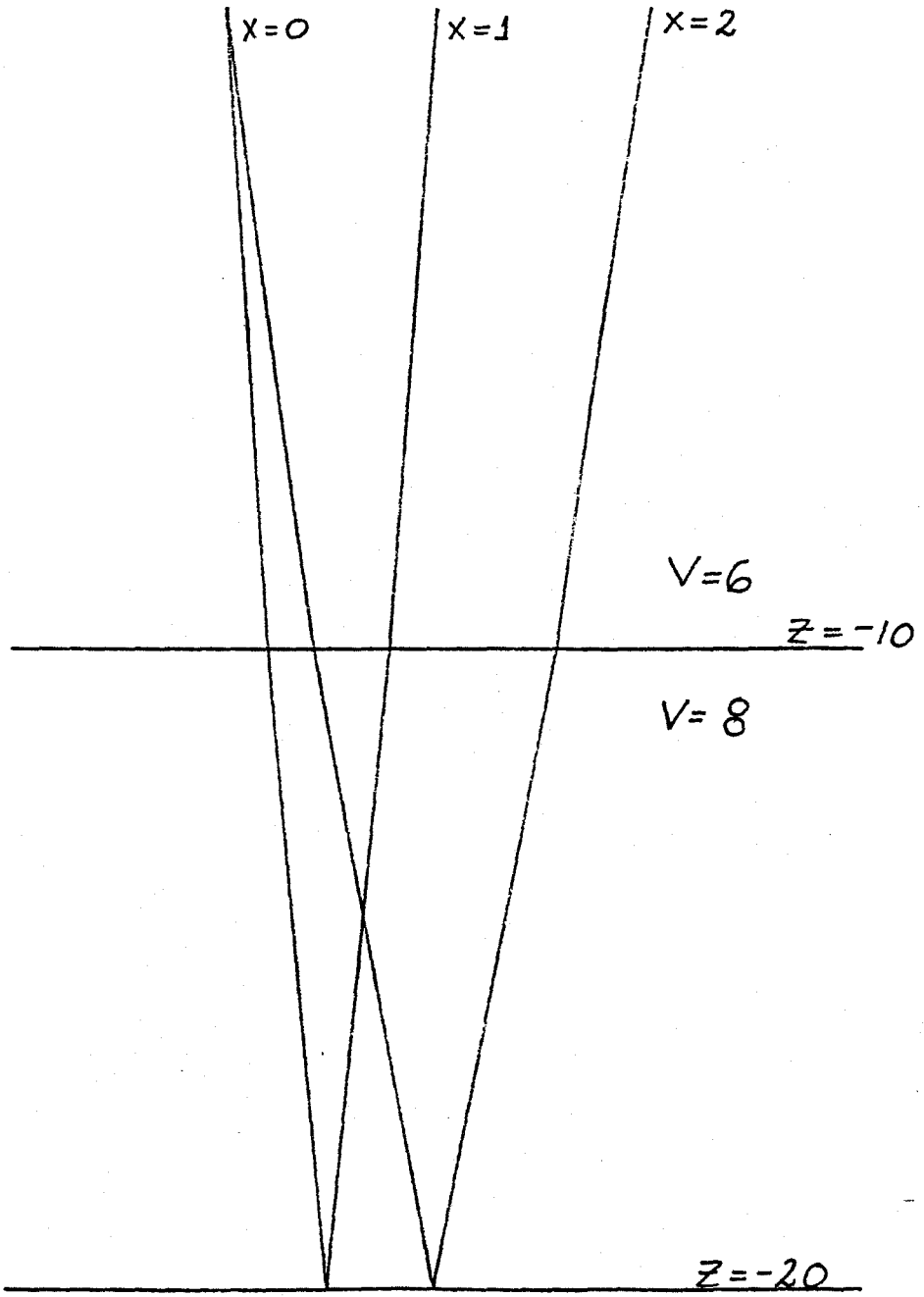


Figure 1.4a Planar Ray Solution
(Note: x,z coordinates have
unequal scalings)

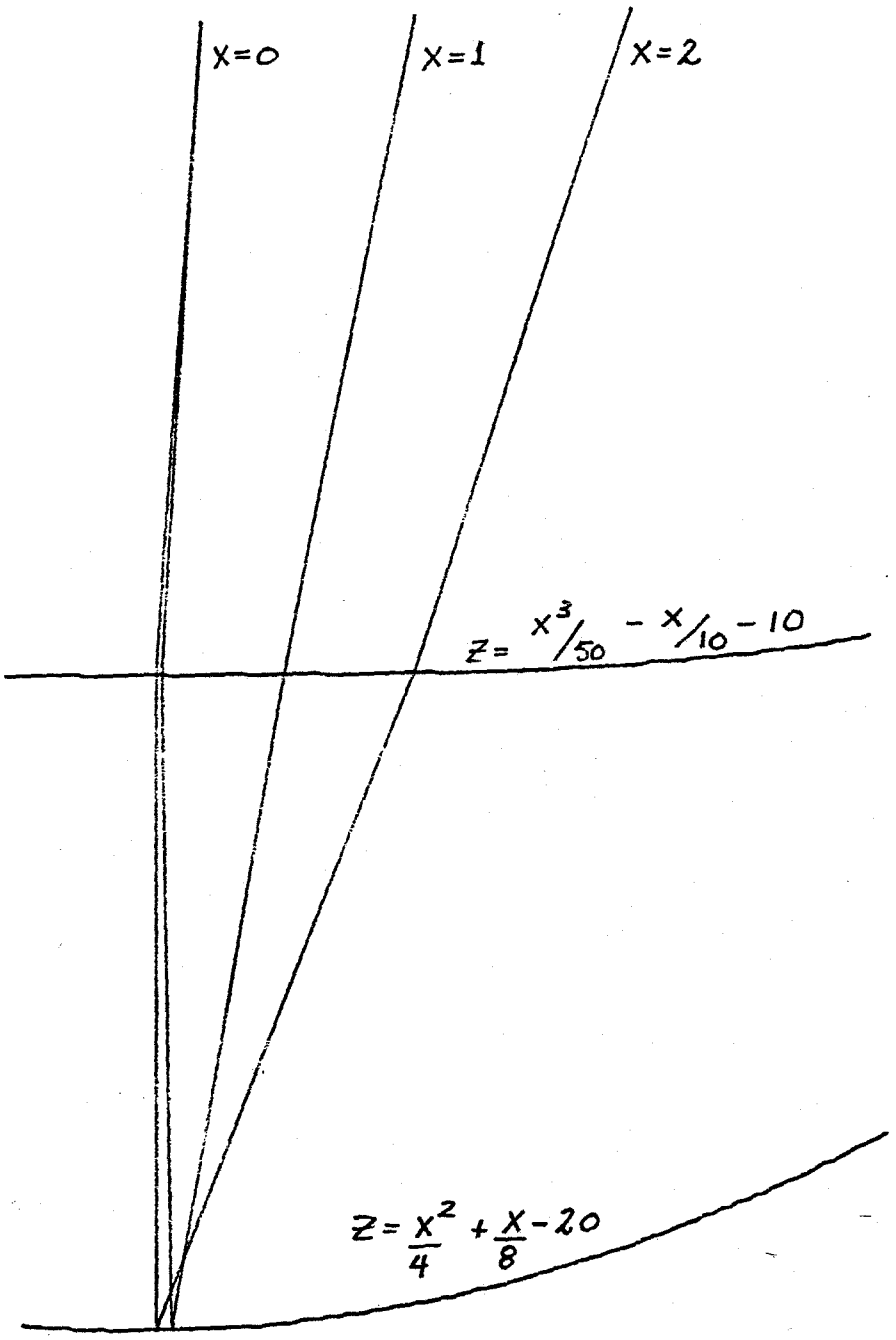


Figure 1.4b Ray Solution $\lambda = 1$; $\Delta\lambda = 1$

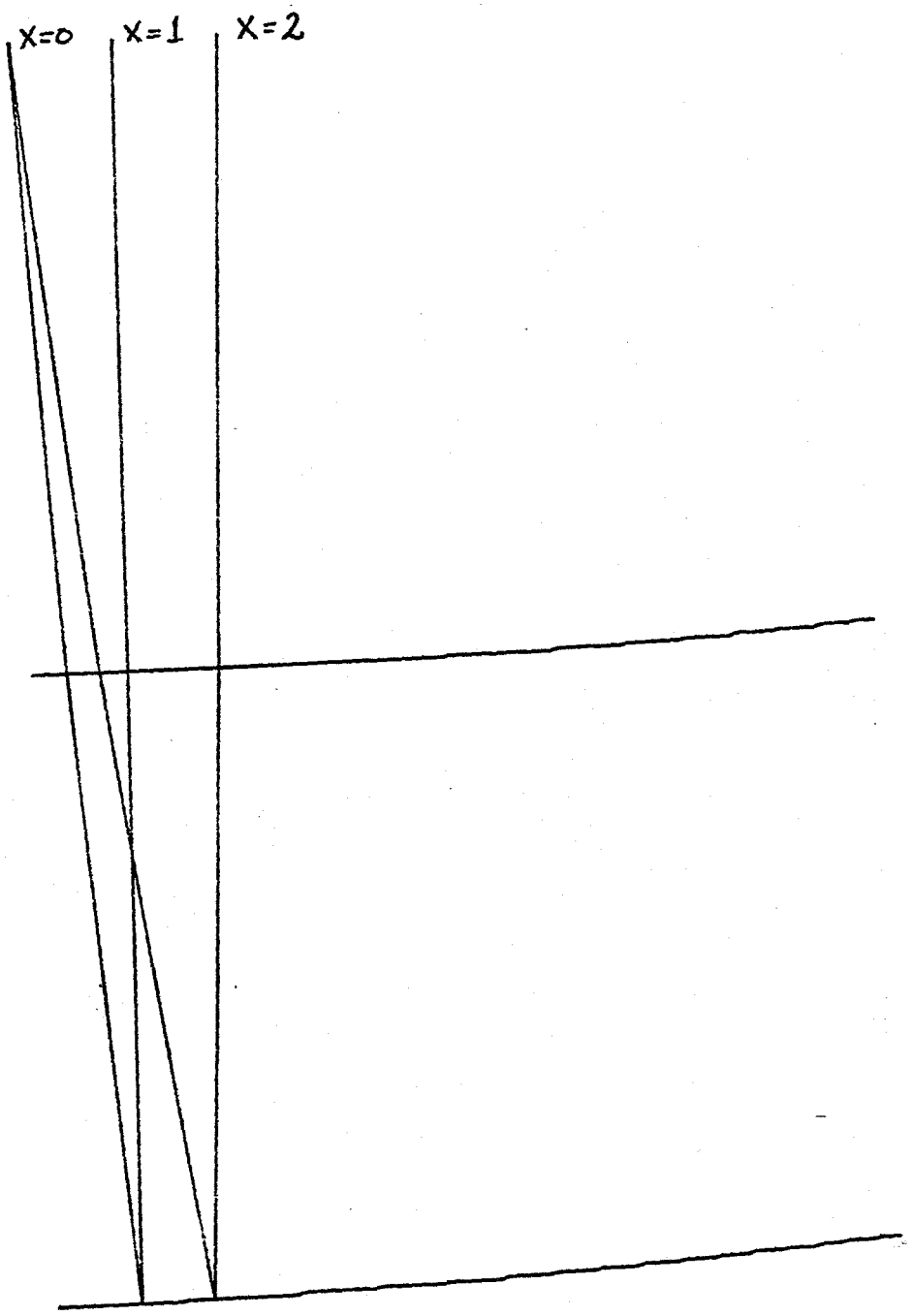


Figure 1.5a Ray Solution $\lambda = \frac{1}{32}$; $\Delta\lambda = \frac{1}{32}$

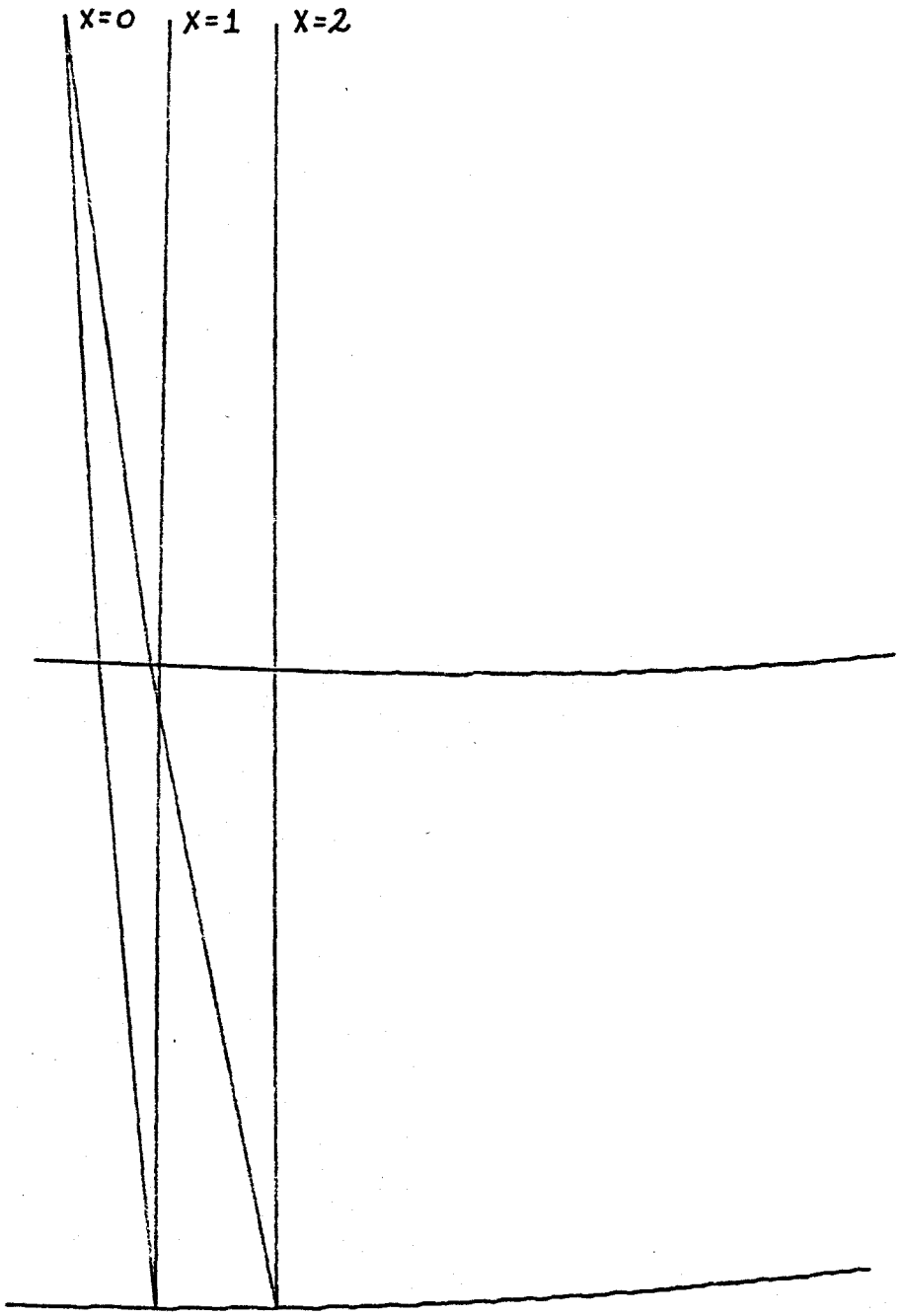


Figure 1.5b Ray Solution $\lambda = \frac{1}{16}$; $\Delta\lambda = \frac{1}{32}$

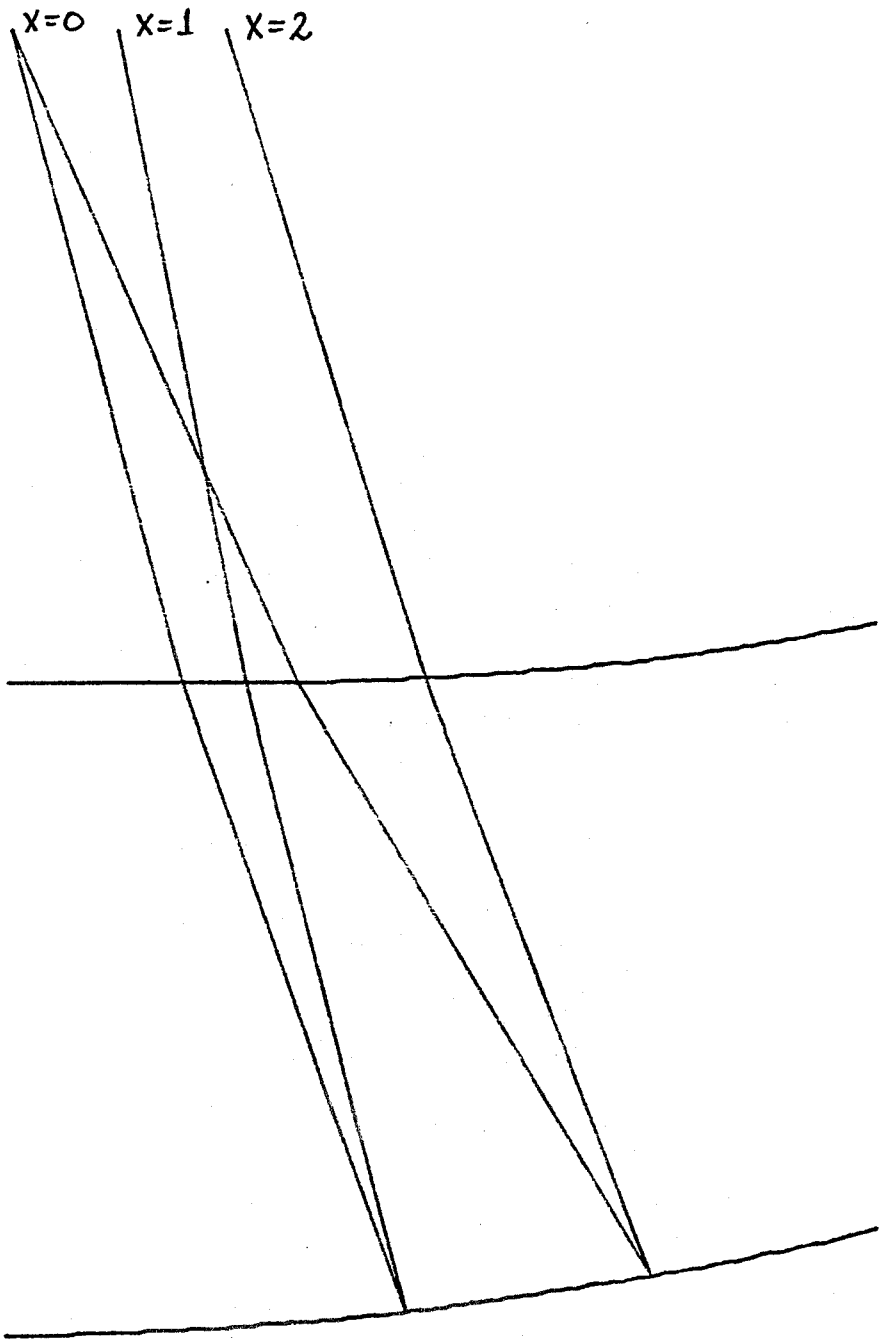


Figure 1.5c Ray Solution $\lambda = \frac{3}{32}$; $\Delta\lambda = \frac{1}{32}$

Section 2. The Calculation of Amplitude.

In the previous section we dealt with the kinematic determination of the seismic ray. From this, we can calculate the travel time along a ray. In this section, we shall discuss the determination of the first order amplitude coefficent $\underline{U}_o$. In general, $\underline{U}_o$ is complex and we can write

$$\underline{U}_o(\underline{X}) = \left|\underline{U}_o(\underline{X})\right| \underline{n}\, e^{i\varepsilon} \tag{2.1}$$

where $\underline{n}$ is the unit vector in the direction of the particle displacement, and ε is the phase. If the seismic source has a time history $F(t)=\int_{-\infty}^{\infty} S(\omega) e^{i\omega t} d\omega$, then according to seismic ray theory, the contribution to the seismic signal at receiver $\underline{X}_R$, $\underline{W}(\underline{X}_R,t)$, from the particular ray with travel time $\sigma(\underline{X}_R)$, is given by (see Cerveny and Ravindra [3]):

$$\underline{W}(\underline{X}_R,t) = \underline{n}\left|\underline{U}_o(\underline{X}_R)\right|\left(\frac{\cos\varepsilon}{\pi} Re\int_0^{\infty} e^{i\omega(t-\sigma(\underline{X}_R))} S(\omega)\, d\omega - \frac{\sin\varepsilon}{\pi} Im\int_0^{\infty} e^{i\omega(t-\sigma(\underline{X}_R))} S(\omega) d\omega\right) \tag{2.2}$$

The terms $-\frac{1}{\pi} Im\int_0^{\infty} e^{i\omega(t-\sigma(\underline{X}_R))} S(\omega)\, d\omega$ and $\frac{1}{\pi} Re\int_0^{\infty} e^{i\omega(t-\sigma(\underline{X}_R))} S(\omega) d\omega$ are a Hilbert Transform pair. From (2.2) we see that the phase "ε" has an important effect on the pulse shape.

We will break the calculation of $\underline{U}_o$ into three factors: 1) the calculation of transmission/reflection coefficents at each interface; 2) the calculation of the geometrical spreading factor for the ray tube; 3) the location of internal caustics. The formulae, which we use, are not new and have been derived by various authors. We have gathered together these various formulae required for numerical computation, for the sake of

completeness.

2.1 Calculation of Transmission/Reflection Coefficents at an Interface.

We consider an interface $z=f_{i_k}(x,y)+c_{i_k}$, separating two layers, A and B, with different elastic constants λ_A, μ_A, ρ_A ($c1=\sqrt{\frac{\lambda_A+2\mu_A}{\rho_A}}$, $c2=\sqrt{\frac{\mu_A}{\rho_A}}$) and λ_B, μ_B, ρ_B ($c3=\sqrt{\frac{\lambda_B+2\mu_B}{\rho_B}}$, $c4=\sqrt{\frac{\mu_B}{\rho_B}}$). At the node x_k, y_k, z_k there is an incident compressional or shear displacement. At this point on the interface, the following boundary conditions must hold: 1) continuity of particle displacement across the interface; 2)continuity of tractional forces across the interface. For these conditions to hold in general, requires that there be a transmitted compressional displacement $\underline{R}_3$, a transmitted shear displacement $\underline{R}_4$, a reflected compressional displacement $\underline{R}_1$, and a reflected shear displacement $\underline{R}_2$ as well as the incident displacement vector at the node. We now follow the paper of H.B.Keller[11].

Consider, at each interface, the following coordinate system. $\underline{\xi}_1$ is the unit normal at the k'th interface , and the orientation is chosen so that $\underline{\xi}_1 \cdot (\underline{X}_k - \underline{X}_{k-1}) > 0$. $\underline{\xi}_2$ is the unit tangent vector. It is in the plane of incidence determined by $(\underline{X}_k - \underline{X}_{k-1})$ and $\underline{\xi}_1$, and is orthogonal to $\underline{\xi}_1$. Its orientation is chosen so that $\underline{\xi}_2 \cdot (\underline{X}_k - \underline{X}_{k-1}) > 0$. $\underline{\xi}_3$ is defined as $\underline{\xi}_1 \times \underline{\xi}_2$. For the reflected and transmitted compressional displacements, $\underline{R}_1$ and $\underline{R}_3$, the displacement vector lies along the direction of their ray, and hence entirely in the plane of incidence and we can write:

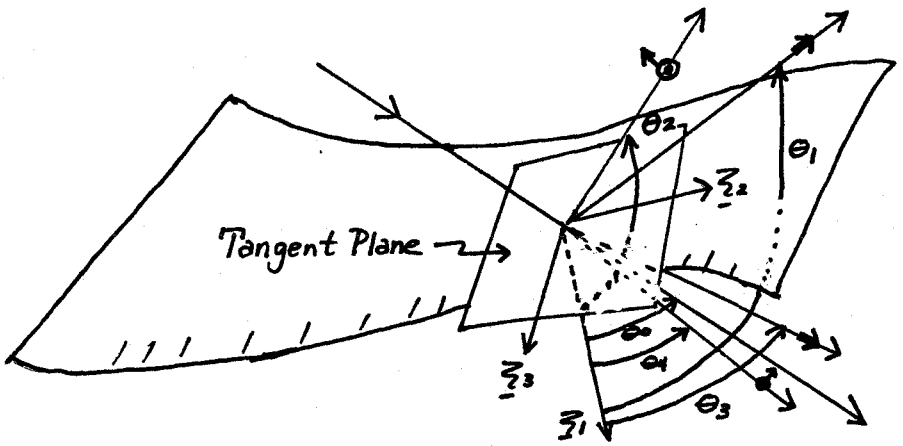
$$\underline{R}_1 = \alpha_1 \cos\theta_1 \underline{\xi}_1 + \alpha_1 \sin\theta_1 \underline{\xi}_2 \qquad (2.3a)$$

$$\underline{R}_3 = \alpha_3 \cos\theta_3 \underline{\xi}_1 + \alpha_3 \sin\theta_3 \underline{\xi}_2 \qquad (2.3b)$$

The shear displacements, $\underline{R}_2$ and $\underline{R}_4$, lie orthogonal to their rays and have a component in the plane of incidence (amplitude α_2 and α_4 respectively), the SV component, and in general a component normal to the plane of incidence, the SH component (β_2 and β_4), and we can write for these displacements:

$$\underline{R}_2 = \alpha_2 \sin\theta_2 \underline{\xi}_1 - \alpha_2 \cos\theta_2 \underline{\xi}_2 + \beta_2 \underline{\xi}_3 \qquad (2.3c)$$

$$\underline{R}_4 = \alpha_4 \sin\theta_4 \underline{\xi}_1 - \alpha_4 \cos\theta_4 \underline{\xi}_2 + \beta_4 \underline{\xi}_3 \qquad (2.3d)$$



(the small arrows indicate particle displacement direction)

Figure 1.6 Coordinate System Used at Each Interface

The incident wave is of the form

$$\underline{R}_o = \alpha_o \cos\theta_o \underline{\xi}_1 + \alpha_o \sin\theta_o \underline{\xi}_2 \quad (P) \tag{2.4a}$$

or

$$\underline{R}_o = \alpha_o \sin\theta_o \underline{\xi}_1 - \alpha_o \cos\theta_o \underline{\xi}_2 + \beta_o \underline{\xi}_3 \quad (S) \tag{2.4b}$$

From Snell's law, we can calculate all the necessary trigonometrical quantities; we get:

$$\sin\theta_j = \frac{c_j}{c_o} \sin\theta_o$$

$$\cos\theta_j = \sigma_j \sqrt{1-\sin^2\theta_j} \quad \text{or} \quad -\sigma_j i \sqrt{\sin^2\theta_j - 1} \qquad \sin\theta_j > 1 \tag{2.5}$$

$$\sigma_j = -1 \quad j = 1 \text{ or } 2 ; \quad \sigma_j = 1 \quad j = 3 \text{ or } 4$$

We note that for some of these "splitting" angles we may have to use complex trigonometrical quantities. At an interface, the boundary conditions decouple into separate boundary conditions for the P-SV coefficents (α_1 , α_2 , α_3 , α_4) and the SH coefficents (β_2 , β_4). If the incident ray is shear, then the SH amplitude coefficents (i.e., β_2 or β_4) are given by:

$$\beta_2 = -\beta_o \frac{\left(\frac{\mu_B}{c_4}\cos\theta_4 - \frac{\mu_A}{c_2}\cos\theta_2\right)}{\Delta}$$

$$\beta_4 = -\frac{2\rho_A c_2 \cos\theta_2}{\Delta}\beta_o \tag{2.6}$$

where:

$$\Delta \equiv \frac{\mu_B}{c_4}\cos\theta_4 - \frac{\mu_A}{c_2}\cos\theta_2$$

For the P-SV coefficents, Snell's law and the boundary conditions at the interface yield the following 4 by 4 system:

$$\begin{bmatrix} \cos\theta_1 & \sin\theta_2 & -\cos\theta_3 & -\sin\theta_4 \\ \sin\theta_1 & -\cos\theta_2 & -\sin\theta_3 & \cos\theta_4 \\ -\rho_A c_1 \cos 2\theta_2 & -\rho_A c_2 \sin 2\theta_2 & \rho_B c_3 \cos 2\theta_4 & \rho_B c_4 \sin 2\theta_4 \\ -\frac{\mu_A}{c_1}\sin 2\theta_1 & \rho_A c_2 \cos 2\theta_2 & \frac{\mu_B}{c_3}\sin 2\theta_3 & -\rho_B c_4 \cos 2\theta_4 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} = \begin{bmatrix} \alpha_0 K_1 \\ \alpha_0 K_2 \\ \alpha_0 K_3 \\ \alpha_0 K_4 \end{bmatrix} \tag{2.7}$$

Here for:

a) incident compressional displacement

$$K_1 = \cos\theta_1 \quad K_2 = -\sin\theta_1 \quad K_3 = \rho_A c_1 \cos 2\theta_2$$
$$K_4 = -\frac{\mu_A}{c_1}\sin 2\theta_1 \tag{2.8a}$$

b) incident shear wave

$$K_1 = -\sin\theta_2 \quad K_2 = -\cos\theta_2 \quad K_3 = -\rho_A c_2 \sin 2\theta_2$$
$$K_4 = -\rho_A c_2 \cos 2\theta_2 \tag{2.8b}$$

For a free surface, the system reduces to:

a) incident compression wave

$$\alpha_1 = -\alpha_0\left(\frac{c_1^2\cos^2 2\theta_2 + c_2^2 \sin 2\theta_2 \sin 2\theta_0}{c_1^2\cos^2 2\theta_2 - c_2^2 \sin 2\theta_2 \sin 2\theta_0}\right) \tag{2.9a}$$

$$\alpha_2 = \alpha_0 \frac{(2 c_1 c_2 \cos 2\theta_2 \sin 2\theta_0)}{(c_1^2 \cos^2 2\theta_2 - c_2^2 \sin 2\theta_2 \sin 2\theta_0)}$$

b) incident shear wave

$$\alpha_1 = -\alpha_0 \frac{(2 c_1 c_2 \cos 2\theta_0 \sin 2\theta_0)}{(c_1^2 \cos^2 2\theta_0 + c_2^2 \sin 2\theta_1 \sin 2\theta_0)}$$

$$\alpha_2 = \alpha_0 \frac{(-c_1^2 \cos^2 2\theta_0 + c_2^2 \sin 2\theta_1 \sin 2\theta_0)}{(c_1^2 \cos^2 2\theta_0 + c_2^2 \sin 2\theta_1 \sin 2\theta_0)} \qquad (2.9b)$$

$$\beta_2 = \beta_0$$

Thus we have the formula for the transmission and reflection coefficents at an interface in terms of the local coordinate system defined above. Unlike the two dimensional case, where the plane of incidence is constant for all interfaces, the plane of incidence, and hence the definition of the SV and SH components, changes at each interface for the three dimensional problem. Also, due to the local nature of our coordinate system, the definition of positive shear displacement direction can be reversed from one interface to the next. Thus, for the case of a transmitted or reflected shear wave, we have to recompute α_0 and β_0 for the coefficent calculations at the next interface. For a compressional ray this is not a consideration as the displacement is in the direction of the ray segment and hence is in the new plane of incidence. To calculate the new quantities α_0, β_0 for a shear wave, we simply multiply $\underline{R}_2$ or $\underline{R}_4$ by the matrix which relates the two coordinate systems, and find the new values.

2.2 Calculation of Geometrical Spreading Factor.

The Geometrical Optics approximation conserves energy in an infinitesimal ray tube about any ray. Thus, the amplitude of the

disturbance at a point $\underline{X}$ is (in Geometrical Optics) inversely proportional to the square root of the cross-sectional area of the tube at $\underline{X}$. If r_ξ and r_η are the principal radii of curvature of the wavefront at $\underline{X}$, the area of the ray tube $d\sigma$ is proportional to $r_\xi r_\eta$. If a sequence of interfaces is encountered we have a situation like that shown below in figure 1.7.

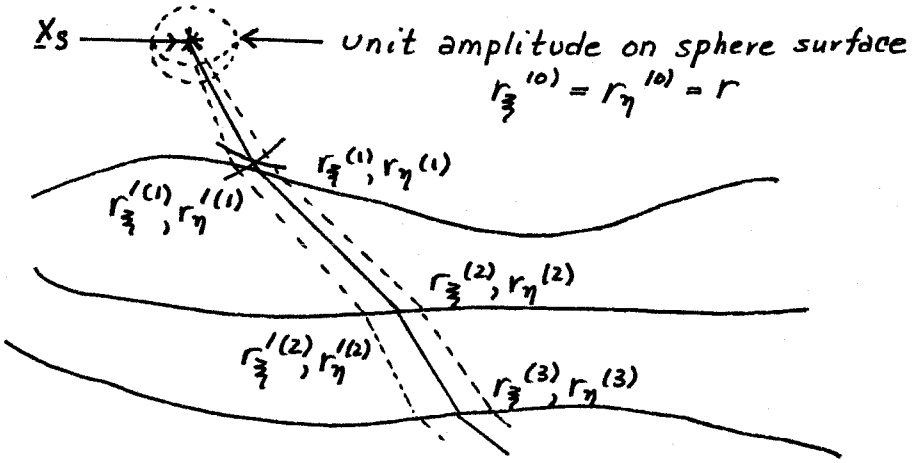


Figure 1.7 Ray Tube at Interfaces.

At interface 1 the amplitude is proportional to $(r_\xi^{(0)} r_\eta^{(0)} / r_\xi^{(1)} r_\eta^{(1)})^{1/2}$; thus at the second interface the amplitude is proportional to $\left[\left(\frac{r_\xi^{(0)} r_\eta^{(0)}}{r_\xi^{(1)} r_\eta^{(1)}}\right)\left(\frac{r_\xi'^{(1)} r_\eta'^{(1)}}{r_\xi^{(2)} r_\eta^{(2)}}\right)\right]^{1/2}$, or in general we have that the geometrical spreading factor at $\underline{X}_R$, $R(\underline{X}_S, \underline{X}_R)$ is given by:

$$\frac{1}{R(\underline{X}_S, \underline{X}_R)} = \left(\frac{\Gamma_\xi^{(0)} \Gamma_\eta^{(0)}}{\Gamma_\xi^{(FINAL)} \Gamma_\eta^{(FINAL)}} \prod_{i=1}^{N} \frac{(r_\xi'^{(i)} \Gamma_\eta'^{(i)})}{(\Gamma_\xi^{(i)} \Gamma_\eta^{(i)})} \right)^{1/2} \qquad (2.10)$$

Thus to calculate this factor we must know the radii of curvature of the wavefront prior to reflection/transmission and be able to calculate the curvatures after. The formulae for this are given by several authors; we

follow the work of Stavroudis[17].

Three local coordinate systems are defined for the incident wave, the interface, and the transmitted/reflected wave. For the incident wave we define $\underline{r}_1$, $\underline{P}$, and $\underline{Q}$; for the interface we use the coordinates $\underline{N}$, $\underline{P}$, and $\underline{Q}_i$, and for the emergent wave we use $\underline{r}_2$, $\underline{P}$, and $\underline{Q}_2$. The vectors $\underline{r}_1$ and $\underline{r}_2$ are the normalized incident and emergent rays. The vector $\underline{N}$ is the unit normal to the interface, chosen such that $\underline{r}_1 \cdot \underline{N} \geq 0$. We use the definitions:

$$\underline{P} \equiv (\underline{r}_1 \times \underline{N}), \quad \underline{Q}_1 \equiv \underline{r}_1 \times \underline{P}, \quad \underline{Q}_i \equiv \underline{N} \times \underline{P},$$
$$\underline{Q}_2 \equiv \underline{r}_2 \times \underline{P} \tag{2.11}$$

The incident and emergent layer velocities are denoted by V2 and V1. These coordinates are shown below in figure 1.8.

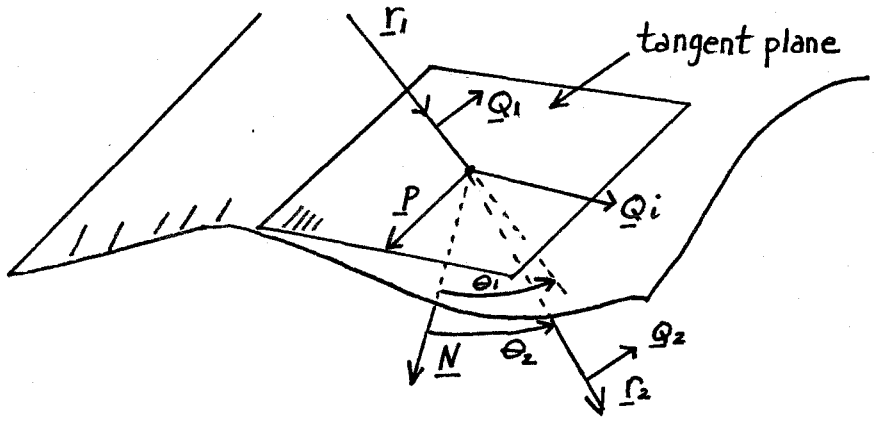


Figure 1.8 Local Coordinate Systems.

We now suppose that the two principal radii of curvature of the incident wavefront (the infinitesimal patch of the wavefront along the ray), r_ξ , r_η , and one of the principal direction vectors $\underline{V}_\xi$ are known. The cosine

of the angle between $\underline{V}_{\xi}$ and $\underline{P}$ is found from $\cos\theta = \underline{V}_{\xi} \cdot \underline{P}$. If we denote the radii of curvature of the wavefront in the $\underline{P}$ and $\underline{Q}$ directions by r_p , and r_q we have from Euler's thereom in differential geometry:

$$\frac{1}{r_p} = \frac{\cos^2\theta}{r_\xi} + \frac{\sin^2\theta}{r_\eta} \qquad (2.12a)$$

$$\frac{1}{r_q} = \frac{\sin^2\theta}{r_\xi} + \frac{\cos^2\theta}{r_\eta} \qquad (2.12b)$$

Denoting the torsion of the geodesic in the $\underline{Q}$ direction as $1/\sigma$ we can write:

$$\frac{1}{\sigma} = \frac{1}{2}\left(\frac{1}{r_\xi} - \frac{1}{r_\eta}\right)\sin 2\theta \qquad (2.13)$$

Adding (2.12a) and (2.12b) we obtain:

$$\frac{1}{r_p} + \frac{1}{r_q} = \frac{1}{r_\xi} + \frac{1}{r_\eta} \qquad (2.14)$$

Subtracting

$$\frac{1}{r_p} - \frac{1}{r_q} = \frac{\cos 2\theta}{r_\xi} - \frac{\cos 2\theta}{r_\eta} \qquad (2.15)$$

From (2.14) and (2.15) we obtain that:

$$\tan 2\theta = \frac{2}{\sigma}\,\frac{1}{\frac{1}{r_p} - \frac{1}{r_q}} \qquad (2.16)$$

By taking directional derivatives of Snell's vector law in the $\underline{P}$ and $\underline{Q}$ directions, we get:

$$\frac{1}{r_p'} = \frac{n}{r_p} + \frac{b}{R_p} \tag{2.17a}$$

$$\frac{\cos\theta_2}{\sigma'} = \frac{n\cos\theta_1}{\sigma} + \frac{b}{\bar{\sigma}} \tag{2.17b}$$

$$\frac{\cos^2\theta_2}{r_q'} = \frac{n\cos^2\theta_1}{r_q} + \frac{b}{R_q} \tag{2.17c}$$

where:

$b = \cos\theta_2 - n\cos\theta_1$

$n = V2/V1$

r_p', r_q' are the radii of curvature in the $\underline{P}, \underline{Q}_2$ directions for the transmitted/reflected wavefront.

σ' is the corresponding torsion

$Rp, Rq, \bar{\sigma}$ are the interface quantities.

In our notation, curvatures will be positive if the deviation of the surface from its tangent plane is in the direction of the defined normal vector. Here our normal vectors are $\underline{r_1}$, $\underline{N}$, and $\underline{r_2}$. A positive wavefront curvature corresponds to a converging wave. To begin the calculation of the spreading function we know that at the first interface

$$r_p = r_q = r_{initial} - d \tag{2.18}$$

where "d" is the distance from the source to the first interface. We can

take $\underline{V}_\xi = \underline{P}$ and hence $\theta = 0$ and $1/\sigma = 0$. From equations (2.17) we now calculate $\frac{1}{r_p'}$, $\frac{1}{r_q'}$, and $\frac{1}{\sigma'}$. An angle θ' between a principal direction and can be calculated from (2.16). We now calculate the new principal direction vector, $\underline{V}_\xi'$, and the principal curvatures $\frac{1}{r_\xi'}$ and $\frac{1}{r_\eta'}$.

$$\underline{V}_\xi' = \underline{P}\cos\theta' + \underline{Q}_2 \sin\theta' \tag{2.19a}$$

$$\frac{1}{r_\xi'} = \frac{\cos^2\theta'}{r_p'} + \frac{\sin^2\theta'}{r_q'} + \frac{\sin 2\theta'}{\sigma'} \tag{2.19b}$$

$$\frac{1}{r_\eta'} = \frac{\cos^2\theta'}{r_q'} + \frac{\sin^2\theta'}{r_p'} - \frac{\sin 2\theta'}{\sigma'} \tag{2.19c}$$

To transfer to the next interface we have the relation for the principal curvatures

$$\frac{1}{r_{\xi\,(new)}} = \frac{1}{r_\xi' - d} \quad ; \quad \frac{1}{r_{\eta\,(new)}} = \frac{1}{r_\eta' - d} \tag{2.20}$$

where "d" here is the distance travelled from the one interface to the next. The new angle θ_{new} is calculated from

$$\cos\theta_{new} = \underline{P}_{new} \cdot \underline{V}_\xi' \tag{2.21}$$

and then we determine $\frac{1}{r_p}$, $\frac{1}{r_q}$, and $\frac{1}{\sigma}$ from (2.12) and (2.13) and continue.

At each ray node, we must calculate the principal curvatures of the interface and the angle between P and a principal direction. We make the following definitions:

$$E = 1 + \left(\frac{\partial f}{\partial x}\right)^2 \tag{2.22a}$$

$$F = \left(\frac{\partial f}{\partial x}\right)\left(\frac{\partial f}{\partial y}\right) \tag{2.22b}$$

$$G = 1 + \left(\frac{\partial f}{\partial y}\right)^2 \tag{2.22c}$$

$$L = \frac{\frac{\partial^2 f}{\partial x^2}}{\sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}} \tag{2.22d}$$

$$M = \frac{\frac{\partial^2 f}{\partial x \partial y}}{\sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}} \tag{2.22e}$$

$$N = \frac{\frac{\partial^2 f}{\partial y^2}}{\sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}} \tag{2.22f}$$

Here, we are taking the interface to be of the form z=f(x,y). The principal curvatures and the associated principal directions can be found from the eigenvalues and eigenvectors of the system (see Stoker[18]):

$$(\underline{\underline{A}} - k\underline{\underline{B}})\,\underline{X} = \underline{O} \tag{2.23}$$

where:

$$\underline{\underline{A}} \equiv \begin{pmatrix} L & M \\ M & N \end{pmatrix} \tag{2.24a}$$

$$\underline{\underline{B}} \equiv \begin{pmatrix} E & F \\ F & G \end{pmatrix} \tag{2.24b}$$

The eigenvalues k give the principal curvatures and the principal directions are determined from the corresponding eigenvectors. Using the notation $\underline{x}=(x,y)$, the principal direction is given by

$$\underline{X}_P = \frac{x\left(1, 0, \frac{\partial f}{\partial x}\right) + y\left(0, 1, \frac{\partial f}{\partial y}\right)}{\sqrt{x^2 + y^2 + \left(x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y}\right)^2}} \qquad (2.25)$$

Using the above described methods we can calculate the exact geometric spreading function for a ray as given by (2.10).

2.3 Calculation of the Location of Caustics.

In section 2.1, we pointed out that in general, we have a complex valued system to solve at each interface, which leads to complex reflection/ transmission coefficents and a phase shift of the seismic signal. The other source of phase shifts along the ray is when the ray tube passes through a point where the ray tube has zero cross-sectional area. Equivalently, this point is where one of the 2 principal radii of curvature of the wavefront is zero. We call such points caustics. Near these points, $|\underline{U}_0|$ becomes infinite and the ray theory is not valid. However, along the ray, away from such a point, we can "patch up" the theory by introducing a phase shift of $\pi/2\ \mathrm{sgn}\,\omega$ to the signal. The location of these caustics is very easy to find in our formulation. As outlined in the previous section, at each interface we calculate the wavefront curvatures r_ξ and r_η. If r_ξ or r_η is positive (say r_ξ) then the wavefront is converging along the principal direction. At the next interface (or at the receiver) the radius will be $r_\xi - d$. Hence if $r_\xi - d$ is

negative then the ray tube has passed through a caustic in this layer. The location of this caustic point is given by

$$\underline{X}_{caustic} = \underline{X}_k + \frac{r_1}{d}(\underline{X}_{k+1} - \underline{X}_k) \qquad (2.26)$$

Here "d" is the distance $||\underline{X}_{k+1} - \underline{X}_k||_2$ and $\underline{X}_{k+1}, \underline{X}_k$ are the "k+1" and "k" ray node points. It is possible for there to be two different caustics within a layer corresponding to both radii vanishing at different points within the layer. If a principal radius of curvature is negative at the k'th node, then a caustic along the next ray segment, corresponding to this principal direction, is not possible.

2.4 Modifications For Receiver at Free Surface

If a receiver is situated right on a free surface, then the observed seismic response is the sum of the incident ray and the reflected and converted displacements at the surface. Formulae (conversion vectors) relating the incident displacement and the observed free surface response are given in Cerveny et al.[3]. Unless otherwise stated, we will, in our numerical examples, be calculating only the incident amplitudes.

Section 3. Numerical Implementation and Examples

3.1 Numerical Implementation

Here we outline some of the characteristics of the ray-tracing codes developed. The codes were all written in Fortran-IV and implemented on a VAX 11/780. The user inputs analytic expressions for the interfaces encountered by the ray into a subroutine. The analytic expressions for $\partial f_k/\partial x$, $\partial f_k/\partial y$, $\partial^2 f_k/\partial x^2$, $\partial^2 f_k/\partial y^2$, and $\partial^2 f_k/\partial x \partial y$ are also user supplied for each interface. The mean levels $\{c_k\}$ of the interfaces, the sequence of the interfaces encountered, the layers' shear and compressional velocities, source and receiver positions and other required information are input at the terminal at the beginning of execution of the program.

Using the values $\{c_k\}$, the parallel plane problem is quickly solved for the purely compressional ray. The details of this problem are outlined in Appendix A. At this step we form the Jacobian matrix J, for the system formulation used, with $\lambda = 0$, $\frac{\partial f_k}{\partial x} = 0$, $\frac{\partial f_k}{\partial y} = 0$ etc. We then find the L-U decomposition of $\underline{\underline{J}}$ and calculate $d\underline{x}/d\lambda$. The Jacobian for all formulations (see figure 1.9) for N>1, has 7 bands, and thus we use a banded system solver. For many of the other continuation steps used in the programs, the L-U decomposition of $\underline{\underline{J}}$ already exists, from previous Newton iterations, and hence can be reused. The quantity

$$\underline{X}^{(0)} = \underline{X}(0) + \frac{d\underline{X}}{d\lambda} \Delta\lambda \qquad (3.1)$$

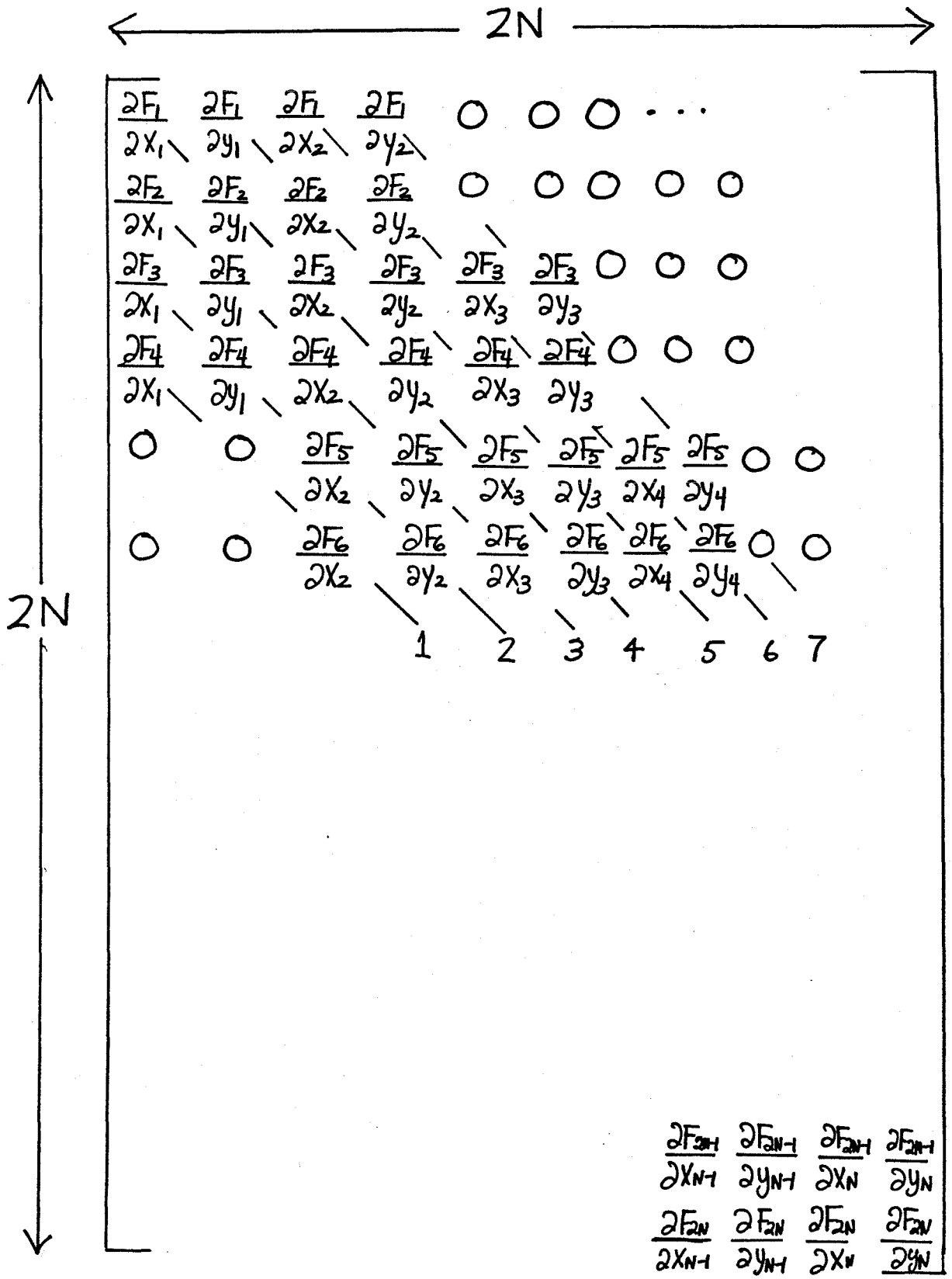


Figure 1.9 Banded Structure of Jacobian

with $\Delta\lambda = 1$ is now formed. The continuation parameter λ is set equal to 1, the Jacobian matrix and the residual vector of (1.17) or (1.18) is formed, and the Newton iterations begin. If the number of iterations exceeds 7 without the L_2 norm of the residual vector being less than .0001, we return to the previous value of λ, and recalculate (3.1) with $\Delta\lambda = \Delta\lambda/2$. We will now have to take two of these continuation steps. This process is continued until either we reach a $\Delta\lambda$ that is smaller than an user input minimum or we solve our problem to within a residual less than .0001 .

Once we have found the purely compressional ray, there is the option to generate the rays for the various P-S velocity permutations. Velocity continuation is used between the successive velocity vectors to generate these rays. The algorithm used sequentially arranges 2^K velocity vectors, $0 < k < N$, corresponding to the first k+1 layers, so that each differs from the previous one by only one component. As an example for k=N=3 we calculate for [P,P,P,P], [P,S,P,P], [P,S,S,P], [P,P,S,P], [P,P,S,S], [P,S,S,S], [P,S,P,S], and [P,P,P,S] (we use the notation that "P" denotes a compressional ray segment and "S" a shear segment). Here we have assumed that we wish the first ray segment always to be compressional in the calculations. The parameter k is input at the terminal at the beginning of the execution of the program. One can also return to the saved purely compressional ray solution and use continuation in receiver position to generate a seismic gather. For each ray, using the formulae of section 2, we can quickly calculate the ray's amplitude and phase.

We checked the amplitude/phase calculations by running some simple examples where analytic answers are known. For example, for a stack of

parallel plane layers, an analytic expression for the geometrical spreading function is known (see Ceverny,Ravindra[4]). We also checked to make sure that the 2 different formulations of the kinematic problem ((1.17) and (1.18)) gave the same variational derivatives and solutions for the same problems. The program will, if desired, print out the ratios of the cartesian components of $\frac{V_{k+1}}{d_k}(\underline{X}_k - \underline{X}_{k-1}) \times \underline{N}_k$ and $\frac{V_k}{d_{k+1}}(\underline{X}_{k+1} - \underline{X}_k) \times \underline{N}_k$ for k=1 to N. These 3 ratios should all numerically be equal to one, if in fact a physical ray has been found.

3.2 Numerical Examples.

Example 1.

In example 1, we shall show the various stages of calculations of a purely compressional ray [Vp,1;Vp,2;Vp,2;Vp,1] and the ray [Vp,1;Vs,2;Vp,2;Vp,1] for the following interfaces:

```
z1: X**3/10 + Y**3/20 -X/5 + Y/4 -5
                                                    (3.2)
z2: -X**2/20 - Y**2/10 - 10
```

The velocities for this model are: Vp,1=6, Vs,1=3, Vp,2=8, Vs,2=4. Here, we take the sequence of interfaces to be as shown below in figure 1.10.

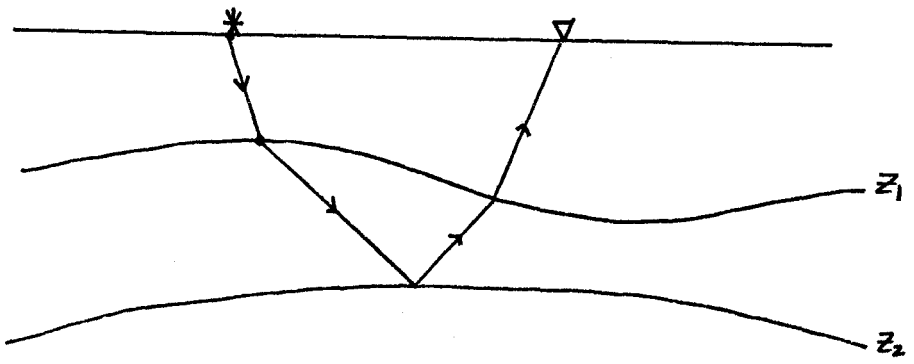


Figure 1.10 Schematic Cross Section for Ray

We calculate the rays from the source (x=0,y=0,z=0) to receiver (x=1,y=1,z=0). The calculations, shown below, used formulation (1.18). The initial planar estimate , shown below in Table 1a, for the purely compressional ray, required 1 bisection and 2 scalar Newton iterations (see Appendix A).

