ASYMPTOTIC WAVE METHODS IN HETEROGENEOUS MEDIA

by

Wafik B. Beydoun

S.M., Massachussetts Institute of Technology (1982)

M.S.T. Géophysique et Géotechniques, Univ. P.& M. Curie, Paris (1980)

SUBMITTED TO THE DEPARTMENT OF
EARTH, ATMOSPHERIC, AND PLANETARY SCIENCES
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

at the

May 2, 1985

Signature of Author.....
Department of Earth, Atmospheric, and Planetary Sciences
May 1985

Certified by.....
M. Nafi Toksoz
Thesis Advisor

Accepted by.....
Theodore R. Madden
Chairman
Departmental Committee on Graduate Students

Asymptotic Wave Methods in Heterogeneous Media

by

Wafik B. Beydoun

submitted to the Department of Earth, Atmospheric, and Planetary Sciences
on May 2, 1985 in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy in Geophysics

Abstract

The limitations of asymptotic wave theory and its geometrical manifestations are newly formalized and scrutinized in Chapter II. Necessary and sufficient conditions for the existence of acoustic and seismic rays and beams in general inhomogeneous media are expressed in terms of new physical parameters: the threshold frequency ω_0 associated with the P/S decoupling condition, the cut-off frequency ω_c associated with the radiation-zone condition, the total curvature of the wavefront and the Fresnel-zone radius. The analysis is facilitated with the introduction of a new ancillary functional - the hypereikonal which is capable of representing ordinary as well as evanescent waves. The hypereikonal is the natural extension of the eikonal theory. With the aid of the above new parameters, simple conditions are obtained for the decoupled far field, the decoupled near field, two point dynamic ray tracing, paraxial wave fields and Gaussian beams.

Chapter III deals with a canonical problem. The Green's function, in a constant gradient medium, is presented, for an explosive point source, in frequency and time domains. The analytical dynamic ray tracing (DRT) solution is re-derived with conditions stated in Chapter II. The Gaussian beam (GB) solution is investigated and new beam parameters are defined. Comparisons between exact and approximate solutions are made; for both methods, DRT and GB, conditions of validity are explicit and quantitative. An accuracy criterion is defined in the time domain, and measures a global relative error. The range of validity is expressed in the form of two inequalities for the dynamic ray tracing method and of five inequalities for the Gaussian beam method. Results remain accurate at ray turning points. For the type of medium considered, the breakdown of the dynamic ray tracing method is smoother and better behaved than that of Gaussian beams. As examples, a vertical seismic profiling configuration, and a shallow earthquake are modeled, using Gaussian beams.

Chapter IV describes the paraxial ray method, and its uses in modeling seismic waves. It is a flexible and fast method for computing asymptotic Green's functions. The method is an extension of the standard ray method, and a degenerate case of Gaussian beams. Accuracy is controllable, within ray and paraxial conditions. Comparison of results with finite difference and discrete wavenumber are very satisfactory. Examples for different heterogeneous media are shown.

A full-waveform inversion is then presented in Chapter V. A new approach, using tensor algebra formalism, is presented. Combined data sets (eg. VSP and surface

reflection data) with prior information are simultaneously handled. The forward model is generated by the paraxial ray method. The inversion is performed in the frequency domain, for interface and layer parameters. Sensitivity analysis is studied for each parameter. Data generated by finite difference is inverted and obtained estimates are accurate. VSP field data is inverted to estimate local geologic structure.

Thesis Advisor: M. Nafi Toksöz
Title: Professor of Geophysics

ACKNOWLEDGEMENTS

I am happy to acknowledge the help I have received throughout the work of this thesis. My advisor, Nafi Toksöz, provided me with guidance and direction during all my stay at MIT. His skeptical eye and solid sense of geophysics have been of invaluable help.

I express my gratitude to Ari Ben-Menahem for providing me the opportunity to work with him during his stay at MIT. Chapters II and III of this thesis were done in collaboration with him. His extensive experience in theoretical seismology and insights, along with his enthusiasm will always be remembered.

I have benefited from numerous discussions with Gilles Garcia. I learned about (exhaustive) quasi-Newton's methods for solving least squares problems, and hieroglyphs. He freely shared with me his stockpile of articles concerning the subject covered in Chapter V of this thesis.

I am grateful to Vernon Cormier and Roger Turpening, who have been a continual source of information and encouragement. I thank Greg Duckworth for reviewing the thesis and for his positive criticism, and Carol Blackway for spending her time acquainting me with the Burch data set, and providing me with information for the inversion of the field data. Interactions with Kei Aki, Michel Bouchon, Arthur Cheng, Raul Madariaga, Ted Madden, Ivan Pscencik, and Gene Simmons broadened my geophysical knowledge.

I am grateful to many of my colleagues over the years at MIT. To Tim Keho, for all the discussions, ranging from geopolitics to word decoding of WILD radio station songs, including science and sports. His humor and friendship are very appreciated. Numerous

fundamental discussions with Gengiz Esmersoy concerning wave versus ray methods, or symbol notations, will remain in memory. Dominique Dubucq, Christophe Gonguet, Marc Larrere, Frederic Mathieu, Michael Mellen, Bob Nowack, Fico Pardo-Casas, Benoit Paternoster, Michael Prange, Rob Stewart, Ken Tubman, Ru Shan Wu, Kiyoshi Yomogida, and many others, provided useful discussions.

Finally, I thank my parents, who encouraged me all the way through this long education, and for their altruistic guidance.

This work was supported by the M.I.T. Full Waveform Acoustic Logging and Vertical Seismic Profiling Consortia, the Compagnie Générale de Géophysique, and the Air Force Office of Scientific Research.

To my parents, and sister,

Helen, Bulind and Mouna

P.S. **KW, BP, NB.**

TABLE OF CONTENTS

Table of contents

Table of contents

I. GENERAL INTRODUCTION

Symbolism, language, scientific formulae here are all synonymous.

- J. Bronowski

The notion of waves is one of the unifying concepts of modern physics. Their physical manifestation can be of very different nature (elastic, electromagnetic, quantum mechanic, gravitational), but their behavior remains describable mathematically, in common terms. The basic properties of waves is that they carry energy from point to point in a medium, and like moving matter, they have velocity and momentum.

Seismology deals with the generation and propagation of elastic waves in the earth. The data are seismograms, which is a measure of the disturbance caused by the wave during its passage through observers (seismographs) placed in contact with the earth. Typically these are records of particle displacement, or particle velocity, or particle acceleration, or pressure (in a fluid medium), as a function of time. The recorded seismic wavefield depends generally on three main mechanisms: (1) the wave generation (the source of energy), (2) the propagation through the complex media (scattering, diffraction, attenuation etc.), and (3) the seismograph measurement bias (quality of the coupling with the earth, partial information if it has less than three spatial components of recording, transfer function, etc.). These seismic records are very valuable if information regarding any of these three mechanisms is sought. Most importantly, they provide means of probing the earth interior. The advent of computers revolutionized seismology; data acquisition has improved enormously and sophisticated data processing is now possible. It is the second great tool, after analytical

mathematics, for theoretical developments.

Modeling

Modeling is a very useful approach in understanding seismic wave propagation in heterogeneous media and consequently helping the interpretation of data. For example, assume we have a specific earth structure in which we place sources and receivers. From the physical laws governing the behavior of seismic waves, one can determine, analytically or numerically, "synthetic" seismograms of the earth model. Modeling methods for complex media are numerous. There is no universal method that is applicable to all media. Each method is best suited for a given model, depending on the model's structure, source/receiver configuration, and the allocated computer time. Further, each method has its own assumptions and validity conditions that must be fully understood before its practical use.

Hermann and Wang (1985), compare synthetics seismograms of several methods developed for plane-layered media, and list references of basic methods that deal with this problem. These are generally full-wave methods, in the sense that the full effects of the media are simulated and recorded. They are computationally expensive, and cannot be simply extended to handle more complicated geometries. Another set of full-wave methods, able to handle diffractions from sharp interfaces, concerns boundary integral methods. The media is limited to homogeneous layers separated by arbitrarily smooth interfaces, although extension for complex media can be made via asymptotic wave methods (see below). The methods treat each interface separately and with the continuity conditions at the interface, integral equations are set up for the wavefield. References for these methods can be found in Kennett and Harding (1985). Methods which consist in solving numerically the differential elastic wave equation in complex media, are called direct numerical methods. They are full-wave methods, and are

expensive computationally. Finite difference (Alford *et al*, 1974; Kelly *et al*, 1976; Stephen, 1984) and finite elements (Marfurt, 1978) are the two main techniques.

The last set of methods is based on simple geometrical considerations. They are variations of ray theory, and will be called more generally, asymptotic (high frequency) wave methods. They are versatile, flexible and are generally used either directly in modeling, or indirectly in other methods requiring approximate Green's functions in heterogeneous media. These methods are not full-wave but are explicit, in the sense that individual wave types can be propagated separately. This offers the possibility of constructing progressively the full-wave character of the field. However if many wave types are sought, variations of these methods (Keller and Perozzi, 1983) must be considered.

Hybrid methods, combining advantages of compatible methods, are now under progress. Modeling real earth structures which would handle (1) multiples, diffraction and scattering effects, (2) critical region effects (caustics, shadow zones, etc.), and (3) interface waves (surface, head, etc.) is not an impossible task. Asymptotic wave methods are compatible with boundary integral methods, and can be an example of such a hybrid method.

The aim of this thesis is to establish, in some quantitative manner, the range of validity of asymptotic wave methods. Necessary conditions in obtaining Helmholtz wave equations from the elastodymanic wave equation are explicit. A hypereikonal is introduced that leads naturally to the ray solution for high frequencies. The focus is on dynamic ray tracing, Gaussian beam methods and, particularly on an intermediate method, called paraxial ray. We will not cover Maslov asymptotic theory (Chapman and Drummond, 1982), which can be obtained as a degenerate case of Gaussian Beams (Madariaga, 1984; Klimes, 1984). The paraxial ray method is shown to be a fast,

flexible and robust method. It is the method that we choose for modeling.

Inversion

The other fundamental task in seismology, is the extraction of medium or source parameters from data. Given a data set, estimating the parameters, require some assumptions on their prior values and constraints. Numerous approaches exist and, here again, each method has its own limitations. The three basic sets of methods are direct inversion, approximate direct inversion, and iterative inversion. They all yield, as end product, an image of the subsurface, and are sometimes referred to in general as inversion or imaging methods.

Direct inversion methods require the solution of an inverse operator such that when applied to observed data, it reconstructs medium properties exactly. These operators are difficult to obtain, particularly for more than one space dimension. Approximations, such as the Born approximation, simplify the problem and enable a direct solution to be found. A review of these methods with references can be found in Esmersoy (1985). They generally assume the inhomogeneity to be included in a uniform background medium for which the Green's function is known exactly. However, extension to more complex background media can be done using asymptotic wave methods that provides approximate Green's functions.

Iterative inversion requires, generally, a forward model that is repeatedly used to generate synthetics. These synthetics are compared to data, and within a defined norm, the task is to minimize the norm of the data-synthetic discrepancy vector for medium parameters. Basic references on the subject are Beck and Arnold (1977), Luenberger (1973) and Aki and Richards (1980). Travel-time inversion has been, so far, widely used in seismology. But with the improvement in data acquisition systems, full-waveform

inversion can now be attempted, since more information about the medium or source parameters is added. Examples of full-wave inversion for three dimensional (3D) medium parameters are found in Thomson (1983); for source parameters in Nabelek (1983); and for 1D medium parameters with vertical seismic profiling (VSP) in Stewart (1983).

The last chapter of this thesis deals with full-wave iterative inversion of combined VSP, surface reflection, multi-offset VSP or crosshole data. Forward model synthetics are generated by the paraxial ray method. The norm considered is the L_2 or "energy" norm. The problem is presented within the framework of tensor algebra. The Gauss-Newton method is re-derived in this context. The hèterogeneous media contains homogeneous layers separated by smooth interfaces. Sensitivity of the inversion to medium parameters is studied. Inversion of field data is presented.

NOMENCLATURE OF CHAPTERS II-IV

A_0	amplitude function of local plane wave
B	parameter governing initial half-width of Gaussian beam
C	constant factor calibrating ray methods to an exact solution
c_0	point source radiation pattern including source strength
c_1	source strength, ratio c_0 to source radiation pattern
e_A	relative error of maximum amplitude of approximate signal
e_{ph}	relative error of maximum amplitude time of approximate signal
e_τ	relative error of travel time of approximate signal
E	relative power error of approximate signal on the entire trace
E^W	E on wavelet time window
F_0	$c_1/\rho_0^{1/2}$, strength of point source (dimensionless)
f	general source radiation pattern including source strength
FFC	far field amplitude condition
$FRN1$	Fresnel 1-condition
$FRN2$	Fresnel 2-condition
$\vec{g}_\alpha$	P wave coupling vector
$G(\vec{r}\,\vert\,\vec{r}_0)$	scalar Green's function
G_r	radial Green's function in (r,z)
G_z	vertical Green's function in (r,z)
HFC	high frequency condition
h	η^{-1}, vertical distance between $\alpha(0)$ and $\alpha=0$
I	unit dyadic (tensor)
J	Jacobian of cartesian to ray coordinates

Nomenclature

J^S	surficial Jacobian of cartesian to ray coordinates
j_0	ray parameter, takeoff angle, bi-spherical coordinate
j	ray angle of incidence
K	Gaussian wavefront curvature
K_R	R^{-1}, local ray curvature
K_w	$\alpha P/Q = R_w^{-1}$, total curvature of wavefront (or phase front)
K_w^b	$\alpha P/Q$, local curvature of beam wavefront (complex)
k	ω/α, P wavenumber
l_0	α/ω_0, medium's characteristic length
L	local half-width of Gaussian beam
m_0	source seismic moment
M	receiver / observer location
M_0	point source location
NHI	non-horizontal incidence of ray condition
n	cartesian coordinate perpendicular to the ray
p	$\sin j/\alpha$, ray horizontal slowness, ray parameter
$\hat{p}$	$\equiv\hat{\xi}$, unit vector tangent to the ray
P	$\alpha^{-1}\, dQ/ds$, functional in the eikonal
PRX	paraxial ray condition
Q	functional in the eikonal
r	radial distance from vertical z axis
r_t	$(z_0+h)\cot j_0$, radial coordinate of turning point of ray
RCC	regularity of ray-centered coordinate system condition
R	local ray radius of curvature
R_w	local wavefront radius of curvature
s	arclength along a specified ray

Nomenclature

S	hypereikonal
t	time
T	generalized travel time
$\vec{U}$	particle displacement vector in time or in frequency
u_ξ	Green's function along ξ in (ξ,z)
u_z	Green's function along z in (ξ,z)
v	P or S intrinsic velocity of medium
V_F	Fresnel volume
W	P wave vertical displacement in (r,z) coordinate system
W^b	W computed by the Gaussian beam method
W^r	W computed by the dynamic ray tracing method
z	vertical axis
z_t	$R-h$, depth of turning point of ray
α	compressional (P) wave velocity of the medium
α_0	P wave velocity at the source location
β	shear (S) wave velocity of the medium
γ	α^2/β^2
$\gamma_{1,2}$	ray parameters
Δ	source-receiver distance in a homogeneous medium
ε	cumulative error along the ray
ε_α	P wave coupling scalar
Λ	wavelength
λ	First Lame elastic parameter
μ	Second Lame elastic parameter, shear modulus
ξ	τ/t_0, bi-spherical coordinate, curvilinear coordinate along a ray
$\hat{\xi}$	unit vector tangent to the ray

Nomenclature

ρ	density of the medium
ρ_0	density of the medium at the source location
$\rho_{1,2}$	principal radii of wavefront curvature
σ	point source time function
$\overline{\sigma}$	point source spectrum
τ	source - receiver travel time
Φ	weight function for the Gaussian beam
φ_1	P wave potential
φ_2	S wave potential
ψ	Helmholtz potential
ψ^{par}	Parabolic potential
ω	angular frequency
ω_0	medium's threshold frequency
ω_c	medium's cut-off frequency

II. RANGE OF VALIDITY OF RAY AND BEAM METHODS

Nature is not a gigantic formalizable system. In order to formalize it,
we have to make some assumptions which cut out some parts.
We then lose the total connectivity.

- J. Bronowski

1. INTRODUCTION

Asymptotic wave theory is subjected to a number of fundamental restrictions which put severe limitations on its applicability. These restrictions involve three types of physical parameters: frequency, distance and gradients of structural elements. The conditions are formulated in the form of inequalities. If any of these inequalities is violated, ray theory becomes (progressively or rapidly) invalid and we must resort to the full-wave theory or other valid approximations. The limitations of ray theory fall into several categories, each arising in connection with a different type of asymptotic approximation. Each category renders its own conditions for the validity of seismic ray theory.

There exist today a few basic *ray-methods* which are interrelated. The oldest is the eikonal method (e.g., Born and Wolf, 1964; Luneburg, 1964) upon which the entire field of geometrical optics is based. It became a very useful tool in studies of seismic wave propagation (e.g., Singh and Ben-Menahem, 1969, Cerveny *et al.*, 1977; Aki and Richards, 1980; Cerveny and Hron, 1980) Next came the WKBJ and the saddle-point approximations at high frequencies (e.g., Bremmer, 1951) with some new development

and variants (e.g., Kravtsov, 1968, and Kravtsov and Feizulin, 1969, Chapman, 1978). In recent years, the advent of computers allowed seismologists to look for fast and efficient algorithms to calculate collimated wave fields in inhomogeneous media. The paraxial wave approximation, especially in a ray-centered coordinate system was found to offer some advantages (e.g., Landers and Claerbout, 1972; Corones, 1975; McDaniel, 1975; Palmer, 1976, 1979; DeSanto, 1977; McCoy, 1977; Bastiaans, 1979; Cerveny *et al.*, 1982; and Haus, 1983)

In addition, certain efforts were made to harness the Kirchhoff-Helmholtz integral (e.g., Baker and Copson, 1939) to the evaluation of seismic fields (e.g. Kravtsov and Feizulin, 1969; Scott and Helmberger, 1983; and Carter and Frazer, 1984) There is now a growing need, both in earthquake seismology and in seismic exploration, for computational methods that can render sufficiently accurate solutions to wave propagation problems in three dimensions. In this chapter we shall examine the validity of the various approximations involved in the asymptotic wave theory.

2. DECOUPLING CONDITION FOR GENERAL INHOMOGENEOUS MEDIA

The elastodynamic vector equation for isotropic media, in the absence of body forces, can be put in the compact form

$$\rho \frac{\partial^2 \vec{U}}{\partial t^2} = \nabla[(\lambda + 2\mu)\nabla \cdot \vec{U}] - \nabla \times (\mu \nabla \times \vec{U}) + 2\nabla\mu \cdot \nabla \times (I \times \vec{U}) , \qquad (2.1)$$

where I is the unit dyadic, $\vec{U}$ denotes the particle displacement vector at point $M(\vec{r})$ and time t, ρ is the density, and λ and μ are the Lamé elastic parameters of the medium. $I \times \vec{U}$ is the vector product of the unit dyadic and the displacement vector and

$$\nabla\times(I\times\vec{U}) \equiv \frac{1}{2}I\times\nabla\times\vec{U} - I\nabla\bullet\vec{U} + \frac{1}{2}(\nabla\vec{U} + \vec{U}\nabla).$$

Introducing the notation

$$\rho\alpha^2 = \lambda + 2\mu, \quad \rho\beta^2 = \mu, \tag{2.2}$$

and applying the vector identities

$$\nabla(\psi\varphi) = \psi\,\nabla\varphi + \varphi\,\nabla\psi, \tag{2.3}$$

$$\nabla\bullet(\rho\vec{A}) = \rho\nabla\bullet\vec{A} + (\nabla\rho)\bullet\vec{A},$$

$$\nabla\times(q\vec{B}) = q\,\nabla\times\vec{B} + (\nabla q)\times\vec{B},$$

we may recast Eq (2.1) in the form

$$\frac{\partial^2\vec{U}}{\partial t^2} = \nabla(\alpha^2\nabla\bullet\vec{U}) - \nabla\times(\beta^2\nabla\times\vec{U}) + \frac{\nabla\rho}{\rho}(\alpha^2\nabla\bullet\vec{U}) - \frac{\nabla\rho}{\rho}\times(\beta^2\nabla\times\vec{U}) \tag{2.4}$$

$$+ 2\beta^2\left(\frac{\nabla\beta^2}{\beta^2} + \frac{\nabla\rho}{\rho}\right)\bullet\nabla\times(I\times\vec{U}).$$

Taking the divergence of Eq. (2.4), applying to it the Fourier transform over t and defining

$$N = \alpha^2\nabla\bullet\vec{U}, \tag{2.5}$$

$$\vec{A} = \beta^2\nabla\times\vec{U}, \tag{2.6}$$

$$\vec{g}_\alpha = \frac{\nabla\alpha^2}{\alpha^2}, \quad \vec{g}_\beta = \frac{\nabla\beta^2}{\beta^2}, \quad \vec{g}_\rho = \frac{\nabla\rho}{\rho}; \qquad k_\alpha = \omega/\alpha. \tag{2.7}$$

it is possible to transform Eq. (2.4) into

$$\nabla^2 N + \vec{g}_\rho\bullet\nabla N + k_\alpha^2 N = (2\vec{g}_\beta + \vec{g}_\rho)\bullet\nabla\times\vec{A} - \vec{W}, \tag{2.8}$$

where

$$\vec{W} = N\nabla\bullet\vec{g}_\rho - \vec{A}\bullet\nabla\times\vec{g}_\rho + \nabla\{2\beta^2(\vec{g}_\beta+\vec{g}_\rho)\} : \nabla\times(I\times\vec{U}) + 2\vec{g}_\rho\bullet\{\vec{g}_\beta\times\vec{A}\}. \tag{2.9}$$

The advantage of writing Eq (2.1) in the form (2.8) is obvious: It separates the terms of

the equation into three distinct groups according to the order of the derivatives of the constitutive parameters ρ and β. At this point we discard the vector $\vec{W}$ which is composed solely of *second order terms in* $\vec{g}_\rho$ and $\vec{g}_\beta$ as compared with $\vec{g}\cdot\nabla N$ and $(2\vec{g}_\beta + \vec{g}_\rho)\cdot\nabla\times\vec{A}$. Further change of variable

$$b = \rho^{1/2}N, \quad \vec{B} = \rho^{1/2}\vec{A},$$

followed by neglection of terms of the order $(\vec{g}_\rho)^2$, leads us to the scalar wave equation

$$\nabla^2 b + k_\alpha^2 b = (2\vec{g}_\beta + \vec{g}_\rho)\cdot\nabla\times\vec{B} + O(\varepsilon_\alpha^2;\ \varepsilon_\beta^2;\ \varepsilon_\rho^2), \tag{2.10}$$

where

$$\varepsilon_\alpha = \left|\vec{g}_\alpha\frac{\alpha}{\omega}\right| \quad \varepsilon_\beta = \left|\vec{g}_\beta\frac{\beta}{\omega}\right|; \quad \varepsilon_\rho = \left|\vec{g}_\rho\frac{\alpha}{\omega}\right|. \tag{2.11}$$

Similarly, by taking the curl of Eq(2.4), we obtain to first order in $\vec{g}_\alpha$, $\vec{g}_\beta$ and $\vec{g}_\rho$

$$\nabla^2\vec{A} + (\vec{g}_\rho - \vec{g}_\beta)\cdot\nabla\vec{A} - \vec{g}_\beta\times\nabla\times\vec{A} + k_\beta^2\vec{A} =$$

$$= \left(\frac{-\beta^2}{\alpha^2}\right)\left[2\vec{g}_\beta + \left(2 - \frac{\alpha^2}{\beta^2}\right)\vec{g}_\rho\right]\times\nabla N + O(\varepsilon_\alpha^2;\ \varepsilon_\beta^2;\ \varepsilon_\rho^2), \tag{2.12}$$

$$\nabla^2\vec{B} - \nabla\vec{B}\cdot\vec{g}_\beta + k_\beta^2\vec{B} =$$

$$= -\left(\frac{\beta^2}{\alpha^2}\right)\left[2\vec{g}_\beta + \left(2 - \frac{\alpha^2}{\beta^2}\right)\vec{g}_\rho\right]\times\nabla b + O(\varepsilon_\alpha^2;\ \varepsilon_\beta^2;\ \varepsilon_\rho^2). \tag{2.13}$$

Note that since (2.1) does not have an *explicit* dependence on $\nabla\lambda$, as it does on $\nabla\mu$, the gradient of the compressional velocity α does not enter explicitly into Eqs. (2.4) - (2.12). It is, however, implicit via Eq.(2.5) and the gradients of N and b in Eqs. (2.8), (2.10) and (2.12). For that reason we must require, as we indeed do in Eq. (2.14) and (2.18), that $\varepsilon_\alpha \ll 1$.

We shall adopt the following terminology:

1. *Smooth inhomogeneous media:* second order terms in the coupling vectors are neglected but first terms are kept. The equations of motion are coupled [Eq. (2.10), (2.13)].

2. *Weak inhomogeneous media:* first and second order terms in the coupling vectors are dropped. The equations of motion are totally decoupled .

Eqs. (2.10) and (2.13) are the first order elastodynamic equations for the coupled shear and compressional wave motion in general inhomogeneous isotropic media. The entities $\vec{g}_\rho$, $\vec{g}_\alpha$ and $\vec{g}_\beta$ are the *P-SV coupling vectors*. Landers and Claerbout (1972) and Landers (1974) numerically integrated equations similar to our (2.10) and (2.13) in two dimensions to recover the seismic wavefield. In the context of *ray-theory* we assume a *total decoupling* of P-SV motion. A necessary condition is

$$\varepsilon_\alpha \ll 1, \quad \varepsilon_\beta << 1, \quad \varepsilon_\rho \ll 1 . \tag{2.14}$$

In media that obey (2.14), the material properties change slowly over distances of order of a wavelength. The corresponding equations of motion in such media will then simplify to

$$\nabla^2 b + k_\alpha^2 b = 0, \quad \nabla^2 \vec{B} + k_\beta^2 \vec{B} = 0. \tag{2.15a}$$

Thus, for weak inhomogeneous media, the functions

$$b = \rho^{1/2}\alpha^2 \nabla\cdot \vec{U} \qquad \text{and} \qquad \vec{B} = \rho^{1/2}\beta^2 \nabla\times \vec{U} , \tag{2.15b}$$

obey the respective scalar and vector Helmholtz equations. In earthquake and exploration seismology, there are many instances where this approximation is sufficient. If first order coupling terms of P and SV waves are retained, then one must solve Eqs. (2.10) and (2.12) simultaneously. In either case, the spectral displacement field $\vec{U}(M,\omega)$ is recovered by further integration and differentiation,

$$4\pi\vec{U}(\vec{r}) = -\nabla_{\vec{r}} \int \frac{\nabla\cdot\vec{U}(\vec{\xi})}{|\vec{r}-\vec{\xi}|}\, d^3\xi + \nabla_{\vec{r}} \times \int \frac{\nabla\times\vec{U}(\vec{\xi})}{|\vec{r}-\vec{\xi}|}\, d^3\xi\,, \tag{2.16}$$

where $\nabla_{\vec{r}}$ is the del operating on the $\vec{r}$ coordinates of the integrand. The decoupling conditions (2.14) can also be written in the alternative form (Λ = wave length, $k = \frac{2\pi}{\Lambda}$ = wave number, and $k = \frac{\omega}{\alpha}$ for P waves and $k = \frac{\omega}{\beta}$ for S waves)

$$\Lambda \ll \min\left\{\left|\frac{\alpha}{\nabla\alpha}\right|, \left|\frac{\beta}{\nabla\beta}\right|, \left|\frac{\rho}{\nabla\rho}\right|\right\} \equiv \frac{l_0}{2} \quad \text{or} \quad k_\alpha l_0 \gg 1\,, \tag{2.17}$$

or

$$\left|\frac{\nabla\alpha}{\omega}\right| \ll 1, \quad \left|\frac{\nabla\beta}{\omega}\right| \ll 1, \quad \left|\alpha\frac{\nabla\rho}{\rho\omega}\right| \ll 1\,, \tag{2.18}$$

where l_0 is a characteristic length at a given location, and may vary from point to point in the medium. Condition (2.18) refers to a *local property* of the elastic medium. It can be shown in a vertically or radially varying medium that $\left|\frac{\beta}{\nabla\beta}\right|$ is radius of curvature at the lowest point of the ray (turning point).

The *mode decoupling condition* that we shall use is

$$\omega \gg \omega_0, \quad \omega_0 \equiv \frac{1}{2}\max\left\{|\nabla\alpha|,\ |\nabla\beta|, \alpha\frac{|\nabla\rho|}{\rho}\right\}. \tag{2.19}$$

This condition defines a virtual threshold frequency which must be surpassed by the wave frequency ω. Note that the threshold frequency ω_0 is tied to the characteristic length l_0 via the relation $l_0\omega_0 = \alpha$. We shall show later that (2.19) is physically more meaningful than (2.18) in the sense that ω_0 exists even when exact decoupling renders (2.18) meaningless.

Note that in homogeneous media with material discontinuities, l_0 assumes the geometrical meaning of radius of curvature of the discontinuity. If, for example, we

consider compressional wave propagation in a sphere of radius a, (2.17) will be replaced by

$$k_\alpha a \gg 1 \,, \tag{2.20}$$

which means that the wavelength must be much smaller than the sphere's radius. Note that this condition does not mean that, at the interface, P and S wave are not coupled by boundary conditions, but that their propagation, away from the interface, remains uncoupled (independent Helmholtz equations for P and S waves).

The inequalities (2.14), (2.17) and (2.18) are only *necessary conditions* for the elastodynamic equations to yield approximate decoupling of P and SV waves. The following note is appropriate in this connection: The smallness of ε_α, ε_β and ε_ρ does *not* imply that $\left|\frac{\nabla\alpha}{\alpha}\right|$, $\left|\frac{\nabla\beta}{\beta}\right|$ and $\left|\frac{\nabla\rho}{\rho}\right|$ cease to play a role in shaping the amplitudes of seismic waves in the earth. On the contrary, these entities turn out to be essential in the seismic theory of Gaussian beams and dynamic ray tracing. Indeed, if we solve the Helmholtz equations (such as Eqs. (2.15a) in coordinate systems where the metric scale factors depend on the velocities α and β (such as the ray-centered coordinate system or the intrinsic coordinate system)), second-order derivatives of α and β reappear in the Helmholtz equation.

In weakly inhomogeneous media we assume a decomposition of the displacement field analogous to the decomposition in the case where the medium parameters depend on a single coordinate (see for example Ben-Menahem and Singh, p.417, 1981)

$$\vec{U} = \rho^{-1/2}\,\nabla\psi_1 + \rho^{-1/2}\,\nabla\times\nabla\times(\hat{\alpha}\psi_2) + \mu^{-1/2}\,\nabla\times(\hat{\alpha}\psi_3)\,, \tag{2.21}$$

where $\hat{\alpha}$ is a (unit) vector subjected to certain restrictions, and the potentials ψ_i obey the respective Helmholtz equations $(\nabla^2 + k_i^2)\psi = 0 \;\; i = 1, 2$. The condition under which

(2.21) holds is that of Eq. (2.19) in vector separable coordinate systems. Substitution of Eq. (2.21) into Eqs.(2.15) links the potential systems (b, $\vec{B}$) and (ψ_1, ψ_2, ψ_3). The relations are $b = -\omega^2 \psi_1$ and $\vec{B} = -\omega^2 \nabla\times(\hat{a}\psi_2) + \beta\nabla\times\nabla\times(\hat{a}\psi_3)$.

Vertically Inhomogeneous Media

In the special case where the inhomogeneity of the elastic medium depends on one cartesian coordinate z, the analysis becomes simpler and mathematical tools are available to solve the equations of motion. Three important results can then be derived:

1. Exact decoupling of the P and SV wave motions is possible only in special cases where the constitutive parameters satisfy a pair of nonlinear ordinary differential equations (Hook, 1961, 1962; Alverson *et al.*, 1963; Lock, 1963). We shall deal with an example of this category in chapter III.

2. Under less restrictive conditions (keeping *first order* terms in ε_α , ε_β, and ε_ρ) the field can still be represented by an expression similar to (2.21)

$$\vec{U} = \frac{1}{f_1}\nabla(f_2\,\varphi_1) - \frac{1}{f_3}\nabla\times\nabla\times(\hat{e}_z f_4\,\varphi_2)\,. \tag{2.22}$$

Substituting this expression in Eqs. (2.8) and (2.13) and keeping terms up to first order derivatives in λ, μ, and ρ, a straightforward, though lengthy, analysis yields two coupled equations for the potentials φ_1 and φ_2

$$\nabla^2\varphi_1 + \frac{\rho'}{\rho}\frac{\partial\varphi_1}{\partial z} + k_1^2\varphi_1 = -\eta_{12}\left[\nabla^2 - \frac{\partial^2}{\partial z^2}\right]\varphi_2\,, \tag{2.23}$$

$$\nabla^2\varphi_2 + \frac{\rho'}{\rho}\frac{\partial\,\varphi_2}{\partial z} + k_2^2\varphi_2 = \eta_{21}\varphi_1\,. \tag{2.24}$$

Here,

$$\eta_{12} = \frac{\mu}{\lambda + 2\mu}\left(\frac{2\mu'}{\mu} - \frac{\rho'}{\rho}\right), \quad \eta_{21} = \left(\frac{2\mu'}{\mu} - \frac{\lambda + 2\mu}{\mu}\frac{\rho'}{\rho}\right), \tag{2.25}$$

are scalar coupling factors, and primes indicate derivatives with respect to z. A perturbation scheme for the solution of (2.23) and (2.24) is presented in Appendix A.

Lock (1963) derived an exact solution of the elastodynamic vector equation for

$$\mu(z) = \mu(0)e^{\alpha z}; \quad \lambda(z) = \lambda(0)e^{\alpha z}; \quad \rho(z) = \rho(0)e^{\alpha z} . \tag{2.26}$$

In this case $\vec{g}_\alpha = \vec{g}_\beta = 0$, $\vec{g}_\rho = \alpha\hat{e}_z$ = const. and Eq. (2.4) reduces exactly to our Eqs. (2.23) to (2.24) with $\rho'/\rho = \alpha$, $\eta_{12} = \alpha\gamma^{-1}$, $\eta_{21} = \alpha(2-\gamma)$, $\gamma = (\lambda + 2\mu)/\mu$. Since the velocities have fixed values, the rays are straight lines and the coupling is effected through the constant density gradient alone.

From this point on we shall assume a total P/S decoupling. Hence, we do not have to deal any more with the elastodynamic vector equation itself, but rather focus our attention on approximate solutions of the scalar Helmholtz equation.

3. HYPEREIKONAL

Let us now examine the Helmholtz equation for general inhomogeneous media. A convenient mathematical vehicle is afforded by a general orthogonal coordinate system where the only constraint is that one of its coordinates is the wave travel time τ. This system will be denoted (τ,γ_1,γ_2), with corresponding scale factors h_τ, h_1 and h_2. The scalar Helmholtz equation $\nabla^2\psi + \frac{\omega^2}{v^2}\psi = 0$ in this system is

$$\frac{1}{h_\tau h_1 h_2}\left[\frac{\partial}{\partial\tau}\left(\frac{h_1 h_2}{h_\tau}\frac{\partial\psi}{\partial\tau}\right) + \frac{\partial}{\partial\gamma_1}\left(\frac{h_\tau h_2}{h_1}\frac{\partial\psi}{\partial\gamma_1}\right) + \frac{\partial}{\partial\gamma_2}\left(\frac{h_\tau h_1}{h_2}\frac{\partial\psi}{\partial\gamma_2}\right)\right] + \frac{\omega^2}{v^2}\psi = 0 \tag{3.1}$$

We take ψ to be of the form $\psi = A_0(\tau,\gamma_1,\gamma_2;\omega)\, S(\tau;\omega)$. and substitute into (3.1). After

some intermediate steps, (3.1) reduces to

$$\frac{v^2}{h_\tau^2}\frac{d^2S}{S\,d\tau^2} + \omega^2 + v^2\frac{\nabla^2 A_0}{A_0} + \frac{v^2}{S}\frac{dS}{d\tau}\left[\nabla^2\tau + 2\frac{\nabla A_0}{A_0}\cdot\nabla\tau\right] = 0. \tag{3.2}$$

Assume that the terms involving first derivatives with respect to τ are of different order than the rest. This yields the two equations

$$\frac{v^2}{h_\tau^2}\frac{d^2S}{S\,d\tau^2} + \omega^2 + v^2\frac{\nabla^2 A_0}{A_0} = 0\,, \tag{3.3}$$

and

$$\nabla^2\tau + 2\frac{\nabla A_0}{A_0}\cdot\nabla\tau = 0. \tag{3.4}$$

Consider the class of inhomogeneous media such that

$$\nabla^2 A_0 + \frac{\omega_c^2}{v^2}A_0 = 0\,, \tag{3.5}$$

where ω_c and v^2/h_τ^2 are either constants or a slowly varying functions of the coordinates. In cases where both are constants, an exact solution of Eq. (3.3) is

$$S(\tau;\omega) = e^{i\omega\tau\frac{h_\tau}{v}(1-\omega_c^2/\omega^2)^{1/2}}\,, \tag{3.6}$$

The S functional in (3.6) is defined as the *hypereikonal*. Equation (3.4) is the *transport equation* with an exact solution

$$A_0(\tau,\gamma_1,\gamma_2;\omega) = f(\gamma_1,\gamma_2;\omega)\left[\frac{h_\tau}{h_1\,h_2}\right]^{1/2}. \tag{3.7}$$

where $f(\gamma_1,\gamma_2;\omega)$ is a function independent of τ. This solution assumes $h_1\,h_2 \neq 0$. In the case where γ_1 and γ_2 are the two spherical angles, f can be interpreted as the product of the radiation pattern of a source located at the origin of the spherical coordinate system times the source strength. Thus, for example $f(\gamma_1,\gamma_2;\omega) = c_0$

(constant) the source is a point explosion. Let J be the Jacobian of the transformation from cartesian to the general orthogonal coordinate system (τ,γ_1,γ_2).

$$J \equiv \frac{D(x,y,z)}{D(\tau,\gamma_1,\gamma_2)} \equiv h_\tau h_1 h_2 . \tag{3.7a}$$

Define the surficial Jacobian as $J^S \equiv J/h_\tau \equiv h_1 h_2$. Equation (3.7) then reads

$$A_0 = f \left[\frac{h_\tau}{J^S}\right]^{1/2} . \tag{3.7b}$$

J^S is sometimes called the spreading function.

The solution of the Helmholtz equation can then be written in the compact form

$$\psi \equiv A_0 S = f(\gamma_1,\gamma_2;\omega) \left[\frac{h_\tau}{J^S}\right]^{1/2} e^{i\omega\tau \frac{h_\tau}{v}(1-\omega_c^2/\omega^2)^{1/2}} . \tag{3.8}$$

For $\omega_c = 0$, or $\omega \gg \omega_c$, this solution reduces to the ray solution (see section 4) or the WKBJ solution of the wave equation (see section 6), respectively. In media where (3.8) holds, there exist a *cut-off frequency* ω_c below which wave propagation is not possible. The medium cut-off frequency, defined as, $\omega_c \equiv v(\nabla^2 A_0/A_0)^{1/2}$, may be complex. Media for which ω_c is complex will not be treated here and a special investigation will follow in a sequel paper. The factor $(1-\omega_c^2/\omega^2)^{1/2}$ in equation (3.8) indicates the presence of a second order effect of dispersion which is usually neglected for seismic body waves at high frequencies.

The hypereikonal formulation for the Helmholtz equation can be extended in a straightforward way to the vector elastodynamic wave equation (2.1). In Appendix B, we show how this can be achieved for the special case of vertically inhomogeneous media.

Vertically Inhomogeneous media

As an application we shall derive in closed form the results given above in the case where the medium's properties are only allowed to vary along the z direction. A convenient coordinate system is the *intrinsic coordinate system* (Yosiyama, 1940). In a vertically inhomogeneous medium, we define a cartesian coordinate system centered at O (figure 1). The cylindrical coordinates of a general point M are (r, ϑ, z). A seismic source is placed on the z axis at $M_0(0,0,z_0)$. We define new coordinates $M(\tau,j_0,\vartheta)$ relative the origin M_0, where ϑ is the azimuth angle of the cylindrical system. The transformation-equations linking the coordinates (z,r) and (τ,j_0) are given by the implicit integral relations

$$r = \int_{z_0}^{z} (g^2 - p^2)^{-1/2}\, p\, dz \qquad ; \qquad \tau = \int_{z_0}^{z} g^2\, (g^2 - p^2)^{-1/2} dz\ , \tag{3.9}$$

where

$$g = \frac{1}{v(z)}, \qquad p = g(z_0)\sin j_0\ . \tag{3.10}$$

These are the standard travel time equation, τ, and horizontal slowness, p, in Herglotz-Wiechert formulas (Aki and Richards, 1980); j_0 is the takeoff angle. It is understood that p is held fixed (z independent) during the integrations. Thus, although j_0 is indeed a coordinate, it acts as a fixed parameter during the intergration. v denotes either α or β.

For any given velocity distribution $v(z)$, Eq. (3.9) defines the relations $r = r(z,p)$ and $\tau = \tau(z,p)$. For any given values of z and r, these two relations define the pair (τ,j_0) and vice versa. For a fixed value of p (i.e., j_0), the r integral in (3.9) defines a curve $r = r(z)$ in the plane $\vartheta = const.$ and the τ integral defines the travel time along this curve from z_0 to z.

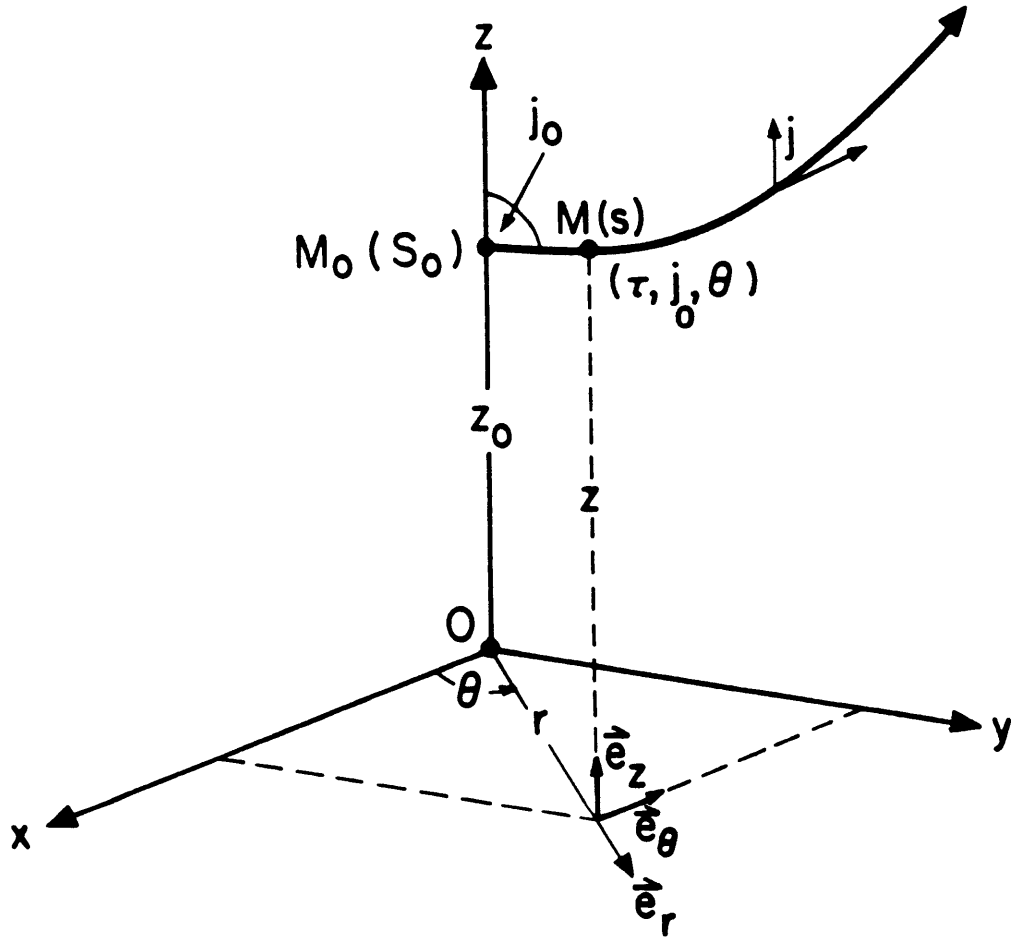


Figure 1. The intrinsic coordinate system

The angle j between the vertical and the tangent at any point on this curve is obtained readily by differentiating r with respect to z for fixed p. This is the ray angle of incidence.

$$\left(\frac{\partial r}{\partial z}\right)_p = p\ (g^2 - p^2)^{-1/2} = \tan j\ . \tag{3.11}$$

Solving for p and comparing with (3.10) we find

$$p = \frac{\sin j}{v(z)} = \frac{\sin j_0}{v_0}, \tag{3.12}$$

where $v_0 = v(z_0)$. Eq. (3.10) is the *ray equation* and Eq. (3.12) is *Snell's law*. For a fixed value of j_0, the ray equation is given by a relation $r = r(z;j_0)$. With these definitions, one can verify in a straightforward manner the following statements:

A. The line element of the new system is

$$ds^2 = h_\tau^2 d\tau^2 + h_\vartheta^2 d\vartheta^2 + h_{j_0}^2 dj_0^2\ , \tag{3.13}$$

where the explicit expressions for the scale factors are:

$$h_\tau = v(z);\ h_\vartheta = r;\ h_{j_0} = p\,\frac{(g^2 - p^2)^{1/2}}{g}\cot j_0 \int_{z_0}^{z} g^2 (g^2 - p^{2)}{}^{-\frac{3}{2}} dz\ . \tag{3.14}$$

Hence, the coordinate system (τ, j_0, ϑ) is orthogonal.

B. The wavefront equation is obtained by eliminating p between the relations $\tau_0 = \int_{z_0}^{z} g^2 (g^2 - p^2)^{-1/2} dz$ and $r = \int_{z_0}^{z} p\,(g^2 - p^2)^{-1/2} dz$. At any given time $\tau = \tau_0$ the wavefront is a *surface of revolution* generated by the plane curve (figure 2)

$$z = z(r; z_0; \tau_0)\ . \tag{3.15}$$

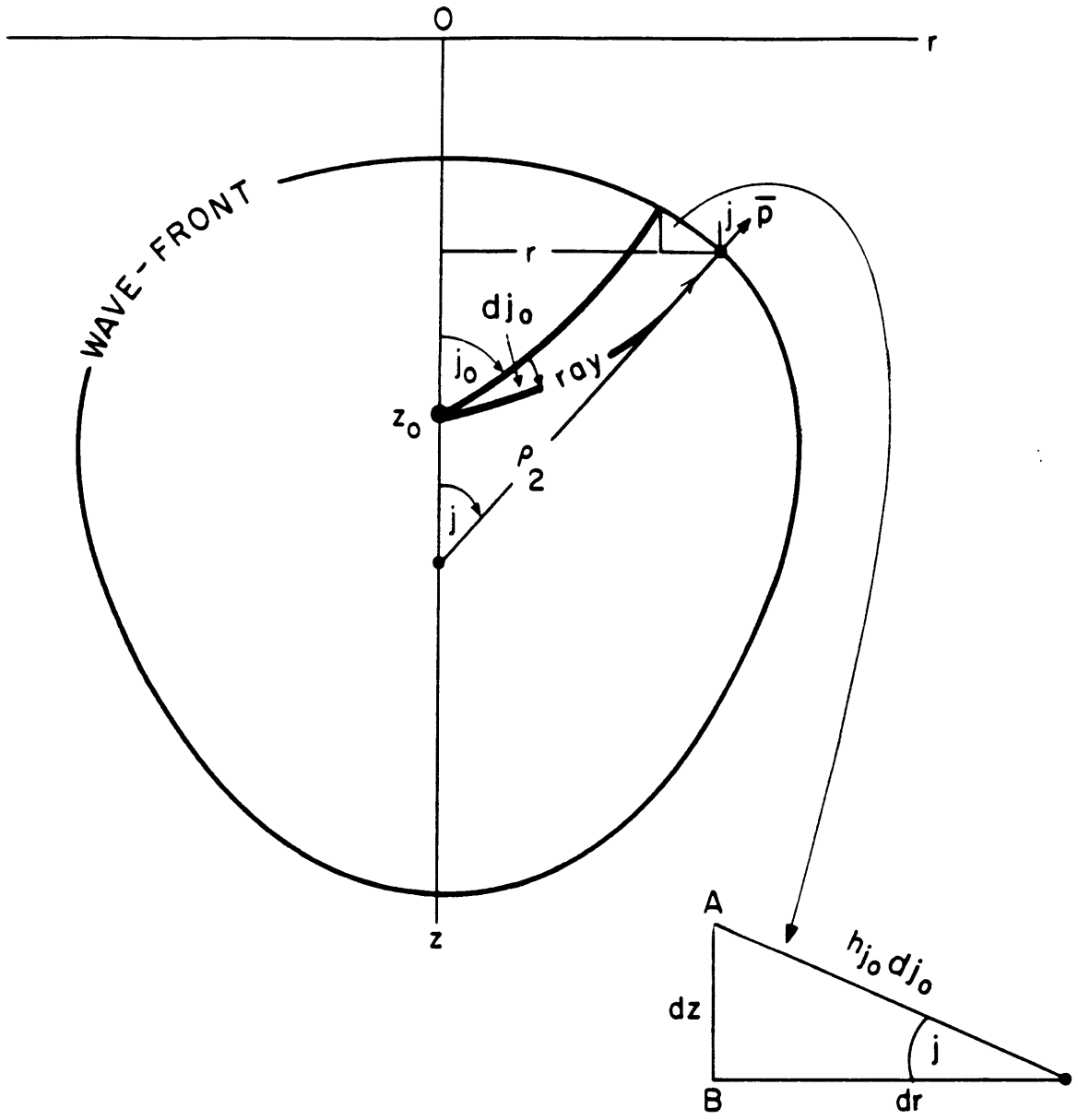


Figure 2. Cross-section of wavefront at $\vartheta = const$.

Since on this surface $d\tau = 0$, we have from Eq. (3.13) and from the geometry of a line element on a general surface of revolution

$$ds^2 = h_\vartheta^2 d\vartheta^2 + h_{j_0}^2 dj_0^2 = r^2 d\vartheta^2 + \left[1 + \left(\frac{dz}{dr}\right)^2\right] dr^2 .$$

Therefore, on the wave front, up to a sign

$$\frac{dz}{dr} = \tan j ; \quad h_{j_0} = \frac{1}{\cos j}\frac{dr}{dj_0} = \frac{1}{\sin j}\frac{dz}{dj_0} ; \tag{3.16}$$

The geometric interpretation of Eq. (3.16) is evident from figure 2. An alternative interpretation of h_{j_0} is obtained when we differentiate r in Eq. (3.9) with respect to j_0 at fixed z and use Eq. (3.14)

$$h_{j_0} = \cos j \left(\frac{\partial r}{\partial j_0}\right)_z . \tag{3.17}$$

C. The *Gaussian curvature* of a surface of revolution is defined as $K = (\rho_1\rho_2)^{-1}$ where ρ_1 *and* ρ_2 are the wavefront *principal radii of curvature* (eg., Eisenhart, 1909). In our present case

$$\rho_1 = \frac{\left[1 + \left(\frac{dz}{dr}\right)^2\right]^{\frac{3}{2}}}{\frac{d^2z}{dr^2}} = \left[\frac{d\sin j}{dr}\right]^{-1} , \tag{3.18}$$

$$\rho_2 = r\frac{\left[1 + \left(\frac{dz}{dr}\right)^2\right]^{1/2}}{\left(\frac{dz}{dr}\right)} = \frac{r}{\sin j} ,$$

$$K = \frac{\frac{dz}{dr}\frac{d^2z}{dr^2}}{r\left[1 + \left(\frac{dz}{dr}\right)^2\right]} = \frac{\sin j}{r}\frac{d\sin j}{dr} = \frac{\sin j \cos j}{r}\frac{dj}{dr} ; \quad j = j(r) . \tag{3.19}$$

D. The *divergence coefficient* (e.g., Ben-Menahem and Singh, 1981) is given by the expression $X = \left[\frac{v_0\rho_0}{v\rho}\,\frac{\sin j_0}{r\cos j}\left(\frac{\partial j_0}{\partial r}\right)_z\right]^{1/2}$, where $|u| = X\,e^{i\omega\tau}\,|u_0|$, u being the particle displacement. In terms of the scale factors and with the aid of Eqs. (3.7) and (3.17), it is found that

$$X = \left(\frac{\rho_0}{\rho}\right)^{1/2}\frac{v_0}{v}\,\frac{A_0}{c_1}, \qquad A_0 = c_1\left(\frac{\sin j_0}{v_0}\right)^{1/2}\left(\frac{h_\tau}{h_\vartheta h_{j_0}}\right)^{1/2}. \tag{3.19a}$$

The point source constant introduced earlier, c_0, is here equal to $c_1\,(\sin j_0/v_0)^{1/2}$, where c_1 is the source strength. Using Eq. (3.12), we can also write

$$X = \left(\frac{\rho_0}{\rho}\right)^{1/2}\frac{v_0}{v}\,(\rho_2 h_{j_0})^{-1/2}. \tag{3.20}$$

If we substitute for $\sin j$ in (3.16) from Snell's law (3.12), and evaluate the $\frac{d}{dr}$ derivative at the wavefront, we obtain

$$\frac{\cot j}{\rho_1} = \frac{\cot j_0}{h_{j_0}} + \frac{1}{R}, \tag{3.21}$$

where $R = \left(\frac{\sin j}{v}\frac{dv}{dz}\right)^{-1}$ is the radius of curvature of the ray at the wavefront. In homogeneous media $\rho_1 = h_{j_0}$. In weak inhomogeneous media we have approximately

$$\rho_1 \approx \frac{\tan j_0}{\tan j}\,h_{j_0}. \tag{3.22}$$

The Jacobian of the transformation from cartesian to intrinsic coordinates is $J = D(x,y,z)/D(\tau,j_0,\vartheta) = h_\tau h_\vartheta h_{j_0}$ (3.17a). If we consider a homogeneous unit focal sphere around the source, Eq. (3.14) will render for this sphere $(h_\tau)_0 = v_0$, $(h_\vartheta)_0 = 1$, $(h_{j_0})_0 = \sin j_0$. Denoting the Jacobian on the focal sphere, $J_0 = v_0 \sin j_0$, we find from (3.19a)

$$X = \left[\frac{\rho_0 J_0}{\rho J}\right]^{1/2} \quad ; A_0 = c_1 \frac{v}{v_0}\left[\frac{J_0}{J}\right]^{1/2} . \tag{3.23}$$

Note that in weak vertically inhomogeneous media, the *total curvature* (Eisenhart, 1909) of the wavefronts is

$$K_w = (\rho_1\rho_2)^{-1/2} \approx \left[\frac{\tan j}{\tan j_0}\right]^{1/2} \frac{A_0}{c_1} . \tag{3.24}$$

We thus see that A_0 is a fundamental quantity that is linked to divergence coefficient, the Jacobian and proportional to the total curvature of the wavefront. In cases where $(\tan j / \tan j_0)^{1/2}$ is close to unity, A_0 is close to the total wavefront curvature.