Table 1a. Planar Solution

node	"x"	"y"
0	0.000000	0.000000
1	0.214111	0.214111
2	0.500000	0.500000
3	0.785889	0.785889
4	1.000000	1.000000

Table 1b. Initial Ray Estimate

node	"x"	"y"
0	0.000000	0.000000
1	-0.163438	0.182553
2	-0.067111	-0.001901
3	0.539834	0.824263
4	1.000000	1.000000

The initial ray estimate: $\underline{X}^{(0)} = \underline{X}(0) + \frac{d\underline{x}}{d\lambda}$ is shown above in Table 1b. With this initial estimate we have a residual, $||F(X)||_2 = 1.80$. The two successive Newton iterations had residuals: 1. 1.255E-2 2. 6.557E-6. The resulting ray solution is shown below in Table 1c.

Table 1c. Ray Solution velocities 6,8,8,6

node	"x"	"y"	"deviation from plane"
0	0.000000	0.000000	0.000000
1	-0.069961	0.326969	0.097448
2	0.169430	0.329889	-0.012318
3	0.594940	1.026166	0.212640
4	1.000000	1.000000	0.000000

Table 1d. Estimated Ray 6,4,8,6

node	"x"	"y"
0	0.000000	0.000000
1	0.201051	-0.037784
2	0.209590	0.178123
3	0.612635	0.947115
4	1.000000	1.000000

We now use the solution of Table 1.c to continue to the ray for the

velocity sequence [Vp,1;Vs,2;Vp,2;Vp,1]. Using a single Euler step, the estimated ray for this velocity sequence is shown above in Table 1d. With this estimate, we have an initial residual of .3633 and the residuals for the following two Newton iterations are: 1. 1.6537E-2 2. 3.0816E-5. The resulting ray solution is shown below in Table 1e.

Table 1e. Ray Solution velocities 6,4,8,6

node	"x"	"y"	"deviation from plane"
0	0.000000	0.000000	0.000000
1	0.346168	-0.214290	-0.119150
2	0.298160	0.069099	-0.004922
3	0.661044	0.891464	0.154966
4	1.000000	1.000000	0.000000

We also calculated the compressional ray to the receivers (x=1,y=1,z=0), (x=1.5,y=1.5,z=0), and (x=2,y=2,z=0). We show the calculations for the interfaces corresponding to λ =0, λ =.5 and λ =1 in figures 1.11a-1.11c. We also used our other formulation, (1.17), to calculate the above Tables. There were, outside of small numerical effects, no differences between the results. At each stage, formulation (1.18) always required as few, or one less Newton iteration, than did formulation (1.17).

Example 2.

In this section, we illustrate the use of continuation in receiver location. Our interfaces are analytically:

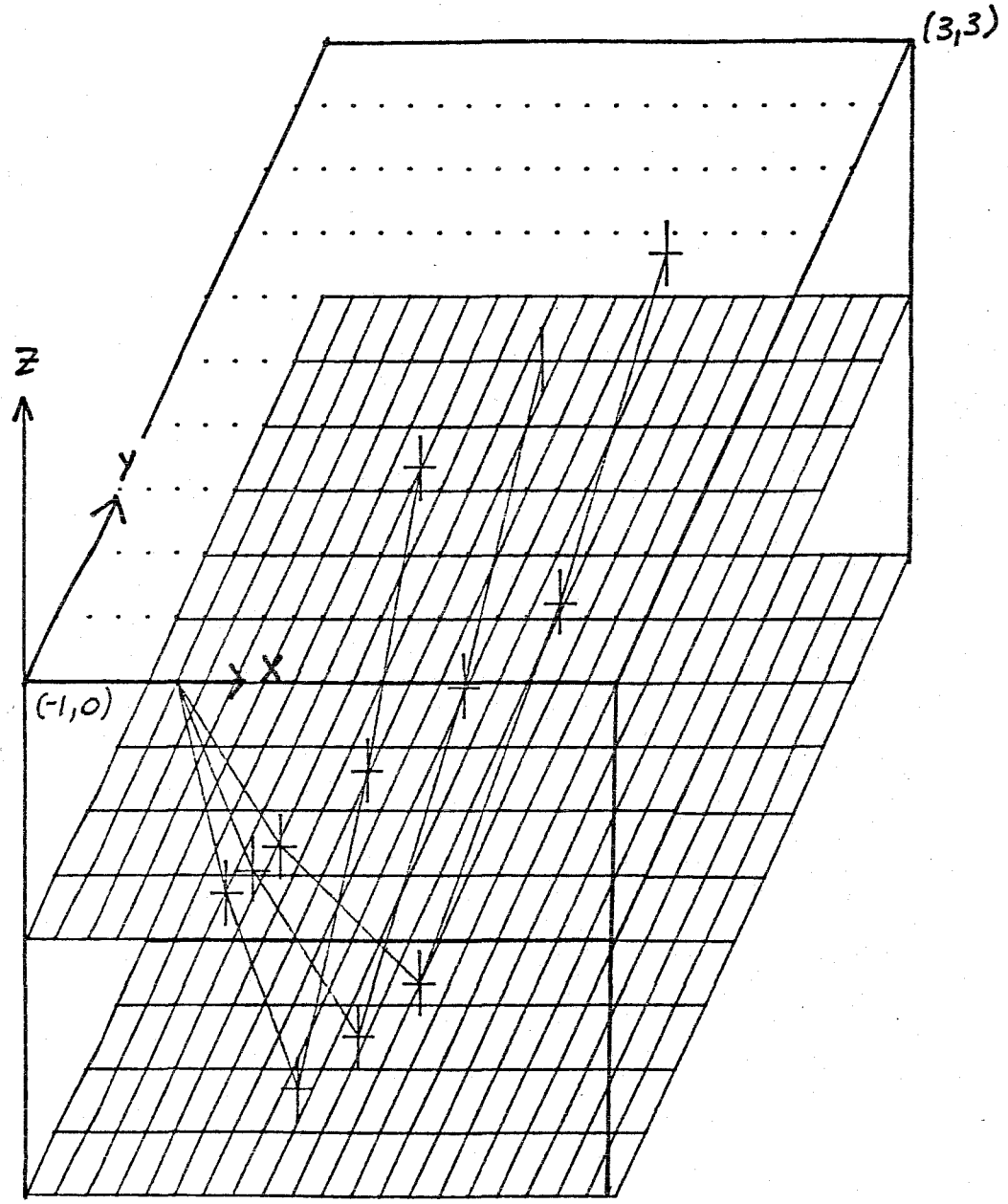


Figure 1.11a
Rays for Example 1 $\lambda = 0$
(Note: different scales for x,y,z coordinates)

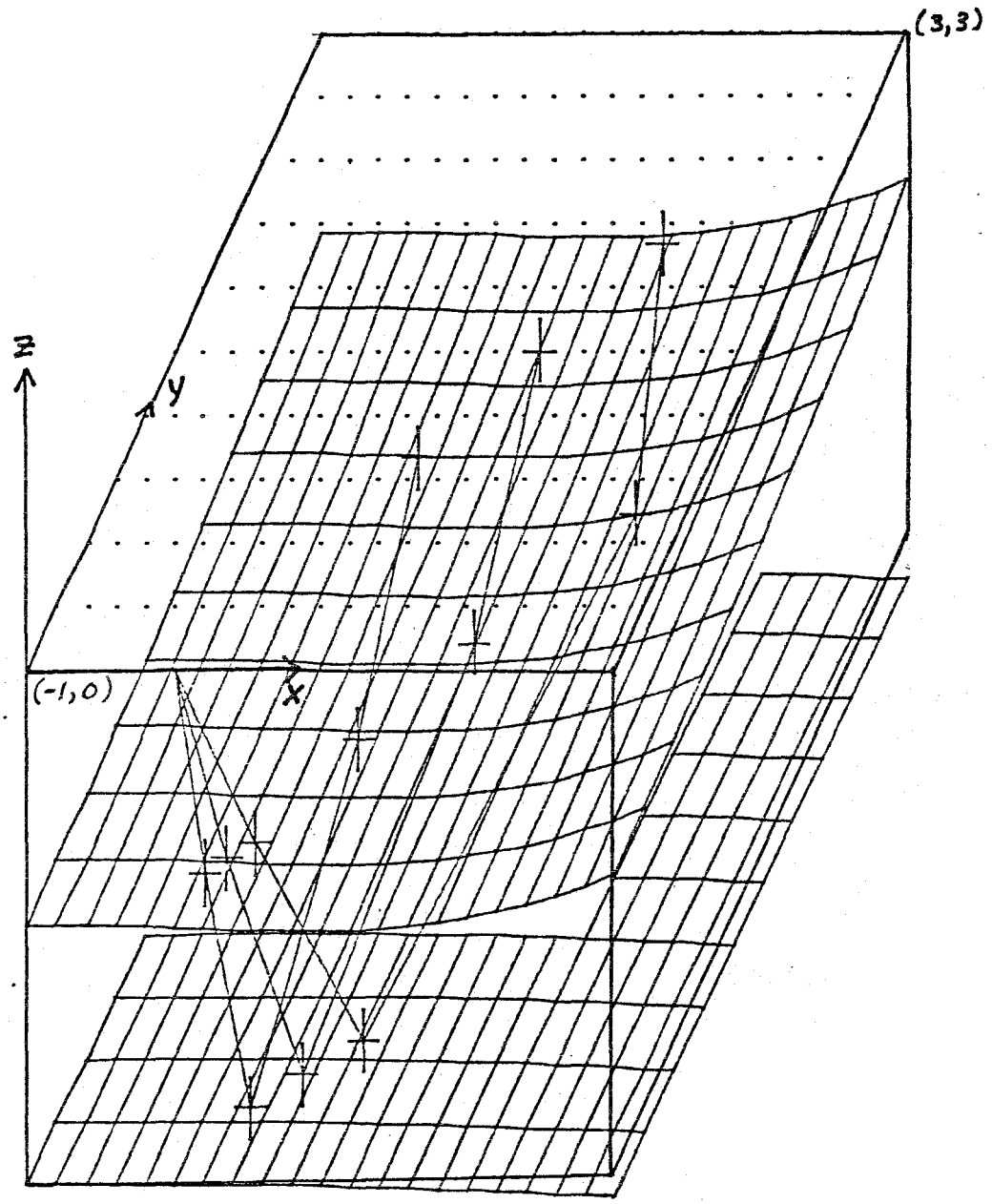


Figure 1.11b
Rays for Example 1 $\lambda = \frac{1}{2}$

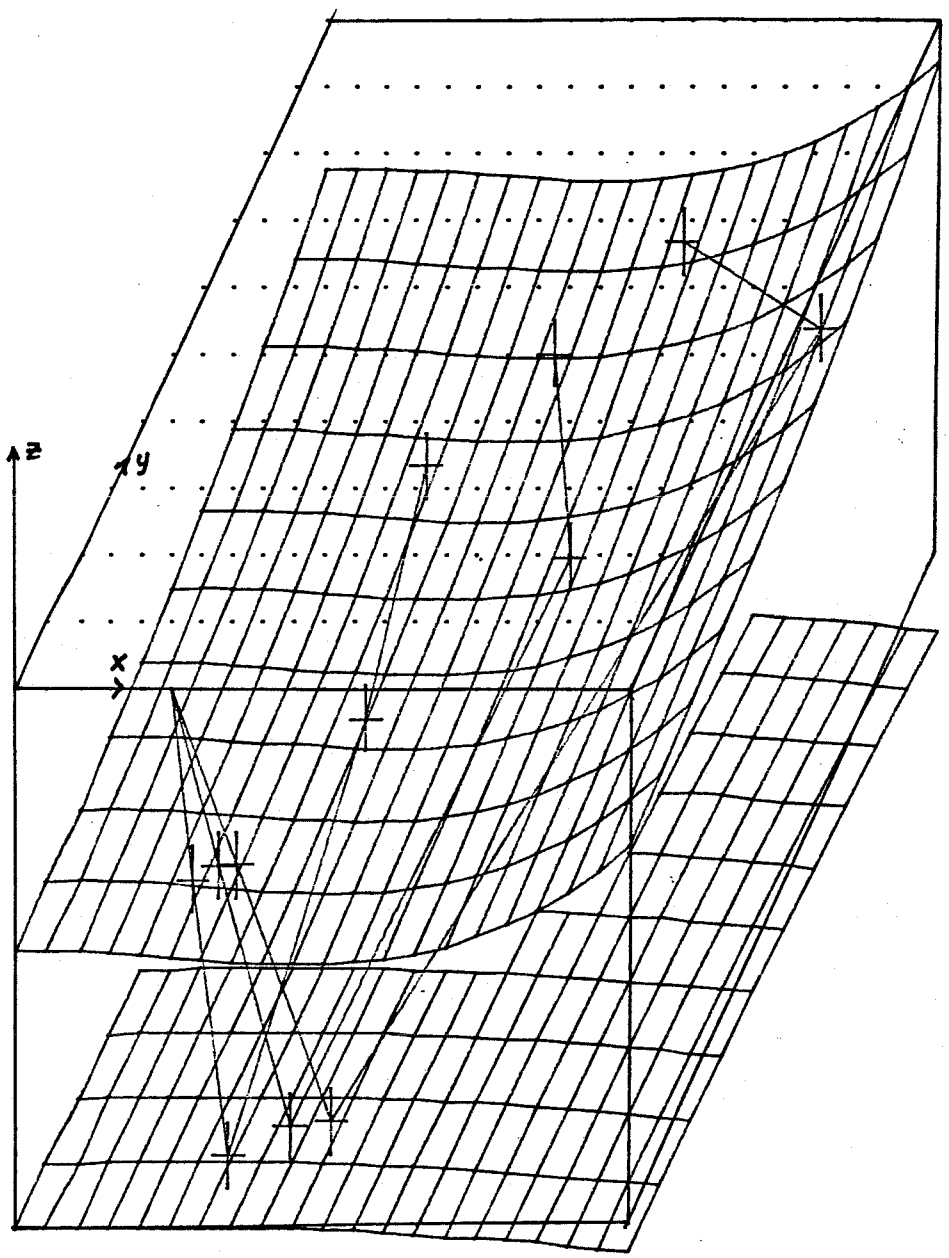


Figure 1.11c
Rays for Example 1 $\lambda = 1$

z1: -x/20 +y/15 -5

z2: sin(x/4) * sin(y/8) -10 (3.3)

The receivers are located at (X_i=1+.5i, y_i=2+i, z=0; i=1,4) and the source is located at (0,0,0). The interface sequence is shown below in figure 1.12.

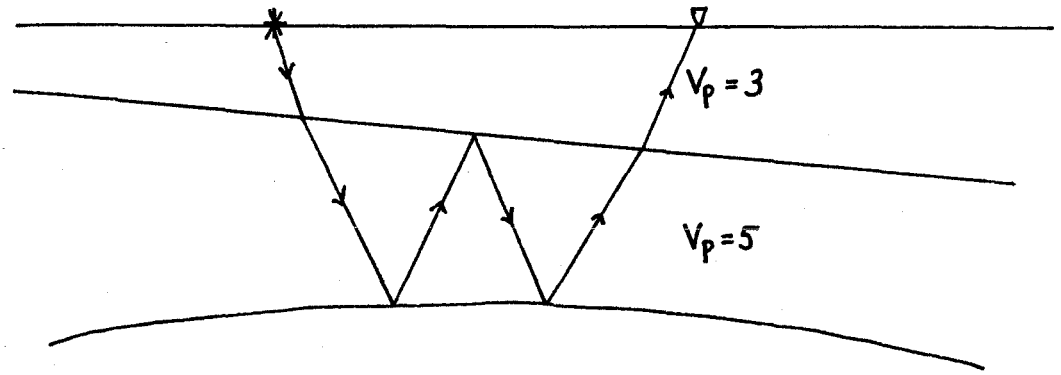


Figure 1.12 Schematic Cross Section for Ray

In Tables 2a-2d we show the calculated ray solutions and Tables 3a-3c show the estimates obtained from

$$\underline{X}(\underline{X}_{R,i+1}) = \underline{X}(\underline{X}_{R,i}) + \frac{d\underline{X}}{d\lambda} \qquad (3.4)$$

$$\underline{X}_R(\lambda) = (1-\lambda)\underline{X}_{R,i} + \lambda \underline{X}_{R,i+1}$$

The ray solutions are plotted in figure 1.13b and the ray solutions for the parallel plane interfaces are shwon in figure 1.13b.

Table 2a. Ray Solution velocities 3,5,5,5,5,3

node	"x"	"y"	"deviation from plane"
0	0.000000	0.000000	0.000000
1	0.355616	0.399356	0.008843
2	1.122941	0.843203	0.029148
3	1.629240	.940733	-0.018746
4	1.641815	1.692839	0.083807
5	1.175562	1.969386	0.072514
6	1.000000	2.000000	0.000000

Table 3a. Predicted Ray Solution

node	"x"	"y"
0	0.000000	0.000000
1	0.516316	0.601642
2	1.559414	1.384762
3	2.181631	1.689298
4	2.311005	2.640265
5	1.739174	2.948629
6	1.500000	3.000000

Table 2b. Ray Solution velocities 3,5,5,5,5,3

node	"x"	"y"	"deviation from plane"
0	0.000000	0.000000	0.000000
1	0.502850	0.591524	0.014292
2	1.525755	1.359347	0.062949
3	2.140103	1.660176	0.003673
4	2.261189	2.604634	0.171337
5	1.720275	2.936187	0.109732
6	1.500000	3.000000	0.000000

Table 3b. Predicted Ray Solution

node	"x"	"y"
0	0.000000	0.000000
1	0.637028	0.773567
2	1.895959	1.850239
3	2.611098	2.350272
4	2.832910	3.480368
5	2.248235	3.888600
6	2.000000	4.000000

Table 2c. Ray Solution velocities 3,5,5,5,5,3

node	"x"	"y"	"deviation from plane"
0	0.000000	0.000000	0.000000
1	0.625378	0.764804	0.019718
2	1.866992	1.828899	0.101978
3	2.575636	2.325787	0.026271
4	2.790357	3.450616	0.268561
5	2.232003	3.878166	0.146944
6	2.000000	4.000000	0.000000

Table 3c. Predicted Ray Solution

node	"x"	"y"
0	0.000000	0.000000
1	0.737077	0.929828
2	2.181398	2.278574
3	2.978514	2.968482
4	3.280462	4.268752
5	2.729237	4.809729
6	2.500000	5.000000

Table 2d. Ray Solution velocities 3,5,5,5,5,3

node	"x"	"y"	"deviation from plane"
0	0.000000	0.000000	0.000000
1	0.727788	0.922709	0.025125
2	2.158582	2.261891	0.143351
3	2.950965	2.949275	0.049070
4	3.247656	4.245577	0.367254
5	2.717052	4.800898	0.184207
6	2.500000	5.000000	0.000000

From these tables we can see that the estimated solutions for the next receiver location are very good,and we only require one or two Newton iterations to sufficiently refine this estimate.

Example 3.

In this example, we consider the interfaces:

```
z1:  -5
z2:  -10                                              (3.5)
z3: x**2/10.+ y**2/20.-20.
```

The sequence of interfaces encountered is shown in figure 1.14. Due to the parabolic nature of the reflecting interface, we expect that there might be multiple solutions to this ray tracing problem. For a source at the origin, and a receiver at (x=2.0,y=2.0,z=0), we find that using the two

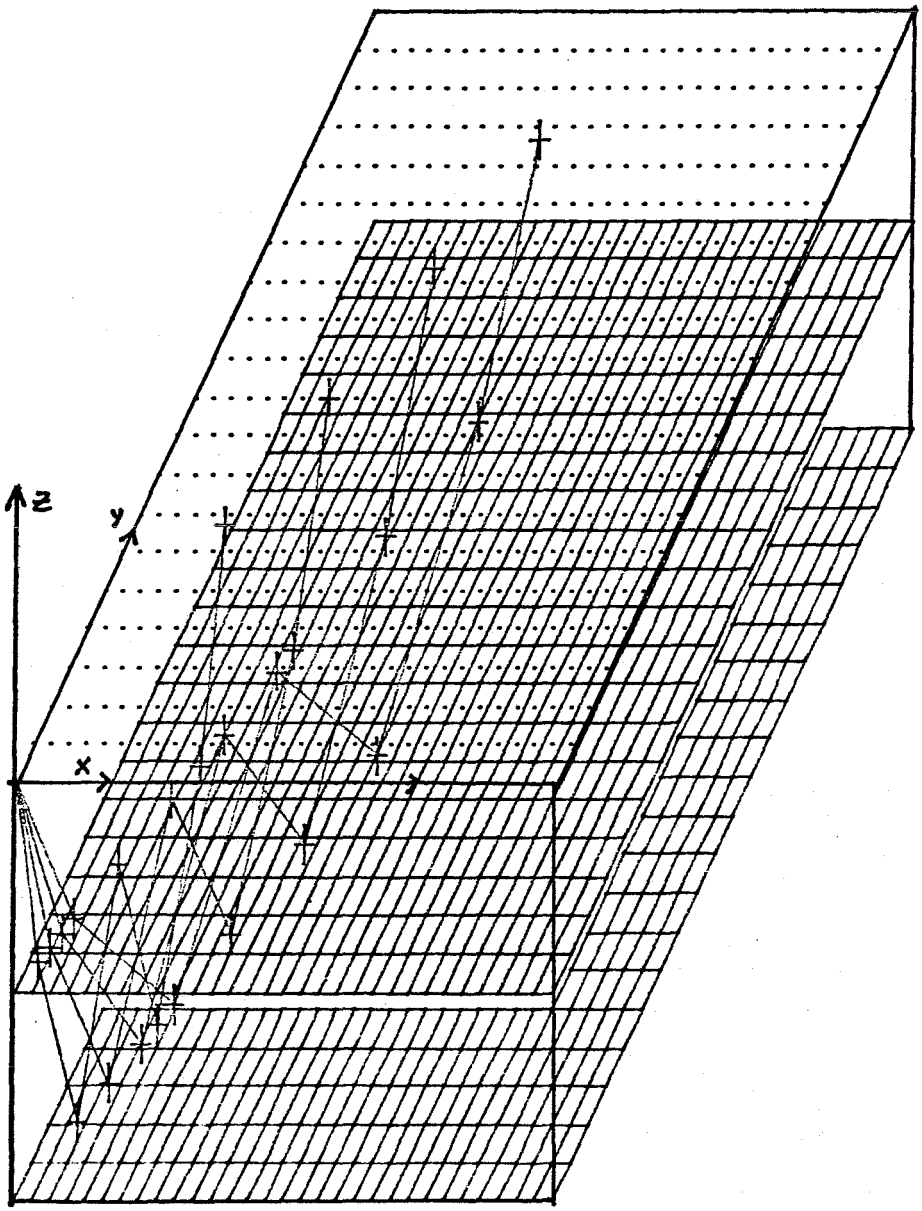


Figure 1.13a
Rays for Example 2 $\lambda = 0$

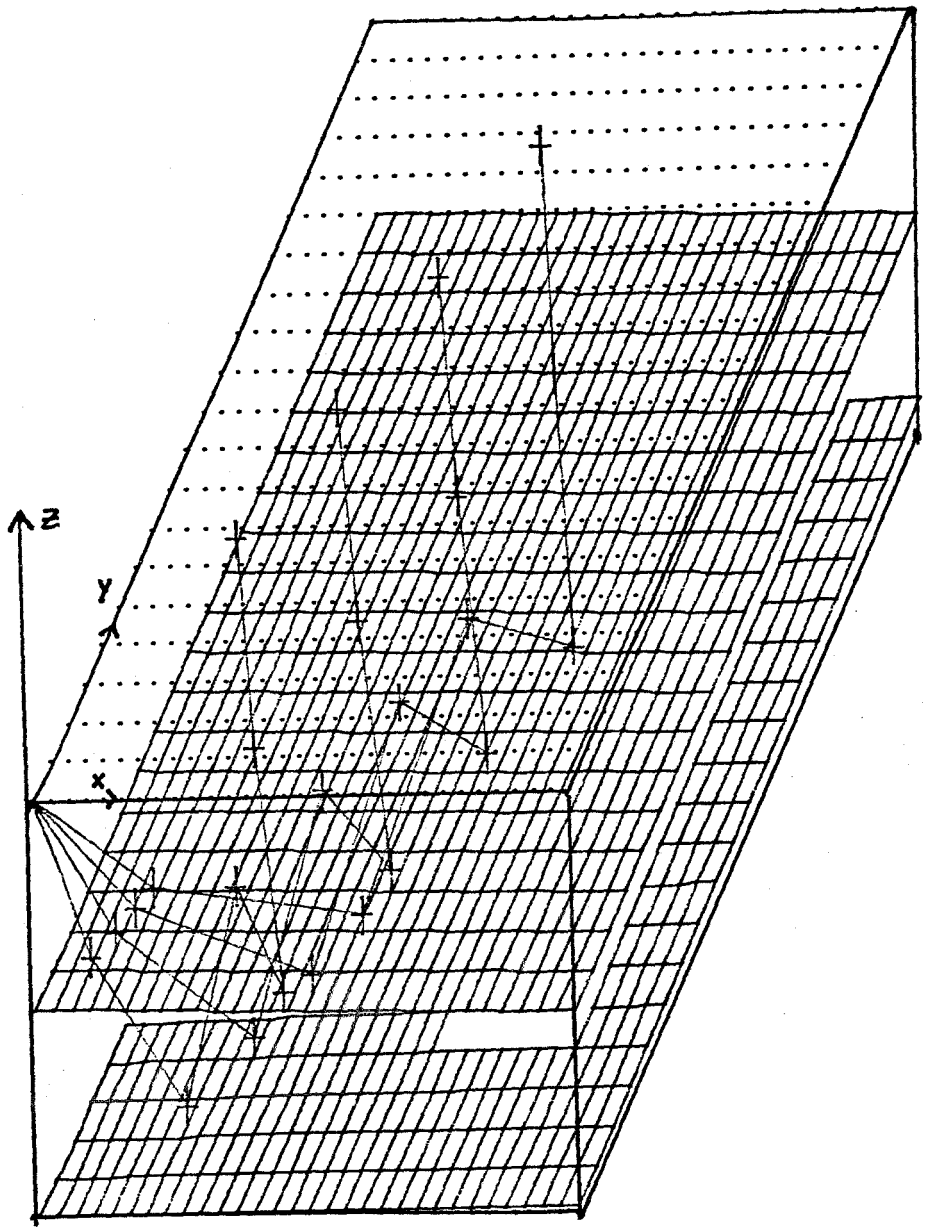


Figure 1.13b
Rays for Example 2 $\lambda = 1$

different formulations (1.17) and (1.18) for the problem lead to very different ray solutions. Formulation (1.18) converges with a step length of $\Delta\lambda = 1$ to find a compressional ray. Formulation (1.17) uses $\Delta\lambda = 1/8$. before the ray solution at $\lambda = 1$ can be found. The two different ray solutions are shown below in Tables 4a and 4b.

Table 4a. Ray Solution velocities 6,8,12,12,8,6 formulation (1.18)

node	"x"	"y"	"deviation from plane"
0	0.000000	0.000000	0.000000
1	-0.144600	-1.568824	0.000000
2	-0.345304	-3.746355	0.000000
3	-0.782553	-8.490256	3.665462
4	0.833719	-2.396894	0.000000
5	1.516813	0.178380	0.000000
6	2.000000	2.000000	0.000000

Table 4b. Ray Solution velocities 6,8,12,12,8,6 formulation (1.17)

node	"x"	"y"	"deviation from plane"
0	0.000000	0.000000	0.000000
1	-0.256438	2.042790	0.000000
2	-0.623406	4.966071	0.000000
3	-1.289890	10.275302	5.445473
4	0.344558	6.164054	0.000000
5	1.313466	3.726889	0.000000
6	2.000000	2.000000	0.000000

Neither of the ray solutions above correspond to the solution path that originates at $\lambda = 0$. To properly follow this path we force the two formulations to take very small steps in λ ($\Delta\lambda = 1/1024$ and $1/2048$), and we are lead to the ray solution at $\lambda = 1$ shown in Table 4c.

Table 4c. Ray Solution velocities 6,8,12,12,8,6

node	"x"	"y"	"deviation from plane"
0	0.000000	0.000000	0.000000
1	2.566463	0.531342	0.000000
2	6.425517	1.330293	0.000000
3	9.296496	1.924680	8.827704
4	7.573200	1.942469	0.000000
5	4.270069	1.976567	0.000000
6	2.000000	2.000000	0.000000

We can find yet another ray solution to the problem by first calculating the solution for the receiver at (x=.5,y=.5,z=0) and then using continuation in receiver position to generate, in one continuation step (both formulations), the ray in Table 4d.

Table 4d. Ray Solution velocities 6,8,12,12,8,6

node	"x"	"y"	"deviation from plane"
0	0.000000	0.000000	0.000000
1	-0.074337	-0.282214	0.000000
2	-0.173585	-0.659000	0.000000
3	-0.467114	-1.773354	0.179059
4	1.095777	0.617026	0.000000
5	1.614200	1.409934	0.000000
6	2.000000	2.000000	0.000000

For the ray of Table 4d, we show the calculated amplitude information. The densities and velocities used are shown in figure 1.14. We assume a point compressional source with unit energy per unit surface area, and the

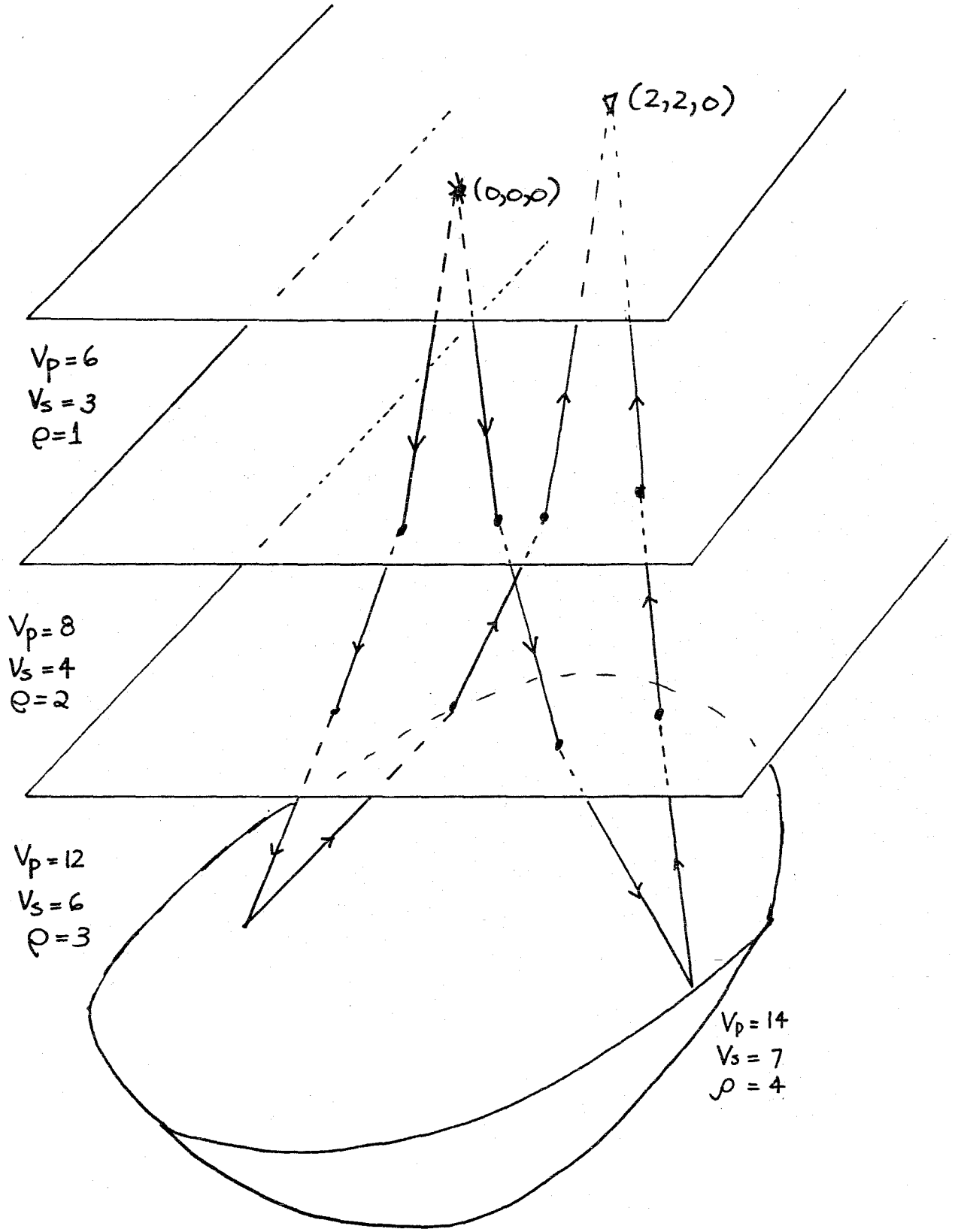


Figure 1.14 Schematic Diagram of Geometry for Example 3.

calculated travel time and amplitude is as follows:

travel time: 4.62

displacement: ex: 1.68E-4

ey: 2.57E-4

ez: 2.18E-3

caustics located at: x=.0025;y=-1.055;z=-16.87

x=.742;y=.0763;z=-12.22

We note that the two caustics are located along the reflected ray in the bottom layer. This is expected because of the focussing properties of a parabolic reflector.

Example 4.

In the following example, we will use velocity and receiver continuation to generate 32 rays at 2 receivers. We consider the following interfaces:

$$\begin{aligned} z1&: x/10\ \sin(y/4) - 5 \\ z2&: x/8 - y/16 - 10 \\ z3&: xy/40 - 20 \end{aligned} \tag{3.6}$$

with the ray sequence as shown below in figure 1.15.

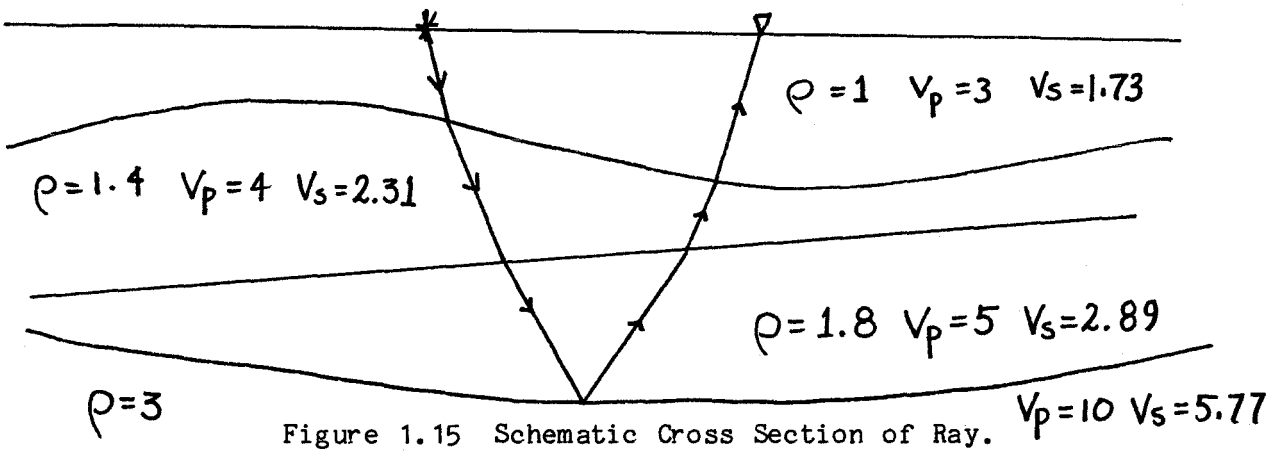


Figure 1.15 Schematic Cross Section of Ray.

We take the source at (x=0,y=0,z=0) and calculate the 32 velocity permutation (we take the first ray segment as always compressional) rays for receivers at (x=1,y=1,z=0),and (x=2,y=2,z=0). Amplitude of the L2 norm of the displacement vector (log base 10) is plotted as a function of travel time for these two stations in figures 1.15a-b. Using our continuation methods for velocity continuation and receiver continuation, both formulation (1.18) required on average two Newton iterations per ray to bring the residual below the designated 1E-5 level. Thus, the generation of large numbers of seismic rays can be made very efficient.

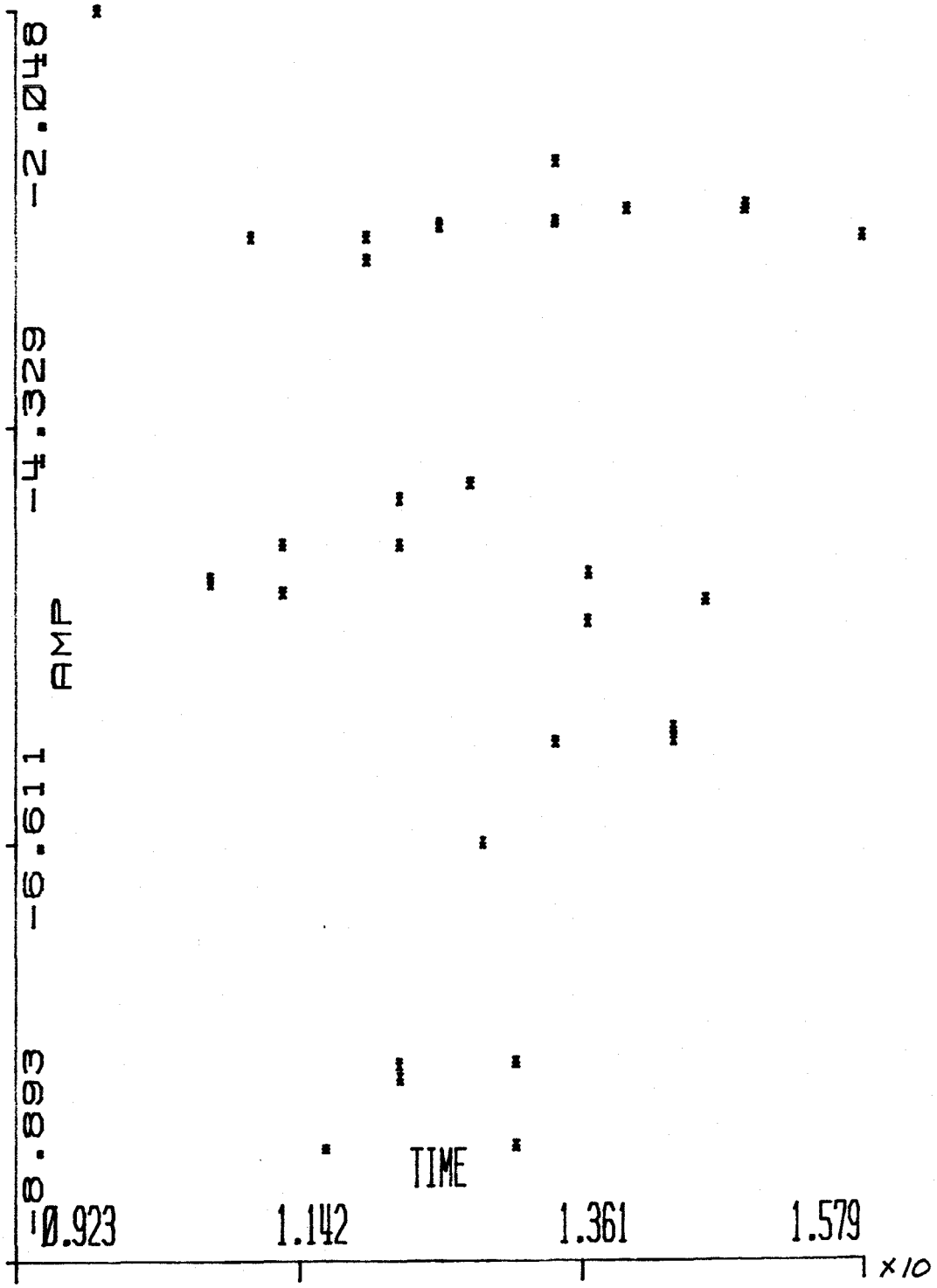


Figure 1.16a Amplitude vs. Arrival Time--Station 1.

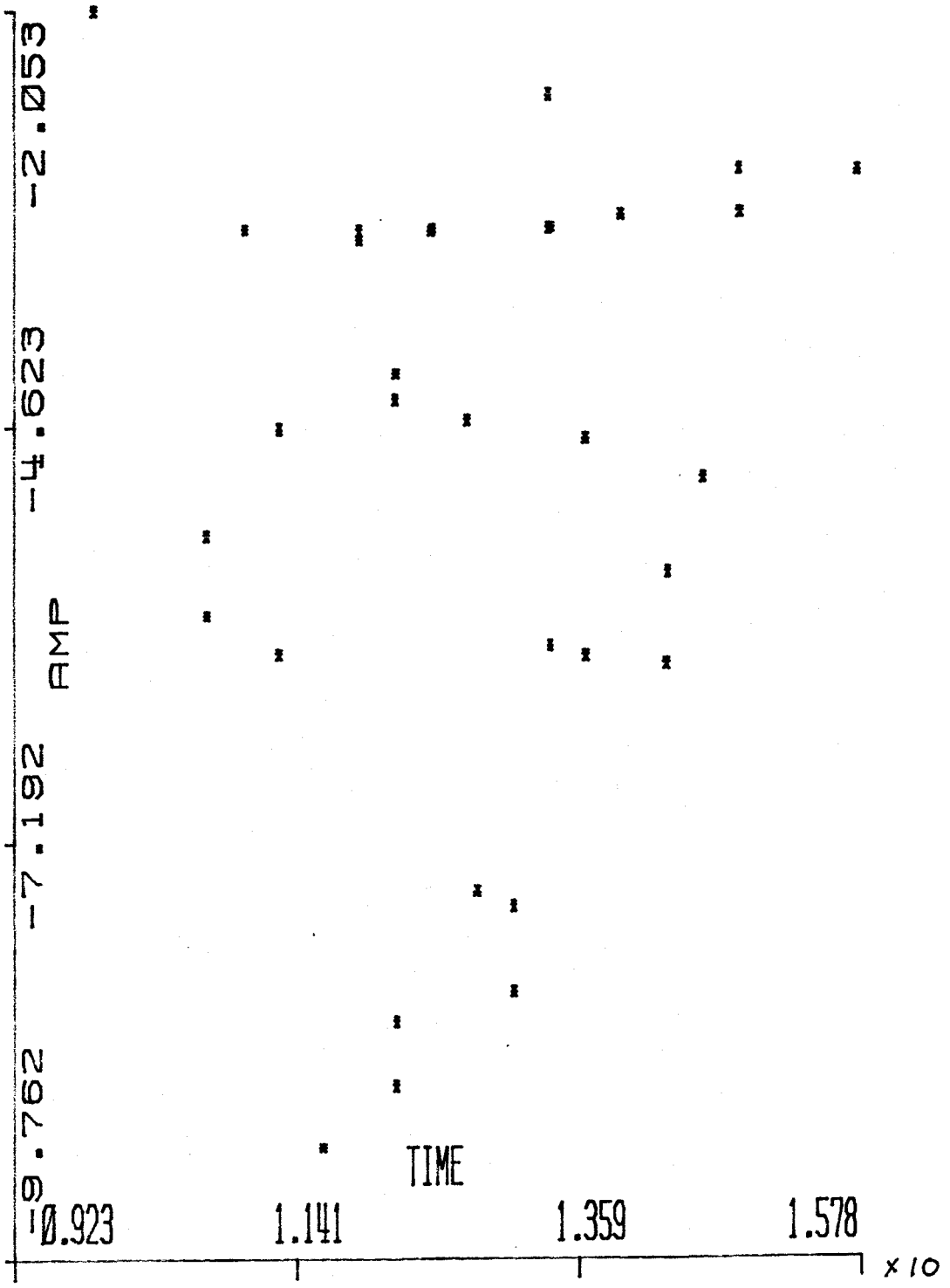


Figure 1.16b Station 2
Amplitude vs. Travel time

Chapter 2. Non-Linear Least Squares Estimation of Elastic and Interface Parameters From Observed Ray Data

Introduction

In Chapter 1 we dealt with a forward problem in seismology. Given all the necessary elastic and geometric parameters of a model of a section of the earth, we showed how to calculate the ray theoretic seismograms for various source/ receiver configurations. In this chapter we use our ray-tracing methods as the nucleus of a non-linear least squares algorithm for the estimation of elastic and interface parameters from the observed seismic ray amplitudes and travel times. Sometimes, there may be some unknown but required information about some portion of the observed data. The example we shall consider is when the ray types of some observations are unknown. To correctly estimate the medium's parameters, it is sometimes necessary to include these observations and we show how to do this in such a way that we simultaneously identify the ray types of the observations and estimate the parameters. In general, the observed ray data will be corrupted with noise, and so we examine the stability properties of the inversion, and show how the Singular Value Decomposition of a matrix can be used to determine the subspace of parameter space, that can be resolved in a stable fashion.

The starting point for all methods of "solution" of inverse problems is to relate the observations as a mathematical function of the unknown parameters:

$$\underline{K} = \mathcal{L}\,\underline{\alpha} \tag{0.1}$$

$\underline{K}$ represents the observations, $\underline{\alpha}$ the unknown parameters (or a quantity from which the parameters can be recovered), and $\mathcal{L}$ is an operator. In this chapter $\underline{K}$ will be a vector of observed travel times and/or amplitudes; $\underline{\alpha}$ will be the unknown layer velocities and/or interface parameters, and $\mathcal{L}$ will be the non- linear operator relating $\underline{K}$ and $\underline{\alpha}$. For our examples, (0.1) will be an overdetermined system of non-linear equations. We will linearize (0.1) and solve a sequence of linear least squares problems. At each iteration of the algorithm we find the pseudo-inverse $\mathcal{L}_n^+$ of the linearized operator $\bar{\mathcal{L}}_n$, and determine the parameter estimate $\bar{\alpha}_n = \mathcal{L}_n^+ \underline{K}_n$. Under certain circumstances $\{\bar{\alpha}_n\}$ will converge as $n \to \infty$ to a $\underline{\alpha}^*$ which we take as a local minimum of $\| \underline{K} - \mathcal{L}\underline{\alpha} \|_2$. This answer is usually non-unique as it is only a local estimate, and for each linear problem, we resolve $\bar{\alpha}_n$ only in the row space of the matrix $\mathcal{L}_n$ (the minimal length solution).

The solution of each linear least-squares problem involves solving the forward ray tracing problem. That is, for each parameter estimation $\bar{\alpha}_n$, we calculate all the rays that correspond to the observations. Thus the entire inversion procedure can involve finding many rays. Hence, the efficiency and accuracy of the ray tracing method used can be very important for this type of parameter estimation scheme.

Section 1 Non-linear Least Squares Inversion

We will rely on an optimization method to estimate the parameter values $\underline{\alpha}^*$ of (0.1). In general, in optimization techniques we often wish to constrain the unknown parameters, $\underline{\alpha}^*$, to lie within physically acceptable bounds. If the constraints are in the form of inequalities, then the problem is in the realm of linear or non-linear programming. If the constraints are in the form of linear relationships between the unknown parameters, then using Lagrange multipliers, we can still cast our problem in the form of a linear or non-linear least squares problem.

In this section, there are implicit constraints on the unknown parameters. For example, we require that velocities be positive, and also the compressional velocity for a layer should exceed the layer's shear velocity. However, we will ignore these constraints and we will find that our estimates satisfy the appropriate physical inequalities.

Theory of Non-Linear Least Squares Estimation (1.1).

In this section we review the portions of the theory of non-linear least squares estimation that are relevant to the understanding of our numerical methods. We consider a function $\Phi(\underline{p})$ which depends non-linearly upon the set of parameters $\underline{p} \equiv (p_1, p_2, \cdots, p_N)^T$. We suppose that $\Phi(\underline{p})$ is of the form

$$\Phi(\underline{p}) = \frac{1}{2} \underline{F}^T(\underline{p}) \underline{F}(\underline{p}) \qquad F_i \quad i = 1, \ldots, M \tag{1.1}$$

We wish to minimize Φ with respect to the parameters $\underline{p}$. If $\underline{p}_0$ is a

parameter vector such that $\nabla\underline{\Phi}(\underline{p}_0)=0$ and $\underline{p}_0 = \underline{p} + \Delta\underline{p}$ ($\Delta\underline{p}$ small), we obtain the approximation:

$$\underline{0} = \nabla\underline{\Phi}(\underline{p}_0) \simeq \nabla\underline{\Phi}(\underline{p}) + \underline{\underline{\nabla^2\Phi}}(\underline{p})\cdot\Delta\underline{p} \qquad (1.2)$$

Thus a Newton type method would be to take:

$$\Delta\underline{p} = (-\underline{\underline{\nabla^2\Phi}}(\underline{p}))^{-1}\nabla\underline{\Phi}(\underline{p}) \qquad (1.3)$$

We call $\underline{\underline{\nabla^2\Phi}}$ the Hessian. Some algebra shows that

$$\underline{\underline{\nabla^2\Phi}} = \underline{F}^T\frac{\partial^2\underline{F}}{\partial\underline{p}^2} + \underline{\underline{J}}^T\underline{\underline{J}} \qquad (1.4)$$

where $\underline{\underline{J}}$ is the Jacobian matrix $J_{ij} \equiv \partial F_i/\partial p_j$. If the residual of the problem (i.e., $\|\underline{F}(\underline{p}_0)\|_2 \simeq 0$) is small, we can make the approximation:

$$\Delta\underline{p}^{(n)} = -(\underline{\underline{J}}^T\underline{\underline{J}})_n^{-1}\underline{\underline{J}}_n^T\underline{F}(\underline{p}^{(n)}) \qquad (1.5)$$

This is called the Gauss-Newton method and is applicable for small residual problems. The method is used iteratively, converging (hopefully) to a point of zero gradient. We write:

$$\underline{p}^{(n+1)} = \underline{p}^{(n)} + \Delta\underline{p}^{(n)} \qquad (1.6)$$

where $\Delta\underline{p}^{(n)}$ is given by (1.5).

For a small residual problem we could also consider a sequence of linear problems. At parameter value $\underline{p}$ we seek $\Delta\underline{p}$ so that:

$$0 \simeq \underline{F}(\underline{p}) + \frac{\partial \underline{F}}{\partial \underline{p}} \Delta \underline{p} \tag{1.7}$$

or the problem at each iteration is to find $\underline{\Delta p}$ from the minimization:

$$\underset{\underline{\Delta p}}{\text{minimize}} \quad \| \underline{\underline{J}}\, \underline{\Delta p} + \underline{F} \|_2 \tag{1.8}$$

This sequence of linear problems is equivalent to (1.5), (1.6) for $\underline{\underline{J}}^T \underline{\underline{J}}$ nonsingular.

Solution of the Linear Problem (1.1a).

As mentioned above, we will often be solving a sequence of linear least squares problems, so we outline here the important aspects of the linear problem. In standard notation, the problem we wish to solve is:

$$\min_{\underline{x}} \quad \| \underline{\underline{A}}\, \underline{x} - \underline{b} \|_2 \tag{1.9}$$

$\underline{\underline{A}}$ is a mxn matrix (m > n)
$\underline{x}$ is a vector of length n
$\underline{b}$ is a vector of length m

It is well known that A possesses the following decomposition (the singular value decomposition) (Wiggins [20]):

$$\underset{mxn}{\underline{\underline{A}}} = \underset{mxk}{\underline{\underline{U}}}\ \underset{kxk}{\underline{\underline{\Lambda}}}\ \underset{kxn}{\underline{\underline{V}}^T} \tag{1.10}$$

The matrix $\underline{\underline{U}}$ contains k eigenvectors $\underline{u}$ of $\underline{\underline{A}}\,\underline{\underline{A}}^T$. These eigenvectors are of length "m" and hence are associated with the

observation space ($\underline{b}$). $\underline{\underline{\Lambda}}$ is a diagonal matrix containing k (k< n) non-zero eigenvalues. The matrix $\underline{V}$ contains k eigenvectors (of length n) of $\underline{\underline{A}}^T\underline{\underline{A}}$. These eigenvectors, because of their length, are associated with the parameter space ($\underline{X}$). These matrices (Wiggins [20]) satisfy the relations:

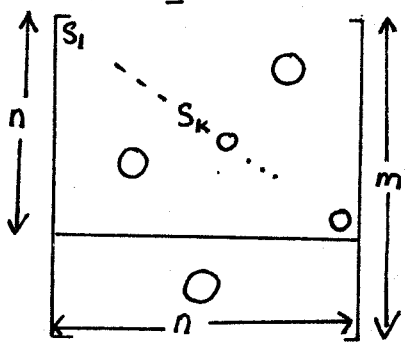
$$\underline{\underline{A}}\,\underline{V} = \underline{U}\,\underline{\underline{\Lambda}} \qquad \underline{\underline{A}}^T\underline{\underline{A}}\,\underline{V} = \underline{V}\,\underline{\underline{\Lambda}}^2$$

$$\underline{\underline{A}}^T\underline{U} = \underline{V}\,\underline{\underline{\Lambda}} \qquad \underline{\underline{A}}\,\underline{\underline{A}}^T\underline{U} = \underline{U}\,\underline{\underline{\Lambda}}^2 \tag{1.11}$$

It is also common (Strang [19]) to write:

$$\underline{\underline{A}} = \underset{(m\times m)}{\underline{U}_E}\ \underset{(m\times n)}{\underline{\underline{\Lambda}}_E}\ \underset{(n\times n)}{\underline{V}_E} \tag{1.12}$$

where $\underline{U}_E$ contains m orthogonal eigenvectors of $\underline{\underline{A}}\underline{\underline{A}}^T$ (the first k columns correspond to the k non-zero singular values) and $\underline{V}_E$ contains n orthogonal eigenvectors of $\underline{\underline{A}}^T\underline{\underline{A}}$. $\underline{\underline{\Lambda}}_E$ now has the structure:



(1.13)

The minimization of $\|\underline{\underline{A}}\underline{x} - \underline{b}\|_2$ can now be written as:

$$\min_{\underline{x}} \left\| \underline{U}_E\,\underline{\underline{\Lambda}}_E\,\underline{V}_E\,\underline{x} - \underline{b} \right\|_2 \tag{1.14}$$

$\underline{U}_E^T$ is unitary, and the transformation defined by $\underline{U}_E^T$ preserves the euclidean norms of a vector. We can write (1.14) after multiplying through by $\underline{U}_E^T$ as:

$$\min_{\underline{x}} \| \underline{\underline{\Lambda}}_E \underline{\underline{V}}_E^T \underline{x} - \underline{\underline{U}}_E \underline{b} \|_2 \tag{1.15}$$

Defining $\underline{\underline{V}}_E^T \underline{x} = \underline{y}$ we have

$$\min_{\underline{y}} \| \underline{\underline{\Lambda}}_E \underline{y} - \underline{\underline{U}}_E \underline{b} \|_2 \tag{1.16}$$

Since $\underline{\underline{\Lambda}}_E$ has the structure shown in (1.13), we find that the value of y_j, $j>k$ has no effect on (1.16), so it can be arbitrary. Taking these components to be zero, we obtain:

(a) $\underline{y} = \underline{\underline{\Lambda}}_E^{-1} \underline{\underline{U}}_E^T \underline{b}$

or

(b) $\underline{x} = \underline{\underline{V}}_E \underline{\underline{\Lambda}}_E^{-1} \underline{\underline{U}}_E^T \underline{b}$

or

(c) $\underline{x} = \underline{\underline{V}} \underline{\underline{\Lambda}}^{-1} \underline{\underline{U}}^T \underline{b}$

$$\underline{\underline{\Lambda}}_E^{-1} \equiv \begin{bmatrix} 1/s_1 & & & \\ & \ddots & & \\ & & 1/s_k & \\ & & & 0 \\ & & & \ddots \\ & & & 0 \end{bmatrix} \quad \begin{matrix} \leftarrow m \rightarrow \\ \\ \\ \\ \\ \\ \end{matrix} \quad \Big\} n \tag{1.17}$$

If k=n then we find that the pseudo-inverse which we define as:

$$\underline{\underline{A}}^+ = \underline{\underline{V}}_E \underline{\underline{\Lambda}}_E^{-1} \underline{\underline{U}}_E^T \tag{1.18a}$$

is also equal to

$$\underline{\underline{A}}^+ = \left(\underline{\underline{A}}^T \underline{\underline{A}} \right)^{-1} \underline{\underline{A}}^T \tag{1.18b}$$

The geometrical interpretation of the pseudo-inverse is simple. The data vector $\underline{b}$ is projected into the column space of $\underline{\underline{A}}$, to a vector $\underline{p}$. Then, the solution $\underline{x}$ is found in the row space of $\underline{\underline{A}}$ such that $\underline{A}\underline{x} = \underline{p}$. Since we take all components of $\underline{x}$ in the null space of $\underline{\underline{A}}$ to be zero, we have found the "minimum length" solution.