A class of media for which Eq. (3.2) holds is found in the following way: we put

$$A_0 = c_1 \left[\frac{v}{v_0}\right]^{1/2} \frac{1}{D}, \quad D = \left[\frac{r\, h_{j_0}}{\sin j_0}\right]^{1/2} , \tag{3.25}$$

where D is a certain distance. It then follows that Eq. (3.5) assumes the new form

$$v^2 \frac{\nabla^2 A_0}{A_0} = v^2 D \nabla^2 \frac{1}{D} + \left[-\frac{1}{4}\left[\frac{dv}{dz}\right]^2 + \frac{1}{2} v \frac{d^2 v}{dz^2}\right] - \frac{v}{D}\frac{\partial D}{\partial z}\frac{dv}{dz} \equiv -\omega_c^2 . \tag{3.26}$$

In a homogeneous medium $D\,\nabla^2 \frac{1}{D} \equiv 0$, $v' = v'' = 0$ and $\omega_c = 0 \equiv \omega_0$. In a medium with a constant velocity gradient $v = v(0)(1 + \eta z)$, one proves that $v^2\, D\nabla^2 \frac{1}{D} - \frac{v}{D}\frac{\partial D}{\partial z}\frac{dv}{dz} \equiv 0$. Eq. (3.26) then yields the exact result

$$\omega_c = \frac{1}{2}\frac{dv}{dz} = \frac{\eta}{2} v(0) \equiv \omega_0 . \tag{3.27}$$

In a medium with a linear gradient $v = v(0)\,(1 + \eta_1 z + \eta_2 z^2)$

$$v^2 \frac{\nabla^2 A_0}{A_0} = -\frac{1}{4} v^2(0)[\eta_1^2 - 4\eta_2] + O\left[\frac{1}{D^2}\right] . \tag{3.28}$$

So, at least in the far field, the cut-off frequency is at

$$\omega_c = \frac{v(0)}{2}(\eta_1^2 - 4\eta_2)^{1/2} . \tag{3.29}$$

However $\omega_0 = \frac{v(0)}{2}(\eta_1 + 2z\eta_2)$. This includes Eq. (3.27) as a special case wherever $\eta_2 = 0$.

The threshold frequency, ω_0, characterizes the heterogeneous medium. In order for asymptotic wave methods to be applicable in such a medium, the wave angular frequency ,ω, must be greater than ω_0. This is the mode decoupling condition derived in section 2, and expressed in (2.19). Another medium characteristic frequency is the cutt-off frequency, ω_c, introduced in (3.8). The wave frequency must be greater than ω_c so that second order effect of dispersion can be neglected.

We thus see that in media where the velocity profile can be represented by a polynomial of degree one, $\omega_0 = \omega_c$. In general, ω_c and ω_0 depend upon the coordinates, and their equality is not obvious. In the special case of the paraxial approximation (and ray theory) their equality is postulated on the basis of dimensional analysis.

4. RAY-SERIES: EIKONAL AND TRANSPORT EQUATIONS

Geometrical optics is based on the first term of the asymptotic-series solution of the Helmholtz equation

$$\psi(M,\omega) = e^{i\omega\tau(M)} \sum_{n=0}^{\infty} \frac{A_n(M)}{(-i\omega)^n} . \tag{4.1}$$

The substitution of

$$\psi(M,\omega) = A_0(M)\, e^{i\omega\tau(M)} , \tag{4.2}$$

into the Helmholtz wave equation [sometimes known as the *ansatz of Sommerfeld and Runge* (Cornbleet, 1973)] leads to

$$\left[\nabla^2 + \frac{\omega^2}{v^2}\right]\psi = \left\{\frac{\nabla^2 A_0}{A_0} + \omega^2\left[\frac{1}{v^2} - (\nabla\tau)^2\right] + i\omega\left[\nabla^2\tau + 2\frac{\nabla A_0}{A_0}\cdot\nabla\tau\right]\right\}\psi = 0 . \quad (4.3)$$

If A_0 is a sufficiently slowly varying function of the coordinates over a wavelength scale, we have

$$\omega \gg \omega_c . \quad (4.4)$$

This is defined as the *high frequency condition.* Then, equating real and imaginary parts on both sides of Eq. (4.3) we obtain the well-known *eikonal* and *transport* equations respectively

$$(\nabla\tau)^2 = \frac{1}{v^2} ; \quad (4.5)$$

$$\nabla^2\tau + 2\frac{\nabla A_0}{A_0}\cdot\nabla\tau = 0. \quad (4.6)$$

The characteristic equations of the eikonal (4.5) yield rays Ω. The solution of the eikonal has the physical meaning of the travel time along a ray Ω connecting the source $M_0(s_0)$ to a receiver at $M(s)$, where s is the arclength along the ray measured from a given fiducial point. It is given explicitly by the equation

$$\tau(s) = \int_{s_0}^{s} v(s)^{-1}\, ds . \quad (4.7)$$

The transport equation (4.6) involves the travel time $\tau(s)$ and a function $A_0(s)$ with the physical meaning of an amplitude of a local plane wave along the ray coordinate s. It is the same as that derived in the previous section (3.4). The surface $\tau(s) = const.$ yields the wavefront (or phase front). The solution of the transport equation is given explicitly in (3.7).

An equivalent approach in obtaining the eikonal is to equate to zero the coefficient of

ω^2 in (4.3). The transport equation is of lower order in frequency, since it is obtained by equating to zero the coefficient of ω. For high frequencies the remaining term in (4.3) is assumed of order one (high frequency approximation).

Having defined the concept of a ray, the coordinate system introduced and used in the previous section remains valid for characterizing rays. The coordinate system (τ,γ_1,γ_2) introduced in the beginning of section 3, will henceforth be called *ray coordinates*. A special case is the intrinsic coordinate system (τ,j_0,ϑ), which will be used in the present context.

However, it is convenient to solve the eikonal equation in the *ray-centered coordinate system* (s,q_1,q_2). The process of solving the eikonal equation in this orthogonal system is known as *dynamic ray tracing*. s measures the arclength along the ray Ω from an arbitrary reference point to the receiver position $M(s)$ and q_1,q_2 are the cartesian coordinates in the plane perpendicular to Ω, with origin at Ω. Details of this system and its regularity are described in Cerveny (1983[a]). The scale factors of this coordinate system with respect to the cartesian reference frame are

$$h_s = 1 + q_1 \frac{1}{v}\frac{\partial v}{\partial q_1} + q_2 \frac{1}{v}\frac{\partial v}{\partial q_2}; \quad h_{q_1} = h_{q_2} = 1 .$$

The partial derivatives are evaluated at $q_i=0$.

Defining K_w as the local wavefront (or phase front) curvature matrix (in the vicinity of Ω), the eikonal equation in the ray centered coordinate system of Ω is written in the form of an ordinary non-linear differential equation of the first order of the Ricatti type for K_w, which cannot be generally solved by analytical techniques. Letting $K_w(s) \equiv v(s)\, P(s)\, Q^{-1}(s)$, the eikonal is expressed as a coupled first-order ordinary linear differential system (Cerveny, 1983[a])

$$\frac{dQ(s)}{ds} = v(s)\,P(s)\,, \tag{4.8}$$

$$\frac{dP(s)}{ds} = -v(s)^{-2}\,V(s)\,Q(s)\,,$$

where v is the P or S wave velocity and $V_{ij} = \left[\frac{\partial^2 v}{\partial q_i \partial q_j}\right]_{qi=qj=0}$. Solving this system, with specified initial conditions (point source or line source), for the 2 x 2 matrices $P(s)$ and $Q(s)$ determines $K_w(s)$. $K_w(s)$ is symmetrical, and in 2-D media, it reduces to the scalar K_w (total wavefront curvature).

It is shown that in ray centered coordinates, for a given point s on Ω, the transport equations for the P wave principal component $\vec{u}^{\alpha}$ (along the tangent vector to Ω) and for S wave principal components $\vec{u}^{\beta 1}$ (along $\vec{q}_1$) and $\vec{u}^{\beta 2}$ (along $\vec{q}_2$) are independent. The ray centered coordinates "untwists" the ray of its torsion, at every point s of Ω, to its initial value at M_0. The unit vectors that constitute this coordinate system are sometimes called the polarization vectors. In a weak heterogeneous medium, the analytical solution of the transport equation for a receiver $M(s)$ is given by

$$A_0(s) = c_0\left[\frac{v(s)}{J^S(s)}\right]^{1/2}, \tag{4.9}$$

where c_0 is a constant characterizing an explosion source. For other sources embedded in a homogeneous focal sphere, c_0 must be replaced by $f(\gamma_1,\gamma_2;\omega)$, as defined in equation (3.7). This is nothing else than (3.7b) expressed on the ray Ω. For P waves, $A_0 \equiv |\vec{u}^{\alpha}|$ and $v(s) = \alpha(s)$, and for S waves $A_0 \equiv |\vec{u}^{\beta 1}|$, or $A_0 \equiv |\vec{u}^{\beta 2}|$, and $v(s) = \beta(s)$. J^S is the surficial Jacobian of the transformation from cartesian to ray coordinates (τ,γ_1,γ_2). It is computed via $J^S(s) = \det[Q(s)]$, and measures the cross sectional area of an elementary ray tube surrounding Ω . J^S is related to J with the identity given above (3.7b). Here $h_\tau \equiv v h_s$. On the ray we have $h_\tau \equiv v$. The Jacobian J^S, in the dynamic ray tracing method, results from the product of two Jacobians

$$J^S \equiv h_\tau^{-1} \frac{D(x,y,z)}{D(\tau,\gamma_1,\gamma_2)} = h_\tau^{-1} \frac{D(x,y,z)}{D(\tau,q_1,q_2)} \frac{D(\tau,q_1,q_2)}{D(\tau,\gamma_1,\gamma_2)},$$

or in terms of scale factors

$$J^S \equiv h_{\gamma_1} h_{\gamma_2} = h_{q_1} h_{q_2} \det[Q] = \det[Q]. \tag{4.9a}$$

Q can be viewed as the Jacobian matrix of the transformation from ray centered to ray coordinates, and consequently $\det[Q]$ is the Jacobian of this transformation.

The displacement is recovered via (2.21). Recalling that A_0 and τ are only functions of s, and that in the case of P waves the displacement is along the ray $\hat{p}$, we have

$$\vec{U} = \rho^{-1/2} \nabla\psi = c_0 \rho^{-1/2} \left(\frac{i\omega}{v} A_0 + \frac{dA_0}{ds} \right) e^{i\omega\tau} \hat{p}. \tag{4.10}$$

Since the ray solution is the far field contribution of $\vec{U}$, (condition (5.9)) we have for the ray solution

$$\vec{U}^r = c_0 (\rho v J^S)^{-1/2} i\omega e^{i\omega\tau} \hat{p}. \tag{4.11}$$

Expressed in terms of a divergence coefficient (3.23) we have

$$\vec{U}^r = C \left(\frac{\rho_0 v_0 J_0^S}{\rho v J^S} \right)^{1/2} i\omega e^{i\omega\tau} \hat{p}, \tag{4.12}$$

where $C \equiv U_0^r / i\omega = c_1 / (\rho_0^{-1/2} v_0)$. Parameters with subscript 0 denote their value at the homogeneous unit focal sphere surrounding the source. For a point source $J_0^S = \sin j_0$ (see 3.23).

Cumulative Error

So far we have not raised the question of the error produced by replacing the

exact solution of the Helmholtz equation by the first term in the geometrical optics approximation. Let then

$$\psi = \psi_0 + \varepsilon = \psi_0\left[1 + \frac{\varepsilon}{\psi_0}\right], \tag{4.13}$$

where $\psi_0 = A_0 e^{i\omega\tau(1-\omega_c^2/\omega^2)^{1/2}}$ as derived in equation (3.8), $\psi = A_0 e^{i\omega\tau}$, and ε is the correction term.

With the provision $\omega \gg \omega_c$ we have from (4.13)

$$\left|\frac{\varepsilon}{\psi_0}\right| \approx \left| e^{\frac{i\tau\omega_c^2}{2\omega}} - 1 \right| \approx \frac{\omega_c^2\,\tau}{2\omega}. \tag{4.14}$$

Therefore, in order to have small relative error we must have

$$\omega \gg \omega_c^2\tau. \tag{4.15}$$

This condition is defined as the *Fresnel 1-condition*. It is re-derived, within the paraxial approximation in section 7, for a homogeneous medium.

Note that

$$\lim_{\omega\to\infty}\left|\frac{\varepsilon}{\psi_0}\right| = 0. \tag{4.16}$$

This means that in the radiation zone of inhomogeneous media (i.e. when eq. (4.4) is satisfied), the error due to ray theory tends to decrease with the increase of the frequency.

We can view (4.14) as a *cumulative* error along the ray and condition (4.15) ensures that this error stays small. This can be shown by substituting (4.13) into the Helmholtz equation and using the transport equation (4.4). We find that ε obeys its own Helmholtz equation with a forcing term

$$\nabla^2\varepsilon + \frac{\omega^2}{v^2}\varepsilon = -\frac{\nabla^2 A_0}{A_0}\psi_0 \,. \tag{4.17}$$

Direct integration of Eq. (4.17) renders

$$\varepsilon = \frac{1}{4\pi}\int \frac{\nabla^2 A_0}{A_0}\psi_0 G(\vec{r}\,|\,\vec{r}_0)d^3\vec{r}_0 \,, \tag{4.18}$$

where the Green's function G satisfies

$$\nabla^2 G + \frac{\omega^2}{v^2}G = -4\pi\delta(\vec{r} - \vec{r}_0) \,. \tag{4.19}$$

In order to estimate the integral in (4.18) we apply the mean value theorem over a finite volume element which is assumed to be the Fresnel volume. As we shall see in section 7, this volume is a paraboloid of revolution. The source-receiver distance will be taken to be $\bar{v}\tau$. The volume is then

$$V_F = \frac{\pi}{2} r_0^2 \, \bar{v}\tau \,, \tag{4.20}$$

where $\bar{v}$ is the average velocity along the ray, $r_0^2 = \Lambda v \tau$ is the first Fresnel zone radius at the observer.

Eq (4.18) then becomes

$$\varepsilon \approx \frac{1}{4\pi}\left|\frac{\nabla^2 A_0}{A_0}\psi_0 \; G\right|_{average} V_F \,. \tag{4.21}$$

with Fresnel volume, V_F, given by (4.20).

However, the absolute value of G is approximately equal to the wavefront curvature, R_w^{-1}, which is of same order as $(\bar{v}\tau)^{-1}$, and $\nabla^2 A_0 / A_0 = l_0^{-2}$. Substituting these values into (4.21) we find that the relative error ε/ψ_0 is small, if (4.15) is satisfied (within a factor $\frac{\pi}{2}$). The relative error can thus be seen as a cumulative error

along the ray.

Even within the limits of the above restrictions, the ray-series will fail to represent the solution of the Helmholtz equation at critical points where the Jacobian in (4.9) vanishes. For a vertically heterogeneous medium, according to Eq. (3.20), this failure can be avoided if (see figure 2)

$$h_{j_0} \neq 0 \quad ; \quad \rho_2 = \frac{r}{\sin j} \neq 0 \quad . \tag{4.22}$$

5. NEAR AND FAR FIELDS

We shall first derive intuitive conditions for inhomogeneous media from extrapolation of the homogeneous case. Then, we shall define formally the far field condition directly from the eikonal and transport equations. In a homogeneous elastic medium, the displacement field of seismic point source such as an explosion is given explicitly by the expression (e.g. Ben-Menahem and Singh, 1981)

$$\vec{U} = \frac{-m_0}{4\pi\rho\alpha^2}\nabla\left[\frac{e^{ik_\alpha\Delta}}{\Delta}\right], \tag{5.1}$$

$$= \frac{-m_0}{4\pi\rho\alpha^2}\left[ik_\alpha - \frac{1}{\Delta}\right]\frac{e^{ik_\alpha\Delta}}{\Delta}\hat{e}_\Delta,$$

where m_0 is the source seismic moment, $k_\alpha = \omega/\alpha$ is the compressional wavenumber, Δ is the source-receiver distance and $\hat{e}_\Delta$ is a unit vector along Δ.

If $k_\alpha\Delta \gg 1$ or $\Delta \gg \Lambda$, or $\tau \gg$ wave period, the expression

$$\vec{U} \approx \frac{-im_0k_\alpha}{4\pi(\lambda + 2\mu)}\left[\frac{e^{ik_\alpha\Delta}}{\Delta}\right]\hat{e}_\Delta, \tag{5.2}$$

represents the far-displacement field. Since $k_\alpha\Delta = \omega\Delta/\alpha = \omega\tau$ where τ is the P-wave travel-time along the ray, it is expected that the far field condition in inhomogeneous media will read accordingly

$$\tau \gg t_0, \tag{5.3}$$

where τ given by (4.7) is the travel-time along the ray with intrinsic velocity $v(s)$ (v denotes either a shear velocity β *or* a compressional velocity α) and t_0 is some *threshold time* which is characteristic of the particular inhomogeneous medium under consideration. This threshold-time must be related to the *threshold frequency* of Eq. (2.19) through the relation

$$\omega_0 t_0 \approx 1 \,. \tag{5.4}$$

For wave propagation in decoupled inhomogeneous media, both (2.19) and (5.3) must hold simultaneously. Therefore, combination of both conditions yield

$$\omega\tau \gg \omega_0 t_0 \approx 1 \,. \tag{5.5}$$

Similarly, the *near field* of (5.1) can be defined as $k_\alpha \Delta \ll 1$. The corresponding analogy for an inhomogeneous medium, taking again into account the decoupling condition, becomes

$$\omega_0 \, \tau \ll \omega \, t_0 \,, \tag{5.6}$$

or

$$\omega\tau \ll \left[\frac{\omega}{\omega_0}\right]^2 . \tag{5.7}$$

For a given velocity distribution $v(s)$, $A_0(s)$ and $\tau(s)$ can be calculated by solving the eikonal and transport equations. The solution in (4.7) is physically meaningful only if both the unit slowness vector $\hat{p} = v \, \nabla\tau(s)$ and the local plane wave amplitude $A_0(s)$ are slowly varying functions of the coordinates. These limitations can be expressed in mathematical form as follows:

Let s be the local direction of propagation along a ray in a ray centered coordinate system. Then if Λ is a wavelength, the local change of the wave amplitude over a distance of one wavelength can be estimated from

$$A_0(s + \Lambda) = A_0(s) + \Lambda \frac{dA_0}{ds} + \ldots = A_0(s)\left[1 + \Lambda \frac{dA_0 / ds}{A_0} + \ldots\right] . \tag{5.8}$$

A slow variation of $A_0(s)$ means that

$$k\,A_0 \gg |\hat{p}\bullet\nabla A_0| = \left|\frac{dA_0}{ds}\right| . \qquad (5.9)$$

This is defined as the *far field condition*. Similarly, a slow variation of $\hat{p}$ means that $\Lambda \ll |\hat{p}/\nabla\hat{p}|$, or expressed along the ray

$$\omega \gg \left|\frac{dv}{ds}\right| , \qquad (5.10)$$

where $\hat{p}$ is the local unit slowness vector tangent to the ray (normal to the wavefront) (figure 2). Since, up to here we have assumed that the decoupling condition (2.19) was satisfied and $|\nabla v| \geq |dv/ds|$, we conclude that the condition (5.10) is always included in (2.19).

In vertically inhomogeneous media, we have shown in (3.24) that A_0 is equal approximately to the total curvature of the wavefront. Denoting the wavefront radius of curvature by R_w, we may write condition (5.9) in a new form

$$k\,R_w \gg |\hat{p}\bullet\nabla R_w| . \qquad (5.11)$$

In homogeneous media $A_0 = 1/\Delta$, where Δ is a coordinate along the ray. Then, $|\hat{p}\bullet\nabla A_0| = 1/\Delta^2$ and we fall back on the familiar far field condition $k\Delta \gg 1$.

Thus, in addition to our former requirements, $\omega \gg \omega_0$, which secures the decoupling of the vector elastodynamic equation (2.1), and small variations the slowness vector, and $\omega \gg \omega_c$, which guaranties the high frequency approximation, Eq. (5.9) secure small variations of the local plane wave amplitude.

The far field condition (5.9) can be re-derived extending the error analysis used at the end of the previous section to vector waves. Restricting our analysis to compressional waves only, we define the error ε_p along the ray,

$$\varepsilon_p \equiv \hat{p}\bullet\nabla\psi - \hat{p}\bullet\nabla\psi_0 , \qquad (5.12)$$

where $\hat{p}$ is the unit vector tangent to the ray.

To estimate ε_p we again postulate

$$\psi_0 = A_0 e^{i\omega\tau\left[1-\frac{\omega_c^2}{\omega^2}\right]^{1/2}} \quad , \quad \psi = A_0 e^{i\omega\tau} . \tag{5.13}$$

Carrying out the differentiation as indicated in (5.12) with the condition (4.4), and recalling that $ds = v\ d\tau$, we obtain

$$\varepsilon_p \approx \left[ik + \frac{1}{A_0}\frac{dA_0}{ds}\right]\varepsilon , \tag{5.14}$$

where ε is the error given in (4.14). In order for the error ε_p be small we must achieve (4.4), (5.9) and (4.15). The Fresnel 1-condition (4.15), when multiplied on both sides by ω, yields

$$\omega^2\ \omega_c^{-2} \gg \omega\ \tau . \tag{5.15}$$

This relation has a simple physical interpretation. The far field region cannot be too large, since it might interfere with the high frequency condition, introducing errors in phase. Therefore, given a frequency, a receiver cannot be arbitrarily far away from the source and is constrained by the Fresnel 1-condition (upper bound on source-receiver distance).

6. WKBJ APPROXIMATIONS

We have examined so far the ray-series method through which the scalar Helmholtz equation was approximated. A second approach which also leads to ray theory is based on the following three steps: Assume first that the velocity depends only on z.

(1) Apply a double integral transform to the wave equation, leaving an ordinary differential equation in the z coordinate.

(2) Obtain WKBJ solutions to the said equation.

(3) Approximate the inversion integral for the horizontal slowness, p, by saddle point methods.

We shall provide here a case of sufficient generality such that the limitations of the method become apparent while its interrelations to the other two methods can be recognized.

We apply the Fourier-Bessel transform to the equation

$$\nabla^2\psi + \frac{\omega^2}{v^2(z)}\psi = -2\, c_1\, \delta(z - z_0)\, \frac{\delta(r)}{r}, \tag{6.1}$$

where the transform is defined as

$$G\,(p,z;\omega) = \int_0^\infty \psi(M,\omega)\, J_0(\omega p r)\, r dr\ ,$$

where $M(r,z)$ and J_0 is the Bessel function of the first kind and of order zero.

Since G is an even function in p (axial symmetry with respect to the z axis) the solution is obtained from the inverse Fourier-Bessel transform (Chapman, 1978), and has the integral representation

$$\psi(M,\omega) = \frac{\omega|\omega|}{2}\int_{-\infty}^{\infty} G\,(p,z,z_0;\omega)\, H_0^{(1)}(\omega p r) p dp\ , \tag{6.2}$$

where $H_0^{(1)}$ is the Hankel function of the first kind and order zero.

G then obeys the ordinary differential equation

$$\frac{d^2G}{dz^2} + \omega^2 q^2 G = -2 c_1 \delta(z - z_0), \tag{6.3}$$

where $q^2(z) = v^{-2}(z) - p^2$. p and q are the horizontal and vertical components of slowness, respectively.

The solution of (6.3) is given in terms of the two independent solutions $\bar{\psi}_{1,2}(z)$ of the homogeneous equation of (6.3),

$$G(z \mid z_0;p,\omega) = \frac{c_1 \bar{\psi}_1(z_<) \bar{\psi}_2(z_>)}{\bar{\psi}_1'(z_0) \bar{\psi}_2(z_0) - \bar{\psi}_2'(z_0) \bar{\psi}_1(z_0)}, \tag{6.4}$$

where $z_< = \min(z,z_0)$, $z_> = \max(z,z_0)$. We approximate G by the WKBJ method (Bender and Orszag, 1978) and obtain

$$G_{WKBJ} \approx \frac{i c_1}{\omega} \left[q(z) q(z_0)\right]^{-1/2} e^{i\omega\tau(p)}, \tag{6.5}$$

with

$$\tau(p) = \int_{z_0}^{z} q(\zeta)\, d\zeta . \tag{6.5a}$$

The dependence of p in q is implicit, see (6.3).

Then, $H_0^{(1)}(\omega pr)$ is approximated by its asymptotic (i.e. far field) representation: $H_0^{(1)}(\omega pr) \approx (2/(\pi\omega pr))^{1/2} \exp(i\omega pr - i\pi/4)$ as $\omega pr \to \infty$. Equation (6.2) takes intermediate form

$$\psi(M,\omega) \approx \int_{-\infty}^{\infty} A(z \mid z_0;p,\omega)\, e^{i\omega T(p)}\, dp , \tag{6.6}$$

with

$$A(z \mid z_0;p,\omega) = i c_1 e^{-i\pi/4} \left[\frac{\omega p}{2\pi r\, q(z)\, q(z_0)}\right]^{1/2}, \tag{6.7}$$

$$T(p) = \tau(p) + pr \ . \tag{6.8}$$

We then apply the stationary phase approximation to (6.6) and find that the final result is

$$\psi(M,\omega) \approx B\, e^{i\omega\tau}, \tag{6.9}$$

where

$$B = c_1 \left| \frac{p_r}{r\, q_r\, q_{0r}\, |T''(p_r)|} \right|^{1/2},$$

with $p_r = \sin j / v = \sin j_0 / v_0$, $q_r = \cos j / v$, and $q_{0r} = \cos j_0 / v_0$, where the angles j's have been defined in section 3, and τ is the ray travel time, equal to (3.9) or (4.7).

B is equal to A_0 in (3.19a), since $h_{j_0} = p_r \cos j \cot j_0 \, |T''(p_r)|$. Thus WKBJ solution yields results identical to those obtained from geometrical ray theory via the eikonal and transport equations (see section 4). Note that if we use a double Fourier-transform instead of the Fourier-Bessel transform, the integral in (6.6) becomes a double integral over p_1, and p_2 where the wavefront equation is

$$T = p_1 x + p_2 y + \int_{z_0}^{z} \left(\frac{1}{v^2(z)} - p_1^2 - p_2^2 \right)^{1/2} dz. \tag{6.10}$$

Here (p_1, p_2) are the components of the slowness vector in the (x,y) directions respectively.

The stationary phase principle is equivalent to Fermat principle of stationary time and the equation for the stationary-point is in fact Snell's Law in disguise. The value of the integral at the stationary point yields the amplitude and the phase of the seismic ray.

If the velocity v depends on all three cartesian coordinates, the present method

fails. We may nevertheless obtain an *approximate* integral representation for weak inhomogeneous media similar to (6.6) if we write

$$\psi(M,\omega) \approx \int_{-\infty}^{\infty}\int A(\tau,\gamma_1,\gamma_2;\omega)\, e^{i\omega T}\, \Phi(\gamma_1,\gamma_2)d\gamma_1 d\gamma_2 \,, \tag{6.11}$$

where $T = T(M,\tau,\gamma_1,\gamma_2)$. This representation may have a number of different interpretations depending on the approximation used. The simplest way is to think of the integrand as a superposition of eikonal solutions with ray parameters (γ_1,γ_2) . $\Phi(\gamma_1,\gamma_2)$ is a weighting function.

The WKBJ approximation of (6.11) which yields the geometrical ray is (Titchmarsh, 1967),

$$\psi(M,\omega) \approx \frac{2\pi i A(\gamma_1^{(0)},\gamma_2^{(0)};\omega)}{R_w}\, e^{i\omega T(\gamma_1^{(0)},\gamma_2^{(0)})} \,. \tag{6.12}$$

Here $\gamma_1^{(0)},\gamma_2^{(0)}$ are roots of $\frac{\partial T}{\partial \gamma_1} = 0, \frac{\partial T}{\partial \gamma_2} = 0$, and

$$R_w = \omega \left[\frac{\partial^2 T}{\partial \gamma_1^2}\frac{\partial^2 T}{\partial \gamma_2^2} - \left(\frac{\partial^2 T}{\partial \gamma_1 \partial \gamma_2} \right)^2 \right]^{1/2} = \frac{\omega}{v} K^{-1/2} \,. \tag{6.13}$$

Caustics are given by the equation $R_w = 0$. The *geometrical ray* connecting the source to point M has a travel time given by $T(M,\tau,\gamma_1^{(0)},\gamma_2^{(0)}) \equiv \tau$.

The expression

$$K = \frac{1}{v^2} \left\{ \frac{\frac{\partial^2 T}{\partial \gamma_1^2}\frac{\partial^2 T}{\partial \gamma_2^2} - \left(\frac{\partial^2 T}{\partial \gamma_1 \partial \gamma_2} \right)^2}{\left[1 + v^2 \left(\frac{\partial T}{\partial \gamma_1} \right)^2 + v^2 \left(\frac{\partial T}{\partial \gamma_2} \right)^2 \right]^2} \right\}^{-1}_{\gamma_1 = \gamma_1^{(0)}, \gamma_2 = \gamma_2^{(0)}} \,, \tag{6.14}$$

is the *Gaussian curvature* of the surface $T = T(\gamma_1,\gamma_2)$ at $\gamma_1 = \gamma_1^{(0)}, \gamma_2 = \gamma_2^{(0)}$ (e.g.

Eisenhart, 1909). Clearly, (6.14) is a generalization of (3.19) for a surface which is not necessarily a surface of revolution.

7. PARAXIAL APPROXIMATIONS

In the ray series approximation we have replaced the Helmholtz wave equation by the eikonal and transport equations which give a fair account of the field everywhere at high frequencies. A different approximation is obtained if we restrict *a priori* the wavenumber such that the wave is forced to be collimated in the direction of a given axis (or a space curve). The simplest way to derive this approximation is by means of the following artifice: we may formally uncouple the Helmholtz equation $(\nabla^2 + \frac{\omega^2}{v^2})\,\psi = 0$ into two equations, governing energy flow in the positive and negative directions respectively. Thus, waves propagating in the forward z direction satisfy

$$\left\{ i\,\frac{\partial}{\partial z} + \left[k^2 + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right]^{1/2} \right\} \psi = 0 \quad , \qquad k = \omega / v \ . \tag{7.1}$$

This formal factorization can be justified rigorously and the radical can be given a precise interpretation as a *pseudo differential operator* (Bastiaans, 1979; Corones, 1975). If we expand the radical in Taylor series assuming

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \ll k^2 \,, \tag{7.2}$$

we shall have

$$\left[k^2 + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right]^{1/2} \approx k + \frac{1}{2k} \left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right] + \textit{higher terms.}$$

Consequently (7.1) will yield the *paraxial wave equation*

$$\left\{\left[\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right]+2ik\frac{\partial}{\partial z}+2k^2\right\}\psi^{par}=0, \tag{7.3}$$

where the potential ψ^{par} is a solution of (7.3), and not necessarily of the Helmholtz wave equation.

Substituting in (7.3) the ansatz $\psi^{par}=e^{ikz}u(x,y,z)$, one finds that u satisfies the equation

$$\left\{\left[\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right]+2ik\frac{\partial}{\partial z}\right\}u=0. \tag{7.4}$$

This equation is sometimes called the *parabolic (Schrödinger type) wave equation.*

Consider first, the wave propagation in homogeneous media: an exact solution of Eq. (7.3) is given by the *Fresnel diffraction integral* (Haus, 1984)

$$u(x,y,z)=\int_{-\infty}^{\infty}\int u_0(x_0,y_0)\,h(x-x_0,y-y_0,z)\,dx_0\,dy_0\,, \tag{7.5}$$

where $u_0(x_0,y_0)=u(x_0,y_0,0)$ and h is the impulse response function

$$h(x,y,z)=\frac{-ik}{2\pi z}\,e^{\frac{ik}{2z}(x^2+y^2)} \tag{7.6}$$

In the limit $z\to 0$, h goes to $\delta(x)\,\delta(y)$ and Eq. (7.5) reduces to an identity. Moreover, as long as $h(x,y,z)$ is *close* to a delta function we shall have $\psi^{par}(x,y,z)\approx e^{ikz}u_0(x,y)$ which means that there will practically be little distortion of the original signal.

In order for $h(x,y,z)$ to behave like a delta function its phase must have a strong variation over a scale-length, say l_0, in both x and y directions. Taking for simplicity, $x_0=y_0=0$, the z extent of the region over which u_0 does not vary appreciably (probably z-independent), is

$$\frac{k}{z} l_0^2 \gg 1 , \tag{7.7}$$

or

$$z \Lambda \ll 2\pi l_0^2 , \tag{7.8}$$

Λ being the wavelength.

The entity $2\pi l_0^2 / \Lambda$ defines a region (upper bound on distance z) in which the cumulative distortion of the signal is small. Recalling that $r_0 = (\Lambda z)^{1/2}$ is the radius of the *first Fresnel zone* (e.g., Savarensky, 1975), the condition (7.7) can be recast in the more physically-significant form

$$r_0^2 \ll 2\pi l_0^2 . \tag{7.9}$$

This is defined as the *Fresnel condition* .

Thus, in the paraxial approximation, a ray can be visualized as a central thread surrounded by a volume (known as the *Fresnel volume*), made of the first Fresnel zone at each point. If we enlarge the cross section of this volume at each point beyond $(2\pi)^{1/2} l_0$, the field associated with a given ray is distorted and the simple ray concept will cease to apply. The condition for the first Fresnel zone is obtained (at any point near the axis) from the requirement that the phase of a collimated signal relative to its phase on the axis does not exceed π (no destructive interference).

Since the equi-phase surfaces of $\psi^{par}(x,y,z)$ are (Eq. (7.6) with $x_0 = y_0 = 0$)

$$\Sigma = kz + \frac{k}{2z}(x^2 + y^2) = \text{const} , \tag{7.10}$$

the above condition reads

$$\Sigma(x,y,z) - \Sigma(0,0,z) = \frac{k}{2z}(x^2 + y^2) < \pi . \tag{7.11}$$

Denoting $x^2 + y^2 = r_0^2$, the inequality in (7.11) defines the *Fresnel paraboloid*

$\frac{kr_0^2}{2z} = \pi$ with a cross section $r_0 = (\Lambda z)^{1/2}$. Inequality (7.9) thus states that ray methods are not applicable whenever the distance between the source and the observer is beyond a certain threshold. We shall later extend this restriction to general inhomogeneous media as well.

Note that the impulse response $h(x,y,z)$ can be derived directly from the Green's function solution $\psi = \frac{1}{\Delta} e^{ik\Delta}$ of the scalar wave equation $\nabla^2\psi + k^2\psi = -4\pi\delta(\vec{\Delta})$ if we restrict the motion to the vicinity of the z axis

$$\Delta = (z^2 + (x-x_0)^2 + (y-y_0)^2)^{1/2} \approx z + \frac{(x-x_0)^2 + (y-y_0)^2}{2z} + O\left[\frac{\{(x-x_0)^2 + (y-y_0)^2\}^2}{8z^3}\right], \tag{7.12}$$

and

$$\frac{e^{ik\Delta}}{\Delta} \approx \left[\frac{e^{ikz}}{z}\right] e^{\frac{ik}{2z}[(x-x_0)^2 + (y-y_0)^2]} = \frac{2\pi i}{k} e^{ikz} h(x-x_0, y-y_0, z) . \tag{7.13}$$

A *sufficient* condition for the accuracy of Eq. (7.13) requires that the maximum phase change contribution by the next higher order term in the binomial expansion of $k\Delta$ to be much less than one radian. This condition will be met if the distance z satisfies

$$z^3 \gg \frac{\pi}{4\Lambda}[(x-x_0)^2 + (y-y_0)^2]^2 . \tag{7.14}$$

This requirement is, however, *not a necessary one*. For the Fresnel approximation to remain valid, it is only required that the higher order terms of the expansion not change the value of the superposition integral (7.5), and for this to be true the maximum value of the added phase factors need not be much less than one radian. If the distance z is small enough to violate (7.14), the quantity $\frac{k}{2z}$ will be so

large that the primary contribution to the integral in Eq. (7.5) will arise from points near $x = x_0$, $y = y_0$ such that (7.14) is automatically satisfied.

The Fresnel 1-condition was defined in (4.15). In a homogeneous medium containing a body of scale length l_0, condition (4.15) [with $\omega_c = \alpha / l_0$, and $z = \alpha\tau$ as the source-receiver distance], yields (7.9). The Fresnel 1-condition was derived by imposing the hypereikonal to be approximately equal to the eikonal. However, when paraxial rays are considered, simple investigation of (7.6) suggest that another Fresnel condition must take into account the wavefront curvature effect on the travel time extrapolation. The argument of the exponent is written in 2-D heterogeneous media

$$\frac{ik}{2} \frac{n^2}{R_w(s)}, \tag{7.15}$$

where n is the ray-receiver distance, and $R_w \equiv K_w^{-1}$ is the wavefront radius of curvature . By an approach similar to the homogeneous case we obtain the *Fresnel 2-condition*

$$R_w \Lambda \ll 2\pi l_0^2 \tag{7.16}$$

which yields (7.8) in the homogeneous limit. Therefore, we can conclude that in weak heterogeneous medium the generalization of the Fresnel condition in homogeneous media, is

$$\max\left(R_w , \alpha\tau \right) \Lambda \ll 2\pi l_0^2 , \tag{7.17}$$

A new definition of the first Fresnel zone radius could be postulated (see 7.9) in heterogeneous media; it is

$$r_0^2 = \Lambda \max(R_w , \alpha\tau) .$$

The solution of the parabolic wave equation in ray centered coordinate (Cerveny,

1983[a]) is

$$\psi^{par}(\omega) = c_0 \left[\frac{v}{J^S}\right]^{1/2} e^{[i\omega\tau + \frac{ik}{2} K_w n^2]}, \tag{7.18}$$

where

$$K_w = vP/Q, \tag{7.19}$$

is the total wavefront curvature, s is the arc length parameter, v is the intrinsic wave velocity on the ray and n is the normal distance off the ray. Functions P and Q are those defined in (4.8). The paraxial condition (7.2) constrains the distance n. In Appendix C we show that in a weak heterogeneous medium (7.2) can be expressed as

$$n\ K_w \ll 1, \tag{7.20}$$

which will be called the *paraxial condition*.

In media where (7.18) is valid, the inequalities in (7.16) will hold with R_w given by (7.19) and variable $v(s)$. However, since ray-centered coordinates are subjected to the restriction $n \ll R$ (regularity of the system), $R = K_R^{-1}$ being the ray local radius of curvature, the additional condition imposed on the paraxial approximation is

$$n\ K_R \ll 1. \tag{7.21}$$

8. GAUSSIAN BEAMS

An important contribution to ray theory in weak inhomogeneous media was made in recent years (Cerveny, 1982) by adapting the Gaussian beam concept to the asymptotic solution of the elastodynamic wave equation in a ray-centered coordinate system. Previously, Deschamps (1971) and Felsen (1976) have shown that a Gaussian beam can be mathematically realized by allowing the source coordinates to be complex.

The ensuing Green's function then leads naturally to the Gaussian beam concept. Gaussian beams have been used to model fault zones (Cormier and Spudich, 1984), and to model surface waves in smooth laterally varying media (Yomogida, 1984; Yomogida and Aki, 1984). We shall give now a brief summary of this method and formulate its validity conditions.

Gaussian beams represent high-frequency asymptotic solutions of the elastodynamic equation (2.1). These solutions are concentrated close to rays of P and S waves. The amplitude distribution of the principal component of the beam displacement in a plane perpendicular to the central ray is bell-shaped (Gaussian) with its maximum on the ray. The Gaussian beam solution can be viewed as an analytic continuation of the ray solution described in section 4, by allowing the phase function to be complex valued and the solution to be one way (in the direction of increasing s). Equation (2.1) reduces to the parabolic equation (similar to eq. (7.4)) in the high frequency approximation.

Rays $\{\Omega\}$ and travel times $\{\tau\}$ serve as support for the beam computations, except that rays are not required to hit the receiver. Ray centered coordinates are used. For a given ray Ω, the trial form of solution for the principal component vector of displacement, in the frequency domain, of an individual (elementary) Gaussian beam at a receiver $M(s,q_1,q_2)$ is

$$\vec{U}^{\gamma_1\gamma_2}(M,\omega) = \vec{u}^{\gamma_1\gamma_2}(s)\, e^{i\omega[\tau(s)+\frac{\vec{q}^T K_w^b(s)\vec{q}}{2v(s)}]}, \tag{8.1}$$

where $\vec{q}^T = (q_1,q_2)$, K_w^b is a 2 x 2 complex valued matrix, γ_1 and γ_2 are the ray parameters of Ω. Solution (8.1) is of high-frequency type (zero order term of an asymptotic series in inverse square-root powers of frequency).

For a given ray Ω, the eikonal equation is of the same form as (4.8), but now P,

Q and J^S are complex valued ($K_w^b = v\ PQ^{-1}$ and $J^S = \det Q$) . Re $\{K_w^b\} \equiv K_w$ controls the geometric properties of the Gaussian beam local wavefront curvature, whereas Im$\{K_w^b\}$ governs the amplitude profile of the Gaussian beam in the plane perpendicular to Ω . In this fashion the Gaussian beam solution has no singularity since J^S is complex and never vanishes even at caustics. A useful quantity is the matrix of the half-width of the Gaussian beam related to Im (K) in the following way

$$L = \left(\frac{k}{2} \text{Im}\{K\} \right)^{-1/2} , \qquad k = \omega / v . \tag{8.2}$$

We specify two linearly independent sets of real initial conditions for $P(s_0)$ and $Q(s_0)$. Generally we choose a point source (or line source) and a plane wave as independent initial conditions. We then solve separately the two real eikonal equations and reconstruct the complex solution by superimposing the two real solutions with appropriate complex-valued constant. In three dimensional heterogeneous medium, six real-valued parameters completely specify the beam. In two dimensions the number of parameters reduces to two.

The transport equation is of the same form as in the dynamic ray tracing formulation, and its solution is similar to (4.9). In terms of a divergence coefficient and for the displacement we get (see 4.12)

$$\vec{u}^{\gamma_1 \gamma_2}(s) = \Phi(\gamma_1, \gamma_2) \left(\frac{\rho_0\, v_0\, J_0^S}{\rho(s)\, v(s)\, J^S(s)} \right)^{1/2} i\omega\, \hat{p} , \tag{8.3}$$

where $\Phi(\gamma_1, \gamma_2)$ is the complex weight factor of each elementary beam, and $\hat{p}$ is the unit tangent vector to the ray Ω.

The high frequency displacement at a point M is then reconstructed by superimposing Gaussian beam elementary solutions. For the the P wave displacement $\vec{U}^b$ we

have, in the frequency domain,

$$\vec{U}^{b}(M,\omega) = \int\int \vec{U}^{\gamma_1\gamma_2}(M,\omega)\, d\gamma_1 d\gamma_2 , \tag{8.4}$$

the integration being done over all rays passing through some neighborhood of M .

If $\bar{\sigma}(\omega)$ is the spectrum of a point source, the P wave displacement in time $\vec{U}^{b}(M,t)$ is computed following

$$\vec{U}^{b}(M,t) = \frac{1}{\pi}\mathrm{Re}\left\{\int_0^{\infty} \bar{\sigma}(\omega)\,\vec{U}^{b}(M,\omega)e^{-i\omega t}\, d\omega\right\}. \tag{8.5}$$

Integrating first with respect to the ray parameters , then performing the frequency integration corresponds to the so-called spectral approach. If the frequency integration is done first the method is called the wave packet approach. Since ray parameters , γ_i, are related to horizontal slowness, the wave packet method is equivalent to the slownwess method (Chapman, 1978).

In the two dimensional case matrices P and Q are scalar complex functions as well as K_w^{b} and L . The factor $\Phi(\gamma_1,\gamma_2)$ of each beam is reduced to $\Phi(\gamma_1)$. Two real free parameters are then at our disposal. The so-called initial beam parameter B , and the distance d of the minimum of the half-width function $L(s)$ from the initial point M_0. This distance corresponds to the location at which the beam matches the wave field. We define a complex parameter $\varepsilon \equiv d - iB$. This parameter weights the contribution of the plane wave initial conditions with respect to the point source initial conditions. The existence condition of the beam imposes $B > 0$.

Equation (8.4) may be visualized as a superposition of WKBJ solutions (6.11) (Popov, 1982). It is shown by Madariaga (1984) and Klimes (1984) that, for the limiting case $B\to\infty$, the Gaussian beam solution (8.4) tends to the WKBJ solution as given by Chapman and Drummond (1983).

The determination of the factor $\Phi(\gamma_1)$ of each beam is done by matching the asymptotic form of an exact known solution to the saddle point contribution of (8.4).

Gaussian beams are subjected to all the limitations imposed so far on ray theory and three more:

(1) The coordinate system regularity condition (7.21)

$$n\, K_R \ll 1\,, \tag{8.6}$$

(2) The paraxial condition (7.20)

$$n\, K_w \ll 1\,, \tag{8.7}$$

(3) The Fresnel condition for a beam (7.17)

$$\max\left(|\, K_w^{\flat} \,|^{-1}\,, \omega\tau \right) \Lambda \ll 2\pi\, l_0^2\,. \tag{8.8}$$

The modulus of $K_w^{\flat}$ is taken since we expect that the beam width $L(s)$ must be constrained by this condition. Written as it is, we can have $L \ll l_0$, which is an appropriate condition since we are only interested in high frequency Gaussian beams (no diffraction).

9. CONCLUSION

The limitations of asymptotic wave theory in weak inhomogeneous media were formulated in a systematic way. New physical parameters characterizing medium properties have been defined. Conditions of validity were derived explicitly. They can be summarized as follows:

(1) Mode decoupling condition $\omega\, \omega_0^{-1} \gg 1$, where ω_0 is the medium *threshold frequency*.

(2) High frequency (radiation zone) condition $\omega\,\omega_c^{-1} \gg 1$, where ω_c is the medium *cut-off frequency*.

(3) Fresnel 1-condition $\omega^2\,\omega_c^{-2} \gg \omega\,\tau$, τ being the ray travel time.

(4) The far field condition $k\,A_0 \gg |\hat{p}\bullet\nabla A_0|$, where A_0 is the local plane wave amplitude.

In the paraxial approximation, especially for Gaussian beams, the validity conditions are (1-4) and:

(5) Regularity condition $n\,K_R \ll 1$. where n is the ray receiver distance, and K_R is the ray curvature

(6) Paraxial condition $n\,K_w \ll 1$, where K_w is the total wavefront curvature (real).

(7) Fresnel 2-condition $\Lambda\,|K_w^{\flat}|^{-1} \ll 2\pi\,l_0^2$, where $K_w^{\flat}$ is the total beam wavefront curvature (complex), and l_0 the medium characteristic length.

Although condition (1) and (2) were assumed to be one and the same in the present paper, we have kept them as separate conditions mainly because each of them have a different physical interpretation. Additional exact solutions for other realistic inhomogeneous media should be sought in order to enhance our understanding of wave, ray and beam propagation in the Earth. There is also an urgent need to develop new methods for the approximate analytical quadrature of the superposition beam integral.

10. APPENDIX A: Perturbation solution of (2.23) and (2.24)

A further change of variable $\varphi_i = Y_i \rho^{-1/2}$ $i = 1, 2$, reduces Eqs. (2.23) and (2.24) into the simpler form

$$\nabla^2 Y_1 - \frac{1}{v_1^2}\frac{\partial^2 Y_1}{\partial t^2} = \varepsilon_2\left[\nabla^2 - \frac{\partial^2}{\partial z^2}\right]Y_2 \qquad \varepsilon_2(z) = -\eta_{12}\,, \tag{A.1}$$

$$\nabla^2 Y_2 - \frac{1}{v_2^2}\frac{\partial^2 Y_2}{\partial t^2} = \varepsilon_1 Y_1 \qquad \varepsilon_1(z) = \eta_{21}\,. \tag{A.2}$$

In the derivation of this equation we have again neglected a term $\frac{1}{4}\rho^{-\frac{5}{2}} Y_i[(\rho')^2 - 2\rho\rho'']$ on the left hand side of each equation, in accord with our previous scheme.

Eqs. (A.1) and (A.2) represent a coupled system of wave equations for the unknown potentials Y_1 and Y_2. One may obtain a single equation of the fourth order for Y_2 by simply substituting Y_1 from (A.2) into (A.1) and dropping terms of second order in the derivatives of ρ, μ and v_2. This approach, however, complicates the solution unnecessarily. Let us assume that ε_1 and ε_2 are small everywhere and that a *perturbation scheme* of the first order is feasible such that

$$Y_1 = Y_1^0 + \varepsilon_1 Y_1^1 \;; \qquad Y_2 = Y_2^0 + \varepsilon_2 Y_2^1\,. \tag{A.3}$$

Eqs. (A.1) and (A.2) then yield

$$\nabla^2 Y_i^0 - \frac{1}{v_i^2}\frac{\partial^2 Y_i^0}{\partial t^2} = 0, \qquad \iota = 1, 2\,. \tag{A.4}$$

$$\nabla^2 Y_1^1 - \frac{1}{v_1^2}\frac{\partial^2 Y_1^1}{\partial t^2} = -\left[\nabla^2 - \frac{\partial^2}{\partial z^2}\right]Y_2^0 \;; \tag{A.5}$$

$$\nabla^2 Y_2^1 - \frac{1}{v_2^2}\frac{\partial^2 Y_2^1}{\partial t^2} = Y_1^0\,, \tag{A.6}$$

where the Y_i are calculable to order ε_i^2.

The corresponding displacements are obtained via a combined use of Eqs. (2.22) and (A.3) . For the compressional wave we have

$$\vec{U}_1 = \rho^{-1/2}\left[\nabla Y_1^0 + \varepsilon_1 \nabla Y_1^1 + \hat{e}_z Y_1^0\left(\frac{\rho'}{2\rho} - \frac{2\mu'}{\lambda + 2\mu}\right)\right], \tag{A.7}$$

and a similar equation for $\vec{U}_2$. For $\varepsilon_1 \ll 1$, $\vec{U}_1 + \vec{U}_2$ in (A.5) reduces to $\vec{U}$ in (2.21).

11. APPENDIX B: Hypereikonal of the elastodynamic equation

An exact solution of the vector elastodynamic equation for SH waves in general vertically inhomogeneous media (Ben-Menahem and Singh, 1981) is $\vec{U} = \nabla \times (\hat{e}_z \varphi)$ where $\Theta = \eta^{1/2}\varphi$ obeys the equation

$$\nabla^2 \Theta + \left(\frac{\omega^2}{v^2} - \frac{\omega_1^2}{v^2}\right)\Theta = 0, \tag{B.1}$$

and

$$\omega_1^2 = \frac{1}{2\rho}\frac{d^2\mu}{dz^2} - \frac{1}{4\mu\rho}\left(\frac{d\mu}{dz}\right)^2 . \tag{B.2}$$

The corresponding solution of Eq. (B.1) for this case is

$$\Theta \approx \left[\frac{\mu v \sin j_0}{r h_{j_0}}\right]^{1/2} e^{i\omega\tau(1 - \omega_c^2/\omega^2 - \omega_1^2/\omega^2)^{1/2}} . \tag{B.3}$$

Similar solutions exist for P and SV wave motion in weak vertically inhomogeneous media in which decoupling is maintained. For P waves:

$$\omega_2^2 = -\frac{\lambda + \mu}{\rho}\frac{dg_1}{dz} - \frac{1}{\rho}\frac{d\mu}{dz}g_1 + \frac{\lambda + 2\mu}{2\rho^2}\frac{d^2\rho}{dz^2} - \frac{\lambda + 2\mu}{4\rho^3}\left(\frac{d\rho}{dz}\right)^2 , \tag{B.4}$$

$$g_1 = \frac{\lambda + 2\mu}{\lambda + \mu}\left(\frac{1}{\rho}\frac{d\rho}{dz} - \frac{2}{\lambda + 2\mu}\frac{d\mu}{dz}\right).$$

For SV waves:

$$\omega_3^2 = \frac{\lambda + \mu}{\rho}\frac{dg_2}{dz} - \frac{1}{\rho}\frac{d\lambda}{dz}g_2 + \frac{\mu}{2\rho^2}\frac{d^2\rho}{dz^2} - \frac{\mu}{4\rho^3}\left(\frac{d\rho}{dz}\right)^2, \tag{B.5}$$

$$g_2 = \frac{\mu}{\lambda + \mu}\left(\frac{2}{\mu}\frac{d\mu}{dz} - \frac{1}{\rho}\frac{d\rho}{dz}\right).$$

12. APPENDIX C: Paraxial wave condition

We assume that the medium is locally homogeneous, and that its properties do not differ between the central ray and the observer. In equation (7.18) the phase factor can be viewed as a second order Taylor expansion of the travel time away from the central ray

$$\tau(s,n) = \tau(s,0) + \frac{n^2}{2}\frac{K_w}{v}. \tag{C.1}$$

Figure 3 leads to a geometrical value of $\tau(s,n)$

$$\tau(s,n) = \tau(s,0) + \frac{\delta R_w}{v}, \tag{C.2}$$

where $R_w \equiv K_w^{-1}$. Thus $\delta R_w = n^2 K_w / 2$. The expansion in (C.1) assumes $\delta R_w \, K_w \ll 1$, which is satisfied if

$$n \, K_w \ll 1. \tag{C.3}$$

In 2-D media, the paraxial condition (7.2) is written in the frequency-wavenumber space as

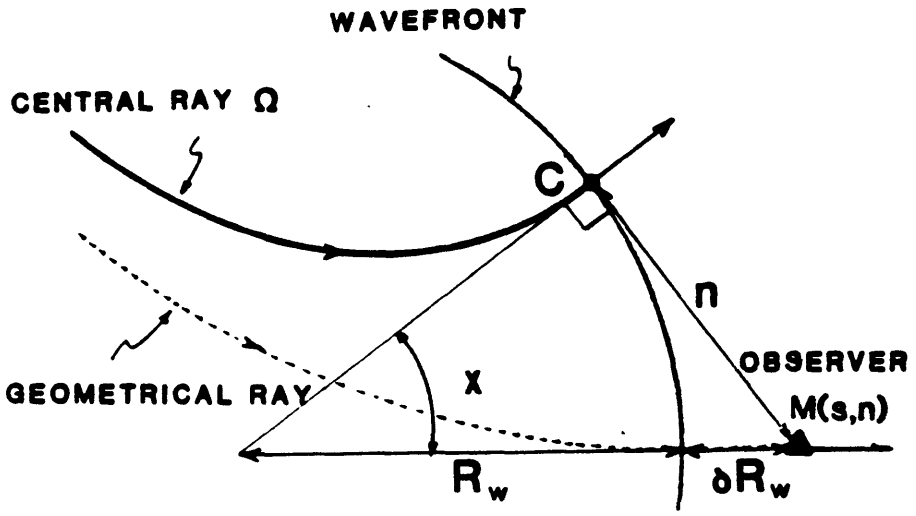
$$k_n^2 \ll k^2 \equiv \frac{\omega^2}{v^2}, \tag{C.4}$$

where k_n is the wavenumber component along the n direction (normal to the ray).