When the rank of $\underline{\underline{A}}$ is equal to n, then (1.18b) is equivalent to the "Normal Equations":

$$\underline{\underline{A}}^T \underline{\underline{A}}\, \underline{x} = \underline{\underline{A}}^T \underline{b} \tag{1.19}$$

This system can frequently be ill-conditioned and it is often advisable to work directly with (1.9). This is because the singular values of $\underline{\underline{A}}^T \underline{\underline{A}}$ are the square of those of $\underline{\underline{A}}$, and hence the condition number of (1.19) is the square of that for (1.9). As mentioned above, multiplying (1.9) through by unitary transformations, does not affect the problem.. Golub [6] showed that by multiplying (1.9) by a sequence of Householder matrices $\{\underline{\underline{P}}_j\}_{j=1}^{n}$, one could triangularize $\underline{\underline{A}}$ to the form:

$$\underline{\underline{R}} = \underline{\underline{P}}_n \underline{\underline{P}}_{n-1} \cdots \underline{\underline{P}}_1 \underline{\underline{A}} = \begin{bmatrix} T \\ O \end{bmatrix} \begin{matrix} \uparrow n \\ \downarrow m \end{matrix} \quad \left(\longleftarrow n \longrightarrow \right) \tag{1.20}$$

The new problem

$$\min_{\underline{x}} \left\| \underline{\underline{R}}\, \underline{x} - (\underline{\underline{P}}_n \underline{\underline{P}}_{n-1} \cdots \underline{\underline{P}}_1)\, \underline{b} \right\|_2 \tag{1.21}$$

is now trivially solved. If the matrix $\underline{\underline{A}}$ is very large and storage is a concern, $\underline{\underline{A}}$ can be partitioned and a sequence of triangularizations used. This is discussed in Lawson and Hanson [12].

Finally, we note that it is sometimes of interest to calculate the singular value decomposition of a matrix. The estimate, $\underline{x}$, is determined

from (1.17c) . This method is also numerically better conditioned than forming the normal equations (1.19), for the same reasons discussed above for the QR algorithm. Also, much valuable information can be obtained from this decomposition. We used the singular value decomposition package of IMSLIB (LSVDF) which uses the algorithm of Lawson and Hanson [12]. This algorithm in turn is based on the work of Golub and Reinsch [7]. As discussed above, we can associate null-vectors Vj (length n) of the parameter space with the singular values of $\underline{\underline{A}}^T\underline{\underline{A}}$. For eigenvectors vj associated with very small singular values, the variance of the parameter estimate in the space $\tilde{G}$ spanned by these vectors is large. In other words, a small change in the data can lead to very large changes in the parameter estimates in $\tilde{G}$. The extreme case is for a zero singular value, where the parameter estimate in the direction of the null-vector can be completely arbitrary. With a singular value decomposition we avoid the problem of a rank deficient matrix by working only in the row space of the matrix. However, even for very small non-zero singular values, it is often best to consider these values to be zero, and the associated space $\tilde{G}$ to be a subspace of the null space. There are two main reasons to do this. For realisitic data with noise(or for synthetic data, numerical truncation and round off) the parameter estimates in the space $\tilde{G}$ may be totally unreliable and wild. In our iterative method, wild parameter estimates can lead to situations where ray solutions fail to exist, and our method fails. Secondly, the numerical algorithm can only find the singular values of the matrix within some tolerance, so that very small singular values are somewhat inaccurate and even theoretically zero singular values may be very small non-zero numbers. As will be discussed later, our Jacobian matrix $\underline{\underline{J}}$

is a numerical approximation to the "true" Jacobian, as our ray solutions are found numerically (residual<.0001) and for amplitude inversions we approximate derivatives with finite differences. The singular values are found to within a tolerance of order $||A||_2 \mathcal{E}$, where $||A||_2$ is the maximum singular value and $\mathcal{E}$ is the machine precision (here, about .000005). In general, in our calculations the norm of the Jacobian matrix is O(1) and we usually set the pseudo-rank tolerance at .0001 (i.e., singular values less than this are set to zero, and the dimension of the problem reduced). In most cases this choice seemed to work well.

Modifications for the Non-linear Problem (1.1b).

As discussed above, the non-linear least squares problem can be solved by a Newton iteration scheme (1.3), if the initial parameter estimate is sufficiently close to the local minimum. For the case where the problem has a small residual at this minimum, we can further simplify the method to (1.5). However, when we do not start close to the minimum it is very likely that Newton's method will not converge. A method which has been devised to circumvent the problems of Newton's iterations is the Levenberg-Maquardt algorithm, which we used and we now describe.

If we consider an iteration method, where at each step we take $\Delta \underline{p}$ as a small step in the $-\nabla \underline{\Phi}^T$ (using the notation of 1.1) direction so that $\|\underline{F}(\underline{p})\|_2$ decreases, we have the gradient method. This method will, in general, decrease the residual of the problem, but often the decrease will be slow. Combining the gradient and Gauss-Newton methods, we solve the system:

$$\left(\underline{J}_n^T \underline{J}_n + \lambda_n \underline{I}\right)\underline{\Delta p} = \underline{J}_n^T \underline{F}(\underline{p}) \tag{1.22}$$

or this is equivalent to the minimization of

$$\left\| \begin{bmatrix} \underline{J}_n \\ \lambda_n \underline{I} \end{bmatrix} \underline{\Delta p} + \begin{bmatrix} \underline{F} \\ \underline{0} \end{bmatrix} \right\|_2 \tag{1.23}$$

Equations (1.22) or (1.23) are solved at each step of the L-M iterative scheme. However, some selection criterion for λ_n is needed. As $\lambda_n \to \infty$, $\| \Delta \underline{p} \|_2 \to 0$ and $\Delta \underline{p}$ rotates toward the negative gradient vector $-\nabla \Phi(\underline{p})$. Thus we are guaranteed that if we choose λ_n large enough, we will decrease the residual of the system. However, the stepsize may be exceedingly small. The selection criterion for λ_n is somewhat adhoc, and we found in our numerical experiments that the best criterion changes from problem to problem. In most circumstances, for zero or small residual problems, it was found that if the Gauss-Newton method converged, it would converge much more rapidly than the L-M method with $\lambda_n \neq 0$. Thus the procedure we used was to start with $\lambda_n = 0$. Then after the Newton step, if the residual of the problem had increased, we recalculated the first step with

$$\lambda_1 \longleftarrow 2\lambda_1 + 1 \tag{1.24}$$

This doubling procedure was continued until a change in the parameters was found that decreased the residual. For the new parameter estimate we did one of two things:

(i) set $\lambda_2 = \lambda_1 / 2$

or (1.25)

(ii) set $\lambda_2 = 0$

For the case where our initial estimate was close enough to the minimum so that only one or two gradient steps were required before the Newton method would converge with decreasing residuals, then (ii) was the best choice. When the initial guess was "far" from the local minimum, then for the first several iterations (i) was the better criterion. Starting with a value for λ_2, we repeat the procedure (1.24). In general, if we use the procedure as outlined above, then at the n'th iteration we may have to solve (1.22) for a few (and sometimes several) values of λ_n before we obtain a decrease in the value of $\|\underline{F}(p)\|_2$. However as we show now, most of the computing is for the solution of (1.22) with the first value of λ_n. Successive inversions of (1.22) for other values of λ_n can be done "cheaply."

From the singular value decomposition (1.12) of $\underline{J}$ we can write

$$\underline{J}^T\underline{J} = \underline{V}_E \underline{\Lambda}_E^T \underline{\Lambda}_E \underline{V}_E^T \qquad \underline{\Lambda}_E^2 \equiv \underline{\Lambda}_E^T \underline{\Lambda}_E$$

$$(\underline{J}^T\underline{J} + \lambda_n \underline{I}) = \underline{V}_E (\underline{\Lambda}_E^2 + \lambda_n \underline{I}) \underline{V}_E^T$$

$$(\underline{J}^T\underline{J} + \lambda_n \underline{I})^{-1} = \underline{V}_E (\underline{\Lambda}_E^2 + \lambda_n \underline{I})^{-1} \underline{V}_E^T \qquad (1.26)$$

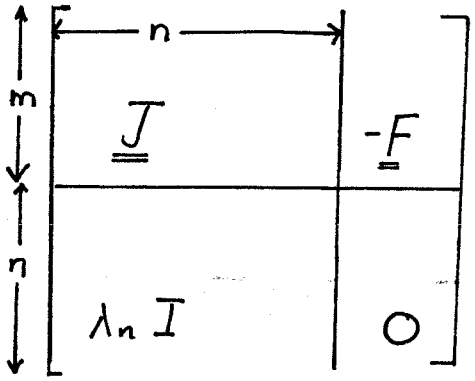
Thus for $\lambda_n \neq 0$ the pseudo-inverse becomes

$$(\underline{J}^T\underline{J} + \lambda_n \underline{I})^{-1} \underline{J}^T = \underline{V}_E (\underline{\Lambda}_E^2 + \lambda_n \underline{I})^{-2} \underline{\Lambda}_E^T \underline{U}_E^T \qquad (1.27)$$

This differs from the pseudo-inverse (1.18) only in the elements of the diagonal matrix $\underline{\Lambda}_E$ .. Thus once we have computed the singular-value

decomposition of the Jacobian matrix at the n'th iteration, we simply replace the singular values $\{s_i\}$ by $(s_i^2 + \lambda_n)/s_i$, for the solution of (1.22). This is the method we use in our calculations.

However, in many practical situations one might be using the Householder triangularization method of Golub (1.20). We can write the coefficent matrix of the minimization problem (1.23) in partitioned form:



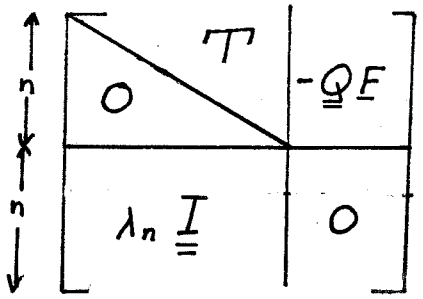
(1.28)

Then instead of calculating the QR decomposition of the full m + n by n matrix for each value of λ_n , we can first calculate:

$$Q\,J = R \qquad R = \begin{bmatrix} T \\ O \end{bmatrix} \begin{matrix} n \\ m \end{matrix}$$

(1.29)

This decomposition is then used for all values of λ_n , and we consider instead the problem with coefficent matrix:



(1.30)

For m>>n and for several values of λ_n , formulation (1.30) can be very economical. This method is discussed in Osborne [14].

economical. This method is discussed in Osborne [14].

Finally we mention that for large residual problems (i.e., $\|\underline{F}(\underline{p})\|_2 \gg 0$, where $\|\nabla\Phi(\underline{p})\|_2 = 0$), we may have to use the true Hessian for the problem (1.2). This greatly changes the numerical approaches outlined above. A survey article on large residual methods can be found in Nazareth [13]. In the numerical examples presented below, we will rely on the Gauss-Newton/ Levenberg-Maquardt algorithm. Sufficient conditions for the convergence of the Levenberg-Maquardt algorithm and the pseudo-inverse Newton's method are given by Osborne [14] and Ben-Israel [1]. Perozzi [15] implemented a simple Gauss-Newton algorithm in a ray data inversion example for a two dimensional geometry. Many of the formulae that we give below for the least squares Jacobian's elements can also be found in Perozzi [15].

Section 2 The Seismological Problem.

Theory of Travel Time Inversion (2.1).

We consider the layered media of chapter 1, and use the ray notation of that part of the thesis. Thus we can write for the travel time of a disturbance between the source $\underline{X}_S$ and receiver $\underline{X}_R$:

$$tr(\underline{X}_S, \underline{X}_R) = \sum_{j=1}^{N+1} \frac{D_j}{V_j} \qquad (2.1)$$

Here: tr denotes travel time, N is the number of interface intersections , Dj is the length of the j'th ray segment, and Vj is the velocity on the j'th ray segment. Equivalently, we have:

$$tr(\underline{X}_S, \underline{X}_R) = \sum_{j=1}^{N+1} \frac{\sqrt{(x_j - x_{j-1})^2 + (y_j - y_{j-1})^2 + (z_j - z_{j-1})^2}}{V_j} \qquad (2.2)$$

We suppose that one knows several travel times for known source-receiver configurations ($\underline{X}_S$, $\underline{X}_R$). Also, for the time being, we assume that the ray types, e.g., (PSP.., PPS..., etc.), are known. Here, we use the notation that "P" denotes a compressional ray segment, and "S" a shear ray segment. Later, we will examine the problem of determining ray type.

We denote the vector of known travel times as $\underline{obs}$, and in the notation of (1.1), tr_i is defined as the calculated travel time for the i'th ray. We set:

$$F_i \equiv tr_i(\underline{X}_{S,i}, \underline{X}_{R,i}) - obs_i \qquad (2.3)$$

The Jacobian $J_{ij} \equiv \partial F_i / \partial p_j$, where $\underline{p}$ is the vector of unknown parameters $(P_1, \ldots, P_n)^T$, and thus:

$$J_{ij} = \frac{\partial tr_i}{\partial p_j} + \frac{\partial tr_i}{\partial \underline{X}} \frac{\partial \underline{X}}{\partial p_j} \qquad (2.4)$$

where $\underline{X}$ denotes the vector of the coordinates of the ray's intersections with the interfaces. However, as seen in chapter 1, from Snell's law, which the rays satisfy at each interface, $\frac{\partial tr_i}{\partial \underline{X}} = 0$. Thus (2.4) is simply

$$J_{ij} = \frac{\partial tr_i}{\partial p_j} \qquad (2.5)$$

For the specific examples which we examine, we will write down the formula (2.5) for these cases. If one wishes to calculate the true Hessian (1.4) for the problem, we must also calculate the term

$$\frac{d^2 tr_i}{dp_k dp_j} = \frac{\partial^2 tr_i}{\partial p_k \partial p_j} + \frac{\partial}{\partial \underline{X}} \left(\frac{\partial tr_i}{\partial p_j} \right) \left(\frac{\partial \underline{X}}{\partial p_k} \right) \qquad (2.6)$$

Thus for this term we must calculate the variation of the ray solution with respect to the parameters. However, as shown in chapter 1, this can be done analytically using the Jacobian of the ray solution. Usually, however, we will need only (2.5).

Inclusion of Amplitude Information (2.2).

The ray-tracing method of chapter 1 also calculates the amplitude of the disturbance (assuming unit energy at unit distance from source). As

discussed in that chapter, the amplitude is a complicated function of the elastic parameters and the interface geometry. Hence we do not attempt to analytically calculate the Jacobian, but instead we use a finite difference approximation. We will use amp_j to denote the amplitude of the j'th ray. In our numerical examples we will take the amplitude to be the modulus of the euclidean norm of the displacement vector, but, in general, there are various definitions to use for the amplitude. We will write the derivatives as:

$$\frac{\partial amp_i}{\partial p_j} = \frac{(amp_i(\underline{p} + \underline{e}_j) - amp_i(\underline{p}))}{\Delta} \qquad (2.7)$$

where $e_j^T = (0,0,0,\ldots,\Delta_j,0,\ldots,0)$. Thus to approximate the derivatives, we do a series of n ray traces (n is the number of unknown parameters), to calculate the perturbed amplitudes. These ray traces are very fast, as the ray solution for parameters $\underline{p}$, provides an extremely good initial guess for the problem for $\underline{p} + \underline{e}_j$. We make a suitable choice for Δ in terms of the order of magnitude of the parameters, and the number of significant figures on the computer.

In general, if we have M rays, and an amplitude and a travel time for each ray, we can write our observation vector as:

$$\begin{aligned} obs &= \text{observed traveltime}_i \qquad i = 1,M \\ obs &= \text{observed amplitude}_i \qquad i = 1,M \end{aligned} \qquad (2.8)$$

and

$$\begin{aligned} F &= A\,(tr_i - obs_i) \qquad i = 1,M \\ F &= B\,(amp_i - obs_i) \qquad i = 1,M \end{aligned} \qquad (2.9)$$

Equation (2.9) allows us to weigh differently the travel time and

amplitude observations (A and B).

General Numerical Implementation (2.3).

For the numerical examples that follow, in order to examine our inversion schemes, we generated the "observations" using a known model and the ray-tracing techniques of chapter 1 (formulation 1.17). A random number generator was used to simulate noise in the data. The computer generated numbers in the interval [0,1]. Random noise was generated by the formula:

$$\text{noise}_i = \alpha_k/100 \times \text{ave},_k \times 2 \times (\text{ran}_i - .5) \qquad (2.10)$$

$k = 1$	refers to travel times
$= 2$	refers to amplitudes
α,k	is input percentage
ave,k	is average of observations
ran,i	is i'th random number $\in$ [0,1]
noise,i	is number we add to obs

Here, the random numbers are generated by the machine in the interval [0,1]. By subtracting off .5 and multiplying by 2, we change the interval to [-1,1]. Then algorithmically we proceed as shown below in Table 1.

Table 1. Least-Squares and Ray Tracing Algorithm.

i) input observations (perhaps with noise) from data file
ii) input initial parameter guess
    input desired gradient smallness; eps
    $\sigma = 0$.
iii) do for $i = 1$, total number of rays (M)
    a) trace ray$_i$ with current parameter estimates p
    b) calculate travel time

if using amplitudes - calculate amplitude
c) calculate residual F_i , (and F_{i+M})
d) calculate row i in least squares Jacobian analytically using (2.5)
e) if amplitudes included then calculate row i + M in Jacobian using finite difference approx.

iv) calculate norm of gradient
if $\|\text{gradient}\|_2 <$ eps print out parameter estimates - stop

v) calculate residual $\underline{F}^T \underline{F}$

vi) if residual has increased from previous iteration then
$\sigma \leftarrow 2\sigma + 1$; itry = 2
else
$\sigma \leftarrow \sigma/2$ or $\sigma \leftarrow 0$; itry = 1

vii) solve problem $(\underline{J}^T \underline{J} + \sigma \underline{I})\ \underline{p} = -\underline{J}^T \underline{F}$ with singular value decomposition with current value of σ
a) if itry = 2 use previous decomposition of Jacobian with previous $\underline{F}$ and new σ
b) if itry = 1; do new composition with new σ

viii) if singular values S_i fall below S_{MIN} set S_i equal to zero

ix) form new values of parameters

x) go to iii) if less than maximum number of allowed iterations- else stop

The nucleus of this algorithm is the ray-tracing. If we include amplitudes in the problem, and if there are M rays and N parameters, then we must do (N + 1)M ray-traces (for travel time inversion only M) per iteration. Thus it is important to make the ray-tracing efficient. As discussed in chapter 1, most of the computation involved in our method is for the calculation of the first ray in a series of rays. The other rays in the series can be calculated very quickly by using successive continuations. We calculated the rays in inversion examples using both ray tracing formulations discussed in chapter 1. These two different formulations made no difference to the inversion results.

To illustrate how the ray-tracing can effectively be included in the inversion algorithm, we suppose that we have two arrays on the earth's

surface, with three receivers each. Further, we assume in this simple example that there is a primary reflector and we have recorded the information for the PP and PS rays at each receiver. Then schematically our ray-tracing algorithm could be as shown in figure 2.1.

Thus for each array, the "most" work is often put in at the first iteration, for the first station's purely compressional ray Then continuation methods quickly generate all the other rays for the array. For our methods, a continuation method can mean using the previous solution as the initial guess for the new velocity sequence's or new receiver's ray solution(etc.), or we can use a one derivative Euler correction:

$$\underline{X}(\lambda) = \underline{X}(o) + \frac{d\underline{X}}{d\lambda}\,\Delta\lambda \qquad (2.11)$$

The method of calculating $\frac{d\underline{x}}{d\lambda}$ is discussed in chapter 1. From one iteration to the next in the non-linear least squares method (especially near the minimum) the parameter change is often small, so we use the saved rays from the previous iteration as initial ray estimates for the first compressional ray at the new parameter values.

In general, one might have to use more than one continuation step in the above techniques, but in our computations one step was sufficient. We did not use the first derivative corrections for the various continuations for these problems, but these can be simply and cheaply included if desired. Finally, we mention that, in general, there exists more than one ray solution for a specific ray type. Thus we want the ray solutions calculated by the non-linear least squares program to lie on the same solution branch as the observed data. This could perhaps cause problems, but in our computations this problem did not often arise.

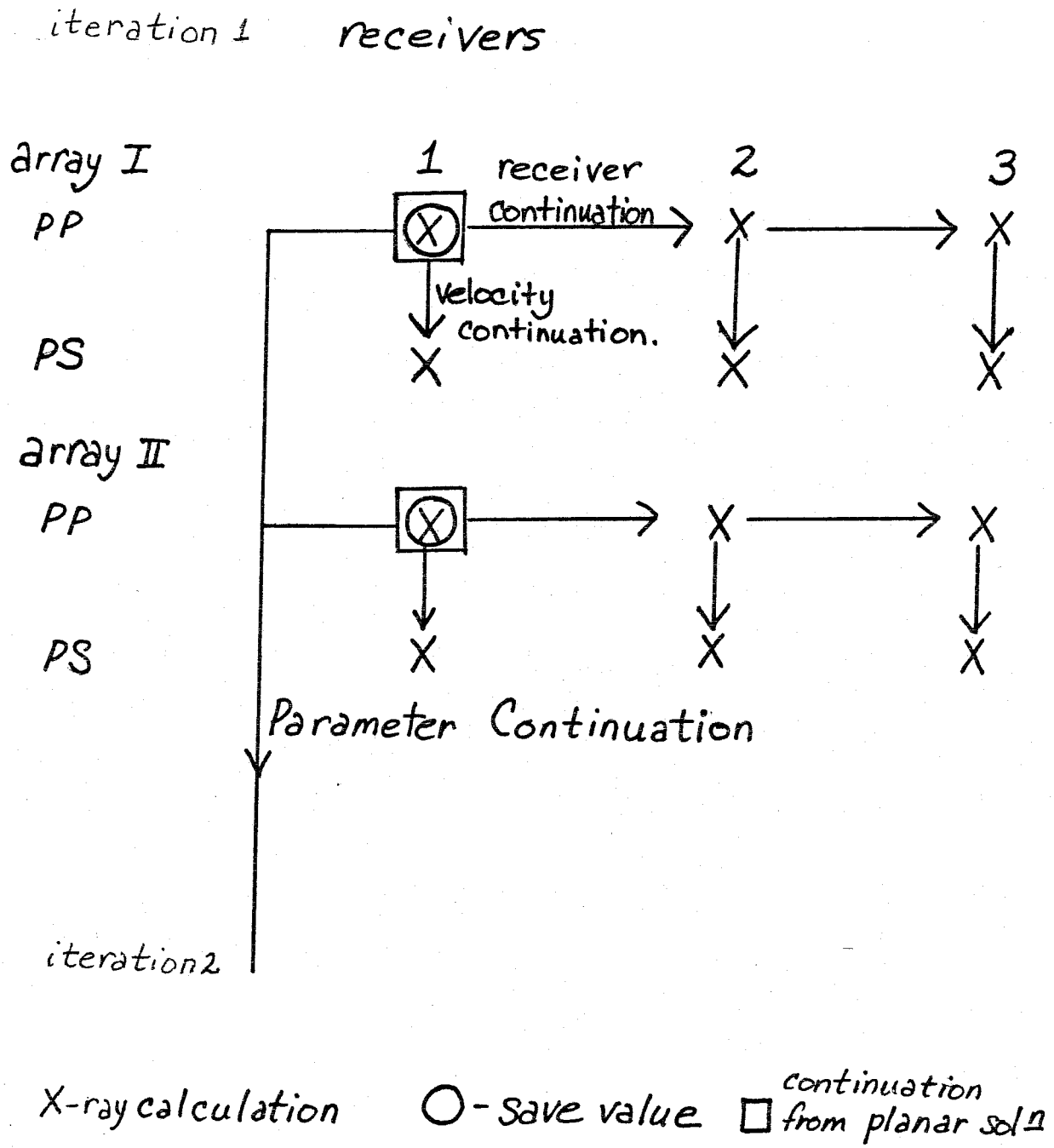


Figure 2.1 Schematic of Continuations That can be Used in Least--Squares Inversion.

Section 3. Inversion Examples.

Inversion of Travel Times for Layer Velocity Estimates (3.1).

For the estimation of layer velocities, the analytic expression for the elements of the least-square's jacobian (2.5) is:

$$J_{iK} = \frac{\partial tr_i}{\partial V_K} = -\sum_{j=1}^{N+1} \frac{D_j}{V_j^2}\, \delta_{j,K} \qquad (3.1)$$

($\delta_{j,K}$ =1 if v corresponds to the same layer and velocity type (P or S) as v_K ; $\delta_{j,K}$ =0 otherwise.)

For example, for the geometry shown in figure 2.2, with v =v2,s we have: $\delta_{1,K}$ =0, $\delta_{2,K}$ =1, $\delta_{3,K}$ =0, $\delta_{4,K}$ =0, $\delta_{5,K}$ =1, and $\delta_{6,K}$ =0.

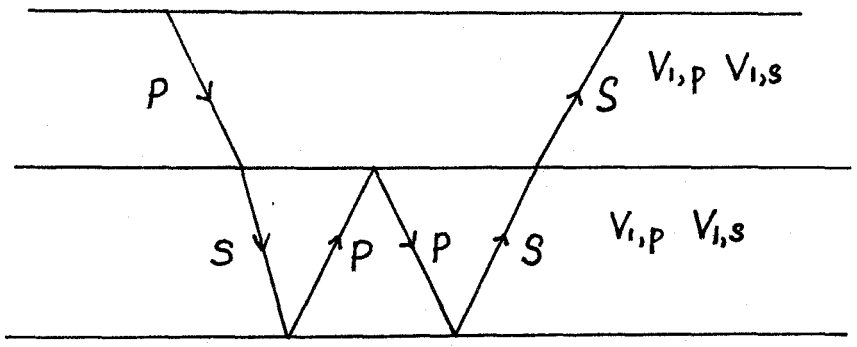


Figure 2.2 Model Geometry for Example

As a first example, we consider the planar interface structure shown below in figure 2.3.

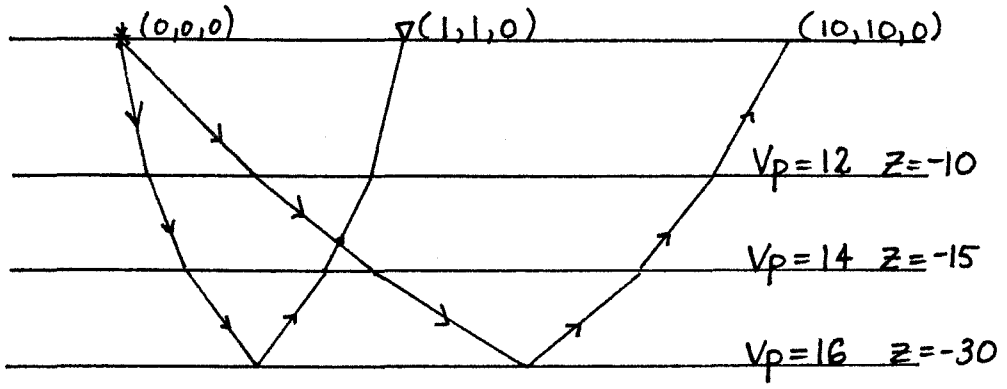


Figure 2.3 Model Geometry For First Inversion Example

We have used 10 stations and the purely compressional ray at each, as shown above. This problem is numerically singular, with a minimum singular value of O(1E-6). Thus even when we start close to our known answer of (12,14,16), we do not neccessarily converge to this estimate,as our method deletes one parameter direction from the problem at each step. Starting with an initial velocity estimate of (11.00,13.00,17.00) (residual .304, gradient .197), we converge in 3 iterations to (12.21,13.20, 16.13) (residual 8.23E-6,gradient 4.83E-6). It is not obvious why the Jacobian is nearly singular but, as we show, this is because the angles of incidence of the rays for this geometry (even for the receiver at x=10,y=10) are relatively small. The three columns of the Jacobian matrix are from (3.1):

$$J_{i,1} = -\frac{D_{1,i}}{V_1^2} - \frac{D_{6,i}}{V_6^2} = -\frac{2D_{1,i}}{V_1^2}$$

$$J_{i,2} = -\frac{D_{2,i}}{V_2^2} - \frac{D_{5,i}}{V_5^2} = -\frac{2D_{2,i}}{V_2^2} \qquad (3.2)$$

$$J_{i,3} = -\frac{D_{3,i}}{V_3^2} - \frac{D_{4,i}}{V_4^2} = -\frac{2D_{4,i}}{V_3^2}$$

Now for small offsets we have that

$$D_{K,i} = \sqrt{c_k^2 + h_{K,i}^2} \triangleq c_k + \frac{1}{2}\frac{h_{K,i}^2}{c_k} \qquad (3.3)$$

where C_K is the thickness of the k'th layer and $h_{K,i}$ is the horizontal distance travelled by the i'th ray in the k'th layer (here $h_{K,i} = \sqrt{2}\,\Delta X_{K,i}$).
The ray solution satisfies the scalar Snell laws:

$$V_1 \sin\theta_2 = V_2 \sin\theta_1$$
$$V_1 \sin\theta_3 = V_3 \sin\theta_1 \qquad (3.4)$$

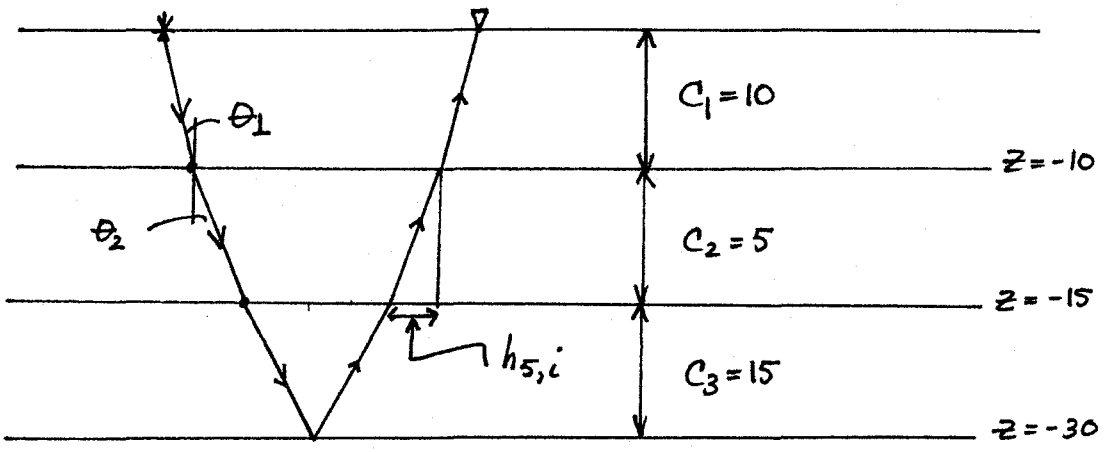


Figure 2.4 Notation Used For (3.4)

For small θ, we make the approximation:

$$V_1 \tan\theta_2 = V_2 \tan\theta_1$$
$$V_1 \frac{h_{i,2}}{5} = \frac{V_2\, h_{i,1}}{10}\,; \quad h_{i,2} = \frac{1}{2}\frac{V_2}{V_1} h_{i,1} \qquad (3.5)$$

Similarly,

$$h_{i,3} = \frac{15}{10}\frac{V_3}{V_1}\, h_{i,1}$$

Using these approximations, we can write the columns (3.2) as:

$$\begin{aligned} J_{i,1} &= \frac{-2}{V_1^2}\left\{10 + \frac{h_{i,1}^2}{20}\right\} \\ J_{i,2} &= \frac{-2}{V_2^2}\left\{5 + \left(\frac{V_2}{V_1}\right)^2 \frac{h_{i,1}^2}{40}\right\} \\ J_{i,3} &= \frac{-2}{V_3^2}\left\{15 + \left(\frac{V_3}{V_1}\right)^2\left(\frac{15}{10}\right)^2 \frac{h_{i,1}^2}{30}\right\} \end{aligned} \qquad (3.6)$$

For the case V1=11,V2=13, and V3=17 we find (a,b,c) such that

$$a\,\underline{J}_{i,1} + b\,\underline{J}_{i,2} + c\,\underline{J}_{i,3} = 0 \qquad (3.7)$$

Normalized, we find that this triplet is (.245,-.957,.157). The calculated eigenvector (associated with s=2.46E-6) from the singular value decomposition is:

$$(.247,-.957,.152) \qquad (3.8)$$

Thus the near linear dependance of the columns arises mathematically from the fact that the small angle approximations are valid here. Physically this means that velocity perturbations in the direction (3.8) will effectively have no effect on the travel times. If we consider the model of figure 2.4 with the interface depths at z=-2,z=-4, and z=-6, we no

longer have a numerically singular matrix.(the angles of incidence are no longer small) With an initial velocity estimate of (14.00,18.00,22.00) (residual =.716, gradient= .067) we arrive at the estimate (11.99,13.99,16.02) in three iterations.

Returning to our original model, we now include a second ray in the inversion process (figure 2.5) This problem (at least for estimates we considered) was non-singular, and for "sufficiently close" guesses (e.g., V= (13.00,16.00,18.00)) we converged rapidly to the zero V=(12.00,14.00,16.00). However, when we start farther off we do not neccessarily converge to this zero. The results for this two ray inversion are shown below in Table 2a.

Table 2a. Two Ray Family Inversion

it	velocities	residual	gradient
0	(20.0,11.0,6.0)	20.16	119.27
1/2	(26.47,1.32,9.75)	34.80	
1*	(20.07,11.11,9.3)	7.53	18.65
2	(20.2,11.34,12.09)	2.26	3.20
3	(20.19,11.32,13.4)	0.84	.495
4	(19.9,10.87,13.89)	0.66	.104
5	(19.38,10.02,14.31)	0.51	.085
6	(18.78,8.91,14.94)	0.30	.104
7	(18.4,8.09,15.6)	0.12	.083
8	(18.29,7.8,15.93)	0.025	.015
9	(18.26,7.76,16.)	0.009	.0008
10	(18.25,7.75,16)	0.0088	2.78E-5

*-increase σ; $\sigma = 2\sigma + 1$

Table 2b. Three Ray Family Inversion

it	velocities	residual	gradient
0	(20.0,11.0,6.0)	22.1	
1	(7.45,13.26,9.74)	15.07	47.40
2	(10.28,13.96,13.55)	4.136	6.87
3	(11.76,14,15.62)	.536	.69
4	(11.99,14,15.97)	1.17E-2	1.4E-2
5	(12,14,16)	6.78E-6	7.8E-6

Table 2c. Four Ray Family Inversion

it	velocities	residual	gradient
0	(20.0,11.0,6.0)	22.97	
1	(6.73,13.42,9.76)	20.56	90.95
2	(9.68,13.99,13.56)	5.97	13.22
3	(11.55,14,15.63)	.89	1.43
4	(11.98,14,15.99)	2.71E-2	4.13E-2
5	(12,14,16)	3.10E-5	4.68E-5

For the three ray and four ray inversions, the results of which are shown above in Tables 2b and 2c, we have added the rays between the source and receivers (x=i,y=x : i=1,10) shown below in figure 2.5.

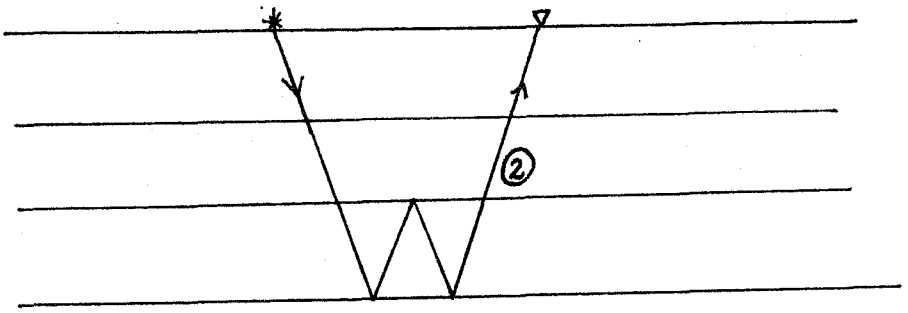


Figure 2.5 Second Ray Added

For the other two inversions, the rays are shown below in figure 2.6.

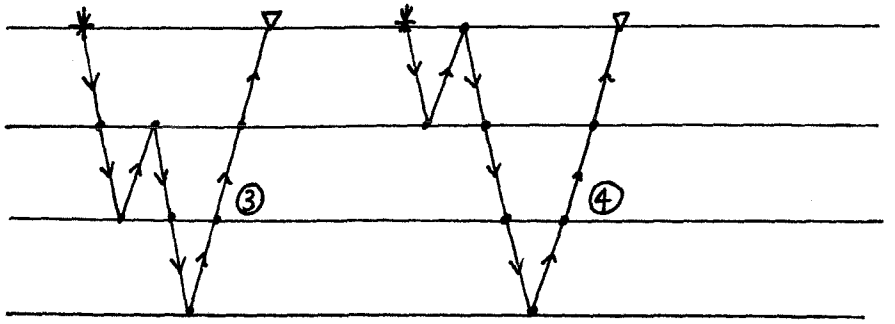


Figure 2.6 Third and Fourth Rays Added

For the two ray inversion, the bottom layer is transversed the most frequently with ray segments, and the velocity estimate for this layer is accurate. The three ray inversion forces the velocity estimate to the zero residual value of (12,14,16). The inclusion of a fourth ray does not seem to greatly improve the inversion process.

We now investigate the effect of noise on the inversion process. Table 3a shows the synthetically generated travel time data for the 4-ray case. Table 3b shows the same data but with a 2% level of random noise added. We also generated data sets with .1% and 5% noise. The inversion results for the noise corrupted data are shown below in Table 4.

Table 3a. Theoretical Travel Time Data

Ray I				
4.257115	4.260601	4.266404	4.274514	4.284917
4.297596	4.312529	4.329691	4.349054	4.370587
Ray II				
6.131699	6.133937	6.137665	6.142880	6.149580
6.157757	6.167406	6.178520	6.191091	6.205108
Ray III				
4.971238	4.974237	4.979229	4.986211	4.995172
5.006102	5.018986	5.033809	5.050553	5.069198
Ray IV				
5.923528	5.926254	5.930795	5.937145	5.945300
5.955250	5.966986	5.980497	5.995770	6.012790

Table 3b. Data With 2% Random Noise Added
seeds=(21,23)

4.150548	4.155611	4.167043	4.194750	4.272062
4.295577	4.416106	4.326548	4.255034	4.477531
6.119734	6.170942	6.038863	6.074052	6.197310
6.206653	6.244928	6.204394	6.291408	6.145629
4.997062	5.021716	5.031688	5.087910	4.918983
4.919220	4.969139	5.088148	4.968181	5.157207
5.907382	5.894323	5.884533	5.946950	5.991971
5.932769	6.054843	6.067191	5.939481	5.966097

Table 4. Effect of Noise on Inversion

noise added to data	velocities	residual	gradient
.1% (1,3)	11.998,13.994,16.003	2.16E-2	4.90E-5
(2,7)	11.998,14.014,16.004	1.76E-2	4.75E-5
2% (17,4)	12.126,13.644,16.169	.347	2.63E-5
(21,23)	12.038,13.980,15.970	.431	3.86E-5
5% (4,4)	12.099,14.878,15.760	.482	2.72E-5
(6,1)	12.308,13.983,15.780	1.040	1.05E-5

(i1,i2) refer to input seeds to the random number generator

It is hard to qualitatively measure the stability of a problem, but here even at the 5% noise level we are still obtaining useful velocity estimates. One would expect this to be a stable problem as the condition

number (ratio of largest to smallest singular value) is only about 60 here.

Our above example was for a plane geometry and we inverted compressional travel times for the compressional velocities. We consider now an inversion for both compressional and shear velocities for a fully three dimensional problem. The interfaces are given below in (3.9).

```
z1: -x**2/18 + y/20 -10                                    (3.9)
z2: x/16 -y**3/10    -20
```

Synthetically, we generated the rays (P,P),(P,S),(P,P,P,P),and (P,S,P,P) for the source at the origin and the receivers at x=i,y=i :i=1,5 This model is shown schematically in figure 2.7 below.

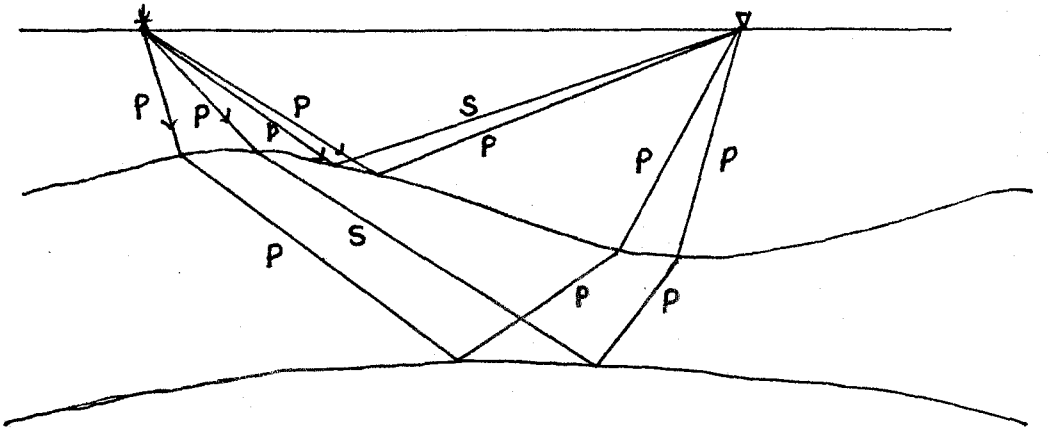


Figure 2.7. Rays generated

The layer velocities we used to synthetically generate the data were:

$$V_{P,1} = 6 \; ; \; V_{S,1} = 3 \; ; \; V_{P,2} = 8 \; ; \; V_{S,2} = 4.6$$

We show in Table 5 that there are two ray solutions for the purely compressional ray from source x = 0, y = 0, z = 0 to receiver x = 2, y = 2, z = 0.

Table 5 Ray Coordinates of Two Solutions

velocities: 6.000, 8.000, 8.000, 6.000

ray coordinates:

interface	x	y	deviation from plane
1	0.000000	0.000000	0.000000
2	0.653699	-0.174863	-0.032483
3	1.755344	-0.571727	0.128397
4	1.630442	0.953441	-0.100014
5	2.000000	2.000000	0.000000

travel time = 5.844254 Family I

ray coordinates:

interface	x	y	deviation from plane
1	0.000000	0.000000	0.000000
2	0.654532	0.235391	-0.012031
3	1.762799	0.379107	0.104726
4	1.635942	1.367517	-0.080308
5	2.000000	2.000000	0.000000

travel time = 5.833599 Family II

Thus we see that the 2 ray solutions have very similar x-coordinates and very similar travel times. The solution which the ray tracing program finds in this example is very dependent upon the initial estimate for the ray. For the ray from the source to receiver (x = 1, y = 1, z = 0), using the "planar solution" (see chapter 1) as an initial estimate resulted in the convergence of Newton's method to family II. Using the "planar solution" with a continuation derivative correction term, Newton's method converges to family I.

The non-linear least squares program, used in this example, calculated

rays which corresponded to family II. We also used the reflections from the first interface in our inversion, but there was no problem of multi-solutions for these rays. Thus for our inversion to make sense, the inversion program should calculate the branch of solutions that corresponds to the data. In this numerical example, we make our synthetic data be of the form of family II. In practice, however, the problem of "close" multisolutions, when it arises, may not be so easily avoided. In Table 6.a, below we show the results of our inversion for the 2 layers' P and S velocities.

Table 6a. Inversion for Shear and Compressional Velocities

it	velocities	residual	gradient
0	(4.00,2.00,10.00,6.00)	7.685	46.347
1	(5.33,2.67,7.55,4.23)	2.642	8.938
2	(5.93,2.96,7.98,4.57)	.246	.686
3	(6.00,3.00,8.00,4.60)	2.78E-3	7.62E-3
4	(6.000000,3.000000, 7.999999,4.600000)	2.02E-6	2.29E-6

As a matter of interest, we use family I in the observational data, and the inversion results are shown in Table 6b.

Table 6b. Inversion Using Family I.

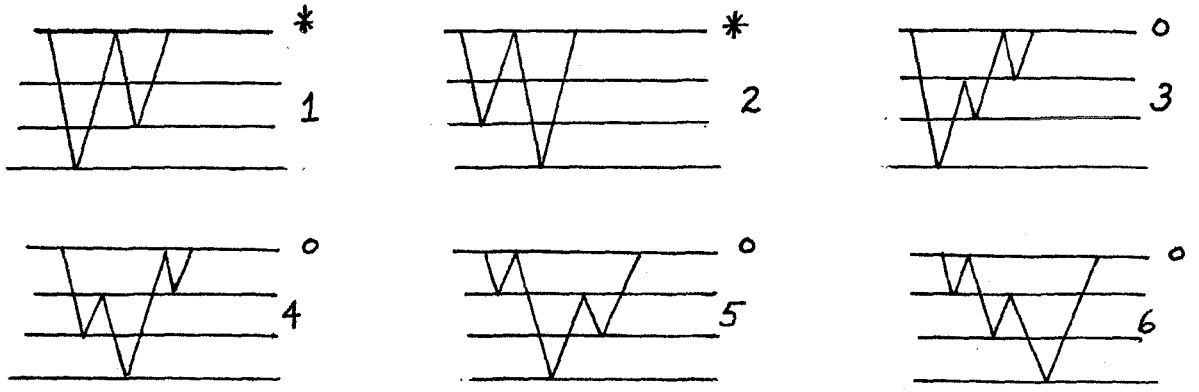
it	velocities	residual	gradient
0	(4.00,2.00,10.00,6.00)	7.661	46.196
1	(5.33,2.66,7.45,4.21)	2.673	9.037
2	(5.93,2.96,7.91,4.56)	.568	1.257
3	(6.00,3.00,7.94,4.59)	3.25E-2	7.66E-3
4	(5.999054,3.000239, 7.940210,4.590474)	3.24E-2	2.20E-6

Thus we see that, for this example, computing on the wrong branch of solutions does not destroy the velocity estimates.

The Ray Labelling Problem (3.2).

Thus far in our inversion examples, we have assumed that from the field data we are able to properly identify the ray types of the different arrivals. Also, we have not addressed the problem of how to "pick" travel times. If one had fairly good estimates of the layer velocities (e.g., from well log surveys), using our forward modelling program (chapter 1) could help determine the ray types which correspond to the different travel times. However, we will consider the problem where one wishes to include various travel times in an inversion but is unsure of the corresponding ray type (e.g., does a particular arrival time correspond to a PPPP,PSPP,or PPSP disturbance, etc.).

For a medium consisting of a stack of parallel layers, there are often several rays with exactly the same travel times and amplitudes. We call a group of rays that have identical travel times for all offsets a kinematically equivalent family. Similarly, a group of rays with identical amplitude characterisitics is called a dynamically equivalent family. From Hron[8], the group of compressional rays shown below in figure 2.8 is an example of a kinematically equivalent family. We also show dynamically equivalent sub-families.



*, o - dynamically equivalent subfamilies.

Figure 2.8 A Kinematically Equivalent Family of Rays

We will characterize a class of rays for the planar case with the vector $(n_{1p}; n_{1s}, n_{2p}; n_{2s}, n_{3p}; n_{3s}, \ldots\ldots\ldots, n_{Jp}; n_{Js})$ where n_{ip} denotes the number of compressional ray segments in the i'th layer and n_{is} the number of shear ray segments in this layer. For the case, where there are no P-S (or S-P) conversions (i.e., $n_{is} = 0$ or $n_{ip} = 0$ for all i), Hron [8] has given the number of kinematically equivalent rays in a class. This number is:

$$N(n_{1p}, \ldots, n_{Jp}) = \prod_{i=1}^{J-1} C_{\bar{n}_{i+1}}^{\bar{n}_i + \bar{n}_{i+1} - 1}$$

$$C_{\bar{n}_{i+1}}^{\bar{n}_i + \bar{n}_{i+1} - 1} = \frac{(\bar{n}_i + \bar{n}_{i+1} - 1)!}{\bar{n}_{i+1}!\,(\bar{n}_i - 1)!} \qquad (3.10)$$

$$\bar{n}_i \equiv n_{ip}/2$$

These concepts of similarity are very useful for forward modelling, using parallel plane layers. When the travel times or the amplitude of a particular ray has been calculated, one has effectively calculated for an

entire family of rays. However, we can clearly see a non-uniqueness in assigning ray types to arrival times or amplitudes, in the inverse problem for a planar geometry. For slightly non-planar interfaces, the travel times for the rays of a kinematically equivalent family will, in general, now all be slightly different.

To illustrate our technique for including unknown ray types in the inversion problem, we consider the following interfaces:

```
z1=-x1**2/18 +y1/20 -5                                    (3.11)
z2=x2/16 -y2/10 -15
```

The velocities used to generate the data were $V_{p,1} = 5$, $V_{s,1} = 2.9$ $V_{p,2} = 7$, and $V_{s,2} = 4$. For our first example, we shall include the rays from the primary reflector (P,P) and (P,S), and from the second reflector (P,P,P,P), (P,S,P,P), (P,S,S,P), and (P,P,S,P). The source is at x = 0, y = 0, z = 0 and the receivers are located at (x = i, y = i, z = 0) (i = 1,5). The synthetic travel time data are shown below in Table 7.

Table 7. Ray Type and Travel Times

PP	PS	PP	PS	PP
2.011079	2.736598	2.070584	2.816977	2.172013
2.956534	2.309438	3.147441	2.476372	3.381163
PPPP	PSPP	PSSP	PPSP	PPPP
4.843695	5.899878	6.954338	5.899583	4.869617
PSPP	PSSP	PPSP	PPPP ...	etc
5.929223	6.979764	5.920315	4.910124	5.974652
7.015685	5.948861	4.964833	6.035735	7.061820
5.985007	5.033229	6.111891	7.117856	6.028548

As we can see, the travel times for the (P,S,P,P) and (P,P,S,P) rays are very close. If the interfaces had been parallel planes, then in fact these times would be identical. In practice, the accurate determination of travel times is a major problem, and whether one could in fact, resolve the small time differences of some of these rays is doubtful, but we will not concern ourselves with this problem. These examples simply serve to illustrate a type of method that can be used to include unknown ray types in the inversion process. However, there are still many problems of travel time determination and ray type identification which we are avoiding. We suppose that we know which of the observations correspond to the (P,P), (P,S) and (P,P,P,P) types of rays, but the other "labellings" we are unsure of, although we assume here that they lie in the group (P,S,P,P), (P,P,S,P) and (P,S,S,P). However, in order to determine the second layer's shear velocity, we wish to include these rays of unknown type in our inversion.

In the inversion program, when we calculate the travel time, for example, for the (P,S,P,P) ray for a particular station, we do not know with which of the three observed times to associate it. In particular, for the residual vector $\underline{F}$ ($F_i \equiv$ calculated traveltime$_i$ - observation$_i$), we do not know which observation to subtract from the calculated travel time. Conversely, we can consider the i'th observation as fixed, and the calculated travel time to associate with it undetermined. We will consider the problem from the viewpoint of trying to determine the observation to use with a fixed calculated travel time.

To formulate this mathematically, we denote the calculated travel times for (P,S,P,P), (P,S,S,P) and (P,P,S,P) rays at a particular receiver as tr_{i_1}, tr_{i_2}, and tr_{i_3}, respectively. We denote the group of observations

for these rays as obs_{i_1}, obs_{i_2}, and obs_{i_3}, and now we wish to determine which of obs_{i_1}, obs_{i_2} and obs_{i_3} to associate with tr_{i_1} etc. Explicitly putting this indeterminancy into the problem we form:

$$F_{i_1} = tr_{i_1} - \alpha_1 obs_{i_1} - \alpha_2 obs_{i_2} - \alpha_3 obs_{i_3}$$

$$F_{i_2} = tr_{i_2} - \beta_1 obs_{i_1} - \beta_2 obs_{i_2} - \beta_3 obs_{i_3} \qquad (3.12)$$

$$F_{i_3} = tr_{i_3} - \gamma_1 obs_{i_1} - \gamma_2 obs_{i_2} - \gamma_3 obs_{i_3}$$

Ideally, we would wish the triplets $(\alpha_1, \alpha_2, \alpha_3)$, $(\beta_1, \beta_2, \beta_3)$ and $(\gamma_1, \gamma_2, \gamma_3)$ to be of the form {(0,1,0), (1,0,0), (0,0,1)} or some permutation of this. In general, the values of $\alpha_1, \alpha_2, \alpha_3, \beta_1$, etc. will not be integral, but one hopes that they will be sufficiently close to zero or one, so that the ray labels can be easily determined. Thus we have reformulated an integer problem as a continuous variable problem. Hopefully, the continuous answer will correctly indicate which integer nodes to examine for the optimal solution. With this in mind, we rewrite (3.12) as:

$$F_{i_1} = tr_{i_1} - \alpha_1 obs_{i_1} - \alpha_2 obs_{i_2} - (1-\alpha_1-\alpha_2) obs_{i_3}$$

$$F_{i_2} = tr_{i_2} - \beta_1 obs_{i_1} - \beta_2 obs_{i_2} - (1-\beta_1-\beta_2) obs_{i_3} \qquad (3.13)$$

$$F_{i_3} = tr_{i_3} - (1-\alpha_1-\beta_1) obs_{i_1} - (1-\alpha_2-\beta_2) obs_{i_2} - (\alpha_1+\beta_1+\alpha_2+\beta_2) obs_{i_3}$$

These same parameters are introduced into the corresponding terms for the

other receivers. By using the same coefficents at each station, we are assuming that by some means (perhaps by establshing some trends in the data) we have been able to arrange the data into groups of unknown ray type. In general, we may wish to introduce new coefficents at specific receivers, indicating that for these stations we are unsure of with which group of travel times to associate the station's unknown ray times. We have introduced four new parameters into the problem and four new columns into the least squares Jacobian. The new Jacobian entries are as follows:

$$\frac{\partial F_{i1}}{\partial \alpha_1} = obs_{i3} - obs_{i1} \qquad \frac{\partial F_{i1}}{\partial \alpha_2} = obs_{i3} - obs_{i2}$$

$$\frac{\partial F_{i2}}{\partial \beta_1} = obs_{i3} - obs_{i1} \qquad \frac{\partial F_{i2}}{\partial \beta_2} = obs_{i3} - obs_{i2} \qquad (3.14)$$

$$\frac{\partial F_{i3}}{\partial \alpha_1} = obs_{i1} - obs_{i3} \qquad \frac{\partial F_{i3}}{\partial \alpha_2} = obs_{i2} - obs_{i3}$$

$$\frac{\partial F_{i3}}{\partial \beta_1} = obs_{i1} - obs_{i3} \qquad \frac{\partial F_{i3}}{\partial \beta_2} = obs_{i1} - obs_{i2}$$

Thus this modified Jacobian has the structure:

		$\frac{\partial}{\partial V_{p,1}}$	$\frac{\partial}{\partial V_{s,1}}$	$\frac{\partial}{\partial V_{p,2}}$	$\frac{\partial}{\partial V_{s,2}}$	$\frac{\partial}{\partial \alpha_1}$	$\frac{\partial}{\partial \beta_1}$	$\frac{\partial}{\partial \alpha_2}$	$\frac{\partial}{\partial \beta_2}$
station 1	(P,P)	x	o	o	o	o	o	o	o
	(P,S)	x	x	o	o	o	o	o	o
station 2	(P,P)	x	o	o	o	o	o	o	o
	(P,S)	x	x	o	o	o	o	o	o
.									
.									
station 5	(P,P)	x	o	o	o	o	o	o	o
	(P,S)	x	x	o	o	o	o	o	o
station 1	(P,P,P,P)	x	o	x	o	o	o	o	o
	(P,S,P,P)	x	o	x	x	x	x	o	o
	(P,S,S,P)	x	o	o	x	o	o	x	x
	(P,P,S,P)	x	o	x	x	x	x	x	x
.									
.									
station 5	(P,P,P,P)	x	o	x	o	o	o	o	o
	(P,S,P,P)	x	o	x	x	x	x	o	o
	(P,S,S,P)	x	o	o	x	o	o	x	x
	(P,P,S,P)	x	o	x	x	x	x	x	x

Figure 2.9 Elements of the Least Squares Jacobian
x-denotes non-zero entry

For the case $V_{p,1} = 5$, $V_{s,1} = 2.9$, $V_{p,2} = 7$, $V_{s,2} = 4$, $\alpha_1 = 1$, $\alpha_2 = 0$, $\beta_1 = 0$, $\beta_2 = 1$, the numerical inversion results are as follows:

Table 8. Inversion for Velocity and Ray Type

IT	VELOCITIES	LABELS	RESIDUAL	GRADIENT
0	(8.00,4.00)	(.300,.300)	10.38	15.46
	(12.00,6.00)	(.300,.300)		
1	(3.20,2.49)	(.977,1.45E-4)	16.08	124.19
	(3.43,3.00)	(1.32E-2,1.000)		
2	(4.35,2.85)	(.989,1.86E-4)	4.57	18.65
	(7.86,5.67)	(-2.25E-2,1.000)		
3	(4.92,2.90)	(.995,9.74E-5)	.728	2.134
	(6.53,3.99)	(-8.34E-3,1.000)		
4	(5.00,2.90)	(1.00,1.02E-5)	3.73E-2	9.36E-2
	(6.97,4.00)	(-1.28E-3,1.00)		

Final estimate: velocities: (5.000000,2.900000,6.999860,
4.000000) labels:(1.000008,2.164E-7,-1.095E-5,.9999999)
residual: 1.56E-4 gradient:3.78E-4

We note that in this example, we allowed Newton's method to continue

unchanged even after an increase in the residual at the first step. In this problem there is a moderate amount of ill-conditioning. Physically, this arises because the travel times for the (P,S,P,P) and (P,P,S,P) labels are so similar, yet we find distinct labels to assign to these times. Thus small perturbations to these travel times can give rise to large variations in the resultant labels. With one percent random noise added to the observational data we obtain the following estimates:

```
velocities: (5.04,2.91,6.97,3.98)
labels:(.769,1.46E-2,.1464,.996)
residual=.147       gradient=4.41E-6
(Seeds for random number generator =(17,2)
```

Thus with this noise level, which can be greater than the time difference between the (P,S,P,P) and (P,P,S,P) rays for most of the receivers, we cannot clearly distinguish between these ray types. However, simple arithmetic shows that we do weight the observation for the (P,S,S,P) ray correctly, and our velocity estimates are good.