Define k_s the wavenumber component along the s direction (tangent to the ray), then $k^2 = k_n^2 + k_s^2$. The angle χ between the two vectors $\vec{k}$ and $\vec{k_s}$ links the paraxial condition (C.4) to the geometry of the wavefront (figure 3). We have

$$\sin\chi = \frac{k_n}{k} = \frac{n}{R_w + \delta R_w} . \qquad \text{(C.5)}$$

The paraxial condition is then $\sin\chi \ll 1$, and since we impose in the expansion (C.1) that $R_w \gg \delta R_w$, we fall back to (C.3). Therefore (C.3) justifies the travel time expansion (C.1) and guarantees the paraxial condition (7.2).



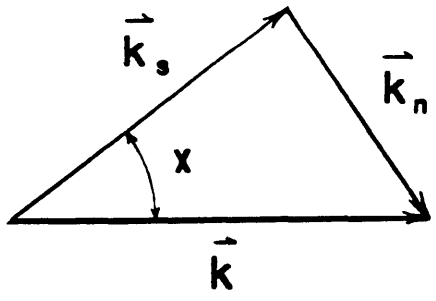


Figure 3. Geometry of the Paraxial wave condition

III. A CANONICAL PROBLEM

Science is an attempt to represent the known world as a closed system with a perfect formalism.

- J. Bronowski

1. INTRODUCTION

Asymptotic wave methods (AWM) are able to give an approximate answer to many problems of seismic body wave propagations in generally heterogeneous media (Cerveny *et al*, 1977; Aki and Richards, 1980; Ben-Menahem and Singh, 1981). They are fast, robust and, usually reduce to simple geometrical considerations. Their accuracy is based on asymptotic conditions that have been derived and described in chapter II. The results are considered reliable as long as the "asymptotic conditions are satisfied". However, when a condition "starts" failing, it is difficult to know how bad or how fast the breakdown occurs.

The most direct approach to test these conditions would be to compare AWM with an exact solution (Green's function) in the type of media considered. For example, the ray solution in a homogeneous medium can easily be derived from the Green's function (chapter II). However, the homogeneous medium is a degenerate case because it has no characteristic length, frequency etc. Closed form Green's functions are very difficult to obtain for media other than the homogeneous case. And when they are found (Hook, 1961, 1962; Lock, 1963), they represent an exact solution of the elastodynamic wave equation requiring specific conditions on the medium that often are not physical.

Chapter III

The next approach is to compare AWM with wave methods obtained by numerical procedures (finite difference, finite elements, reflectivity method etc.). But here again, numerical conditions (medium discretization, source periodicity etc.) must be fulfilled in order to get accurate results. Further, the comparison can be undertaken economically only for small-sized medium (near-field methods).

The present chapter deals with a comparison of AWM using the first approach. A second canonical problem (other than the homogeneous case) is considered. The inhomogeneous medium is that of a constant gradient model where exact mode-decoupling exists. Hook (1962) treated several problems concerning the separability of the elastodynamic wave equation, and showed that Green's functions, for an impulsive point source, can be obtained in this medium. The aim is to define clearly the range of validity of AWM in this medium. Two asymptotic wave methods are described. The dynamic ray tracing (DRT) method developed by Cerveny (1981), and its generalization the Gaussian beam (GB) method are set up in this medium. The DRT solution obtained analytically from the ray formulation is re-derived from the Green's function with the ray conditions developed in chapter II. Gaussian beam parameters are defined and discussed. Beam conditions derived in chapter II are shown to be necessary and sufficient. Comparisons between DRT and GB are shown. As an application of the beam method, a vertical seismic profiling (VSP) section with offset from an explosive source at the surface and a surface seismic section from a shallow earthquake are presented.

2. CONSTANT GRADIENT MODEL

Considering the propagation of elastic waves in isotropic media with physical parameters that depend on the cylindrical coordinate z alone (oriented downwards) and

with constant Poisson ratio, the elastodynamic wave equation reduces to (Hook, 1962)

$$\rho \frac{\partial^2 \vec{U}}{\partial t^2} = \gamma \nabla(\mu \nabla \cdot \vec{U}) - \nabla \times (\mu \nabla \times \vec{U}) + 2\mu' [\vec{U}' - \hat{e}_z \nabla \cdot \vec{U} + \hat{e}_z \times (\nabla \times \vec{U})] + \rho \vec{F} , \quad (2.1)$$

where $\gamma = (\lambda+2\mu)/\mu = \alpha^2 / \beta^2$ is constant, $\hat{e}_z$ is the unit vector in the z direction, and a prime denotes partial differentiation with respect to z. $\vec{F}$ is a body force distribution per unit mass.

Looking for a particular solution of equation (2.1) let us assume that it has the form

$$\vec{U} = \frac{1}{f_1} \nabla (f_2 \varphi_1) - \frac{1}{f_3} \nabla \times [f_4 \nabla \times (\hat{e}_z \varphi_2)] - \nabla \times (\hat{e}_z \varphi_3) , \quad (2.2)$$

and that the source is specified as an explosion at $\vec{r}_0 (r=0, z=z_0)$, with strength F_0 (dimensionless quantity), where r denotes the radial distance of a point to the z axis. We have

$$\vec{F}(\vec{r};t) = \frac{4\pi F_0 \alpha^2}{f_1} \nabla \left(f_2 \delta(\vec{r} - \vec{r}_0) \right) \delta(t) , \quad (2.3a)$$

or equivalently

$$\vec{F}(r,z;t) = \frac{2F_0 \alpha^2}{f_1} \nabla \left(f_2 \frac{\delta(r)}{r} \delta(z - z_0) \right) \delta(t) . \quad (2.3b)$$

Here f_i's are dimensionless functions of z, and φ_i's are potentials whose dependence on the coordinates and on time are unrestricted. If the seismic moment is m_0, then $m_0 = -4\pi \rho_0 \alpha_0^2 F_0 \times unit\ volume$. Hook (1961) has shown that there exist a wide class of elastic media for which the above representation is feasible. In these media the potentials satisfy the coupled differential equations

$$\nabla^2\varphi_1+\xi_1\varphi_1{}'+\xi_2\varphi_1-\alpha^{-2}\frac{\partial^2\varphi_1}{\partial t^2}=\gamma^{-1}\varepsilon_2(\nabla^2-\frac{\partial^2}{\partial t^2})\varphi_2-\frac{2F_0}{r}\delta(r)\delta(z-z_0)\delta(t), \quad (2.4)$$

$$\nabla^2\varphi_2+\xi_1\,\varphi_2{}'+\xi_3\,\varphi_2-\beta^{-2}\frac{\partial^2\varphi_2}{\partial t^2}=\varepsilon_1\,\varphi_1\,, \quad (2.5a)$$

$$\nabla^2\varphi_3+\frac{\mu'}{\mu}\varphi_3{}'-\beta^{-2}\frac{\partial^2\varphi_3}{\partial t^2}=0\,, \quad (2.5b)$$

where ∇^2 denotes the Laplacian operator, the ξ_i are functionals of the constitutive medium parameters and their derivatives, and ε_1, ε_2 are coupling factors with similar functional dependence. If $f_1\equiv f_2$ and $f_3\equiv f_4$, then $\xi_1=\rho'/\rho$, the prime denoting partial differentiation with respect to z. The waves represented by φ_3 are called SH waves, and are not excited by the source (2.3).

In the present chapter we shall be interested in a medium where the compressional and shear wave velocities are linear functions of the coordinate z (figure 1). For a parameter $\eta > 0$ and $\alpha, z, z_0 \geq 0$

$$\begin{aligned}
\alpha(z) &= \alpha(0)\,(1+\eta z)\,,\\
\beta(z) &= \beta(0)\,(1+\eta z)\,,\\
\rho(z) &= \rho(0)\,(1+\eta z)^{\alpha}\,,\\
\mu(z) &= \mu(0)\,(1+\eta z)^{\alpha+2}\,,\\
\lambda(z) &= \lambda(0)\,(1+\eta z)^{\alpha+2}\,,\\
\gamma &= \frac{3+\alpha}{1+\alpha/2}\,,
\end{aligned} \quad (2.6)$$

where $\alpha(0) = \dfrac{\alpha_0}{1+\eta z_0}$ and $\rho(0) = \dfrac{\rho_0}{(1+\eta z_0)^{\alpha}}$.

$\mu(0)$, $\lambda(0)$, and $\beta(0)$ values follow from relationships between the elastic constants.

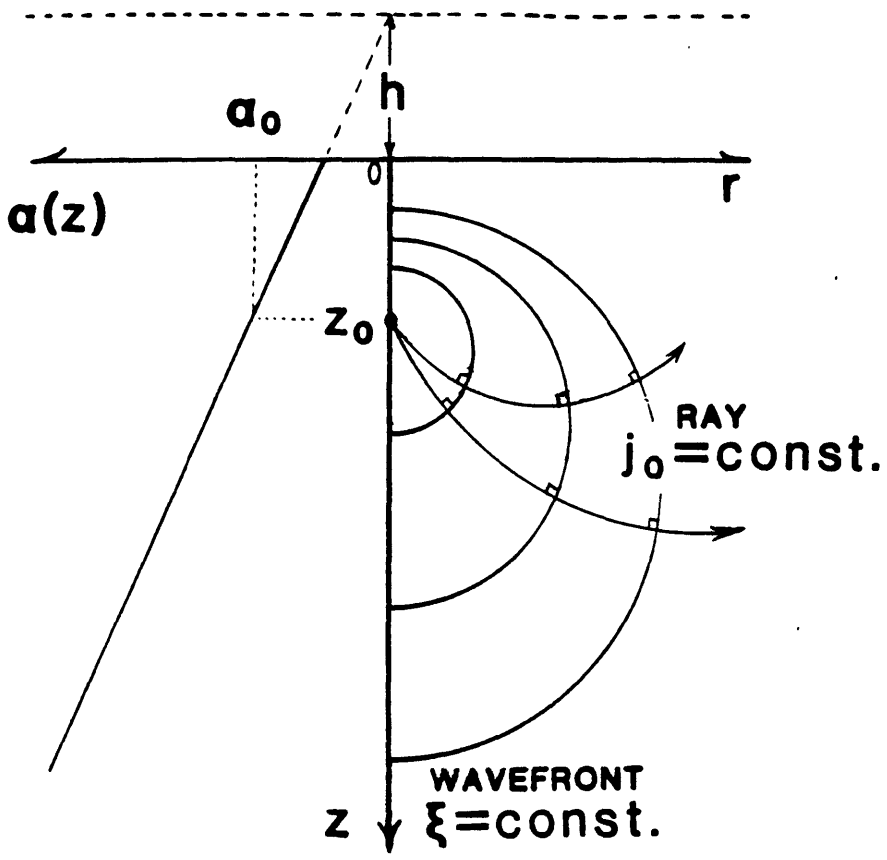


Figure 1. Wavefront and rays in the canonical model

The source is located at $(r=0, z=z_0)$ where the P velocity is α_0 and the density is ρ_0. For convenience of notation we set $h = \eta^{-1}$. The surface $z=0$ is not a free surface. The region $z \geq 0$ is the restriction of a medium whose properties still obeys equations (2.6) for $-h \leq z < 0$ (figure 1). Therefore the solutions do not include waves other than the direct wave (no surface waves, multiply reflected waves etc.). Note that $\alpha(-h)=0$.

In this medium $\xi_1 = \mu'/\mu$ and $\xi_2 = \xi_3 = \varepsilon_1 = \varepsilon_2 \equiv 0$, and equation (2.4) reduces to

$$\nabla^2 \varphi_1 + \frac{2+a}{z+h} \varphi_1' - [\alpha(0)(1+\eta z)]^{-2} \frac{\partial^2 \varphi_1}{\partial t^2} = - \frac{2F_0}{r} \delta(r)\delta(z-z_0)\delta(t) . \qquad (2.7)$$

The prime denoting partial differentiation with respect to z. Here waves represented by φ_2 in equation (2.5a) are the SV waves and are not excited by the source function (2.3). The P wave source generates only P waves (perfect decoupling), and since no boundaries are present, only P waves will propagate. The displacement vector in (2.2) reduces to

$$\vec{U} = \frac{1}{f_1} \nabla (f_2 \varphi_1) , \qquad (2.8)$$

where $f_1 = f_2^2$ and $f_2 = \frac{z_0+h}{z+h}$.

Define the Fourier transform of a time function $\varphi(t)$ as

$$\overline{\varphi}(\omega) = \int_{-\infty}^{\infty} \varphi(t)\, e^{i\omega t}\, dt ,$$

and conversely, the inverse Fourier transform of a spectrum $\overline{\varphi}(\omega)$ as

$$\varphi(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \bar{\varphi}(\omega)\, e^{-i\omega t}\, d\omega.$$

Transforming equation (2.7) into the frequency domain, the equation for the Fourier-transformed potential $\bar{\varphi}_1$ becomes

$$\nabla^2 \bar{\varphi}_1 + \frac{2+a}{z+h} \bar{\varphi}_1' + \kappa_a^2 \bar{\varphi}_1 = -\frac{2F_0}{r} \delta(r)\delta(z-z_0), \tag{2.9}$$

with $\kappa_a^2 = \left[\frac{\omega f_2}{\alpha_0}\right]^2$.

The explicit solution of equation (2.9) is expressed in terms of the bi-spherical coordinate ξ and the cylindrical coordinate z (Pekeris, 1946) and is found to be

$$\bar{\varphi}_1(\omega) = F_0 \left[\frac{z_0+h}{z+h}\right]^{(a+3)/2} \frac{e^{i\omega\tau(1-\omega_c^2/\omega^2)^{1/2}}}{R_w}, \tag{2.10}$$

where

$t_0 = \frac{z+h}{\alpha(z)} = [\eta\ \alpha(0)]^{-1} = (2\omega_0)^{-1}$,

$\omega_c = (1+a)\ \omega_0$,

$R_w = (z_0+h)\ \sinh\xi$,

$\tau = \xi\ t_0$.

Rays and wavefronts are exactly determined in terms of the bi-spherical coordinate system (ξ, j_0). The third bi-spherical coordinate denoting the angle of rotation about the z-axis is not present because of the axial symmetry of the problem. The travel-time along the ray from the source to the receiver is $\tau = \xi\ t_0$. The ray

curvature at the receiver is expressed as $R^{-1} = \sin j_0/(z_0+h)$. Rays are curves with j_0 = constant, since in a medium with constant velocity gradient the ray curvature is constant. These are arcs of circles with centers at $(r_t = (z_0+h)\cot j_0, z=-h)$ and radii R. Consequently, the coordinate of the ray turning point is $(r_t, z_t = R-h)$. Wavefronts are surfaces with ξ = constant (constant travel time). These are spheres with centers at $(r=0, z=z_c)$ and radii R_w. The coordinate ξ can be viewed as a normalized arclength $(d\xi = ds/(z+h))$ or a normalized travel time (τ/t_0) along a specified ray with takeoff angle j_0.

Define

$$R_1 = [r^2 + (z-z_0)^2]^{1/2}, \quad \text{and} \quad R_2 = [r^2 + (z+z_0+2h)^2]^{1/2}.$$

The relations between the two systems (ξ, j_0) and (r, z) are given by the equations

$$r = \frac{(z_0+h)\sin j_0}{\coth\xi - \cos j_0},$$

$$z + h = \frac{z_0 + h}{\sinh\xi(\coth\xi - \cos j_0)},$$

$$\xi = 2\tanh^{-1}\left[\frac{R_1}{R_2}\right],$$

$$j_0 = \sin^{-1}\left[\frac{2r(z_0 + h)}{R_1 R_2}\right],$$

$$R_w = \frac{R_1 R_2}{2(z + h)},$$

$$R = \frac{R_1 R_2}{2r}.$$

The geometrical meaning of R_1, R_2, and R is shown in figure 2.

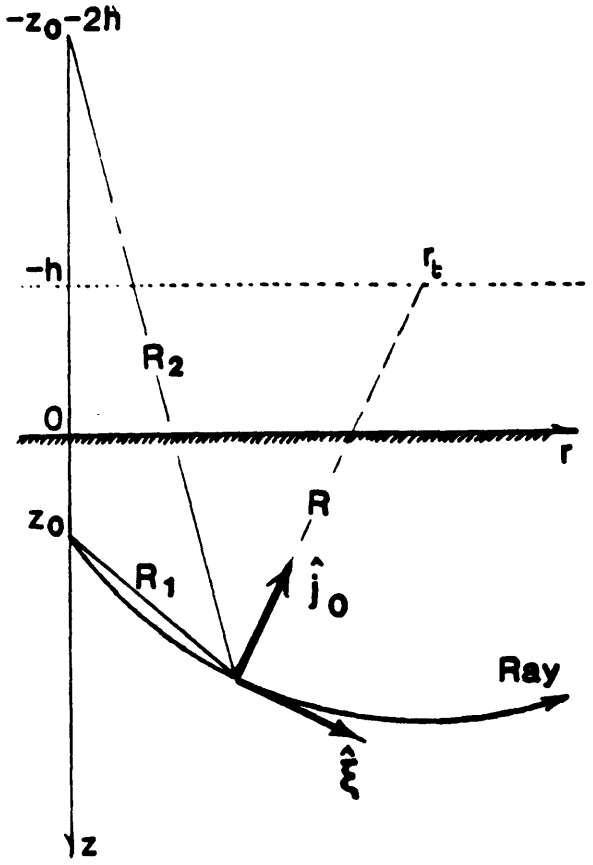


Figure 2. Geometrical representation of quantities in the canonical model

$R_w^{-1} \equiv K_w$ is defined in part I as the total curvature of the wavefront. Here, since only this definition is used, we shall refer to it as the wavefront curvature. The two medium's characteristic frequencies ω_0 and ω_c are the same as those defined in chapter II. The definition, in chapter II, of the threshold frequency is

$$\omega_0 \equiv \frac{1}{2}\max\left(|\nabla\alpha|, |\nabla\beta|, \alpha\frac{|\nabla\rho|}{\rho}\right) = \frac{\eta}{2}\,\alpha(0)\,\max(1,\gamma^{-1/2},\alpha)\,.$$

Anticipating that the density variations are small ($\alpha \leq 1$), we obtain $\omega_0 = \alpha(0)\,\eta/2$. The cut-off frequency is defined, in chapter II equation (3.5), as $\omega_c \equiv \alpha\,|\nabla^2 A_0 / A_0|^{1/2}$, where A_0 is the amplitude function of the local plane wave, appearing in the transport equation. ω_c is shown to be equal in Chapter II equation (3.27), to ω_0 in a constant gradient medium where the density is constant ($\alpha = 0$) which is directly verified in this context. The medium characteristic length $l_0 \equiv \alpha/\omega_0$, is equal to $2(z+h)$.

The inverse Fourier transform of equation (2.10) yields (Abramowitz and Stegun, 1972)

$$\varphi_1(t) = F_0 \left(\frac{z_0+h}{z+h}\right)^{(\alpha+3)/2} \frac{g_0(t)}{R_w}\,, \qquad (2.11)$$

where

$$g_0(t) = \delta(t-\tau) - \frac{(\alpha+1)^2\,\tau}{4}\,G_1(\zeta)\,H(t-\tau)\,,$$

$$G_1(\zeta) = J_1(\zeta)/\zeta\,,$$

with $\zeta = -\frac{\alpha+1}{2}\,(t^2-\tau^2)^{1/2}$.

Here δ and H are, respectively, the Dirac delta and Heaviside unit-step generalized functions and J_1 is the Bessel function of the first kind and of order one.

The second term in $g_0(t)$ is attributed to the presence of dispersion in inhomogeneous media. The dispersive properties of this medium are controlled by the cut-off frequency ω_c.

The displacement vector is obtained directly from equation (2.8) by applying the gradient operator to $\varphi_1(t)$ or to $\overline{\varphi}_1(\omega)$. Following Hook (1962), expressing this vector in terms of the mixed, non-orthogonal, coordinate system, yields

$$\vec{U} = \frac{1}{f_1}\left(h_\xi^{-1}\,\frac{\partial}{\partial\xi}\,\hat{\xi} + \frac{\partial}{\partial z}\,\hat{z}\right)(f_2\,\varphi_1)\,, \tag{2.12}$$

where $h_\xi = z + h$ is the scale factor for the coordinate ξ.

In the frequency domain

$$\vec{U}(\omega) = \overline{u}_\xi(\omega)\,\hat{\xi} + \overline{u}_z(\omega)\,\hat{z}\,, \tag{2.13}$$

where

$$\overline{u}_\xi(\omega) = F_0\left(\frac{z_0+h}{z+h}\right)^{(a+3)/2}\frac{i\omega t_0\,(1-\omega_c^2/\omega^2)^{1/2}-z_c/R_w}{R_w\,(z_0+h)}e^{i\omega\tau\,(1-\omega_c^2/\omega^2)^{1/2}}, \tag{2.14}$$

and

$$\overline{u}_z(\omega) = -F_0\left(\frac{z_0+h}{z+h}\right)^{(a+3)/2}\left(\frac{a+5}{2}\right)\frac{1}{R_w\,(z_0+h)}\,e^{i\omega\tau\,(1-\omega_c^2/\omega^2)^{1/2}}. \tag{2.15}$$

In the time domain

$$\vec{U}(t) = u_\xi(t)\,\hat{\xi} + u_z(t)\,\hat{z}\,, \tag{2.16}$$

where

$$u_\xi(t) = -F_0 \left(\frac{z_0+h}{z+h} \right)^{(a+3)/2} \frac{g_1(t) + z_c\, g_0(t)/R_w}{R_w\,(z_0+h)}, \tag{2.17}$$

and

$$u_z(t) = -F_0 \left(\frac{z_0+h}{z+h} \right)^{(a+3)/2} \left(\frac{a+5}{2} \right) \frac{g_0(t)}{R_w\,(z_0+h)}, \tag{2.18}$$

and where

$$z_c = (z_0+h)\cosh\xi, \tag{2.19}$$

$$g_1(t) \equiv t_0 \frac{\partial g_0(t)}{\partial t} = t_0\delta'(t-\tau) - t_0 \frac{(a+1)^2}{4} \left[\frac{\tau}{2}\delta(t-\tau) - G_2(\zeta)\, H(t-\tau) \right],$$

$$G_2(\zeta) = G_1(\zeta) + \left[\frac{a+1}{2\zeta}\tau \right]^2 J_2(\zeta),$$

where $\delta'(t)$ denotes the derivative of $\delta(t)$, and J_2 is the Bessel function of the first kind and of order two.

Note the strong resemblance of (2.14) with the frequency domain solution of the Green's function in a homogeneous medium (see II.5.1).

The components of the vector particle displacement of the P wave in an orthonormal coordinate system (r,z) is of interest. Therefore, if $\vec{U}$ is expressed in the non-orthogonal coordinate system (ξ,z), the vertical component in the orthonormal coordinate system (r,z) is the covariant component of $\vec{U}$ along the z axis

$$G_z \equiv \vec{U} \cdot \hat{z} = u_\xi\, g_{z\xi} + u_z, \tag{2.20}$$

and the radial component is the covariant component of $\vec{U}$ along the r axis

$$G_r \equiv \vec{U} \cdot \hat{r} = u_\xi g_{r\xi}, \tag{2.21}$$

where

$$g_{z\xi} = \hat{\xi} \cdot \hat{z} = \frac{1}{z+h} \frac{\partial z}{\partial \xi},$$

and

$$g_{r\xi} = \hat{\xi} \cdot \hat{r} = \frac{1}{z+h} \frac{\partial r}{\partial \xi}.$$

From the relations between z, r and ξ it follows that

$$g_{z\xi} = \frac{\cosh\xi \cos j_0 - \sinh\xi}{(\cosh\xi - \sinh\xi \cos j_0)}, \tag{2.22}$$

$$g_{r\xi} = \frac{r}{R_w}. \tag{2.23}$$

Note that if $j \equiv \cos^{-1}(\hat{\xi} \cdot \hat{z})$ is the angle of incidence of the ray, $\cos j \equiv g_{z\xi}$ and $\sin j \equiv g_{r\xi}$. At the ray turning point we have $R=z+h$, leading to $g_{z\xi}=0$, therefore the only vertical contribution in (2.20) comes from u_z.

If $\sigma(t)$ specifies the point source time function, then the P wave component of the particle displacement $W(M,t)$ at a point $M(r,z)$ is given by

$$W(M,t) = G_z(M,t) * \sigma(t), \tag{2.24}$$

where * denotes a convolution operation. We have a similar equation for the radial component replacing G_z by G_r.

We shall compute the synthetic seismogram resulting from (2.24), and compare it to that calculated by asymptotic wave theory.

Chapter III

3. THEORETICAL SEISMOGRAMS BY THE DYNAMIC RAY TRACING METHOD

In this section the eikonal and transport equation will be solved analytically for the medium described in section 2. Then, we will show that, with the ray conditions developed in chapter II, the Green's function derived in section 2 reduces to the ray solution.

We shall refer to the in-plane as the plane spanned by the ray trajectory, and the out-of-plane the plane perpendicular to it. The in-plane dynamic ray tracing system, in the constant velocity gradient medium described in section 2, for P waves and for a point source, can be written as (see II.4.8)

$$\frac{dP(s)}{ds} = 0 , \qquad (3.1)$$

$$\frac{dQ(s)}{ds} = \alpha(s)\, P(s) .$$

The out-of-plane dynamic ray tracing system in this type of medium is exactly of the same form as (3.1) replacing Q by $Q^{\perp}$ and P by $P^{\perp}$. The initial conditions of this system for a point source are:

$$P(s_0) = P^{\perp}(s_0)/\sin j_0 = \alpha_0^{-1} , \qquad (3.2)$$

$$Q(s_0) = Q^{\perp}(s_0) = 0 .$$

Solving this system, yields:

$$P(s) = P^{\perp}(s)/\sin j_0 = \alpha_0^{-1} , \qquad (3.3)$$

$$Q(s) = Q^{\perp}(s)/\sin j_0 = \alpha_0^{-1} \int_{s_0}^{s} \alpha(s)\, ds .$$

The functions Q and P have the dimensions of $length$ and $time^{-1}$, respectively.

If this integral is evaluated, the P wave ray displacement in the frequency domain is given by (II.4.12)

$$\vec{U}^r(\omega) = C \left[\frac{\rho_0 \alpha_0 J_0^S}{p \alpha J^S} \right]^{1/2} i\omega e^{i\omega\tau} \hat{\xi}, \tag{3.4}$$

where C is a constant depending on source initial conditions, $\hat{\xi}$ is the unit vector tangent along the ray, and τ the travel time. The source neighborhood is assumed to be homogeneous at least up to the unit sphere surrounding the source. We have (II.3.23) $J_0^S \equiv \sin j_0$, and $(J^S / J_0^S)^{1/2} = (Q\, Q^{\perp} / \sin j_0)^{1/2} = Q$. For the medium properties described in section 2, (3.4) yields

$$\vec{U}^r(\omega) = C \left[\frac{z_0 + h}{z + h} \right]^{(\alpha+1)/2} \frac{i\omega e^{i\omega\tau}}{Q} \hat{\xi}. \tag{3.5}$$

Rewriting the equation for Q in terms of the angle of incidence of the ray, j, yields

$$Q(j) = \alpha_0^{-1} \int_{j_0}^{j} \alpha(s) \frac{ds}{dj} dj, \tag{3.6}$$

ds / dj is the radius of curvature of the ray. In vertically heterogeneous media the ray parameter $p = \sin j / \alpha$ is constant along a ray. Thus

$$d\alpha = \cos j \frac{dj}{p} = \frac{dz}{ds} \frac{dj}{p},$$

yielding

$$\frac{dj}{ds} = p \frac{d\alpha}{dz}.$$

Since $\alpha(z) = \alpha(0)(1+\eta z)$ we have

$$\frac{ds}{dj} = \frac{z_0+h}{\sin j_0},$$

which is constant. It is easily verified that ds/dj is indeed equal to R given in section 2. Snell's law reads $\sin j_0/\alpha_0 = \sin j/\alpha$, and applying it to equation (3.6) renders

$$Q(s) = \frac{R}{\sin j_0}(\cos j_0 - \cos j) = \frac{r}{\sin j_0}. \tag{3.7}$$

It is equivalent to specify the function Q as a function of s or of ξ, since both are curvilinear coordinates along a specified ray. The radial distance $r(\xi)$ contains the only dependence on ξ.

Since $\sin j_0 = (z_0+h)/R$ and $Rr = (z+h)R_w$, equation (3.5) becomes

$$\vec{U}^r(\omega) = C\left(\frac{z_0+h}{z+h}\right)^{(a+3)/2} \frac{i\omega e^{i\omega\tau}}{R_w}\hat{\xi}. \tag{3.8}$$

In the time domain,

$$\vec{U}^r(t) = -C\left(\frac{z_0+h}{z+h}\right)^{(a+3)/2} \frac{\delta'(t-\tau)}{R_w}\hat{\xi}. \tag{3.9}$$

As the main focus is on the vertical component of the P wave, the orthogonal projection of vector $\vec{U}^r$ on $\hat{z}$ yields $\vec{U}^r\cdot\hat{z} = U^r \cos j$ where $\cos j = \hat{\xi}\cdot\hat{z}$. From Snell's law we get $\sin j = \frac{r}{R_w}$, leading to

$$\cos j = \left[1 - \left(\frac{r}{R_w}\right)^2\right]^{1/2}. \tag{3.10}$$

It is easily verified that $\cos j \equiv g_{z\xi}$ which is given in equation (2.22).

If the point source time function is $\sigma(t)$, and for an isotropic source radiation

pattern, the P wave vertical displacement $W^r(M,t)$ at a point $M(r,z)$ is computed via

$$W^r(M,t) = -C \left[\frac{z_0+h}{z+h}\right]^{(a+3)/2} \frac{\sigma'(t-\tau)}{R_w} \cos j. \qquad (3.11)$$

The radial component of displacement is obtained simply by replacing $\cos j$ in (3.11) by $\sin j$.

The ray curvature is in general equal to

$$K_R = -\frac{1}{\alpha}\frac{\partial \alpha}{\partial n},$$

where n is along the direction normal to the ray, oriented $\hat{n} = d\hat{\xi}/d\xi$. This direction is identical to the j_0 direction introduced in section 2. Recalling that the scale factor for the j_0 coordinate is $h_{j_0} = (z_0+h)(\coth\xi - \cos j_0)^{-1}$, it can be verified that $K_R = R^{-1}$. Further note that, in general, the curvature of the wavefront $K_w \equiv \alpha P/Q$ is here equal to $\alpha(\alpha_0 Q)^{-1}$ and is indeed equal to R_w^{-1}.

The exact solution (Green's function) for the displacement field in the constant gradient model described in section 2 (equations (2.14) and (2.15)) reduces to (3.8) using the ray conditions (see II). These are the following: (1) mode decoupling condition $\omega/\omega_0 \gg 1$; (2) high frequency (radiation zone) condition $\omega/\omega_c \gg 1$; (3) Fresnel 1-condition $\omega^2 \omega_c^{-2} \gg \omega\tau$; (4) far field condition $k A_0 \gg |dA_0/ds|$; where $A_0 = c_0 (\alpha/J^S)^{1/2}$ (see II.4.9) or explicitly

$$A_0 = c_0 \left[\frac{\alpha_0}{r R_w}\right]^{1/2},$$

where c_0 is an integration constant (see II.4.9).

In this special medium we have perfect P/S decoupling, and condition (1) is linked to condition (2) by $\omega_c = (1+\alpha)\omega_0$. Therefore violating (1) will not, in this case, yield mode coupling but will violate (2), thus rendering non-oscillatory waves. Condition (2) approximates $(1-\omega_c^2/\omega^2)^{1/2}$ by 1. Taylor-expanding the argument in the exponent of (2.14), with condition (2), we obtain (3). This is the Fresnel 1-condition that we have derived in chapter II. It is necessary since condition (2) must also approximate $i\omega\tau(1-\omega_c^2/\omega^2)^{1/2}$ by $i\omega\tau$. Expressing condition (4) in the coordinate system $(\hat{\xi},\hat{z})$, and recalling that $dA_0/ds = (z+h)^{-1}\, dA_0/d\xi$, yields

$$\omega t_0 \gg \frac{z_c}{R_w}, \tag{3.12}$$

where $t_0 = (z+h)/\alpha$.

This condition implies that we neglect the terms z_c/R_w in (2.14) (near field term) with respect to ωt_0 (far field term). Equation (3.12) can be written in terms of $\omega_0 = (2t_0)^{-1}$, and since the right hand side of (3.12) is always greater or equal to 1, this condition includes ray conditions (1) and (2) (provided $\alpha << 1$). Therefore, failure of condition (2) (or (1)) would necessarily fail condition (4). This can be explained by the fact that in equation (2.14) and (2.15), as ω gets closer to ω_0, first the near field terms cannot be neglected compared to the far field term. Consequently, the approximation of the phase term by $\omega\tau$ in the exponent gradually breaks down. Therefore increasing the velocity gradient (ω_0) leads to a far field breakdown followed by a phase distortion, due to medium's dispersion properties. In the time domain, this phase distortion effect is caused by the approximation of $g_0(t)$ by $\delta(t-\tau)$.

At a ray turning point $\cos j = 0$. The vertical displacement (3.11) vanishes. But the exact solution gives a non-zero contribution (in 2.20). This does not mean that the ray result is inaccurate in this zone. It is easily shown that, if the far field and the high

frequency conditions are satisfied, the magnitude and direction of the exact vector displacement is correctly represented by DRT. We shall verify this statement numerically, in section 7.

The decoupling condition (1) can be thought as requiring the ray radius of curvature to be much greater than the wavelength, the high frequency condition (2) as imposing the displacement field for P waves to be along the ray, and the far field condition (4) can be seen as requiring the wavefront curvature change along the ray times the wavelength to be much smaller that the wavefront curvature. The Fresnel 1-condition requires that the far field region must not exceed a certain upper limit governed by the high frequency condition (2). This imposes a limit on how far the observer should be. In the present medium, conditions (1) and (2) are equivalent (since $\omega_c = (1+\alpha)\omega_0$). Whether or not the threshold frequency is of same order as the cut-off frequency in a general medium, requiring therefore only one medium characteristic frequency, is still unknown. However we have postulated in chapter II that these two values are equal in the context of ray and paraxial ray (or beam) methods.

Comparison of the asymptotic solution of the Green's function with the ray solution determines the value of the constant C introduced in (3.4)

$$C = F_0 \frac{t_0}{z_0 + h} = \frac{F_0}{\alpha_0} .$$

This relation links F_0 to c_0 and c_1 defined in Chapter II (4.11-12): $c_1 = F_0 \rho_0^{1/2}$.

The Fresnel 2-condition $R_w \Lambda \ll 2\pi l_0^2$, was not necessary in deriving the ray solution from the Green's function. We shall test this condition in section 7.

4. THEORETICAL SEISMOGRAMS BY THE GAUSSIAN BEAM METHOD

The Gaussian beam formulation will be set-up for the medium of section 2. Various beam parameters and initial conditions are defined. Beam conditions of validity are presented. The elementary Gaussian beam solution, for a point source in a two-dimensional medium, can be written in the form (see II.8.1 and II.8.3)

$$\vec{U}^{j_0}(M,\omega) = \Phi \left[\frac{\rho_0 \alpha_0 J_0^S}{\rho \alpha J^S} \right]^{1/2} i\omega e^{i\omega [\tau + n^2 P/2Q]} \hat{\xi}_{j_0} , \tag{4.1}$$

where j_0 is the ray parameter (takeoff angle) characterizing the central ray, τ is the travel time along the ray, $\hat{\xi}_{j_0}$ is the unit vector tangent to the central ray, n is the distance from the receiver to the central ray. $J_0^S = \sin j_0$, and the complex Jacobian J^S is equal to $Q\ Q^{\perp}$ and

$$\begin{aligned} P(s) &= Z\,(P_2(s) + \varepsilon\, P_1(s)) , \\ Q(s) &= Z\,(Q_2(s) + \varepsilon\, Q_1(s)) , \\ P^{\perp}(s) &= P_2^{\perp}(s) , \\ Q^{\perp}(s) &= Q_2^{\perp}(s) , \end{aligned} \tag{4.2}$$

where subscripts 1 and 2 in (4.2) denote two linearly independent sets of real initial conditions and solutions of the complex eikonal, Z is a complex constant and $\varepsilon = d - i\,B$ (both values are dimensionless). The two beam parameters d and B are those introduced in chapter II, and will be specified later in this section. Note that $B > 0$ (existence condition of beam). In this case

$$\Phi = C \left[\frac{\omega\, Q(s_0)}{2\pi\alpha_0} \right]^{1/2} e^{i\pi/4} ,$$

is the weight factor for a point source, determined by matching the steepest descent contribution of (4.13) to the Green's function of a point source in a homogeneous medium.

P_2, Q_2, $P^{\perp}$ and $Q^{\perp}$ are the point source solutions of the dynamic ray tracing system with same initial conditions and solutions as in section 3. P_1 and Q_1 are the plane wave solutions of the dynamic ray tracing with initial conditions:

$$P_1(s_0) = 0\,, \tag{4.3}$$

$$Q_1(s_0) = z_0 + h\,,$$

referred to as the plane wave initial condition (constant spreading and zero wavefront curvature). The solution of this system is

$$P_1(s) = 0\,, \tag{4.4}$$

$$Q_1(s) = z_0 + h\,.$$

Madariaga (1984) proposed a modified plane wave initial condition, satisfying the WKBJ solution, which we correct for dimension

$$P_1^M(s_0) = -\frac{(z_0+h)\sin^2 j_0}{\alpha_0^2 \cos j_0}\left[\frac{\partial \alpha}{\partial z}\right]_{s=s_0} = -\frac{\sin^2 j_0}{\alpha_0 \cos j_0}\,, \tag{4.5}$$

$$Q_1^M(s_0) = z_0 + h.$$

Here j_0 is the takeoff angle of the ray measured from the z-axis. The solution of this system is

$$P_1^M(s) = -\frac{\sin^2 j_0}{\alpha_0 \cos j_0}\,, \tag{4.6}$$

$$Q_1^M(s) = z_0 + h - \tan j_0\, r\,.$$

The solution (4.6) can generaly be constructed by a linear superposition of the two solutions (P_1,Q_1) and (P_2,Q_2). The plane wave caustic occurs at $r = (z_0+h)\cot j_0$ ($Q_1^M=0$). This is the radial distance of the ray turning point. Indeed, this distance is identical to r_t given in section 2. In the neighborhood of the ray turning point, $Q \approx Z\ Q_2$, the beam propagation is mainly due to the point source contribution. But the initial conditions (4.5) are set up for quasi-plane waves. These conditions are therefore invalid at the ray turning point. Note that the solution (4.6) is singular at $j_0 = \pi/2$ (horizontal takeoff angle). Rays that are near-horizontal should not be taken into account. This condition can be expressed, using Snell's law in the canonical model, as $\sin j_0 = (z_0+h)/R \ll 1$ or equivalently

$$z_0+h \ll R, \tag{4.7}$$

The (complex) wavefront curvature of the beam follows from the natural generalization of the ray concept

$$K_w^{\flat} = \alpha \frac{P}{Q} \equiv K_w + \frac{2i}{kL^2}, \tag{4.8}$$

with

K_w = Re($K_w^{\flat}$) being the local wavefront curvature, and

L = [$2k^{-1}$ Im($K_w^{\flat}$)]$^{1/2}$ the beam half width.

Cerveny *et al* (1982), suggest the following beam parameter B, for a beam-receiver pair $[j_0,M(s,n)]$,

$$B_M(j_0) \equiv \left| \frac{Q_2(s)}{Q_1(s)} + d \right|. \tag{4.9}$$

In a homogeneous medium, this value yields a minimum value of the beam half-width $L(s)$

at the receiver $M(s,n)$, for ray j_0. For the computations $d = 0$ is used, which means that the distance at which the matching between the waves emanating from the source and the superposition of Gaussian beams is done at the source location. If $d \neq 0$ the matching is done on a sphere of radius $d \times unit\ length$, with center at the source.

If N receivers are present ($M_i, i=1,N$), we define three values of the beam parameter in addition to $B_{Mi}(j_0)$

$$B_m = \max_{(i=1,N),\{j_0\}} B_{Mi}(j_0), \tag{4.10}$$

$$B_{j_0} = \max_{(i=1,N)} B_{Mi}(j_0), \tag{4.11}$$

$$B_i = \max_{\{j_0\}} B_{Mi}(j_0), \tag{4.12}$$

where $\{j_0\}$ is the set of rays satisfying (4.18) for the receiver M_i.

The first value for B is constant throughout the computations (constant for all rays and receivers), the second value is constant for each ray but changes from ray to ray and the third is constant on each receiver but varies from receiver to receiver. Some justification will be brought in section 7 on the choice of these values. Use of the initial conditions (4.5) imposes a large value of the parameter B (quasi-plane waves) in order for the solution to be a complex perturbation of the WKBJ method (Madariaga, 1984). The initial beam parameter must usually be much larger than that obtained using B_m (4.10).

Thus, the P wave vertical displacement is

$$\overline{W}^b(M,\omega) = \int_D \overline{U}^{j_0}(M,\omega) \cos j \, dj_0. \tag{4.13}$$

The integration is being done over all rays passing through some neighborhood D of M. This neighborhood is determined by conditions of validity of the beam (4.18) and (4.19).

If $\bar{\sigma}(\omega)$ is the spectrum of a point source time function $\sigma(t)$, the P wave vertical displacement in time $W^b(M,t)$ is computed following

$$W^b(M,t) = \frac{1}{\pi} \text{Re} \left\{ \int_0^\infty \bar{\sigma}(\omega)\ \bar{W}^b(M,\omega)\ e^{-i\omega t}\ d\omega \right\}. \tag{4.14}$$

We shall be using the wave-packet approach in computing (4.14). We shall therefore first evaluate the Fourier transform (4.14), then integrate with respect to takeoff angles, j_0 (4.13). The numerical integration of (4.13) is done by the rectangle formula

$$W^b(M,t) \approx \sum_k U_z^k(M,t)\ \Delta j_0\ , \tag{4.15}$$

where $U_z = U^{j_0} \cos j$.

The results with initial conditions (4.3) will be compared to that using (4.5). Sensitivity on the results due to ray sampling Δj_0 and summation limits will be investigated. For comparison purposes, we shall numerically integrate (4.13) by the trapezoidal formula

$$W^b(M,t) \approx \sum_k \left[\ U_z^{k-1}(M,t) + U_z^k(M,t)\ \right] \frac{\Delta j_0}{2}\ , \tag{4.16}$$

and by Simpson's rule

$$W^b(M,t) \approx \sum_k [\ U_z^{k-2}(M,t) + 4\, U_z^{k-1}(M,t) + U_z^k(M,t)\ ]\ S(k)\ \frac{\Delta j_0}{3}\ , \tag{4.17}$$

where $S(k) \equiv [1 + (-1)^k]/2$.

As in section 3, the radial component of displacement is obtained by replacing the term $\cos j$ by $\sin j$. The Gaussian beam conditions of validity described in chapter II are:

(1) Ray conditions used in section 3, since Gaussian beams are high-frequency asymptotic solutions of the elastodynamic wave equation and that ray serve as support for the beams.

(2) Regularity of the ray centered coordinate system

$$n\, K_R \ll 1 . \tag{4.18}$$

(3) The paraxial condition (rays considered not far from geometric ray)

$$n\, K_w \ll 1 . \tag{4.19}$$

(4) The Fresnel 2-condition (beams within the first Fresnel's zone radius)

$$\Lambda \left| K_w^{\flat} \right|^{-1} = \Lambda \left[K_w^2 + (\frac{2}{k\, L^2})^2 \right]^{-1/2} \ll 2\pi\, l_0^2 , \tag{4.20}$$

where Λ is the wavelength, l_0 is the characteristic length of the medium, $L(s)$ the half-width of the beam, $K_w(s)$ the wavefront curvature, and n the ray-receiver distance.

Since no analytic, closed form, expression of the GB solution exist at the present time, we can only check these conditions numerically. The validity of GB at the ray turning point will also be investigated.

5. ACCURACY CRITERIA

We will compare synthetic seismograms computed using asymptotic wave theory (equations (3.11) and (4.14)) with the exact synthetics from the canonical problem (equation (2.24)). Four numerical parameters are defined to test the differences

between the two traces.

Let $\tilde{W}(M,t)$ be the compressional vertical displacement computed from asymptotic wave theory ($\tilde{W} = W^{\tau}$ *or* W^{b}). A tilde denotes the value computed using the approximate method. For a given receiver M define E as

$$E(M) = 100 \frac{\int [W(M,t) - \tilde{W}(M,t)]^2 \, dt}{\int W^2(M,t) \, dt} . \tag{5.1}$$

E is the power ratio (in percent) of the signals' difference to the exact signal. It is an L_2 norm of good-fit. The value E is independent of the source signal considered, and measures a relative global error. The time interval in the integration includes the whole waveform. Practically, we will put an upper limit on E, say E^{Max}, above which we will consider the results of the approximate method to be invalid.

Let

$$e_{\tau}(M) = 100 \left[\frac{\tau(M) - \tilde{\tau}(M)}{\tau(M)} \right], \tag{5.2}$$

$$e_{A}(M) = 100 \left[\frac{A(M) - \tilde{A}(M)}{A(M)} \right], \tag{5.3}$$

$$e_{ph}(M) = 100 \left[\frac{T(M) - \tilde{T}(M)}{T(M)} \right], \tag{5.4}$$

where τ, A, and T are respectively the wave travel time, the signal maximum amplitude and the time of the maximum signal amplitude at the receiver M. These three good-fit parameters are local time error measurements and are complementary. They are useful when the error E becomes large and additional information is required to determine locally the lack of accuracy. e_{τ} measures the relative error in travel time. e_A is sensitive to amplitude errors whereas e_{ph} represents, in some sense, the relative error

in phase of the signal.

6. MODEL PARAMETERS AND NUMERICAL DETAILS

The source function used is the Gabor wavelet

$$\sigma(t) = e^{-[\omega_g (t-t_i)/\Gamma]^2} \cos[\omega_g (t-t_i) + \psi], \tag{6.1}$$

where ω_g is the center angular frequency of the signal ($\omega_g = 2\pi f_g$), t_i is the initial time shift realizing the causality of the signal, Γ controls the half-width of the signal's envelope in time, and ψ is the initial phase.

The Fourier transform in equation (4.14) is computed using an IMSL discrete Fourier transform routine. Equation (2.24) is computed using a standard convolution in time program.

The reference model (Model 0) is the following:

α = 0

h = 10 km

z_0 = 0 km

$\alpha(0)$ = 3 km/sec

ρ_0 = 3 g/cm^3

f_g = 10 Hz

Γ = 5

ψ = 0 $radians$

t_i = 0.2 sec

The reference model assumes the density to be constant throughout the model

(α =0). The models that will be considered in this study are :

FAR FIELD BREAKDOWN for DYNAMIC RAY TRACING

Model 1 Model 0 with receivers at the surface z =0, and
r_1=0.1 km to r_4=0.4 km (Δr=0.1 km)
r_4=0.4 km to r_{15}=2.6 km (Δr=0.2 km)

HIGH FREQUENCY BREAKDOWN for DYNAMIC RAY TRACING

Model 2 Model 0 with one receiver at (r =100 km,z =0)
with $\alpha(0)$=10−170 km/sec ($\Delta\alpha$=40 km/sec)
with $\alpha(0)$=170−1450 km/sec ($\Delta\alpha$=160 km/sec)

PARAXIAL / FAR FIELD BREAKDOWN for GAUSSIAN BEAMS

Model 3 Model 0 with receivers at the surface z =0, and
r_1=0.4 km to r_9=3.6 km (Δr=0.4 km)

FRESNEL / HIGH FREQUENCY BREAKDOWN for GAUSSIAN BEAMS

Model 4 Model 0 with one receiver at (r =100 km,z =0), $\alpha(0)$=50 km/sec
with f_g=10, 40, 100 Hz

VERTICAL SEISMIC PROFILING SIMULATION with GAUSSIAN BEAMS

Model 5 Model 0 with a frequency of 40 Hz , and t_i=0.05 sec,
receivers at r=2 km, and z_1=0.1 km to z_{15}=1.5 km (Δz =0.1 km)

TURNING POINT INVESTIGATION

Model 6 Model 5 with receivers at z_1=0.15 km to z_{10}=0.24 km (Δz =0.01 km)

EARTHQUAKE SIMULATION with GAUSSIAN BEAMS

Model 7 Model 0 with a frequency of 5 Hz , and t_i=0.4 sec
source at z_0=7 km and receivers at z =0,
r_1=10 km to r_{11}=20 km (spacing 1 km)

The ray tracing program RAY81, written by I. Psencik (1983), was modified to compute Gaussian beam synthetics in two dimensional heterogeneous media. In the calculation we are not required to have boundaries close to the receiver line, since the medium properties are known even above z =0. This precaution is taken to avoid any additional approximation. The interpolation of the ray contribution to the receiver, not inherent in the method, would possibly deteriorate results. The approximate analytic Fourier transform of the Gabor wavelet (Cerveny, 1983[b]) is not used for similar reasons.

7. RESULTS AND DISCUSSION

Each asymptotic condition will be written in the form of a breakdown parameter that must be much greater than one. The upper limit on the global error E^{Max} is chosen to be at about 20%. This subjective value has been chosen by estimating in numerous synthetics the error above which the approximate results become unsatisfactory. Below 5%, the results are considered very accurate. The frequency used in the breakdown parameters is $f = f_g$ or $\omega = \omega_g$ this is because the spectra of the signal is well centered around this frequency. It has been verified that the global error E depends slowly on the source time function considered. A Kelly source (Kelly *et al*, 1976) was used and yielded same error trend and about the same errors (within 10%) as with the Gabor wavelet. The seismograms plotted are the vertical component of displacement in the (r,z) coordinate system. The radial component will be introduced when necessary.

Dynamic Ray Tracing

We shall first study the range of validity of DRT. The first condition to be tested is the far field condition (3.12) $k\ A_0 \gg | dA_0 / ds |$, which is expressed as

$$FFC \equiv \frac{R_w}{2\ z_c}\ \omega\ \omega_0^{-1} \gg 1 . \qquad (7.1)$$

Figure 3 shows the synthetics calculated with the exact analytical solution (left) and the synthetics computed with the DRT solution (right) from Model 1 parameters. The error analysis is presented in Table 1. For $FFC \geq 8$, the DRT results are reliable. Below this value, the near field term not taken into account in the DRT becomes predominant.

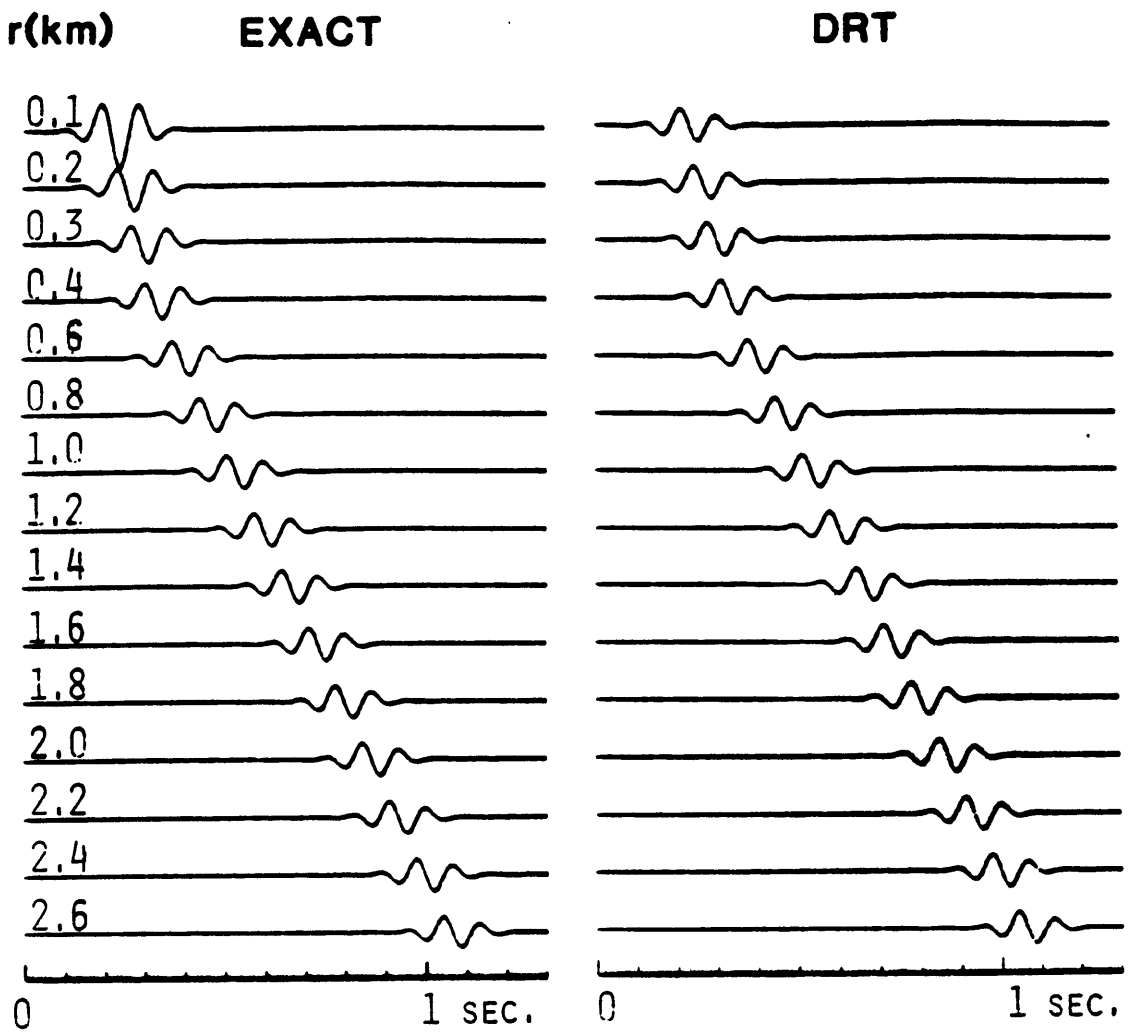


Figure 3. (left) exact synthetics of Model 1 (far field breakdown), and (right) DRT synthetics of Model 1.