We now try another numerical example. Our interface geometry is as above (3.11) and the velocities are the same as the above example. We calculate the rays (P,P),(P,P,P,P),(P,S,P,P) (tri1),(P,S,S,P) (tri2), (P,P,S,P) (tri3), and (P,P,S,S) (tri4) for the source at x=0,y=0,z=0 and receivers at x=.05+(i-1),y=x and z=0 (i=1,6). For this case, we assume we know the purely compressional observational travel times. However,we consider that for the remaining four observations for each station, we do not know whether they are of (P,S,P,P),(P,S,S,P),(P,P,S,P) or (P,P,S,S) type. Including the unknown ray types in the inversion, introduces nine additional parameters. Following (3.14), we can write for this case:

$$F_{i1} = tr_{i1} - \alpha_1 obs_{i1} - \alpha_2 obs_{i2} - \alpha_3 obs_{i3} - (1-\alpha_1-\alpha_2-\alpha_3) obs_{i4}$$

$$F_{i2} = tr_{i2} - \beta_1 obs_{i1} - \beta_2 obs_{i2} - \beta_3 obs_{i3} - (1-\beta_1-\beta_2-\beta_3) obs_{i4}$$

$$F_{i3} = tr_{i3} - \gamma_1 obs_{i1} - \gamma_2 obs_{i2} - \gamma_3 obs_{i3} - (1-\gamma_1-\gamma_2-\gamma_3) obs_{i4} \qquad (3.15)$$

$$F_{i4} = tr_{i4} - (1-\alpha_1-\beta_1-\gamma_1) obs_{i1} - (1-\alpha_2-\beta_2-\gamma_2) obs_{i2} - (1-\alpha_3-\beta_3-\gamma_3) obs_{i3} + (\alpha_1+\alpha_2+\alpha_3+\beta_1+\beta_2+\beta_3+\gamma_1+\gamma_2+\gamma_3-2) obs_{i4}$$

Numerical inversion results for this example are shown below in Table 9. (the observations are such that: $\alpha_1=1, \alpha_2=0, \alpha_3=0, \beta_1=0, \beta_2=1, \beta_3=0, \gamma_1=0, \gamma_2=0$, and $\gamma_3=1$).

Table 9. Inversion for Velocity and Ray Type: Example 2

IT	VELOCITIES	LABELS	RESIDUAL	GRADIENT
0	(3.0,1.4)	(0,0,1),(1,0,0)		
	(5.0,3.0)	(0,1,0)	17.70	167.13
1	(4.20,2.21)	(1.41,.225,-.372)		
	(6.43,3.71)	(9.24E-2,1.075,-.179)	4.63	24.79
		(.180,9.56E-2,.802)		
2	(4.87,2.79)	(1.08,4.67E-2,-7.82E-2)		
	(6.95,3.96)	(2.92E-2,1.02,-4.97E-2)	.585	2.62
		(4.15E-2,2.21E-2,.954)		
3	(5.00,2.90)	(1.00,1.07E-3,-1.89E-3)		
	(7.00,4.00)	(1.24E-3,1.00,-1.92E-3)	1.25E-2	5.44E-2
		(8.55E-4,4.74E-4,.999)		
4	(5.00,2.90)	(1.000,9.64E-5,-1.61E-4)		
	(7.00,4.00)	(-7.68E-5,1.000,6.08E-5)	8.64E-6	3.69E-5
		(-2.66E-4,-1.26E-4,1.00)		

The condition number of $\underline{\underline{J}}$ is approximately 5E3. Adding 1% noise destroys the labels, but we still have fair velocity estimates. We note that if we had left out the rays (P,S,P,P),(P,S,S,P), (P,P,S,P) and (P,P,S,S) we could not have determined either of the layers' shear velocities. With the same initial estimates as above but with 1% random noise (seeds=(1,3)), we

converge to the following estimates:

velocities (5.115,2.513,6.842,4.064)
labels (.605,2.11E-2,.444) (-.732,.272,.654)
(-.627,-.350,1.581)
residual: .122 gradient: 7.21E-6

or: $\alpha_1 = .605$ $\alpha_2 = .0211$ $\alpha_3 = .444$ $\alpha_4 = (1-\alpha_1-\alpha_2-\alpha_3) = -.070$

$\beta_1 = -.732$ $\beta_2 = .272$ $\beta_3 = .654$ $\beta_4 = (1-\beta_1-\beta_2-\beta_3) = .806$

$\gamma_1 = -.627$ $\gamma_2 = -.350$ $\gamma_3 = 1.581$ $\gamma_4 = (1-\gamma_1-\gamma_2-\gamma_3) = .396$

$(1-\alpha_1-\beta_1-\gamma_1) = 1.754$ $(1-\alpha_2-\beta_2-\gamma_2) = 1.057$

$(1-\alpha_3-\beta_3-\gamma_3) = -1.678$ $(1-\alpha_4-\beta_4-\gamma_4) = -.132$

3.3 Inversion for Interface Shape and Layer Velocities.

Thus far in our examples we have assumed that the depths and the shapes of the interfaces are known. Now we will include the depths and shapes of the reflectors as parameters in the inversion scheme. We wish the reflectors to be represented in some form that can be easily parametrizable. For example, one might represent the interfaces by a collection of bicubic splines with unknown coefficents. In our case to simplify the programming, we take the interfaces to be arbitrarily dipping planes. Thus $z_i = \alpha_i x + \beta_i y + \gamma_i$ (i=1, number of interfaces). We shall also take the layers' velocities as unknown.

A least squares approach to this type of problem has also been formulated by Gyoystdal and Ursin [5],but their method is different than ours. They utilize timemaps at the surface (a timemap is the function tr(x,y) where tr(x,y) is the zero offset travel time at source-receiver location (x,y) for a specific sequence of interfaces). This allows them, with an initial velocity estimate, to find the rays' intersections with the

interfaces. Cubic splines are then fitted to these intersection points to approximate the interfaces. Non-zero offset travel times are then used with the interfaces found from above to estimate the layer velocities, and the procedure is then repeated iteratively.

We do not require timemaps, and we invert simultaneously for all the unknown parameters. In the section on layer velocity estimates we had the formula for the Jacobian elements:

$$J_{ik} = \frac{\partial tr_i}{\partial V_k} = -\sum_{j=1}^{N+1} \frac{D_j}{V_j} \delta_{j,k} \qquad (3.16)$$

($\delta_{j,k}=1$, if Vj corresponds to the same layer and velocity type (P or S) as Vk; otherwise $\delta_{j,k}=0$). To invert for the plane's parameters $(\alpha_k, \beta_k, \gamma_k)$, we must calculate $\partial tr_i/\partial \alpha_k$, $\partial tr_i/\partial \beta_k$, and $\partial tr_i/\partial \gamma_k$. From

$$tr_i = \sum_{j=1}^{N+1} \frac{\sqrt{(x_{j,i}-x_{j-1,i})^2+(y_{j,i}-y_{j-1,i})^2+(z_{j,i}-z_{j-1,i})^2}}{V_j} \qquad (3.17)$$

(where a subscript zero refers to the source location $\underline{X}_S$
a subscript N+1 refers to the receiver location $\underline{X}_R$)

we obtain:

$$\frac{\partial t r_i}{\partial \alpha_k} = \sum_{j=1}^{N+1} \frac{\Delta z_j \left(\tau_{j,k} x_j - \Omega_{j,k} x_{j-1} \right)}{D_j V_j}$$

$$\frac{\partial t r_i}{\partial \beta_k} = \sum_{j=1}^{N+1} \frac{\Delta z_j \left(\tau_{j,k} y_j - \Omega_{j,k} y_{j-1} \right)}{D_j V_j} \qquad (3.18)$$

$$\frac{\partial t r_i}{\partial \gamma_k} = \sum_{j=1}^{N+1} \frac{\Delta z_j \left(\tau_{j,k} - \Omega_{j,k} \right)}{D_j V_j}$$

where $\tau_{j,k}$ =1 if $\underline{X}_j$ is on the k'th interface;
0 otherwise
$\Omega_{j,k}$ =1 if $\underline{X}_{j-1}$ is on the k'th interface;
0 otherwise.

With this expanded Jacobian, we are ready to include the interface parameters in the inversion. For our first numerical example, we use the following model:

$$z1=.08\ x1 - .1\ y1 - 5 \qquad (3.19)$$
$$z2=.20\ x2 + .1\ y2 - 10$$

$$Vp,1=12 \quad Vs,1=6 \quad Vp,2=14 \quad Vs,2=8.2$$

For the first inversion example, we shall use one array: source at x=0,y=0, z=0 and receivers x=i,y=i,z=0,i=1,5. We synthetically generate the travel time data for the (P,P),(P,S),(P,P,P,P), and (P,S,P,P) rays at each receiver. We observe that there is a stability problem with this type of inversion using a single linear array. For example, if we make the interface parameter estimates $\alpha_1 = \beta_1$ and $\alpha_2 = \beta_2$, then the calculated ray solutions are such that Xi=Yi for all i, and it is easy to see from (3.18) that there are two pairs of identical columns. Thus the dimension of the right null-space of J is two and is spanned by the vectors $\underline{P}_1 = [0,0,0,0,\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0,0,0,0]^T$ and $\underline{P}_2 = [0,0,0,0,0,0,0,\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0]^T$. This is physically sensible. For this case, the ray solutions all lie in the

vertical plane determined by the line x=y as shown below in figure 2.10.

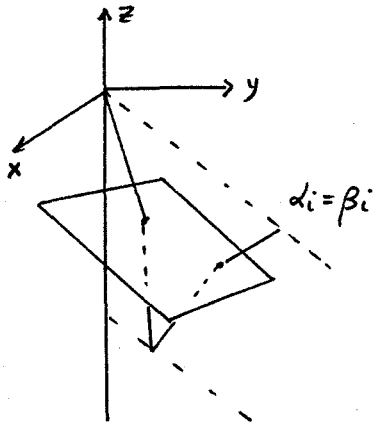


Figure 2.10 Rays Lie in Plane Determined by Line x=y

Changing the parameters α_1 , β_1 so that $\alpha_1 = \beta_1$ has no effect on the depth of the interface along the line x=y; similarly for the second interface and α_2 and β_2. Thus to first order in α_1 , β_1 , α_2 , and β_2 , changes in the parameters in the above form have no effect on the travel times for the array along x=y; z=0. Using a singular value decomposition of the Jacobian matrix, we find the estimate which is orthogonal to the null-space(i.e., we stay in the space $\alpha_1 = \beta_1$ $\alpha_2 = \beta_2$). Thus if our initial parameter estimate is such that $\alpha_1 = \beta_1$, and $\alpha_2 = \beta_2$ we will always remain in this space at each iteration in the non-linear least squares algorithm.

Numerical Inversion Example. Table 10

```
iteration 1

velocities 8,4,10,5
interface (0,0,-4)
          (0,0,-8)
residual 1.792 gradient  3.908
singular values 2.868,1.782,.467,.370,.340,.193,
                2.31E-2,3.99E-3,5.00E-9,5.99E-16
```

```
iteration 2

velocities 10.59,5.30,13.40,7.37
interface  (-6.338E-3,-6.341E-3,-4.61)
           (9.878E-2,9.878E-2,-9.397)
residual  .433 gradient  .888
singular values 2.180,1.953,.346,.232,.164,.115,
                1.23E-2,2.19E-3,9.13E-8,8.80E-8

iteration 3

velocities 11.82,5.91,14.08,8.20
interface  (-9.785E-3,-9.753E-3,-4.92)
           (.148,.148,-9.93)
residual 3.03E-2  gradient 5.87E-2
singular values 2.208,1.927,.309,.200,.127,9.62E-2,
                9.70E-3,1.97E-3,1.06E-6,3.81E-8

iteration 4

velocities 12.00,6.00,13.98,8.20
interface  (-1.001E-2,-9.899E-3,-4.96)
           (.1496,.1496,-9.97)
residual 4.17E-4 gradient 7.798E-4
singular values 2.247,1.900,.303,.196,.127,9.61E-2,
                9.39E-3,1.96E-3,4.02E-6,5.88E-8

iteration 5

velocities 12.00,6.00,13.98,8.19
interface  (-1.001E-2,-9.899E-3,-4.96)
           (.1497,.1496,-9.97)
residual 3.16E-5 gradient 4.66E-7
```

For the more general case of $\alpha_1 \neq \beta_1$, and $\alpha_2 \neq \beta_2$, there still seems to be a null-space. That is, one can find a direction of parameter changes so that to first order the travel times along the array x=y are unchanged.

Table 11. Singular Velocity/Interface Inversion

i	velocities	interfaces	resid	grad	Smin
0	(8.000,4.000) (10.00,5.000)	(.1000,0.000,-4.00) (.0000,.1000,- 8.00)	1.565	3.854	2.4E-7
1	(10.519,5.259) (13.394,7.077)	(.456,-.427,-4.730) (7.04E-2,.156,-9.37)	.371	.758	6.0E-8
2	(11.653,5.829) (13.3497,7.077)	(.219,-.236,-5.00) (.104,.199,-10.01)	2.88E-2	4.23E-2	6.4E-8
3	(11.983,5.992) (14.020,8.258)	(.125,-.145,-5.01) (.165,.134,-9.96)	2.52E-2	4.59E-2	6.6E-8
4	(12.000,6.000) (14.000,8.200)	(9.7E-2,-.117,-5.01) (.191,.108,-9.99)	4.37E-3	8.80E-3	4.6E-8
5	(12.000,6.000) (14.000,8.200)	(8.6E-2,-.106,-5.00) (.201,9.9E-2,-10.00)	6.50E-4	1.29E-3	4.5E-8
6	(12.000,6.000) (14.000,8.200)	(8.3E-2,-.103,-5.00) (.203,9.7E-2,-10.00)	3.48E-5	6.78E-5	

When we add data from a second linear array we stabilize the problem. To numerically illustrate this we use the same model as above, but also generate the travel times for the array of receivers x=-i,y=i, z=0, i=1,5. A numerical inversion example is shown below in Table 12.

Table 12. Velocity/Interface Inversion: 2 Arrays

it	velocities	interfaces	resid	grad	Smin
0	(8.000,4.000) (16.00,7.000)	(.000,0.000,-4.00) (.000,0.000,-8.00)	1.28	3.81	3.6E-3
1	(10.56,5.283) (15.99,9.526)	(.053,.062,-4.59) (.196,.118,-10.08)	.287	.598	2.3E-3
2	(11.82,5.917) (14.09,8.259)	(.077,-.097,-4.94) (.197,.124,-10.03)	.034	.065	2.7E-3
3	(12.00,5.999) (14.00,8.203)	(.080,-.100,-5.00) (.200,.010,-10.00)	1.6E-3	3.6E-3	2.6E-3
4	(12.000,6.000) (14.000,8.2000)	(.080,-.100,-5.00) (.200,.100,-10.00)	4.6E-4	9.0E-4	1.7E-3

Final estimates: velocities:(12.00001,6.000006,13.99997,8.1999982)
interfaces:(8.0000073E-2,-.1000006,-5.000005)
(.1999997,9.9999577E-2,-9.999994)
residual:2.02E-6 gradient:2.18E-6

As can be seen below in Table 13, these estimates are sensitive to noise. Finally we consider another model. The interfaces are:

$$z1: z1=.11\ x1 + .1\ y1 - 5. \qquad (3.20)$$

$$z2: z2=.05\ x2 + .1\ y2 - 10.$$

and the velocities Vp,1=6 Vs,1=3 Vp,2=8 Vs,2=4. Once again we generated the travel time data for the (P,P),(P,S),(P,P,P,P), and (P,S,P,P) rays at the receivers of the two linear arrays described above. A numerical inversion example is given below in Table 14.

Table 13. Effect of Noise on Estimates

%	velocity est.	interface estimate	final resid.	final grad.
.1	(12.00 ,6.00)	(.088 ,-.102 ,-4.99)	4.6E-3	3.2E-7
1,3	(13.70 ,8.02)	(.197 ,.101 ,-9.89)		
.1	(11.97,5.99)	(.080,-.099,-4.98)	4.7E-3	4.8E-7
13,5	(13.82,8.10)	(.198,.102,-9.93)		
.5	(11.83,5.93)	(.082 ,-.104,-4.91)	2.1E-2	7.6E-7
6,9	(16.91,9.99)	(.229,.0854,-10.97)		
.5	(12.14,6.07)	(.085,-.108,-5.05)	2.5E-2	3.3E-6
14,1	(14.22,8.36)	(.200,.084,-10.16)		
.8	(11.66,5.78)	(.083,-.083,-4.83)	3.8E-2	1.8E-6
17, 19	(13.04,7.73)	(.186,.110,-9.24)		
.8	(11.41,5.66)	(.082,-.065,-4.73)	3.7E-2	1.1E-5
3,5	(12.08,7.09)	(.183,.115,-9.07)		
1	(11.59,5.75)	(.080,-.092,-4.76)	4.2E-2	2.8E-6
47, 22	(11.94,7.05)	(.191,.115,-9.12)		
1	(11.81,5.90)	(.087,-.106,-4.88)	3.8E-2	3.2E-5
3,2	(17.14,10.25)	(.221,.100,-11.06)		

ι_1 , ι_2 are the input seeds to the random number generator

Table 14. Model (3.20) Estimates

it	velocities	interfaces	resid	grad	Smin
0	(6.00 ,5.00)	(0.000,0.000,-5.00)	2.47	5.169	6.5E-3
	(6.00 ,5.00)	(0.000,0.000,-10.00)			
1/2	(5.82 ,1.80)	(.163,.204,-4.92)	2.60		
	(7.91 ,4.10)	(.064,.103,-10.14)			
1*	(6.30 ,4.55)	(.139,-.278,-5.29)	1.454	1.924	7.5E-3
	(6.36 ,4.67)	(.054,.144,-9.90)			
1.5	(5.55 ,2.34)	(.072,.412,-4.05)	2.17		
	(7.88 ,4.00)	(.029,.297,-9.89)			
2*	(6.48 ,3.98)	(.116,-.044,-5.45)	.973	1.240	6.1E-3
	(6.56 ,4.45)	(.077,.044,-9.87)			
3	(5.65 ,2.61)	(.103,.182,-4.68)	.263	.890	7.6E-3
	(8.61 ,4.38)	(.041,.145,-10.18)			
4	(5.91 ,2.94)	(.108,.115,-4.93)	4.2E-2	.168	7.8E-3
	(8.06 ,4.03)	(.047,.104,-9.99)			
5	(6.00 ,3.00)	(.109,.100,-5.00)	1.2E-3	5.5E-3	7.7E-3
	(8.00 ,4.00)	(-.05,.100,-10.00)			

* - Levenberg-Maquardt parameter set to one.

Final Estimates
Velocities: 5.999988,2.999993,8.000032,4.00017
Interfaces: (.1099999,.1000013,-4.99999)
(.04999926,.09999909,-10.00001)
residual: 2.8E-6 gradient 7.67E-6

In this example we used a non-zero Levenberg-Maquardt parameter twice during the iterations to force the residual down. In this example we reset the parameter to zero after the residual had decreased.

3.4 Inclusion of Amplitude Information.

Thus far we have dealt only with the inversion of travel time data. We

now discuss the inclusion of amplitude information in the inversion process. There are several different definitions of amplitude that can be used. By using seismometers with different orientations at the same recording point, one can record the three amplitudes of the particle displacement in the three cartesian directions, or amplitude can refer to some scalar function of these amplitudes. In the work that follows we shall take "amplitude" to mean the euclidean norm of the displacement vector. We will for these examples calculate and invert using the true observed free surface amplitude (see Chapter 1 2.4).

As mentioned in (2.2), we will calculate the derivatives of amplitude with respect to variations in the parameters by using finite difference approximations. The derivatives of the travel time function are calculated analytically as outlined above. There is a fundamental problem associated with the inversion of amplitude information. The derivatives $\partial amp/\partial \gamma$, where γ is some parameter, are often ill-behaved and for some values of γ discontinuous. (where rays are becoming critical). Thus one can experience problems in the iterative inversion scheme when parameter values are near or pass through these critical regions.

In the numerical example that follows, we avoided this problem by considering arrays with very small offsets, thus keeping the angles of incidence of the calculated rays very small. However, other calculations with larger offsets showed slow convergence, often to parameter values not used in the synthetic data generation. The model for the synthetic data generation was one of those used in the previous section and is shown below in figure 2.11.

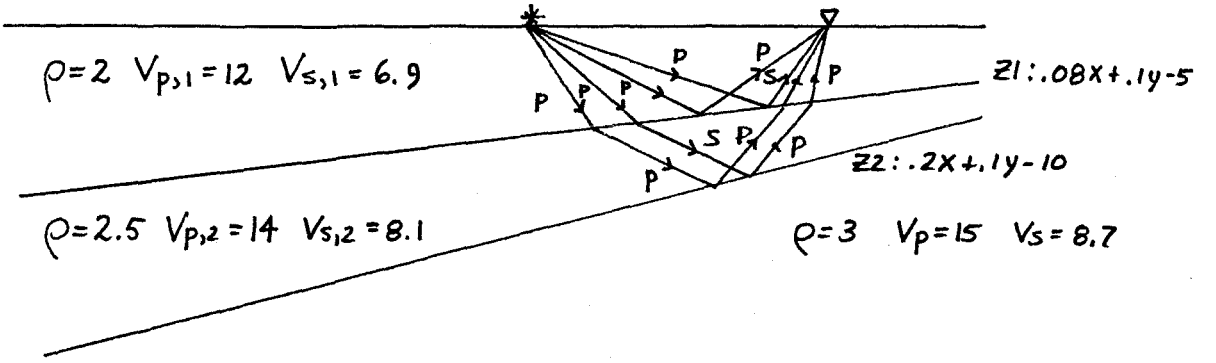


Figure 2.11 Schematic Model for Synthetic Data Generation

We generated the rays (P,P),(P,S),(P,P,P,P), and (P,S,P,P) for the source at x=0,y=0,z=0 and receivers at x=.05 +(i-1)/10 y=x (i=1,3) and x=-.05-(i-1)/10,y=.05+(i-1)/10 (i=1,3). As mentioned previously, our residual function is now weighted,

$$\frac{1}{2}\underline{F}^{T}\underline{F} = \sum_{i=1}^{24} (rt)^2(tr_i - obs_i)^2 + (ra)^2(amp_i - obs_{i+24})^2 \qquad (3.21)$$

and the Jacobian $J_{ij} = \frac{\partial F_i}{\partial y_j}$ is modified accordingly. The results of the inversion for layer velocities for the top 2 layers (we assume that the parameters of the lower half space are known; we also know the layer densities) and the interface parameters are shown below in Table 15.

Table 15. Effect of Weighting Amplitudes

rt	ra	final vel.	final interfaces	resid.	grad.	i
1	0	(9.38,5.39) (8.21,4.78)	(6.23E-2,-6.56E-2 -3.89) (.130,4.83E-2,-6.77	1.86E-4	1.07E-5	3
100	1	(12.00,6.90) (14.00,8.10)	(.0800,-.100,-5.00) (.200,.100,-10.00)	3.67E-4	2.67E-2	4
1	1	(12.00,6.90) (14.00,8.10)	"" ""	1.26E-4	1.71E-6	4
1	5	"" ""	(.0796,-.103,-5.00) (.201,.107,-10.00)	5.59E-4	9.00E-6	4

"i" is the number of required iterations.

For all four of the above runs, the initial estimate was Vp,1 =9, Vs,1 =5, Vp,2 =8, and Vs,2 =4 with interfaces (0.,0.,-4.),and (0.,0.,-8.). We see that for this model we require non-zero weights for the amplitudes to force the estimated parameters to the model values.

As a final example of the use of amplitude information, we consider the following planar example shown below in figure 2.12.

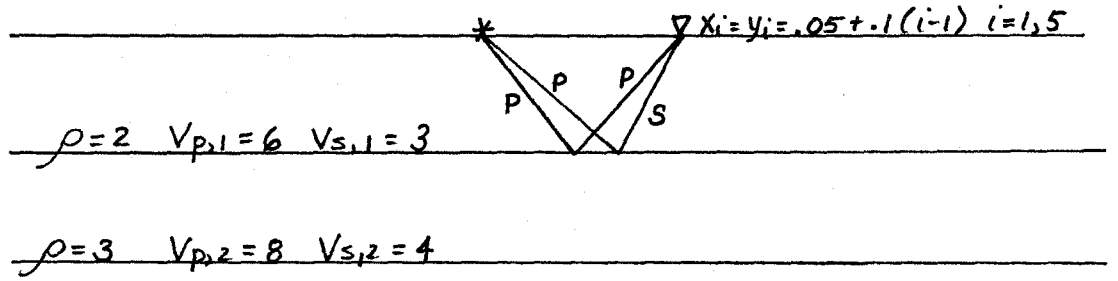


Figure 2.12 Model for Travel Time/Amplitude Inversion

We generate the travel times and amplitudes for the primary reflections,

(P,P) and (P,S). We have no rays passing through the second layer, but by utilizing the amplitude information we hope to estimate this layer's seismic velocities as well. If we set rt=1, ra=0, then at each iteration we have 2 singular values corresponding to Vp,2 and Vs,2. Starting at (Vp,1=10, Vs,1 =5, Vp,2 =7, Vs,2 =3) we arrive at a final velocity estimate of (6,3,7,3). For the case rt=1, ra=1 our results are shown below in Table 16.

Table 16. Velocity Inversion Using Amplitudes

it	velocities	residual	gradient	Smin
0	(9.000,5.000) (10.000,6.000)	2.452	1.140	5.76E-3
.5	(4.5000,1.666) (6.7788,2.548)	3.813		
1	(8.4621,4.2399) (10.001,6.0000)	1.961	1.178	6.20E-3
2	(4.9900,2.4875) (7.0802,3.1534)	1.3729	2.3911	9.72E-3
3	(5.8299,2.9125) (7.8271,3.8658)	0.1989	0.2534	8.53E-3
4	(5.9952,2.9974) (7.9948,3.9935)	5.559E-3	6.708E-3	8.31E-3
5	(6.0000,3.0000) (8.0000,3.9966)	1.267E-5	5.90E-6	

Appendix A - Solution of the Parallel Planar Problem

We consider here the three-dimensional plane problem for a source point $\underline{X}_S$ and receiver at $\underline{X}_R$. For any velocity sequence, V, for the layers, there will always be an unique solution. We immediately satisfy the coplanarity condition by considering all rays to lie in the plane, P, shown in figure A-1, determined by $\underline{X}_R - \underline{X}_S$ and the z-axis.

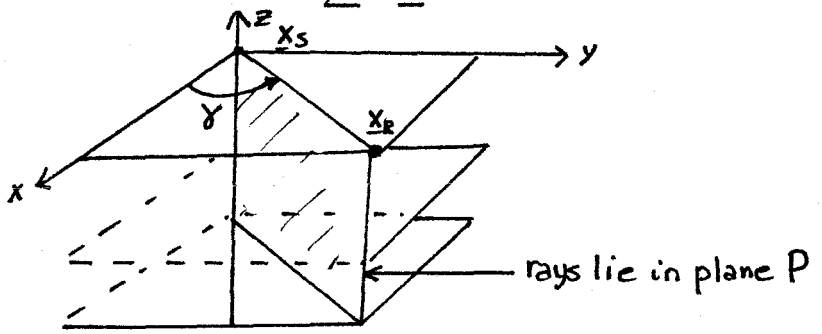


Figure A-1 Geometry for Planar Problem.

Thus we consider only two coordinates: the z-axis, and define the x-axis to lie in the plane P, orthogonal to the z-axis. Looking into the P-plane we have the model shown below in figure A-2.

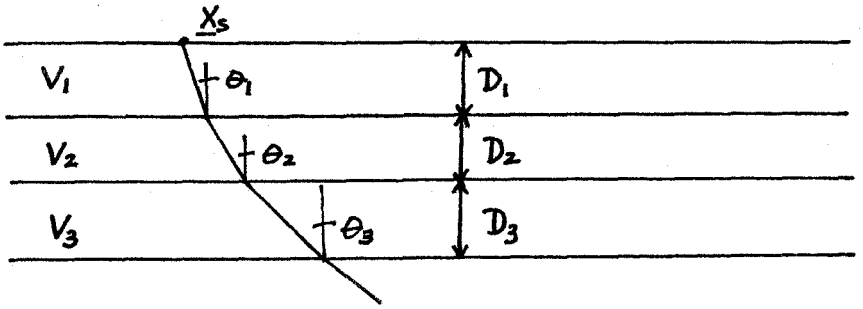


Figure A-2 P-Plane Cross Section of Figure A-1

We let V_K denote the velocity of the k'th layer; D_K is its thickness, and θ_K is the angle that the ray in the k'th layer makes with the z-axis.

From Snell's Law, $\sin\theta_k = V_k/V_1 \sin\theta_1$. The "x" distance travelled by the ray in the k'th layer is $D_k \tan\theta_k = D_k \sin\theta_k / \sqrt{1-\sin^2\theta_k}$, or

$$d_k \equiv D_k \frac{V_k}{V_1} \sin\theta_1 \Big/ \sqrt{1-\left(V_k/V_1\right)^2 \sin^2\theta_1} \qquad \text{(A-1)}$$

Thus, over N layers

$$d \equiv \sum_{k=1}^{N} D_k \alpha_k \sin\theta_1 \Big/ \sqrt{1-\alpha_k^2 \sin^2\theta_1} \qquad \text{(A-2)}$$

where $\alpha_k = V_k / V_1$. Now we consider

$$\hat{F}(x) \equiv \sum_{k=1}^{N} D_k \alpha_k \Big/ \sqrt{1-\alpha_k^2 x^2} \qquad \text{(A-3)}$$

where we restrict "x" ($=\sin\theta_1$) such that $0<x<\min 1/\alpha_k$. In this range, $\hat{F}(x)$ is a continuous function of x ;at x=0, $\hat{F}(0)=0$, and as $x \longrightarrow \min \frac{1}{\alpha_k}$, $\hat{F}(x) \longrightarrow \infty$ Thus from the "Intermediate Value Theorem," $\hat{F}(x)$ must at some point attain the value "d," where "d" is the horizontal distance between source and receiver. Hence, we have a ray solution. Also we can calculate

$$\hat{F}'(x) = \sum_{k=1}^{N} \frac{D_k \alpha_k}{(1-\alpha_k^2 x^2)^{3/2}} \qquad \text{(A-4)}$$

$$\hat{F}''(x) = \sum \frac{3x\alpha_k^3 D_k}{(1-\alpha_k^2 x^2)^{5/2}} \qquad \text{(A-5)}$$

We see that $\hat{F}(x)$ is strictly increasing and hence $\hat{F}(x)$ is one-to-one from $(0, \min \frac{1}{\alpha_k})$ to $(0, \infty)$.

We know that $\hat{F}(0) < d$ and that as $x \longrightarrow \min \frac{1}{\alpha_k}$ $\hat{F}(x) > d$. Thus we can

divide the interval [0, .9 min $\frac{1}{\alpha_k}$] into 9 increments, Δx, and starting with x1, locate the first xi such that F(xi) > d. Now using the bisection method we can reduce the residual of F(x)-d to less than one (typically for our problems two bisections). Applying Newton's method, this residual can quickly (2 or 3 iterations) be reduced to .00001. With this value for x, $x = \sin\theta_1$, we calculate d_k, in the plane of incidence from (A-1). These coordinates are then converted to cartesian coordinates from

$$x_k = x_s + \left(\sum_{j=1}^{k} d_j\right)\cos\gamma$$

$$y_k = y_s + \left(\sum_{j=1}^{k} d_j\right)\sin\gamma \qquad \text{(A-6)}$$

It is interesting to note that, for this problem, we can say when Newton's method is guaranteed to converge. From the Fourier conditions (see Burden et al. [2]), Newton's method will converge in the interval (a,b) if the following conditions hold:

(i) $F(a)\,F(b) < 0$

(ii) $F'(x)$ is either strictly increasing or decreasing on (a,b)

(iii) $F''(x)$ does not change sign on (a,b)

(iv) if "c" is the endpoint of (a,b) at which $|F'(x)|$ is smallest, then $\left|\frac{F(c)}{F'(c)}\right| < b-a$

In our case $F(x)=\sum_{k=1}^{N}\frac{D_k d_x x}{\sqrt{1-\alpha_k^2 x^2}} - d$, and (a,b) =(0, min $\frac{1}{\alpha_k}$) and F(x) has the

following properties.

(i) $F(a)\,F(b) < 0$ for $d > 0$

(ii) $F'(x)$ is strictly positive

(iii) $F''(x)$ is non-negative on (a,b)

(iv) 0 is the endpoint at which $|F'(x)|$ is smallest and $|F(0)/F'(0)| = d/\sum_{k=1}^{n} D_k \alpha_k$.

Thus if $d/\sum_{k=1}^{n} D_k \alpha_k < \min \frac{1}{\alpha_k}$, we have guaranteed convergence of Newton's method for any initial guess in $(0, \min \frac{1}{\alpha_k})$.

References

1. A. Ben-Israel. "A Newton-Raphson Method for the Solution of Systems of Equations," Journal of Mathematical Analysis and Applications, 15(1966), pp. 243-252.

2. R.L. Burden, J.D. Faires, and A.C. Reynolds. Numerical Analysis (1978), Prindle, Weber & Schmidt, Boston.

3. V. Cerveny, I. Molotkov, and I. Psencik. Ray Method in Seismology (1977), Univerzita Karlova, Praha.

4. V. Cerveny and R. Ravindra, Theory of Seismic Headwaves (1971), University of Toronto Press, Toronto.

5. H.Gjoystadl and B.Ursin. "Inversion of Reflection Times in Three Dimensions," Geophysics, 46, No.7(July 1981),

6. G.H. Golub. "Numerical Methods for Solving Linear Least Squares Problems," Numerische Mathematik, 7(1965), pp. 206-216.

7. G.H. Golub and C.Reinsch. "Singular Value Decomposition and Least Squares Solutions," Numerische Mathematik, 14(1978), pp. 403-420.

8. F.Hron. "Numerical Methods of Ray Generation in Multilayered Media," Methods of Computational Physics, 12(1972), Academic Press, New York, pp. 1-34.

9. H.B. Keller and D.J. Perozzi. "Fast Seismic Ray Tracing," preprint (1983).

10 H.B. Keller."Global Homotopies and Newton Methods," Recent Advances in Numerical Analysis (ed., C.de Boor and G.H. Golub) (1978), Academic Press, New York.

11. H.B. Keller. "Propagation of Stress Discontinuities in in Inhomogeneous Elastic Media," SIAM Review, Volume 6., Number 4 (October 1964), pp. 356-382.

12. C.L.Lawson and R.J.Hanson. Solving Least Squares Problems. Prentice Hall, Eaglewood Cliffs, 1974.

13. L.Nazareth. "Some Recent Advances to Solving Large Residual Non-Linear Least Squares Problems," SIAM Review, 22, No.1(Jan.1980), pp. 1-12.

14. M.R.Osborne. "Some Aspects of Non-Linear Least Square Calculations.", Conference on the Numerical Methods for Non-Linear Optimization, University of Dundee, 1971. (ed. F.A. Lootsma) (1972), Academic Press, London, pp. 171-191.

15. D.J. Perozzi. "I. Seismic Ray Tracing in Piecewise Homogeneous Media, II. Analysis of Optimal Step Size Selection in Homotopy and Continuation Methods," Ph.D Thesis, California Institute of Technology, Pasadena (1980).

16. P.M. Shah. "Ray Tracing in Three Dimensions," Geophysics, Vol.38, No.3, (June 1973), pp. 600-604.

17. O.N. Stavroudis. "Simpler Derivation of the Formulas for Generalized Ray Tracing," Journal of the Optical Society, Vol.66, No.12 (December 1976), pp. 1330-1333.

18. J.J. Stoker. Differential Geometry (1969), Wiley-Interscience, New York.

19. G.Strang. Linear Algebra and its Applications. Academic Press, New York, 1980.

20. R.Wiggins. "The General Linear Inverse Problem: Implication of Surface Waves and Free Oscillations For Earth Structure," Reviews of Geophysics and Space Physics, 10, No.1(Feb.1972), pp.251-285.

PART II INVERSE SCATTERING AND CURVED RAY TOMOGRAPHY WITH APPLICATIONS TO SEISMOLOGY.

II.0 Introduction Part II

In Part II of this thesis, we discuss two methods of approaching inverse problems in seismology. As discussed in the introduction to chapter 2, Part I, we are interested in the determination of parameters $\underline{\alpha}$ from a knowledge of some observations $\underline{K}$

$$\underline{K} = \mathcal{L}\,\underline{\alpha} \tag{0.1}$$

In chapter 2 Part I we estimated $\underline{\alpha}$ by the use of non-linear least squares. In Part II, we deal with the exact and approximate inversion of equations of the form (0.1) by more analytical methods.

In chapter 1, we discuss the use of inverse scattering theory to determine the velocity and density (or sometimes just the impedance profile) for physical models where the parameters vary only with depth. Here $\underline{K}$ is the observed particle velocity, $\underline{\alpha}$ is the velocity and density profiles and $\mathcal{L}$ is an integral operator relating $\underline{K}$ and $\underline{\alpha}$. In chapter 2, we consider the situation where the velocity field of the medium may vary three dimensionally. We linearize (0.1) by taking the velocity field to be a perturbation from a known background field. $\underline{K}$ will be a vector of observed travel times (actually, travel time perturbations), $\underline{\alpha}$ will be the unknown perturbation to the slowness field (reciprical to the velocity field) and $\mathcal{L}$ will be the operator which projects $\underline{\alpha}$ along rays of the known background field. We will take the background field to vary only with depth and derive approximate inversion formulae for (0.1).

Chapter 1. Applications of Inverse Scattering Theory to Seismology.

1.0 Introduction

In this chapter, we discuss the application of inverse scattering methods to the problem of parameter determination in elasticity. For problems described by a scalar wave equation, where the medium parameters vary only with depth, a knowledge of the surface response of the medium to an impulsive source yields the profile of some parameter or combination of parameters exactly. By an impulsive source we mean a source function which is a spatial delta function at the surface, and in time contains all harmonics. We will also take the source to be a delta function in time. Thus, for this case, the impulse response is simply the Green's function for the problem evaluated at the surface of the medium.

There are several approaches which have been developed in inverse scattering theory. We will examine some of the aspects of the theory of Gelfand and Levitan[11] and its time domain extensions (Symes [18], and Burridge [4]). Although this theory was developed for one dimensional Schroedinger operators in the spectral and time domains, the wave operators which we will consider can be brought to this form through suitable transformations. Very recent work (e.g., Bube and Burridge [3], Santosa and Schwetlick [16]) deal with the wave equation directly, without any transformations to the Schroedinger form. However, the theory of Gelfand and Levitan is still of fundamental importance to the problems of inverse scattering.

For a multidimensionally varying medium, the inverse problem is more

difficult. One can not always transform the problem to Schroedinger form (Jacobs [12]). Approximate perturbation methods have been developed. If we assume that the parameters are perturbations from known parameter fields (the background fields), we can express the Green's function in terms of the Green's function for the background wave operator. This yields the Lippman-Schwinger series. With a knowledge of the Green's function at the surface this series can be iteratively inverted to produce the true parameter fields. (Clayton and Stolt [7],[8], and Stolt and Jacobs [17]) This iterative procedure can be shown (Prosser [14]) to converge, under certain circumstances, to consistent parameter values (i.e., the reconstructed operator will produce the observed data.)

For all these problems the stability of the methods are very important. There are many practical problems that arise in the determination of the true impulse response. The data contains noise, and often the data has to be deconvolved of the source and receiver frequency characteristics. If the source or receiver is bandlimited (finite frequency response), then it is not possible to recover the impulse reponse function.

In the work that follows, we will examine various aspects of the Gelfand- Levitan theory. Some of our approaches are somewhat different from those mentioned in the literature and our numerical examples are new. But this is basically a survey and, in some cases, an extension of the methods already discussed in the literature.

Section 1 Inverse Scattering for the Schroedinger Operator.

1.1 The Integral Equations of Gelfand and Levitan.

Gelfand and Levitan [11] considered the inverse problem of determining q(x) and the boundary parameter h for the half-line eigenvalue problem ((1.1a) with boundary condition (1.1b)), from a knowledge of the spectral function (defined below) for this Schoedinger problem. Later, we shall consider the finite interval inverse problem with the second boundary condition (1.1c).

$$\phi'' + (\lambda - q(x))\phi = 0 \tag{1.1a}$$

$$\phi'(0) + h_1 \phi(0) = 0 \tag{1.1b}$$

$$\phi'(L) + h_2 \phi(L) = 0 \tag{1.1c}$$

The starting point of Gelfand and Levitan's inversion procedures are the representations:

$$\phi(x,\lambda) = \cos\sqrt{\lambda}x + \int_0^x \mathcal{L}(x,t)\cos\sqrt{\lambda}t\,dt \tag{1.2a}$$

or

$$\cos\sqrt{\lambda}x = \phi(x,\lambda) + \int_0^x K(x,t)\phi(t,\lambda)\,dt \tag{1.2b}$$

Now we assume a knowledge of the spectral function $\rho(\lambda)$. The spectral function is defined such that formally

$$\int_{-\infty}^{\infty} \phi(x,\lambda)\,\phi(\xi,\lambda)\,d\rho(\lambda) = \delta(x-\xi) \tag{1.3}$$

We take $\rho(\lambda)$ to be monotonically increasing and of the form

$$\rho(\lambda) = \begin{cases} \frac{2}{\pi}\sqrt{\lambda} + \sigma(\lambda) & \lambda \geq 0 \\ \sigma(\lambda) & \lambda < 0 \end{cases} \tag{1.4}$$

and assume that the integral

$$f(x,y) \equiv \int_{-\infty}^{\infty} \cos\sqrt{\lambda}\,x \cos\sqrt{\lambda}\,y \, d\sigma(\lambda) \tag{1.5}$$

exists. Here, we are assuming that $\rho(\lambda)$ is the spectral function for a Schroedinger equation. Later, we shall give sufficient (and less restrictive than the existence of (1.5)) conditions for an arbitrary function to be the spectral function for some Schroedinger equation and boundary condition. Two integral equations are derived by Gelfand and Levitan; a linear equation:

$$f(x,y) + \int_0^x f(y,t)\,\mathcal{L}(x,t)\,dt + \mathcal{L}(x,y) = 0$$

$$0 \leq y < x \qquad q(x) = 2\frac{d\mathcal{L}(x,x)}{dx}\,; \quad h_1 = \mathcal{L}(0,0) \tag{1.6}$$

and a non-linear integral equation:

$$f(x,y) = K(x,y) + \int_0^y K(t,x)K(t,y)dt \qquad (1.7)$$

$$0 \le y \le x \qquad q(x) = -2\frac{d}{dx}K(x,x) \qquad h_1 = -K(0,0)$$

Equations (1.6) and (1.7) were originally derived using the theory of the spectral properties of the Schroedinger operator. However, it is very useful to think of (1.7) (or we could similarly consider (1.6)) in the time domain. Here we follow closely the papers of Burridge[4] and Symes[18]. Following Symes, and using his notation, we consider the initial value problem.

$$\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} + q(x)\right) U(x,t) = 0$$

$$\frac{\partial U}{\partial x}(0,t) + h_1 U(0,t) = 0 \qquad (1.8a)$$

$$U(x,0) = \delta(x) \qquad \frac{\partial U}{\partial t}(x,0) = 0 \qquad (1.8b)$$

We will call the solution of this problem the Riemann function R(x,t;0,0). The solution for the same problem but with initial conditions (1.8b) replaced by:

$$U(x,0) = 0 \qquad \frac{\partial U}{\partial t}(x,0) = \delta(x) \qquad (1.8c)$$

is given by

$$G(x,t;0,0) = \int_0^t R(x,\sigma;0,0)\,d\sigma \qquad (1.9)$$

From Duhamel's principle, G(x,t ;0,0) is a Green's function for (1.8a).

In Symes [18] and Burridge [4] it is shown that the unknown kernel K(x,t) in (1.7) is simply the regular part, K(x,t;0,0), of the Riemann function R(x,t;0,0). The function $f(x,y)=\int_{-\infty}^{\infty} \cos\sqrt{\lambda}x \cos\sqrt{\lambda}y \; d\sigma(\lambda)$ is recovered from a knowledge of the Riemann's function at x=0.

$$f(x,y) = \frac{1}{2}\left(K(0,x+y;0,0) + K(0,|x-y|;0,0)\right) \tag{1.10}$$

If instead of the initial value problem (1.8b) we have quiescent initial conditions and a point source $\delta(x)\ \delta(t)$ on the right hand side of (1.8a) then from (1.9) the Riemann function R(x,t;0,0) is simply the time derivative of the Green's function G(x,t;0,0). In our applications this time derivative will correspond to particle velocity in seismology. Solving equation (1.6) or (1.7), we can reconstruct the potential q(x) and the boundary condition parameter h.

1.2 Modifications of the Theory for a Finite Interval.

A major advantage of the time domain formulation of the Gelfand and Levitan theory is that in order to determine the potential q(x) for x [0,T), we need only know the impulse response function K(x=0,t;0,0,) for a time interval $0<t<2T$. To recover the spectral information from the signal would require a knowledge of the signal for all time.

Let us now consider in the spectral domain the problem (1.1) with a second boundary condition imposed at x=L. Assuming this operator has a discrete spectrum, then the integral relations (1.6) and (1.7) are still valid for $0<x,y<L$ but now:

$$f(x,y)=\sum_{k=0}^{\infty}\frac{1}{\rho}\cos\sqrt{\lambda_k}\,x\cos\sqrt{\lambda_k}\,y-\frac{(2-\delta_{0,k})}{L}\cos\frac{k\pi x}{L}\cos\frac{k\pi y}{L} \quad (1.11a)$$

where $\{\lambda_k\}, k=1,\infty$ are the eigenvalues and $\rho_k = \int_0^L \phi^2(\lambda_k, x')\,dx'$, where $\phi(\lambda_k,x)$ is the eigenfunction and $\phi(\lambda_k,0)=1$. The second boundary condition parameter, h2, is found by constructing the ratio $\phi'(\lambda_k,L)/\phi(\lambda_k,L)$ if the kernel $\mathcal{L}(x,t)$ has been determined.

Once again, it is easy to see formally, from the separation of variables, that f(x,y) is simply formed from the regular part of the Riemann's function:

$$\mathcal{K}(0,t;0,0)=\sum_{k=0}^{\infty}\frac{1}{\rho_k}\cos\sqrt{\lambda_k}\,t-\frac{(2-\delta_{0,k})}{L}\cos\left(\frac{k\pi t}{L}\right)$$

$$f(x,y)=\frac{1}{2}\left(\mathcal{K}(0,x+y;0,0)+\mathcal{K}(0,|x-y|;0,0)\right) \quad (1.11b)$$

More fundamentally, from causality, the second boundary condition, has no effect on the observed signal at x=0 until time t=2L. Thus for a finite length of time, 0<t<2L, there are an infinite number of discrete spectral representations of the observed signal at x=0, corresponding to finite interval problems on 0<x<L with different boundary conditions at x=L. After time t=2L, the second boundary will influence the observed signal, but this later portion of the signal is not needed for the potential determination. As we shall see later, it is sometimes useful to consider a finite length of signal as a discrete spectral sum.

1.3 Solvability and Stability of the Gelfand-Levitan Equations.

We first consider the non linear integral equation:

$$f(x,t) = K(x,t) + \int_0^x K(s,x)K(s,t)\,ds \qquad (1.7)$$
$$0 \le x \le t \le T$$

Symes [18] has shown that (1.7) can be put in the operator form, for $0<x<t<T$

$$I + \Omega(x,t) = (I + K^T)(I + K) \qquad (1.12)$$

and:

$$\Omega(x,t) \equiv \begin{cases} f(x,t) & x < t \\ f(t,x) & x > t \end{cases} \qquad (1.13a)$$

$$\Omega\phi(x) \equiv \int_0^T \Omega(x,t)\phi(t)\,dt \qquad (1.13b)$$

$$K\phi(x) \equiv \int_x^T K(x,t)\phi(t)\,dt \qquad (1.13c)$$

$$K^T\phi(x) \equiv \int_0^x K(t,x)\phi(t)\,dt \qquad (1.13d)$$

The operator decomposition (1.12) of (1.7) leads to the following.

Thereom 1 (Symes)

The integral equation (1.7) has a solution K(x,y) in the space $W_1^m(C_T)$ if and only if K(0,t;0,0) satisfies:

(i) $K(0,t;0,0) \in W_1^m(0,2T)$

(ii) there exists an $\epsilon_1 > 0$ such that for any $\phi \in L_2[0,T]$

$$\int_0^T |\phi|^2 + \int_0^T dx \int_0^T dt\, \phi(x)\bar{\phi}(t) f(x,t) \geq \epsilon_1 \int_0^T |\phi|^2\, dt \qquad (1.14)$$

where:

$$f(x,t) \equiv \frac{1}{2}\left\{ K(0, t+x; 0, 0) + K(0, |t-x|; 0, 0) \right\}$$

Here W_1^m is the Sobolev space of functions with m absolutely integrable derivatives. The domain C_T is defined as $\{(x,t): 0 \leq x \leq t \leq T\}$ and $\mathcal{W}_1^m$ denotes the subspace of W_1^m with a well-defined restriction to each line in $\mathbb{R}^2$, lying in W_1^m of functions on the line.

As mentioned above, the kernel K(x,t) is the regular part of the Riemann function and hence must satisfy the appropriate wave equation in the sector x<t. Symes [18] showed that there is a complete equivalence between the solution of (1.7) and the non-linear hyperbolic problem:

$$\begin{aligned} &\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} + q(x)\right) K(x,t) = 0 \\ &\frac{\partial K}{\partial x}(0,t) + h\, K(0,t) = 0 \\ &K(0,0) = h \\ &K(0,t) = F(t) \qquad q(x) = -2\,\frac{d K(x,x)}{dx} \end{aligned} \qquad (1.15)$$

This problem can be discretized and solved numerically and Symes [18] has proved stability results for various discretizations of (1.15).

We now discuss the stability of the inverse problem. If we perturb an impulse response function K(0,t;0,0), is the resulting function an impulse

response function, and if so what can we say about the difference between the two potentials? We can easily answer the first question. Symes[18] shows that condition (ii) of Theorem 1 is stable to L_∞ perturbations of F(t)=K(0,t;0,0) In other words, if condition (ii) holds for an impulse response function F1(t), there exists a "δ" such that if $|F1(t)-F2(t)|_\infty<\delta$ (0<t<2T), then condition (ii) holds for F2(t). Thus, F2(t) is also an impulse response function for some potential q(x) and boundary parameter h, if it is sufficiently smooth (condition (i)) We now answer the second question from above.

Thereom 2.

We consider a pair of impulse response functions Fj(t), j=1,2 0<t<2T, with corresponding potentials and boundary parameters (qj(x),hj), xε[0,T]. For any ϵ >0, there exists a $\delta(\epsilon)$ such that if $||F1(t)-F2(t)||_\infty<\delta$, then $|h1-h2|+||\int_0^x (q_1(x')-q_2(x'))dx'||<\epsilon$.