TABLE 1 - FAR FIELD BREAKDOWN Model 1 - DYNAMIC RAY TRACING - HFC = 419						
$r(km)$	τ(sec)	e_τ(%)	e_{ph}(%)	e_A(%)	E(%)	FFC
0.1	0.03	0	-6	57	78	2
0.2	0.07	0	-2	33	47	4
0.3	0.10	0	-2	22	28	6
0.4	0.13	0	-2	15	18	8
0.6	0.20	0	0	7	9	13
0.8	0.27	0	0	5	5	17
1.0	0.33	0	0	3	3	21
1.2	0.40	0	0	3	2	25
1.4	0.47	0	0	2	2	29
1.6	0.53	0	0	1	1	33
1.8	0.60	0	0	1	1	37
2.0	0.67	0	0	0	1	41
2.2	0.73	0	0	0	1	45
2.4	0.80	0	0	1	1	49
2.6	0.86	0	0	1	0	53

TABLE 2 - HIGH FREQUENCY BREAKDOWN Model 2 - DYNAMIC RAY TRACING - (r,z) = (100 km,0)						
$\alpha_0(km/sec)$	τ(sec)	e_τ(%)	e_{ph}(%)	e_A(%)	E(%)	HFC
10	4.625	0	0	1	0	126
50	0.925	0	4	3	1	25
90	0.514	0	5	6	5	14
130	0.356	0	-1	9	9	10
170	0.272	0	7	12	14	7
330	0.140	0	-3	28	39	4
490	0.094	0	-3	41	58	3
650	0.071	0	11	51	71	2
810	0.057	0	11	58	79	1.6
970	0.048	0	12	64	85	1.3
1130	0.041	0	10	69	88	1.1
1290	0.036	0	10	72	91	1.0
1450	0.032	0	11	75	92	0.9

Table 1. Far field breakdown of Model 1 with DRT; Table 2. High frequency breakdown of Model 1 with DRT.

The global error is due first, to an error in amplitude (e_A), and second to an error in phase (e_{ph}). Note that the travel time calculation is accurate everywhere. The breakdown is seen clearly in the synthetics and is represented quantitatively in Table 1.

The following condition is the high frequency condition

$$HFC \equiv \omega\, \omega_c^{-1} \gg 1 . \tag{7.2}$$

Model 2 synthetics are displayed in figure 4 for three values of HFC(7,3 and 0.9), and the error analysis in Table 2. The minimum value of the key parameter HFC is about 7. For $HFC \leq 7$, the high frequency breakdown occurs in a similar fashion as the far field breakdown, as expected and discussed in section 3. We have chosen a receiver position such that $\xi \approx 4.6$, thus $z_c / R_w \approx 1$. Then (7.1) yields $FFC \approx 0.5\, \omega\omega_0^{-1}$. But expressing here (7.2) in terms of $\omega_0 = \omega_c$, yields $HFC \approx 2\, FFC$. For an HFC of about FFC, the errors are of same order as the far field breakdown errors. The velocity gradient values are quite unrealistic but this is because we have chosen to keep the frequency constant and vary the cut-off frequency ω_c. Keeping the velocity gradient in a realistic range and varying the frequency would yield identical results, the important parameter here being HFC. Therefore, we can conclude that the high frequency breakdown in this type of medium is similar to the far field breakdown. The phase distorsion effect due to the dispersive characteristics of the medium is not predominant here.

The Fresnel 1-condition will be defined as

$$FRN1 = HFC^2 (\omega \tau)^{-1} \gg 1 .$$

We have chosen a model 0 with $h=1km$, $\alpha(0)=5km/sec$, and receivers up to $4000km$. The minimum $FRN1$ calculated was about 3.5. This condition is difficult to test independently of the others, and because of numerical limits, we have not been able presently to get below $FRN1 \approx 3.5$. The error at this point was still negligible. In Table 2, values of $FRN1$ are shown. The breakdown is due to the high frequency condition violation. Therefore the Fresnel 1-condition is not independently tested in Model 2, and cannot serve as a criteria. The Fresnel 2-condition (see chapter I) is not a necessary condition. In the limit of numerical calculations this condition was violated by 3 orders of magnitude without introducing noticeable error.

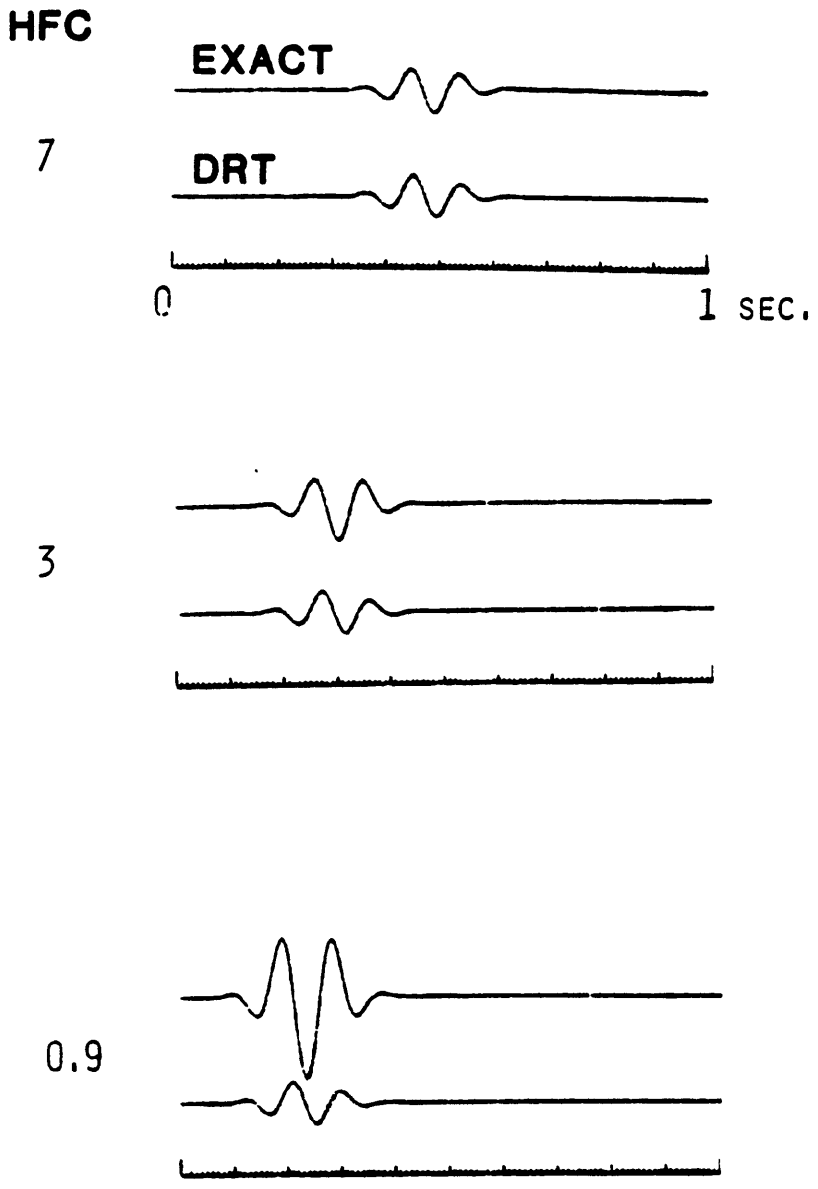


Figure 4. Three pairs of synthetics of Model 2 (high frequency breakdown). In a pair: top trace is exact and bottom trace is from DRT.

Gaussian Beams

In the beam error analysis it is important to define another global error E^W. The definition is similar to (5.1) but the integration is done on a time window containing the wavelet alone. The reason being that with the superposition of beams, error due to truncation of (4.13) or of the sampling interval will increase E without necessarily increasing E^W. Further, it is possible to determine how much of the global error is due this "noise". The local error e_τ on the travel time cannot be computed since we superpose different signals that contribute to the final signal without direct calculation of the GB travel time. We shall impose the ray-centered coordinate system regularity condition (4.18)

$$RCC \equiv (n\ K_R)^{-1} \gg 1\ , \tag{7.3}$$

and the paraxial condition (4.19)

$$PRX \equiv (n\ K_w)^{-1} \gg 1\ . \tag{7.4}$$

The lower limit of RCC (not attained) has be chosen to be 5, and that of PRX, 4. These two values numerically constrain the conditions appropriately. The beam parameter d = Re(ε) will be set equal to 0 throughout the computations.

The first group of tests are on beam parameter definitions, the numerical integration of (4.10), and the initial conditions of the complex dynamic ray tracing system. These tests will be done in a far field breakdown context. Models 3 and 4 will be implemented here. The beam parameters that are considered for presentation are given by eqns. (4.9-11). Other beam parameters have been tested (constant value, average value, minimum value, etc. instead of maximum value of eqns. (4.10-11)) but were unreliable. Either they were not automatic (for example constant) and therefore

needed to be redefined if the medium or receivers were modified or else yielded to high an error. The exact seismograms of Model 3 are shown in figure 5a (top left). The ray diagram used in the beam computations is displayed in figure 5a (bottom). Initial conditions (4.3) are used. The fan of equidistant ray takeoff angles is $[-45^{\circ},50^{\circ}]$, oriented top right to bottom left. 20 rays are sufficient to yield satisfactory results with the beam parameter B_i (4.12). Figure 5a (top right) shows the synthetics of Model 3 for the B_i parameter and Table 3 the error analysis. Other beam width parameters are investigated in Appendix D. Results show that the most stable parameter is the one used here. The global error is due to the far field breakdown . A detailed examination of the breakdown between r=0.4 km and r=0.8 km showed that the critical (lower bound) FFC is at about 13. Thus, the far field condition requires $FFC \geq 13$. Details on superposition of beams, ray density, summation limits and results with initial conditions (4.3) are presented in Appendix E.

The last test is on the Fresnel condition (4.20)

$$FRN2 \equiv 2\pi\, l_0^2\, f \left[\left(\frac{K_w}{\alpha}\right)^2 + \left(\frac{1}{\pi f\, L^2}\right)^2 \right]^{1/2} \gg 1 , \qquad (7.5)$$

and the high frequency condition (7.2). Model 4 is implemented where four frequencies are being considered 10;40;100 and 200 Hz. The fan of rays is $[77^{\circ},81^{\circ}]$. 20 total rays are used with conditions (4.3). The error analysis is in Table 4 . Note that the minimum critical $FRN2$ is at about three. Therefore for $FRN2 \geq 3$, we are guaranteed that the Fresnel condition is satisfied.

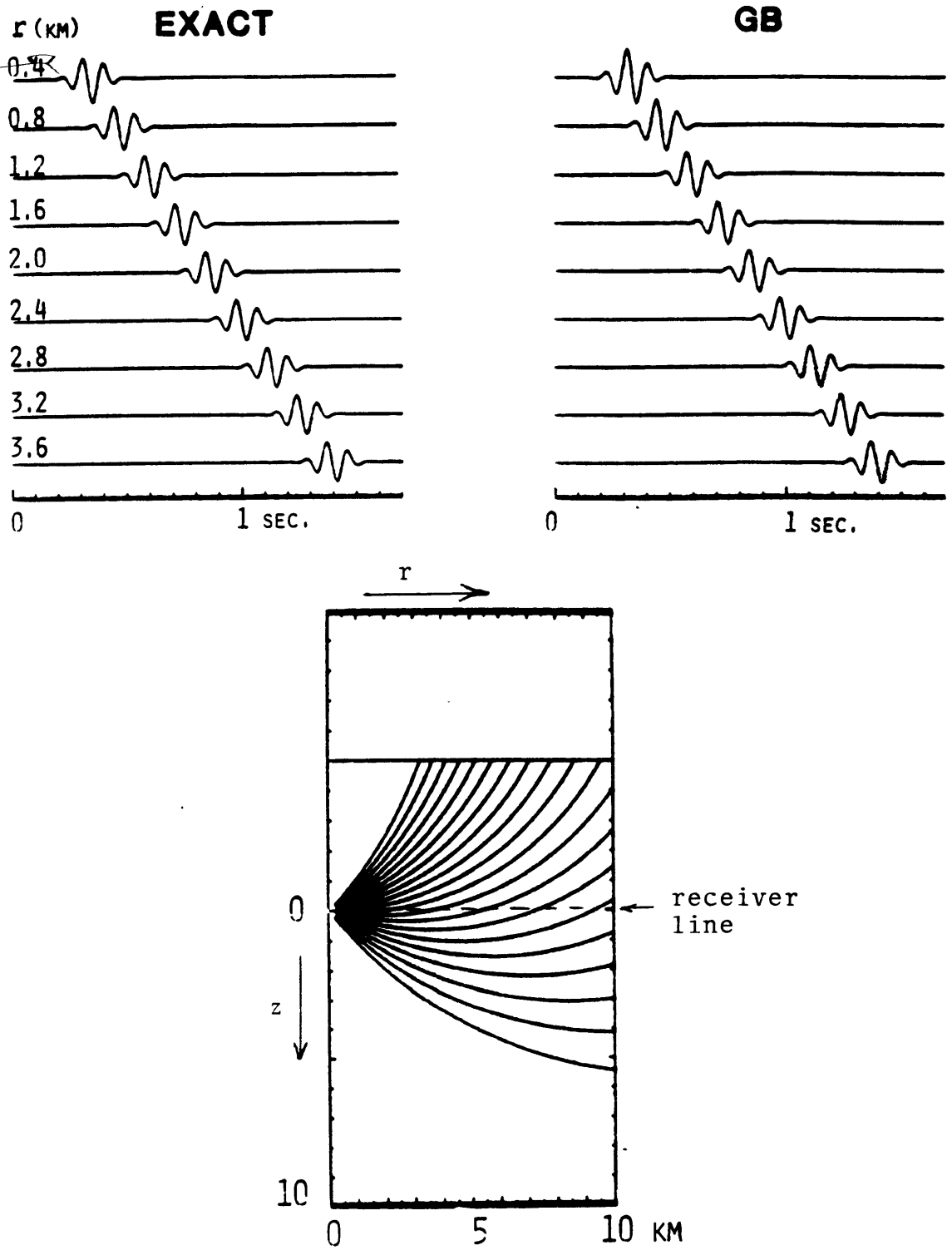


Figure 5a. (top left) exact synthetics of Model 3 (far field breakdown), (top right) GB synthetics of Model 3, and (bottom) ray diagram for GB of Model 3.

The minimum $FRN2$ in Model 3 is 402 and that of Model 3x, 226. In beam methods, the Fresnel condition might alter the results even if the high frequency condition is satisfied. For this reason the high frequency condition in the beam method is not sufficient. Note that we are still in the far field region since $FFC \approx 0.5 HFC$ for this receiver. We have seen from the DRT range of validity that the high frequency breakdown is similar to the far field breakdown. Since both methods are asymptotic, we shall assume this result to be true for GB too, as long as the Fresnel condition is satisfied. Thus the condition for the high frequency condition is $HFC \geq 13$.

The standard ray method (DRT) has the advantage of breaking down more smoothly than the Gaussian beam method. Further, it requires less conditions for its applicability. Figure 5b clearly shows this difference for the far field condition and the high frequency condition breakdowns. The Gaussian beam computations here require an extension of the actual medium necessary for the extrapolation (distance ray-receiver, Q and P corrections , etc.) of the wave field. Approximate techniques are possible when the medium cannot be extrapolated naturally (Cerveny, 1983[a]). The possible advantage of GB in the present case is mainly computational. If the DRT solution were to be computed by solving numerically the DRT system (as the GB solution), the computer time would be greater than with GB, due to iterations for finding the source-receiver ray.

TABLE 3 - FAR FIELD BREAKDOWN
Model 3 - GAUSSIAN BEAMS
B_i - 20 rays

$r(km)$	τ(sec)	e_{ph}(%)	e_A(%)	E(%)	E^W(%)	FFC
0.4	0.13	10	-15	39	39	8
0.8	0.27	10	-16	9	9	17
1.2	0.40	0	-6	1	1	25
1.6	0.53	0	-4	2	2	33
2.0	0.67	0	-4	1	1	41
2.4	0.80	0	-1	1	1	49
2.8	0.93	0	-2	1	1	57
3.2	1.06	0	-2	1	1	65
3.6	1.19	0	-2	1	1	72

TABLE 4 - FRESNEL / HIGH FREQUENCY BREAKDOWN
Model 4 - GAUSSIAN BEAMS
(r,z)=(100 km,0) - B_i - 20 rays

f Hz	e_{ph}(%)	e_A(%)	E(%)	E^W(%)	HFC	$FRN2$
10	3	53	42	42	25	0.5
40	1	20	16	16	101	2
100	0	3	6	6	251	5
200	0	6	2	2	503	11

Table 3. Far field breakdown of Model 3 with GB; Table 4. Fresnel / High frequency breakdown of Model 4 with GB.

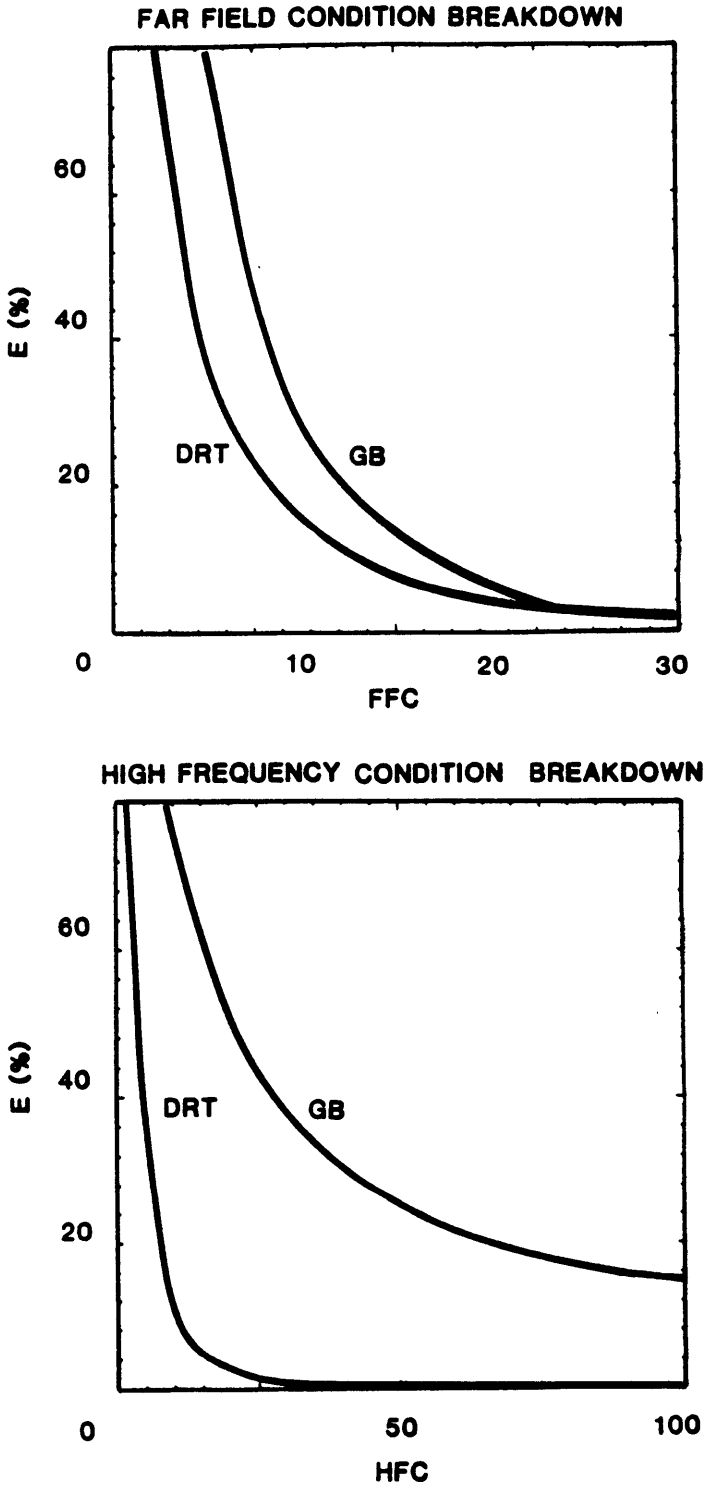


Figure 5b. (top) far field condition breakdown for DRT and GB, and (bottom) high frequency condition breakdown for DRT and GB.

Two examples

We shall simulate a typical vertical seismic profile with an offset source at the surface. The receivers are positioned in a vertical borehole. The Gaussian beam method is used with 20 rays and a fan $[-10^o,50^o]$. Model 5 represents this case, Table 5 shows the error analysis , figure 6 (top left) the exact analytical synthetics and figure 6 (top right) the Gaussian beam synthetics. Since for the choice of beam (4.14) there is no difference between E and E^W, we shall omit the latter in the following tables. The ray diagram is shown in figure 6 (bottom). The receiver array crosses a ray turning point location (around receiver 2) where the vertical component vanishes in the approximate methods. DRT error is about 90 % at this point, whereas GB is about 307 %. The results are reliable and very accurate outside this zone. The synthetics plotted are not scaled in depth. The scale is conserved between a given approximate trace and the corresponding exact trace.

We wish to determine precisely if the ray turning point is a singular region for DRT or GB. It is important at this point to consider the radial component of displacement, since the vertical component vanishes for the approximate methods. If the radial displacement is accurate and much greater than the exact vertical displacement at the ray turning point, we can conclude that the ray method under consideration is accurate around and at the ray turning point. Model 6 is implemented. For GB, the same rays as in Model 5 are used. Errors for DRT are shown in Table 6, and those of GB in Table 7. The maximum vertical amplitude, in a normalized scale, of the exact solution is shown under Ex.Max.Vert. The maximum vertical amplitude of the approximate solution can be derived from that of the exact solution via (5.3). They are roughly of same order.

TABLE 5 - VERTICAL SEISMIC PROFILING Model 5 - GAUSSIAN BEAMS B_i - 20 rays				
$z(km)$	τ(msec)	e_{ph}(%)	e_A(%)	E(%)
0.1	663	2	-5	1
0.2	662	0	27	307
0.3	663	-2	6	1
0.4	666	-2	3	0
0.5	669	0	1	0
0.6	675	-2	2	0
0.7	682	0	1	0
0.8	690	-1	1	0
0.9	699	0	1	0
1.0	709	0	0	0
1.1	721	-1	1	0
1.2	733	0	1	0
1.3	746	0	-1	0
1.4	761	0	-2	0
1.5	775	0	-3	0

TABLE 6 - TURNING POINT in VSP Model 6 - DYNAMIC RAY TRACING						
$z(km)$	e_{ph}(%)	e_A(%)	E(%)	Rad E(%)	Ex.Max.Vert	Max.Rad
0.15	0	3	1	0	3	114
0.16	0	4	2	0	2	114
0.17	0	5	4	0	2	114
0.18	0	10	9	0	1	114
0.19	1	25	34	0	0.6	114
0.20	1	71	90	0	0.4	113
0.21	-1	16	20	0	0.8	113
0.22	-1	8	7	0	1	113
0.23	-1	5	3	0	2	113
0.24	-1	4	2	0	2	113

Table 5. Vertical seismic profiling of Model 5 with GB; Table 6. Turning point in Model 6 with DRT.

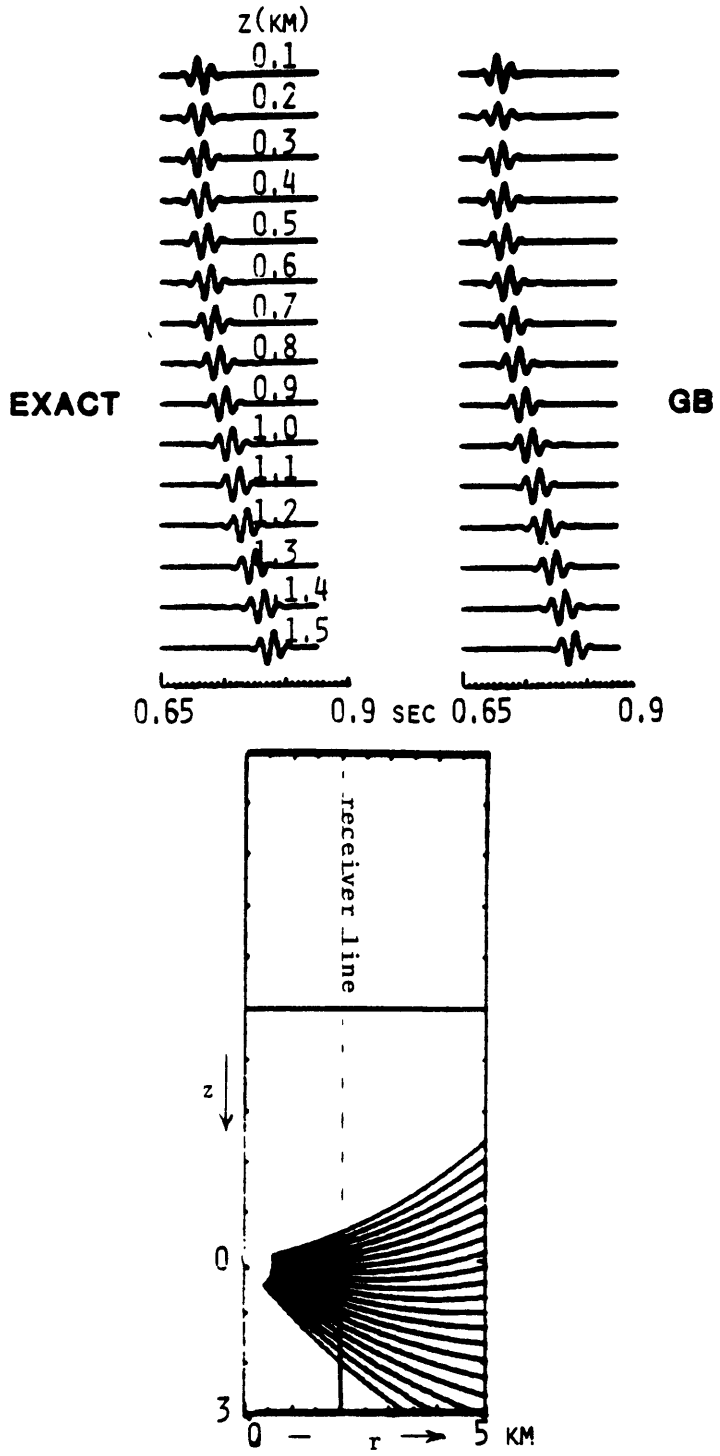


Figure 6. (top left) exact synthetics of Model 5 (vertical seismic profiling), (top right) GB synthetics of Model 5, and (bottom) ray diagram for GB of Model 5.

Maximum radial amplitude of approximate solutions, normalized to the same scale, is under the column Max.Rad. The error on the radial component is represented by Rad E. The results are close to perfect for the radial component of displacement for both, DRT and GB. Further, the maximum amplitude in the radial direction is at least fifty times the maximum amplitude in the vertical direction, around the ray turning point. Therefore DRT and GB are accurate at ray turning points since the magnitude and direction of the displacement are the same, within our criteria, as those given by the exact solution.

The last example simulates an earthquake, where the source is at a depth of 7 km. Model 7 represent this configuration. 20 rays are used in the Gaussian beam method, with a fan of $[-90^{o},20^{o}]$. Figure 7 (top) displays the Gaussian beam synthetics, figure 7 (bottom) the ray diagram, and Table 8 the error analysis. Results are close to perfect everywhere.

As a summary, we have the following conditions that have to be met in order for the ray and beam methods to be reliable and accurate (i.e. $E < 20\%$):

Dynamic Ray Tracing:

$$FFC \equiv kA_0 \left| \frac{dA_0}{ds} \right|^{-1} > 8 ,$$

$$HFC \equiv \omega \omega_c^{-1} \geq 7 .$$

Gaussian Beams with conditions (4.3) and B_i:

$$FFC \geq 13 ,$$

$$HFC \geq 13 ,$$

$$RCC \equiv (n\ K_R)^{-1} > 5,$$

$$PRX \equiv (n\ K_w)^{-1} > 4,$$

$$FRN2 \equiv 2\pi l_0^2\ \Lambda^{-1}\ |K_w^b| \geq 3.$$

The decoupling condition in the canonical model ($\omega\omega_0^{-1} \gg 1$) is automatically satisfied with the high frequency condition, since the two characteristic frequencies are related. In the case of structures which are more complex than a medium with constant velocity gradient, the above conditions remain valid and applicable as long as ω_0 and ω_c exist and are real. The Fresnel 1-condition must further be investigated; so far for $FRN1 \equiv HFC^2(\omega\ \tau)^{-1} \geq 3.5$, we have not detected any error. The sensitivity of GB to the parameter d is presently studied. More complicated types of media where an exact solution can be found are being investigated. Exact solutions (Cagniard's problem) in media with an interface should be compared quantitatively to check the accuracy of head waves obtained using GB with large B (Nowack and Aki, 1984). A more representative source function might be considered such as that used by Madariaga and Papadimitriou (1984), close to being a numerical delta function.

TABLE 7 - TURNING POINT in VSP
Model 6 - GAUSSIAN BEAMS
B_i - 20 rays

$z(km)$	$e_{ph}(\%)$	$e_A(\%)$	$E(\%)$	Rad $E(\%)$	Ex.Max.Vert	Max.Rad
0.15	1	-8	4	0	3	114
0.16	1	-10	7	0	2	114
0.17	1	-12	13	0	2	114
0.18	1	-27	42	0	1	114
0.19	1	-27	121	0	0.6	114
0.20	0	26	307	0	0.4	113
0.21	-1	46	61	0	0.8	113
0.22	-1	28	21	0	1	113
0.23	-1	19	10	0	2	113
0.24	-1	15	6	0	2	113

TABLE 8 - EARTHQUAKE SIMULATION
Model 7 - GAUSSIAN BEAMS
B_i - 20 rays

$r(km)$	τ(sec)	$e_{ph}(\%)$	$e_A(\%)$	$E(\%)$	$E^{\Psi}(\%)$
10	3.02	0	1	0	0
11	3.21	2	-2	0	0
12	3.40	2	-1	0	0
13	3.60	2	2	0	0
14	3.79	0	1	0	0
15	3.99	0	-2	0	0
16	4.18	0	-1	0	0
17	4.38	0	0	0	0
18	4.57	0	-2	1	1
19	4.76	0	0	0	0
20	4.95	2	-1	0	0

Table 7. Turning point in Model 6 with GB; Table 8. Earthquake simulation in Model 7 with GB.

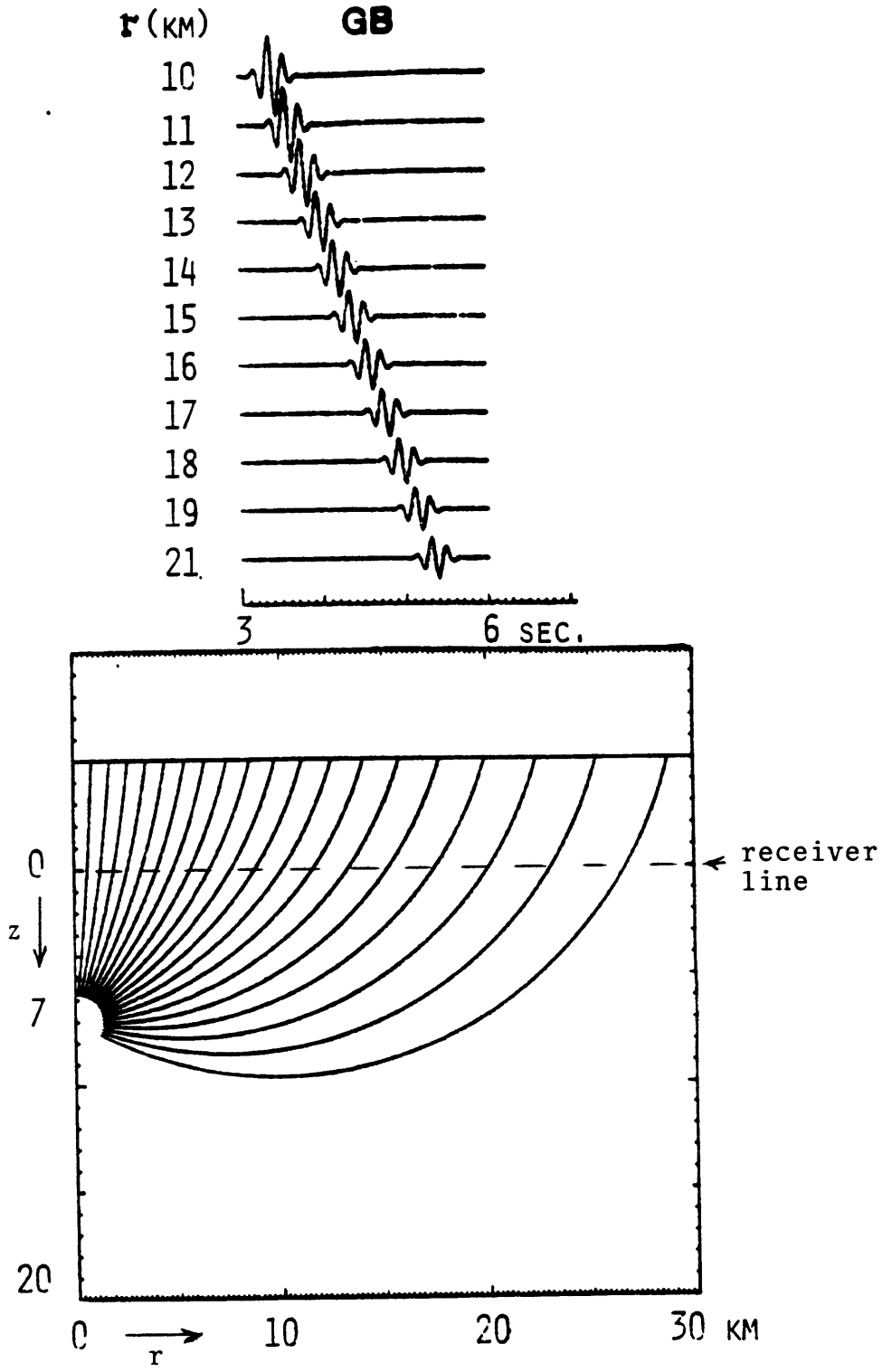


Figure 7. (top) GB synthetics of Model 7 (earthquake simulation), and (bottom) ray diagram for GB of Model 7.

8. CONCLUSION

Range of validity of two asymptotic wave methods, the dynamic ray tracing method and the Gaussian beam method, have been established. The medium considered is that of a constant gradient model where the Green's function is known analytically in time and frequency domains.

Two conditions must be met in the dynamic ray tracing method: (1) far field, and (2) high frequency conditions. The Gaussian beam method requires three additional conditions: (3) regularity of the ray centered coordinate system, (4) paraxial rays, and (5) rays within the first Fresnel radius conditions. Quantitative values are shown for each condition, and examples exhibit clear breakdowns. The two methods are accurate at ray turning points.

Two examples show the application of the Gaussian beam method to different types of seismic problems. For the canonical medium considered, the Gaussian beam method is less accurate than dynamic ray tracing, particularly when a breakdown occurs. The validity conditions are applicable for complex structures as long as the two medium's characteristic frequencies ω_0 and ω_c exist and are real.

9. APPENDIX D: Details on beam width parameters

Other beam width parameters are investigated with Model 3 (far field breakdown). Error analysis with the beam parameter B_m (4.10) is shown in Table D1. The calculated constant value equals B_m = 4. Results with parameters B_M (4.9) and B_{j_0} (4.11) are shown in Table D2 and D3, respectively. Although all four beam parameters (4.9-11) yield about the same results at $r \geq 1.2\ km$, the most smooth breakdown (compared to DRT) is achieved with the beam parameter B_i (4.12). Our basic assumption being that the beam method should not behave that differently from DRT in regions where DRT is regular. A possible explanation is that, for this choice, the beams are adapted to the receiver location keeping the maximum allowable (constant) value of the beam parameter on each receiver, and are renormalized from receiver to receiver. Note that for this choice we can have a very wide spread of receivers. The beam parameter (4.12) B_i is the beam of choice.

TABLE D1 - FAR FIELD BREAKDOWN
Model 3 - GAUSSIAN BEAMS - B_m - 20 rays

$r(km)$	τ(sec)	e_{ph}(%)	e_A(%)	E(%)	E^W(%)	FFC
0.4	0.13	10	-352	988	988	8
0.8	0.27	10	3	58	57	17
1.2	0.40	0	0	2	2	25
1.6	0.53	0	-4	1	1	33
2.0	0.67	0	-3	0	0	41
2.4	0.80	0	-2	0	0	49
2.8	0.93	0	-2	1	1	57
3.2	1.06	0	-2	1	1	65
3.6	1.19	0	-2	1	1	72

TABLE D2 - FAR FIELD BREAKDOWN
Model 3 - GAUSSIAN BEAMS - B_M - 20 rays

$r(km)$	τ(sec)	e_{ph}(%)	e_A(%)	E(%)	E^W(%)	FFC
0.4	0.13	10	-8	42	42	8
0.8	0.27	9	14	30	30	17
1.2	0.40	6	-4	10	10	25
1.6	0.53	5	-13	12	12	33
2.0	0.67	4	-14	6	6	41
2.4	0.80	4	-11	3	3	49
2.8	0.93	3	-9	2	2	57
3.2	1.06	3	-7	2	2	65
3.6	1.19	0	-10	2	2	72

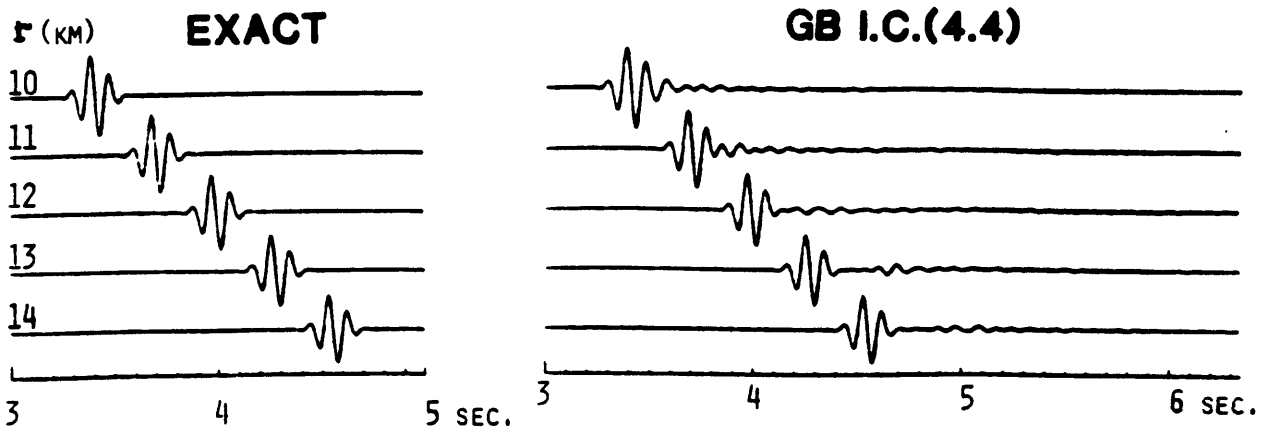
TABLE D3 - FAR FIELD BREAKDOWN
Model 3 - GAUSSIAN BEAMS - B_{j_0} - 20 rays

$r(km)$	τ(sec)	e_{ph}(%)	e_A(%)	E(%)	E^W(%)	FFC
0.4	0.13	08	-342	1153	1153	8
0.8	0.27	10	-34	53	52	17
1.2	0.40	0	-14	3	3	25
1.6	0.53	0	-14	3	3	33
2.0	0.67	5	-11	2	2	41
2.4	0.80	4	-9	1	1	49
2.8	0.93	3	-8	2	2	57
3.2	1.06	0	-7	2	2	65
3.6	1.19	0	-10	2	2	72

Table D1, Table D2 and Table D3: Testing of other beam parameters in GB.

10. APPENDIX E: Beam integration and initial conditions analysis

The method of integrating (4.13) yields different results depending on the method. The following Model 3x is considered: Model 0 with receivers at the surface z=0, and r_1=10 km to r_5=14 km (spacing 1 km). This model has no breakdowns and yield very good results with 20 rays and a fan of [1°,65°], where the exact synthetics are shown in figure 8 (top left), and the ray diagram in figure 8 (bottom). The error analysis is shown in Table E1. The synthetics are very accurate and are identical to those in figure 8 (top left) (E=0). In order to test the method of integration we decided to shoot only eight rays with the same fan. There is no distinguishable differences between the rectangular, (4.15) Table E2, and trapezoidal rule, (4.16) yielding same errors as the rectangular rule. However, results become less accurate with Simpson's rule, (4.17) Table E3. As a consequence, there is no need presently to use more advanced methods of integration (other than the rectangular rule 4.15) to get reliable and accurate results. The fan of rays is required to cover conditions (7.3) and (7.4) so that they become active. This can be easily implemented by checking how many of the rays contribute to the final seismogram. At each receiver, the number of rays contributing to the seismogram should be less than the total number of rays. Increasing the number of rays (i.e. ray density) improves the accuracy. Of the 20 rays, about 15 contribute to an individual seismogram, yielding a very good accuracy. Even with 8 total rays, (about 6 rays per receiver) the results remain satisfactory (i.e. $E < E^{Max}$). The number of rays that will contribute exactly for each seismogram is not known a priori. That is why the total number of rays is not required to be even when Simpson's integration rule is used.



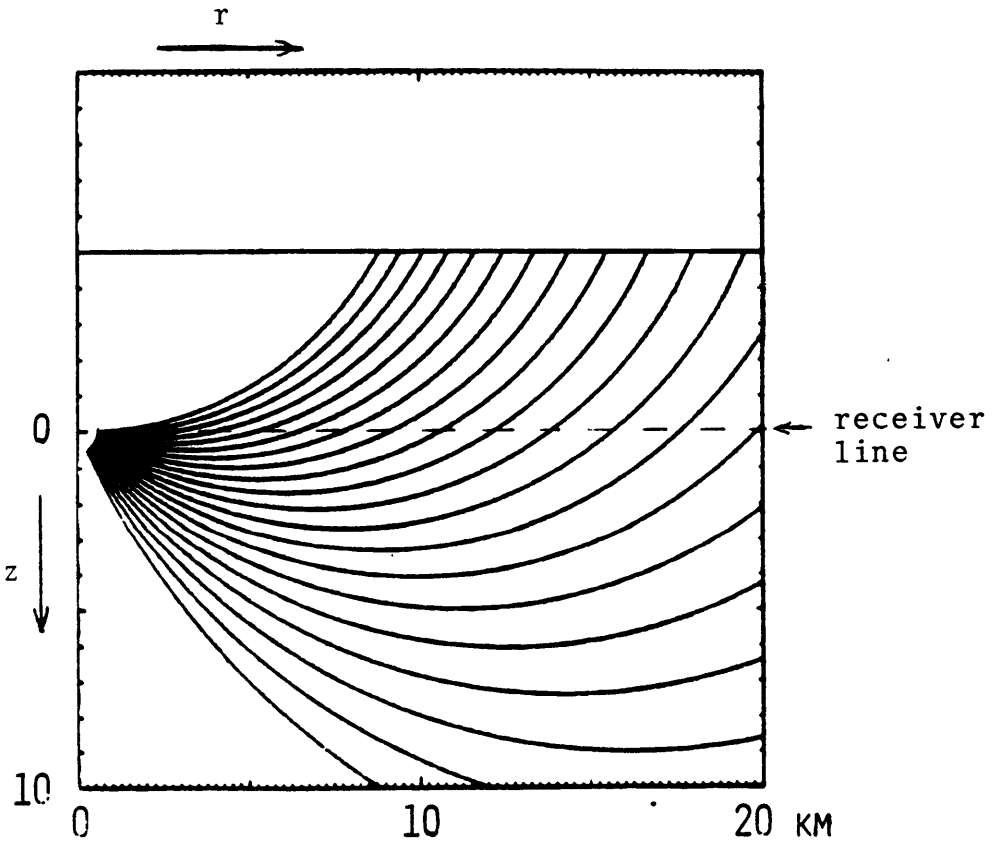


Figure 8. (top left) exact synthetics of Model 3x, (top right) GB synthetics of Model 3x with conditions (4.5), and (bottom) ray diagram for GB of Model 3x.

If the initial conditions (4.5) are used instead of (4.3), there are additional conditions for the applicability of GB. The Non horizontal incidence condition (4.7) is written in the form

$$NHI \equiv R\,(z_0+h)^{-1} \gg 1 .$$

Since results become singular as the takeoff angle reaches the horizontal (see 4.6), the corresponding rays distort the seismograms considerably. Receivers close to the source, but still within the far field limit, cannot be considered. Further, a higher ray density is required to obtain an accuracy similar to that with initial conditions (4.3). We shall use the constant value of the beam parameter of B = 44100, with 100 rays. These two initial conditions are compared for Model 3, with same fan of rays, and the error analysis displayed in Table E4. This table should be compared to Table 3. With Model 3x, the synthetics with conditions (4.5) are shown in figure 8 (top right) and the error analysis is displayed in Table E5 and are to be compared to that of Table E1. The results with the new conditions are "noisier" due to the contributions of near-horizontal rays . This is seen in the synthetics and quantified by the differences between E and E^W in the tables. For the furthest receiver in Model 3 the parameter NHI is equal to 1.016, and in Model 3x, NHI = 1.2. Therefore for NHI greater than say 1.2, use of the initial conditions (4.5) in GB would yield reliable results provided the ray density is high enough.

TABLE E1
Model 3x - GAUSSIAN BEAMS - B_i - 20 rays

$r(km)$	τ(sec)	e_{ph}(%)	e_A(%)	E(%)	E^W(%)	FFC
10	3.21	0	0	0	0	156
11	3.50	0	0	0	0	164
12	3.79	0	0	0	0	170
13	4.07	0	3	0	0	176
14	4.35	0	3	0	0	181

TABLE E2
Model 3x - GAUSSIAN BEAMS - B_i - 8 rays

$r(km)$	τ(sec)	e_{ph}(%)	e_A(%)	E(%)	E^W(%)	FFC
10	3.21	0	-10	1	1	156
11	3.50	-1	10	6	6	164
12	3.79	0	8	3	3	170
13	4.07	1	-23	10	10	176
14	4.35	0	-14	13	13	181

Table E1. Model 3x with GB and B_i with 20 rays; Table E2. Model 3x with GB and B_i with 8 rays.

TABLE E3 - Model 3x - GAUSSIAN BEAMS B_i - 8 rays - Simpson's rule						
$r(km)$	τ(sec)	e_{ph}(%)	e_A(%)	E(%)	E^W(%)	FFC
10	3.21	0	-22	8	8	156
11	3.50	-1	-20	16	16	164
12	3.79	1	-24	12	12	170
13	4.07	1	-34	14	14	176
14	4.35	0	-6	25	25	181

TABLE E4 - FAR FIELD BREAKDOWN **Model 3 - GAUSSIAN BEAMS** B=44100 - 100 rays - In.Cond. (4.5)						
$r(km)$	τ(sec)	e_{ph}(%)	e_A(%)	E(%)	E^W(%)	FFC
0.4	0.13	13	-1274	****	****	8
0.8	0.27	3	-257	1029	1028	17
1.2	0.40	1	-87	352	345	25
1.6	0.53	2	-56	245	230	33
2.0	0.67	1	-38	117	87	41
2.4	0.80	1	-28	92	54	49
2.8	0.93	1	-30	79	48	57
3.2	1.06	1	-31	79	51	65
3.6	1.19	-2	-29	71	53	72

TABLE E5 - Model 3x - GAUSSIAN BEAMS B=44100 - 100 rays - In.Cond. (4.5)						
$r(km)$	τ(sec)	e_{ph}(%)	e_A(%)	E(%)	E^W(%)	FFC
10	3.21	0	3	4	3	156
11	3.50	0	3	3	2	164
12	3.79	0	2	2	0	170
13	4.07	0	1	3	0	176
14	4.35	0	2	2	0	181

Table E3. Model 3x with Simpson's integration rule in GB; Table E4. Far field breakdown of Model 3 with GB and initial conditions (4.5); Table E5. Model 3x with GB and initial conditions (4.5) with large initial beam width parameter.

IV. MODELING WITH THE PARAXIAL RAY METHOD

There is no permanence to scientific concepts because they are only our interpretations of natural phenomena.

- J. Bronowski

1. INTRODUCTION

Central rays introduced in the preceding chapters individually sample the medium. All extrapolations and computations of the high frequency wavefield are performed in this local frame, that is, in the ray centered coordinate system. One may express these operations in a global reference frame in order to represent all the operations in the same coordinate system (Cerveny, 1983), but this procedure does not introduce additional information about the medium.

The displacement field calculated in standard ray method (two-point ray tracing) corresponds to a boundary value problem, where the central ray connects the source to the receiver (see II.4.12). Such a ray is called geometric. An iterative shooting scheme is implemented to solve this problem. Examples of this method are shown for surface seismics in Cerveny, Moloktov and Psencik (1977), and in vertical seismic profiling in Mellen (1984).

The Gaussian beam method generally requires more than eight beams in order to reconstruct the high frequency wavefield (see II.8.4). The beam support is a central ray and the beams superposed are required to sample the medium in the vicinity of the receiver. It is viewed as an initial value problem, where rays do not necessarily cross

receiver locations.

Original formulation of the parabolic-paraxial wave theory was done in 1946 by Leontovich and Fock. A brief historical survey of the method with numerous references and applications to many areas in physics can be found in Tappert (1977). A formal derivation and extension of the theory for scalar Helmholtz equation is presented in Fishman and McCoy (1984). Several authors considered the application of the theory to elastic waves (Landers and Claerbout, 1972; McCoy, 1977; Hudson,1980; Corones *et al.*, 1982). Wales and McCoy (1983), compared the results of some of these different approaches, in a weak-scattering context.

Within our context, the paraxial-parabolic ray method is an intermediate method situated between the standard ray method and the Gaussian beam method. It is described briefly in section II.7. It assumes that the central ray(s) is (are) within the neighborhood of the receiver. Considering a virtual ray passing through the receiver (geometric ray) with wavenumber vector $\vec{k}$, and for a central ray in the vicinity of the receiver with wavenumber vector $\vec{k}_s$, the central ray is paraxial if the angle between $\vec{k}$ and $\vec{k}_s$, χ, is sufficiently small so that $\sin\chi \approx \chi$ (i.e. $\chi \ll 1$). If this is satisfied then $|\vec{k}| \approx |\vec{k}_s|$.

The parabolic approximation assumes a privileged direction of propagation, which in our case is the ray direction (along s). This approximation neglects the $\partial^2/\partial s^2$ term in the Helmholtz wave equation. The dispersion relation of the Helmholtz wave equation in a homogeneous medium which is represented by a circle, is now approximated by a parabola (Claerbout, 1976). The vertex of the parabola is the unique point of contact with the circle, and is tangent to it. Its axis is parallel to $\vec{k}_s$. This approximation leads to $|\vec{k}| \approx |\vec{k}_s|$ (see Appendix II.C).

A ray that is paraxial satisfies therefore the parabolic approximation and vice-versa. The denominations paraxial and parabolic describe the same type of approximations but in different spaces; the former determines the receiver neighborhood in the spatial domain ($nK_w \ll 1$), and the latter imposes an upper bound of the wavenumber vector orthogonal to $\vec{k}_s$ ($|\vec{k}_n| \ll |\vec{k}|$) in the frequency-wavenumber space (see II.C.4). We shall, henceforth, choose the term *paraxial ray approximation* to describe the restriction of the energy flux to $\vec{k}_s$ directions that fall within a narrow angle of the $\vec{k}$ axis.

The paraxial ray method can be derived by extension of the standard ray method, or as a limiting case of Gaussian beams. The Taylor expansion of the travel time away from the central ray, followed by a scaling and rotating (similarity operation) of the ray vector amplitude yield the paraxial ray. In the Gaussian beam formulation, setting the complex parameter ε (III.4.2) equal to zero, the superposition integral (II.8.4 or III.4.10) reduces to an average of paraxial rays. This limit requires that the weight factor Φ be equal to $1/N$, where N is the total number of rays contributing in the integral.

The advantage of the paraxial ray method over the standard ray method is mainly computational. The method does not require two point ray tracing. The method is robust and flexible in the sense that receivers can be placed anywhere in the medium. Results can be as accurate as required, with computation time as tradeoff. One paraxial ray can be used to extrapolate the field at many receivers, reducing then the total number of rays. This number is, generally, less than the total number of receivers. The advantage over the Gaussian beam method is that only one ray is sufficient to extrapolate the high frequency wavefield to the observer, whereas Gaussian beams require a minimum number (on the order of eight), requiring then a greater medium coverage around the receiver. Even though Gaussian beams are restricted to the paraxial region and

therefore uses the paraxial approximation, the two concepts are of different nature. The disadvantage over Gaussian beams is that it is less robust and cannot handle caustics, shadow zones, head waves or other interface waves. We shall present in the following section the development of the paraxial ray method in a straightforward and systematic manner.

2. THEORY

The theoretical steps that are performed in the paraxial ray method are presented. We will mainly be concerned with two dimensional medium heterogeneities. Derivations will implicitly assume that medium properties do not change in one space direction, say in the y direction. Equations for three dimensional heterogeneous media can be found in Cerveny (1983) and Cerveny, Klimes and Psencik (1984).

In "forward" modeling, it is natural to assume that we are given a subsurface model. A global reference frame $(\hat{x}, \hat{z})$ is defined and shown in figure 1. The following inputs are specified: (0) The model domain of definition, in which we define (1) velocities $\alpha(x,z)$ and $\beta(x,z)$, (2) densities $\rho(x,z)$, (3) quality factors $Q(x,z)$ (for near-elastic media i.e. $Q>>1$, Aki and Richards (1980); Toksoz and Johnston (1981); Ben Menahem and Singh (1981); White (1983)), and (4) interfaces $z_i = z_i(x)$, $i=1,I$ separating $I-1$ layers each having a minimum thickness of h_j, $j=1, I-1$. Layer parameters and interfaces are at least of class C^2, that is twice continuously differentiable.