Proof:

We consider the integral equation

$$f(x,y)+\int_0^x \mathcal{L}(x,s)\,f(s,y)\,ds+\mathcal{L}(x,y)=0 \qquad 0\le y\le x\le T \qquad (1.6)$$

Here f(x,y) is the symmetric kernel (F(x+y)+F(|x-y|)/2. For fixed x, g(y)=f(x,y) and $\phi(s)=\mathcal{L}(x,s)$ we can write:

$$g(y)+\int_0^x f(s,y)\,\phi(s)\,ds+\phi(y)=0 \qquad 0\le y\le x \qquad (1.16)$$

From Symes's positive definiteness condition for the kernel f(x,y) in (1.16) it follows that (1.16) has a unique solution. Now we write (1.16) for the kernels f1(x,y),f2(x,y) derived from F1(t) and F2(t).

$$g_1(y) + (I + \hat{F}_1)\phi_1 = 0 \tag{1.17a}$$

$$g_2(y) + (I + \hat{F}_2)\phi_2 = 0 \tag{1.17b}$$

Here Fj is the operator $\hat{F}_j\phi \equiv \int_0^x f_j(s,y)\,\phi(s)\,ds$. Subtracting (1.17b) from (1.17a) we obtain

$$(g_1-g_2) + (\phi_1-\phi_2) + \hat{F}_1\phi_1 - \hat{F}_2\phi_2 \tag{1.18}$$
$$= (g_1-g_2) + (\phi_1-\phi_2) + \hat{F}_1(\phi_1-\phi_2) + (\hat{F}_1-\hat{F}_2)\phi_2 = 0$$

or

$$(\phi_1-\phi_2) = -(I+\hat{F}_1)^{-1}\{(g_1-g_2) + (\hat{F}_1-\hat{F}_2)\phi_2\} \tag{1.19}$$

From the triangle inequality

$$\|\phi_1-\phi_2\|_\infty \le \|(I+\hat{F}_1)^{-1}(g_1-g_2)\|_\infty + \|(I+\hat{F}_1)^{-1}(\hat{F}_1-\hat{F}_2)\phi_2\|_\infty \tag{1.20}$$

We will now bound the first term in the sum of (1.20). We set:

$$\gamma(y) \equiv (I+\hat{F}_1)^{-1}(g_1-g_2) \tag{1.21}$$

and applying $(I+\hat{F}_1)$ to both sides of (1.21)

$$(I+\hat{F}_1)\,\gamma_1(y) = (g_1-g_2) \tag{1.22}$$

Thus from the definition of $\hat{F}_1$

$$\gamma_1(y) + \int_0^x f_1(s,y)\,\gamma_1(s)\,ds = (g_1-g_2) \tag{1.23a}$$

$$\|\gamma_1(y)\|_\infty \le \|(g_1-g_2)\|_\infty + \|\int_0^x f_1(s,y)\,\gamma_1(s)\,ds\|_\infty \tag{1.23b}$$

Now from the Cauchy-Schwarz inequality

$$\left|\int_0^x f_1(s,y)\,\gamma_1(s)\,ds\right| \le \|\gamma_1(\cdot)\|_2\,\|f_1(\cdot,y)\|_2 \qquad \forall y \in [0,x] \tag{1.24}$$

Thus taking the maximum over y, of both sides of (1.24), we obtain

$$\|\int_0^x f_1(s,y)\,\gamma_1(s)\,ds\|_\infty \le \|\gamma_1(\cdot)\|_2 \max_y \|f_1(\cdot,y)\|_2 \tag{1.25}$$

Now, $\gamma_1(y)=(I+\hat{F}_1)^{-1}(g1-g2)$ or $\|\gamma_1(\cdot)\|_2 < \|(I+\hat{F}_1)^{-1}\|_2 \|g1-g2\|_2$. From $g(y)=f(x,y)=(F(x+y)+F(|x-y|))/2.$, we have that $\|g1-g2\|_2=\left[\int_0^x (g_1-g_2)^2\right]^{1/2} < X^{1/2}\delta$ for $\|F_1(t)-F_2(t)\|_\infty < \delta$, or $\|\gamma_1(\cdot)\|_2 < \|(I+\hat{F}_1)^{-1}\|_2 X^{1/2}\delta$. In (1.25) we have for the other term

$$\max_y \|f_1(\cdot,y)\|_2 = \max_y \left(\int_0^x f_1^2(s,y)\,ds\right)^{1/2}$$
$$\le \max_y X^{1/2} \max_{0<\xi<X} |f(\xi,y)| \le X^{1/2}\|f_1(\xi,y)\|_\infty \tag{1.26}$$

Here, we take the maximum norm over $\xi, y \in [0,T]$. Thus,

$$\| \int_0^x f_1(s,y)\,\gamma_1(s)\,ds \|_\infty \leq x\,\delta\, \|(I+\hat{F}_1)^{-1}\|_2 \,\|f_1(\xi,y)\|_\infty \quad (1.27)$$

or from (1.25)

$$\|\gamma_1(y)\|_\infty \leq \delta\,(1 + x\,\|(I+\hat{F}_1)^{-1}\|_2\,\|f_1(\xi,y)\|_\infty) \quad (1.28)$$

Thus, we have bounded the first term in the sum of (1.20).

For the second term of (1.20), $\|(I+\hat{F}_1)^{-1}(\hat{F}_1-\hat{F}_2)\,\phi_2\|_\infty$, following the reasoning above, we have

$$\|(I+\hat{F}_1)^{-1}(\hat{F}_1-\hat{F}_2)\phi_2\|_\infty \leq \delta(1 + \|(I+\hat{F}_1)^{-1}\|_2\, x\, \|f_1(\xi,y)\|_\infty)\|\phi_2\|_\infty \quad (1.29)$$

Finally, we obtain the bound from (1.28) and (1.29)

$$\|\phi_1-\phi_2\|_\infty \leq \delta(1+\|\phi_2\|_\infty)(1+\|(I+\hat{F}_1)^{-1}\|_2\, x\, \|f_1(\xi,y)\|_\infty) \quad (1.30)$$

Now, we consider F_2 ,g_2, ϕ_2 to be fixed quantities, and take F_1 ,g_1, ϕ_1 to be perturbations about the j=2 quantities. We will now replace all the j=1 quantities on the right hand side of (1.30) by j=2 quantities. From the Banach Lemma we can write $\|(I+\hat{F}_1)^{-1}\|_2 \leq \|(I+\hat{F}_2)^{-1}\|_2 / (1-\|(I+\hat{F}_2)^{-1}\|_2\,\delta)$ and clearly $|f1(\xi,y)|_\infty < |f2(\xi,y)|_\infty + \delta$. Thus we can write for (1.30)

$$\|\phi_1-\phi_2\|_\infty \leq \delta\,(1+\|\phi_2\|_\infty)(1+2\|(I+\hat{F}_2)^{-1}\|_2\, T(\|f_2(\xi,y)\|_\infty+\delta)) \quad (1.31)$$

where we have taken $\delta < 1/(2\,x\,\|(I+\hat{F}_2)^{-1}\|_2)$. This bound is good for all

$x \in [0,T]$, when we replace all the norm bounds in (1.31) by their corresponding maximum value for $X \in [0,T]$. For a sufficiently small δ we can make $\|\phi_1 - \phi_2\|_\infty < \epsilon/5$. With the definition of ϕ_j, we have

$$\phi_j(y) \equiv \mathcal{L}_j(x,y)$$
$$\mathcal{L}_j(x,x) = \tfrac{1}{2}\int_0^x q_j(x')\,dx' + h_j \tag{1.32}$$
$$\| \mathcal{L}_1(x,y) - \mathcal{L}_2(x,y) \|_\infty < \epsilon/5$$

which implies

$$\| \mathcal{L}_1(x,x) - \mathcal{L}_2(x,x) \|_\infty < \epsilon/5$$
$$\left\| \tfrac{1}{2}\int_0^x (q_1(x') - q_2(x'))\,dx' + h_1 - h_2 \right\| < \epsilon/5 \tag{1.33}$$

Evaluating (1.33) at x=0, we have $|h1-h2| < \epsilon/5$ so that

$$\left\| \int_0^x (q_1(x') - q_2(x')\,dx' \right\|_\infty + |h_1 - h_2| < \epsilon \tag{1.34}$$

and the proof is complete.

Symes [18] has shown in a Sobolev space setting that the mapping F:-->(q,h) is continuous in the appropriate norms from W_1^{m+1} to $(W_1^m \times \mathbb{R})$. However, in general, for two impulse response functions very close in the L_∞ norm (0<t<2T), all we can say about their corresponding potentials, is that the potentials' "mean" properties will be close.

In this section, we have given necessary and sufficient conditions (Thereom *1*) for a function F(t) (0<t<2T) to correspond to an impulse response function for a Schroedinger wave operator. In the spectral domain, Gelfand and Levitan proved the following:

Thereom 4.

Let $\rho(\lambda)$ be monotonically increasing on $(-\infty, \infty)$ and let

$$\rho(\lambda) = \begin{cases} \frac{2}{\pi}\sqrt{\lambda} + \sigma(\lambda) & \lambda \geq 0 \\ \sigma(\lambda) & \lambda < 0 \end{cases}$$

Suppose $\rho(\lambda)$ satisfies

(i) $$\int_{-\infty}^{0} e^{\sqrt{|\lambda|}x} d\rho(\lambda) < \infty \qquad 0 \leq x < \infty$$

(ii) $$a(x) = \int_{1}^{\infty} \frac{\cos\sqrt{\lambda}x}{\lambda} d\sigma(\lambda)$$ has continuous fourth derivatives

then there exist a continuous potential q(x) and boundary condition parameter h for which $\rho(\lambda)$ is the spectral function.

For a finite interval problem, with a discrete spectrum, the above integrals are replaced with the analagous infinite sums.

1.4 Analytical Solutions of the Inverse Scattering Problem.

For some impulse response functions F(t), the derived kernel $f(s,t) = \{F(t+s)+F(|t-s|)\}/2$ is such that we can analytically solve the linear integral equation (1.6). As we discussed above, we can express a finite length of the observed impulse reponse function as a discrete spectral sum.

$$f(x,y) = \frac{1}{2}(F(x+y)+F(|x-y|) = \sum_{n=0}^{\infty} \frac{\cos\sqrt{\lambda_n}x \cos\sqrt{\lambda_n}y}{\rho_n} - \frac{(2-\delta_{0,n})\cos\frac{n\pi x}{T}\cos\frac{n\pi y}{T}}{T} \qquad (1.47a)$$

We can truncate this series and write

$$f_{N_1}(x,y) = \sum_{n=0}^{N_1} \frac{\cos\sqrt{\lambda_n}x \cos\sqrt{\lambda_n}y}{\rho_n} - \frac{(2-\delta_{0,n})\cos\frac{n\pi x}{T}\cos\frac{n\pi y}{T}}{T} \qquad (1.47b)$$

In the spectral domain, this series corresponds in the spectral domain to the spectral information

$$(\lambda_0, \rho_0, \lambda_1, \rho_1, \ldots, \lambda_{N_1}, \rho_{N_1}, \frac{(N_1+1)^2\pi^2}{T^2}, \frac{2}{T}, \ldots) \qquad (1.48)$$

Assuming $\lambda_{N_1} < \frac{(N_1+1)^2\pi^2}{T^2}$, (1.48) corresponds from Thereom 3 to an impulse response function and in fact, since $f_{N_1}(x,y)$ is analytic in its arguments, so is the calculated $q_{N_1}(x)$. If we choose N1 large enough and $\|F - F_{N_1}\|_\infty < \gamma$, for some $\gamma > 0$, we can from Thereom 2 reconstruct accurately the mean properties of the potential corresponding to F(t) using $F_{N_1}(t)$. The kernel f_{N_1} is a degenerate kernel and hence the integral equation

$$f_{N_1}(x,y) + \int_0^x \mathcal{L}(x,s) f_{N_1}(s,y)ds + \mathcal{L}(x,y) = 0$$

can be solved algebraically. Following Whitham [21] (pp.593-594) we obtain that:

$$q(x) = -2\frac{d}{dx}\left(\frac{1}{|Q|}\frac{d|Q|}{dx}\right) \qquad (1.49a)$$

where:

$$\underline{\underline{Q}} \equiv \left[\begin{array}{c|c} \underline{\underline{I}} + \underline{\underline{P}}_1 & \underline{\underline{P}}_2 \\ \hline \underline{\underline{P}}_3 & \underline{\underline{P}}_4 - \underline{\underline{I}} \end{array}\right] \begin{array}{l} \uparrow \\ 2N+2 \\ \downarrow \end{array} \qquad (1.49b)$$

$$P_{1\,k,j} = \frac{1}{\rho_k}\int_0^x \cos\sqrt{\lambda_k}s\,\cos\sqrt{\lambda_j}s\,ds \quad j,k = 0,\ldots,N \tag{1.49c}$$

$$P_{2\,k,j} = \frac{1}{\rho_k}\int_0^x \cos\sqrt{\lambda_k}s\,\cos\frac{j\pi s}{T}\,ds \tag{1.49d}$$

$$P_{3\,k,j} = \frac{1}{\rho_k}\int_0^x \cos\frac{k\pi s}{T}\,\cos\sqrt{\lambda_j}s\,ds \tag{1.49e}$$

$$P_{4\,k,j} = \frac{(2-\delta_{0,k})}{\pi}\int_0^x \cos\frac{k\pi s}{T}\,\cos\frac{j\pi s}{T}\,ds \tag{1.49f}$$

For $\lambda_k < 0$, we replace cosine terms with hyperbolic cosine terms in the above formulae. If we consider an operator with spectral information

$$\left(0, \frac{1}{T}+\alpha_0, \frac{\pi^2}{T^2}, \frac{2}{T}+\alpha_1, \ldots, \frac{N^2\pi^2}{T^2}, \frac{2}{T}+\alpha_N, \frac{(N+1)^2\pi^2}{T^2}, \frac{2}{T}, \frac{(N+2)^2\pi^2}{T^2}, \frac{2}{T}, \ldots\right) \tag{1.50}$$

(i.e., only the normalizations of the eigenfunctions, are different from the (q=0,h=0) case), we can reduce the above matrix formulation to

$$q(x) = -2\frac{d}{dx}\left(\frac{1}{|Q|}\frac{d|Q|}{dx}\right) : \underline{\underline{Q}} = \underline{\underline{T}} + \underline{\underline{P}} \tag{1.51a}$$

where:

$$P_{kj} = \alpha_k\left(\frac{x}{2} + \frac{T\sin\left(\frac{2\pi kx}{T}\right)}{4k\pi}\right) \quad k=j\neq 0 \tag{1.51b}$$

$$P_{kj} = \alpha_0 x \qquad k=j=0 \tag{1.51c}$$

$$P_{kj} = \frac{T}{2\pi}\left(\frac{\sin\left(\frac{(k+j)\pi}{T}\right)}{(k+j)} + \frac{\sin\left(\frac{(k-j)\pi}{T}\right)}{(k-j)}\right) \qquad (1.51d)$$

Using these formulae, we can derive analytically some results, which we will refer to, later on.

Example 1.

Consider F(t)=H>0, a constant or F(t)=(1/T + H) cos 0t- 1/T cos 0t. Using the formulae (1.52) we obtain:

q(x)=2/(x+1/H)**2 (1.52a)

h=F(0)=H (1.52b)

Example 2.

Consider $F(t)=1/\rho_0 \cos(\sqrt{\lambda_0}\, t) - 1/\pi$ $(0 \leq t \leq 2\pi)$. Using formulae (1.51) we obtain:

$$\underline{\underline{Q}} = \left[\begin{array}{c|c} \left(\frac{x}{2} + \frac{\sin 2\sqrt{\lambda_0}x}{4\sqrt{\lambda_0}}\right)\frac{1}{\rho_0} + 1 & \frac{1}{\rho_0}\,\frac{\sin\sqrt{\lambda_0}x}{\sqrt{\lambda_0}} \\ \hline \frac{1}{\pi}\,\frac{\sin\sqrt{\lambda_0}x}{\sqrt{\lambda_0}} & \frac{x}{\pi} - 1 \end{array}\right] \qquad (1.53)$$

or

$$|Q| = \frac{1}{2}\frac{x^2}{\pi\rho_0} + \frac{x}{\pi\rho_0}\frac{\sin 2\sqrt{\lambda_0}x}{4\sqrt{\lambda_0}} + \frac{x}{\pi} - \frac{1}{\rho_0\pi}\frac{\sin^2\sqrt{\lambda_0}x}{\lambda_0} \qquad (1.54)$$

$$-1 - \frac{1}{\rho_0}\left(\frac{x}{2} + \frac{\sin 2\sqrt{\lambda_0}x}{4\sqrt{\lambda_0}}\right)$$

Example 3.

We consider $F(t)= \alpha \cos(nt)$ $(0<t<2\pi)$ and apply formulae (1.51) to obtain

$$q(x) = 2\left\{n\alpha \sin 2nx + \frac{n\alpha^2 x \sin 2nx}{2} + \alpha^2\frac{\sin 2nx}{4} + \alpha^2\cos^4 nx\right\} \qquad (1.55)$$
$$\frac{}{\left(\frac{\alpha x}{2} + \frac{\alpha \sin 2nx}{4n} + 1\right)^2}$$

We note that as $\alpha \to 0$ $q(x)=2n\alpha \sin(2nx)$, or if we let $\alpha = 1/n$, for $n \to \infty$ $q(x) = 2\sin(2nx)$. Thus $||F(t)||_\infty \to 0$, but $||q(x)||_\infty = 2$. However, $\left|\left|\int_0^x 2\sin(2nx') + \frac{1}{n}\right|\right|_\infty \to 0$ for $n \to \infty$ as predicted by Theorem 2.

Section 2. Application of Inverse Scattering Theory for Schroedinger Operators to Scalar Equations for Elasticity.

2.1 Transformations of General Wave Operators to Schroedinger Form.

We now show that for media with parameters that vary only with depth, we can transform the governing wave operator to Schroedinger form and hence apply the techniques we have described above. We give two examples of problems, of interest, that can be handled in this manner.

Example 1 (from Burridge[4])

A one-dimensional wave equation describing elastic wave propagation in a medium, or wave motion on a string, with a source term is:

$$\hat{\rho}(z)U_{tt} - (\hat{\rho}(z)\hat{c}^2(z)U_z)_z = \delta(z)\delta(t) \qquad (2.1)$$

Here $\hat{c}(z)$ is the wave velocity and $\hat{\rho}(z)$ is the density. A stress free boundary at z=0 implies that $\partial v / \partial z = 0$. We now make the following change of variable:

$$X \equiv \int_0^z \frac{1}{\hat{c}(z')}\, dz' \qquad (2.2)$$

(i.e., X is the travel time to depth z) Equation (2.1) now becomes with $\hat{c}(z) = c(X)$, $\hat{\rho}(z) = \rho(X)$:

$$\rho(x)\, U_{tt} - \frac{1}{c(x)} (\rho c U_x)_x = \frac{\delta(t)\delta(x)}{c(0)} \qquad (2.3)$$

or multiplying through by c(X)

$$\rho(x)c(x)U_{tt} - (\rho c U_x)_x = \delta(t)\delta(x) \qquad (2.4)$$

Now we define $\eta = (\rho(x)\, c(x))^{1/2}$ and letting the new dependant variable $\xi(x) = \eta(x)\, U(x)$, we obtain:

$$\xi_{tt} - \xi_{xx} + \frac{\eta_{xx}}{\eta}\xi = \frac{\delta(t)\delta(x)}{\eta(0)} \qquad (2.5)$$

The boundary condition becomes:

$$U_z = 0 \Rightarrow U_x = 0 \Rightarrow \left(\frac{\xi}{\eta}\right)_x = 0$$

$$\Rightarrow \xi_x - \xi \frac{\eta_x}{\eta} = 0 \quad \text{at} \quad X = 0 \qquad (2.6)$$

Thus, assuming we know η (z=0), we can apply the inverse-scattering theory to obtain $q(x) = \eta_{xx} / \eta$ and $h = -\eta_x(0) / \eta(0)$. Hence for this problem, we must also solve the ordinary differential equation (initial value problem):

$$\eta_{xx} - q(x)\eta = 0$$

$$\eta(0) = \eta_0 \quad ; \quad \eta_x(0) = -h\eta_0 \qquad (2.7)$$

to recover the impedance profile $(\rho(x)c(x))^{1/2}$. We note that the best we can do for this problem is to determine the product $\rho(x)c(x)$. Our answer is left as a function of the travel time X, and since we cannot

uniquely determine c(X), we cannot convert X back to z.

Example 2.

We consider a two-dimensional problem. Figure 1.1 shows a cross section of a three-dimensional space z>0 with a line source at z=0,x=0.

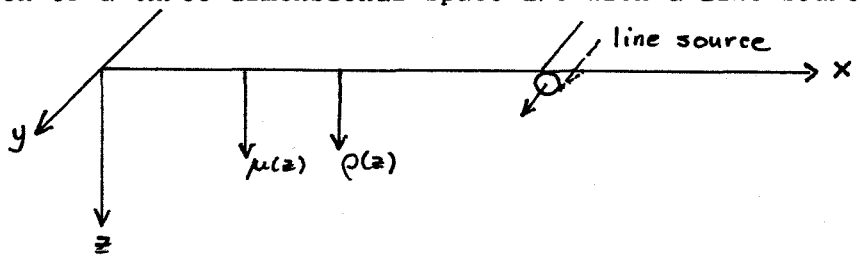


Figure 1.1 SH Line Source in a Half Space.

The wave equation describing the propagation of SH particle motion (displacement in ey direction) is:

$$\hat{\rho}(z)\, U_{tt} - \nabla\cdot(\hat{\mu}(z)\,\nabla U) = \delta(x)\delta(z)\delta(t) \qquad (2.8)$$

with $\partial U/\partial z = 0$ at z=0. We now Fourier Transform (2.8) with respect to "x" to obtain formally:

$$\hat{\rho}(z)\,\tilde{U}_{tt} + \hat{\mu}(z)k_x^2\,\tilde{U} - \frac{\partial}{\partial z}\left(\mu(z)\,\tilde{U}_z\right) = \delta(t)\,\delta(z) \qquad (2.9)$$

We also take the Fourier Transform with respect to "x" of the surface response

$$F(k_x,t) = \int_{-\infty}^{\infty} e^{ik_x x} U(x,z=0,t)\,dx \quad t>0$$

(note that $F(k_x,t)$ is real since U(x,z,t)=U(-x,z,t))

Letting $\hat{c}(z)=(\mu/\rho)^{1/2}$ (2.9) becomes

$$\hat{\rho}(z)\tilde{U}_{tt} + \hat{\rho}(z)\hat{c}^2(z)k_x^2\tilde{U} - (\hat{\rho}\hat{c}^2\tilde{U}_z)_z \qquad (2.10)$$

Now we transform (2.10) as in Example 1 to obtain:

$$\tilde{\xi}_{tt} - \tilde{\xi}_{\tau\tau} + \left(k_x^2 c^2(\tau) + \frac{\eta_{\tau\tau}}{\eta}\right)\tilde{\xi} = \frac{\delta(\tau)\delta(t)}{\eta(0)},$$
$$\tau \equiv \int_0^z \frac{1}{\hat{c}(z')}dz', \quad \eta \equiv (\rho c)^{1/2}, \quad \tilde{\xi} \equiv \eta\tilde{U} \qquad (2.11)$$
$$\tilde{\xi}_\tau + h\tilde{\xi} = 0 \text{ at } \tau = 0, \quad h = -\frac{\eta_\tau(0)}{\eta(0)}$$

For this two dimensional example, by using $F(k_{x,1},t)$ we can recover the potential $k_{x,1}^2 c^2(\tau) + \frac{\eta_{\tau\tau}}{\eta}$; using another wavelength $k_{x,2}$ we can recover $k_{x,2}^2 c^2(\tau) + \frac{\eta_{\tau\tau}}{\eta}$. By using these two potentials (or in general using a fit for several wavelength values) we can recover both $c(\tau)$ and $\eta(\tau)$ and we can convert these profiles to ρ (z) and μ (z).

We can also transform three dimensional problems with point sources to Schroedinger form by applying a Fourier Bessel transform to the wave equation and the observed surface scattering data. Coen [9] has also discussed the use of multi-offset data to recover both the velocity and impedance profiles.

From above, we see that one often has to solve the O.D.E. (2.7) after the inverse scattering problam has been solved. We have shown that the mapping F(impulse response functions):—>(q,h) (potential,boundary condition) is continuous if we measure distance in the (q,h) space by $d((q1,h1),(q2,h2)) = \max_x |(\int_0^x (q_1(x') - q_2(x')) dx')| + |h1-h2|$ and use the L_∞ norm in the space of continuously differentiable impulse response functions. We now prove the following theorem.

Theorem 4.

Let q1(x),q2(x) be continuous functions, 0<x<T, and consider the two solutions of the initial value problems

$$y_1'' - q_1(x) y_1 = 0 \qquad (2.12a)$$

$$y_1(0) = a_1 \qquad y_1'(0) = b_1$$

and

$$y_2'' - q_2(x) y_2 = 0 \qquad (2.12b)$$

$$y_2(0) = a_2 \qquad y_2'(0) = b_2$$

For a fixed positive number ϵ, there exists a "$\delta(\epsilon)$" such that $|| \int_0^x (q_1(x') - q_2(x')) dx' ||_\infty < \delta$, $|b1-b2| < \delta$ and $|a1-a2| < \delta$ implies, that for (q2,a2,b2,y2) the fixed quantities we are perturbing about, $||y1-y2||_\infty < \delta$, and $|| y_1' - y_2' ||_\infty < \delta$.

Proof:

The following proof is based on a method used by Symes [20] for bounding the L_∞ norm of solutions to an O.D.E. of the form (2.12). Firstly, we rewrite (2.12a) as a Volterra integral equation to obtain:

$$y_1' = b_1 + \int_0^x q_1(x')\, y_1(x')\, dx' \qquad (2.13a)$$

$$y_1 = a_1 + b_1 x + \int_0^x (x-x')\, q_1(x')\, y_1(x')\, dx' \qquad (2.13b)$$

Now we let K1(x)= $\int_0^x q_1(x')\, dx'$. Using integration by parts we rewrite (2.13) as:

$$y_1(x) = a_1 + b_1 x - \int_0^x (x-x')\, y_1' K_1(x')\, dx' + \int_0^x y_1(x') k_1(x')\, dx' \qquad (2.14a)$$

$$y_1'(x) = b_1 + K_1(x)\, y_1(x) - \int_0^x K_1(x')\, y_1'(x')\, dx' \qquad (2.14b)$$

We can write similar expressions for y2(x) and y2$'$(x). Subtracting the two expressions we obtain:

$$(y_1-y_2) + \int_0^x (x-x')\, K_1 (y_1-y_2)'\, dx' - \int_0^x K_1 (y_1-y_2)\, dx' \qquad (2.15a)$$
$$= (b_1-b_2)x + \int_0^x (K_1-K_2) y_2 - \int_0^x (k_1-k_2) y_2'(x-x')\, dx' + (a_1-a_2)$$

$$(y_1-y_2)' - K_1(y_1-y_2) + \int_0^x K_1 (y_1-y_2)'\, dx' \qquad (2.15b)$$
$$= (b_1-b_2) + (K_1-K_2) y_2 - \int_0^x (K_1-K_2) y_2'\, dx'$$

Now y2,y2$'$ is the solution of the problem about which we are perturbing,

so we assume that $||y2||_\infty$, $||y2'||_\infty$ are known. By assumption $0<x<T$, $|b1-b2|<\delta$, $||K1-K2||_\infty<\delta$ and $|a1-a2|<\delta$. Clearly, we can make the right hand side of (2.15a) and (2.15b), which we denote as g1(x) and g2(x), small for δ sufficiently small. We now make the substitution $\gamma(x)=(y1-y2)'-K1(y1-y2)$. This then leads to the system of Volterra integral equations:

$$(y_1-y_2)+\int_0^x(x-x')K_1\gamma(x')dx'+\int_0^x K_1(K_1(x-x')-1)(y_1-y_2)=g_1$$

$$\gamma+\int_0^x K_1\gamma dx'+\int_0^x K_1^2(y_1-y_2)=g_2 \tag{2.16}$$

We write this in operator form as

$$\begin{pmatrix} y_1-y_2 \\ \gamma \end{pmatrix}-\begin{bmatrix} F_1 & F_2 \\ F_3 & F_4 \end{bmatrix}\begin{pmatrix} y_1-y_2 \\ \gamma \end{pmatrix}=\begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \tag{2.17}$$

Using the method of Picard iterations we write formally:

$$\begin{pmatrix} y_1-y_2 \\ \gamma \end{pmatrix}=\begin{pmatrix} g_1 \\ g_2 \end{pmatrix}+\begin{bmatrix} F_1 & F_2 \\ F_3 & F_4 \end{bmatrix}\begin{pmatrix} g_1 \\ g_2 \end{pmatrix}+\begin{bmatrix} F_1 & F_2 \\ F_3 & F_4 \end{bmatrix}^2\begin{pmatrix} g_1 \\ g_2 \end{pmatrix}+\cdots \tag{2.18}$$

If we let $\hat{K}=\max\{T||K1||_\infty, ||K1||_\infty+T||K1||_\infty^2, ||K1||_\infty^2\}$, then it is easy to show that the n'th term in the series of (2.18) is bounded by $(2^n x^n/n!\ \hat{K}^n(||g_1||_\infty+||g_2||_\infty))$. Thus the series (2.18) is majorized by the series of $(\ ||g1||_\infty+||g2||_\infty\ )\exp(2\hat{K}T)$. In terms of K_2 we can write $\hat{K}<\max\{T(||K2||_\infty+\delta), ||K_2||_\infty+\delta+T(||K_2||_\infty+\delta)^2, T(||K_2||_\infty+\delta)^2\}$. Thus

$$||y_1-y_2||_\infty \le (||g_1||_\infty+||g_2||_\infty)e^{2||K_2||_\infty T}(1+f(\delta))\equiv\Omega(\delta) \tag{2.19a}$$

$$\Omega(\delta)\to 0 \qquad \delta\to 0$$

$$\|y_1' - y_2'\|_\infty \le \|\gamma\|_\infty + (\|K_2\|_\infty + \delta)\|y_1 - y_2\|_\infty$$
$$\le \Omega(\delta)(1 + \|K_2\|_\infty + \delta) \qquad (2.19b)$$

Hence we can "choose" $\delta(\epsilon)$ so that $\|y1-y2\|_\infty < \epsilon$ and $\|y_1' - y_2'\|_\infty \le \epsilon$ and hence the proof is complete.

Combined with Theorem 2, we have a stability result for the inversion of the observed impulse response function F(t) for the one dimensional impedance profile in example 1, $(\rho(x)c(x))^{1/2}$, where $\rho(x)$ and $c(x)$ have two continuous derivatives. For a fixed $\epsilon > 0$, there exists a "$\delta(\epsilon)$" such that $\|F1(t)-F2(t)\|_\infty < \delta$, for F1(t),F2(t) impulse response functions, and $|\eta_1(0) - \eta_2(0)| < \delta$, implies the impedance profiles are such that $\|\eta_1(x) - \eta_2(x)\|_\infty \le \epsilon$ and $\|\eta_1'(x) - \eta_2'(x)\|_\infty < \epsilon$ over a finite interval $x \in [0,T]$. For the two dimensional example, we can similarly determine $\eta(x)$ in a stable fashion, given the impulse response function F(t,k) for two different wave numbers k1,k2 ($|k1| \neq |k2|$). Similar stability results for this inverse elastic wave equation are given (and derived in a different fashion) by Symes [20]. For the two dimensional SH problem (example 2), we can determine both the impedance profile and the velocity profile from a knowledge of the surface response for two different horizontal wave numbers. However, we do not consider here the stability properties of the mapping $F(k,t) \to (\rho(z), \mu(z))$.

We observe that the comparison of the two impedance profiles is in the

travel times X1= $\int_0^z \frac{1}{\hat{c}_1(z')}\,dz'$, and X2= $\int_0^z \frac{1}{\hat{c}_2(z')}\,dz'$.
Thus, in the physical coordinate, the profiles could look very different if $\hat{c}_1 \neq \hat{c}_2$ and in fact they might not even be defined on the same depth domain.

2.1 Some Practical Problems Arising in Seismological Data.

There are, of course, severe practical limitations to the methods discussed above, for use with exploration seismological data. As discussed in the introduction, it is hard to determine the true impulse response function. The exact frequency characteristics of the source and the receiver are seldom known, and the recorded seismic signal is often corrupted with noise. From our stability results we know, that provided the recorded is not "too far away" from the true impulse response, that we are able to reconstruct accurately the impedance profile. However, for large errors, it is not even clear that the recorded signal is in fact an impulse response function for any continuous profiles. For bandlimited sources and receivers, required spectral information is lost.

In theory, it seems possible, although perhaps difficult, to fit a non-harmonic series of the form

$$F(t) = \sum_{j=L}^{N} \left(\frac{1}{\rho_j} \cos\sqrt{\lambda_j}\,t - \frac{(2-\delta_{0,j})}{T} \cos\left(\frac{j\pi t}{T}\right) \right) \qquad (2.20)$$

to the signal F(t) for0<t<2T, thus extracting a finite amount of spectral data within the possible bandwidth. By assuming the form of the unknown spectral data (e.g., take $\lambda_k = k^2\pi^2/T^2$, $\rho_k = 2/T$ for k>N,k<L), we construct a spectral data set which is known to correspond to some potential. If

because of finite bandwidth we have lost high frequency information in F(t), one would hope that this might lead to only small errors in the signal, and hence the calculated impedance profile would be accurate. The loss of low frequency information, as illustrated by Example 1 (2.53) can be disastrous. For F(t)=H, a D.C. filter would destroy the signal and the reconstructed potential would be q(x)=0 instead of the correct potential q(x)=2/(x+1/H)**2.

Also, in reality the parameter profiles of the Earth vary three dimensionally. Even if a vertically stratified model is a good approximation to the geological section of interest, it is not known how much the three-dimensional properties will affect the inversion scheme. Also, the Gelfand-Levitan method is meant for twice differentiable profiles and it is not clear how discontinuities in the medium will affect the inversion method. However, modifications to the continuous theory, and discrete inverse scattering methods allow for the inclusion of jump discontinuities in the profiles. The paper by Bube and Burridge [3] discusses in detail various other numerical inversion schemes that can be used for impedance profile inversion.

Section 3. Numerical Examples.

To numerically invert the impulse response function F(t) for the potential q(x), we used Syme's [19] second order discretization of the non-linear hyperbolic problem (1.33). The inversion program was written in single precision Fortran and run on an IBM 370 and a VAX-780. The numerical scheme was checked with analytic answers from (1.50). We calculated our answers with various grid sizes to make sure that the discretization had stabilized.

Example 1.

We consider the three dimensional wave equation:

$$\nabla^2 P(x,y,z,t) - \frac{\partial^2 P}{\partial t^2} = \frac{\delta(t)\delta(z)\delta(r)}{2\pi r}$$

$$\frac{\partial P}{\partial z}(z=0,x,y,t) = 0 \qquad r \equiv \sqrt{x^2+y^2} \tag{3.1}$$

If we take the Fourier Bessel Transform of (3.1) we obtain:

$$\frac{\partial^2 \tilde{P}}{\partial z^2} - \frac{\partial^2 \tilde{P}}{\partial t^2} - p^2 \tilde{P} = \frac{\delta(t)\delta(z)}{2\pi} \tag{3.2}$$

The causal solution of (3.1) is:

$$P(z,r,t) = -\frac{1}{2\pi} \frac{\delta(t-R)}{R} H(t) \tag{3.3a}$$

or

$$P(z=0,r,t) = \frac{-1}{2\pi}\frac{\delta(t-r)}{r}H(t) \qquad (3.3b)$$

Here $R=\sqrt{r^2+z^2}$ and H(t) is Heaviside's function. The Bessel transform of the surface data is:

$$\gamma(p,t) = \int_0^\infty r J_0(pr)\left\{\frac{-1}{2\pi}\frac{\delta(t-r)}{r}H(t)\right\}dr \qquad (3.4)$$
$$= -\frac{1}{2\pi}J_0(pt)H(t)$$

Differentiating with respect to time:

$$\frac{\partial\gamma(p,t)}{\partial t} = \frac{1}{2\pi}pJ_1(pt) - \frac{1}{2\pi}J_0(pt)\delta(t) \qquad (3.5)$$

The regular part of the Riemann function, K(x,t) is

$$-\frac{\partial}{\partial t}\int_0^\infty J_0(pr)r\frac{\delta(t-R)}{R}dr = \frac{-t}{\sqrt{t^2-z^2}}pJ_1(p\sqrt{t^2-z^2}) \qquad (3.6)$$

Thus we take $F(t)=-pJ_1(pt)$. According to theory, the inversion should produce a potential of p**2 for this case. The results for three inversions are shown below.

Table 1. $F(t)=-.8J_1(.8t)$ "Δ"=.05

x	q(x)
0.	.6402560
.1	.6400014
.2	.6400035
.3	.6400082
.4	.6400147
.5	.6400195
.6	.6400245
.7	.6400359
.8	.6400475
.9	.6400579
1.0	.6400734

Table 2. $F(t)=-J_1(t)$ "Δ"=.025

x	q(x)
0.	1.000156
.1	1.000002
.2	1.000005
.3	1.000010
.4	1.000030
.5	1.000036
.6	1.000050
.7	1.000081
.8	1.000119
.9	1.000155
1.0	1.000210

Table 3. $F(t)=-10\ J_1(10t)$ "Δ"=.005

x	q(x)
0.0	100.0625
.05	100.0065
.10	100.0262
.15	100.0704
.20	100.1644
.25	100.3588
.30	100.7687
.35	101.6228
.40	103.4054
.45	107.1018
.50	114.7560

The results for p=10 at x=.5 are: "Δ"=.01 , q(.5)=167.4181 and for "Δ"=.02, q(.5)=-17248.02. It is clear from (3.5) and (3.6) that, as "p" increases, smaller and smaller step sizes are needed to properly resolve the oscillations of F(t) and the variation of K(x,t). We note that for this example, the inversions for different values of "p" allow us to properly deduce that η_{xx}/η =0 and c(z)=1.

Example 2.

We now consider the interval $0<z<\pi$ and $0<t<2\pi$. Now the solution of the problem

$$P_{zz} - P_{tt} - p^2 P = \delta(t)\delta(z)$$
$$\frac{\partial P}{\partial z} = 0 \qquad z = 0, \pi \tag{3.7}$$

is

$$P(z,t) = -\sum_{n=0}^{\infty} \frac{\cos nz \sin\sqrt{n^2+p^2}\,t}{\sqrt{n^2+p^2}} \frac{(2-\delta_{0,n})}{\pi} H(t) \tag{3.8}$$

Thus $F(t) = \left(\sum_{n=1}^{\infty} \frac{2}{\pi}(\cos\sqrt{n^2+p^2}\,t - \cos nt)\right) + \frac{1}{\pi}(\cos pt - 1)$. Now from causality the boundary z= π has no effect upon F(t) until $t=2\pi$. Thus for $0<t<2\pi$

$$-J_1(pt) = \sum_{n=0}^{\infty} \frac{(2-\delta_{0,n})}{\pi}(\cos\sqrt{n^2+p^2}\,t - \cos nt) \tag{3.9}$$

Instead of an infinite sum we can consider:

$$F_N(t) = \sum_{n=0}^{N} \frac{(2-\delta_{0,n})}{\pi} \left(\cos\sqrt{n^2+p^2}\,t - \cos nt\right) \qquad (3.10)$$

This corresonds to taking the zero potential (h=0) spectral information (λ_j, ρ_j) for j>N. These approximations will be inaccurate chiefly in the (N+1) frequency component; hence as figures 1.2-1.4 indicate, the approximation oscillates about $-pJ_1(pt)$ with roughly the period $2\pi/\sqrt{\lambda_{N+1}}$. In figure 1.2 we show $F_{10}(t)$ for $0<t<2\pi$; in figure 1.3, $F_{50}(t)$ $0<t<2\pi$, and in figure 1.4 we show $F_{50}(t)$ for $0<t<1$. We now calculate the potentials resulting from F10(t) and F50(t) ("Δ"=.01,"Δ"=.005). These potentials, along with the calculated potential from $F(t)=-J_1(t)$, are shown in figures 1.5 and 1.6. We then used the potentials of figure 1.5, q10(x) and q(x)=1 and solved the initial value problem

$$\eta_{zz} - q(z)\eta = 0$$

$$\eta(0) = 1 \;;\; \eta'(0) = -F(0) \qquad (3.11)$$

Thus, we are considering instead of the three dimensional problem, where $\eta_{zz}/\eta = 0$, the one dimensional problem $\frac{\partial^2 v}{\partial z^2} - \frac{\partial^2 v}{\partial t^2} - \frac{\eta_{zz}}{\eta} v = \delta(t)\delta(z)$. Figure 1.7 shows these impedance profiles obtained from (3.11). They are, within plotting resolution, identical. These results are in agreement with Theorems 3 and 5. The theoretical answer to (3.11) for the correct potential $q(z)=1$ is $\eta(z)=\cosh(z)$.

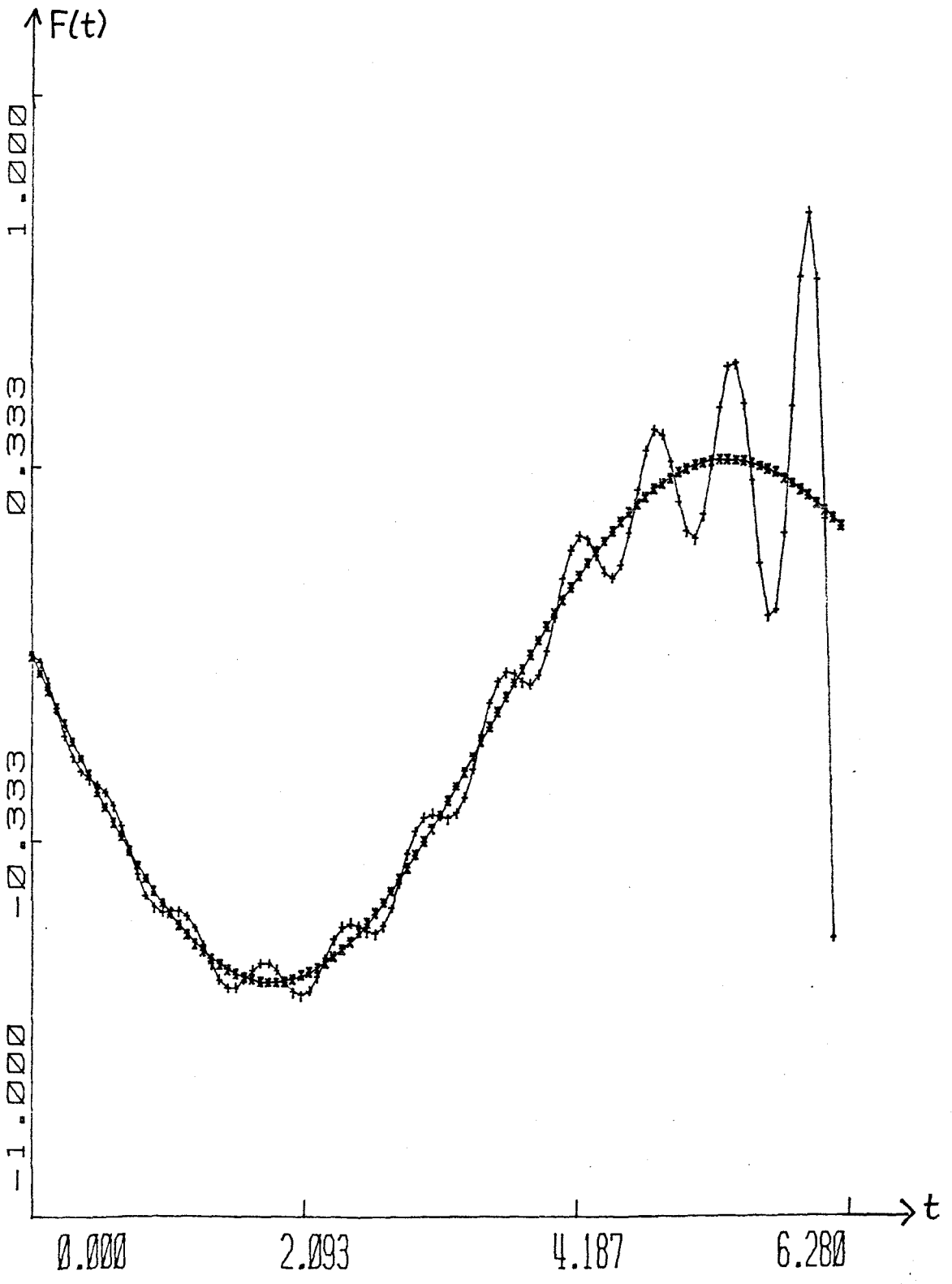


Figure 1.2

N = 10 Approximation to $-J_1(t)$ $0 \le t \le 2\pi$

"*" - - $J_1(t)$ "+" - N = 10 series

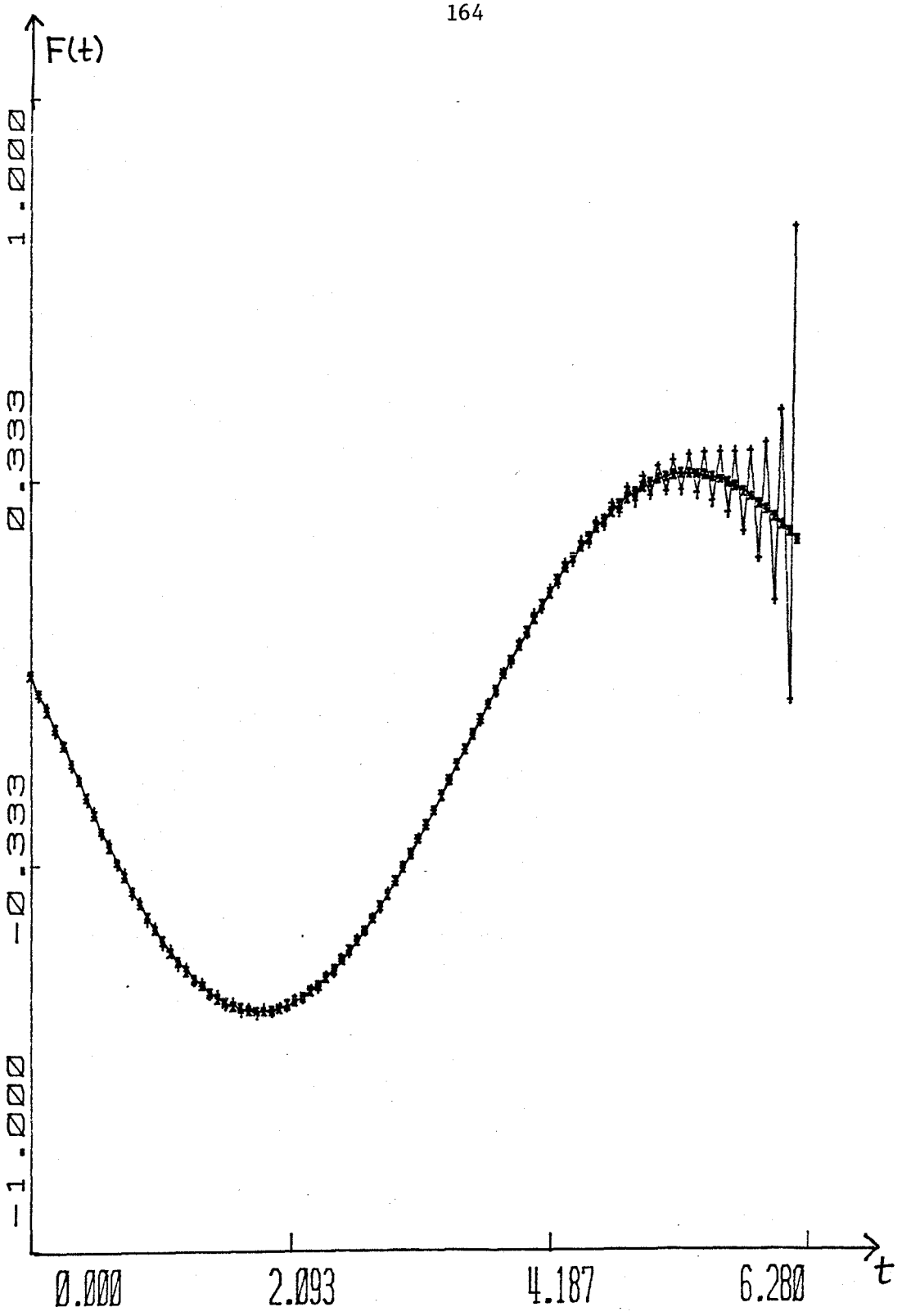


Figure 1.3

N = 50 Approximation to $-J_1(t)$ $0 \leq t \leq 2\pi$

"*" - - $J_1(t)$ "+" - N = 50 series

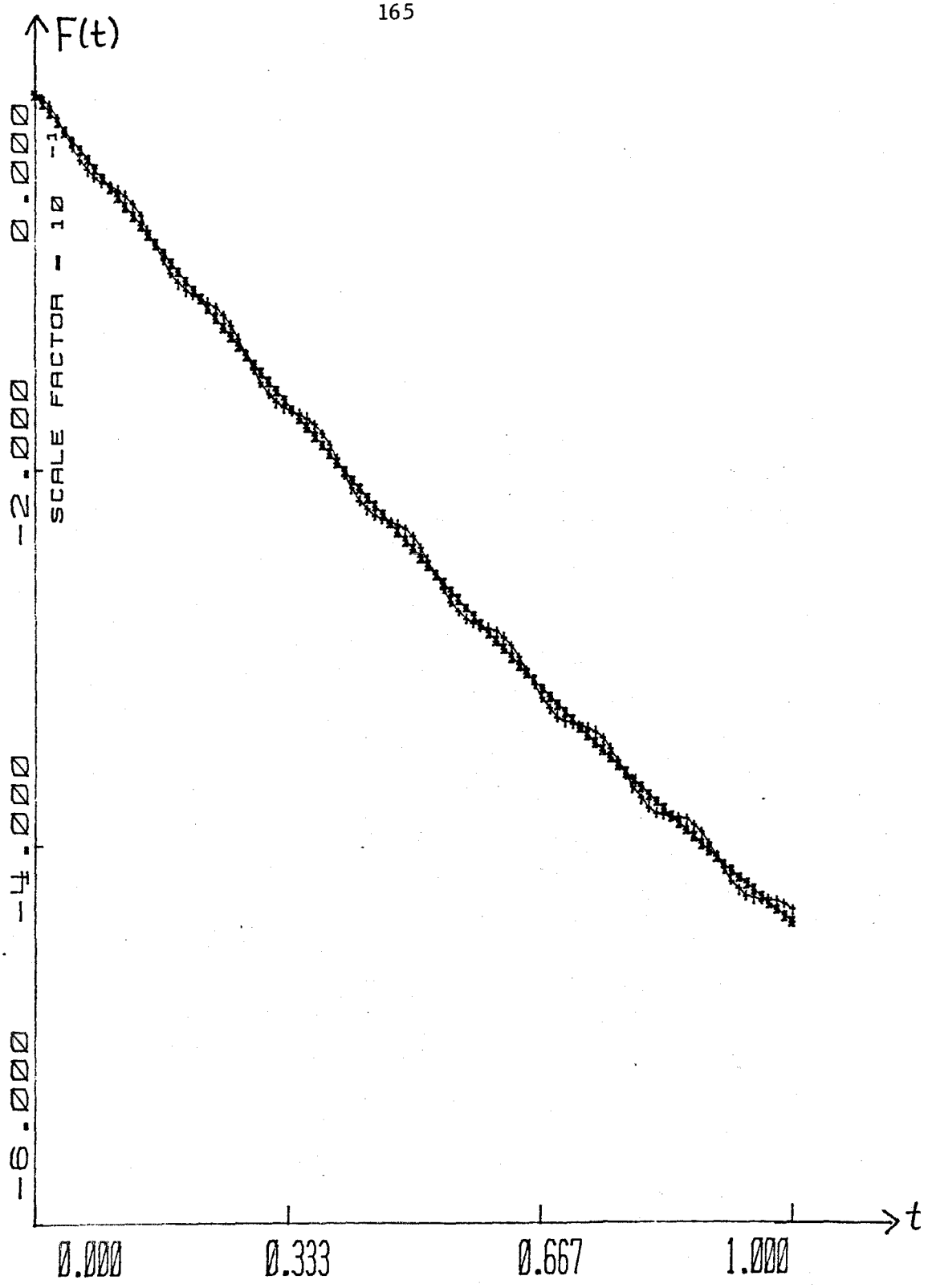


Figure 1.4

N = 50 Approximation to $-J_1(t)$ $0 \leq t \leq 1$

"*" - - $J_1(t)$ "+" - N = 50 series

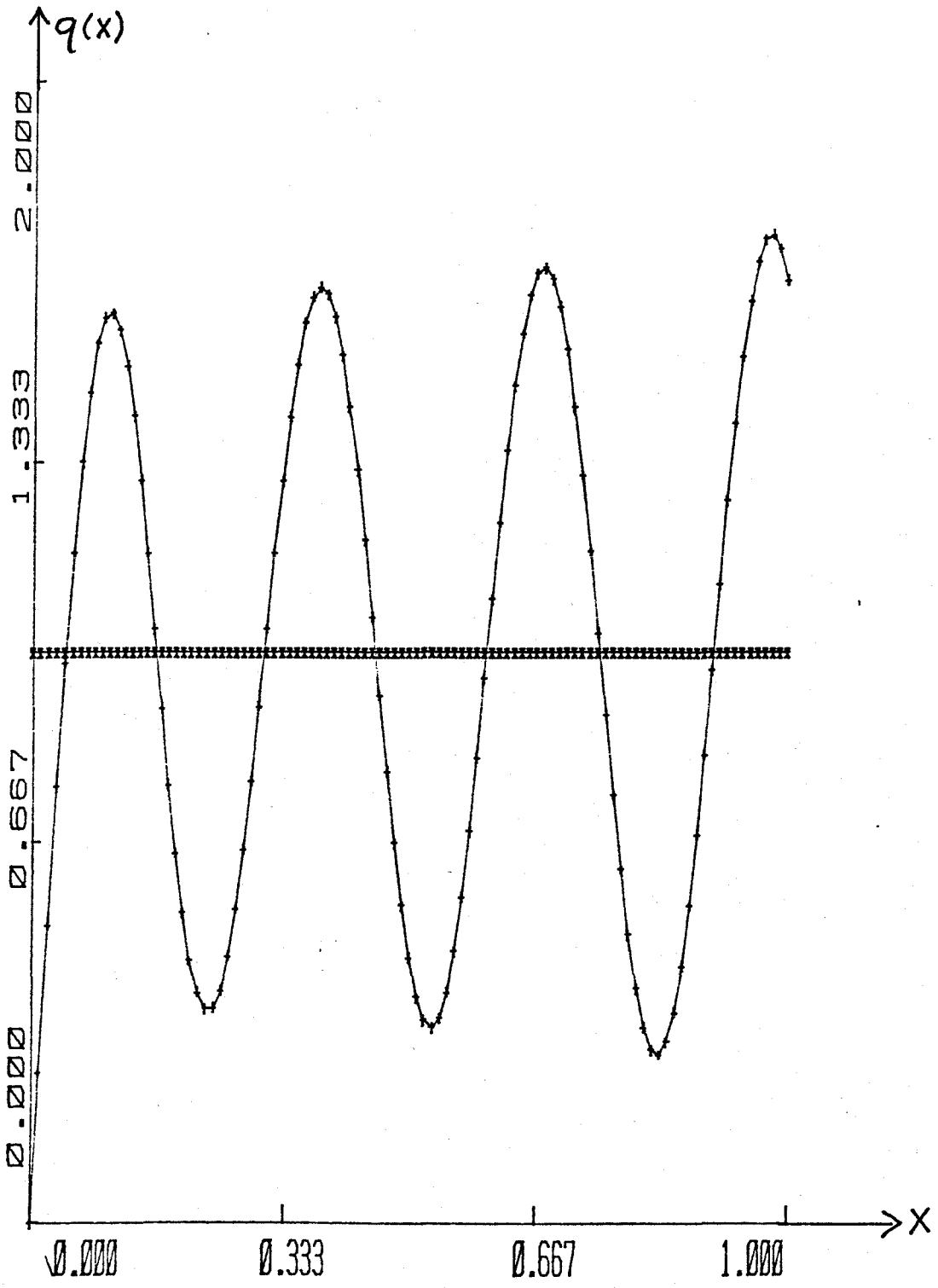


Figure 1.5

Reconstructed potential, N = 10, Δ = .01

"*" - true potential "+" - N = 10 approximation

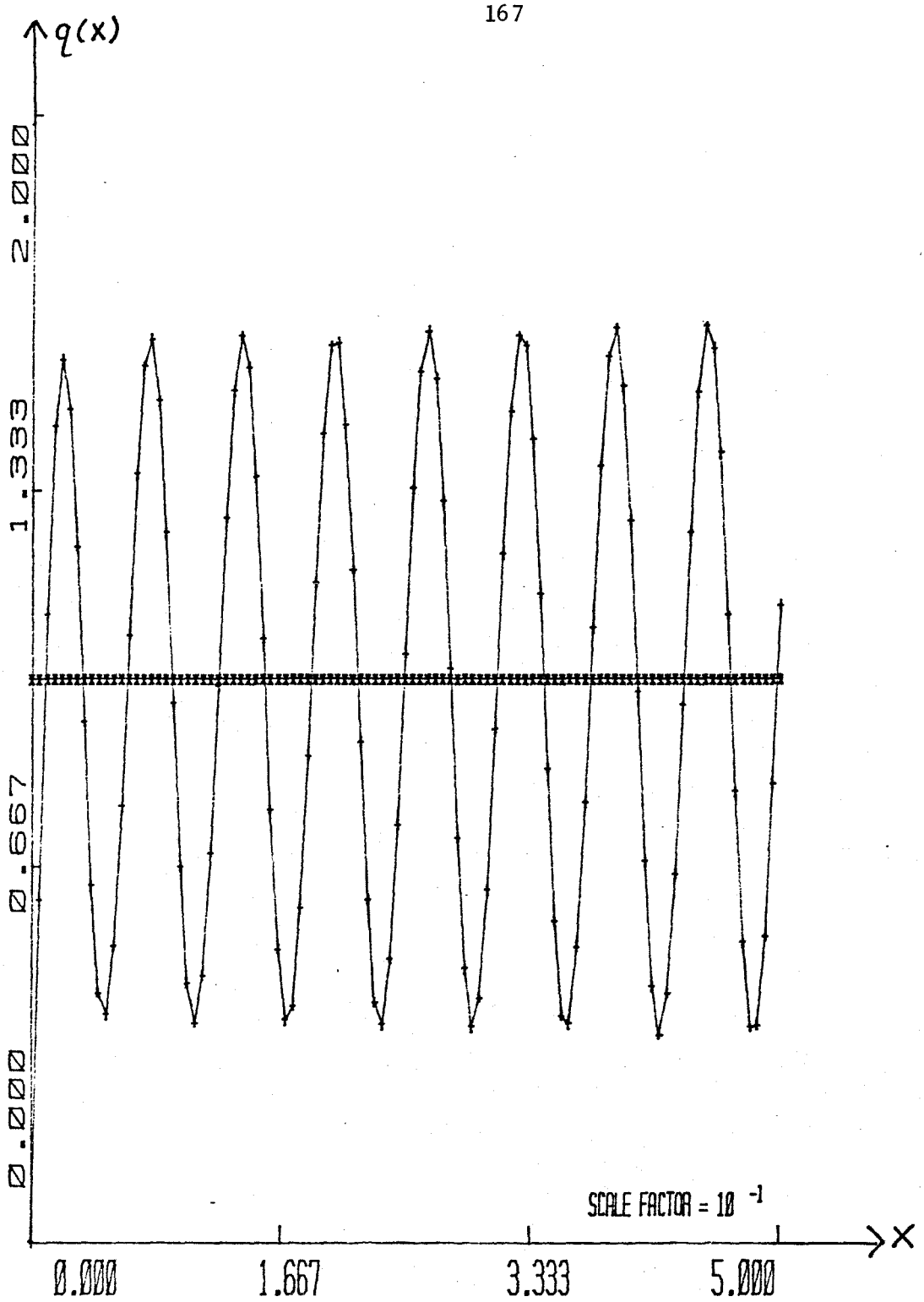


Figure 1.6

Reconstructed potential, N = 50, Δ = .005

"*" - true potential "+" - N = 50, approximation

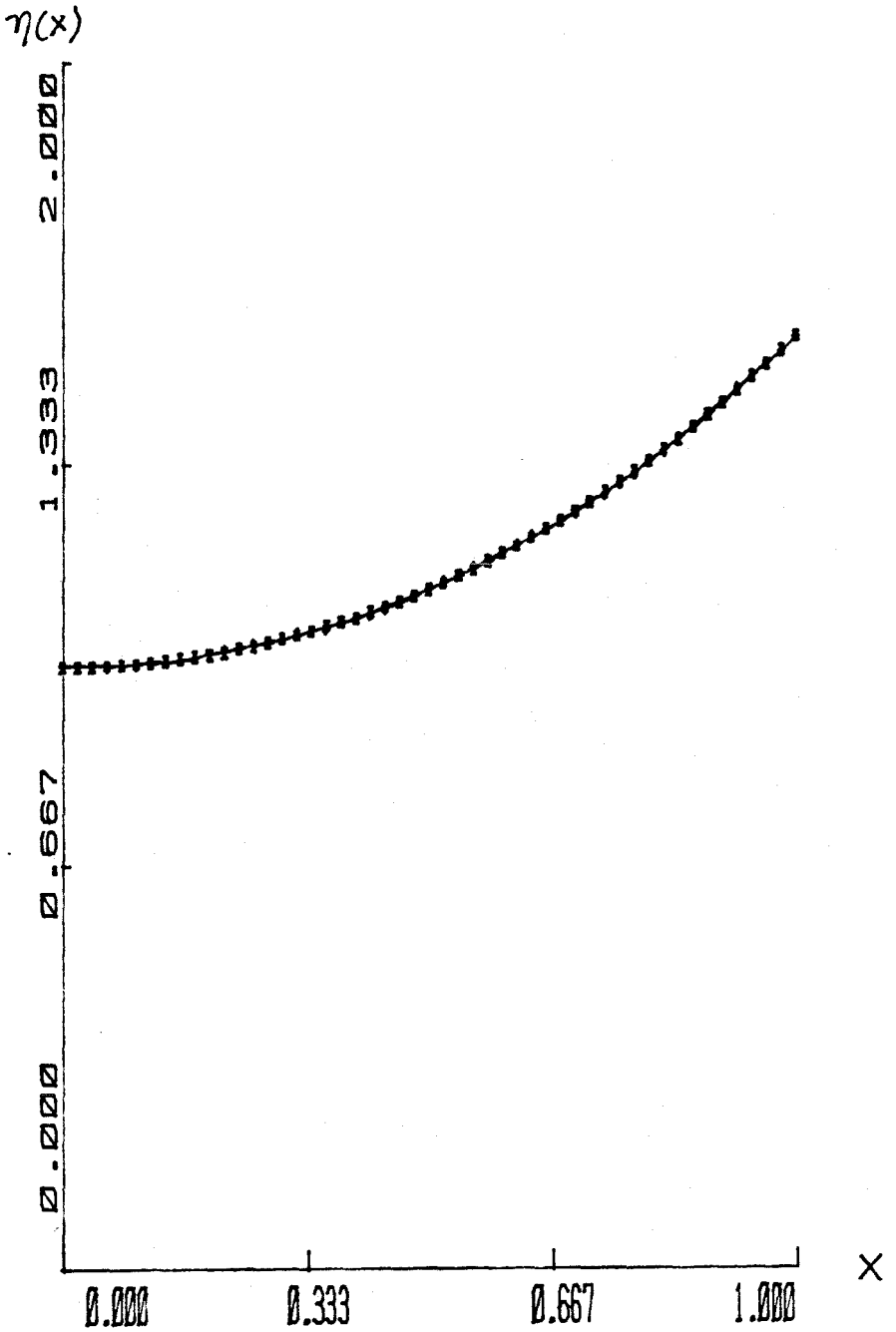


Figure 1.7

Reconstructed Impedance Profile N = 0

"*" - true profile "+" - N = 10 approximation

Example 3.