The source location $M_0(x_0,z_0)$ and its radiation pattern $f(\gamma_1,\gamma_2;\omega)$ are given (see II.3.7). For simplicity we shall consider only one receiver located at $M(x_r,z_r)$. Multiple source/receiver configurations follow by simple extension of this presentation.

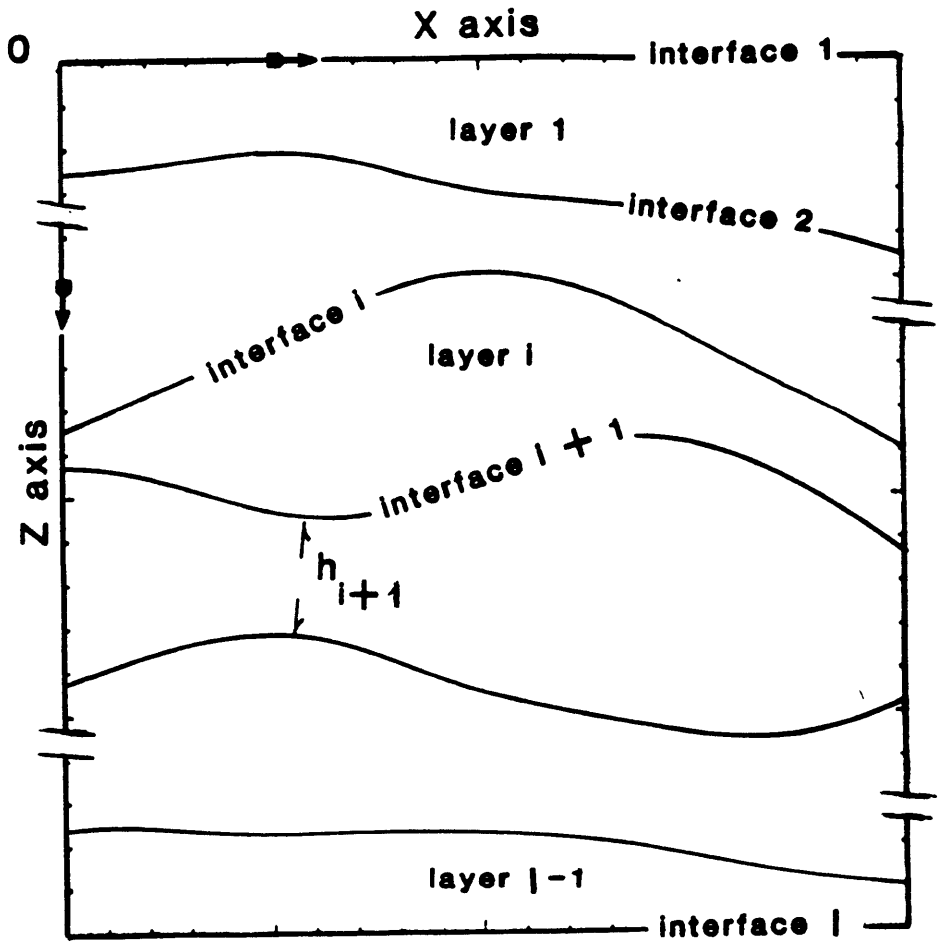


Figure 1. Global reference frame definition in model.

Chapter IV

Applicability conditions of ray theory derived in chapter II are assumed to be satisfied. In the case of layers and interfaces, the characteristic length l_0 (II.2.17) must include the interface radius of curvature and layer minimum thickness. The curvature of interface z_i at point x is given by (Eisenhart, 1909; Vygodski, 1975)

$$K_i^I(x) = \left[1 + \left(\frac{dz_i}{dx}\right)^2\right]^{-3/2} \left|\frac{d^2 z_i}{dx^2}\right|, \tag{2.1}$$

and the interface radius of curvature is equal to $1/K_i^I$. The definition of the threshold frequency introduced in (II.2.19) can be extended accordingly. The mode decoupling condition now reads

$$\omega \gg \omega_0, \quad \omega_0 \equiv \frac{1}{2} \max_{i=1,I;j=1,I-1} \left[|\nabla\alpha|,\ |\nabla\beta|,\ \alpha\frac{|\nabla\rho|}{\rho},\ \alpha K_i^I,\ \frac{\alpha}{h_j}\right]. \tag{2.2}$$

Firstly, we specify a type of wave (or of ray) that we wish to propagate in the model. For example, a P wave source propagating direct P waves up to the z_i interface, then, a reflected S wave from that interface, remaining a transmitted S wave all the way through the model. Such a characterization of a ray Ω is called *the ray code of* Ω. Rays are traced in the medium by solving the characteristics of the eikonal equation (II.4.5). Defining $\vec{p} = \nabla\tau$ and $\vec{r} = \vec{OM'}(x,z)$, these characteristics take the form

$$\frac{d\vec{p}}{d\tau} = -\nabla(\ln v), \tag{2.3}$$

$$\vec{p} = v^{-2}\frac{d\vec{r}}{d\tau},$$

where v is either α or β. This system is called the *ray tracing system*. The system is solved by specifying initial conditions of Ω (initial-value problem) consisting of $\tau=\tau_0$, $\vec{r} = \vec{r}_0 \equiv \vec{OM'}_0$ and $\vec{p} = \vec{p}_0$ satisfying the eikonal $\vec{p}_0^{\,2} = v^{-2}(\vec{r}_0)$. Defining the arclength s as the distance along Ω from M_0 to M', we have

$$s = \int_{M_0}^{M'} ds \quad , \tag{2.4}$$

with $ds = (dx^2 + dz^2)^{1/2}$. The travel time at M' is readily obtained recalling that $ds = v\, d\tau$, and is given by (II.4.7). The arclength at the source is set equal to s_0 (generally equal to zero). We can therefore write $M_0(s_0)$, and $M'(s)$. Across an interface z_i, the ray tracing is stopped. Local phase matching conditions are applied, requiring new initial conditions for the reflected/transmitted ray, depending on the ray code, before resuming computations. Explicit equations are presented in Cerveny, Molotkov and Psencik (1977). The phase manifestation of the travel time is given by $S(\tau) = \exp(i\omega\tau)$.

Secondly, we solve the eikonal equation in the ray centered coordinate system of Ω (II.4.8) known as the dynamic ray tracing system. The two parameters that are derived from it are the local wavefront curvature K_w along Ω and the surficial Jacobian J^S (II.4.9a). Here again, in the presence of interfaces, K_w and J^S must be correctly transformed (Cerveny, 1983). The transport equation (II.3.7) yields then at $M'(s)$ the complex amplitude of the displacement (see II.4.11) in the frequency domain. For a near-elastic medium, and for an explosive point source, it is given by

$$U_0(s,\omega) = \frac{i\omega c_0}{(\rho(s)\, v(s)\, J^S(s))^{1/2}} \exp\left[-\frac{\omega}{2}\int_{s_0}^{s} \frac{ds}{Q(s)v(s)}\right]$$

$$\prod_{i=1}^{I} R_i(D) \left[\frac{\rho(s_D^+)\, v(s_D^+)\, J^S(s_D^+)}{\rho(s_D^-)\, v(s_D^-)\, J^S(s_D^-)}\right]^{1/2} , \tag{2.5}$$

where the R_i is the appropriate reflection/transmission coefficient at interface z_i at the point D. s_D^- denotes the arclength at the point D on the side of the incident wave, and s_D^+ on the side of the reflected/transmitted wave.

In acoustic media, expression (2.5) is slightly modified since it is pressure that is observable (or the potential of U_0 see II.4.9)

$$A_0(s,\omega) = f(\gamma_1,\gamma_2;\omega) \left[\frac{v(s)}{J^S(s)}\right]^{1/2} \exp\left[-\frac{\omega}{2}\int_{s_0}^{s}\frac{ds}{Q(s)v(s)}\right]$$

$$\prod_{i=1}^{I} R_i^{ac}(D) \left[\frac{v(s_D^-)\, J^S(s_D^+)}{v(s_D^+)\, J^S(s_D^-)}\right]^{1/2}, \qquad (2.6)$$

with $v = \alpha$, and the appropriate acoustic reflection/transmission coefficient R_i^{ac}. Here, $v^2 = (\rho\,\kappa)^{-1}$, where κ is the medium adiabatic compressibility (Morse and Ingard, 1968). Considering a line source that will be used in chapter V, the source radiation pattern f, is derived equating (2.6) with the asymptotic expression of the Green's function for an acoustic line source in a homogeneous medium (i.e. $\frac{-i}{4} H_0^{(1)}(\omega\Delta/v_0)$). We have

$$f = -(8\,\pi\,\omega)^{-1/2}\, e^{i\pi/4}. \qquad (2.7)$$

Denoting $\hat{p}$ the unit vector along (tangent to) the ray Ω, and $\hat{n}$ the perpendicular to $\hat{p}$, the displacement vector at $M'(s)$ for P waves is

$$\vec{U}(M',\omega) = U_0(M',\omega)\, e^{i\omega\tau(M')}\, \hat{p}, \qquad (2.8)$$

for S waves, $\hat{p}$ is replaced by $\hat{n}$. In acoustic media the equivalent representation is given by (II.4.2).

Thirdly, the distance from ray Ω to receiver M must be computed. This is achieved by searching for the point $M'(s)$ on the ray such that the product of the ray slope at $M'(s)$ times the slope of the straight line joining the receiver to the ray at $M'(s)$, is equal to -1. If at this point $M'(s)$ on ray Ω the components are (x', z'), then the distance ray-receiver is

$$n = \left[(x_r - x')^2 + (z_r - z')^2\right]^{1/2} . \tag{2.9}$$

The receiver at M can be located in terms of the ray centered coordinate system of Ω, i.e. $M(s,n)$, as long as this system remains regular, that is $n\,K_R \ll 1$ (II.7.21). The assumption of smoothly varying medium is now necessary, since the wavefront is supposed to be about the same at M than at M'. The paraxial angle is then computed following (see figure 2.)

$$\chi = \tan^{-1}(n\,K_w) . \tag{2.10}$$

The next condition for the applicability of the paraxial ray method is the paraxial condition (II.7.20), requiring $|\chi| \ll 1$. An additional condition, that is implicit in the method, which is the paraxial version of condition (II.8.9) is that

$$n \ll l_0 . \tag{2.11}$$

This is a necessary condition. If receiver M is in the the region of medium where the characteristic length is l_0, the paraxial ray that will be used in determining the field at M, must carry the information about l_0. At this point we have rejected rays that do not satisfy the applicability conditions of ray and paraxial ray methods. Let us assume that ray Ω satisfies all the conditions for the extrapolation of the field from Ω at M' to receiver M.

Fourthly, we perform the paraxial corrections. The *kinematic* (or acoustic) paraxial correction for the travel time at M (see fig. 2 and Appendix II.C)

$$\tau(M) = \tau(M') + \frac{\delta R_w}{v} , \tag{2.12}$$

or in terms of the ray centered coordinates

$$\tau(s,n) = \tau(s,0) + n^2 \frac{K_w}{2} . \tag{2.13}$$

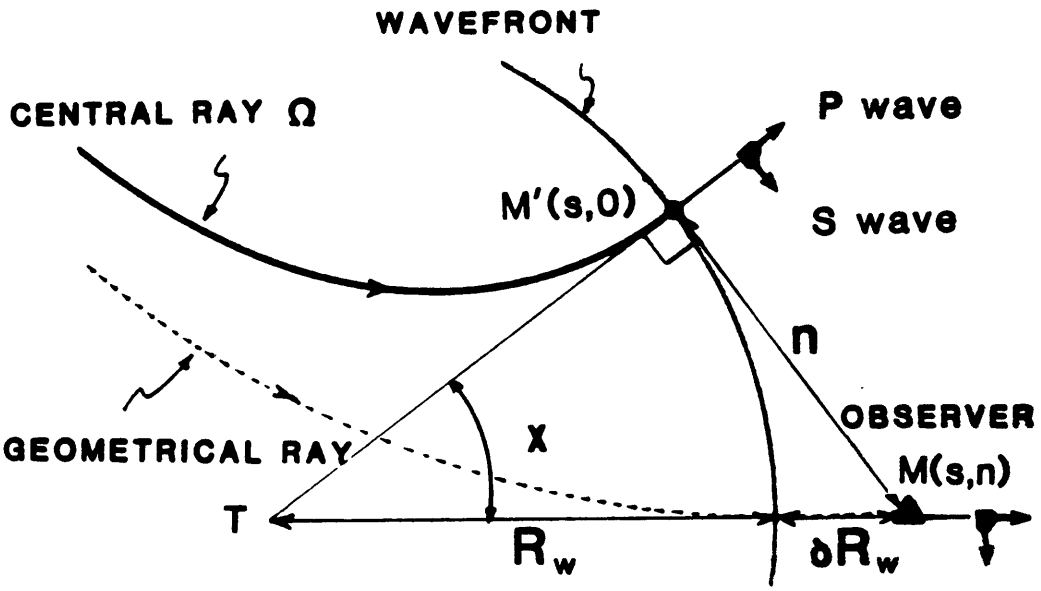


Figure 2. Paraxial ray geometry.

Then, the *dynamic* paraxial correction is performed for the amplitude at M.

$$\vec{U}(M,\omega) = q\ P \bullet \vec{U}(M',\omega)\ e^{i\omega \frac{\delta R_w}{v}}, \tag{2.12}$$

where $q = (1 + \delta R_w / (J^S)^{1/2})^{-1}$ for a point source, and $q = (1 + \delta R_w / J^S)^{-1/2}$ for a line source. The P operation is a rotation of (oriented) angle χ, of center T, from M' to M (figure 2). Denoting $[P]$ the matrix of P, we have

$$[P] = \begin{pmatrix} \cos\chi & \sin\chi \\ -\sin\chi & \cos\chi \end{pmatrix}. \tag{2.13}$$

The amplitude transformation in (2.12) is recognized to be a similarity of scale q , angle of rotation χ and center T. The high frequency Green's function at M is then, for P waves,

$$\vec{U}(M,\omega) = q\ U_0(M',\omega) \exp\left\{ i\omega \left[\tau(M') + \frac{n^2 K_w}{2} \right] \right\} P \bullet \hat{p}\ , \tag{2.14}$$

for S waves, $\hat{p}$ is replaced by $\hat{n}$. The angle χ is very small compared to one, since the paraxial condition is assumed satisfied. Further, in the far field and for a paraxial ray, we may approximate the scale factor q by one (i.e. $\delta R_w \ll (J^S)^{1/2}$ for a point source, and $\delta R_w \ll J^S$ for a line source). Within this approximation the following values are then considered in (2.14)

$$q \approx 1 \quad ; \quad [P] \approx \begin{pmatrix} 1 & \chi \\ -\chi & 1 \end{pmatrix}. \tag{2.15}$$

Lastly, if the source spectrum is equal to $\bar{\sigma}(\omega)$, the seismogram at receiver M, $\vec{u}(M,t)$, is obtained by multiplying the high-frequency Green's function by the source spectrum, and by inverse Fourier transforming the response,

$$\vec{u}(M,t) = FT^{-1} \left\{ \vec{U}(M,\omega).\bar{\sigma}(\omega) \right\}, \tag{2.16}$$

where the convention of Fourier transform is the same as that defined in Chapter III,

section 2.

3. EXAMPLES

Source time function is the Gabor wavelet described in Chapter III equation (6.1). Source parameters are $f_0 = 60$ Hz, $\Gamma = 4$ and $t_i = 0.03$ sec. The paraxial ray computation of (2.14) and (2.16) is done using a modified version of the Gaussian beam code (see Chapter III, section 6). The data generator and testing of the method is performed using a finite difference program written by Esmersoy (1985). It is a constant density acoustic code with absorbing boundaries, where synthetics are computed for a line source. The formulation is heterogeneous with an explicit scheme. We always satisfy the grid dispersion relation, $\Delta d = \alpha_{min}/(5f_{max})$, and the stability condition, $\Delta t \leq \Delta d/(2^{1/2}\,\alpha_{max})$, where Δd is the grid spacing; Δt the time sampling interval; f_{max} the highest frequency present in the source; and where α_{min} and α_{max} are the minimum and maximum velocities in the medium, respectively.

Model 1:

The model is a VSP in a two layer acoustic medium ($Q \to \infty$), with a tilted interface (figure 3). There are no free surface effects, and the 2D line source is at an offset distance of 0.75 km. The borehole is S-shaped with 20 receivers. The interface tilt is about $\Phi \approx 22^o$ (see Appendix G.1). The upper medium velocity is $\alpha_1 = 3.5\ km/sec$ and the lower medium velocity is $\alpha_2 = 4.5\ km/sec$. Pressure synthetics using the paraxial ray method are shown in figure 4. A maximum of ten rays per ray code is enough to construct the wavefield. The corresponding finite-difference synthetics are shown in figure 5. figure 6 displays the difference (residuals) Paraxial - Finite Difference.

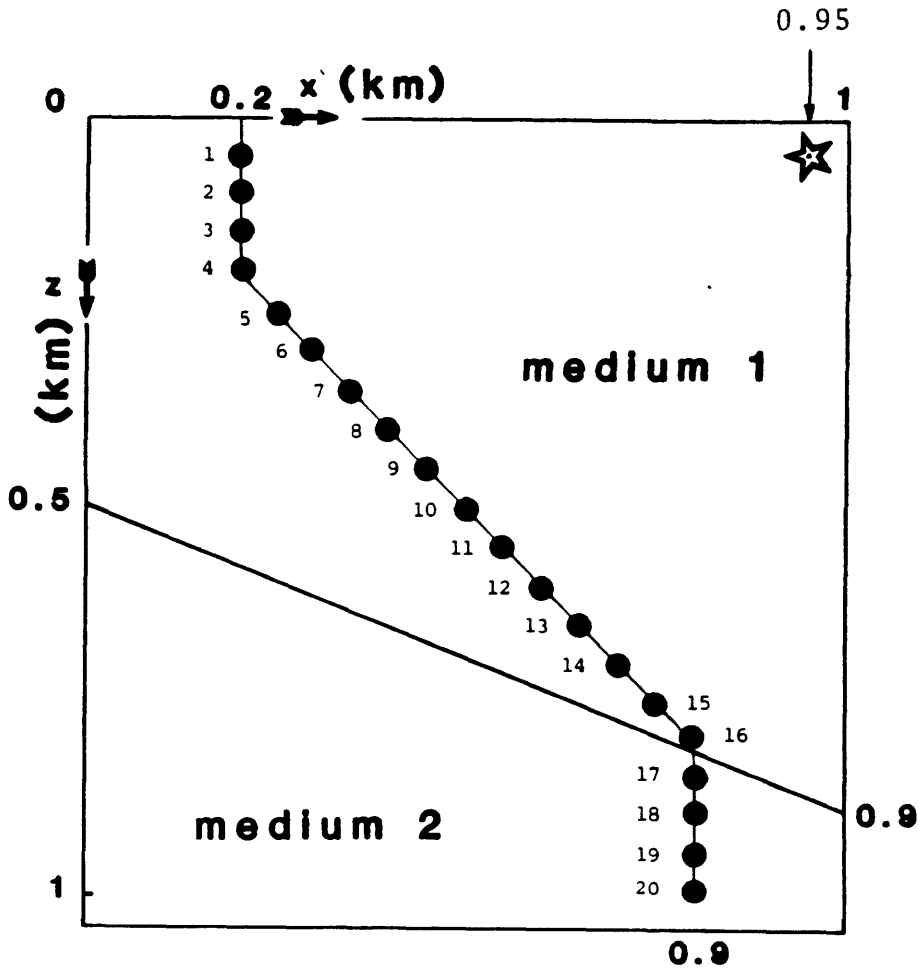


Figure 3. Geometry of Model 1.

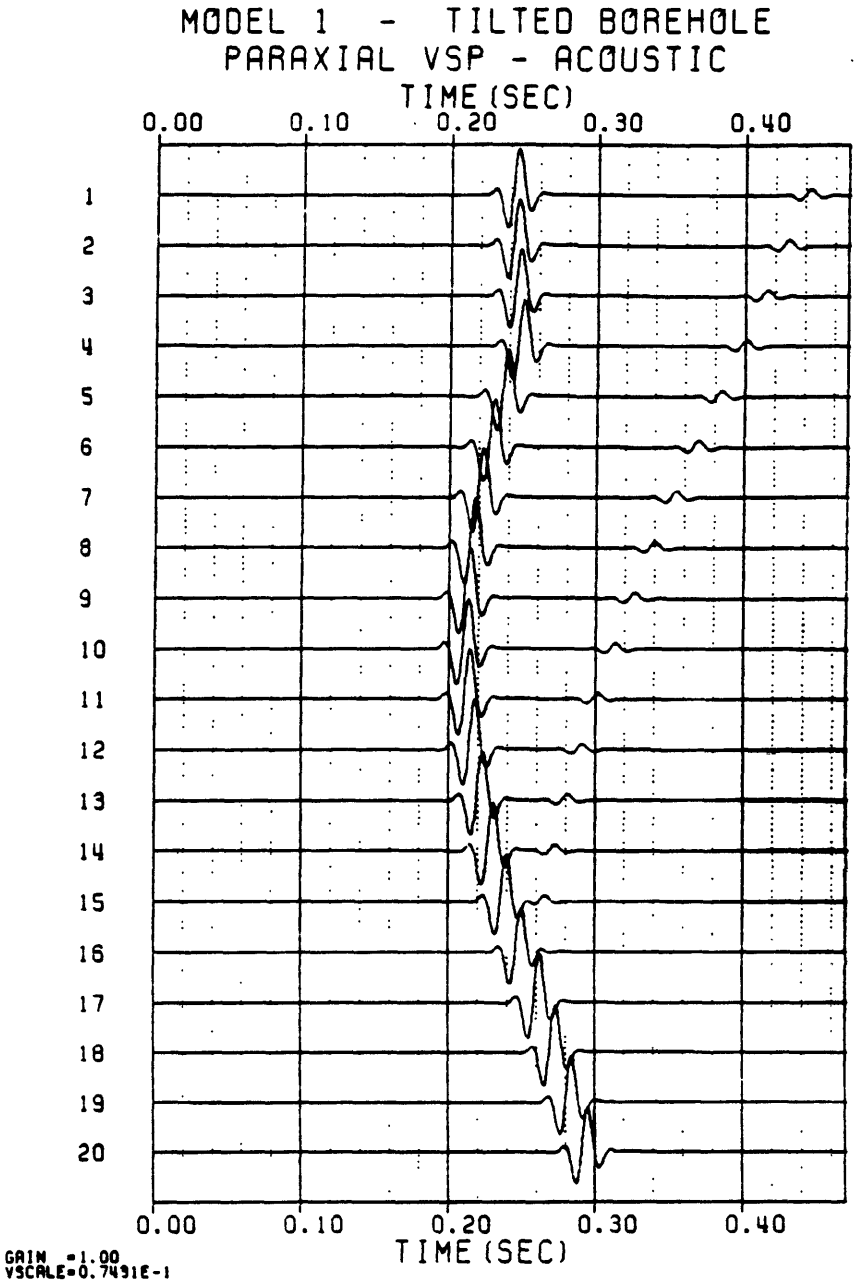


Figure 4. Pressure synthetics of Model 1 (paraxial rays).

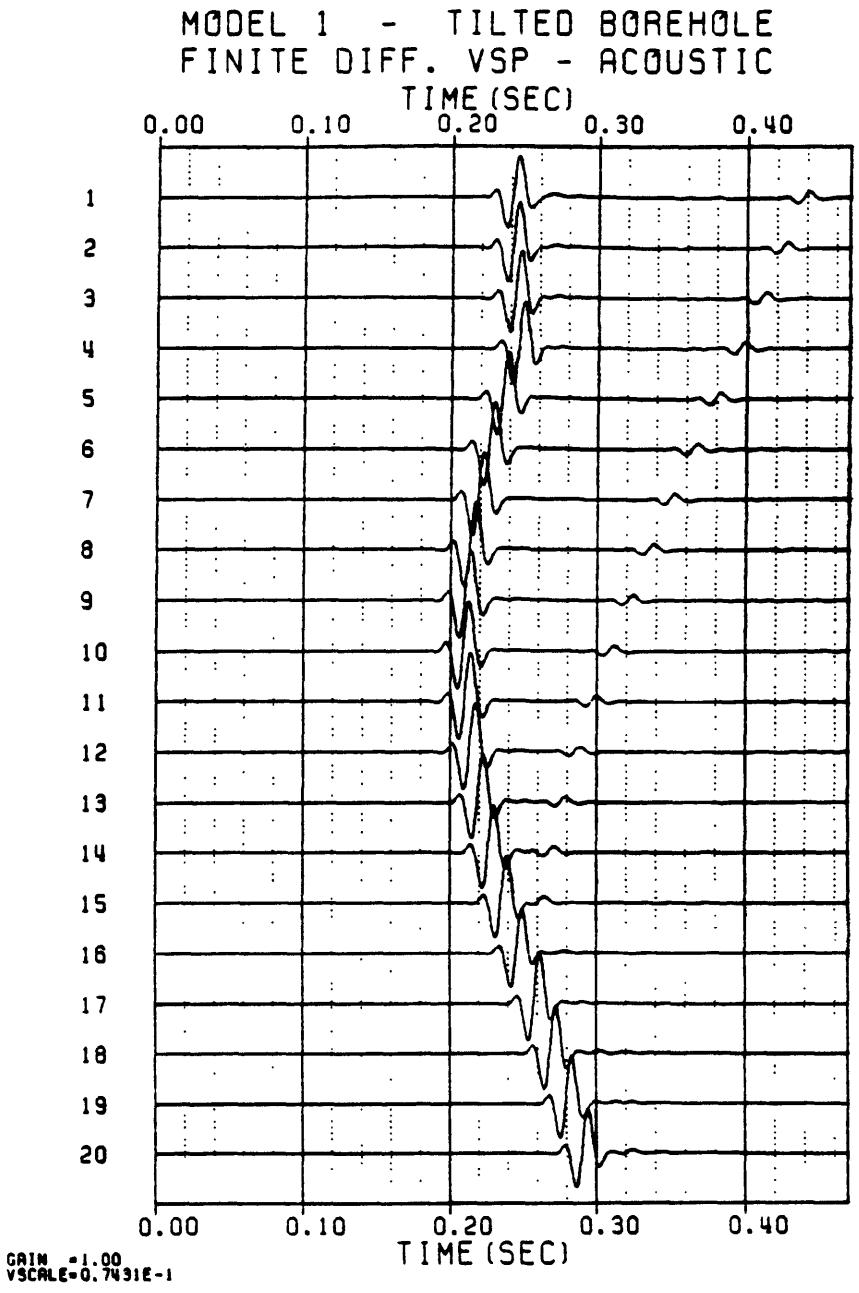


Figure 5. Pressure synthetics of Model 1 with the finite difference method.

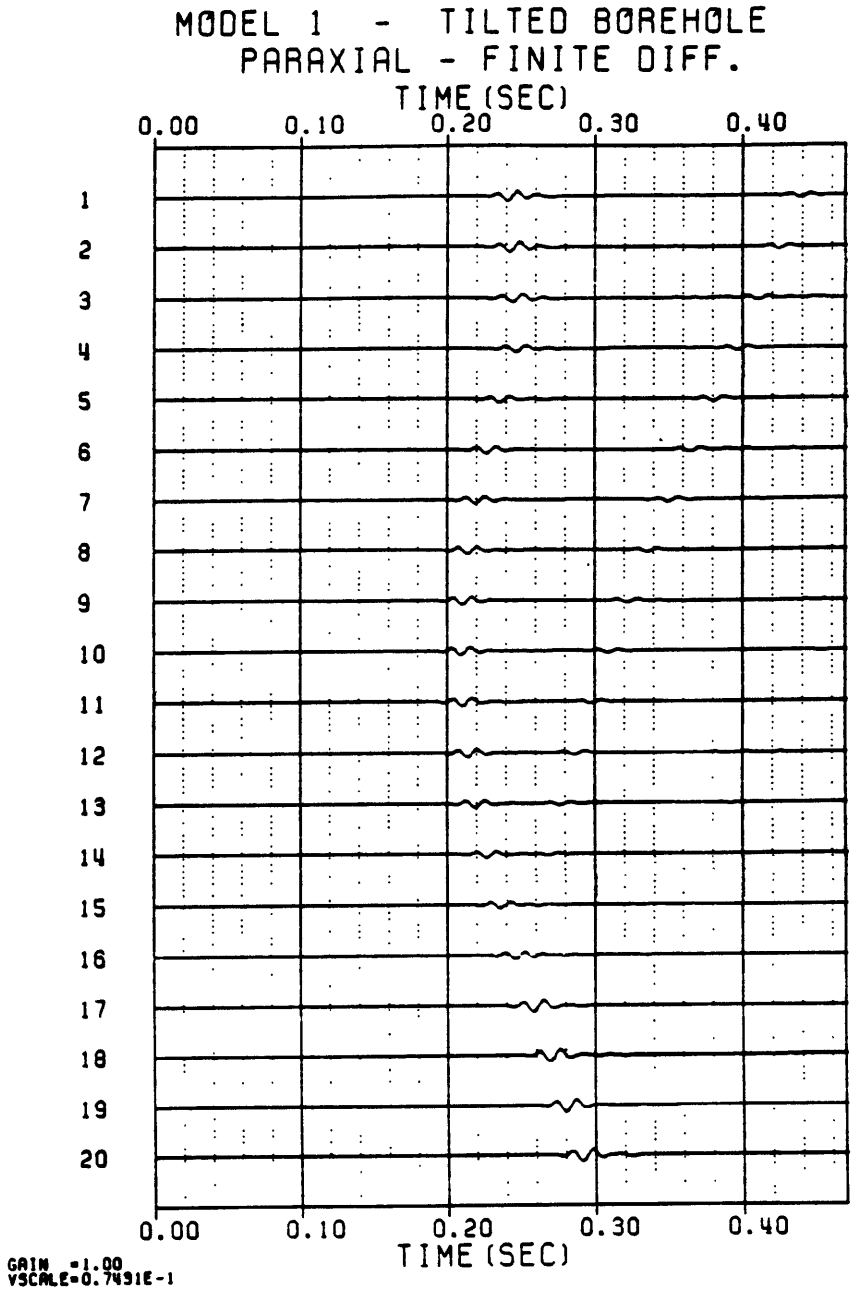


Figure 6. Residual (Paraxial - Finite difference) synthetics of Model 1.

The small differences are attributed to (1) grid dispersion, due to the finite spatial sampling of the medium; this is due, particularly to the propagation in the low-velocity layer; and (2) a small instability resulting from the finite time sampling which is affected by the high velocity layer. This example shows that paraxial ray amplitudes and travel times correctly reproduce the wavefield.

Model 2:

This model is the elastic version of Model 1, with same geometry (figure 3). The source is now an explosion. The factor $i\omega$ in (2.5) is not included because no comparison with an exact method is done. The upper layer parameters are: $\alpha_1 = 3.5\ km/sec$, $\beta_1 = 2\ km/sec$, and $\rho_1 = 2.7\ g/cm^3$. The lower layer parameters are: $\alpha_2 = 4.5\ km/sec$, $\beta_2 = 2.6\ km/sec$, and $\rho_2 = 3.0\ g/cm^3$. Here too, a maximum of ten rays per ray code is enough to construct the wavefield. Four ray codes are considered: (1) The direct P, (2) P to S converted transmitted, (3) P to P primary reflected, and (4) P to S primary converted reflected. Arrivals in the figures are labeled according to this numbering. figure 7 shows the vertical displacement synthetics, and figure 8, the horizontal displacement synthetics. The reflected wavefield comes from the other side of the borehole, with respect to the source, and is correctly taken into account by the method. A two-point ray tracing program would require at least six times more rays, considering that each receiver is found after 3 rays.

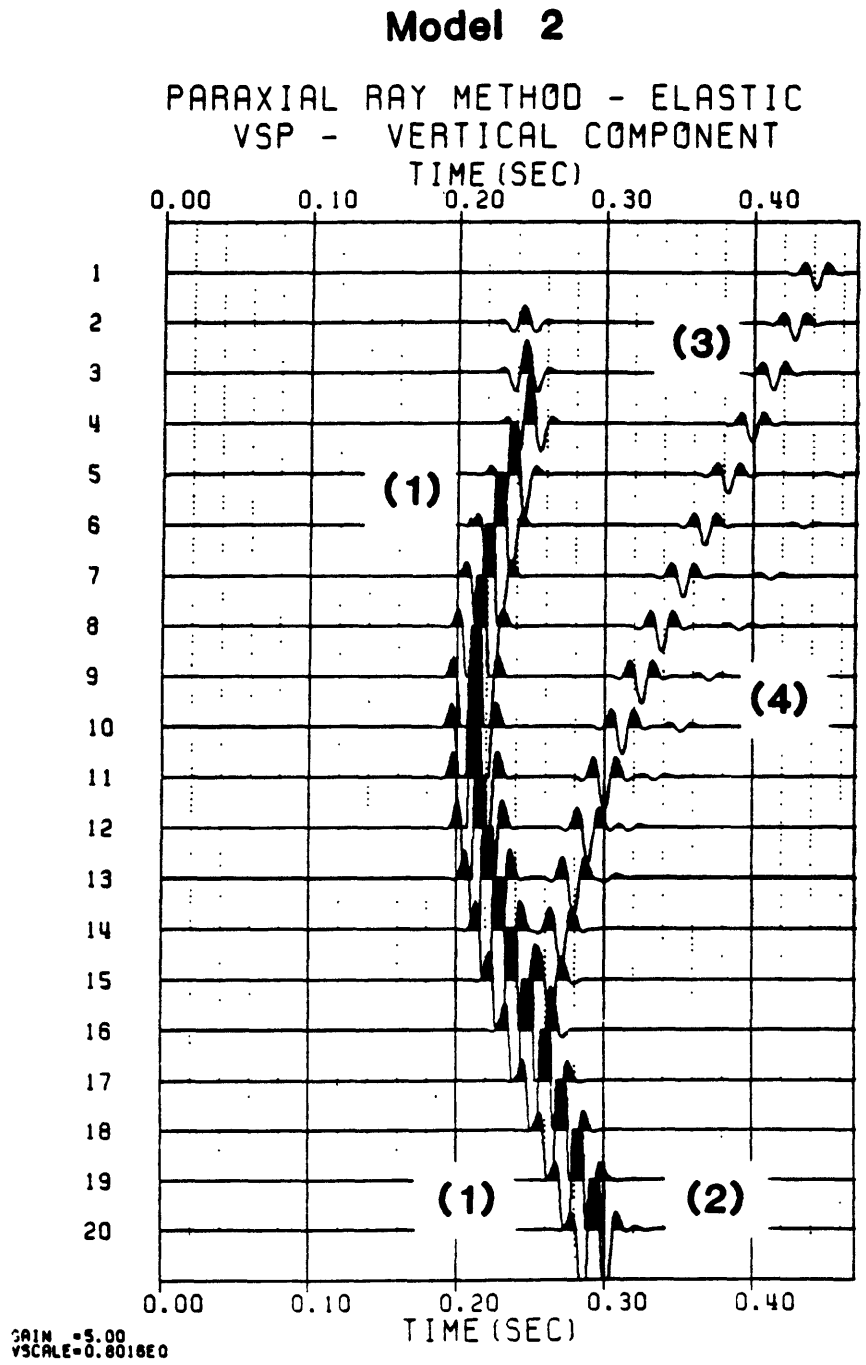


Figure 7. Vertical component of displacement of Model 2 (paraxial rays).

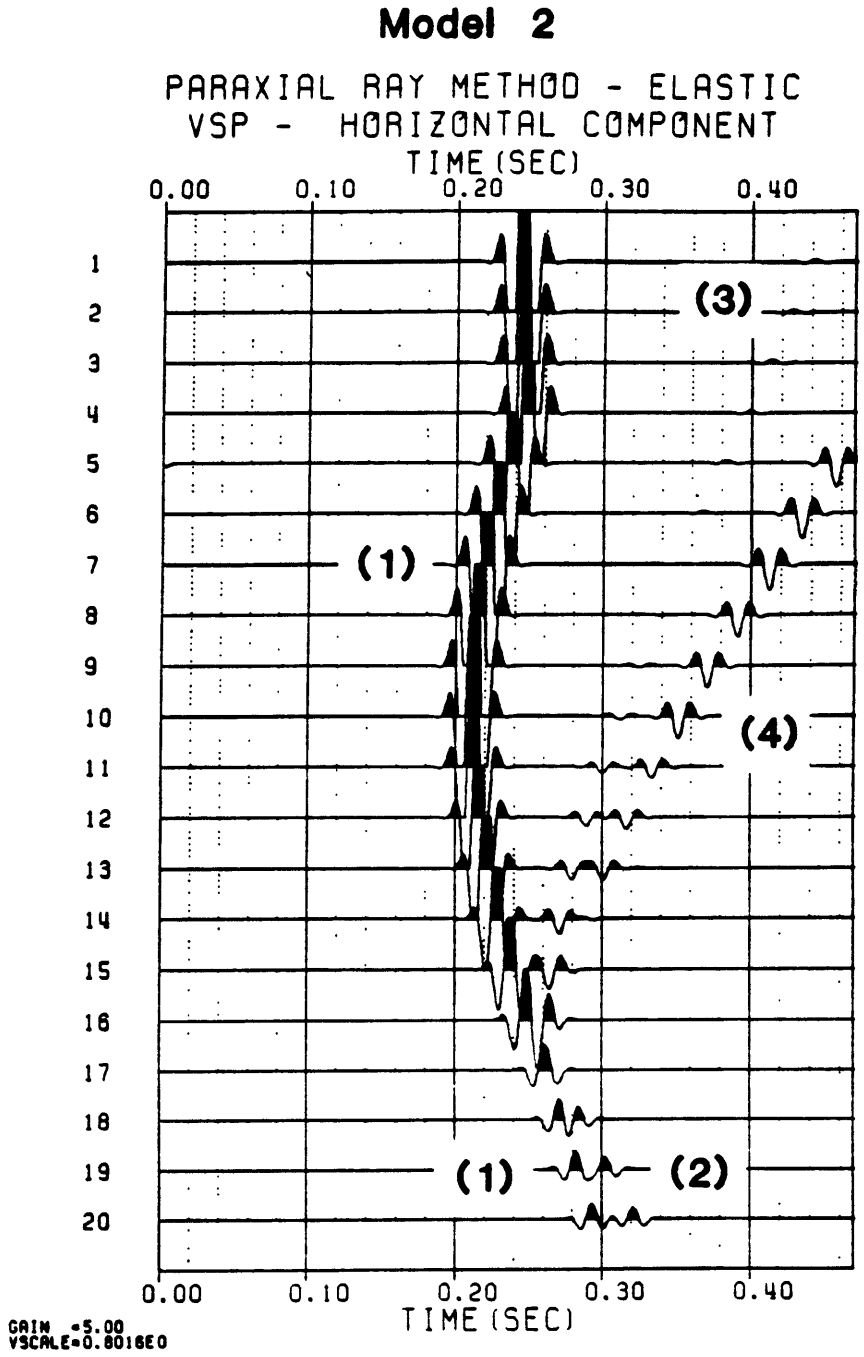


Figure 8. Horizontal component of displacement of Model 2 (paraxial rays).

Model 3:

We compare the paraxial ray method with the discrete wavenumber method (Bouchon and Aki, 1977). The model represents a simplified section of a borehole in Michigan, referred to as model 3. The 11 layer elastic media is subjected to a vertical point force located on the surface, with a source offset of 0.075 km (figure 9). There are 11 receivers. The discrete wavenumber synthetics are computed by Prange (1985). We have used 8 ray codes, with direct P, and primary P reflected rays from interfaces 5 through 11. The source parameters are f_0 = 50 Hz, Γ = 4, and t_i = 0. The medium threshold frequency, ω_0, (2.2) is on the order of 120 Hz x radians, whereas the angular frequency is about 314 Hz x radians. The mean wavelength, λ, is on the order of 100 meters. Within a constant amplitude factor, there is a fairly good agreement between the paraxial synthetics (figure 10) the discrete wavenumber synthetics (figure 11). Here again, even in the presence of thin layers ($h_8 \approx 20$ $meters$), asymptotic results are still representative of the the wavefield.

Model 4:

This last example is a four layer elastic model, containing a reef (figure 12). We combine VSP and surface reflection data. The explosion point source is located on the surface at 525 meters offset, with same source parameters as in Model 1. There is a total of 60 receivers: 40 receivers are placed in the VSP geometry, with initial receiver at depth 150 meters and last receiver at 930 meters (spacing 20 meters). The remaining receivers are on the surface, with 50 meters spacing. Receiver 21 is on interface 4 (the reef), whereas receivers 13 and 14 are separated by interface 3 (dipping interface).

Model 3

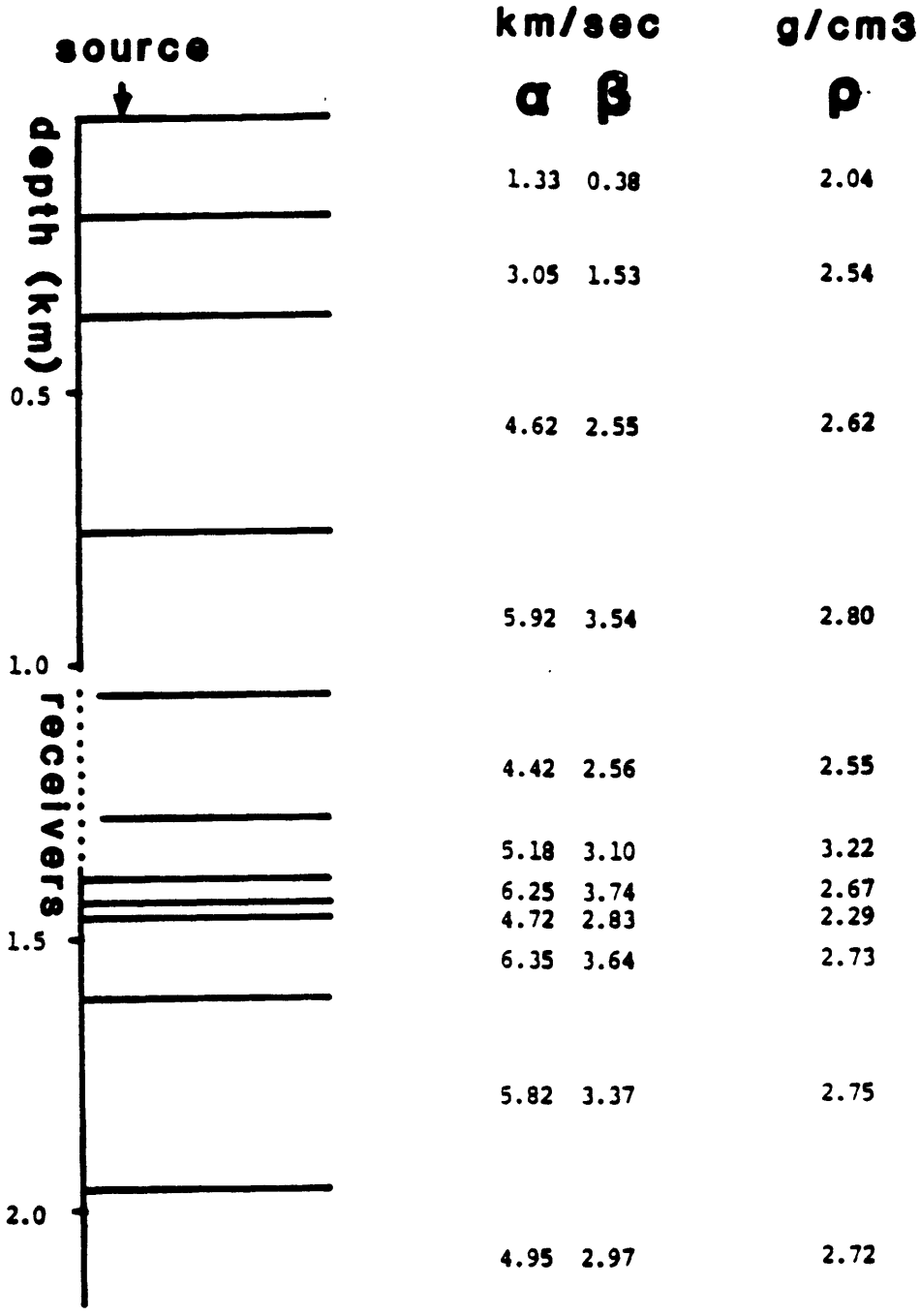


Figure 9. Geometry of Model 3.

Model 3

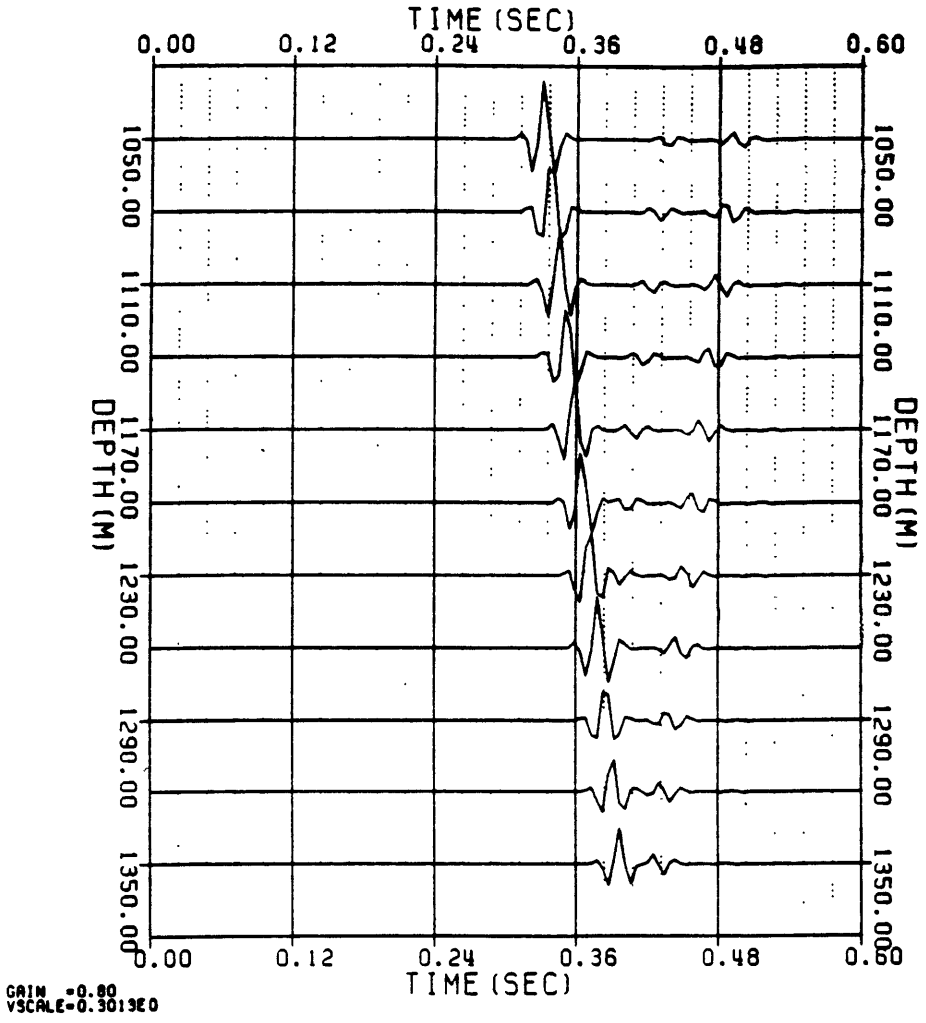


Figure 10. Vertical component of displacement of Model 3 (paraxial rays).

Model 3

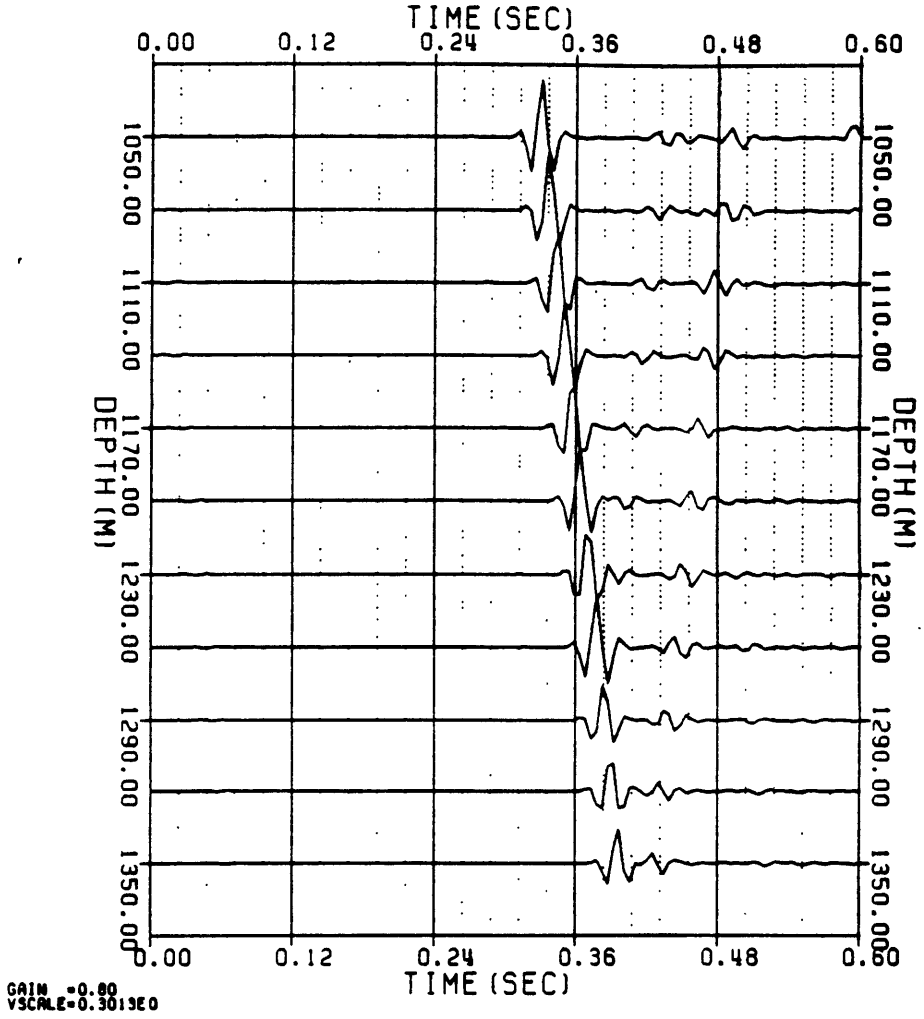


Figure 11. Vertical component of displacement of Model 3 with the discrete wavenumber method.

Layer parameters are the following: densities in g/cm^3,

($\rho_1 = 2.5, \rho_2 = 2.6, \rho_3 = 2.7, \rho_4 = 2.8$),

velocities in km/sec,

($\alpha_1 = 3.3, \alpha_2 = 3.5, \alpha_3 = 4, \alpha_4 = 4.5$).

Shear velocities are $\beta_i = \alpha_i / 3^{1/2}$. 10 ray codes were used, with a maximum of 30 rays per ray code. Figure 13 shows the vertical component of displacement, and figure 14, the horizontal component of displacement. The numbering in the figures corresponds to the following arrivals: (1) direct P wave, (2) P to S conversion at interface 2, (3) P to S conversion at interface 3, (4) P to S conversion at interface 4, (5) P to P primary reflection at interface 2, (6) P to S primary converted reflection at interface 2, (7) P to P primary reflection at interface 3, (8) P to S primary converted reflection at interface 3, (9) P to P primary reflection at interface 4, and finally, (10) P to S primary converted reflection at interface 4.

On surface receivers, arrivals (1), (5) and (6) interfere with each other. The receiver array is not large enough to allow separation of these waves. Arrival (8) is very poorly distinguished on surface data, and is better seen on the VSP. The reef distorts arrival (1) in layer 4 which should have been straight if there were no reef. The ray tracing of arrivals (1) and (9) are displayed in figures 15 and 16, respectively. Note on figure 15, the direct rays that were stopped because of post-critical angle of incidence. The wavefront right below interface 2, is extrapolated to neighboring receivers without difficulty with the paraxial ray method. Whereas, in two point ray tracing, near critical rays, many iterations are required to achieve the same goal.

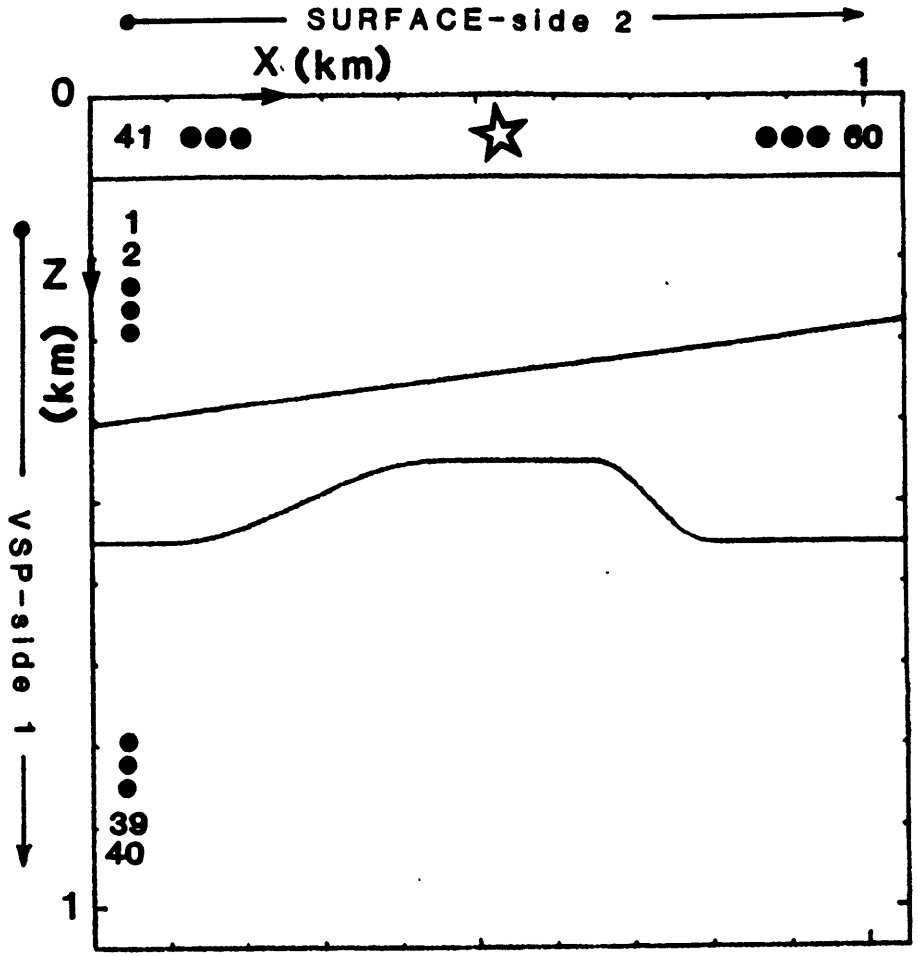


Figure 12. Geometry of Model 4.

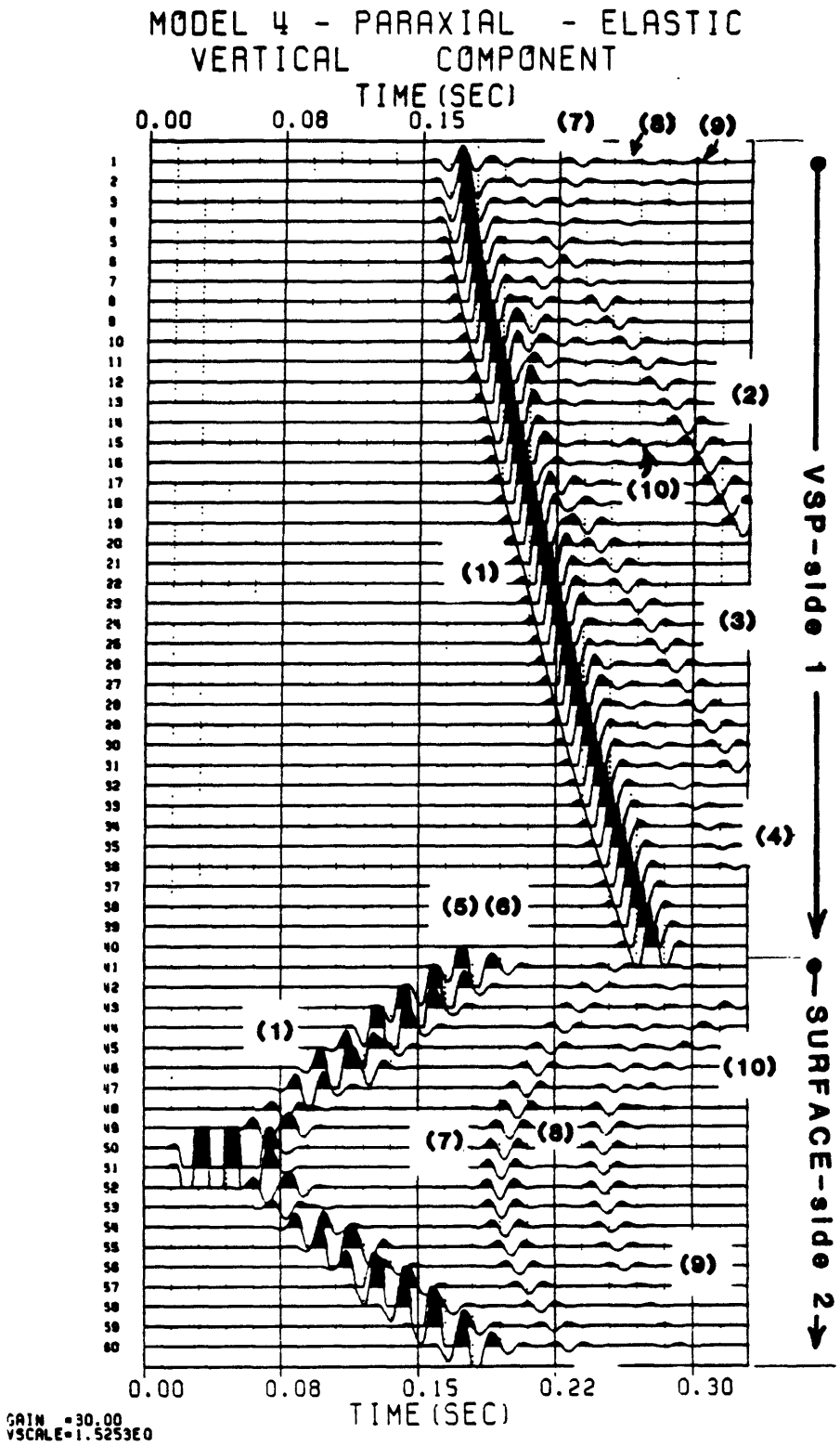


Figure 13. Vertical component of displacement of Model 4 (paraxial rays).

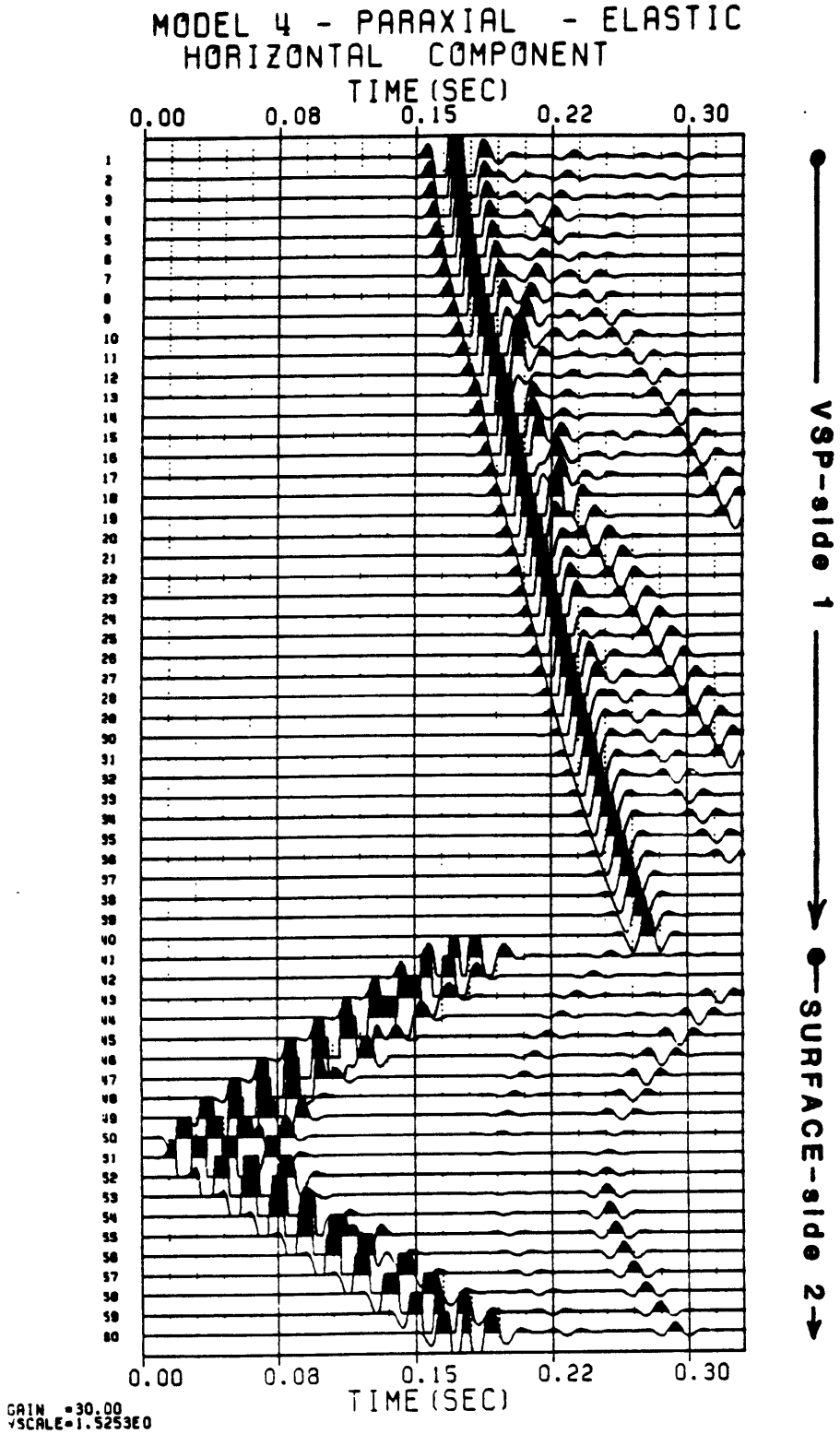


Figure 14. Horizontal component of displacement of Model 4 (paraxial rays).

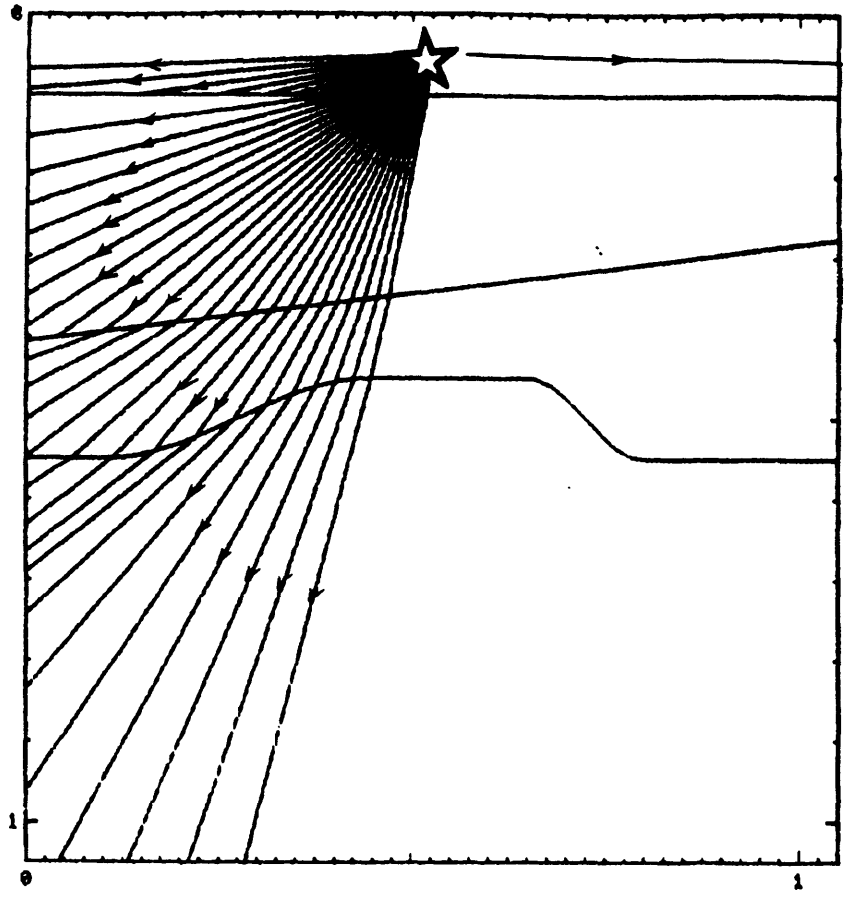


Figure 15. Ray diagram of direct P arrival, for Model 4.

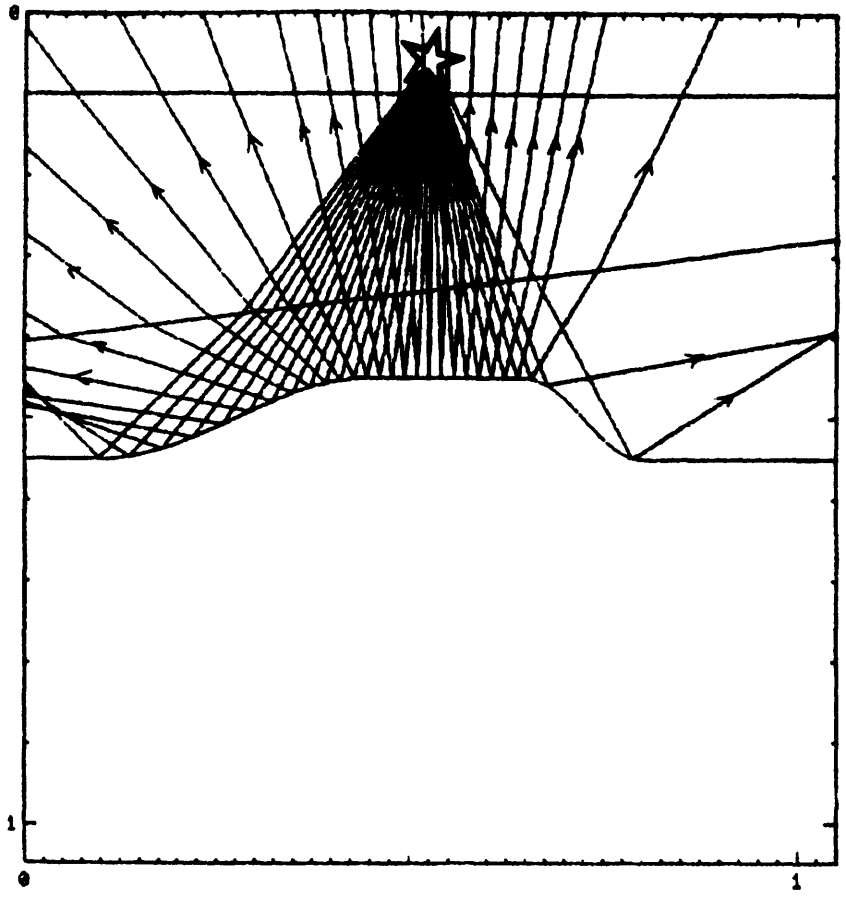


Figure 16. Ray diagram of P to P reflection at interface 4, for Model 4.

4. CONCLUSION

The paraxial ray method is an economical way of computing fast asymptotic Green's function in heterogeneous media. The paraxial corrections are (1) kinematic: approximation of the phase at the receiver, given the phase at the central ray, (2) dynamic: similarity transformation with scale close to unity and angle of rotation equal to χ.

Comparison with acoustic finite difference and elastic discrete wavenumber methods, is very satisfactory. Examples show the flexibility and robustness of the method.

In problems such as imaging and inversion of heterogeneous media, where large numbers of Green's function computations are required, the paraxial ray method offers an attractive way of achieving this task.

V. FULL WAVEFORM INVERSION

Chaque science, chaque étude, a son jargon inintelligible,
qui semble n'être inventé que pour en défendre les approches.
- Voltaire

1. INTRODUCTION

The present chapter deals with a full wave inversion for interface and layer parameters. The goal is to use combined sets of VSP and surface reflection data, multi offset VSP's, or cross borehole data, to estimate subsurface parameters. The paraxial ray method presented in Chapter IV generates the forward model. Amplitude information constrains the inversion uniquely by (1) post critical reflection phase effects, (2) density information, (3) medium density and attenuation factor and (4) the effect of the displacement component.

The inversion is done in the frequency domain since the paraxial results are computed in this space. Further, the data dimension is reduced. necessary, compared to time domain full - wave inversion, while maintaining all signal information. Finally, the assumption of uncorrelated samples is justified in this domain.

The heterogeneous media contain homogeneous layers separated by smooth interfaces. Interfaces are parametrized by simple functions. Prior information includes initial layer parameter estimates, along with their errors in measurement, the type of interfaces that are present (dipping interface, reef type, anticline, fault etc.), and a fixed point on each interface, given for example by a VSP experiment, or any other borehole logging information. This is a realistic consideration, since layer parameters

and interface locations are well determined in the borehole. Other types of prior information can be included such as data from other geophysical sets, or empirical relations between layer parameters.

The non-linear least-norm formulation is presented in the context of tensor algebra. A simple analogy exists between the data covariance operator and the metrical tensor of the basis under which the data is expressed. The inversion is finally reduced to a minimization of a a residual energy, which is the sum of squares of nonlinear functions for parameter estimates. This is then solved using a special case of the Gauss-Newton method known as Levenberg-Marquardt.

We first examine the sensitivity of the residual energy to parameters. Examples of inversion with finite difference data are presented. Finally, an example with field VSP data collected in Michigan, is inverted for velocities and interface dip angles.

2. GENERAL DERIVATION OF THE NON-LINEAR LEAST-NORM FORMULATION

The observables of a given geophysical system (measurements, data, synthetics etc.) can be represented by real vector functions, elements of a vector space which should be specified precisely. It is a real Euclidian space. Its properties are that (1) it is a discrete real linear vector space, (2) N-dimensional and, (3) in which a scalar product has been defined resulting in a Euclidian (L_2) norm. Following the same notations as Aki and Richards (1980), we shall call this space the N-dimensional $data$ space and denote it by **U**. We shall assume that **U** is a subspace of **E**, where **E** is the discrete linear space described in Appendix F.

Given a set of measurements described by a $data$ vector $\vec{d}$ in **U**, the non-linear least-norm inverse problem, consists in the following three steps:

(1) Determine precisely the parameters we wish to extract from the data $\vec{d}$. These parameters are in the form of a vector $\vec{m}$, defined as the *parameter* vector. Denote V the space spanned by the $\vec{m}$'s with dimension M. This is the *parameter* space, a subspace of E. However a necessary assumption is that V is a convex set. This means that for any two vectors $\vec{m}_1$, $\vec{m}_2$ in V, and for a scalar λ within [0,1], the vector $\lambda\vec{m}_1 + (1-\lambda)\vec{m}_2$ is an element of V. Generally, the two spaces U and V are different in nature since in the first case the vectors are formed by data points (eg. amplitude, phase at given time or frequency etc.) and in the second case the vectors contain parameters (eg. velocities, attenuations, interface parameters etc.). However, in some instances (section 3), the data space may contain elements that are present in V space.

(2) Formulate a *model* in the form of algebraic, differential or integral equations (generally non-linear) with its associated initial and boundary conditions. Let this model be represented by a functional $\vec{f}$ that operates on a vector $\vec{m}$ of V space and transforms it into a vector $\vec{f}(\vec{m})$ of U space. This vector is defined as the *synthetics*. It is the model prediction of a data set. At this point, we should identify the parameters that are best suited for the model. There are some models for which it is not possible to uniquely estimate all the parameters from measurements. However, certain functions of the parameters can be estimated. Beck and Arnold (1977) derive a parameter identifiability criterion. A posteriori parameter resolution and variances (section 4) contain information on how good and reliable the parameters are estimated, given the data set and the model.

The *discrepancy* vector $\vec{\varepsilon}$ between the data and the synthetics is a functional defined for any $\vec{m}$ in V,

$$\vec{\varepsilon}(\vec{m}) \equiv \vec{d} - \vec{f}(\vec{m}) . \tag{2.1}$$

This vector is sometimes called the additive error in measurements or the residual vector; in other terms it represents what is not modeled by $\vec{f}(\vec{m})$.

Given a basis in U, with its corresponding metrical tensor $\overleftrightarrow{W}$, the total energy of the discrepancy $\vec{\varepsilon}$ is defined as half the Euclidian norm $s(\vec{\varepsilon},\vec{\varepsilon})$ (see Appendix F.17)

$$E(\vec{m}) \equiv \frac{1}{2} \vec{\varepsilon} \bullet \overleftrightarrow{W} \bullet \vec{\varepsilon} , \tag{2.2}$$

or in matrix form

$$E(\vec{m}) \equiv \frac{1}{2} [\vec{\varepsilon}]^T [\overleftrightarrow{W}] [\vec{\varepsilon}] . \tag{2.3}$$

It is sometimes referred to as the residual energy of the data-synthetics system. The real and symmetric metrical tensor $\overleftrightarrow{W}$ acts as a weighting factor for each component $\varepsilon^i \varepsilon^j$. The data is expressed in the basis of $\overleftrightarrow{W}$. The physical meaning of $\overleftrightarrow{W}$ will be introduced later in this section. A sufficient condition that will be required later is that the functional E be of class C^2, that is twice continuously differentiable.

(3) Finally, assuming that $E(\vec{m})$ is a convex function and that an $\vec{m}^*$ exists in V, such that

$$E(\vec{m}^*) = \min_{\vec{m} \text{ in } V} E(\vec{m}) , \tag{2.4}$$

the last step is to find an estimate of $\vec{m}^*$, say $\hat{m}$, so that $\hat{m} \approx \vec{m}^*$. The function E represents the equation of a hypersurface in V×R space (R denotes the set of real numbers). The convexity of E requires that the set of points in V×R above the surface described by E is convex.

If (2.4) is achieved we then have, for any $\vec{m}$ in **V**, $E(\hat{m}) \leq E(\vec{m})$. The model $\vec{f}$ with parameter vector $\hat{m}$ is said to "best fit" the data $\vec{d}$ within the vector error $\vec{\varepsilon}(\hat{m})$ or with an energy discrepancy of $E(\hat{m})$. This optimization can be interpreted geometrically as searching the deepest or lowest point on the hypersurface of E in **V**×**R** space. $E(\hat{m})$ is the non-linear least-norm (in L_2 sense) of $\varepsilon(\vec{m})$.

For our purposes, the minimization (2.4) can be interpreted as maximizing a Gaussian probability density function for the data vectors (Van Trees, 1968; Tarantola and Valette, 1982), which is a special case of the maximum likelihood method. This assumes that vectors in **U** have a likelihood function of the form

$$L(\vec{d}\,|\,\vec{m}) \;=\; C \exp\left[-\frac{1}{2}[\vec{d} - \vec{f}(\vec{m})]^T\, V^{-1}\; [\vec{d} - \vec{f}(\vec{m})]\right], \qquad (2.5)$$

with observations $\vec{d}$, expected value $\vec{f}(\vec{m})$, and covariance matrix V. $\vec{d}$ is a random sample of size N, and C is a normalization constant factor. This is the probability that event $\vec{d}$ be realized, given $\vec{m}$, thus $\vec{f}(\vec{m})$ and V. The maximum likelihood method searches for an $\vec{m}^*$ that maximizes $L(\vec{d}\,|\,\vec{m}^*)$. The minimization seeks to estimate the parameters $\vec{m}$ present in the expected value $\vec{f}$. Taking the absolute value of the natural logarithm of $L(\vec{d}\,|\,\vec{m})$ converts the maximization of $L(\vec{d}\,|\,\vec{m})$ into the minimization of $\frac{1}{2}[\vec{d} - \vec{f}(\vec{m})]^T\, V^{-1}\; [\vec{d} - \vec{f}(\vec{m})]$. Identifying this representation of the problem with the present formulation (2.1-4) leads to the equality between the matrix $[\mathbf{W}]^{-1}$ and the prior covariance matrix V. The covariance matrix determines the type of basis under which the data is expressed. The minimization handles the prior information as weighting factors for individual data components. More accurate observations are weighted more heavily. Equation (2.5) is then written as $L(\vec{d}\,|\,\vec{m}) = C\exp(-E(\vec{m}))$, where $\vec{d}$ is an implicit variable in E.

If the covariance matrix is diagonal (uncorrelated events), the base vectors are orthogonal. The matrix $[\overset{\leftrightarrow}{W}]$ is diagonal with components reduced to scale factors. Each data component is totally unrelated to, or independent of, another component since the residual energy remains unchanged when cross-term components are added. The energy for each component ε^i is weighted according to the scale factor W_{ii}. If the covariance matrix is equal to a scalar constant w times the identity matrix (white noise if the problem is set in the frequency domain), the base vectors are orthonormal. There is equal prior information about the data variance and consequently there is no privileged weighting ($W_{ii} = w$). In terms of tensor algebra, curiously enough, there is loss of variance: covariant and contravariant components are reduced to usual cartesian components. The minimization is considered as unconstrained by a priori information. The next section will cover in more in detail how a priori information and constraints are handled.

The introduction of the metric tensor $\overset{\leftrightarrow}{W}$ enables the minimization problem to handle more general exponential families than the normal distribution (2.5). The metric tensor (and the corresponding data covariance matrix) is required to be constant and diagonal. Equation (2.2) with (2.1) can be written explicitly as

$$E(\vec{m}) = -(\vec{f}\bullet\overset{\leftrightarrow}{W})\bullet\vec{d} + \frac{1}{2}\,\vec{f}\bullet\overset{\leftrightarrow}{W}\bullet\vec{f} + \frac{1}{2}\,\vec{d}\bullet\overset{\leftrightarrow}{W}\bullet\vec{d}\ . \tag{2.6}$$

Requiring that the distribution has an expected value equal to $\vec{f}$, and setting $\vec{p}(\vec{f}) = \vec{f}\bullet\overset{\leftrightarrow}{W}$, $q(\vec{f}) = \frac{1}{2}\,\vec{f}\bullet\overset{\leftrightarrow}{W}\bullet\vec{f}$, and $g(\vec{d}) = \frac{1}{2}\,\vec{d}\bullet\overset{\leftrightarrow}{W}\bullet\vec{d}$ enables the likelihood function (2.5) to be written in the more general form

$$L(\vec{d}\,|\,\vec{m}) = C\,\exp -\left[\vec{p}(\vec{f})\bullet\vec{d} - q(\vec{f}) - g(\vec{d})\right]. \tag{2.7}$$

The expected value of the distribution is then calculated via the scalar result of Charnes, Frome and Yu, (1976), which we generalize for the vector case,

$<\vec{d}> \equiv (\nabla\vec{p})^{-1} \bullet \nabla\vec{q} = \vec{f}$, the differentiation being with respect to $\vec{f}$, and <> denotes the expected value. We have used identities (F.19-20). Similarly, the diagonal metric tensor is defined formally as

$$\overleftrightarrow{W} = \nabla\vec{p} \ , \tag{2.8}$$

the differentiation being with respect to $\vec{f}$. This is the vector expression of eq.2.3 in Charnes, Frome and Yu, (1976). The *theorem* is:

> If $\vec{d}$ is a random sample of size N satisfying (2.7) with $<\vec{d}> = \vec{f}$, then a maximum likelihood estimate of $\vec{m}$, say $\hat{m}$, will satisfy the equation $\nabla E(\hat{m}) = 0$, provided $\hat{m}$ belongs to V space.

For a likelihood function of general exponential family of type (2.7), and its appropriate metric tensor (2.8) and expected value $\vec{f}$, the linearized stochastic inverse, to be presented in section 4 of this chapter, would yield estimates that are identical to those obtained using the maximum likelihood principle.

Parseval-Plancherel theorem (Bass, 1977; Roddier, 1978) states that the energy of a system is conserved when transformed into the Fourier space. Therefore, for any two conjugate variables (t,ω) with ω being the dual variable of t, and for a N_P-point ($N_P \leq N$) Discrete Fourier Transform (DFT) (Oppenheim and Schafer, 1975),

$$\sum_{i=1}^{N_P} \varepsilon_i^2(\vec{m},t_i) = \frac{1}{N_P} \sum_{j=1}^{N_P} \bar{\varepsilon}_j^2(\vec{m},\omega_j) \ , \tag{2.9}$$

where the bar denotes the Fourier transform of the function considered. For example, if the data vectors depend on time, we can therefore study the minimization process (2.4) in the time or equivalently in the frequency domain, whichever is suitable.

This property is a particular case of the more general property that all unitary operators, such as the DFT, are isometric (i.e. conserve the norm). An operator A is unitary if $A^T = A^{-1}$, where the superscript T denotes the transposed of A, and the complex conjugate transposed if A is complex. Thus if A operates on a vector $\vec{\varepsilon}$ (i.e. $A \bullet \vec{\varepsilon}$) the norm is

$$s(A \bullet \vec{\varepsilon}\,,\, A \bullet \vec{\varepsilon}) \;=\; s(\vec{\varepsilon},\vec{\varepsilon})\,. \tag{2.10}$$

This property, in signal processing, is sometimes referred to as information preserving. Such a transformation has a beneficial effect when observations (signals) are considered as random processes, for which noise is characterized by the covariance matrix, generally non-diagonal. The time for which the autocorrelation function begins to decrease significantly is an approximate measure of the time interval, Δt, for which two events are uncorrelated. In terms of power spectral density (Fourier transform of the autocorrelation function), if the highest signal frequency is f_{max}, the required time interval, Δt, in order to achieve decorrelation, must be greater than f_{max}^{-1}. However, time samples of properly sampled signals (i.e. $\Delta t < (2f_{max})^{-1}$) are correlated. Performing a unitary transformation to the signal redistributes the variance associated with the transformed signal samples, into almost uncorrelated samples (Yechiam and Pearl, 1979), thus performing a quasi-diagonalization of the covariance matrix. Other than being simpler in computation, since cross terms of the covariance matrix are neglected compared to the diagonal terms, this enables the transformed coefficients of the signal to be processed.

3. A PRIORI INFORMATION - CONSTRAINTS

The previous section introduced how prior information on the errors in the data are handled. If some prior information is known about the parameters, such as for instance a

measurement set, their maximum error, upper or lower bounds etc., or results from other types of measurements (gravity data etc.), the present formulation handles these additional information as submodels $\vec{f}_q$ and subdata $\vec{d}_q$ belonging to subspace $\mathbf{U}_q$, $q=1, K$. These subspaces are supplementary in **U**. Since each subspace is characterized by its metrical tensor whose matrix is the inverse of the prior covariance matrix of the a priori data, the statistical inference of this assumption is that that uncertainties in $\mathbf{U}_q$ are uncorrelated with uncertainties in $\mathbf{U}_p$, for $p \neq q$. This is generally a plausible assumption (Aki and Richards, 1980; Tarantola and Valette, 1982). The method could handle the case where the spaces are not disjoint. Introduction of tensors whose components are tensors, would achieve this goal, but this problem is out of the scope of this study.

The presence of prior information adds more robustness and help reduce the variance of the parameters. In probability theory, the way of including prior information utilizes the maximum a posteriori method (Beck and Arnold, 1977). It is a special case of Bayesian estimation (Jackson and Matsu'ura, 1985). The method is based on Bayes's theorem, and estimated value of the parameters are called Bayesian estimates.

The residual energy will take into account separately each type of prior information or constraints. For example, this energy will be required to increase wherever a type of constraint is violated, forcing the optimization to search for the minima in a domain where the constraints produce the least residual energy. It is therefore imperative that this information be specified precisely since it is able to force the optimization to seek an induced minima that could be far from the unconstrained minimum.

Given a parameter vector $\vec{m}$ in **V** space, the model $\vec{f}$ generates synthetics $\vec{f}(\vec{m})$ in **U** space. In order to incorporate in the optimization process the possibility that some or all the parameters are constrained in some manner by *a priori* information it is

sufficient to vector decompose uniquely $\vec{f}(\vec{m})$ into K different subspaces of **U**, K being the total number of different types of information and constraints. The model $\vec{f}$ can be viewed as a linear superposition of K submodels $\vec{f}_q$. The data $\vec{d}$ is decomposed the same way as $\vec{f}(\vec{m})$. The partial residual vector is then $\vec{\varepsilon}_q = \vec{d}_q - \vec{f}_q$. The subspaces $\mathbf{U}_q$, q =1, K, are *supplementary* i.e. their direct sum forms **U**. This can be written as

$$\vec{\varepsilon}(\vec{m}) = \bigoplus_{q=1}^{K} \vec{\varepsilon}_q(\vec{m}) , \tag{3.1}$$

where $\vec{\varepsilon}_q(\vec{m})$ belongs to $\mathbf{U}_q$, and

$$\mathbf{U} = \bigoplus_{q=1}^{K} \mathbf{U}_q , \tag{3.2}$$

where $\oplus$ is the direct sum of linear spaces. Each subspace $\mathbf{U}_q$ is specified by its metrical tensor $\overset{\leftrightarrow}{W}_q$, whose matrix is the inverse of the a priori covariance matrix V_q^{-1}. The corresponding total covariance matrix is then composed of K block-diagonal submatrices. From the supplementarity and linearity of each subspace $\mathbf{U}_q$, the total residual energy will then be the sum of the partial residual energy, that is

$$E(\vec{m}) = \sum_{q=1}^{K} E_q(\vec{m}) , \tag{3.3}$$

where $E_q \equiv \frac{1}{2} \vec{\varepsilon}_q \bullet \overset{\leftrightarrow}{W}_q \bullet \vec{\varepsilon}_q$.

As an illustration, we shall be considering two types of parameter constraints (K=3): (1) A prior information on the parameters requiring an observation vector (prior mean) and a covariance matrix of the parameters with the assumption of independent parameters and normal probability density for the parameter errors. This type of constraint induces minimum weight in the residual energy when the parameters are at their observation value (point wise constraint), and a norm type of increase away from

it. (2) A penalty constraint requiring physical or mathematical bounds on the parameters. This constraint attributes zero weight in the residual energy for all the parameters that are within the finite domain of definition (finite domain constraint) and an increasing weight as they get further away from that region. Other types of point wise or finite domain constraints could be added such as in barrier methods (see Luenberger, 1973). The model $\vec{f}_1$ generating $\vec{f}_1(\vec{m})$ will be the unconstrained model producing synthetic seismograms (amplitude and phase at time or frequency samples). The models $\vec{f}_2$ and $\vec{f}_3$ generate the a priori parameter synthetics $\vec{f}_2(\vec{m})$ and the penalty synthetics $\vec{f}_3(\vec{m})$, respectively. Considering a parameter $\vec{m}$ in **V** space, the submodels $\vec{f}_q$, the synthetics $\vec{f}_q(\vec{m})$ and the data $\vec{d}_q$ can now be explicitly defined.

Seismogram model (Model 1)

Given seismogram data $\vec{d}_1$, $\vec{f}_1$ corresponds to the forward model that generates the synthetic seismograms $\vec{f}_1(\vec{m})$ that is compared to $\vec{d}_1$. The forward model here is implemented using the paraxial ray method. The subspace in which the seismograms belong is $\mathbf{U}_1$, with dimension N_1. The matrix of the metric tensor $\overleftrightarrow{W}_1$ is equal to the inverse of the covariance matrix V_1^{-1}, which is an input.

Parameter prior information model (Model 2)

The data $\vec{d}_2$ for this model is set equal to the given a priori parameter observation ($\vec{d}_2 \equiv \vec{m}_0$). The matrix of the metric tensor $\overleftrightarrow{W}_2$ is equal to the inverse of the covariance matrix V_2^{-1}, which is an input. The synthetics are the parameter themselves, i.e. $\vec{f}_2 = \vec{m}$. The subspace involved is $\mathbf{U}_2$ with dimension N_2=M.

Penalty model (Model 3)

The data $\vec{d}_3$ for this model is equal to zero ($\vec{d}_3 \equiv 0$). Consider a number of functional constraints on the parameters defining a finite domain C in V space. The function $\vec{f}_3$ is such that

$$\begin{aligned} \vec{f}_3(\vec{m}) &= 0 \qquad \text{for all } \vec{m} \text{ in } C, \\ &\neq 0 \qquad \text{for all } \vec{m} \text{ not in } C. \end{aligned} \tag{3.4}$$

In the case where C is defined by a number of inequality constraints:

$$C = \{\vec{m} \text{ in } V : c_i(\vec{m}) \leq 0,\ i=1, N_3\}, \tag{3.5}$$

a useful penalty function is

$$f_3^i(\vec{m}) = \left[\max[0 ; c_i(\vec{m})]\right]^{\frac{\alpha}{2}} \quad ; \qquad i=1, N_3, \tag{3.6}$$

where α an even positive integer (with typical values of 2 or 4). The subspace spanned by $\vec{f}_3$ is U_3 with dimension N_3. The matrix of the metric tensor $\overleftrightarrow{W}_3$ is taken to be diagonal.

The total residual energy (2.2) with (3.3) can be written as

$$E = \frac{1}{2}\left(\vec{\varepsilon}_1 \bullet \overleftrightarrow{W}_1 \bullet \vec{\varepsilon}_1 + \vec{\varepsilon}_2 \bullet \overleftrightarrow{W}_2 \bullet \vec{\varepsilon}_2 + \vec{\varepsilon}_3 \bullet \overleftrightarrow{W}_3 \bullet \vec{\varepsilon}_3\right), \tag{3.7}$$

or, in terms of components

$$\begin{aligned} E(\vec{m}) = \frac{1}{2}\sum_{i,j=1}^{N_1} (W_1)_{ij}\left[d^i - f_1^i(\vec{m})\right]\left[d^j - f_1^j(\vec{m})\right] + \\ + \frac{1}{2}\sum_{i,j=1}^{N_2} (W_2)_{ij}\left[m_0^i - m^i\right]\left[m_0^j - m^j\right] + \frac{1}{2}\sum_{i=1}^{N_3} (W_3)_{ii}\left[f_3^i(\vec{m})\right]^2, \end{aligned} \tag{3.8}$$

with $N_1 + N_2 + N_3 = N$.

This equation shows the effects of constraints in the optimization. The residual energy is weighted according to "how far" the estimated parameters are from their a priori values, in addition to the variance weighting. The different contributions in the residual energy at the right hand side of (3.8) are explicit. The first term corresponds to the seismogram model, the second term to the parameter prior information model and the third to the penalty model. Geometrically, the constraints increase the convexity of the hypersurface of E in V×R space by raising the energy wherever the constraints are active (i.e. violated). This forces the search for the minimum residual energy within the domain defined by the a-priori information.

If other measurements or a priori information were to be implemented in the minimization, they would be associated with models generating synthetics. Their partial residual energy would then be added to (3.7), with the corresponding weights expressed by their metrical tensor.

4. THE LINEARIZED STOCHASTIC INVERSE

The non-linear problem (2.4) described in the preceding section is generally solved by iteration. A typical method of attacking the problem, known as Newton's method (or variations of it), reduces the non-linear system to a sequence of linear least-squares problems. For the sake of simplicity and clarity we shall first review the method in 1-Dimensional V space (m is a scalar), then generalize the results for vector representation in M-Dimensional V space. Two equivalent approaches are possible: Either we linearize the total energy E, then decompose the result into individual contributions, or else decompose the energy via (3.3), then linearize each contribution separately. We shall follow the first derivation, comparing the final result with that in Tarantola and Valette (1982).

One-dimensional problem

Figure 1 shows a typical example of a convex function $E(m)$. There exists a point m^* such that $E(m^*)$ is minimum or equivalently such that

$$E'(m^*) = 0 , \tag{4.1}$$

where the prime denotes the derivative with respect to m.

Following Dennis and Moré (1977), Newton's method can be derived by assuming that a given m_k is an approximation to m^*, and that in a neighborhood of m_k, say at point m_{k+1}, the linearization of $E'(m_{k+1})$ is a good approximation to $E'(m_{k+1})$. We can write

$$E_{k+1}' \approx E_k' + E_k'' (m_{k+1} - m_k) , \tag{4.2}$$

where $E_k \equiv E(m_k)$. This reduces the non-linear problem (4.1) by solving iteratively for k the linear system $E_{k+1}' = 0$, or equivalently

$$-E_k' = E_k'' \Delta m_k , \tag{4.3}$$

with $\Delta m_k \equiv m_{k+1} - m_k$.

Assume $E'' \neq 0$ and that for more generality its inverse is computed approximately. Given an initial a-priori guess m_0 to m^*, Newton's method updates the approximation following

$$\Delta \hat{m}_k = -(E_k'')^{-1} E_k' , \tag{4.4}$$

where $\Delta \hat{m}_k \equiv \hat{m}_{k+1} - m_k$, k being an integer representing the iteration number. The hat represents an approximate estimation of the real value since $(E'')^{-1}$ is computed approximately, i.e $\Delta \hat{m}_k \approx \Delta m_k$.

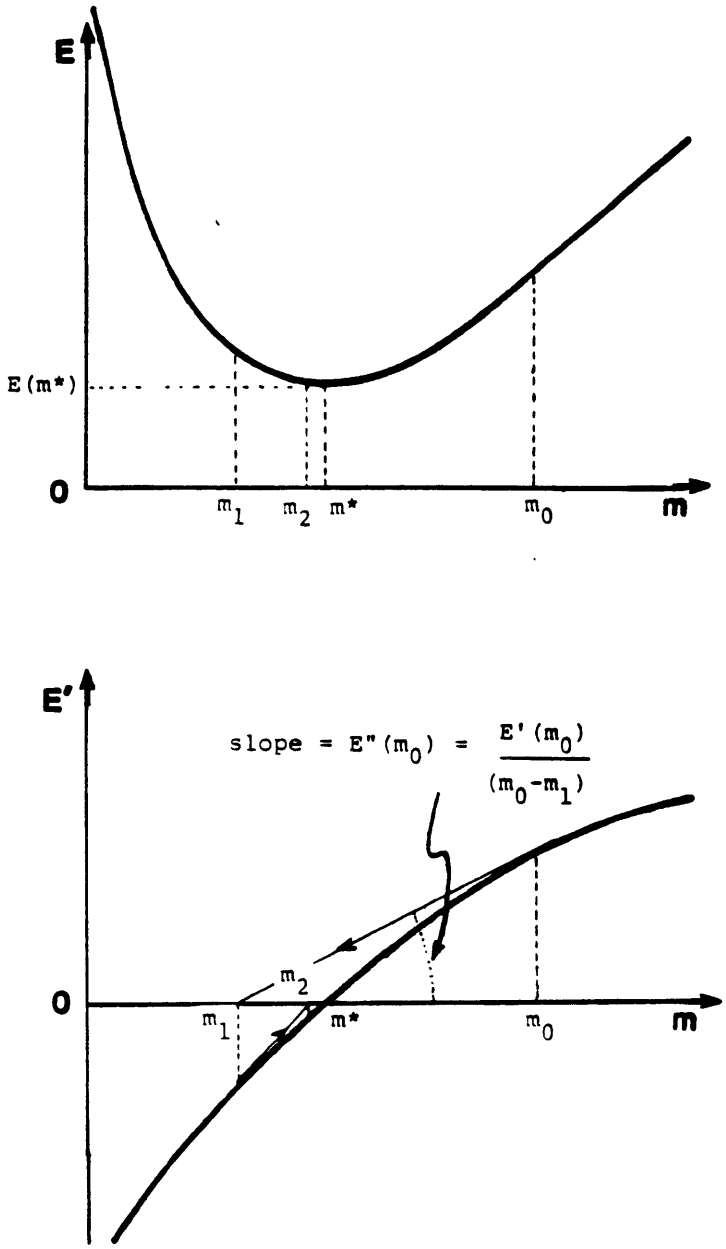


Figure 1. Geometrical representation of a nonlinear 1D minimization.

The algorithm is as follows:

a) enter $\vec{d}$, model $\vec{f}$ and m_0, and exit conditions EXC

b) compute E_k, E_k' and E_k''

c) compute $\Delta\hat{m}$ following (4.4): $m_{k+1} = m_k + \Delta\hat{m}$

d) test EXC, if true STOP: k iterations, $\hat{m} = m_{k+1}$

e) update $k \rightarrow k+1$ go to step b

where $\hat{m}$ is an estimate of m^*. The exit conditions are on m_k and/or E_k and are for example an upper bound on the relative difference between two consecutive iterates of m, and/or of E. Another convergence criterion could the minimum number of digits for which two consecutive iterates of m agree.

Multi-dimensional problem

The vector generalization of (4.4) is straightforward. The scalar m_k is replaced by the M-Dimensional vector $\vec{m}_k$. E_k' and E_k'' are replaced by the vector ∇E and the dyadic $\nabla\nabla E$, respectively. We shall omit, except when needed, the subscript k denoting the iteration number, since, by analogy to the 1-D case, the updating occurs only when all the operations at fixed k are terminated (part e of the algorithm). Recalling that the base vectors are independent of $\vec{m}$ and are fixed for a given a priori covariance matrix (fixed $\overleftrightarrow{W}$), all derivatives with respect to $\vec{m}$ will therefore not affect the basis and its metrical tensor $\overleftrightarrow{W}$. If $\overleftrightarrow{W}$ is dependent on $\vec{m}$, then each iteration would update the a priori covariance matrix, which is a generalization of the present method. This would lead to the computation of Christoffel symbols (last term in F.18), which is straightforward but out of the scope of this study. The data vector $\vec{d}$ is assumed independent of $\vec{m}$ too. Since $\overleftrightarrow{W}$ is symmetric we will use the relation $\overleftrightarrow{W}^T = \overleftrightarrow{W}$ (F.9).

Define the two dyadics

$$\ddot{G}^T \equiv -\nabla\vec{\varepsilon} = \nabla\vec{f} \quad \text{and} \quad H \equiv \nabla\nabla E\,, \tag{4.5}$$

where the superscript T denotes the transposed dyadic, the differentials being evaluated at $\vec{m}_k$. The components of the $\ddot{G}^T$ dyadic are (see F.18),

$$(G^T)_i^j = \nabla_i f^j = \frac{\partial f^j}{\partial m^i}\,. \tag{4.6}$$

Therefore, $G_i^j = \nabla_j f^i$.

Since $E = \frac{1}{2}\vec{\varepsilon}\cdot\ddot{W}\cdot\vec{\varepsilon}$ (2.2), use of (F.19) yields

$$-\nabla E = \ddot{G}^T\cdot\ddot{W}\cdot\vec{\varepsilon}\,, \tag{4.7}$$

and

$$H = -\nabla(\ddot{G}^T\cdot\ddot{W}\cdot\vec{\varepsilon}) = \ddot{G}^T\cdot\ddot{W}\cdot\ddot{G} - \nabla\ddot{G}^T\cdot\ddot{W}\cdot\vec{\varepsilon}\,. \tag{4.8}$$

The components of the triadic $\nabla\ddot{G}^T$ are

$$(\nabla G^T)_{ij}^p = \nabla_i (G^T)_j^p = \frac{\partial^2 f^p}{\partial m^i\,\partial m^j}\,. \tag{4.9}$$

Set $F \equiv -\nabla\ddot{G}^T$, the vectorized equation of Newton's algorithm follows from (4.4) defining $\Delta\vec{m} \equiv \vec{m}_{k+1} - \vec{m}_k$, k being the iteration number, and assuming that H is invertible

$$\Delta\vec{m} = \left(\ddot{G}^T\cdot\ddot{W}\cdot\ddot{G} + F\cdot\ddot{W}\cdot\vec{\varepsilon}\right)^{-1}\cdot\ddot{G}^T\cdot\ddot{W}\cdot\vec{\varepsilon}\,, \tag{4.10}$$

where F, $\ddot{G}$ and $\vec{\varepsilon}$ are evaluated at $\vec{m}_k$.

The $N\times M$ matrix $[\ddot{G}]$ is defined as the *sensitivity* matrix. It is the matrix of partial first-derivatives of the error components with respect to the parameters. It is sometimes called the Jacobian matrix. The $M\times M$ matrix $[H]$ is defined as the *Hessian*

The $M \times M \times N$ matrix $[F]$ involves second partial derivatives of residuals with respect to parameters. The algorithm is identical to that developed for the 1-Dimensional case. The convergence and rate of convergence of the method have been studied extensively (Acton, 1970; Karmanov, 1977; Dennis and More, 1977).

The *Gauss-Newton* methods use the approximation $H \approx \ddot{\vec{G}}^T \bullet \ddot{\vec{W}} \bullet \ddot{\vec{G}}$. Convergence conditions are given in Nazareth (1980). The approximation is justified asymptotically if for example the residual energy is small or if its surface is close to a paraboloid (i.e. quasi-linear residual vector). In Appendix H, we give a one-dimensional example of this approximation, which introduces the concept. Using the Gauss-Newton assumption, equation (4.10) reduces to

$$\Delta \vec{m} = (\ddot{\vec{G}}^T \bullet \ddot{\vec{W}} \bullet \ddot{\vec{G}})^{-1} \bullet \ddot{\vec{G}}^T \bullet \ddot{\vec{W}} \bullet \vec{\varepsilon}, \tag{4.11}$$

which is the familiar weighted least-squares solution of the linear problem

$$\vec{\varepsilon} = -\ddot{\vec{G}} \bullet \Delta \vec{m}, \tag{4.12}$$

with a priori information contained in $\ddot{\vec{W}}$. The energy minimization of $\vec{\varepsilon}$ weighted by $\ddot{\vec{W}}$, yields then, $\Delta \hat{\vec{m}}$.

The linearized *stochastic inverse operator*, at step k, is defined as

$$\ddot{\vec{L}} = (\ddot{\vec{G}}^T \bullet \ddot{\vec{W}} \bullet \ddot{\vec{G}})^{-1} \bullet \ddot{\vec{G}}^T \bullet \ddot{\vec{W}}, \tag{4.13}$$

which is represented by a $M \times N$ matrix $[\ddot{\vec{L}}]$. It transforms the vector $\vec{\varepsilon}$ in **U** space into a vector $\Delta \hat{\vec{m}}$ in **V** space.

As an example, we shall calculate explicitly (4.11) when different models are involved. The two models considered are (1) the seismogram model, Model 1, and (2) the parameter prior information model, Model 2, described in the previous section. The residual vector is then (3.1) $\vec{\varepsilon} = \vec{\varepsilon}_1 \oplus \vec{\varepsilon}_2$. We therefore have from (4.5) $\ddot{\vec{G}} = \ddot{\vec{G}}_1 \oplus \ddot{\vec{G}}_2$.

Similarly, we may write symbolically $\ddot{W} = \ddot{W}_1 \oplus \ddot{W}_2$. From the supplementarity of the two spaces U_1 and U_2 (see 3.2), the dot product of these quantities (ε_i , G_i , W_i) will only pick quantities who have same subscript. Equation (4.11) then reads

$$\Delta\hat{m} = (\ddot{G}_1^T \cdot \ddot{W}_1 \cdot \ddot{G}_1 + \ddot{G}_2^T \cdot \ddot{W}_2 \cdot \ddot{G}_2)^{-1} \cdot (\ddot{G}_1^T \cdot \ddot{W}_1 \cdot \vec{\varepsilon}_1 + \ddot{G}_2^T \cdot \ddot{W}_2 \cdot \vec{\varepsilon}_2) . \qquad (4.14)$$

Since $\vec{\varepsilon}_1 = \vec{d}_1 - \vec{f}_1$, $\vec{\varepsilon}_2 = \vec{m}_0 - \vec{m}$, we have from (4.7) $\ddot{G}_1^T = \nabla\vec{f}$ and $\ddot{G}_2 = \ddot{G}_2^T = \ddot{I}$. $[\ddot{W}_1]$ is the inverse of the data covariance matrix, and $[\ddot{W}_2]$ is the inverse of the parameter covariance matrix. Equation (4.13) takes then the final form

$$\Delta\hat{m} = (\ddot{G}_1^T \cdot \ddot{W}_1 \cdot \ddot{G}_1 + \ddot{W}_2)^{-1} \cdot \left[\ddot{G}_1^T \cdot \ddot{W}_1 \cdot (\vec{d}_1 - \vec{f}_1) + \ddot{W}_2 \cdot (\vec{m}_0 - \vec{m})\right], \qquad (4.15)$$

for which the matrix representation is identical to equation (25) in Tarantola and Valette (1982).

Henceforth, we shall include prior information before the linearization, as shown in (3.8). In general we would minimize the sum of $N_1^2 + N_2^2 + 2N_3 - N$ nonlinear functions (least norm) in M variables (or parameters) by a a special case of the maximum likelihood method, referred to as the Levenberg-Marquardt method. If all the a priori covariance matrices are diagonal, the functions are squared (least squares) and their total number reduce to N. The Levenberg-Marquardt method introduces an additional term in the inverse operator (4.13) of the form $\gamma^2 \ddot{D}$, where $\ddot{D}$ is a (0,2) diagonal dyadic whose components are equal to

$$D_{ii} = (\ddot{G}^T \cdot \ddot{W} \cdot \ddot{G})_{ii} , \qquad (4.16)$$

and $D_{ij} = 0$ for $i \neq j$. γ^2 a positive scalar called the Levenberg-Marquardt parameter, that is updated at every iteration. This particular choice of $\ddot{D}$ has the effect of making each iteration invariant under scale changes in the parameters. A good review of this method is found in Beck and Arnold (1977), Aki and Richards (1980), and in Lines and

Treitel (1984).

The inverse operator in (4.13) is then recast as

$$\ddot{L} = (\ddot{G}^T \bullet \ddot{W} \bullet \ddot{G} + \gamma^2 \ddot{D})^{-1} \bullet \ddot{G}^T \bullet \ddot{W} . \tag{4.17}$$

The search direction and iteration are therefore computed from

$$\Delta \vec{m} = \ddot{L} \bullet \vec{\varepsilon} . \tag{4.18}$$

The method is bounded by the Gauss-Newton method (for γ=0), and by the Steepest-descent method (for $\gamma \gg 1$). This addition results in reducing the size and changing the direction of the step at each iteration. The use of implicitly scaled variables limits the step size $|\Delta \vec{m}|$ in (4.18), in any direction where the residual energy changes rapidly. Depending on the strategy, iterations may start alike the steepest descent method with final iterations close to the Gauss-Newton method. In this case, the parameter γ^2 is chosen adaptively and decreases as the well-behaved minimum is approached. If the neighborhood of the minimum is poorly-conditioned, this parameter remains finite. Among its advantages, it maintains nonsingularity of the inverse operator and improves the search for the minimum (Beck and Arnold, 1977; Nazareth, 1980; Lines and Treitel, 1984). It also tends to reduce oscillations or instabilities.

At the end of iterations, we can estimate how good the linearized inversion performed, given its assumptions. The a priori covariance of the data $[\ddot{W}_1]^{-1}$ in $\mathbf{U}_1$ induces noise in the parameter space **V**, i.e. in the resulting vector $\vec{m}$ in **V**, that is characterized by the final covariance matrix of the parameter noise V_m. The a priori parameter covariance matrix being $[\ddot{W}_2]^{-1}$, we have (Sandell and Shapiro, 1976)

$$V_m = \left([\ddot{G}_1^T][\ddot{W}_1][\ddot{G}_1] + [\ddot{W}_2] + \gamma^2 [\ddot{D}] \right)^{-1} , \tag{4.19}$$

in which we have added the Levenberg-Marquardt term. Investigation of (4.14) allows

an extension of (4.19) in the case where K multi-models are considered,

$$V_m = \left[\sum_{q=1}^{K} [\overset{\leftrightarrow}{G}_q^T][\overset{\leftrightarrow}{W}_q][\overset{\leftrightarrow}{G}_q] + \gamma^2[\overset{\leftrightarrow}{D}] \right]^{-1} . \qquad (4.19a)$$

The generalization of (4.19a) follows from (F.24), where we express the metric tensor in **V** space, $\overset{\leftrightarrow}{W}_m$, from that in **U** space, $\overset{\leftrightarrow}{W}$. From (F.24), (4.5) and the supplementarity of the $\mathbf{U}_q$ spaces, we have

$$\overset{\leftrightarrow}{W}_m = \overset{\leftrightarrow}{G}^T \bullet \overset{\leftrightarrow}{W} \bullet \overset{\leftrightarrow}{G} + \gamma^2 \overset{\leftrightarrow}{D} = \sum_{q=1}^{K} \overset{\leftrightarrow}{G}_q^T \bullet \overset{\leftrightarrow}{W}_q \bullet \overset{\leftrightarrow}{G}_q + \gamma^2 \overset{\leftrightarrow}{D} , \qquad (4.20)$$

with $[\overset{\leftrightarrow}{W}_m]^{-1} = V_m$ in (4.19a).