We now consider the Mathieu Equation

$$\frac{\partial^2 P}{\partial t^2} - \frac{\partial^2 P}{\partial z^2} + 2\cos^2 z\, P = \delta(t)\delta(z) \qquad (3.12)$$

$$\frac{\partial P}{\partial z} = 0 \qquad z = 0, \pi$$

The first eight eigenvalues and normalizations of the associated equation $-\frac{d^2y}{dz^2} + (2\cos^2 z)y = \lambda y$, $\frac{dy}{dz} = 0$ $z=0,\pi$ are from [22].

k	λ_k	ρ_k
0	.87823	5.54317
1	2.46676	1.80500
2	5.10090	1.38777
3	10.01761	1.47255
4	17.00836	1.51848
5	26.00521	1.53814
6	37.00357	1.54841
7	50.00260	1.55446

Table 4. Spectral Information for (3.12)

We now form $F7(t)=\sum_{n=1}^{7}\left(\frac{1}{\rho_n}\cos\sqrt{\lambda_n}t - \frac{2}{\pi}\cos nt\right)+\left(\frac{1}{\rho_0}\cos\sqrt{\lambda_0}t - \frac{1}{\pi}\right)$. The corresponding potentials q7(x) and 2 cos (x) are shown in figure 1.8. We now calculate the corresponding impedances with $\eta(0)=1$ and $\eta'(0)=-F(0)$ $(0< x < 1)$ The two resulting impedance profiles are shown in figure 1.9. We see that once again the tw impedance profiles are extremely close. If we let $2\cos^2(x) = 1 + \cos(2(x))$, the eigenvalues of Table 4 are shifted by -1 and the normalizations are now appropriate for the equation:

$$\frac{\partial^2 P}{\partial t^2} - \frac{\partial^2 P}{\partial z^2} + (\cos(2z))P = \delta(t)\delta(z) \qquad (3.13)$$

$$\frac{\partial P}{\partial z} = 0 \qquad z = 0, \pi$$

We note that the first eigenvalue is negative so that

$$F_7(t) = \sum_{n=1}^{7}\left(\frac{\cos\sqrt{\lambda_n}t}{\rho_n} - \frac{2}{\pi}\cos nt\right) + \left(\frac{\cosh\sqrt{-\lambda_0}t}{\rho_0} - \frac{1}{\pi}\right)$$

The calculated potential and cos(2(x)) are shown in figure 1.10. We have for x>.2 done a better job here than in figure 1.8. in reconstrucitng the approximate potential. This is perhaps because of the asymptotic behaviour of Sturm-Liouville spectra (see Gasymov and Levitan [10]). For $q(x)=2\cos^2 z$, $\int_0^\pi q(z')\,dz' \neq 0$ and we have that:

$$\sqrt{\lambda_k} \sim k + \frac{b_1}{k} + O\left(\frac{1}{k^3}\right) \qquad (3.14a)$$

$$\rho_k \sim \pi/2 + \frac{c_1}{k^2} + O\left(\frac{1}{k^4}\right) \qquad b_1, c_1 \neq 0 \qquad (3.14b)$$

We can now see that the series for F(t) converges very slowly, by writing

$$F(t) = F_7(t) + \sum_{k=N+1}^{\infty}\frac{2}{\pi}\left(\cos\left(k+\frac{b_1}{k}\right)t - \cos kt\right) + \sum_{k=N+1} O\left(\frac{1}{k^2}\right)$$

$$= F_7(t) + \frac{2}{\pi}b_1 t\sum_{k=N+1}^{\infty}\frac{\sin kt}{k} + \sum_{N+1} O\left(\frac{1}{k^2}\right) \qquad (3.15)$$

$$\sum_{k=1}^{\infty}\frac{\sin kt}{k} = \frac{(\pi - t)}{2}$$

Thus F(t) converges like $\sum_{k=1}^{\infty}\frac{\sin kt}{k}$, or, in other words, a finite sum approximation is not very good. The derivatives of the finite sum do not neccessarily correspond to the derivatives of the true impulse response

function. On the other hand, for q(x)=cos(2(x)) $\int_0^\pi q(z')\,dz' = 0$ (h1,h2=0) and $\sqrt{\lambda_k} = K + O(1/k^{**}3)$, so that the resulting finite sum approximations to the impulse reponse function converge absolutely and much more rapidly. The derivative of these sums also converges to the derivative of the impulse response function.

Now we show the results of the loss of low frequency spectral information. We set the lowest eigenvalue $\sqrt{\lambda_0} = 0$ and $\rho_0 = \pi$. The resulting potential and cos(2(x)) are shown in figure 1.11 and the 2 resulting impedance profiles in figure 1.12. As one might expect, the loss of low frequency information, affects the mean level of the calculated impedance profile.

Example 4.

In this example, we consider the formula (1.55). We take $\sqrt{\lambda_0} = .93714$, and $\rho_0 = 5.543$. This is the first eigenvalue and weight from Table 4. As might be expected from the closeness of $\sqrt{\lambda_0}$ and $\sqrt{\lambda_1} = 1$, the potential is ill-behaved. We show the results of the numeric inversion (step size Δ = .028) and the analytic answer below in Table 5.

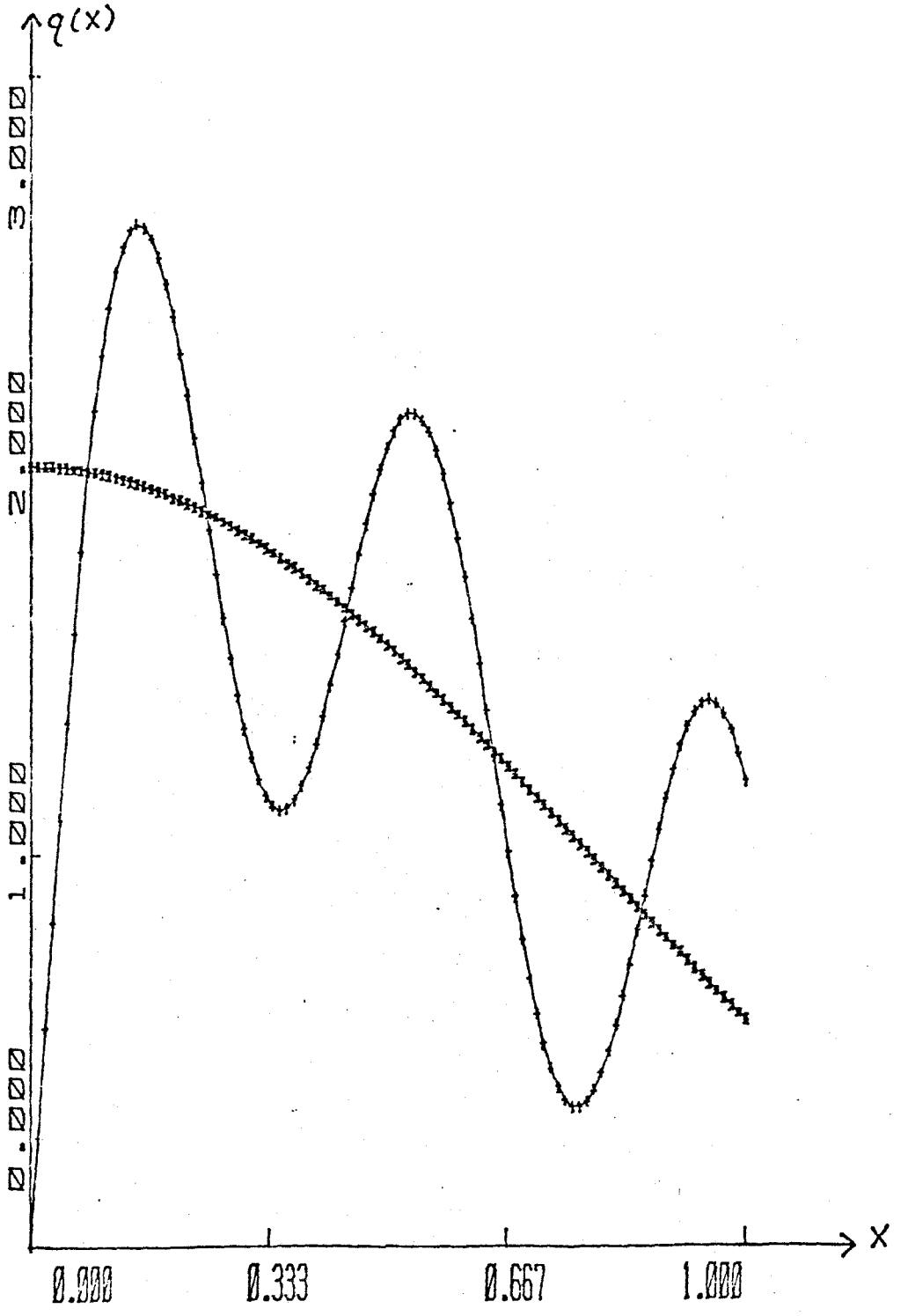


Figure 1.8

Reconstructed Potential - Mathieu Equation

Δ = .01, "*" - $2\cos^2 x$, "+" = N = 7 approximation

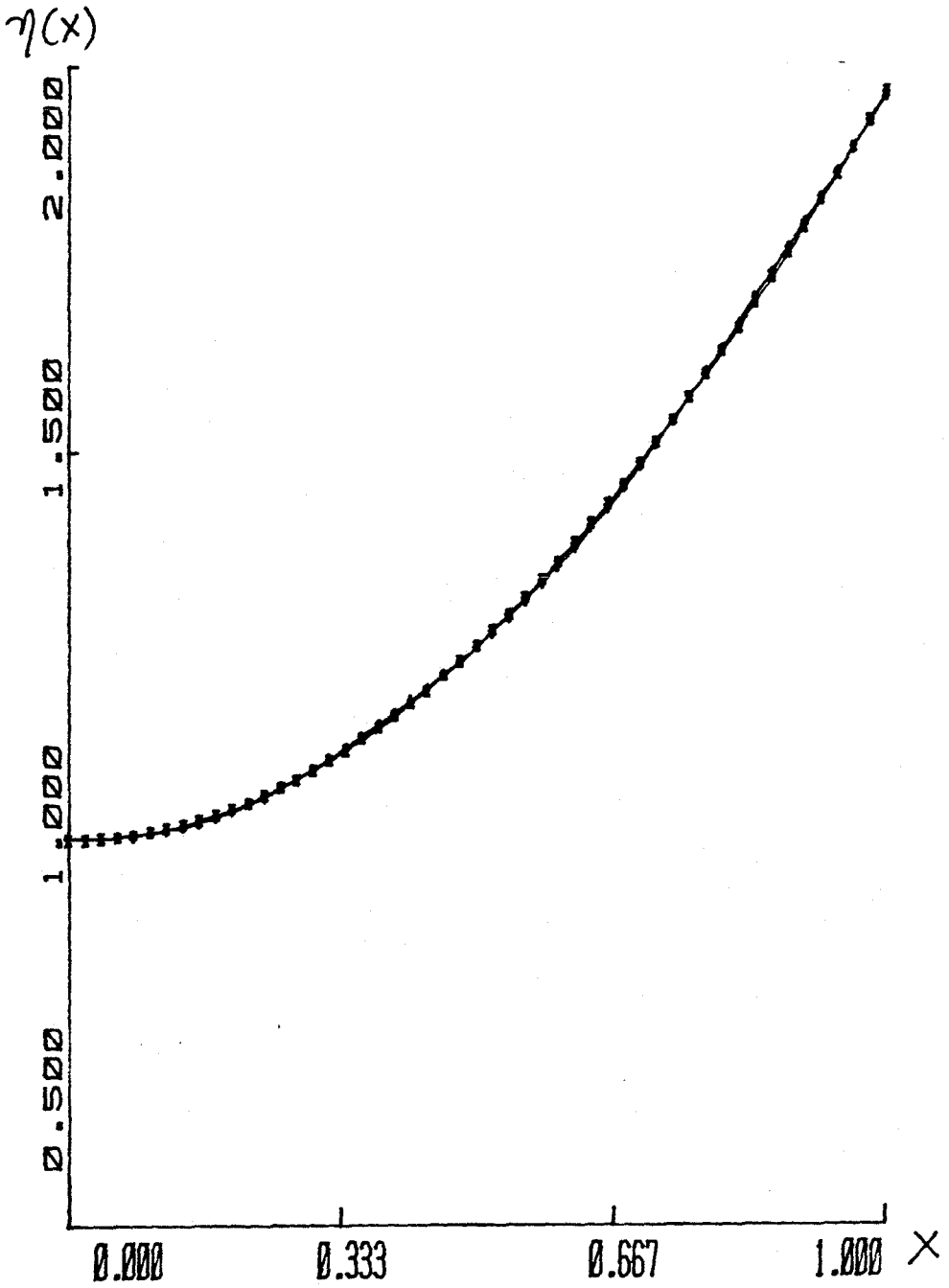


Figure 1.9

Reconstructed Impedance Profiles

"*" - true profile "+" - reconstructed profile

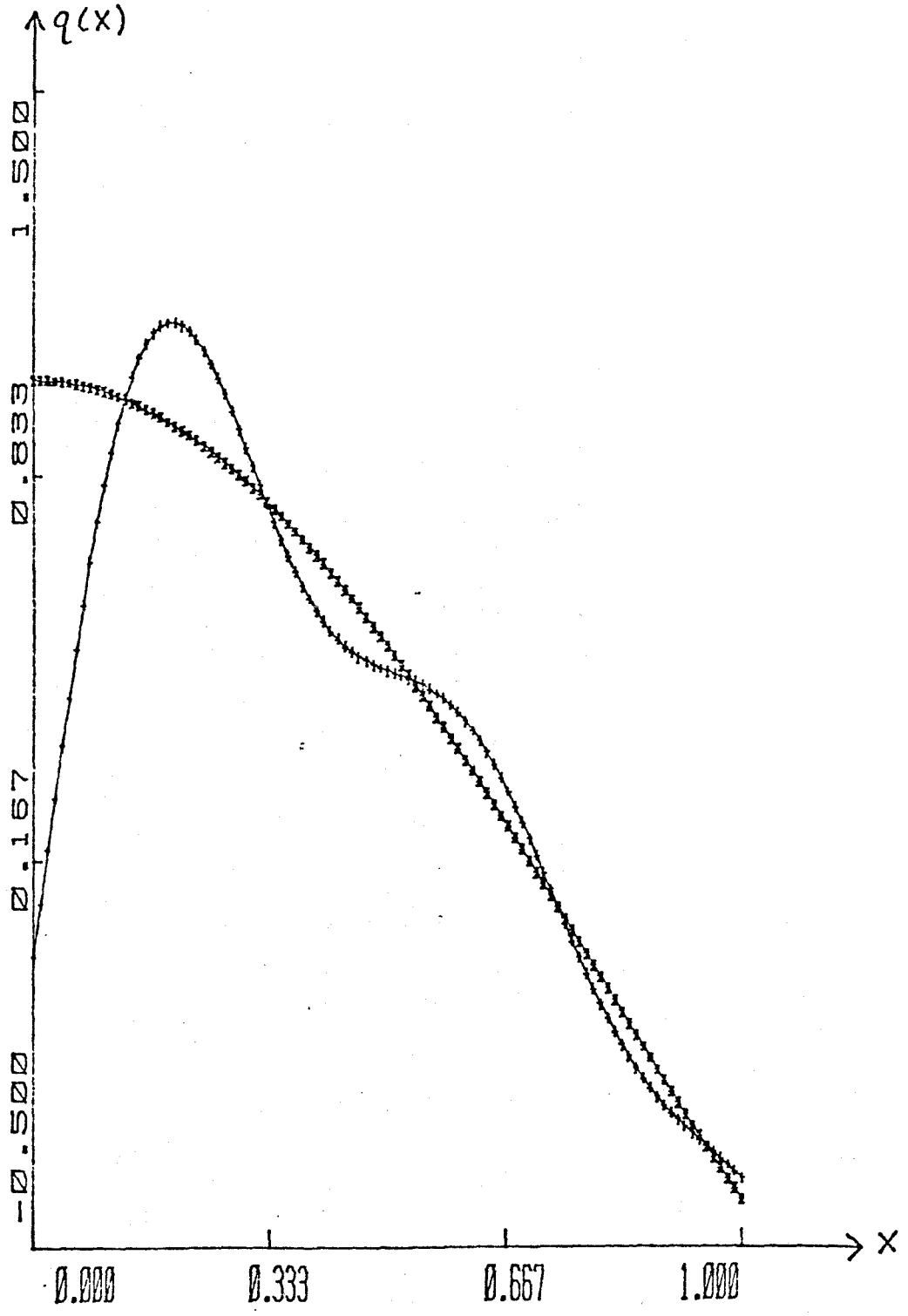


Figure 1.10

Reconstructed Potential - Mathieu Equation

Δ = .01, "*" - cos 2x "+" - N = 7 approximation

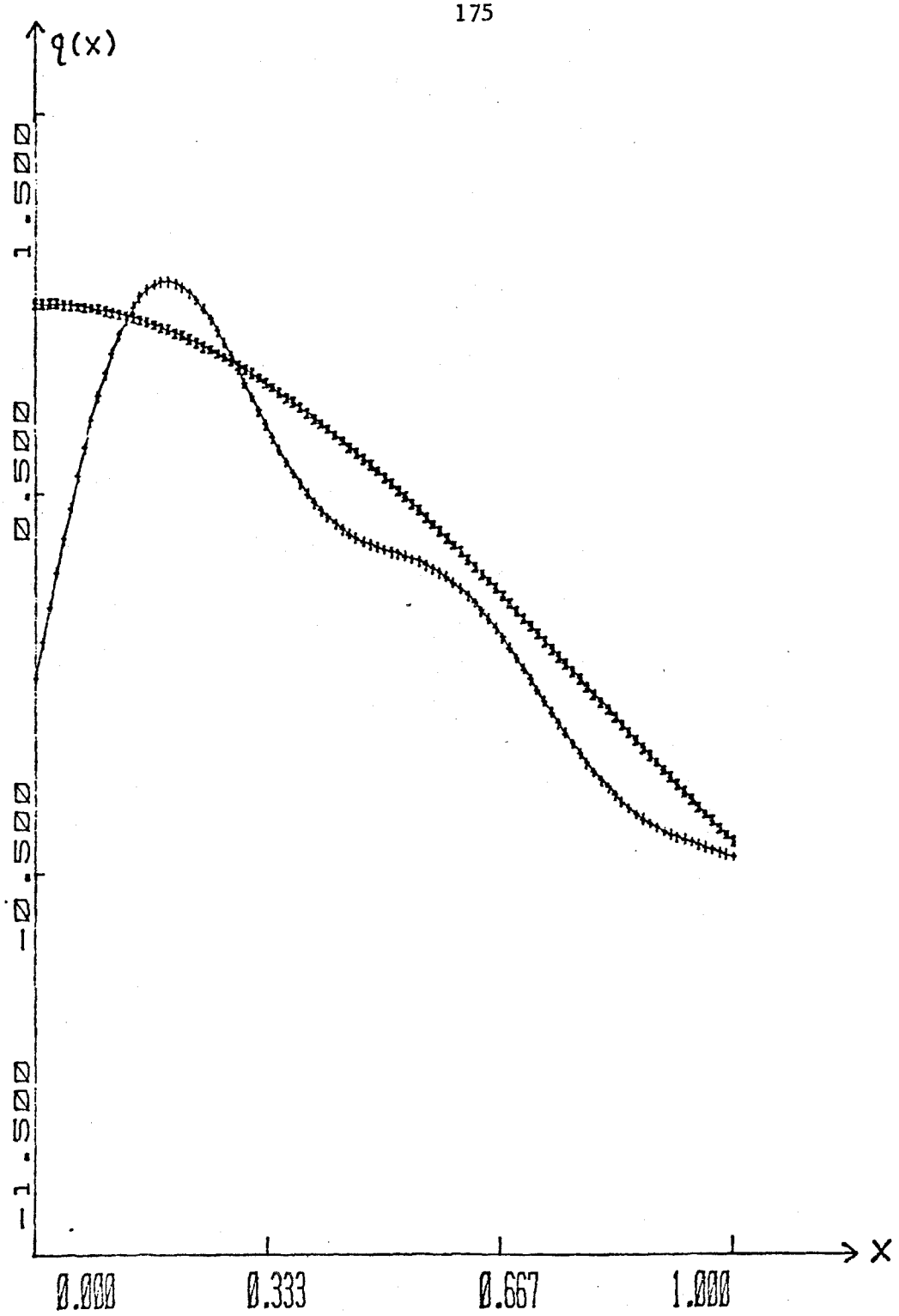


Figure 1.11

Reconstructed Potential - Loss of Low Frequency

Δ = .01, "*" - cos 2x "+" - approximation

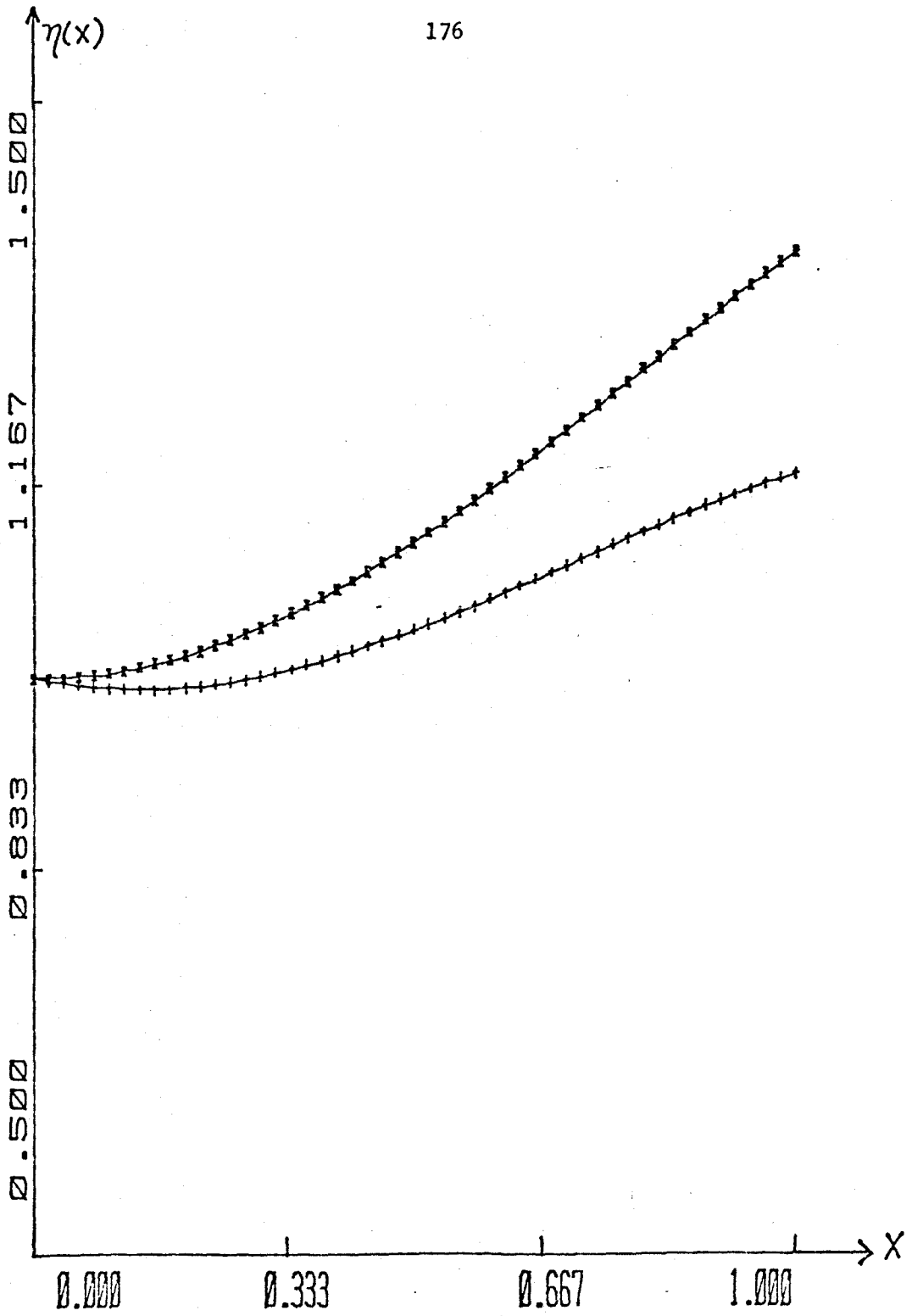


Figure 1.12

Reconstructed Impedance Profile
"*" - true profile "+" - approximation

x	q(x) numeric	q(x) from (1.51)
0.0000	.038038	.038037
.310	.245775	.245827
.620	.443657	.443753
.930	.617327	.617430
1.240	.784602	.784701
1.550	.998352	.998352
1.860	1.373961	1.373796
2.170	2.206190	2.205336
2.480	4.560701	4.555356
2.790	15.76951	15.68920
2.987	80.22237	77.94135

Table 5. Comparison of Numeric and Analytic Calculations.

Chapter 2 Tomographical Reconstruction of Velocity Anomalies.

Introduction (1.0).

Tomography refers to the techniques of reconstructing a field $\rho(\underline{x})$ from a knowledge of $\hat{\rho}(\underline{\xi})$, where $\hat{\rho}(\underline{\xi})$ are the projections (line or surface integrals) of $\rho(\underline{x})$ on surfaces or curves which are parametrized by the coordinates $\underline{\xi}$. For example, in medical x-ray tomography, one can deduce from x-ray attenuation factors the function $P(\theta,t)$ which is equal to $\int_{AB} f(x,y)\,dl$ where the line AB is a distance t from the origin and has slope $dx/dy = -\cot\theta$. The function f(x,y) could be, for example, the density field of the patient's tissue. $P(\theta,t)$ is called the Radon Transform of f(x,y). An inverse transform exists so that from a knowledge of P(0,t) one can reconstruct f(x,y). The Radon Transform and its inverse are extendable to higher dimensions.

In seismology, the field of interest is the "slowness" field, $n(x,y)=1/V(x,y)$ where V(x,y) is the velocity field (compressional or shear). The travel times of disturbances from various sources to various receivers are the "projections" of the slowness field along the rays joining the sources to the receivers. Here, the problem is complicated by the fact that the rays themselves depend upon the unknown velocity field. Thus in general this problem is non-linear. We will linearize the problem by considering the slowness field to be a perturbation from a known background field.

Section 1. Theory

1.1 Linearization of the Inverse Problem.

As suggested above let us suppose that

$$n(x,y)=n_0(x,y)+ n_1(x,y) \qquad (1.1)$$
$$|n_1(x,y)|<<|n_0(x,y)| \quad \forall (x,y)$$

We will take n (x,y) to be piecewise smooth. Similarly, we write for the travel time:

$$\tau(\underline{X}_S, \underline{X}_R) = \tau_0(\underline{X}_S, \underline{X}_R) + \tau_1(\underline{X}_S, \underline{X}_R) \qquad (1.2)$$

where $\tau_0(\underline{X}_S, \underline{X}_R)$ is the travel time from $\underline{X}_S$ to $\underline{X}_R$ for the slowness field $n_0(x,y)$. $\tau_1(\underline{X}_S, \underline{X}_R)$ is the resulting perturbation of the travel time. If we consider $\tau(\underline{X}_S, \underline{X}_R)$ as a function of the field n(x,y) we write formally the first order expansion of τ about $n(x,y)=n_0(x,y)$:

$$\tau(X_S, X_R, n_0(x,y)+\Delta n(x,y)) \simeq \tau_0 + \frac{\delta \tau}{\delta \underset{\sim}{X}} \delta \underset{\sim}{X}(\Delta n) + \frac{\delta \tau}{\delta n} \Delta n \qquad (1.3)$$

where $\underset{\sim}{X}$ represents the ray path joining $\underline{X}_S$ to $\underline{X}_R$ (for a continuous medium we think of $\underline{X}$ as a functional, and for a discretized medium (i.e., the Earth divided into cells) we think of $\underline{X}$ as a vector). From Fermat's principle $\delta\tau / \delta \underset{\sim}{X} = 0$ (this holds in both continuous and discontinuous media). Thus from (1.3)

$$\tau_i(\underline{X}_s, \underline{X}_R) = \frac{\delta\tau}{\delta n}\Delta n = \int_{\text{ray for } n_0(x,y)} n_1(x,y)\, ds \qquad (1.4)$$

Thus the principle which is fundamental to linearized travel time analysis is: the perturbation of the travel time between $\underline{X}_s$ and $\underline{X}_R$ is to first order the projection of the pertubation of the slowness field on the unperturbed ray joining $\underline{X}_s$ and $\underline{X}_R$.

1.2 Formulation and Solution of Linearized Inverse Problem.

The analysis of travel time anomalies has often been formulated as an optimization problem. Let us consider the geometry shown below in figure 2.1.

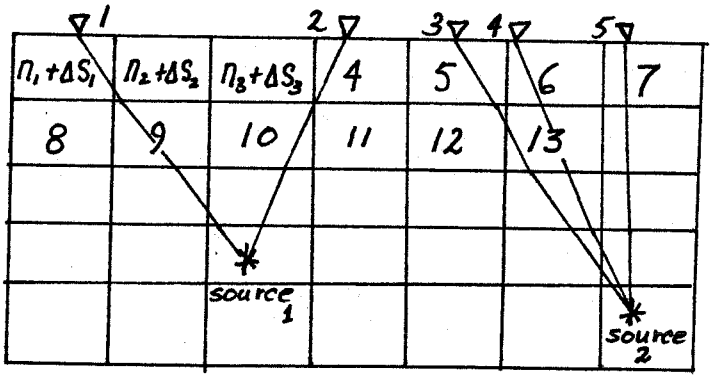


Figure 2.1 Discretized Model of Section of Earth

We consider that from various earthquake sources (source i) we have determined travel time perturbations ti. That is, we have determined the differences of the observed travel times from the theoretical values for an assumed velocity model of the Earth. To model the earth, the earth is discretized into cells that we enumerate. The perturbations of the

slowness field in the cells are denoted by $\Delta S_1, \Delta S_2, \Delta S_3, \Delta S_4, \ldots, \Delta S_N$. The rays are traced from source to receiver using the assumed background velocity field. We denote the length of ray i in cell j as L_{ij}. In matrix formulation we wish to solve the problem

$$\underset{M\times N}{\underline{\underline{L}}} \; \underset{N\times 1}{\Delta \underline{S}} = \underset{M\times 1}{\Delta \underline{t}} \tag{1.5}$$

or for $M \neq N$, the least squares problem

$$\min_{\Delta \underline{S}} \left\| \underline{\underline{L}} \, \Delta \underline{S} - \Delta \underline{t} \right\|_2 \tag{1.6}$$

The methods discussed in Chapter 2, Part I are now applicable to (1.6).

However, the problem for the determination of velocity anomalies from travel time data can be formulated more analytically. For the case, where the velocity field is an increasing function of depth, we do not always need the above assumptions and linearizations. The Wiechert-Herglotz inversion formula (see Aki and Richards [2]) is applicable to this problem. However, we consider the problem for V=V(x,y) and we consider the following reflection geometry. We will denote the background velocity field as C(y) (i.e., we take the background field to vary only with depth) and the basic reflection geometry is shown below in figure 2.2.

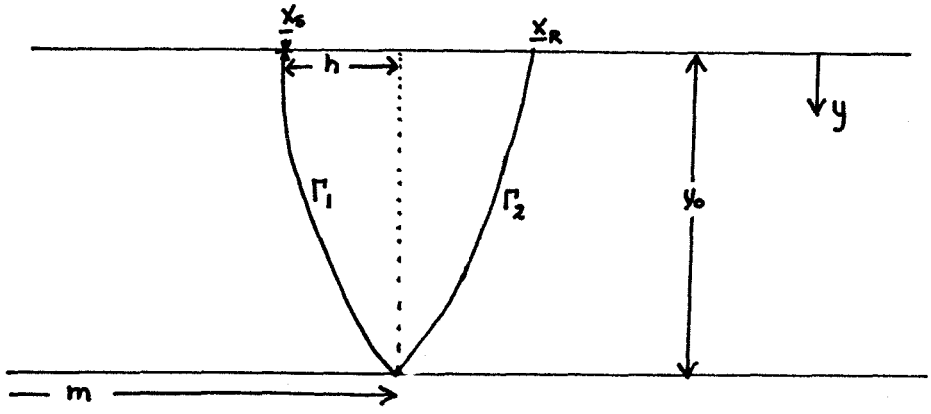


Figure 2.2 Reflection Seismology Geometry

Referring to figure 2.2, there will in general be rays with turning points within the layer. We will not include these rays in our formulation. We consider instead only the rays that reflect back to the surface from the reflector. In keeping with common exploration geophysics' coordinates, we define the offset h and the midpoint m coordinates.

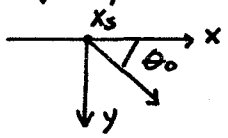
$$h = .5 * (x_R - x_S) \qquad (1.7)$$
$$m = .5 * (x_R + x_S)$$

We will now denote the travel time perturbations $\hat{\tau}(\underline{x}_S, \underline{x}_R)$ as t(m,h) and we denote the perturbed slowness field $n_1(x,y)$ as U(x,y). The ray segments Γ_1 and Γ_2 are parametrized in terms of the coordinate y. The ray solutions for a velocity field that varies only with depth are simple (Whitham [21]).

$$x - x_1 = \int_{y(x_1)}^{y} \frac{p c(y') dy'}{\sqrt{1 - p^2 c^2(y')}} \qquad \frac{ds}{dy} = \frac{1}{\sqrt{1 - p^2 c^2(y)}}$$

$$\tau(x,y;x_1,y_1) = \int_{y_1}^{y} \frac{dy'}{c(y')\sqrt{1-p^2c^2(y')}} \qquad (1.8)$$

$$p(h) \equiv \frac{\cos\theta_0}{c(0)}$$



We will use "p" in this chapter to denote the ray parameter $p=\frac{1}{c(y)}\frac{dx}{ds}$. This variable is constant along a ray and hence is simply related to the "shooting angle" θ_0 at the source, $p=\frac{\cos\theta_0}{c(0)}$. We will consider p as a function of the offset h. Using the notation of (1.8), we can write

$$t(m,h) = \int_0^{y_0} \frac{U(x_1(y'),y')dy'}{\sqrt{1-p^2c^2(y')}} + \int_0^{y_0} \frac{U(x_2(y'),y')dy'}{\sqrt{1-p^2c^2(y')}} \qquad (1.9)$$

where $x_1(y)$ is the "x" coordinate on Γ_1

$$x_1(y) = m - \int_y^{y_0} \frac{pc(y')dy'}{\sqrt{1-p^2c^2(y')}} \qquad (1.10a)$$

and $x_2(y)$ is the "x" coordinate on Γ_2

$$x_2(y) = m + \int_y^{y_0} \frac{pc(y')dy'}{\sqrt{1-p^2c^2(y')}} \qquad (1.10b)$$

This ray formulation follows closely that of Romanov [15], who considers the inverse problem for a continuous medium (not a reflection problem), and formulates the problem with the turning point Z(p) of the rays being an important coordinate. We take the Fourier Transform of both sides of (1.8) with respect to m

$$\hat{t}(k_m,h) = \int_{-\infty}^{\infty} t(m,h)e^{-ik_m m}dm \qquad (1.11)$$

to obtain:

$$\tilde{t}(k_m,h)=2\int_0^{y_0}\frac{\hat{U}(k_m,y')}{\sqrt{1-p^2(h)c^2(y')}}\cos\left\{k_m\int_{y'}^{y_0}\frac{p(h)c(y'')\,dy''}{\sqrt{1-p^2(h)c^2(y'')}}\right\}dy' \qquad (1.12)$$

where we have used the shift formula of Fourier Transforms on the right hand side of (1.8). This is the fundamental relation we wish to solve. However, due to the Fredholm nature of (1.12) it is not known whether in fact a solution to (1.12) exists. Thus, in general, we are attempting to find U(m,y) such that (1.12) is "approximately" satisfied. One possibility would be the discretization of (1.12) and an optimization solution of the resulting problem. The formula (1.12) for the special case C(y)= constant, was derived and used by Kjartansson [13] in his inversion of travel time anomalies. He "inverted" (1.12) by using a least-squares approach.

We now also consider the case C(y)= constant. We will derive directly the Radon-Inversion formula for (1.12). This has been derived before by Clayton [6]. Also, although the usual Radon Transform (e.g., as used in Medical X-ray Tomography) is formulated in terms of different variables, our formula can be obtained from the more standard formula with appropriate assumptions and changes of variables. However, we present our derivation as our more general formula (C=C(y)) will follow from this.

Instead of the reflection problem (figure 2.2), we consider the entire velocity field in the layer to be reflected about the line $y=y_0$ and we consider the equivalent transmission problem shown below in figure 2.3.

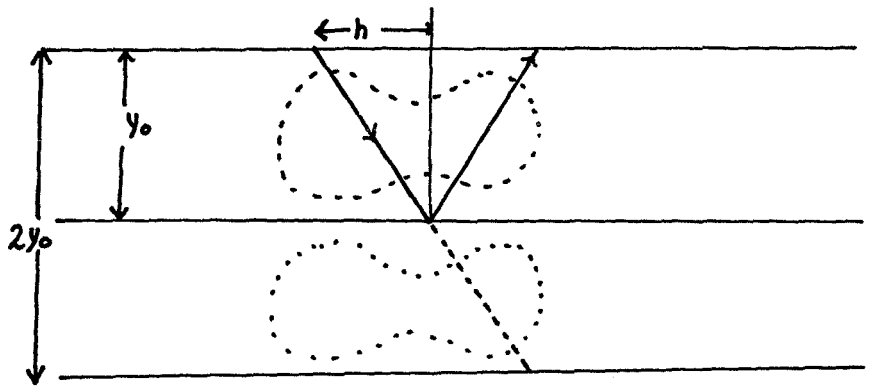


Figure 2.3 Equivalent Transmission Problem

Now we can write:

$$t(m,h) = \int_0^{2y_o} U(x(y),y)\,\frac{\sqrt{h^2+y_o^2}}{y_o}\,dy \qquad (1.13)$$

$$x(y) = m + h\left(\frac{y}{y_o} - 1\right)$$

Defining $\hat{t}(m,h) = \frac{y_o}{\sqrt{h^2+y_o^2}}\, t(m,h)$, and taking the Fourier Transform of both sides of (1.13) with respect to "m" we obtain:

$$\hat{\tilde{t}}(k_m,h) = \int_0^{2y_o} e^{-ik_m h\left(1-\frac{y}{y_o}\right)}\,\tilde{U}(k_m,y)\,dy \qquad (1.14)$$

We now assume a solution for (1.14). Then, we can extend the limits of the integration and consider the new problem

$$\hat{\tilde{t}}(k_m,h) = \int_{-\infty}^{\infty} e^{-ik_m\left(1-\frac{y}{y_o}\right)h}\,\tilde{U}(k_m,y)\,dy \qquad (1.15)$$

We will find the solution to (1.15) by using successive Fourier Transforms

and the answer will be unique (at least for km≠0). Hence, if (1.14) had a solution, then the answer to (1.15) must simply be the solution to (1.14) for $0<y<2y_o$,and zero elsewhere. If, in fact, t(m,h) does not correspond to the rectilinear projection of any anomaly field for $0<y<2y_o$, then we hope that the solution to (1.15) will be a good approximation to (1.14) for $0<y<2y_o$.

Recognizing that the right hand side of (1.15) is simply a Fourier Transform over the "y" coordinate we obtain

$$\frac{1}{2\pi} e^{ik_m h} \tilde{\hat{t}}(k_m, h) = \tilde{\tilde{U}}(k_m, \frac{h k_m}{y_o}) \qquad (1.16)$$

where $\tilde{\tilde{U}}(k_m, k_z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ik_z y} \tilde{U}(k_m, y)\, dy$. Making the change of variable $k_y = h k_m / y_o$ we obtain

$$\frac{1}{2\pi} e^{ik_y y_o} \tilde{\hat{t}}(k_m, \frac{y_o k_y}{k_m}) = \tilde{\tilde{U}}(k_m, k_y) \qquad (1.17)$$

Equation (1.17) illustrates the limitations of our method. In order to determine $\tilde{\tilde{U}}(k_m, k_y)$ for $k_m = 0$, $k_y \neq 0$, we must know $\hat{t}(m,h)$ for infinite offset. In practice we can only measure the travel times to some maximum offset h max. For simplicity, we set t(km,h)=0 for |h|>h max. Thus we cannot hope to resolve the slow lateral variations of the field. Performing the inverse Fourier Transform of (1.17) we obtain:

$$U(m,y) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{ik_m h} \hat{\hat{t}}(k_m, \frac{y_o k_y}{k_m}) e^{ik_m m - ik_y y}\, dk_m\, dk_y \qquad (1.18)$$

Now we let $\eta = k_y y_o / k_m = h$ and obtain:

$$U(m,y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{ik_m(\eta - \eta\frac{y}{y_o} + m)} \tilde{\hat{t}}(k_m, \eta) \frac{|k_m|}{y_o} \, dk_m \, d\eta \qquad (1.19)$$

The factor $|k_m|/y_o$ is the Jacobian of the transformation (km,ky) --> (km, η) and we note that the Jacobian is singular for km=0. However since t(km=0,h)=0 for |h|>hmax, we can extend our integral to the whole space. Now we assume that $\tilde{\hat{t}}(km, \eta)$ decays sufficiently fast as a function of km so that we can approximate $|km|\tilde{\hat{t}}$ by $f(km)\tilde{\hat{t}}$, where f(km) is chosen so that $\mathcal{F}^{-1}(f(km))$ exists. Using the convolution theorem we obtain:

$$U(m,y) = \frac{F(m)}{2\pi y_o} * \int_{-\infty}^{\infty} \hat{t}(m + \eta(1 - \frac{y}{y_o}), \eta) \, d\eta \qquad (1.20)$$

Here, F(m) is $\mathcal{F}^{-1}(f(km))$ and "*" denotes convolution. Numerically we will take f(km) to be the tapered function |k|cos(k/kmax * π/2) for k< kmax and zero elsewhere. This filter function is shown schematically below in figure 2.4.

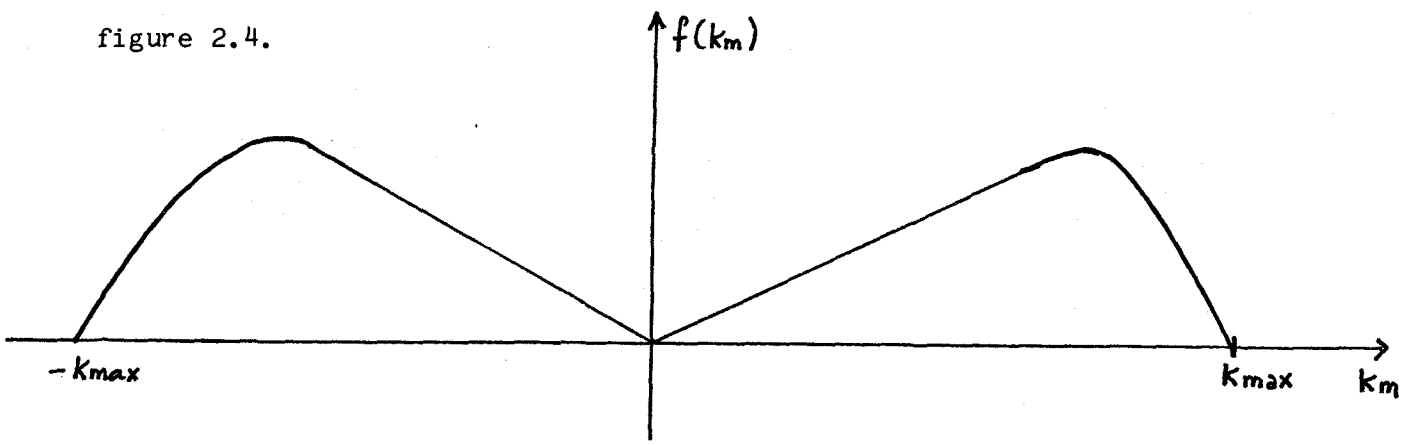


Figure 2.4 Filter Function In Fourier Space

The function F(m) is in "m" space a rapidly oscillating function with zero mean. From simple geometry (figure 2.5) it is easy to show that ignoring

the filter F(m) in (1.20), the integral in (1.20) is simply the integration of all the travel time perturbations over all the rays that pass through the point (m,y). We show below in figure 2.5 two rays that pass through (m,y) and correspond to the same offset "h". We denote their intersections with the reflector at y=y0 as m1 and m2.

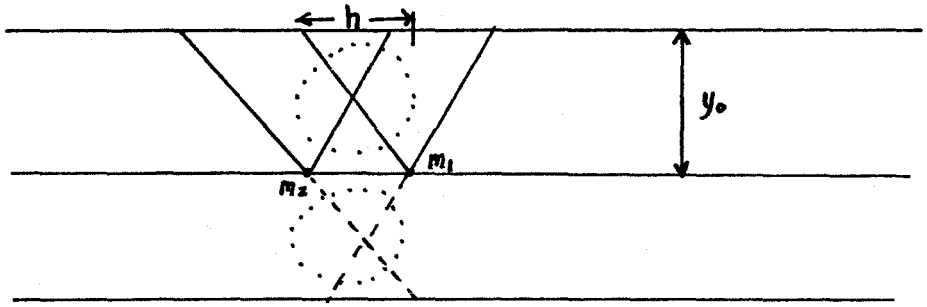


Figure 2.5 Relation Between Point in Space,Offset, and Midpoint

Some geometry shows that m1=h(1-y/y0)+m and m2=m-h(1-y/yo). These are the same two expressions that appear in the integrand of (1.20). We will call the formula of (1.20) without the filter function F(m) the back-projection approximation. We now show how this approximation is related to the "true field." Consider:

$$U_{ap}(m,y) = \int_{-h_{max}}^{h_{max}} \hat{t}\left(m + h\left(1 - \frac{y}{y_o}\right), h\right) dh \qquad (1.21)$$

$$\hat{t}(m,h) = \int_{-\infty}^{\infty} U_{tr}\left(m + h\left(\frac{y}{y_o} - 1\right), y\right) dy \qquad (1.22)$$

"ap"-approximate "tr"-true

Substituting (1.22) into (1.21) we obtain:

$$U_{ap}(m,\bar{y}) = \int_{-h_{max}}^{h_{max}} \int_{-\infty}^{\infty} U_{tr}\left(m + h\,\frac{(y-\bar{y})}{y_o},\, y\right) dh\, dy \qquad (1.23)$$

Let us now change to polar coordinates (r,θ) where $\theta = \tan^{-1}(y\,/\,)$ and $y - \bar{y} = r\sin\theta$. Using this change, we obtain:

$$U_{ap}(m,\bar{y}) = \int_{\tan^{-1}\left(\frac{y_o}{h_{max}}\right)}^{\pi - \tan^{-1}\left(\frac{y_o}{h_{max}}\right)} \int_{0}^{\infty} \frac{U_{tr}(m + r\cos\theta,\, \bar{y} + r\sin\theta)\, r\, dr\, d\theta}{y_o\, r\, |\sin\theta|} \qquad (1.24)$$

$$+ \int_{\pi + \tan^{-1}\left(\frac{y_o}{h_{max}}\right)}^{2\pi - \tan^{-1}\left(\frac{y_o}{h_{max}}\right)} \int_{0}^{\infty} \frac{U_{tr}(m + r\cos\theta,\, \bar{y} + r\sin\theta)\, r\, dr\, d\theta}{y_o\, r\, |\sin\theta|}$$

Thus, when we use the "back projection" method, we have found a field which is given by:

$$U_{ap}(m,y) = \frac{\Omega(m,y)}{y_o} * U_{tr}(m,y) \qquad (1.25)$$

where the filter $\Omega(x,y)$ is shown schematically, below in figure 2.6.