In fact, adding prior information to the parameters (model 2), renders the inverse operator non-singular since the matrix $[\overset{\leftrightarrow}{W}_2]$ is positive definite. Prior information effects are similar to the addition of the Levenberg-Marquardt term. Thus, in this case, the Levenberg-Marquardt term can be eliminated from the operator; its presence being unnecessary because the inverse operator is regular.

Additional effects due to presence of noise in the non-linear problem such as bias in the stochastic inverse ("noisy" operator) are assumed to be small compared with the parameter noise (Box, 1971).

The $M \times M$ *parameter resolution* matrix and the $N \times N$ *data resolution* matrix are defined by

$$R \equiv -[\overset{\leftrightarrow}{L} \bullet \overset{\leftrightarrow}{G}_1] \qquad \text{and} \qquad S \equiv -[\overset{\leftrightarrow}{G}_1 \bullet \overset{\leftrightarrow}{L}] . \qquad (4.21)$$

These two matrices are derived following (4.12) and (4.18) (Hohmann, 1979)

$$[\Delta\hat{m}] = R\ [\Delta m] , \qquad (4.22)$$

$$[\hat{\varepsilon}] = S\ [\varepsilon] , \qquad (4.23)$$

where $\hat{\varepsilon} \equiv -\ddot{G} \bullet \Delta\hat{m}$.

We shall require a measure of the resolution which is the trace of the matrix (Aki and Richards, 1980), i.e.

$$tr(R) = \sum_{i=1}^{M} R_{ii} \,. \tag{4.24}$$

For example if I is the $M \times M$ identity matrix then $tr(\mathbf{I}) = M$.

The closer the parameter resolution matrix is to the $M \times M$ identity matrix, the smaller the $\mathbf{V}_0$ space, i.e. the source of non-uniqueness in determining the parameters from the data ($\ddot{G} \bullet \mathbf{V}_0 = 0$). We can measure this similarity using (4.24), that is compare $tr(R)$ to M. However, if the nonlinear inversion converged in k iterations, the parameter resolution matrix contains information on how good the parameters improved from the $k-1$ iteration. In some instances, we may have a near-singular inverse operator in the neighborhood of the minimum, which would require a $\gamma^2 \neq 0$. It is shown (Aki and Richards, 1980) that the effect of γ^2 reduces the parameter noise but sacrifices the parameter resolution.

The data resolution matrix shows how good, at the end of the iterations, the data is reproduced by the model. Since, in the linearization, the difference data-synthetics (residuals) is performed at each iteration, the data resolution contains global information on the nonlinear inversion. The more the data resolution matrix resembles the $N \times N$ identity matrix, the smaller the $\mathbf{U}_0$ space, i.e. the source of discrepancy between data and synthetics. This can be estimated comparing $tr(S)$ to N. A scalar criteria is the residual energy reduction (RER) at each iteration, considering the initial residual energy as reference. Specifically, if $E_0 = \vec{\varepsilon}(\vec{m}_0) \bullet \ddot{W} \bullet \vec{\varepsilon}(\vec{m}_0) \,/\, 2$ is the initial residual energy, then the residual energy reduction, at a given iteration, is expressed in percent and is equal to RER = 100 $(1 - E/E_0)$. This is commonly called the data variance reduction.

Figure 2 shows the flow chart inverse procedure in its complete form. It is a simple extension of that derived for the 1-Dimensional case.

5. DATA, PARAMETER AND MODEL DESCRIPTION

The original data consist of a set of N_S seismograms denoted $\{ s_l(n) ; n = 1, N_P; l = 1, N_S \}$, where N_P is the total number of time samples in a seismogram. If N_R K-component geophones are simultaneously implemented in the inversion then $N_S = K \times N_R$ (K=2 or 3 generally). Assume the medium to be constituted of near-elastic homogeneous structures separated by smooth interfaces. We will invert for medium parameters, that is, layer velocities, densities and quality factors, and interface parameters. The seismogram model is computed via the paraxial ray method, generating synthetics that will be compared to the data.

The minimization (2.4) is invariant under a Fourier transform (see 2.9). The inversion is done in the frequency domain, and for N_F discrete frequencies situated in the neighborhood of the central frequency of the source wavelet. The justification for treating the problem in the frequency domain rather than in the time domain is due to (1) the modeling results are already in the frequency domain; (2) the full-wave character of the inversion that would require in the time domain N_P points, whereas in the frequency domain the total number is $N_F \ll N_P$, which reduces significantly the computation time. This procedure filters automatically noise with spectrum outside the source's spectrum; (3) Fourier transform of random time samples yields frequency samples that are practically uncorrelated to each other, which goes in favor of the assumption of diagonal covariance matrix for the data samples (end of section 2).

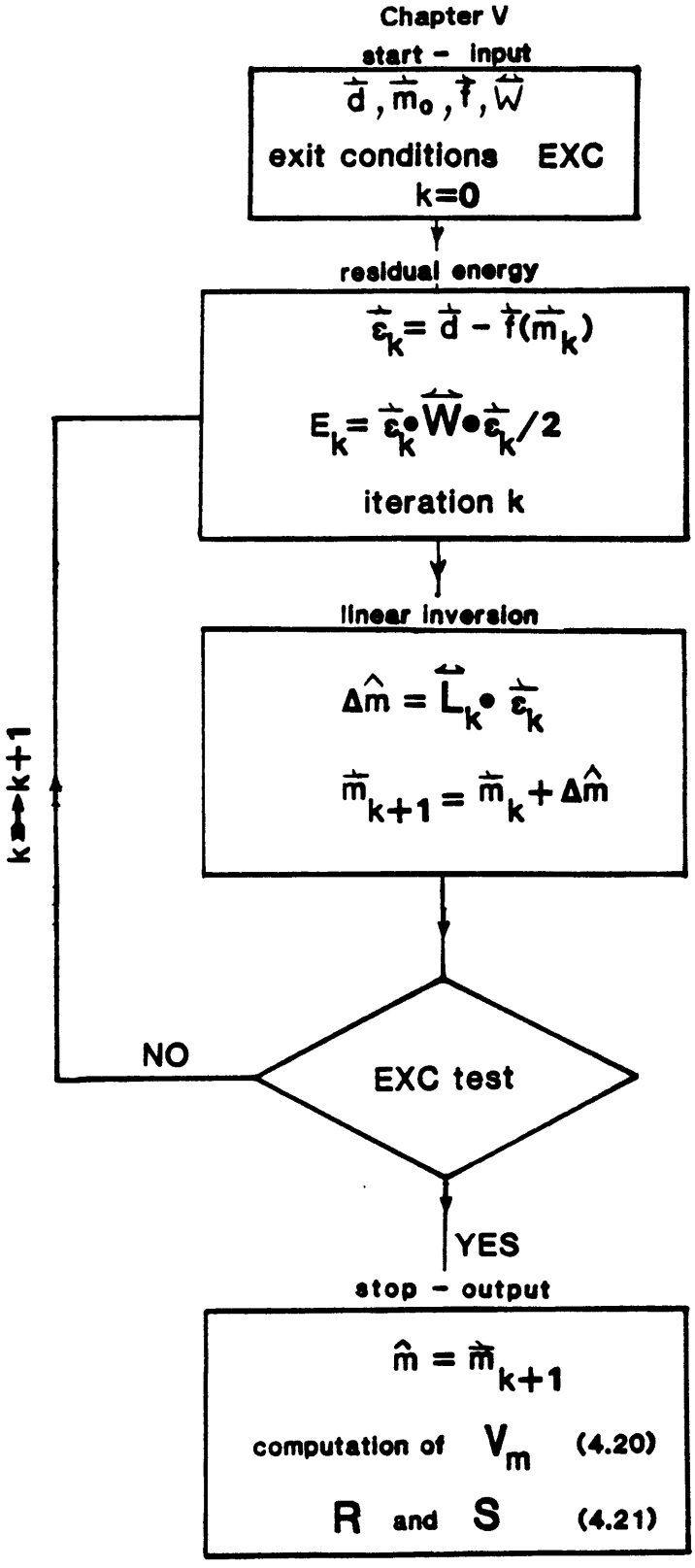


Figure 2. Flow chart of the multidimensional iterative least square inversion.

Define the DFT of $s_l(n)$ as $\bar{s}_l(j)$, j=1, N_P

such that

$$\bar{s}_l(j) = \sum_{j=1}^{N_P} s_l(n) \exp\left[i \left(\frac{2\pi}{N_P}\right) n\, j\right], \tag{5.1}$$

and for Inverse DFT

$$s_l(n) = \frac{1}{N_P} \sum_{j=1}^{N_P} \bar{s}_l(j) \exp\left[-i \left(\frac{2\pi}{N_P}\right) n\, j\right]. \tag{5.2}$$

For a time sampling of Δt, if the time interval of interest lies within $[t_{min}\,;\, t_{max}]$, we have $N_P = \text{Int}\{(t_{max} - t_{min})/\Delta t\} + 1$, and sample n in $s_l(n)$ corresponds to time $t_{min} + (n-1)\Delta t$; The function Int(x) is the integer value of x. The frequency sampling is then $\Delta f = 1/(N_P\, \Delta t)$, the frequency interval $[0;\, (N_P - 1)\Delta f\,]$, and sample j in $\bar{s}_l(j)$ corresponds to frequency $(j-1)\Delta f$. If the source highest frequency is f_{max}, it is assumed that we have achieved 2 Δt $f_{max} < 1$ (Nyquist condition).

If the source frequency spectrum is centered around f_0, the N_F frequency samples are chosen so that they sample adequately the source spectrum and the interval $[f_1\,;\, f_2]$ covers appropriately the spectrum support. This corresponds to a decimation in frequency of the spectrum. Reversing time and frequency domains, if the source time pulse width is t_0, the equivalent Nyquist condition imposes an upper bound on df, and is equal to 2 df $t_0 < 1$. The frequency sampling interval df, must satisfy this condition. Given f_1 and f_2 from the source spectrum, we set $j_1 = \text{Int}(f_1/\Delta f)+1$, $j_2 = \text{Int}(f_2/\Delta f)+1$, and $dj = \text{Int}(df/\Delta f)+1$. The samples are $j_k = j_1 + (k-1)dj$, k=1, N_F, with $N_F = \text{Int}\{(j_2 - j_1 + 1)/dj\}+1$. The last frequency sample is $j_{N_F} = j_2 + 1$, or at frequency $j_2\,\Delta f$, which is at least f_2 and at most $f_2 + \Delta f$.

We now have reduced the original data set $\{s_l(n);\, n=1,\, N_P;\, l=1,\, N_S\}$, to $\{\bar{s}_l(j_1+(j-1)dj);\, j=1,\, N_F;\, l=1,\, N_S\}$, with $N_F < N_P$. Decomposing $\bar{s}_l(j)$ into a real and imaginary part yields

$$\bar{s}_l(j) \equiv a_l^j + i\, b_l^j, \tag{5.3}$$

where $a_l^j = \mathrm{Re}\{\bar{s}_l(j)\}$ and $b_l^j = \mathrm{Im}\{\bar{s}_l(j)\}$. The data vector $\vec{d}_1$ introduced previously can now be explicitly defined in terms of its components. Recall that $\vec{d} = \vec{d}_1 \oplus \vec{d}_2 \oplus \vec{d}_3$, with $\vec{d}_2 = \vec{m}_0$ being an input and $\vec{d}_3 = 0$. Consider the two basis, the first in $\mathbf{U}_1$ space, with dimension $N_1 = 2{\times}N_F{\times}N_S$, and the second in $\mathbf{U}_2$ space, with dimension $N_2 = M$. $\mathbf{U}_3$ space is of dimension N_3.

We have $\vec{d} = \sum_{i=1}^{N_1} d^i \vec{e}_{1_i} + \sum_{i=\alpha}^{M} d^{N_1+i} \vec{e}_{2_i}$, where the d^i's are ordered the following way:

$$[\vec{d}] = \begin{cases} d^1 = a_1^{j_1} \\ d^2 = a_1^{j_1+1} \\ \vdots \\ d^{N_F} = a_1^{j_1+(N_F-1)dj} \\ d^{N_F+1} = a_2^{j_1} \\ \vdots \\ d^{N_F N_S} = a_{N_S}^{j_1+(N_F-1)dj} \\ d^{N_F N_S+1} = b_1^{j_1} \\ \vdots \\ d^{2N_F+N_S} = b_1^{j_1+(N_F-1)dj} \\ d^{2N_F+N_S+1} = b_2^{j_1} \\ \vdots \\ d^{2N_F N_S} = b_{N_S}^{j_1+(N_F-1)dj} \\ d^{N_1+1} = m_0^1 \\ \vdots \\ d^{N_1+M} = m_0^M \\ d^{N_1+M+1} = 0 \\ \vdots \\ d^{N_1+M+N_3} = 0 \end{cases}$$

Chapter V

We now proceed to the description of the parameters that are to be extracted from the data. Let us define a general cartesian reference frame (x,y,z) for the medium as having the z axis oriented downwards and $(\hat{e}_x, \hat{e}_y, \hat{e}_z)$ forming a right handed system. Figure 1 of chapter IV describes this frame. The reference $z=0$ is at the earth surface. Let the medium have I interfaces $\{z_i(x), i=1, I\}$, where each interface, i, is described analytically with I_i parameters $(p_i^j, j=1, I_i)$, by the equation $z_i = z_i(x,y, p_i^1, p_i^2, \ldots, p_i^{I_i})$. An interface parameter p_i^j corresponds to the j-th parameter of interface i. Interfaces must be defined for all (x,y) in the medium, and two interfaces are not allowed to cross each other. The functions z_i must be of class C^2 (Cerveny and Hron, 1980). Each function z_i will be approximated in the program by a cubic spline interpolation scheme which is of class C^2 (De Boor, 1980). Depending on the variations of z_i, a sufficient number of points will be required in the interpolation to reproduce reliably z_i. Two types of interface are described in Appendix G. A one-parameter tilted flat interface (figure 3a), and a five-parameter reef-type interface (figure 3b). The total number of interface parameters is then $M_T = \sum_{i=1}^{I} I_i$. Rules for interfaces and layers are the same as those presented for 2D media in Psencik (1983).

There are $I-1$ homogeneous layers. Each layer i is characterized by (1) the compressional velocity α_i, (2) the shear velocity β_i, (3) the density ρ_i, and (4) the quality factor Q_i (for near-elastic solids, i.e. $Q_i \gg 1$). The total number of layer parameters is $M_L = 4(I-1)$. The parameter vector $\vec{m}$ introduced previously is explicitly defined in terms of its components.

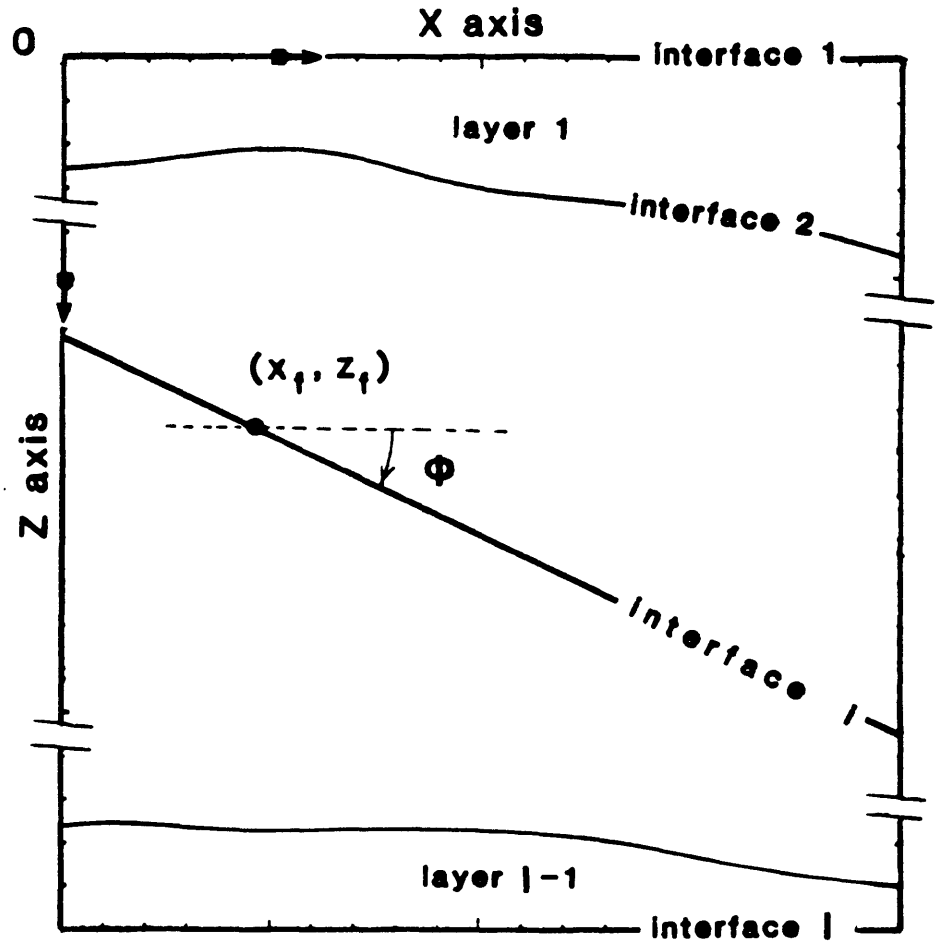


Figure 3a. Geometry and parameter of an analytical flat tilted interface.

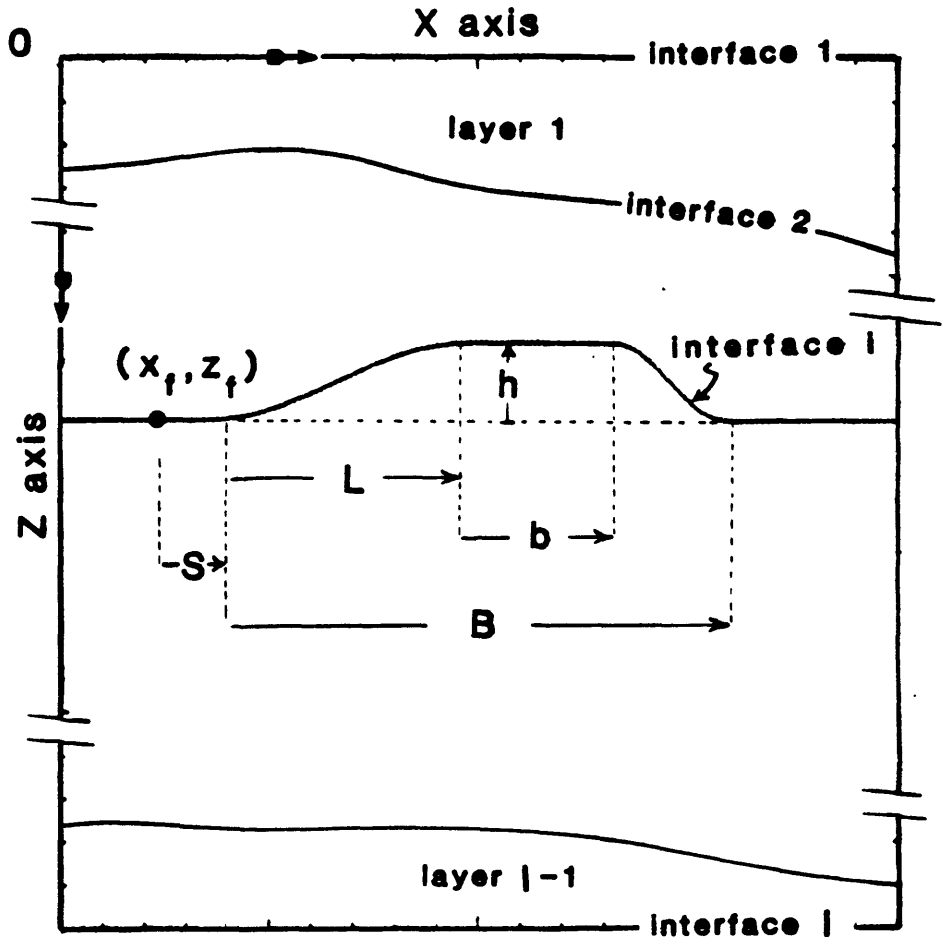


Figure 3b. Geometry and parameters of an analytical reef type interface.

Considering a basis in V space, $\{\vec{q}_i\}$ with dimension $M = M_T + M_L$, we have $\vec{m} = \sum_{i=1}^{M} m^i \vec{q}_i$, with

$$[\vec{m}] = \begin{cases} m^1 = p_1^1 \\ m^2 = p_1^2 \\ \cdot \quad \cdot \quad \cdot \\ m^{Q_1} = p_1^{Q_1} \\ m^{Q_1+1} = p_2^1 \\ \cdot \quad \cdot \quad \cdot \\ m^{M_T} = p_B^{Q_B} \\ m^{M_T+1} = \alpha_1 \\ m^{M_T+2} = \beta_1 \\ m^{M_T+3} = \rho_1 \\ m^{M_T+4} = Q_1 \\ \cdot \quad \cdot \quad \cdot \\ m^{M_T+M_L-3} = \alpha_{M_L} \\ m^{M_T+M_L-2} = \beta_{M_L} \\ m^{M_T+M_L-1} = \rho_{M_L} \\ m^{M_T+M_L} = Q_{M_L} \end{cases}$$

The scaling under which the parameters is represented is of primary importance in the inversion. For example, one may express velocities, v, in terms of slownesses, v^{-1},

or angles Φ, in terms of their tangent $\tan\Phi$, or also taking a logarithmic type of parametrization (Madden and Aki, 1980). The appropriate re-scaling (linear or non-linear) would reduce the number of iterations, and increase the efficiency of the algorithm. The sensitivity analysis would indicate if this the parameter re-scaling is a requirement. The covariance matrix (4.19a) would indicate if the parameters are independent of each other, for the model considered.

Model 1 implements the paraxial ray method producing synthetics. For a given parameter vector $\vec{m}$, this model generates $\vec{f}_1(\vec{m}, \omega)$ that is in the exact same form as $\vec{d}_1$ ($\omega = 2\pi f$). The paraxial results are already in the frequency domain (Chapter IV). The asymptotic Green's function (IV.2.14), $\vec{U}(M,\omega)$, derived previously is multiplied in the frequency domain by the source spectrum $\overline{\sigma}(\omega)$. The source spectrum must be the same as the one in the data $\vec{d}_1$. The displacement vector is then decomposed into components, that must be of same type as to that in the original data. For example, if the data is pressure measurements, the model will generate pressure synthetics. Returning to the same notation as in (5.3), where receiver at M is represented by l, l=1, N_S, and the frequency ω replaced by $j = j_1 + (k-1)dj, k = 1, N_F$, we write

$$\overline{u}_l(j) \equiv \hat{a}_l^j + i\, \hat{b}_l^j , \tag{5.4}$$

where $\hat{a}_l^j = \mathrm{Re}\{U(l,j)\,\overline{\sigma}(j)\}$ and $\hat{b}_l^j = \mathrm{Im}\{U(l,j)\,\overline{\sigma}(j)\}$, with $U(l,j)$ being the component of $\vec{U}(M,\omega)$. The vector $\vec{f}_1$ is then constructed similarly to $\vec{d}_1$, replacing the a's and b's in $\vec{d}_1$, by $\hat{a}$'s and $\hat{b}$'s , respectively. The components of $\vec{d}_1$ are those of $\vec{d}$ from the 1-st component to the $2\times N_F \times N_S$-th.

Model 2 implements the parameter prior information model that simply reproduces the parameter values ($\vec{f}_2 = \vec{m}$). The storage of $\vec{f}_2$ is identical to that in $\vec{d}$. The

parameter component m^i, $i=1, M$, is stored at $f^{N_1+i} = m^i$.

Model 3 implements the penalty model with synthetics (3.6). The storage of $\vec{f}_3$ is below that of $\vec{f}_2$. Element f_3^i, $i=1, N_3$ is stored at $f^{N_1+M+i} = f_3^i$.

The vector $\vec{f}$ is then constructed in a similar way as $\vec{d}$, so that corresponding components are of same type.

Before the inversion, we assume that the a priori covariance matrix of Model 1 is equal to a constant times the unit matrix, and that the a priori covariance matrix of Model 2 is diagonal (independent parameters), i.e.

$$\overset{\leftrightarrow}{W}_1 = w_1 \overset{\leftrightarrow}{I} , \tag{5.5}$$

and

$$(\overset{\leftrightarrow}{W}_2)_{ij} = 0 \quad \text{for } i \neq j .$$

Model 3, also, has a diagonal a priori covariance matrix. We then use as input a normalized residual energy from (3.8)

$$E^N(\vec{m}) \equiv \frac{E(\vec{m})}{w_1} = \frac{1}{2} \sum_{i=1}^{N_1} \left[d^i - f_1^i(\vec{m}) \right]^2 +$$

$$+ \sum_{i=1}^{M} \frac{(W_2)_{ii}}{2w_1} \left[m_0^i - m^i \right]^2 + \sum_{i=1}^{N_3} \frac{(W_3)_{ii}}{2w_1} \left[f_3^i(\vec{m}) \right]^2 . \tag{5.6}$$

This converts the minimization of E into a minimization of a normalized E^N, which is the residual energy, E, "per unit" a priori variance (of Model 1). Set $\sigma^2 = w_1^{-1}$ to be the variance of Model 1.

At the end of the iterations we have an estimated parameter vector $\hat{m}$. σ^2 is estimated from from the residual sum of squares, assuming that the errors are additive,

uncorrelated, and normal and have zero mean and constant variance. This yields (Beck and Arnold, 1977)

$$\hat{\sigma}^2 = \frac{E^N(\hat{m})}{(N-M)}, \tag{5.7}$$

where $\hat{\sigma}$ is an estimate of σ, and $(N-M)>0$ is the degrees of freedom associated to the sum of squares.

The a posteriori parameter covariance matrix follows then from (4.19-20)

$$V_m = \hat{\sigma}^2 \left[\ddot{G}^T \bullet \ddot{W}^N \bullet \ddot{G} + \gamma^2 \ddot{D}^N \right]^{-1}, \tag{5.8}$$

where $\ddot{W}^N$ and $\ddot{D}^N$ are the normalized $\ddot{W}$ and $\ddot{D}$, i.e. each divided by $w_1 = \sigma^{-2}$. Denoting $\vec{m}^*$ the optimal parameter vector (not known generally), and $\hat{m}$ the estimated parameter vector, the expression for the estimated standard deviation (or errors) in the parameters, at the end of the inversion, is

$$m^{*\,i} \approx \hat{m}^i \pm [(V_m)_{ii}]^{1/2}. \tag{5.9}$$

Cross terms in V_m measure the correlation between different parameters. The correlation coefficient between two distinct parameters, m^i and m^j, is calculated via the equation

$$r(\hat{m}^i, \hat{m}^j) = \frac{(V_m)_{ij}}{\left[(V_m)_{ii}\,(V_m)_{jj}\right]^{1/2}}. \tag{5.10}$$

Due to the large dimensions of the data resolution matrix (4.21) (more than 600 × 600), we prefer to plot the computed seismograms from the estimated parameters, and compare that to the original data set.

Noise can be added to test the robustness of the inversion. We have a uniform $(-a,a)$ random distribution noise $e(n)$, characterized by its maximum amplitude a. The

noise is controlled by choosing a reference trace $s_{SP}(n)$, $n=1, N_P$, and computing the integrated power signal to noise ratio. This is achieved by taking the DFT of both $s_{SP}(n)$ and $e(n)$ yielding $\bar{s}_{SP}(j)$ and $\bar{e}(j)$,$j=1,N_P$. We then compute

$$SNR = \frac{\sum_{j=1}^{N_F} |\bar{s}_{SP}(j_1 + (j-1)dj)|^2}{\sum_{j=1}^{N_F} |\bar{e}(j_1 + (j-1)dj)|^2}, \qquad (5.11)$$

which is the power signal to noise ratio. The value 10 $\log_{10}$ (SNR) is the power signal to noise ratio expressed in decibels (dB). We apply a bandpass filter to the noise, with low cut-off frequency sample j_1, and high cut-off frequency sample j_2. This noise is added to traces $s_l(n)$, $l=1, N_S$.

V.6. EXAMPLES

The source time function is the Gabor wavelet described in Chapter III, equation (6.1). Computation of (4.18) is done using the program LMDIF, which is a MINPACK-1 (1980) Fortran subroutine that minimizes the sum of the squares of N nonlinear functions in M variables. The Jacobian matrix, $[\overset{\leftrightarrow}{G}]$, is calculated by a forward-difference approximation. LMDIF is a version of the Levenberg-Marquardt algorithm where the Gauss-Newton method is used as long as the inverse operator (4.13) is non-singular. When this is not the case it computes a Levenberg-Marquardt parameter γ^2 and uses (4.17) as inverse operator.

Model A:

The model consists of two acoustic layers and a dipping interface. Source parameters are: f_0 = 60 Hz, Γ = 4, t_i = 0.03 sec. Figure 4a displays the time signal, and figure 4b its amplitude spectrum. The line source and the receivers are shown in figure 5. There is a total of N_S = 38 receivers, and N_F = 9 frequency samples from 10 Hz to 110 Hz. The number of residuals is N = 689 (where N_2 = 3 residuals are due to Model 2, and N_3 = 2 to Model 3). Three parameters are defined and constitute the parameter vector with components $[\vec{m}]^T = (\Phi\ ,\ \alpha_1\ ,\ \alpha_2)$. Φ is interface 2 dip angle (see G.1) with units in *radians* , α_1 is the upper medium velocity, and α_2 is the lower medium velocity, with units in *km/sec.* The optimal parameter vector is chosen to be to: $[\vec{m}^*]^T$ = (0.197 , 3.5 , 4.5). The optimal interface dip angle is about 11°. The fixed point on the interface has coordinates (x_f = 0.05 *km*, z_f = 0.51 *km*).

The sensitivity analysis consists in studying the variations of the residual energy around an a priori optimal point. Given the model $\vec{f}$ and a parameter vector, $\vec{m}^*$, we set $\vec{d} \equiv \vec{f}(\vec{m}^*)$, and plot E given in (2.2), varying $\vec{m}$. The sensitivity of the model for

specific parameters can then be investigated. A 3D plot of $E = (\alpha_1 ; \alpha_2)$ is shown in figure 6a, constructed with 21x21 points. The sampling interval in each direction is $\Delta\alpha = 0.05 km/sec$. The tilt angle Φ is optimal. A contour plot of the same section is displayed in figure 6b. The trough in the upper layer velocity direction shows that this velocity is much better constrained than that in the lower medium. The global minima is clearly seen in the contour plot. The presence of local minima will be discussed later in this section.

A 3D plot of $E = (\Phi ; \alpha_2)$, with 21x21 points and $\Delta\Phi \approx 0.3^{\circ}$, $\Delta\alpha_2 = 0.05 km/sec$, is shown if figure 7a. Figure 7b is its contour plot. The upper layer velocity α_1 is optimal. The sought resolution on dip angle is a few degrees, and on velocities on the order of 100 m/sec. On this basis, we can compare the sensitivity of the residual energy for these different parameters. The lower medium velocity is better constrained than the interface dip angle. A nice minimum can be seen.

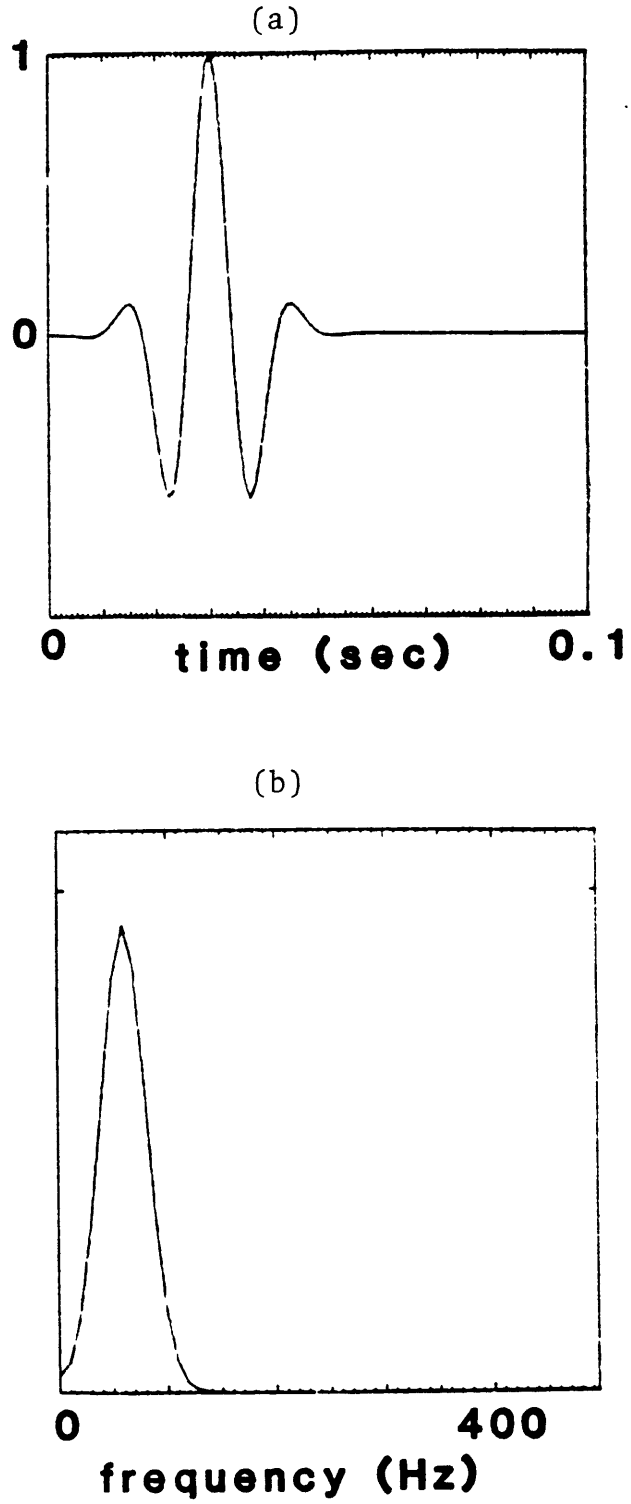


Figure 4. Source signal used in Model A: (a) in time; (b) amplitude spectrum.

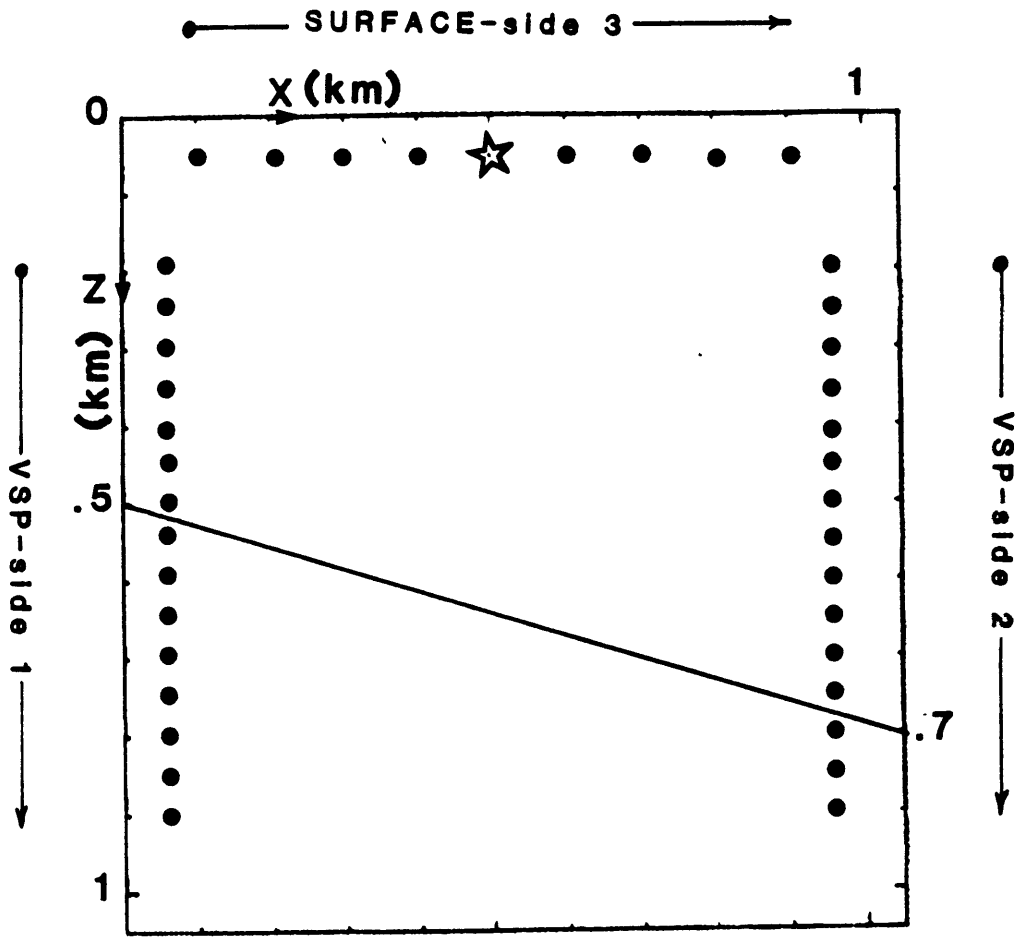


Figure 5. Geometry of Model A.

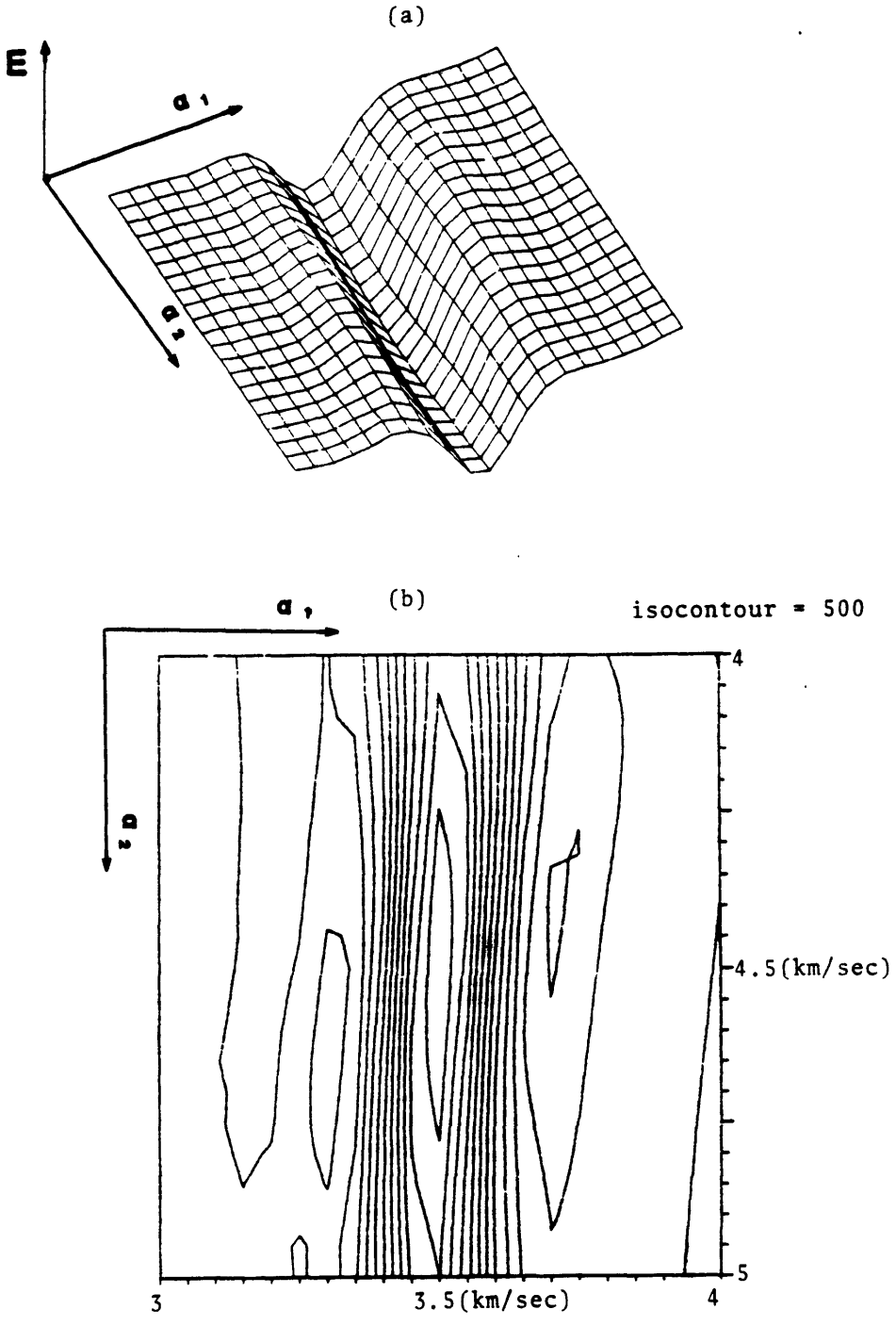


Figure 6. Model A sensitivity as a function of layer velocities: $E = (\alpha_1 ; \alpha_2)$, (a) 3D plot, (b) contour plot.

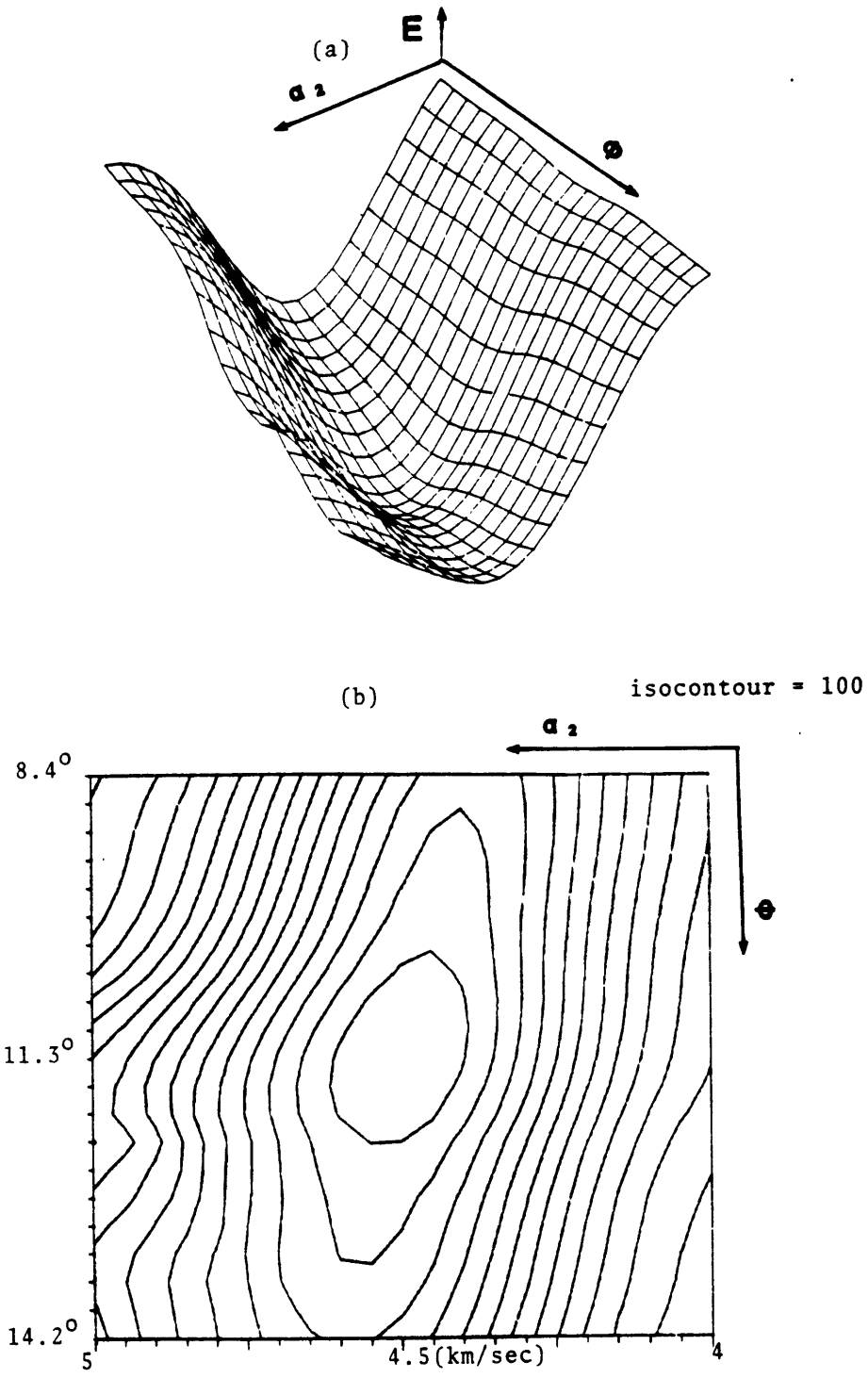


Figure 7. Model A sensitivity as a function of interface slope and lower layer velocity: $E = (\Phi\ ;\ \alpha_2)$, (a) 3D plot, (b) contour plot.

Figure 8a is the 3D plot of $E = (\Phi\ ;\ \alpha_1)$, with 21x21 points, and $\Delta\Phi \approx 0.3^o$, $\Delta\alpha_1 = 0.05 km/sec$. The lower layer velocity α_2 is optimal. The contour plot is displayed in figure 8b. The angle is poorly constrained, compared to the upper medium velocity. The large trough has a global minimum, but the two parameters have such contrasted effects on E that the minimum is poorly resolved in the angle direction. Figure 9a is the 1D plot of $E = (\Phi)$ and figure 9b the plot of $E = (\alpha_1)$. The first figure corresponds to the line passing in the trough of figure 8a, and the second figure is the line perpendicular to the trough. Both clearly shows the minimum along these two directions. Note the relative difference in residual energy between both curves.

Prior information about the parameters would constrain the minimization, and would remove some non-uniqueness in parameter determination. Setting the prior information to $\Phi^{ap} = 0.21$ (about 12^o) with error $(W_2)_{11}^{-1/2} = 0.0175$, and $\alpha_1 = 3.4$ with error $(W_2)_{22}^{-1/2}$, with weight in (5.6) w_1=0.1, yield a 3D sensitivity plot of $E = (\Phi\ ;\ \alpha_1)$, which is plotted in figure 10a. The sampling is the same as in figure 7a, and the lower layer velocity α_2 is optimal. Figure 10b shows the contour plot. Compared to figures 8a and 8b, there is a clear improvement in the determination of a global minimum. The residual energy has increased significantly away from the parameters a priori values. The global minimum is in between the prior values and the optimal values. Since the angle variance is small, the minimum in this direction is closer to 12^o than the optimal value at 11.3^o. Thus prior values must be specified precisely and weighted accordingly, since they force the optimization to search for the minimum in regions that are close to the a priori values.

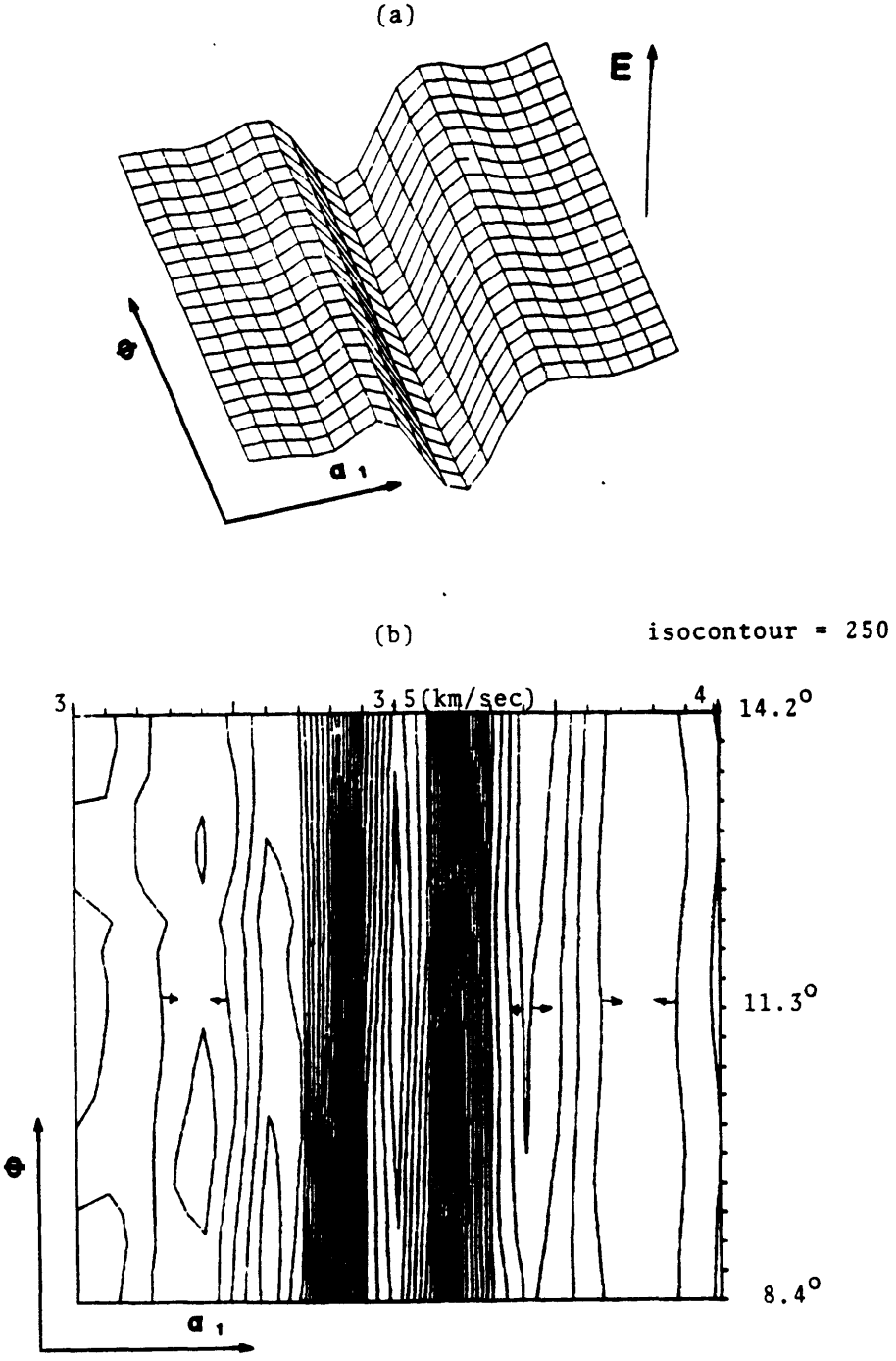


Figure 8. Model A sensitivity as a function of interface slope and upper layer velocity: $E = (\Phi\ ;\ \alpha_1)$, (a) 3D plot, (b) contour plot.

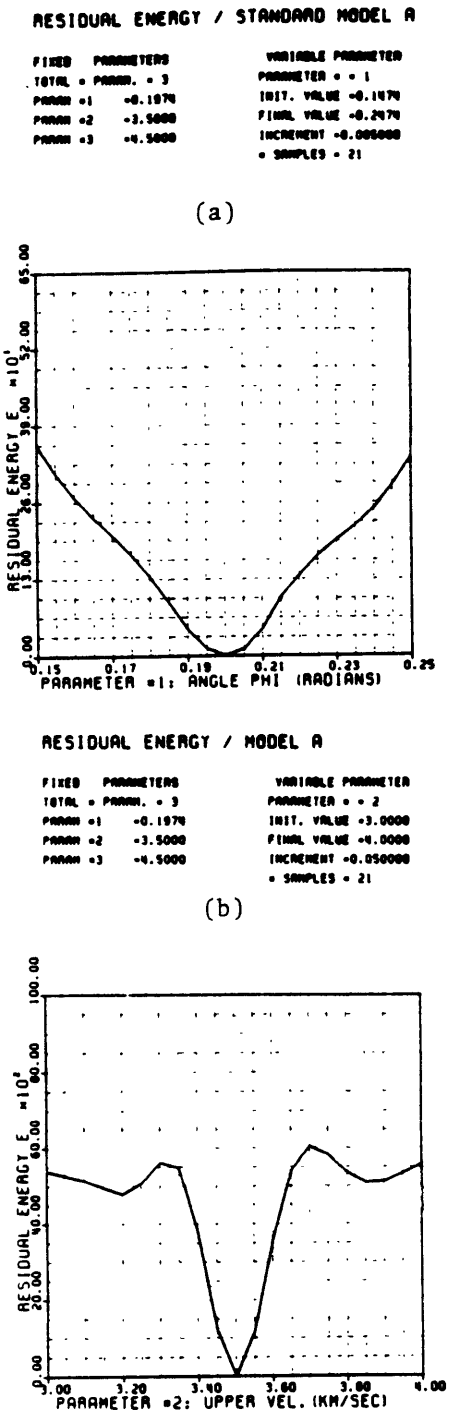


Figure 9. Model A sensitivity as a function of interface slope: (a) $E = (\Phi)$, and as a function of upper layer velocity (b) $E = (\alpha_1)$.