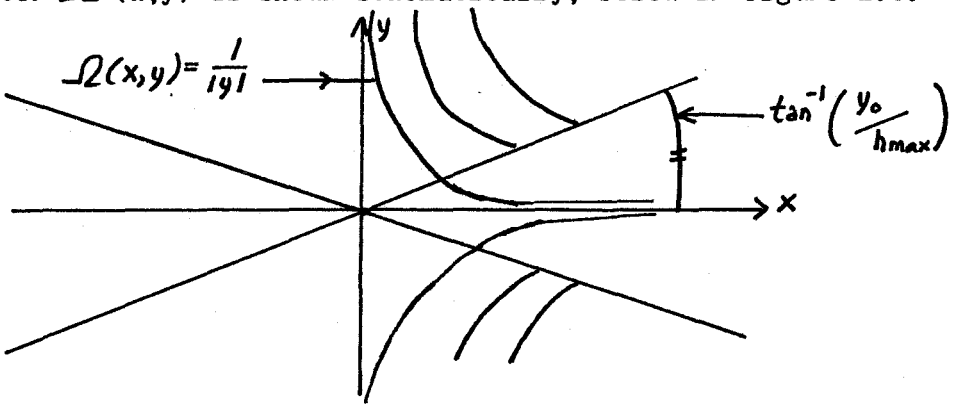


Figure 2.6 Back Projection Filter

We note that the field $U_{tr}(m,y)$ under consideration, in our formulation is the symmetrized field shown below in figure 2.7.

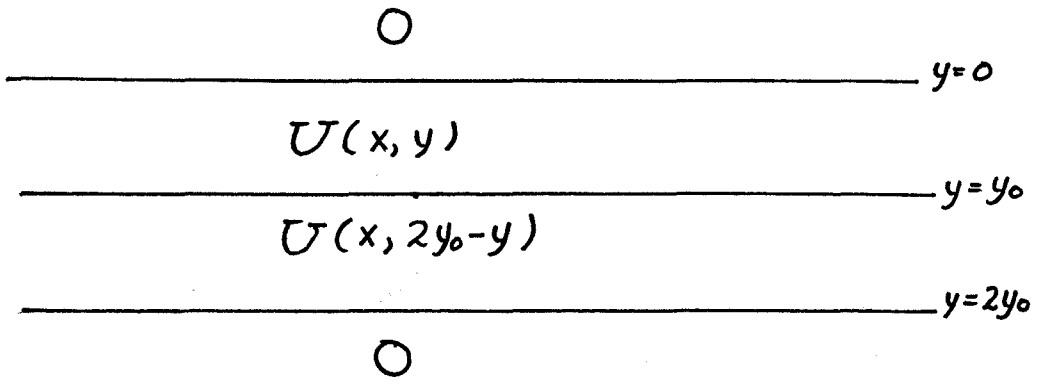


Figure 2.7 Symmetrized Field.

Thus the filter F(m) is chosen to negate the effects of convolution with $\Omega(x,y)$. It is certainly not obvious that $\mathcal{F}(F(m)) = \frac{|k_m|}{2\pi}$ is the appropriate choice. To show that this is indeed the proper F(m), we take the Fourier Transform of both sides of (1.23) with respect to "m" obtaining

$$\tilde{U}_{ap}(k_m, \bar{y}) = \int_{-h_{max}}^{h_{max}} \int_{-\infty}^{\infty} \tilde{U}_{tr}(k_m, y) e^{-ik_m h \frac{(y-\bar{y})}{y_o}} dh\, dy \qquad (1.26)$$

Furthermore taking the Fourier Transform (formally) of both sides of (1.26) with respect to $\bar{y}$ we obtain

$$\tilde{\tilde{U}}_{ap}(k_m, k_y) = 2\pi \int_{-h_{max}}^{h_{max}} \int_{-\infty}^{\infty} \tilde{U}_{tr}(k_m, y) \delta\left(\frac{h k_m}{y_o} - k_y\right) e^{-ik_m h \frac{y}{y_o}} dh\, dy \qquad (1.27)$$

Now we define $\xi = h k_m / y_o$ for $k_m \neq 0$. Thus

$$\tilde{\tilde{U}}_{ap}(k_m, k_y) = 2\pi \int_{-\frac{|k_m| h_{max}}{y_o}}^{\frac{|k_m| h_{max}}{y_o}} \int_{-\infty}^{\infty} \tilde{U}_{tr}(k_m, y) \frac{\delta(\xi - k_y) y_o}{|k_m|} e^{-i\xi y} d\xi\, dy \qquad (1.28a)$$

$$\tilde{\tilde{U}}_{ap}(k_m, k_y) = 2\pi H\left(\frac{|k_m| h_{max}}{y_o} - k_y\right) \int_{-\infty}^{\infty} \frac{\tilde{U}_{tr}(k_m, y) e^{-ik_y y}}{|k_m|} dy \qquad (1.28b)$$

$$\tilde{\tilde{U}}_{ap}(k_m,k_y)=H\left(\frac{|k_m|h_{max}}{y_o}-k_y\right)\frac{2\pi y_o}{|k_m|}\,\tilde{\tilde{U}}_{tr}(k_m,k_y) \qquad (1.28c)$$

$$\tilde{\tilde{U}}_{tr}(k_m,k_y)=\frac{|k_m|}{2\pi y_o}\,\tilde{\tilde{U}}_{ap}(k_m,k_y);\quad \left|\frac{k_y y_o}{k_m}\right|\leq h_{max} \qquad (1.28d)$$

Here H(x) is the Heaviside function. Thus we can start with the back-projection approximation and derive F(m) by finding the required filter to make the approximation as exact as possible.

We now return to the more general problem where the known background field C(y) is not a constant. Once again we consider the transmission problem and we indicate some of our notation below in figure 2.8.

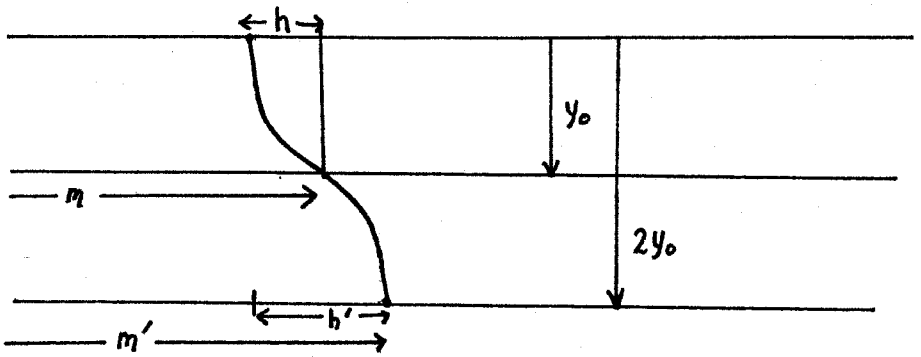


Figure 2.8 Geometry for General Transmission Problem

For a reflection problem, we consider the velocity field and the rays reflected about the line $y=y_o$. Our results are easily extendable to the general transmission problem through a layer of thickness of $2y_o$ where the field is not neccessarily symmetric. However, for this case the coordinates

m,h would be the most appropriate. Let us now suppose that the velocity anomaly lies in a very thin layer which is shown below if figure 2.9.

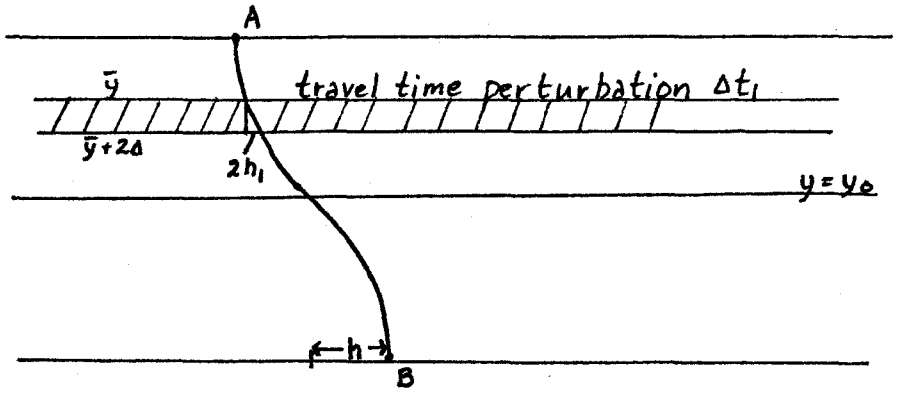


Figure 2.9 Thin Layer Velocity Anomaly For Transmission Problem

The "true" seismic experiment measures the travel time difference from A to B, but instead we consider another "experiment" where the sources are on the surface $y=\bar{y}$, and the receivers are at $y=\bar{y}+2\Delta$. We denote the travel time perturbations between $y=\bar{y}$ and $y=\bar{y}+2\Delta$, along the ray joining the points A and B, as t1, and t1 will equal the perturbation t2 along the entire ray since there is no velocity anomaly outside the region 2Δ. Now we will suppose that Δ is very small and we have only a finite range of offsets, so that the rays are locally very straight in the layer. Using a straight line approximation, we have from our approximate inversion formula (1.20)

$$U(m,y) = \frac{F(m)}{2\pi\Delta} * \int_{-h_{max}}^{h_{max}} \frac{t_1\left(m+h_1\left(1-\frac{(y-\bar{y})}{\Delta}\right),h_1\right)\,dh_1}{ds/dy\,|_{\bar{y}}} \qquad (1.29)$$

In (1.29), h1 refers to the infinitesimal offset of the ray segment in the layer 2Δ , and:

$$h_1 \simeq \Delta \left.\frac{dx}{dy}\right|_{\bar{y}} \; ; \; t_1\left(m+h_1\left(1-\frac{(y-\bar{y})}{\Delta}\right),h_1\right) = t_2\left(m+q\left(y,h(h_1)\right),h(h_1)\right)$$

$$q(y,h) \equiv \int_y^{y_0} \frac{p(h)c(y')dy'}{\sqrt{1-p^2c^2(y')}} \qquad \left.\frac{dx}{dy}\right|_h = -\frac{dq(y,h)}{dy} \tag{1.30}$$

Thus

$$U(m,y) \simeq \frac{F(m)}{2\pi} * \int_{-\left|\frac{dx}{dy}\right|_{max}}^{\left|\frac{dx}{dy}\right|_{max}} \frac{t_2(m+q(y,h),h(h_1))}{ds/dy\,|_{\bar{y}}} d\left(\left.\frac{dx}{dy}\right|_{\bar{y}}\right) \tag{1.31}$$

Now from $\left.\frac{d^2q}{dydh}\right|_{\bar{y}} dh = d\left(\frac{dx}{dy}\right)\Big|_{\bar{y}}$ we have

$$U(m,y) \simeq \frac{-F(m)}{2\pi} * \int_{-h_{max}}^{h_{max}} \frac{t_2(m+q(y,h),h)}{ds/dy\,|_{\bar{y}}} \left.\frac{d^2q}{dydh}\right|_{\bar{y}} dh \tag{1.32}$$

Thus, if indeed our anomaly was restricted to a thin layer, then formula (1.32) would be an accurate approximation at this depth. However, it is not clear that this formula is accurate for the general problem, but it is a limiting formula which we should obtain, when we consider a more general formula.

We now derive another formula, which is a near zero offset approximation. We begin with the integral relation (1.12)

$$\tilde{t}(k_m,h) = \int_0^{2y_0} \tilde{U}(k_m,y') e^{-ik_m q(y';h)} \frac{ds}{dy}(y';h)\,dy' \tag{1.33}$$

and we expand the integrand of (1.33) about h=0 to obtain

$$\tilde{t}(k_m,h) = \int_0^{2y_0} \tilde{U}(k_m,y) e^{-ik_m\left[\int_y^{y_0} c(y')dy'\right]\frac{\partial p}{\partial h}|_{h=0}\,h} dy \tag{1.34}$$

Here we have used that $ds/dy|_{h=0}=1$, $d/dh(ds/dy)|_{h=0}=0$, $q(y,h)|_{h=0}=0$, and $d/dh(q(y,h)|_{h=0} = \left[\int_y^{y_0} c(y')dy'\right]\left.\frac{\partial p}{\partial h}\right|_{h=0}$. Now we rewrite

$$\int_y^{y_o} c(y')dy' = \int_0^{y_o} c(y')dy' - \int_0^{y} c(y')dy' \equiv \tau_o - \tau(y) \qquad (1.35)$$

We change variables in (1.34) from "y" to "τ" and use that $dp/dh|_{h=0} = 1/\tau_o$, (here, we are specifically considering a reflection geometry) and $dy/d\tau = -1/c(y(\tau))$ to obtain:

$$\tilde{t}(k_m,h) = \int_{-\infty}^{\infty} \frac{H(2\tau_o)\tilde{U}(k_m,y(\tau))e^{-ik_m h}\, e^{ik_m \frac{\tau}{\tau_o} h}}{c(y(\tau))}\, d\tau \qquad (1.36)$$

$$H(2\tau_o) \equiv 1 \quad 0 \le \tau \le 2\tau_o\,;\ 0 \text{ otherwise}$$

or

$$H(\Delta h)\,\hat{t}(k_m,h) = H(\Delta h)\,\hat{\tilde{g}}(k_m, \tfrac{k_m h}{\tau_o})$$

$$\hat{\tilde{g}}(k_m,\tau) = \frac{H(2\tau_o)\tilde{U}(k_m,y(\tau))e^{-ik_m h}}{c(y(\tau))} \qquad (1.37)$$

Here H(Δh) signifies that this result is only accurate for 0<h<Δh where Δh is some appropriate small interval. However we will simply let g(km,km h/τ_o) = $\tilde{t}$(km,h) for all h. Now (1.34) is simply the same expression (1.15) we had for the constant velocity case, but here y_o --> τ_o and $\tilde{U}$(km,y) --> $\tilde{U}$(km,y)/c(y). Thus proceeding formally we obtain:

$$\bar{g}(m,y) = \frac{F(m)}{2\pi\tau_o} * \int_{-\infty}^{\infty} t(m + h(1 - \tfrac{\tau(y)}{\tau_o}), h)\, dh \qquad (1.38)$$

Now we let

$$\bar{g}(m,y) = g(m,y) + E(m,y) \qquad (1.39)$$

where E(m,y) is the error from the extension of (1.37) to all "h." Thus

$$U(m,y) = \frac{F(m)}{2\pi\tau_0} c(y) * \int_{-\infty}^{\infty} t(m+h(1-\frac{\tau(y)}{\tau_0}), h)\,dh - c(y)E(m,y) \qquad (1.40)$$

We drop the error term from (1.40). We note that in (1.40), t(m,h) is the "true" travel time (not the normalized travel time as in the constant velocity case) since in the zero offset approximation ds/dy=1. The term $m+h(1-\tau/\tau_0)$ is simply the first order approximation to m+q(y,h) and the expression $c(y)/\tau_0$ is simply $-\left.d^2q/dy\,dh\right|_{h=0}$. If we insert the values of these expressions for arbitrary "h" into (1.40) we obtain

$$U(m,y) = \frac{-F(m)}{2\pi} * \int_{-h_{max}}^{h_{max}} \frac{t(m+q(y,h),h)}{ds/dy} \frac{d^2q}{dy\,dh}\,dh \qquad (1.41)$$

When we restrict (1.41) to small offsets, it is consistent with (1.40) restricted to small offsets (it is consistent to $O(\Delta h^2)$ since $\left.d^2s/dy\,dh\right|_{h=0}=0$ and $\left.d^3q/dy\,dh^2\right|_{h=0}=0$). Formula (1.41) reduces to (1.20) for c(y)=constant. Thus this simple approximation to equation (1.34), and our derived formula (1.32) seem to suggest the formula (1.41) as a possible approximate inversion formula, although so far, we have no strong justification for this.

We now examine (1.42) in more detail. Equation (1.42), neglecting the filter F(m), is a generalized backprojection formula. That is, to reconstruct the field at a spatial point (m,y) we sum up all the "normalized" travel times corresponding to rays that pass through the point. Here the normalization is a local one; we divide the traveltime by the local arclength ds/dy and multiply by the local focussing factor $-d^2q/dy\,dh$. We call the term $-d^2q/dy\,dh$ the focussing factor as this represents the change in the local slope dx/dy of the ray when we change

the offset "h" by a slight amount. We illustrate this concept schematically, below in figure 2.10.

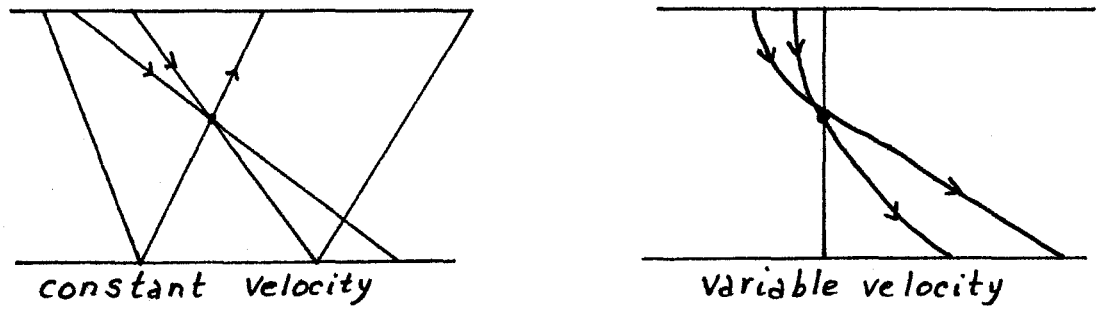


Figure 2.10 Focussing of Rays at a Spatial Point

We see from figure 2.10, that for the constant velocity case there is no focussing of rays and $-d^2q/dydh = 1/y_o$. For a background velocity field that varies with depth there will, in general, be focussing and defocussing of the rays. The focussing factor in the integral of (1.42) simply equalizes the weighting of the angular slope coverage by the rays of the point (m,y).

As in the constant velocity case, we now find an integral relation between the "true" field $U_{tr}(m,y)$ and the "approximate" field $U_{ap}(m,y)$, where $U_{ap}(m,y)$ is the field obtained from the generalized backprojection formula. We have:

$$U_{ap}(m,\bar{y}) = \int_{-h_{max}}^{h_{max}} \frac{t(m+q(\bar{y},h),h)\,\frac{d^2q}{dydh}\Big|_{\bar{y}}\,dh}{ds/dy\,|\,\bar{y}} \qquad (1.43a)$$

and:

$$t(m,h) = \int_0^{2y_o} U_{tr}(m-q(y,h),y)\,\frac{ds}{dy}(y,h)\,dy \qquad (1.43b)$$

Substituting (1.43b) into (1.42) we obtain

$$U_{ap}(m,\bar{y}) = -\int_{-h_{max}}^{h_{max}} \int_{0}^{2y_o} U_{+r}(m-q(y,h)+q(\bar{y},h),y)\, \frac{\frac{ds}{dy}(y,h)}{\frac{ds}{dy}(\bar{y},h)}\, \frac{d^2q}{dydh}\, dy\, dh \quad (1.44)$$

Now we consider a polar coordinate system centred at $(m,\bar{y})$ which we show diagramatically below in figure 2.11

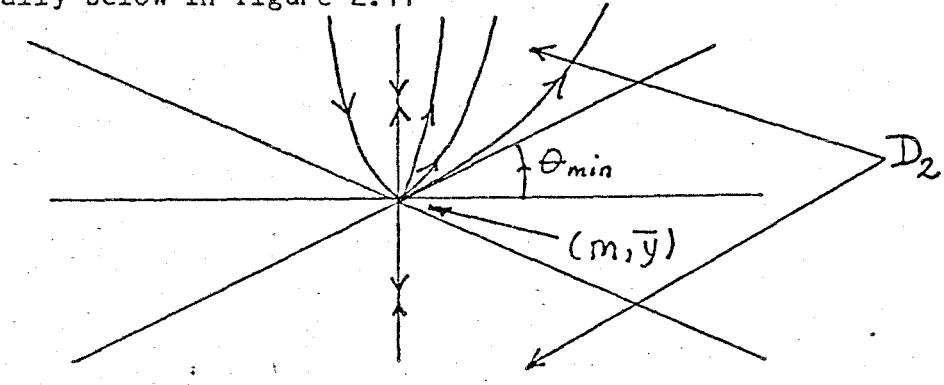


Figure 2.11 Rays and the Polar Coordinate System

Mathematically, D_2 is the domain: $|y/x| > \tan\theta_{min}$ We now change variables from (h,y) to (θ,r) where

$$\theta = \cot^{-1}\left(\frac{dx}{dy}\right) = -\cot^{-1}\left(\frac{\partial q(y,h)}{\partial y}\Big|_{\bar{y}}\right) \quad \frac{ds}{dy}\Big|_{\bar{y}} = \frac{1}{\sin\theta}$$

$$y - \bar{y} = r\sin\theta \qquad x - m = r\cos\theta \quad (1.45)$$

and the Jacobian of the transformation is simply $-1\Big/\left(d^2q/dydh\big|_{\bar{y}}\sin\theta\right)$. Thus we can rewrite (1.44) as

$$U_{ap}(m,\bar{y}) = \int_{D_2} \frac{U_{+r}(m+r\cos\theta+R_1(\theta,r),\ \bar{y}+r\sin\theta)(1+R_2(\theta,r))\,r\,dr\,d\theta}{|y-\bar{y}|} \quad (1.46)$$

Here D_2 is the domain shown in figure 2.11. The terms $R_1(\theta,r)$ and $R_2(\theta,r)$

are remainder terms from the first order expansion of the integrand of (1.45) about $y=\bar{y}$. ($R1(\theta,r), R2(\theta,r) \longrightarrow 0$ as $y \longrightarrow \bar{y}$) Now since the integrand decays like $1/|y-y|$ we assume that a good approximation to (1.46) is the local approximation where we let the "y" limits of D2 go to $\pm\infty$.

$$U_{ap}(m,y) = \int_{D_2} \frac{U_{tr}(m+r\cos\theta, \bar{y}+r\sin\theta)}{|y-\bar{y}|} r\,dr\,d\theta \qquad (1.47)$$

This approximation is most valid for one of two situations; when the anomalies are fairly confined and isolated vertically or as ds/dy and q(y,h) become independent of y (i.e., c(y)--> constant). As the offsets become smaller, ds/dy(y,h) and q(y,h) vary less with depth, and approximation (1.47) becomes increasingly accurate for the small offset information. This is independent of the spatial properties of the anomaly field, and the inversion formula which we derive from (1.44) and (1.47), will be consistent for small "h" with formula (1.40). Alternatively, for any fixed finite maximum offset, if there is only one anomaly region and it is confined to a small vertical thickness 2Δ at $y=y$, then as $\Delta \longrightarrow 0$ the local assumption for (1.47) becomes increasingly valid for $y=\bar{y}$, and our formula (1.50b) will simply be (1.32). For c(y) becoming a constant, then (1.44) becomes (1.47) globally, and from (1.47) we obtain the formula (1.20). In general, there are certainly situations, where the approximation (1.47) may not be good. For example, if the anomaly field is not localized, then our approximation to relation (1.46) ((1.47)) may not be so valid. However, it is true, that per unit area, the integrand of (1.46) is weighted most heavily for $y = \bar{y}$, and more practically, the approximation (1.47) allows us to proceed mathematically. In a more

algorithmic sense, for general background velocity fields, the backprojection approximation to the anomaly field can always be defined. The question that arises is how to best normalize the travel time data for the backprojection. Our algorithm has suggested a very local normalization; we divide the travel time data, corresponding to a particular ray by the local arclength at a depth $\bar{y}$ (to reconstruct the field at $y=\bar{y}$) and compensate for the local focussing. However, for an anomaly field which is global in nature, perhaps it would be better to use another normalization scheme. Clayton [5], has related this problem to the solution of (1.6).When the normal equations for the problem (1.6) are formed, there are various iterative methods of solution which could(at least theoretically) be used to solve the resulting linear system of equations(eg. Jacobi,Gauss-Seidel,etc.) The first iteration (with $\underline{\Delta S}=0$) gives an estimate which can be shown to be very similar to the backprojection approximation. However, there is an arbitrariness in the weighting for a least squares problem, and different weightings lead to different normalizations for the backprojection formula.

Equation (1.47) is almost identical to (1.24) for the constant velocity case, except that the domain of integration changes with depth. That is,from (1.47) it seems as if we are locally convolving the true field with the function $\Omega_{\theta_{min}}(x,y)$ where $\Omega_{\theta_{min}}(x,y)=1/|y|$ for $|y/x|>\tan\theta_{min}$ and $\Omega_{\theta_{min}}(x,y)=0$ otherwise. This function is schematically shown below in figure 2.12.

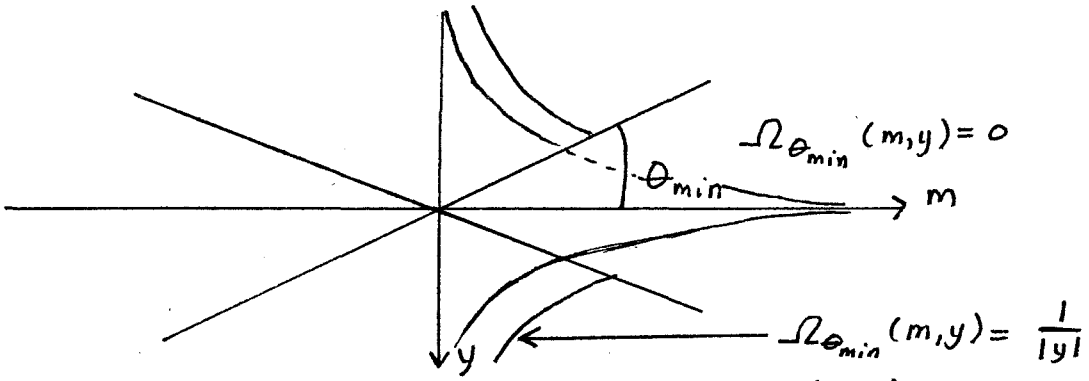


Figure 2.12 Convolution Function $\Omega_{\theta_{min}}(m,y)$

For a constant background field, θ_{min} is constant with depth. In general, however, θ_{min} changes with depth. However, we will still write for the back-projected approximate field $U_{ap}(m,y)$:

$$U_{ap}(m,y) = \Omega_{\theta_{min}}(m,y) * U_{tr}(m,y) + E(m,y) \qquad (1.48)$$

$$E(m,y) \equiv \text{error term}$$

In the Fourier domain, we can write (1.48) as:

$$\tilde{\tilde{U}}_{ap}(k_m,k_y) = \frac{1}{|k_m|}\tilde{\tilde{U}}_{tr}(k_m,k_y) + \tilde{\tilde{E}}(k_m,k_y), \left|\frac{k_y}{k_m}\right| \geq \tan\theta_{min} \qquad (1.49)$$

$$\tilde{\tilde{E}}(k_m,k_y), \left|\frac{k_y}{k_m}\right| < \tan\theta_{min}$$

We now define $\hat{U}_{tr}(m,y)$ such that $\hat{\tilde{U}}_{tr}(km,ky) = \tilde{U}_{tr}(km,ky)$ for $|km/ky| < \cot\theta_{min}$ and $\hat{\tilde{U}}_{tr}(km,ky) = 0$ otherwise. Now we can write (1.49) as:

$$\tilde{\tilde{U}}_{ap}(k_m,k_y) = \frac{1}{|k_m|}\hat{\tilde{U}}_{tr}(k_m,k_y) + \tilde{\tilde{E}}(k_m,k_y) \qquad (1.50a)$$

and

$$\hat{U}_{tr}(m,y) = F(m) * U_{ap}(m,y) - F(m) * E(m,y) \qquad (1.50b)$$

Now within the approximation (1.48), E(m,$\bar{y}$) is independent of "m" so that:

$$\hat{U}_{tr}(m,\bar{y}) = \left(F(m) * U_{ap}(m,y)\right)\Big|_{y=\bar{y}} \qquad (1.51a)$$

and

$$\hat{\tilde{U}}_{tr}(k_m,k_y) = \tilde{U}_{tr}(k_m,k_y) \qquad \left|\frac{k_y}{k_m}\right| > \tan\theta_{min} \qquad (1.51b)$$

We note from (1.27c) that F(m) is independent of θ_{min}. However, how accurately we can resolve $U_{tr}(m,y)$ (i.e., how close $\hat{U}_{tr}$ is to U_{tr}) does depend on θmin. As we have mentioned, in the Fourier domain, we can only resolve $\tilde{\hat{U}}(k_m,k_y)$ for $|k_y/k_m| < \cot\theta_{min}$. Now we consider what this means in the spatial domain. As can be seen below in figure 2.13 , near vertical rays can resolve the extreme lateral boundaries of an anomaly. As the offset is increased, the lateral resolution is increased,but we have little vertical resolution. Similarly, large offsets resolve the vertical variations but do not resolve well the lateral boundaries.

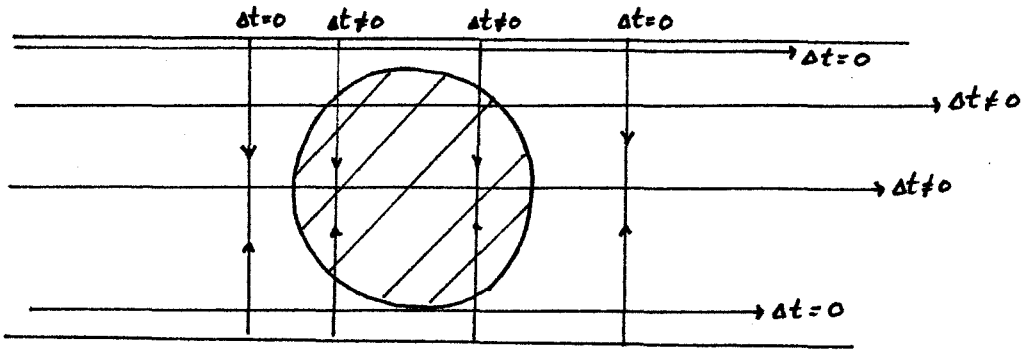


Figure 2.13 Spatial Resolution of the Rays.

Thus, near the top of the layer where the rays are more vertical than at the bottom, we can not resolve well the vertical extent of anomalies, whereas at the bottom of the layer we have both vertical and horizontal resolution.

In summary, our approximate inversion formula, for a vertically varying background velocity is as follows. We use the backprojection formula (1.43a). This means that at a particular point in space (x,y), we have approximately, that the backprojected field (at depth y) is the convolution of a particular function $\Omega_{\theta_{min}}(x,y)$ with the true field. We then showed that deconvolving $U_{ap}(m,y)$ (i.e., convolution with $F(m)/2\pi$) the backprojected field of $t(m,h)$ gave us at depth y an approximation to the true anomaly field within certain resolution limits. For a c(y) increasing with depth, the resolution of our method increases with depth (oppositely for c(y) decreasing). However, the validity of the method depends upon the validity of a localization assumption. (see (1.47) and (1.48)) As anomalies become spatially smaller(vertically wise) and more isolated from each other, or as c(y)--> constant, this assumption should improve.

Section 2 Numerical Implementation.

To test our ideas on tomographic inversion we required two basic programs: one to generate synthetic travel time perturbations for a known anomaly field and background field, and, secondly an inversion program using equations (1.41) or (1.40) to invert the travel time information. All computations were done in single precision Fortran-77 on a VAX computer.

2.1 Details of Numerical Programs.

The data are synthetically generated from the projection of the unperturbed ray through the anomaly field. In other words, the data does not correspond exactly to the perturbations which would be measured in a "true" seismic experiment. However, as discussed above, for relatively small amplitude anomalies, this is a first order accurate approximation. Hence our numerical results, do not test the validity of the linearization of the problem, but instead show the accuracy of the inversion formula (1.41) and (1.40), assuming that the linearized problem is valid. We hope in the future to redo some examples, using finite difference modelling to generate the observed acoustic signals, and address the problem of determining the exact perturbations (accurately "picking" traveltimes is in reality, a major problem in itself), and using this data as input to our inversion program. The background field we consider is of the form $c(y)=ay+b$. We take the velocity anomalies to be disks. Their radii, postion, and the strength of the constant perturbation within the disk are

user input parameters. Due to the linearity of the data generation we can superimpose disks to form other interesting shapes (e.g., an annulus). It can be shown that the rays are the arcs of circles. Some algebra allows us to determine the equation of these circles. The intersection points of a ray y1,i and y2,i with disk "i" are found by applying the quadratic formula for the intersection of two circles. Then the contribution of disk "i" to the perturbation t(m,h) is

$$t(m,h) \leftarrow\!- t(m,h) + V_i \cdot s_i \qquad (2.1)$$

Here Vi is the constant perturbation in the disk i and s_i is the computed arclength of the ray between y1,i and y2,i. We repeat these calculations for the ascending ray segment x2(y). The user specifies the maximum and minimum midpoint, and the percentage of the maximum offset to calculate. The maximum offset, for a velocity profile that increases with depth, corresponds to the ray that has a continuous turning point at $y=y_o$ (for a profile that decreases with depth the maximum offset corresponds to the ray that is horizontal at the surface). Sixty four increments in h, and m are then calculated (we only calculate h>0 as we know that t(m,-h)=t(m,h)) This data file is then stored for future use, and can be displayed by a raster plot in the format which is illustrated below in figure 2.14.

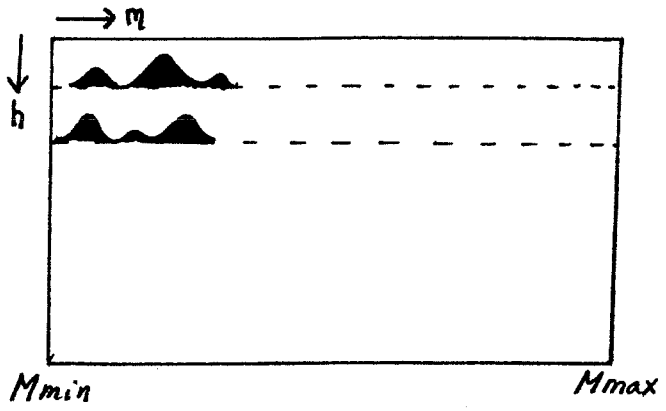


Figure 2.14 Travel Time Data Plot

The two basic formula we wish to examine numerically are equations (1.40) and (1.41):

$$U(m,y) = \frac{F(m)}{2\pi} * \int_{-h_{max}}^{h_{max}} \frac{t(m+q(y,h),h)}{ds/dy} \left| \frac{\partial^2 q}{\partial y \partial h} \right| dh \qquad (2.2a)$$

$$U(m,y) = \frac{c(y)}{\tau_0} F(m) * \int_{-h_{max}}^{h_{max}} t(m+h(1-\frac{\tau}{\tau_0}),h)dh \qquad (2.2b)$$

In fact, we do not actually consider negative offsets but write, for example, (2.2a) as:

$$U(m,y) = \frac{F(m)}{2\pi} * \int_{0}^{h_{max}} \frac{\{t(m+q(y,h),h)+t(m-q(y,h),h)\}}{ds/dy\,(y,h)} \left| \frac{\partial^2 q}{\partial y \partial h} \right| dh \qquad (2.3)$$

Numerically, we use the trapezoidal rule to approximate (2.3), where Δh and h_i are known from the available data. The corresponding midpoints, $m_i+q(y,h_i)$ or $m_i-q(y,h_i)$ for (2.2a), or $m_i \pm h(1-\tau/\tau_0)$ for (2.2b), may not correspond to generated data points, so we use linear interpolation between neighbouring midpoints to determine the appropriate value to use. When $m_i+q(y,h_i) >$ M max or $m_i-q(y,h_i) <$ M min, we set $t(m_i \pm q(y,h_i)) = 0$,

where M max and M min are the limits of the recorded midpoint data. This will lead to artifacts at the edges of the anomaly field, in our reconstruction procedure.

When all the integrations have been done, we have reconstructed the back- projected field. This field can now be convolved with the function F(m) to improve the "map" of the anomaly field. We do this convolution as follows. In the discrete Fourier domain we construct the wavenumber function which is shown schematically below in Figure 2.15.

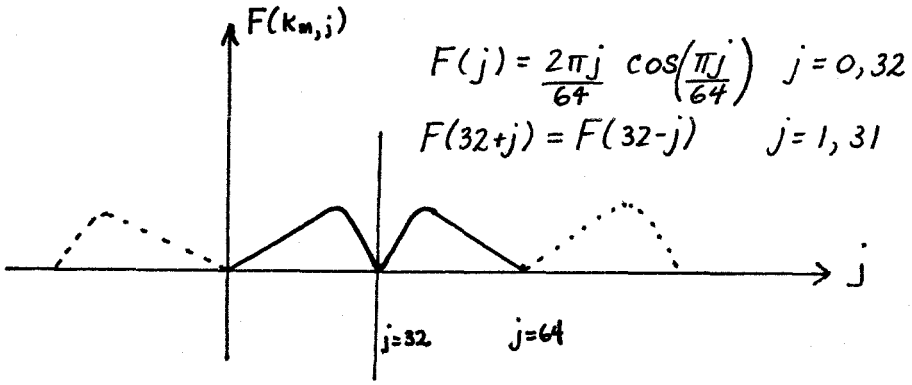


Figure 2.15 Filter Function in Wavenumber Space.

This function is then transformed back to the physical domain for F(m). F(m) is, by construction, a periodic function which is shown schematically below in Figure 2.16.

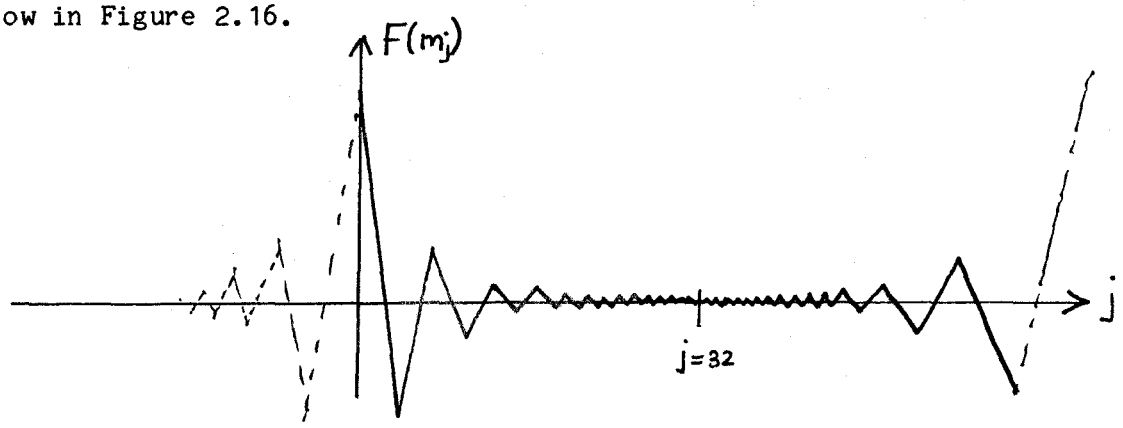


Figure 2.16 Filter Function in Physical Space

If we simply convolve the function shown in figure 2.16 with U(m,y) we will have undesired spatial aliasing effects. To circumvent this, we augment the vector F(mj) j=0,63 with sixty-four zeros, so that schematically F(m) now is as shown below in Figure 2.17.

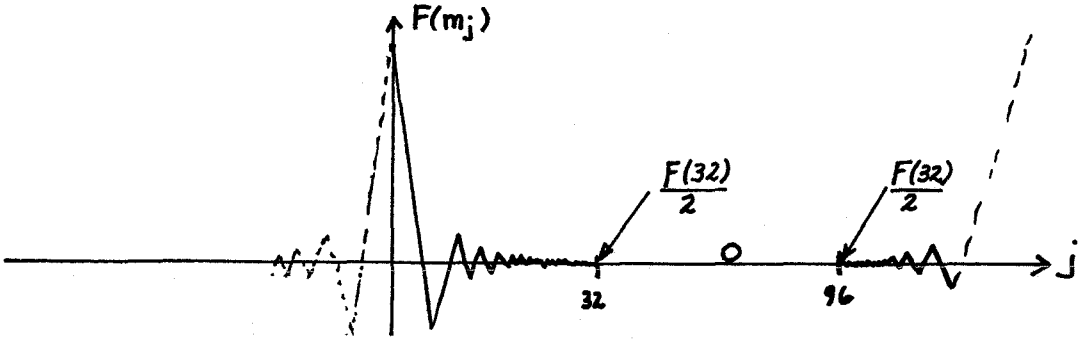


Figure 2.17 Augmented Filter Function

Similarly, we augment the vector $U(m_j,y)$ with zeros. We now take the F.F.T. of each sequences, multiply the Fourier coefficents together and transform back. We can plot the reconstructed fields with a raster plot.

2.2 Numerical Examples.

Example 1.

For our first example, we consider a disk of radius 1, with constant perturbation 1, located at the centre of the field. Our minimum and maximum midpoints are for this example, M min=-5 and M max=5, and the thickness of the layer is d=4. Our basic model for this example is shown schematically below in figure 2.18.

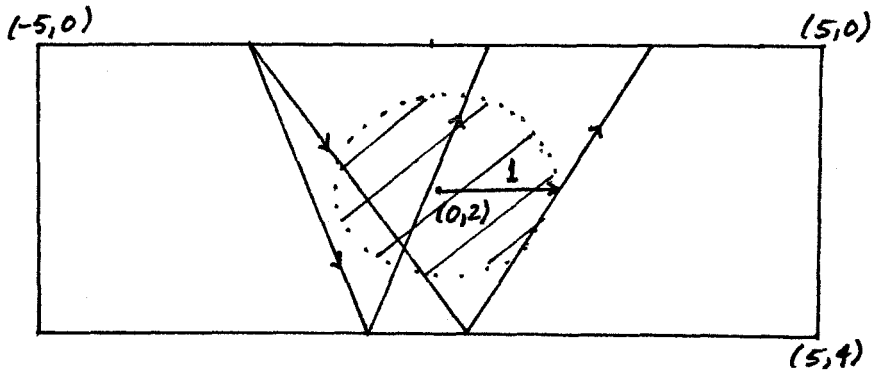


Figure 2.18 Synthetic Anomaly Model

We will vary the velocity profile and also the offset coverage for various examples. In all examples we calculate 64 midpoint positions, where the discrete midpoint positions are given by mj=M min +(j-1)x (M max -M min)/63. To start, we consider the constant velocity case. We shall synthetically generate t(m,h) for mj and hj=(j-1)X8.953/63 (j=1,64). For the inversion of this data file we shall use the constant velocity formula (1.21). Figure 2.19a shows the travel time perturbation data for this problem. The plot layout is as in figure 2.14. We note the two "arms" of data. If we had taken a point anomaly instead of a finite thickness disk, then the arms would be two straight lines, and the slopes of these lines would give the depth to the anomaly. For the point at $(\bar{m},\bar{y})$, then from the two arms of figure 2.19a, we would have from (1.20):

$$\frac{m-\bar{m}}{h} = 1 - \frac{\bar{y}}{y_0} \qquad (2.4)$$

Thus, the slope of these arms leads to a depth estimate for a spatially confined anomaly. However, from (1.41) for the general case, we have that for h =0

$$\frac{m-\bar{m}}{h} = 1 - \frac{\int_0^{\bar{y}} c(y')dy'}{\int_0^{y_0} c(y')dy'} \qquad (2.5)$$

Hence, as we shall see, simply using (1.20) for c(y)=constant will, along with other errors, incorrectly determine depths of anomalies.

Figure 2.19b shows the resulting back-projected field (i.e., no filtering applied). In these plots of the reconstructed field, we see the 64 midpoints, and 16 depth increments yk=(k-1)x4/15 (k=1,16). Figure 2.19b agrees intuitively with the idea of convolving the "true" symmetrized field with $\Omega(m,y)$ (see (1.25)), and we see immediately, that applying a filter which "kills" slow horizontal variations is an appropriate choice. Figure 2.19c shows the results of applying the filter F(m) to Figure 2.19b. Figure 2.19d shows the reconstructed field, when the maximum offset coverage has been reduced to hmax=4.48. The two offset coverages correspond to the angles shown below in figure 2.20.

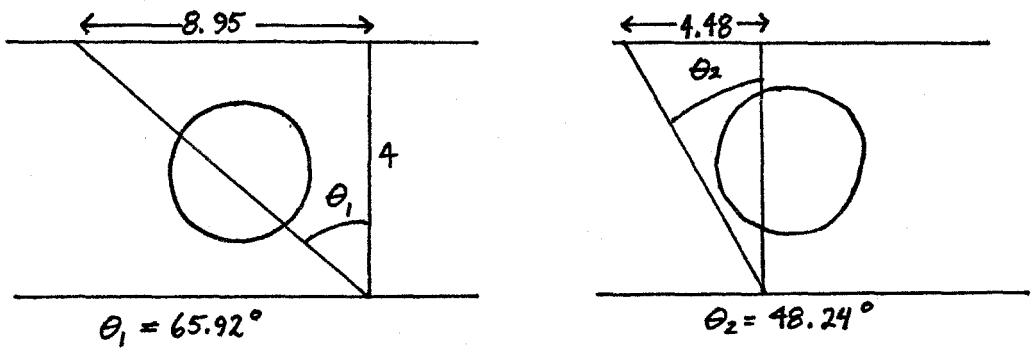


Figure 2.20 Angular-Offset Coverage

We see, as we expected, that the vertical resolution of the anomaly, has decreased. Figure 2.19e shows the synthetic travel time perturbations for C(y)=.2y+1. For this case, the offset corresponding to a turning point at

y=4 is h=7.48, and we numerically take hmax=.8x7.48=5.98. The reconstructed field is shown in figure 2.19f. We note, that as we discussed in the theory, the vertical resolution of the anomaly, particularly at the top of the anomaly has decreased. We used equation (2.3a) for this reconstruction. For our largest gradient example, we take c(y)=1y+1 and compute $t(m_j,h_j);(m_j,h_j)$ for h_j from zero to 3.92. Figure 2.19g, shows the result of inversion, using the constant velocity formula (1.21). As expected from above, we see that the circle is too high in the field.Figures 2.19h and 2.19i show the results of using (2.3a) and (2.3b) respectively.The results are very comparable.

Example 2.

As we have discussed, our method is most accurate for anomalies near the bottom of the layer for velocity profiles which increase with depth. We will consider such a situation, with three disks of varying position, radii, and perturbation strength. Some rays will pass through more than one anomaly. We choose circles of radii (.5,1,.5). Their centres are located at (x1=-3,y1=3), (x2=0,y2=3), and (x3=2,y3=3). The constant perturbations in each are (2,1,2) respectively. Schematically, the anomaly field is as shown below in figure 2.21.

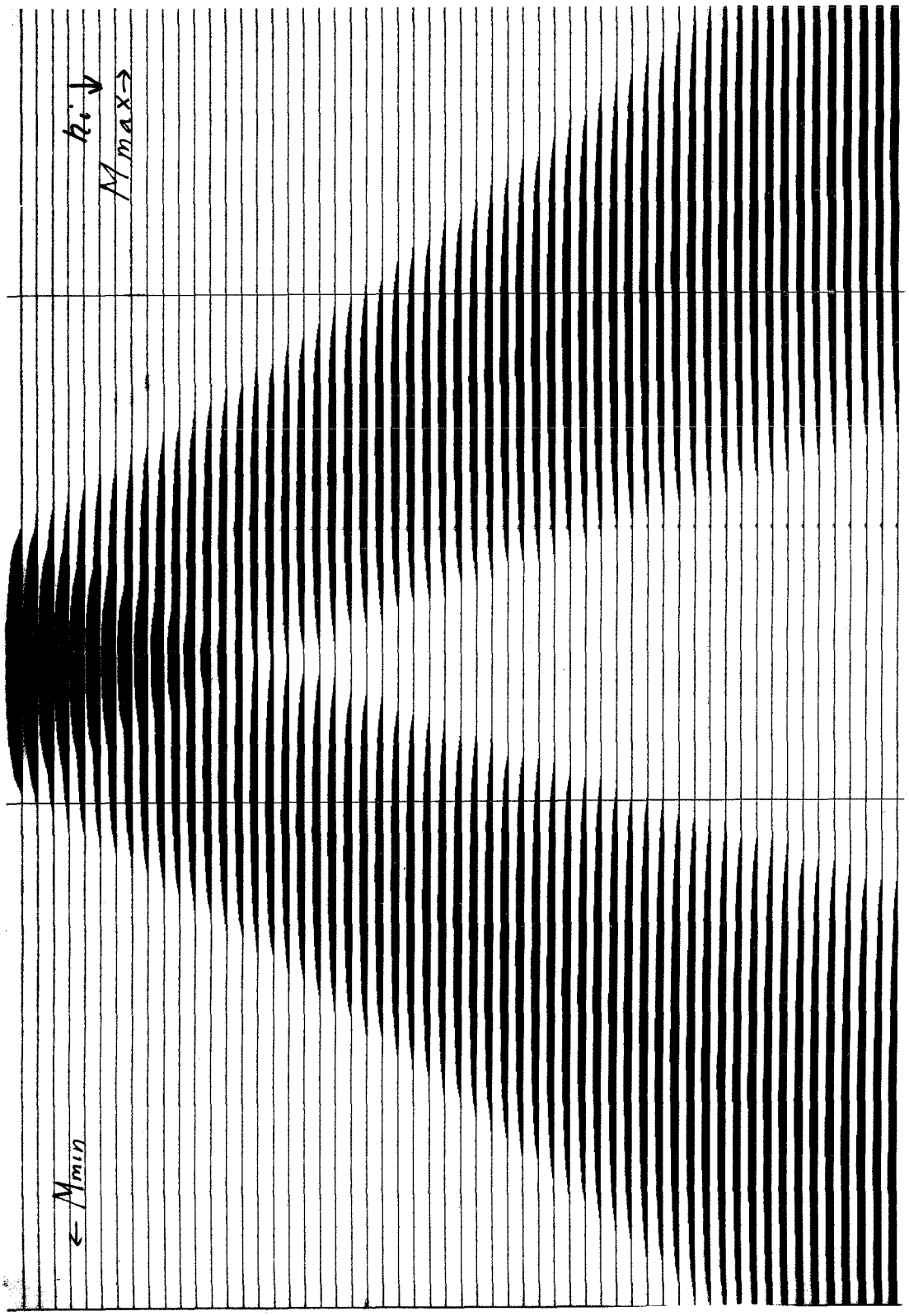


Figure 2.19a Travel Time Data-Constant Velocity
M min=-5, M max=5, h max=8.95

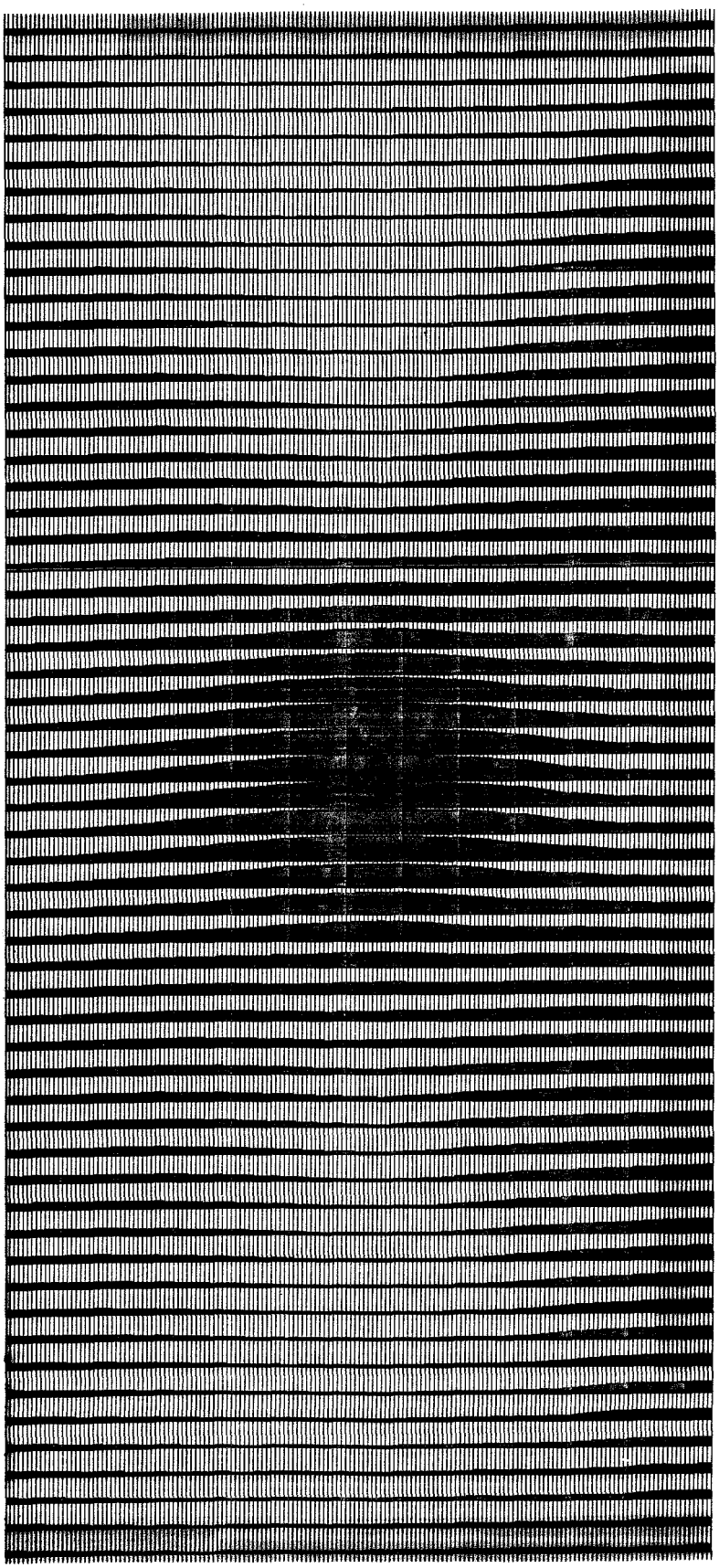


Figure 2.19b Back-Projected Field

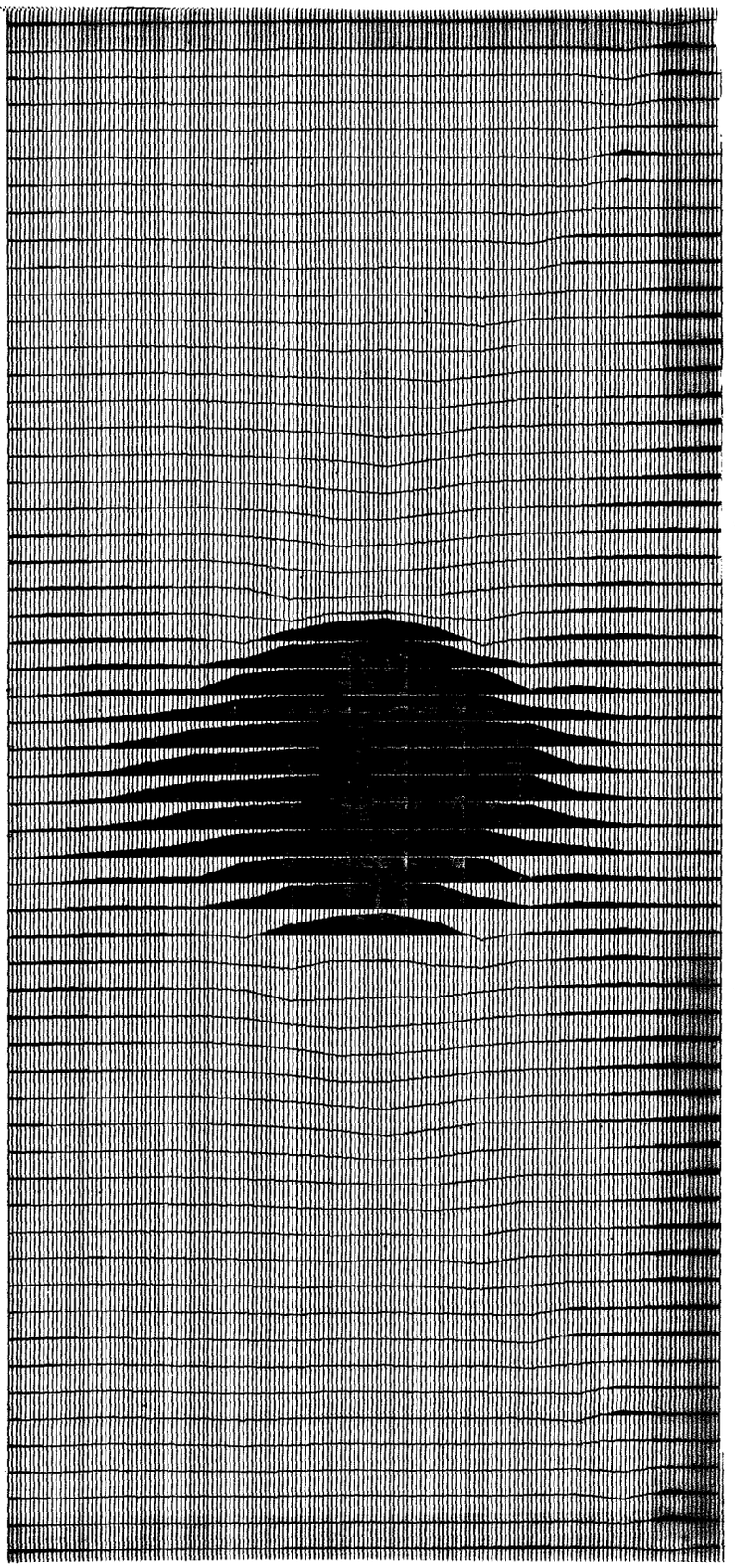


Figure 2.19c Filtered Back-Projected Field

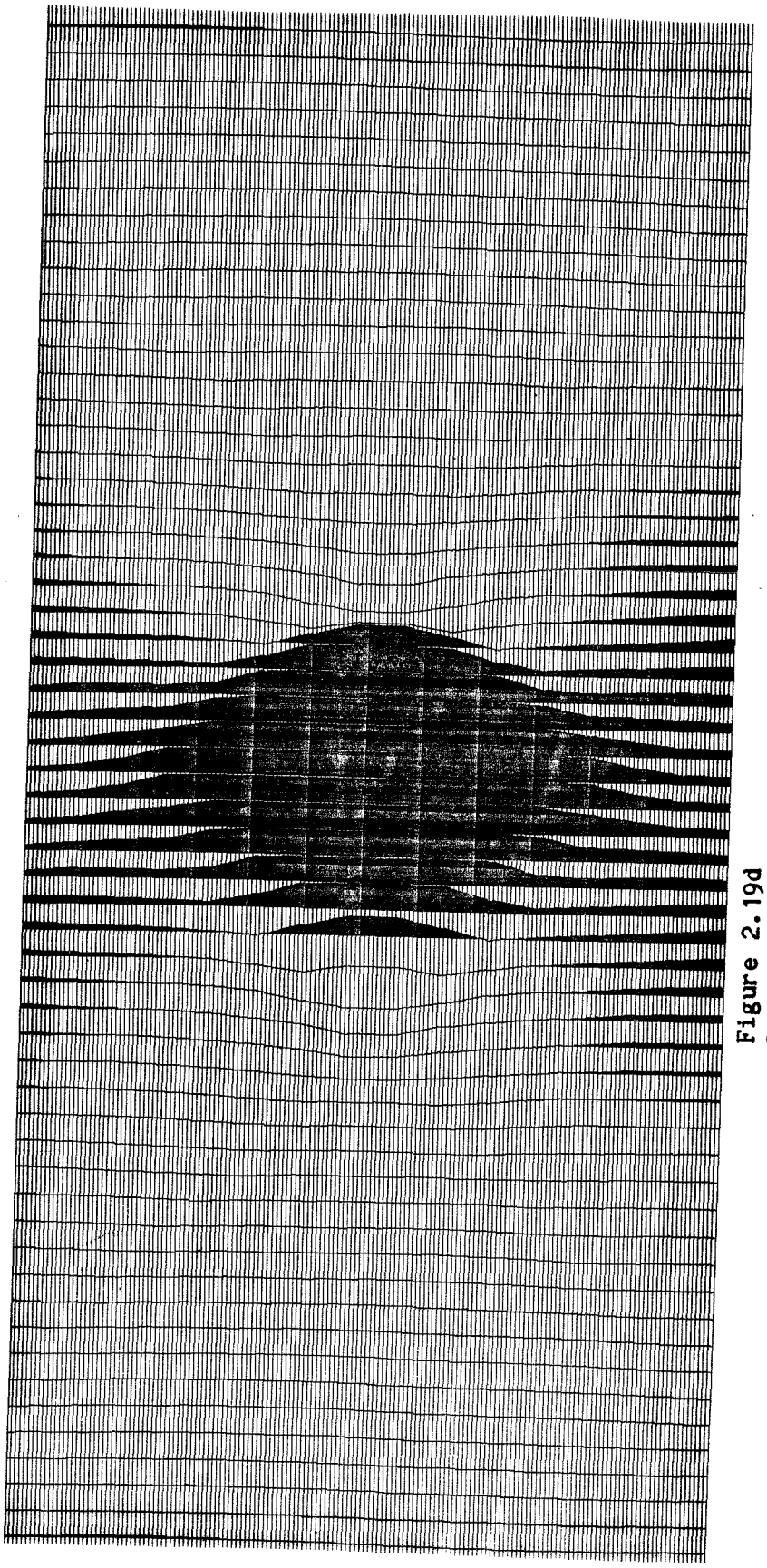


Figure 2.19d
Reconstructed Anomaly Field; h max=4.48

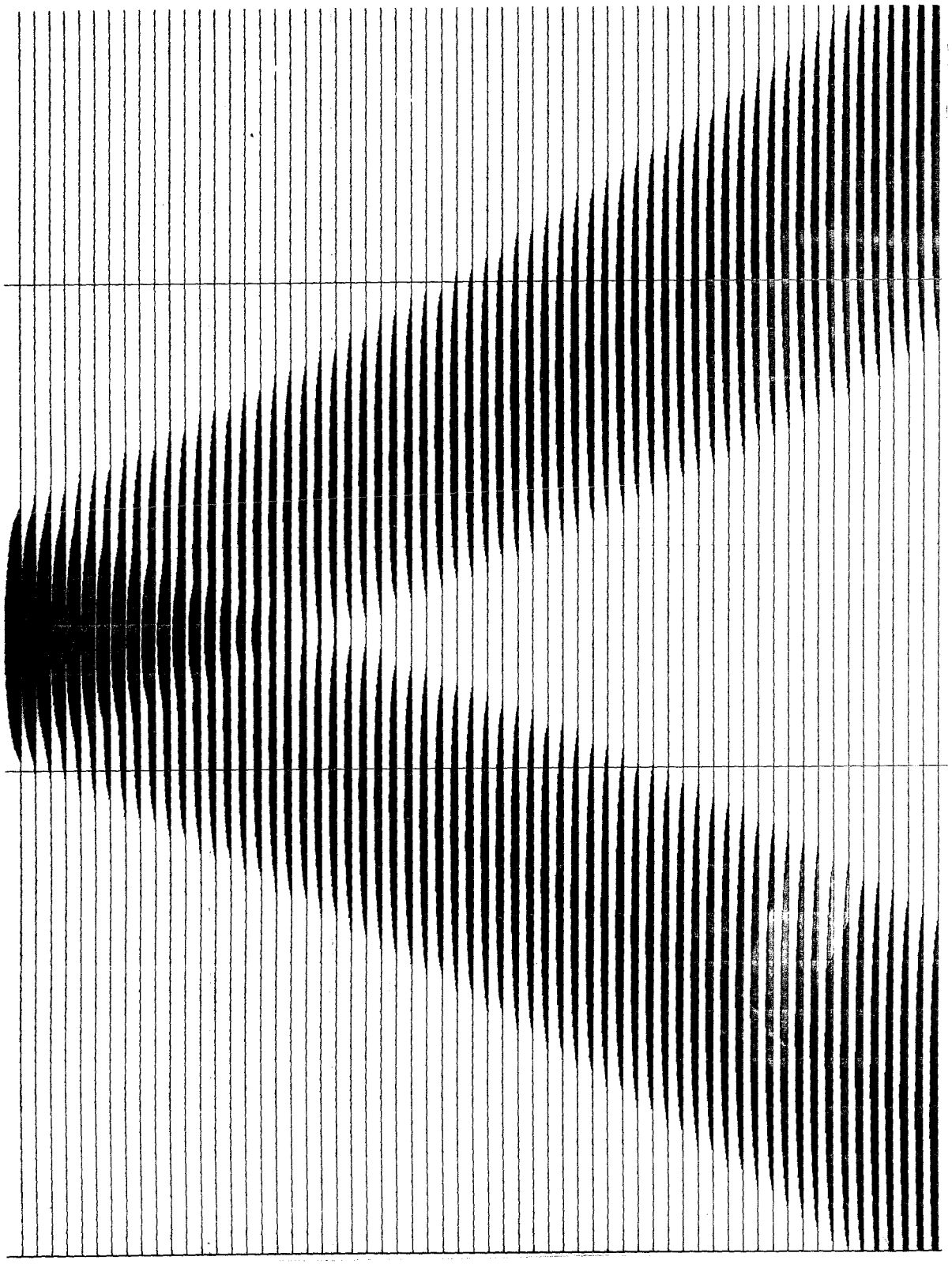


Figure 2.19e
Travel Time Data-C(y)=.2y+1
M min=-5, M max=5, h max=5.98

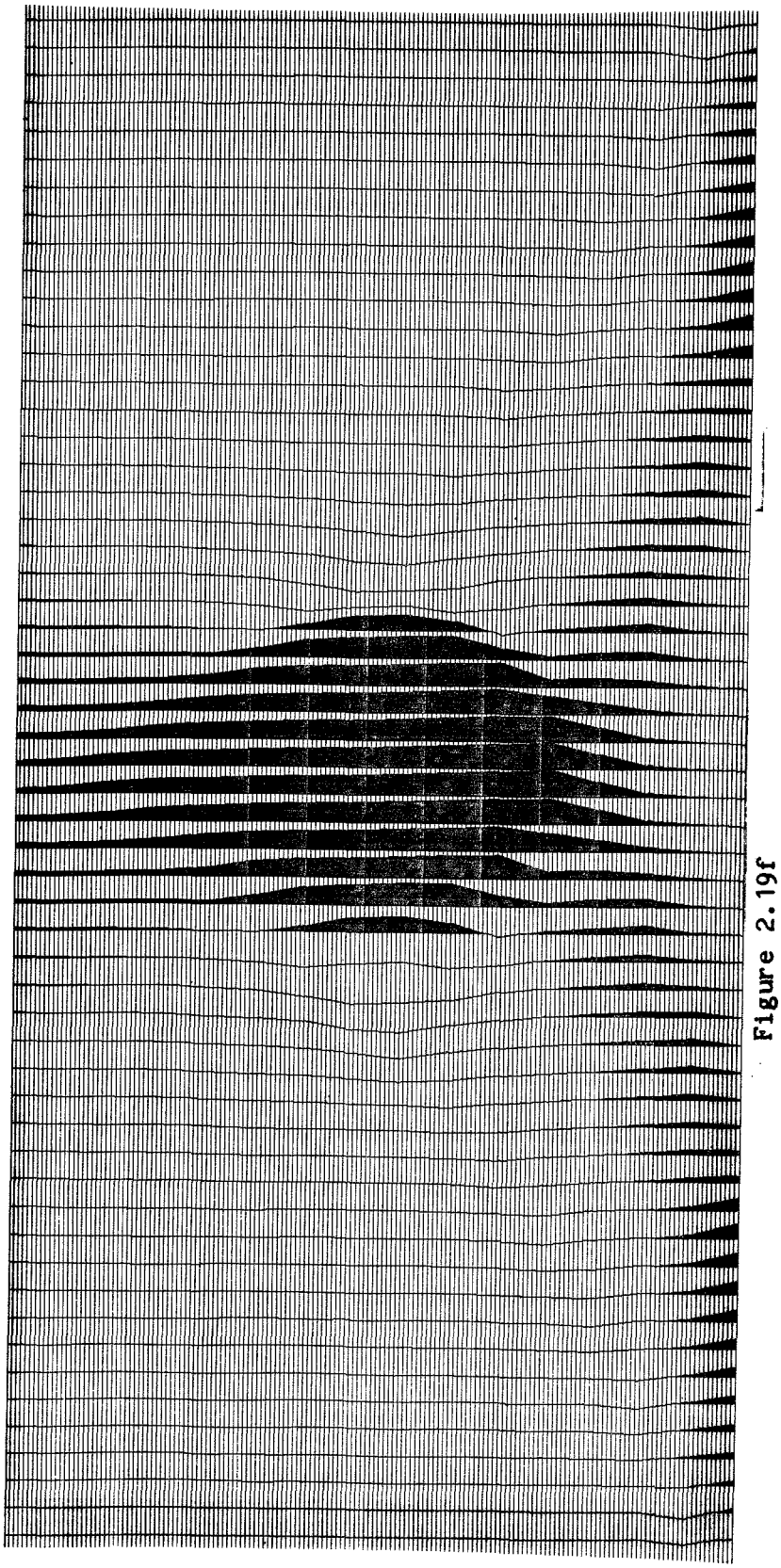


Figure 2.19f
Reconstructed Anomaly Field C(y)=.2y+1

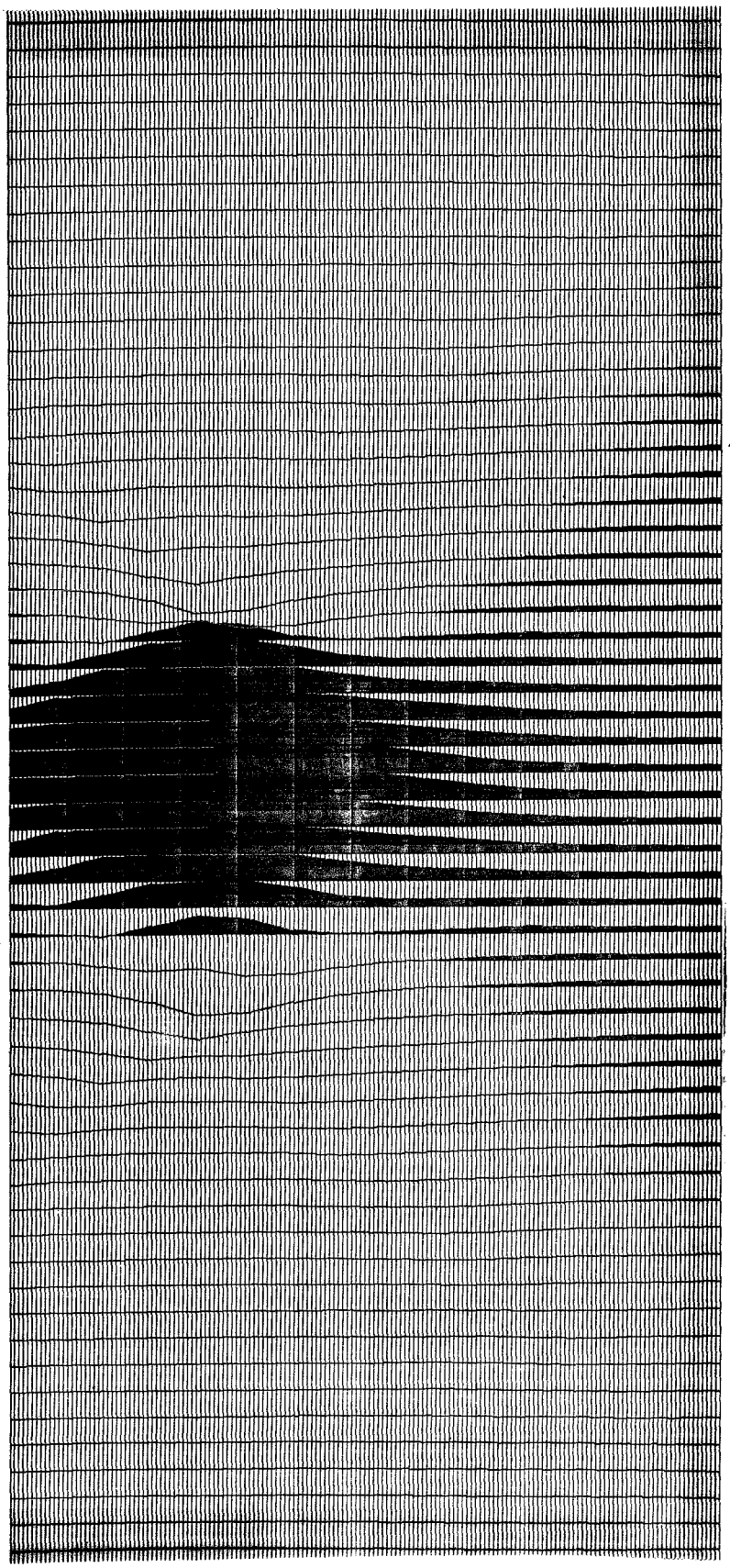


Figure 2.19g
Reconstructed Anomaly Field C(y)=y+1
-using constant velocity inversion formula

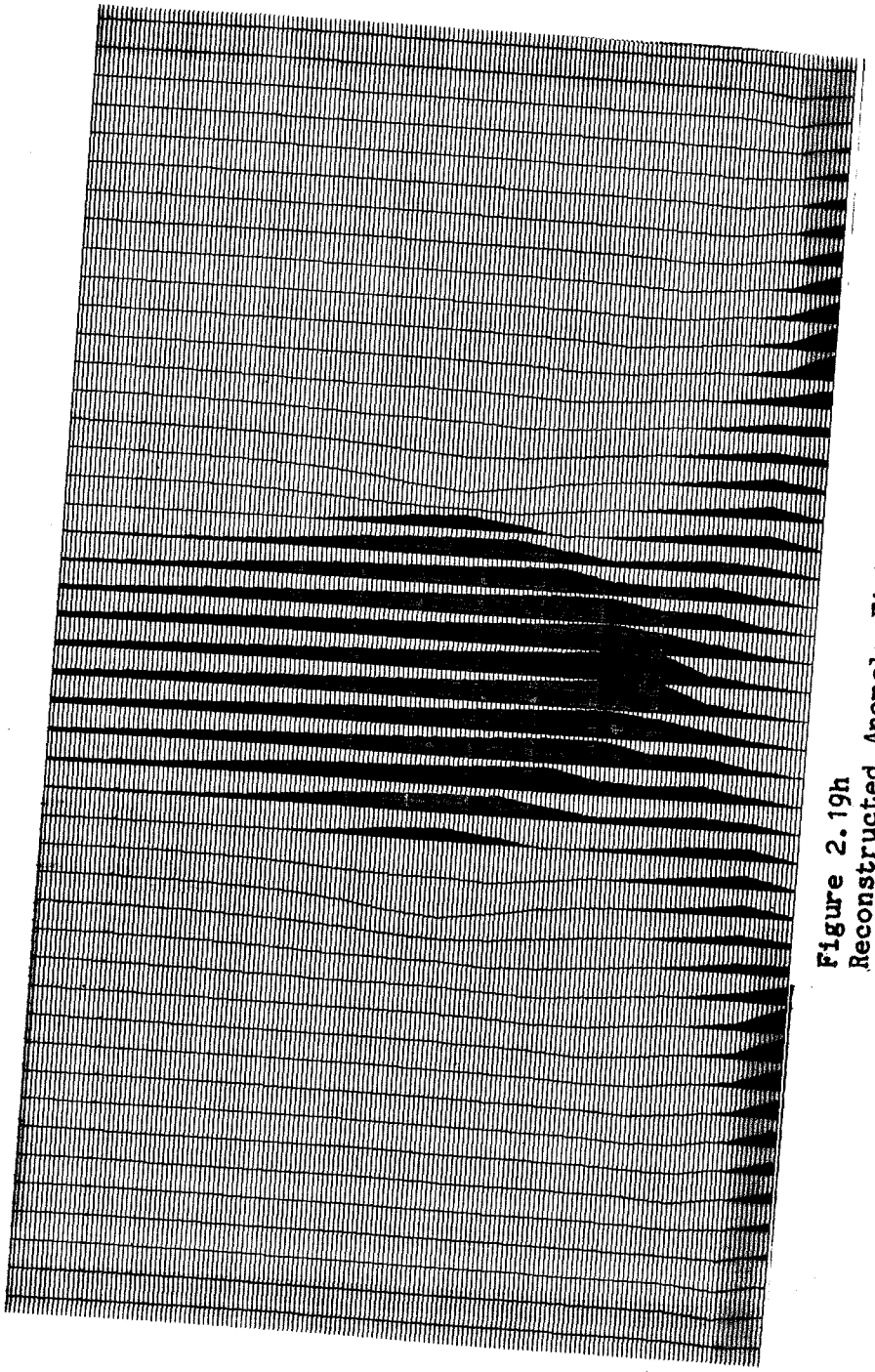


Figure 2.19h
Reconstructed Anomaly Field C(y)=y+1
-Using Inversion Formula (2.2a)

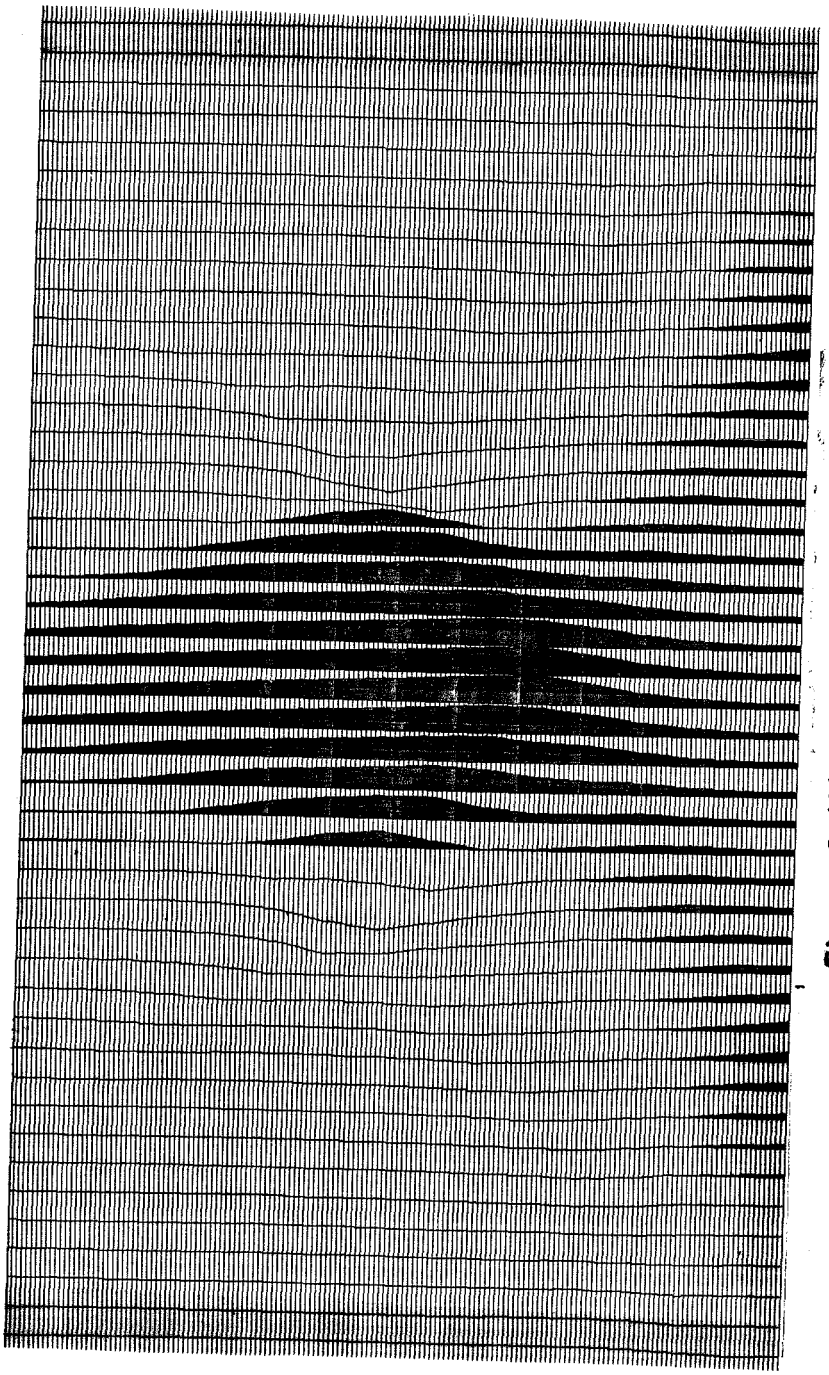


Figure 2.19i
Reconstructed Anomaly Field C(y)=y+1
-Using Inversion Formula (2.2b)

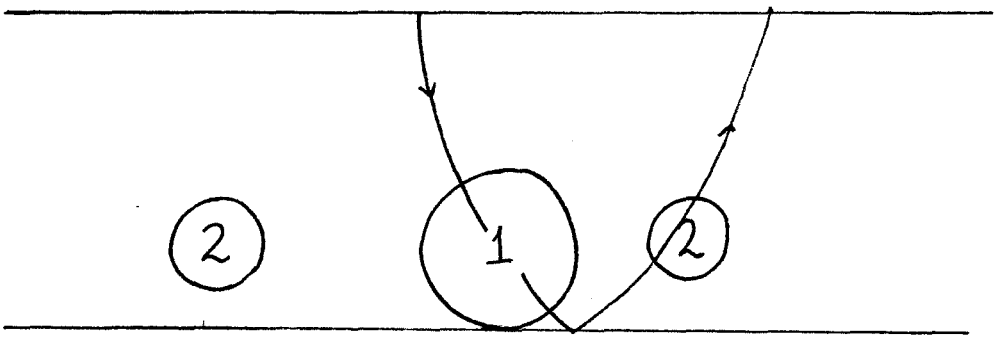


Figure 2.21 Anomaly Model

In figure 2.22a we show the generated travel time data for this problem. Figures 2.22b and 2.22c show the reconstructed anomaly fields using formulae (2.2a) and (2.2b) respectively. Once again, there is not a great deal of difference between the results of using (2.2a) or (2.2b). However, there are some slight differences; formula (2.2a) seems to be more accurate with the vertical resolution of the anomalies. This is to be expected, as formula (2.2a) treats the larger offset rays more correctly than does equation (2.2b).

Example 3.

In this example, the anomaly model consists of three disks, all of radius .5, which we show below in figure 2.23a.

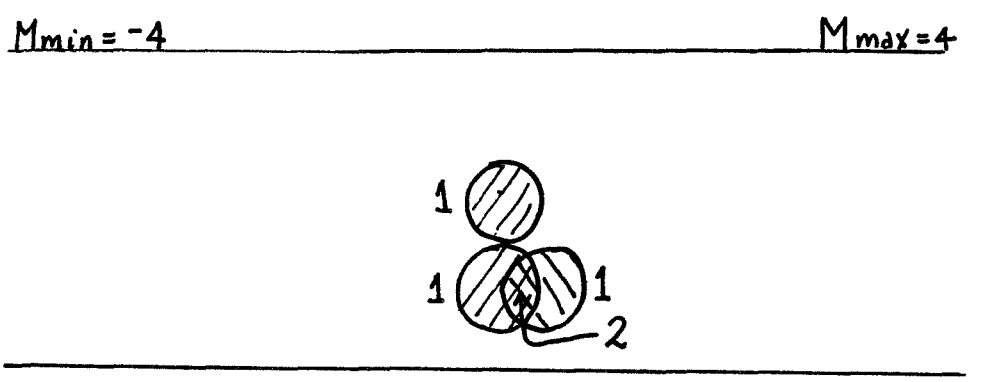


Figure 2.23a Anomaly Field Model I.

For the model shown in figure 2.23a, the disks' centres are at (x1=0,y1=3), (x2=0,y2=2), and (x3=.5,y3=3) and the disks all have slowness perturbations of one. The total anomaly extends over two depth units. The generated travel time data is shown in figure 2.24a and the resulting reconstructed field is shown in figure 2.24b. We now change the anomaly model to be that shown below in figure 2.23b; here we have positioned the overlapping disks higher in the field.

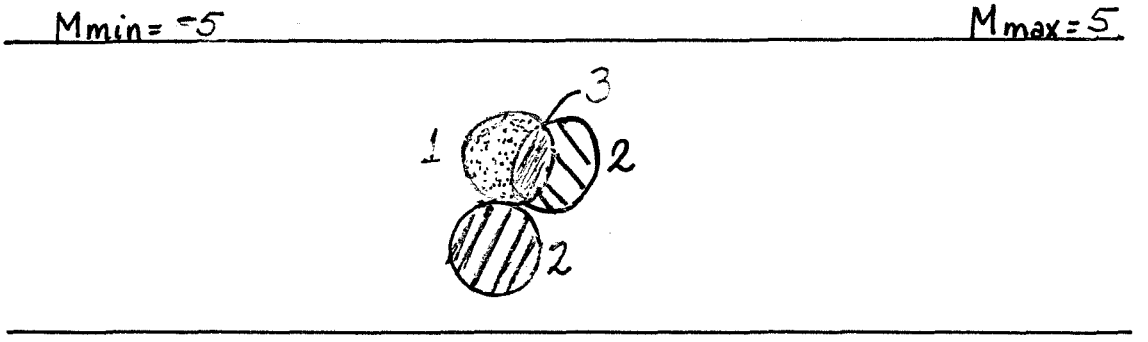


Figure 2.23b Anomaly Field Model II.

The resulting reconstruction is shown in figure 2.24c and it is seen that we have not done as well for this second model as for the first case.

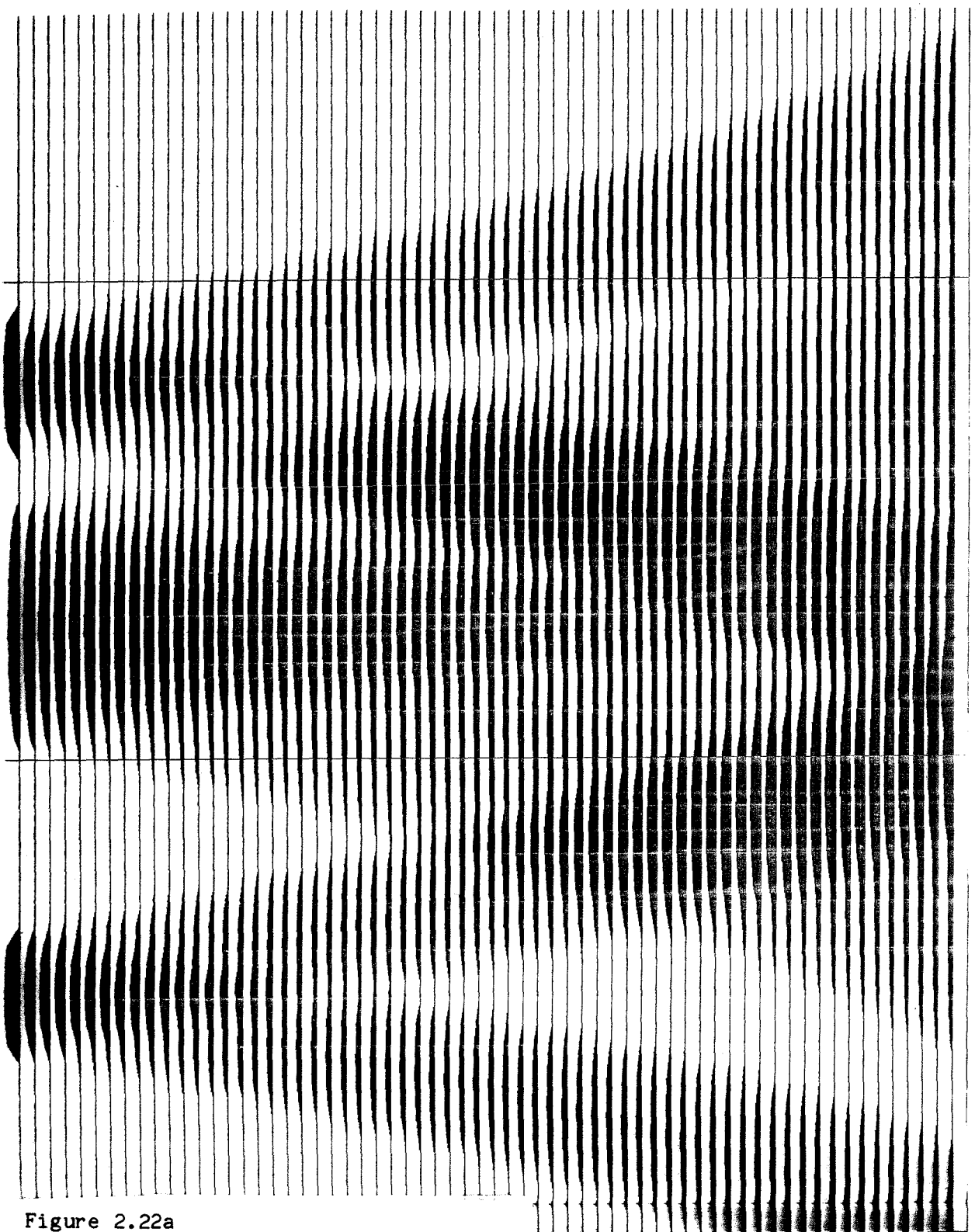


Figure 2.22a
Travel Time Data-3 Circles at Depth y=3.
C(y)=y+1

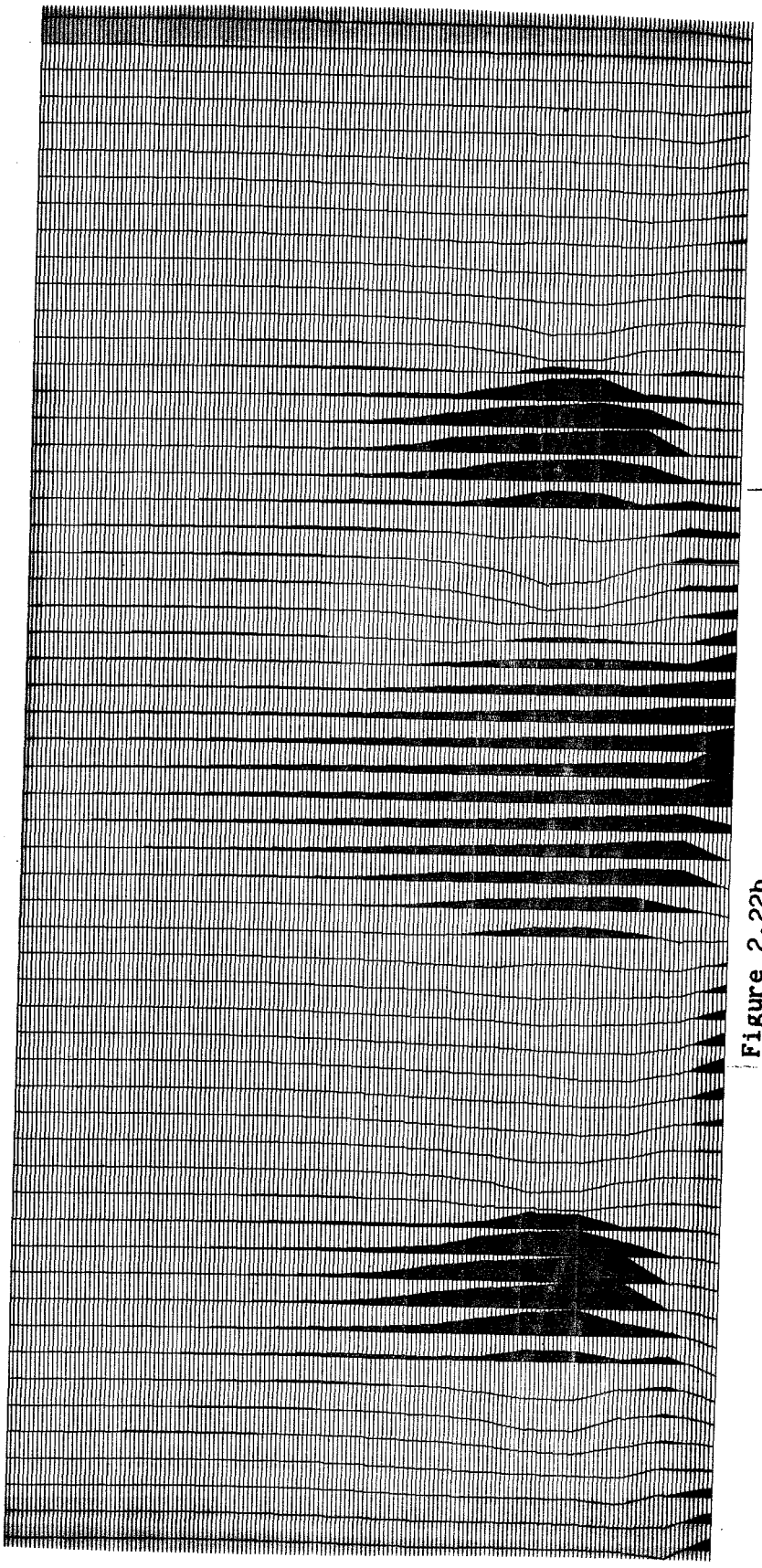


Figure 2.22b
Reconstructed Anomaly Field C(y)=y+1
-Using Inversion Formula (2.2a)

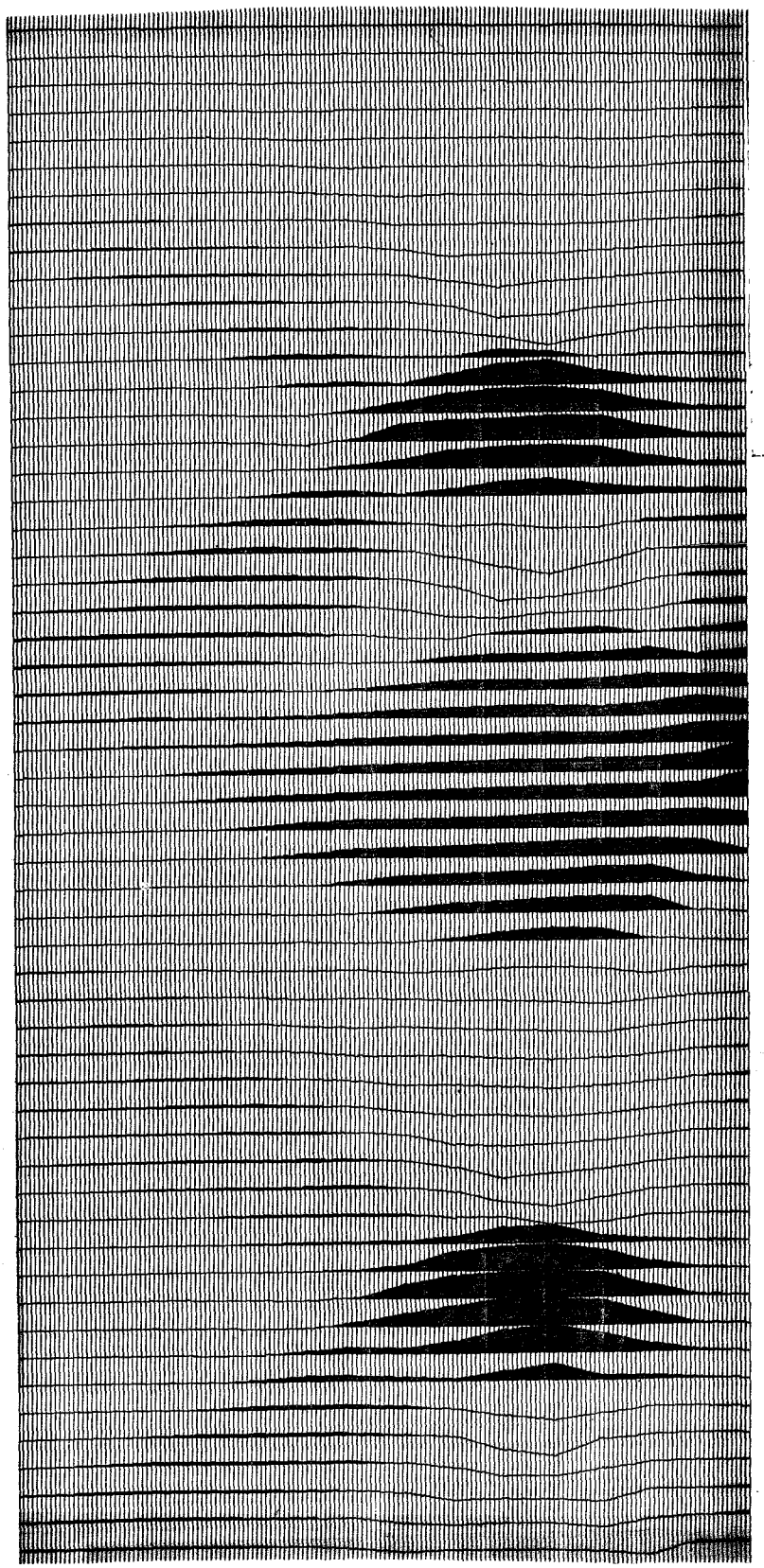


Figure 2.22c
Reconstructed Anomaly Field C(y)=y+1
-Using Inversion Formula (2.2b)

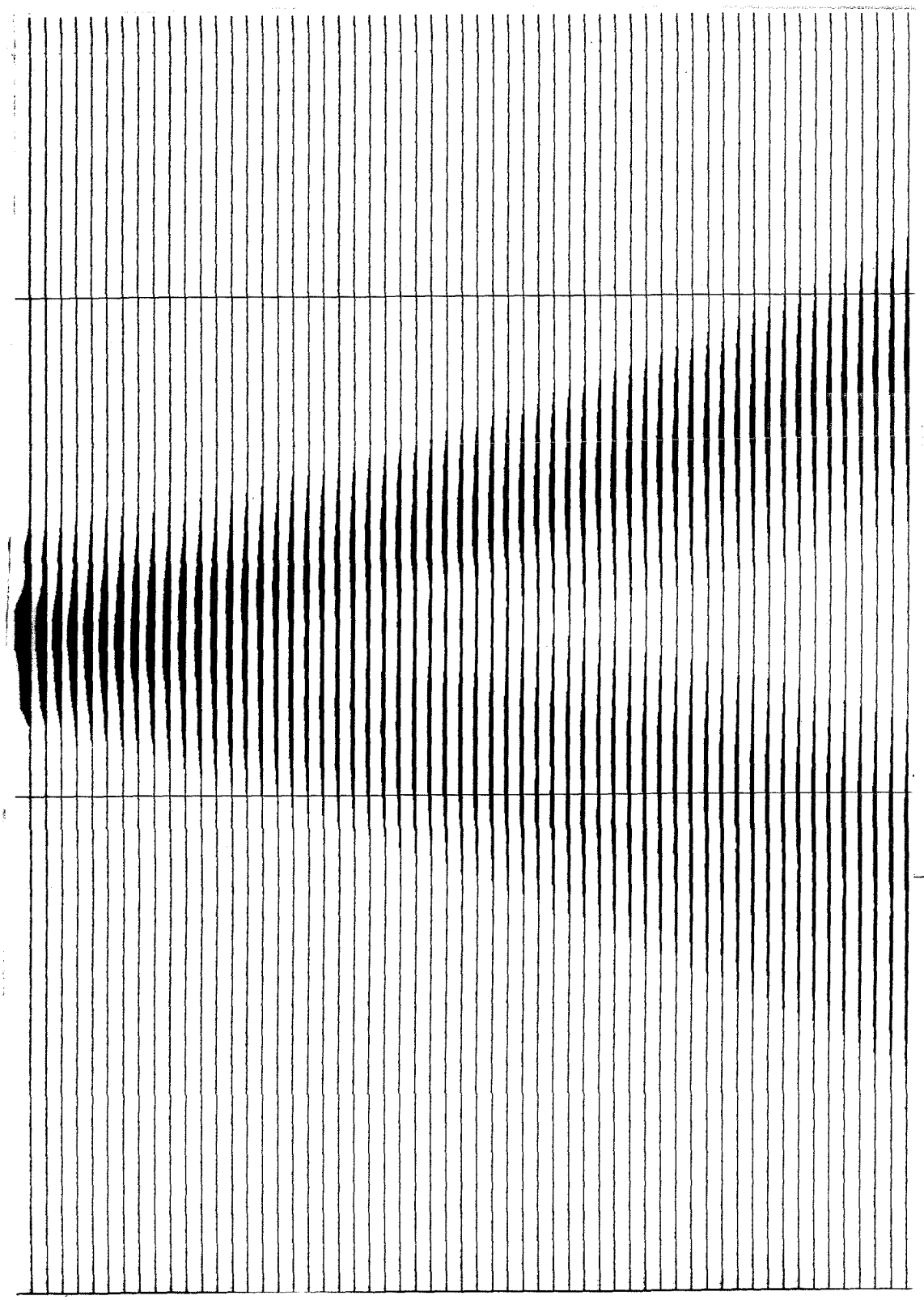


Figure 2.24a
Travel Time Data-Model of Figure 2.23a

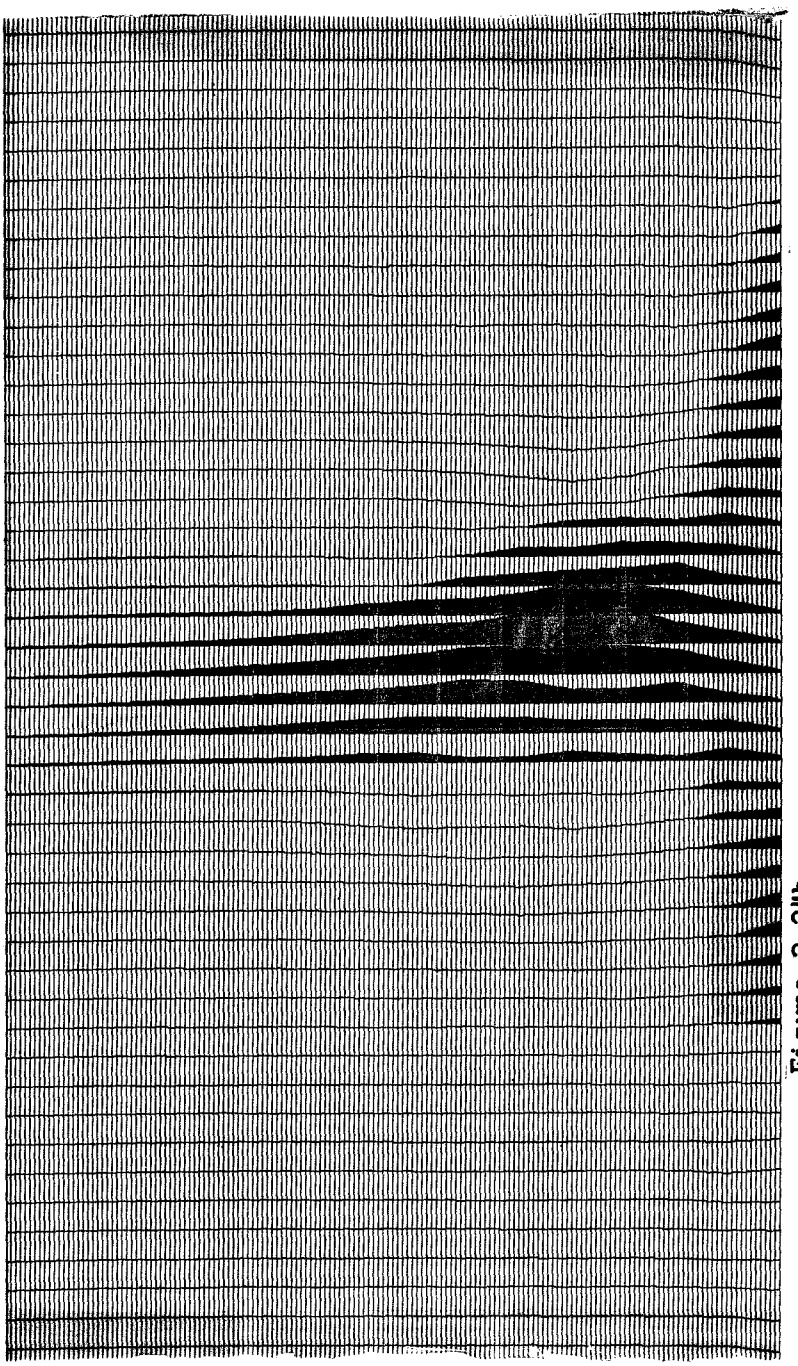


Figure 2.24b
Reconstructed Anomaly Field-Model of Figure 2.23a

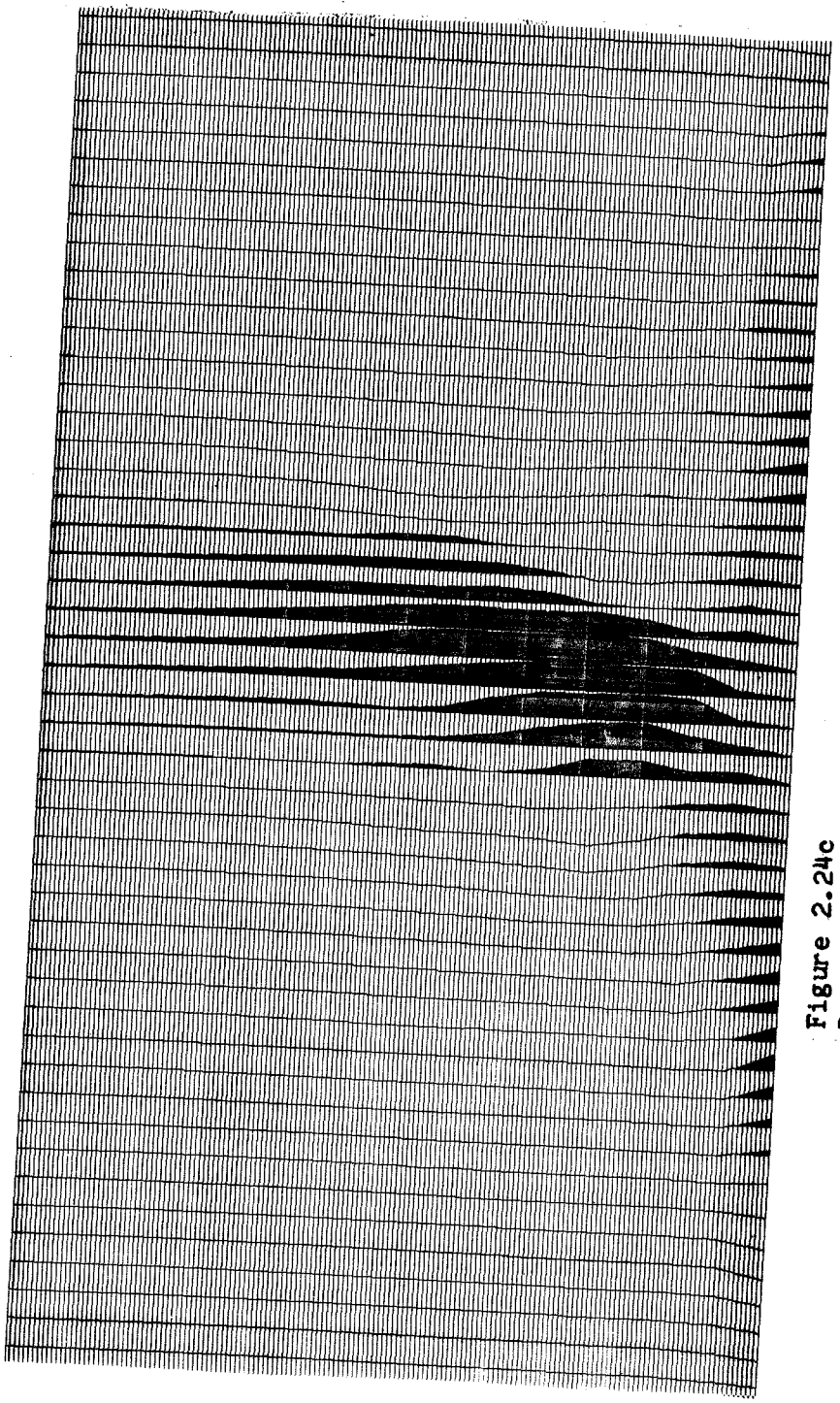


Figure 2.24c
Reconstructed Anomaly Field-Model of Figure 2.23b

References

1. K.Aki,A.Christofferson, and E.S.Husebye. "Determination of the Three Dimensional Seismic Structure of the Lithosphere," Journal of Geophysical Research., 82(1977), pp.277-296.

2. K.Aki and P.Richards. Quantitative Seismology, Theory and Methods Vol.II, W.H.Freeman, San Francisco,1980.

3. K.Bube and R.Burridge. "The One-Dimensional Inverse Problem of Reflection Seismology," preprint 1982.

4. R.Burridge. "The Gelfand-Levitan, The Marchenko, and The Gopinath-Sondhi Integral Equations of Inverse Scattering Theory Regarded in the Context of Inverse Impulse Response Problems," Wave Motion,2(1980), pp.305-323.

5. R.Clayton. Departmental Report, Department of Geological and Planetary Sciences, California Institute of Technology, 1982.

6. R.Clayton. Lecture Notes-"Signal Processing and Exploration Geophysics", Department of Geological and Planetary Sciences, California Institute of Technology, 1981-82.

7. R.Clayton and R.Stolt. "An Inversion Method for Acoustic Wave Fields," Stanford Exploration Project, 24(1980), pp. 57-80.

8. R.Clayton and R.Stolt. "A Born-WKBJ Inversion Method for Acoustic Reflection Data," Geophysics, 46, No.11 (Nov.1981), pp. 1559-1567.

9. S.Coen. "Velocity and Density Profiles of a Layered Acoustic Medium From Common Source-Point Data," Geophysics, 47, No.6(June 1982), pp.898-905.

10. M.G. Gasymov and B.M.Levitan. "Determination of a Differential Equation From Two of its Spectra," Russian Mathematical Surveys, 19(1964), pp.1-63.

11. I.M. Gelfand and B.M.Levitan. "On the Determination of a Differential Equation from Its Spectral Function,"(AMS Translation Series 2, 1(1955), pp.253-304.

12. B.Jacobs. "Non-Existence of a Gelfand-Levitan Coordinate System for the Wave Equation," Stanford Exploration Project, 26(1981), pp.197-204.

13. E.Kjartansson. "Attenuation of Seismic Waves in Rocks," Ph.D Thesis, Department of Geophysics, Stanford University 1980.

14. R.T.Prosser. "Formal Solutions of Inverse Scattering Problems II," Journal of Mathematical Physics., 17, No.10(Oct.1976),pp.1775-1779.

15. V.G.Romanov. Integral Geometry and Inverse Problems for Hyperbolic Equations., Springer, New York, 1974.

16. F.Santosa and H.Schwetlick. "The Inversion of the Acoustical Impedance Profile by Method of Characteristics," Wave Motion, 4(1982), pp.99-110.

17. R.H.Stolt and B.Jacobs. "Inversion of Seismic Data in a Laterally Heterogeneous Medium," Stanford Exploration Project, 24(1980), pp.135-152.

18. W.Symes. "Inverse Boundary Value Problems and a Theorem of Gelfand and Levitan," Journal of Mathematical Analysis and Applications, 71(1979), pp. 379-402.

19. W.Symes. "Numerical Stability in an Inverse Scattering Problem," SIAM Journal of Numerical Analysis, 17(1980), pp.707-732.

20. W.Symes. "Stable Solution of the Inverse Reflection Problem for a Smoothly Stratified Medium," SIAM Journal of Mathematical Analysis, 12(1981), pp.421-453.

21. G.B.Whitham. Linear and Nonlinear Waves. John Wiley, New York,1974.

22. Tables Relating to Mathieu Functions, Characteristic Values, Coeffeicents, and Joining Factors., U.S. Department of Commerce, National Bureau of Standards., 59,(1967).