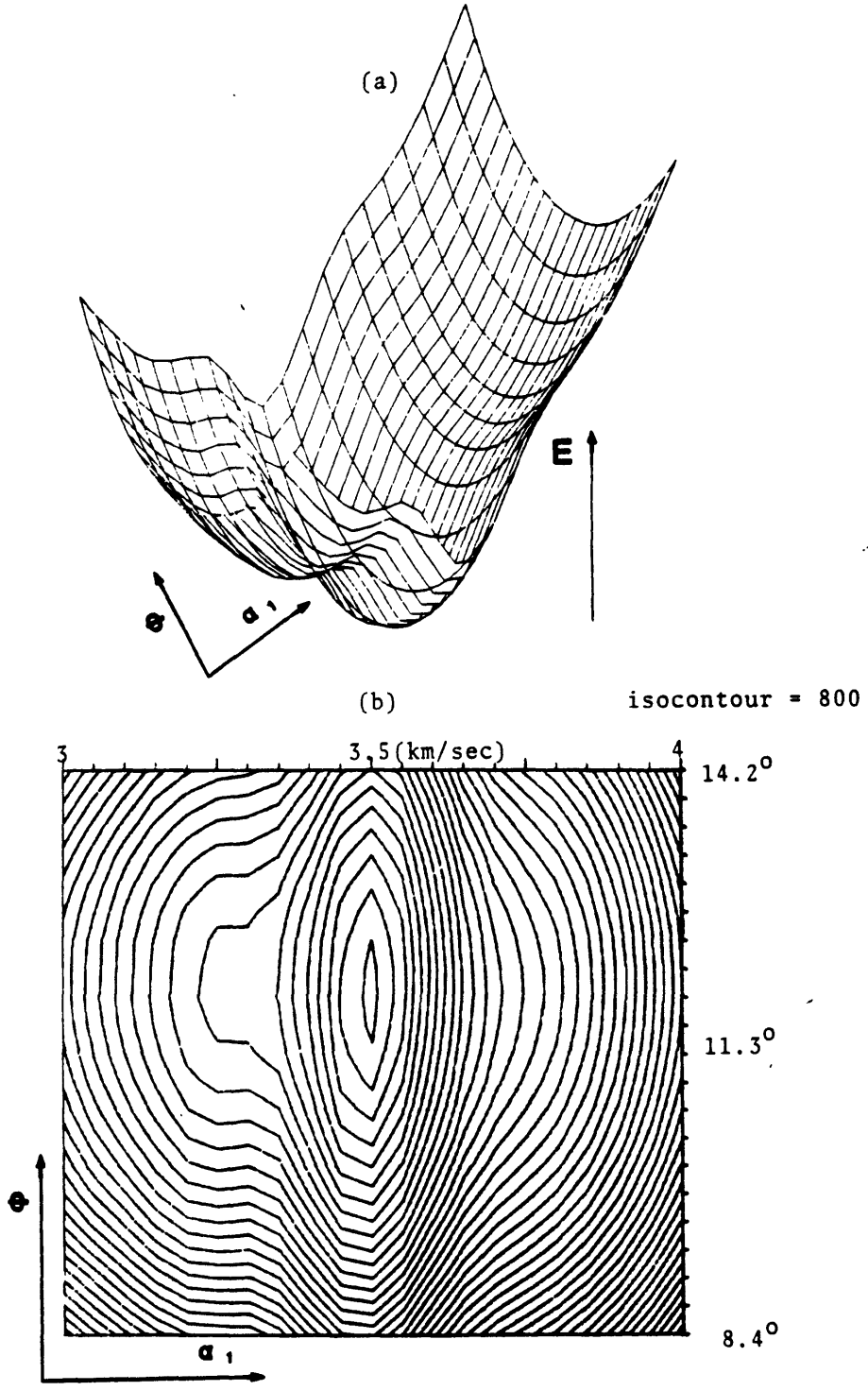


Figure 10. Prior information in Model A parameters; sensitivity as a function of interface slope and upper layer velocity: $E = (\Phi ; \alpha_1)$, (a) 3D plot, (b) contour plot.

Model A '

The model is the same as above, with a greater receiver density. Sensitivity analysis with 62 receivers is studied, with 1116 residuals. The 3D plot of $E = (\Phi\ ;\ \alpha_1)$ is sampled as figure 8a, and is shown in figure 11a, with its contour plot in figure 11b. Compared to figure 8a and 8b, the residual energy increased, without removing local minimas. The wavelength in the upper layer is about 60 meters. In model A, the VSP sampling was 50 meters and the surface sampling, 100 meters. Presently, the VSP sampling is 35 meters and the surface sampling, 50 meters. Increasing the receiver density has the beneficial effect of enhancing the global minimum , but this does not improve the uniqueness of the global minimum. Figures 12 are the equivalent display of figures 9, for Model A'. The energy curvature is increased reducing therefore the parameter variance.

Model A ''

The model is the same as Model A, except that a wider band source is used. Source parameters are $f_0 = 120$ Hz, $\Gamma = 3$ and $t_i = 0.01$ sec. Figure 13a shows the time signal, and 13b its amplitude spectrum. There are 12 frequency samples going from 50 to 300 Hz. Plots of $E = (\Phi\ ;\ \alpha_1)$ with same sampling as figures 8, are shown in figures 14. The lower layer velocity α_2 is optimal. To compare the sensitivity of Model A to that of Model A '', we have corrected the source energy of the latter model so that both models have same source energy. The source energy is dependent on the frequency samples used in the inversion. It is what the inversion ''sees'' as source, through the source spectrum sampling, that is of primary importance. Figures 15 are 1D plots of $E = (\Phi)$, and $E = (\alpha_1)$, respectively. They should be compared to figures 9.

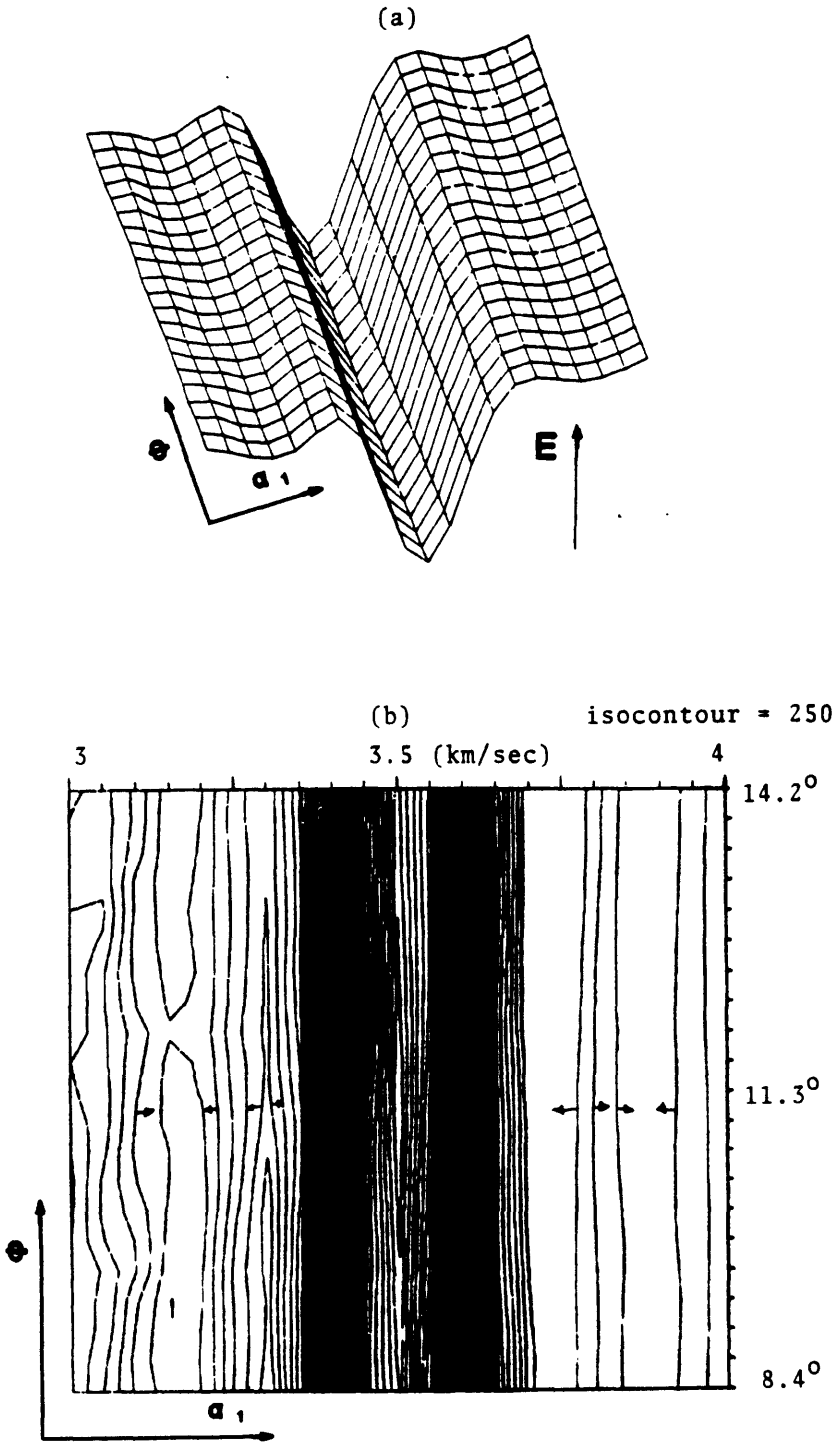


Figure 11. Model A' sensitivity as a function of interface slope and upper layer velocity: $E = (\Phi ; \alpha_1)$, (a) 3D plot, (b) contour plot.

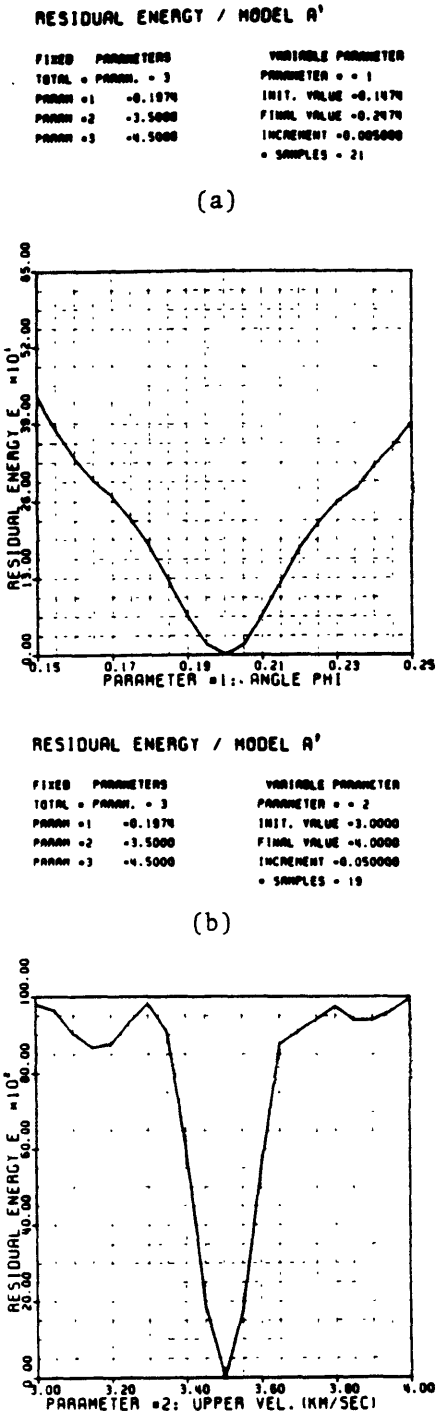


Figure 12. Model A' sensitivity as a function of interface slope: (a) $E = (\Phi)$, and as a function of upper layer velocity (b) $E = (\alpha_1)$.

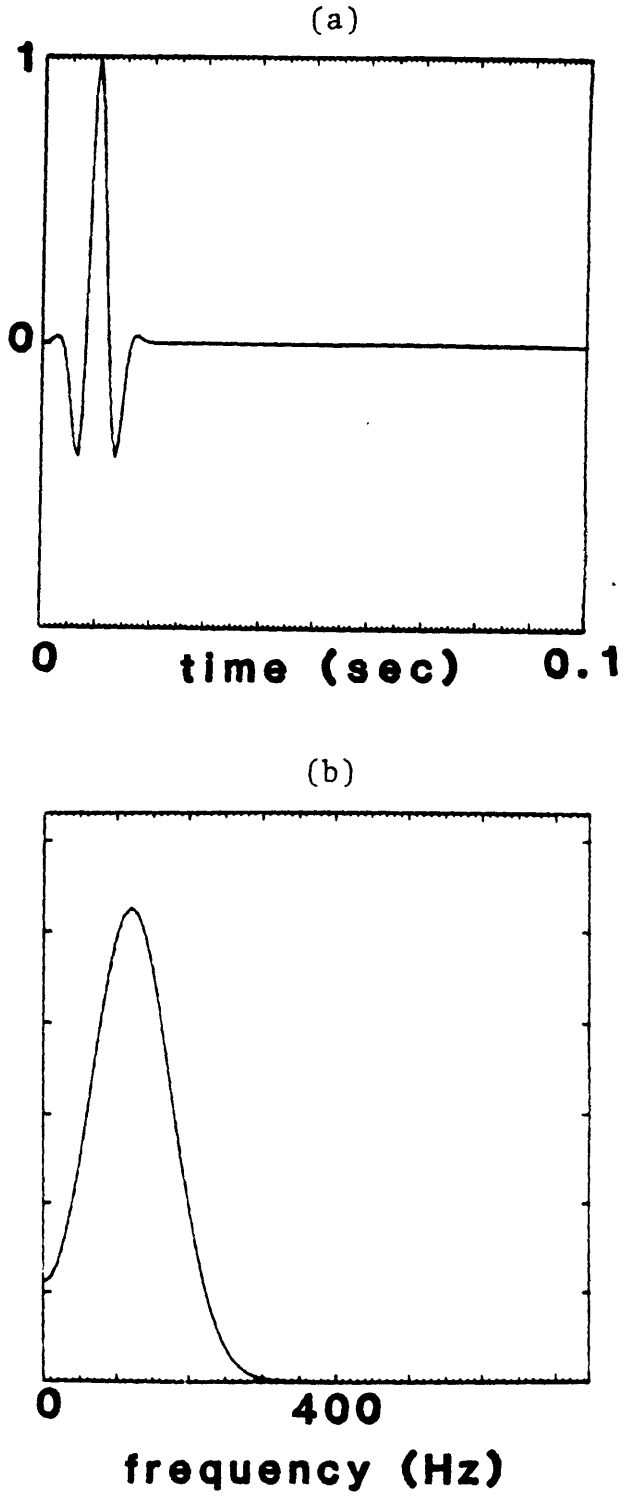


Figure 13. Wide band source signal used in Model A'': (a) in time; (b) amplitude spectrum.

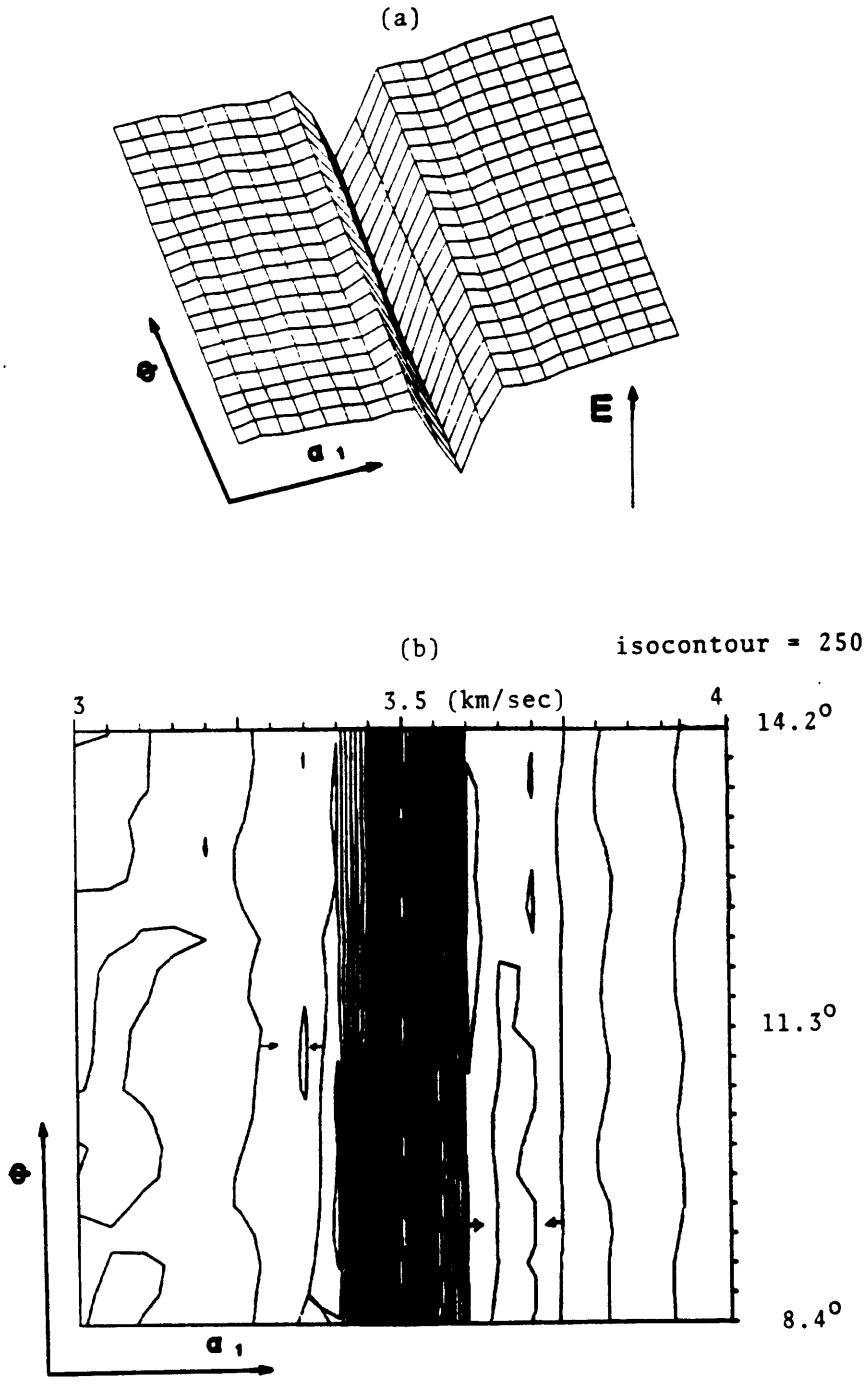


Figure 14. Model A'' sensitivity as a function of interface slope and upper layer velocity: $E = (\Phi ; \alpha_1)$, (a) 3D plot, (b) contour plot.

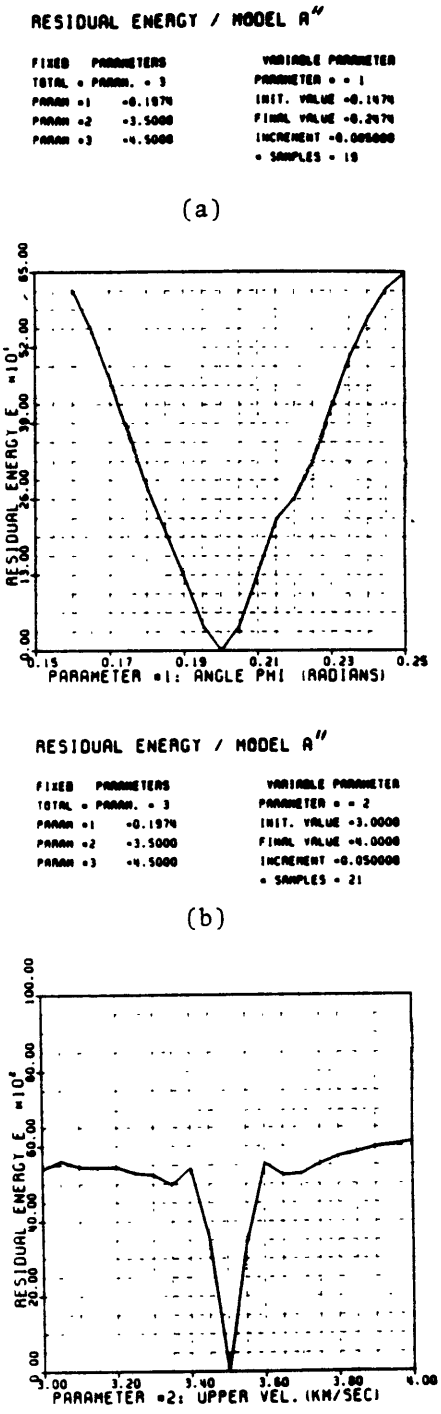


Figure 15. Model A'' sensitivity as a function of interface slope: (a) $E = (\Phi)$, and as a function of upper layer velocity (b) $E = (\alpha_1)$.

The parameter variance has diminished, but on the velocity axis, the region of global minimum has decreased. These effects are mainly due to the source structure and to the full waveform character of the inversion. Specifically, the presence in the source of "sinusoidal" shapes, is the reason of multiple local minima/maxima in residual energy. The velocity sensitivity quantifies changes in E for velocity variations away the optimal value. The primary effect of such a perturbation is a time shift of the wavelet. At a given receiver, a velocity increase is accompanied by a decrease in travel time that shifts the wavelet, with respect to the optimal wavelet, toward earlier times. Due to the fact that the source has zero crossings, the residual energy (of the two traces) is not monotonically increasing as we depart from the optimal trace. Rather, it has a shape that is very similar to that shown in figures 9b, 12b and 15b. The more the sinusoidal-like source is "compressed" (e.g. higher frequency content) the more the sensitivity curve has its local minima/maxima concentrated near the global minimum. This would require good initial estimates for velocities. However, the non-uniqueness could be reduced by transforming the data set, before inversion, to signals that are non-sinusoidal like. An example could be to replace data traces by amplitudes of the analytic signal of the traces (trace envelope). This would reduce the non-uniqueness, but would increase, unfortunately, the parameter variance. Ideally, we would start the inversion with the processed traces, for a given initial parameter guess. Then, we re-invert this time the original traces, with as initial parameters the results of the first inversion. Hopefully, we would get closer to the minimum with the first inversion, without requiring very good initial guesses. Then, improve the estimates with the second inversion, and reduce the parameter variance.

Inversion of Model A

The data $\vec{d}$ results from the finite difference synthetics of Model A (figure 16a). The initial parameter vector is chosen to be $[\vec{m}_0]^T$ = (0 , 3.3 , 4.1), with no constraints on the priori values (i.e. $(W_2)_{ii}$ = 0 in (5.6)). Paraxial synthetics for the a priori values, generated by the forward model, are displayed in figure 16b. Results from the inversion are displayed in Table 1. Parameter updates along with their corresponding residual energy reduction (RER) is presented at each iteration. The Levenberg-Marquardt parameter introduced in (4.17) is referred to as GAMMA**2. At the end of the inversion the following results are directly available: (1) estimated parameters along with their estimated standard deviation (5.9); (2) the parameter covariance matrix (5.8), and the corresponding correlation coefficients (5.10).

We shall refer to observed error, the error on the estimated parameter relative to the optimal parameter that is, here, given (i.e. $1 - \hat{m}^i / m^{*i}$). Estimated parameters are very close to optimal, $[\hat{m}]^T$ = (0.192 , 3.51 , 4.52). The least resolved parameter is the interface dip angle, estimated within observed errors of 3%. Values of estimated parameter standard deviations (see Table 1) must be interpreted as representing parameters errors relative to others. Absolute values between observed and estimated errors can be off by one order of magnitude. But, relatively, they follow same trend. The only noise present is numerical, since the data generated by finite difference method is inverted with the paraxial ray method. The covariance matrix has a fair diagonal trend, that is quantified by the correlation coefficients. Estimated parameters are so close to optimal that there is no distinguishable difference between synthetics in figure 16a and those generated by the model, with estimated parameters.

Adding noise to the finite difference data, with a SNR = 5 dB, results in synthetics that are shown in figure 17. The reference data trace is $s_{SP}(n) \equiv s_1(n)$. Inversion of

this set yields results that are listed in Table 2. Estimated parameters are still very good, $[\hat{m}]^T$ = (0.194 , 3.51 , 4.52). Note that the estimated data variance ,$\hat{\sigma}^2$, is about 52 times that of the previous inversion (Table 1). Absolute values between observed and estimated errors are closer than when no noise is added.

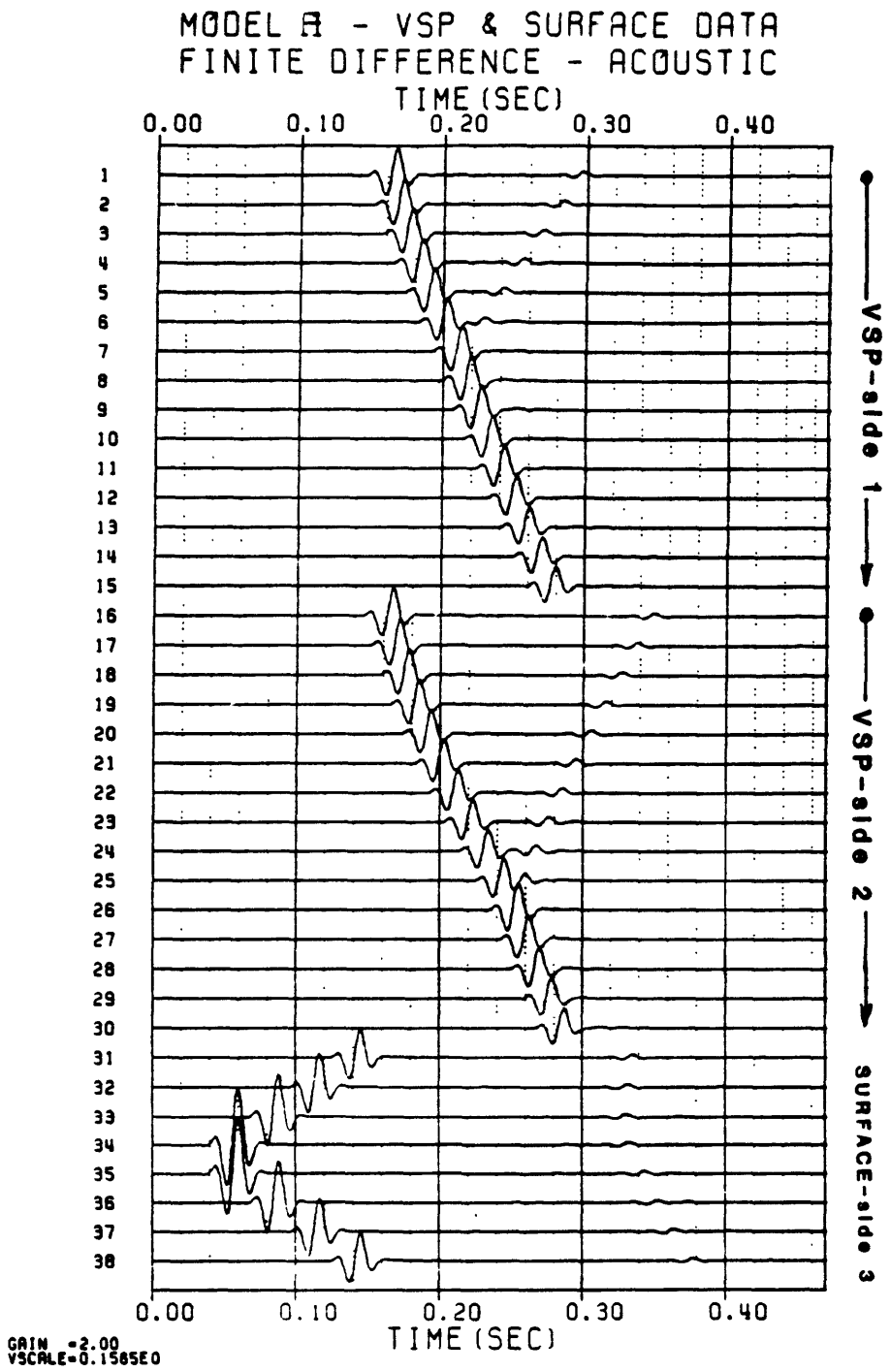


Figure 16a. Model A synthetic data, generated by the finite difference method.

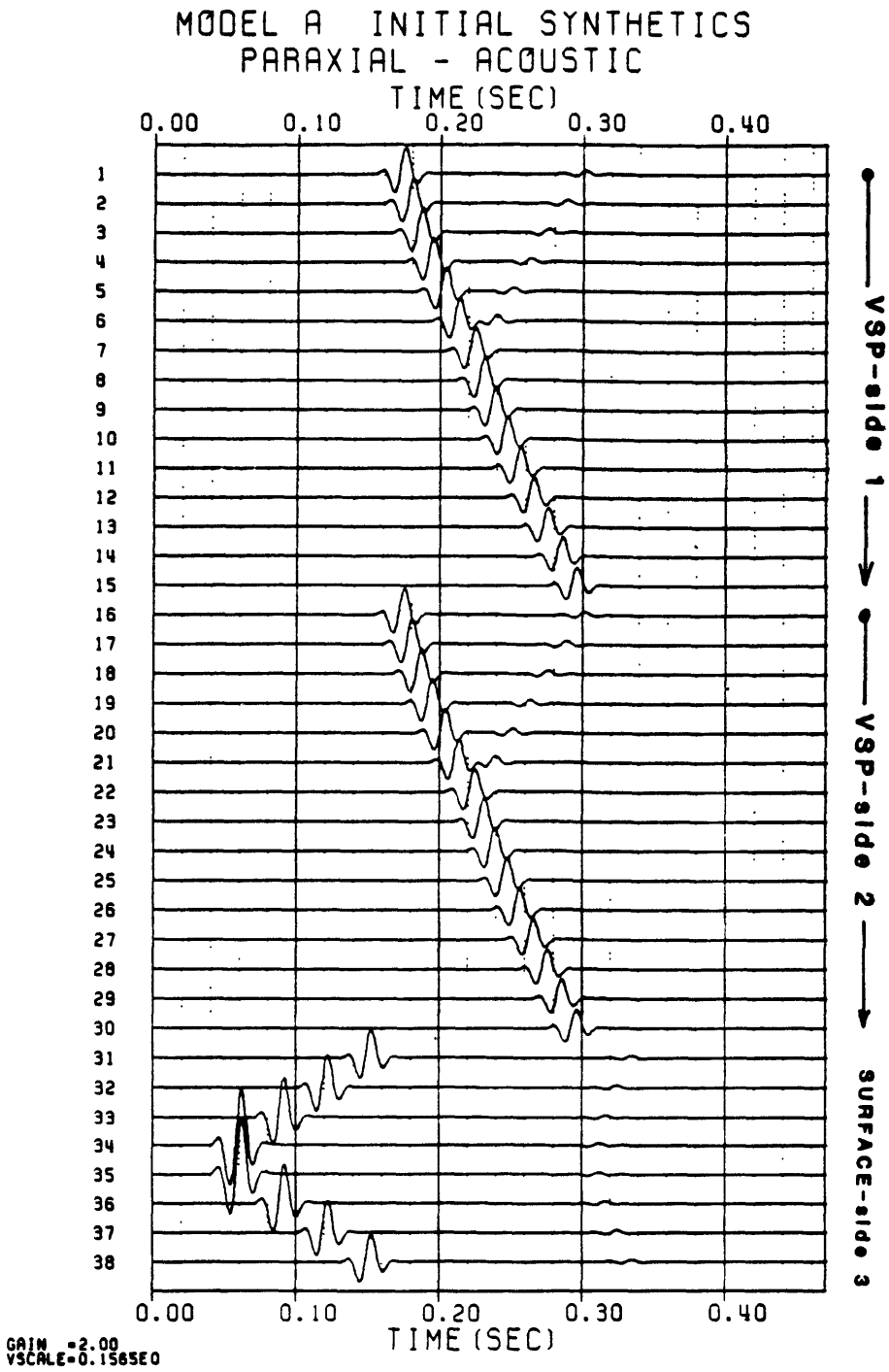


Figure 16b. Model A synthetic data, generated by the paraxial ray method.

```
 INITIAL  PARAMETER VECTOR:  m0(1 -> M)

      0.        3.30000    4.10000

      GAMMA**2 =    0.          E =   26.59917    RER (%)=     0.


----> ITERATION #   1, (CALL#  5):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.00895   3.32720   4.08391
      GAMMA**2 =    0.          E =   25.83407    RER (%)=     2.88
----> ITERATION #   2, (CALL#  9):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.01343   3.37085   4.01638
      GAMMA**2 =    0.          E =   21.94456    RER (%)=    17.50
----> ITERATION #   3, (CALL# 13):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.02534   3.43660   3.89636
      GAMMA**2 =    0.          E =   11.85363    RER (%)=    55.44
----> ITERATION #   4, (CALL# 17):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.03884   3.49967   3.79257
      GAMMA**2 =    0.          E =    5.61612    RER (%)=    78.89
----> ITERATION #   5, (CALL# 21):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.04887   3.51334   3.78540
      GAMMA**2 =    0.          E =    5.50426    RER (%)=    79.31
----> ITERATION #   6, (CALL# 25):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.07753   3.51676   3.80297
      GAMMA**2 =    0.          E =    5.45341    RER (%)=    79.50
----> ITERATION #   7, (CALL# 29):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.09257   3.51622   3.83497

      GAMMA**2 =    0.          E =    5.39309    RER (%)=    79.72
----> ITERATION #   8, (CALL# 33):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.10505   3.51654   3.86752
      GAMMA**2 =    0.          E =    5.32958    RER (%)=    79.96
----> ITERATION #   9, (CALL# 37):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.11617   3.51734   3.90449
      GAMMA**2 =    0.          E =    5.24571    RER (%)=    80.28
----> ITERATION #  10, (CALL# 41):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.12812   3.51711   3.95587
      GAMMA**2 =    0.          E =    5.05466    RER (%)=    81.00
----> ITERATION #  11, (CALL# 45):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.13500   3.51636   4.01935
      GAMMA**2 =    0.          E =    4.65999    RER (%)=    82.48
----> ITERATION #  12, (CALL# 49):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.14607   3.51576   4.10980
      GAMMA**2 =    0.          E =    3.79964    RER (%)=    85.72
----> ITERATION #  13, (CALL# 53):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.15731   3.51458   4.23570
      GAMMA**2 =    0.          E =    2.29807    RER (%)=    91.36
----> ITERATION #  14, (CALL# 57):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.16418   3.51200   4.37266
      GAMMA**2 =    0.          E =    1.01945    RER (%)=    96.17
----> ITERATION #  15, (CALL# 61):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.17087   3.51013   4.44828
      GAMMA**2 =    0.          E =    0.67068    RER (%)=    97.48
----> ITERATION #  16, (CALL# 65):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.17909   3.50917   4.47747
      GAMMA**2 =    0.          E =    0.39147    RER (%)=    98.53
----> ITERATION #  17, (CALL# 69):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.18828   3.50847   4.50384
      GAMMA**2 =    0.          E =    0.10897    RER (%)=    99.59
----> ITERATION #  18, (CALL# 73):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.19235   3.50835   4.51752
      GAMMA**2 =    0.          E =    0.07064    RER (%)=    99.73
----> ITERATION #  19, (CALL# 77):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.19262   3.50840   4.51808
      GAMMA**2 =    0.          E =    0.07068    RER (%)=    99.73
```

Table 1. Information on inversion of Model A: Initial and updated parameters.

```
TOTAL NUMBER OF PARAMETERS:              M =    3
TOTAL NUMBER OF RESIDUALS:               N =  689
LMDIF INVERSION EXIT CONDITION:     INFO =    2
TOTAL NUMBER OF FORWARD MODELS  CALLS =   79
TOTAL NUMBER OF ITERATIONS:              k =   20
FINAL LEVENBERG-MARQUARDT P: GAMMA**2 =    0.      e+00
FINAL RESIDUAL ENERGY:                   E =    0.07064
ESTIMATED DATA VARIANCE:         SIGMA**2 =    0.10298e-03
FINAL RESID. ENERGY REDUCTION: RER(%) =   99.73
MODEL 1 SCALAR WEIGHT:                  w1 =    4.00000

 ESTIMATED PARAMETER VECTOR:  m(1 -> M)

      0.19235    3.50835    4.51752

 PARAMETER STD.DEV. VECTOR: sqrt[Vm(1 -> M)]

      0.00023    0.00025    0.00179

PARAM. COVARIANCE:Vm
                1                2                3

  1    0.510798e-07    0.169078e-07    0.674127e-07

  2    0.169078e-07    0.621151e-07   -0.184516e-06

  3    0.674127e-07   -0.184516e-06    0.322192e-05

 PARAMETER CORRELATION COEFFICIENTS

r( 1, 2)= 0.300 r( 1, 3)= 0.166 r( 2, 3)=-0.412
```

Table 1. Information on inversion of Model A: Final parameters, error information.

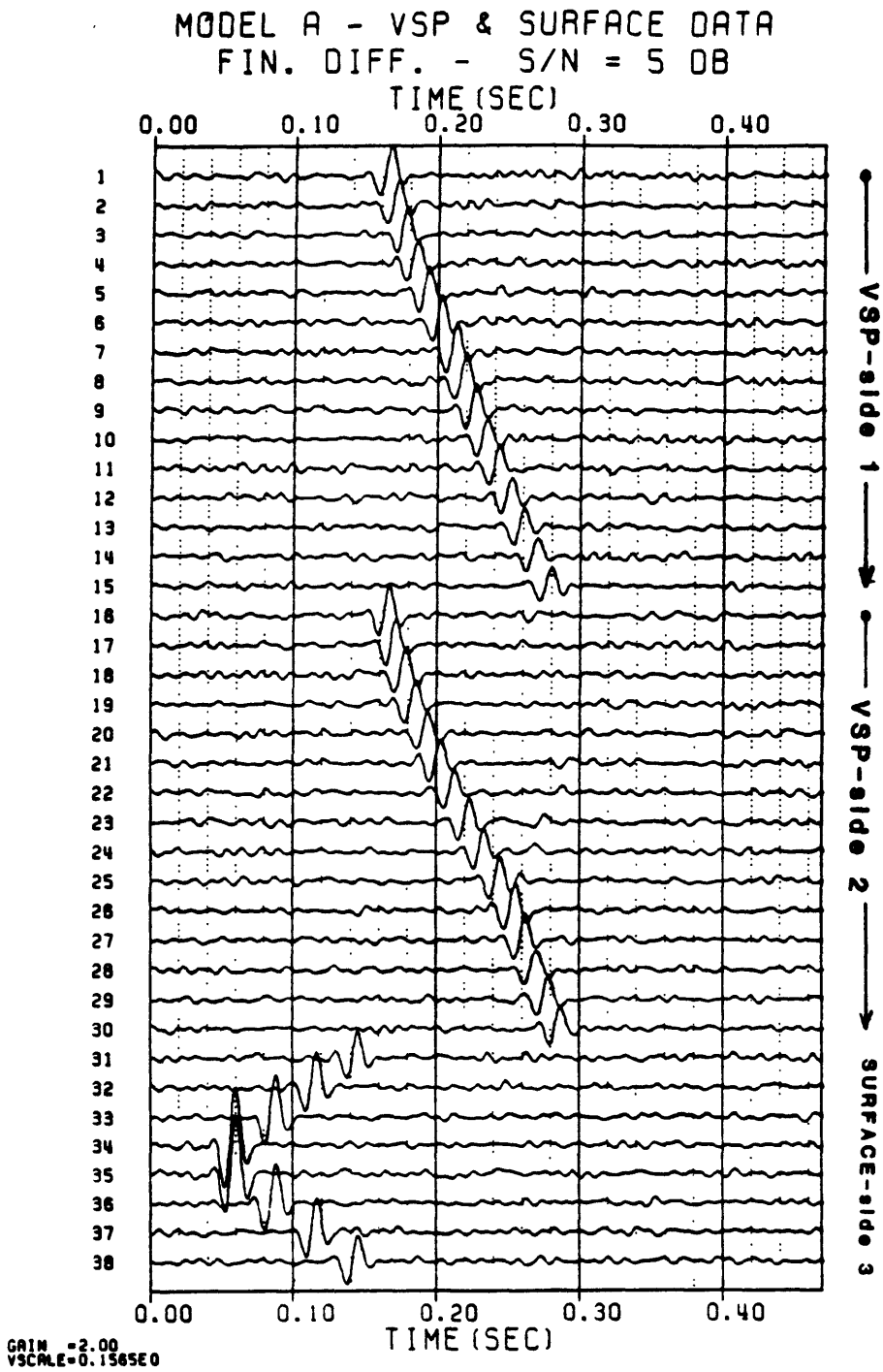


Figure 17. Figure 16a, (Model A) with added noise, SNR=5 dB.

```
 INITIAL  PARAMETER VECTOR:  m0(1 -> M)

      0.        3.30000    4.10000

      GAMMA**2 =     0.          E =    30.96476     RER (%)=      0.

----> ITERATION #   1, (CALL#  5):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.01034   3.32800    4.09175
      GAMMA**2 =     0.          E =    29.97300     RER (%)=      3.20
----> ITERATION #   2, (CALL#  9):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.02124   3.37325    4.03989
      GAMMA**2 =     0.          E =    25.59078     RER (%)=     17.36
----> ITERATION #   3, (CALL# 13):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.02995   3.44150    3.92214
      GAMMA**2 =     0.          E =    14.78338     RER (%)=     52.26
----> ITERATION #   4, (CALL# 17):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.05164   3.50266    3.83479
      GAMMA**2 =     0.          E =     8.98205     RER (%)=     70.99
----> ITERATION #   5, (CALL# 21):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.07743   3.51476    3.84955
      GAMMA**2 =     0.          E =     8.81271     RER (%)=     71.54
----> ITERATION #   6, (CALL# 25):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.09849   3.51592    3.88429
      GAMMA**2 =     0.          E =     8.64515     RER (%)=     72.08
----> ITERATION #   7, (CALL# 29):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.11528   3.51741    3.92627

      GAMMA**2 =     0.          E =     8.54188     RER (%)=     72.41
----> ITERATION #   8, (CALL# 33):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.12288   3.51724    3.97400
      GAMMA**2 =     0.          E =     8.35738     RER (%)=     73.01
----> ITERATION #   9, (CALL# 37):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.13502   3.51658    4.04132
      GAMMA**2 =     0.          E =     7.91324     RER (%)=     74.44
----> ITERATION #  10, (CALL# 41):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.15176   3.51648    4.13633
      GAMMA**2 =     0.          E =     7.04318     RER (%)=     77.25
----> ITERATION #  11, (CALL# 45):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.15667   3.51457    4.25633
      GAMMA**2 =     0.          E =     5.74775     RER (%)=     81.44
----> ITERATION #  12, (CALL# 49):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.16345   3.51280    4.37882
      GAMMA**2 =     0.          E =     4.75150     RER (%)=     84.66
----> ITERATION #  13, (CALL# 53):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.17059   3.51126    4.44588
      GAMMA**2 =     0.          E =     4.41426     RER (%)=     85.74
----> ITERATION #  14, (CALL# 57):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.18101   3.50999    4.47926
      GAMMA**2 =     0.          E =     3.96186     RER (%)=     87.21
----> ITERATION #  15, (CALL# 61):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.19138   3.50886    4.51268
      GAMMA**2 =     0.          E =     3.67103     RER (%)=     88.14
----> ITERATION #  16, (CALL# 65):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.19347   3.50886    4.52124
      GAMMA**2 =     0.          E =     3.66126     RER (%)=     88.18
----> ITERATION #  17, (CALL# 69):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.19401   3.50908    4.52200
      GAMMA**2 =     0.          E =     3.65780     RER (%)=     88.19
----> ITERATION #  18, (CALL# 73):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.19384   3.50906    4.52136
      GAMMA**2 =     0.          E =     3.65653     RER (%)=     88.19
----> ITERATION #  19, (CALL# 77):    UPDATED PARAM. VECTOR: m(1 -> M)
      0.19370   3.50901    4.52088
      GAMMA**2 =     0.          E =     3.65738     RER (%)=     88.19
```

Table 2. Information on inversion of Model A with 5 dB *SNR*: Initial and updated parameters.

```
TOTAL NUMBER OF PARAMETERS:          M =     3
TOTAL NUMBER OF RESIDUALS:           N =   689
LMDIF INVERSION EXIT CONDITION:   INFO =     2
TOTAL NUMBER OF FORWARD MODELS   CALLS =    77
TOTAL NUMBER OF ITERATIONS:          k =    20
FINAL LEVENBERG-MARQUARDT P: GAMMA**2 =     0.      e+00
FINAL RESIDUAL ENERGY:               E =     3.65653
ESTIMATED DATA VARIANCE:      SIGMA**2 =     0.53302e-02
FINAL RESID. ENERGY REDUCTION: RER(%) =    88.19
MODEL 1 SCALAR WEIGHT:              w1 =     4.000000
```

ESTIMATED PARAMETER VECTOR: m(1 -> M)

```
     0.19384    3.50906    4.52136
```

PARAMETER STD.DEV. VECTOR: sqrt[Vm(1 -> M)]

```
     0.00162    0.00179    0.01294
```

PARAM. COVARIANCE:Vm

	1	2	3
1	0.262732e-05	0.863323e-06	0.340159e-05
2	0.863323e-06	0.318958e-05	-0.958286e-05
3	0.340159e-05	-0.958286e-05	0.167408e-03

PARAMETER CORRELATION COEFFICIENTS

```
r( 1, 2)= 0.298 r( 1, 3)= 0.162 r( 2, 3)=-0.415
```

Table 2. Information on inversion of Model A with 5 dB *SNR*: Final parameters, error information.

Inversion of Model B

This is the acoustic version of Model 4 in chapter IV, figure 12. There is a total of $N_S = 60$ receivers, and source and frequency samples are same as Model A. The first interface gives rise to head waves that are not modeled by the paraxial method. Arrivals that are in this range are muted out for surface receivers. First arrivals and primary reflected waves from the second interface are removed, along with head waves. The finite difference data is shown in figure 18. In the inversion, we will assume that parameters of the first layer along with the second interface are known quite accurately. Nine parameters are defined and constitute the parameter vector, with components

$[\vec{m}]^T = (\Phi , S , B , b , L , h , \alpha_2 , \alpha_3 , \alpha_4)$.

Φ is interface 3 dip angle (see G.1) with units in $radians$, (S,B,b,L,h) are the reef parameters of interface 4 (see G.2), with units in km. α_2, α_3 and α_4 are velocities of layers 2, 3 and 4, respectively, with units in km/sec.

The optimal parameter vector is chosen to be:

$[\vec{m}^*]^T = (-0.122 , 0.05 , 0.7 , 0.2 , 0.35 , 0.1 , 3.5 , 4 , 4.5)$.

The optimal interface dip angle is about -7^o. The fixed point on the interface has coordinates ($x_f = 0.05\ km$, $z_f = 0.4\ km$). Initial parameter vector is

$[\vec{m}_0]^T = (-0.2 , 0.04 , 0.77 , 0.22 , 0.32 , 0.11 , 3.4 , 4.2 , 4.3)$,

with corresponding generated synthetics displayed in figure 19. Only 3 ray codes are taken in the forward model, with a maximum of 30 rays per ray code: (1) the direct wave, (2) the primary reflected wave from interface 3 and (3) the primary reflected wave from interface 4.

The inversion yields the following estimates

$[\hat{m}]^T = (-0.13 , 0.04 , 0.77 , 0.07 , 0.34 , 0.09 , 3.5 , 4.0 , 4.49)$.

Velocities have been very well estimated. The experiment does not constrain sufficiently the size of the reef, in particular its large base, B, and its small base b, which is why their estimate are poor. This is seen in the ray diagram in figure 16 of chapter IV. One could possibly constrain these parameters by a priori information. The wavelength being on the order of 70 meters, we are requiring a resolution on S and h that is a fraction of the wavelength. Inversion details are presented in Table 3 and synthetics of estimated parameters are shown in figure 20.

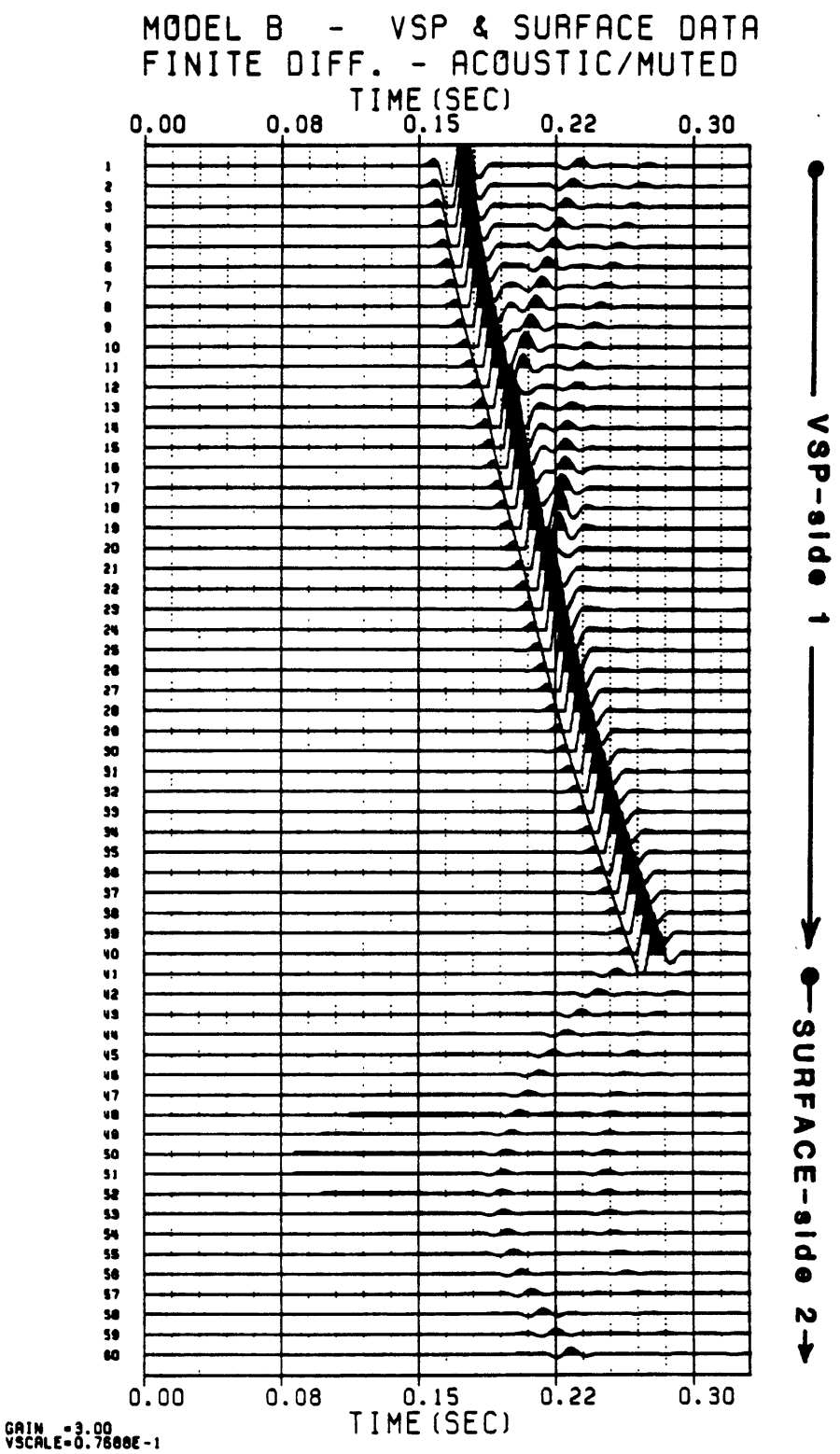


Figure 18. Model B synthetic data, generated by finite difference.

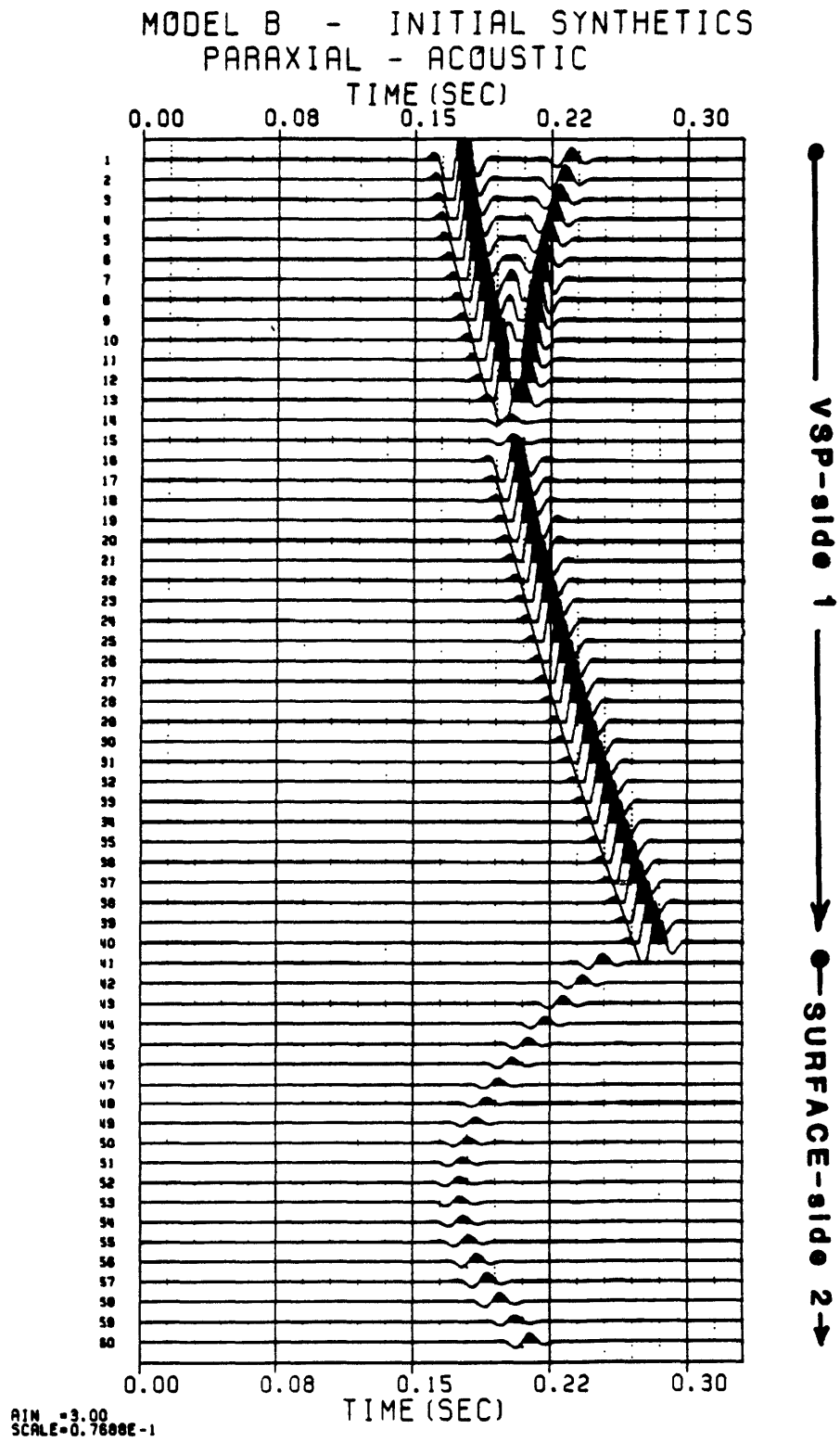


Figure 19. Paraxial synthetics with initial parameters of Model B.

```
 INITIAL  PARAMETER VECTOR:  m0(1 -> M)

    -0.20000   0.04000   0.77000   0.22000   0.32000   0.11000
     3.40000   4.20000   4.30000

     GAMMA**2 =    0.          E =   50.04435    RER (%)=     0.


----> ITERATION #   1, (CALL# 11):   UPDATED PARAM. VECTOR: m(1 -> M)
    -0.18997   0.04537   0.76990   0.07257   0.31528   0.09360
     3.42949   4.19056   4.37177
     GAMMA**2 =    0.          E =   29.29837    RER (%)=    41.46
----> ITERATION #   2, (CALL# 21):   UPDATED PARAM. VECTOR: m(1 -> M)
    -0.19098   0.03835   0.76872   0.01302   0.37168   0.08872
     3.45047   4.12335   4.42908
     GAMMA**2 =    0.          E =   17.90876    RER (%)=    64.21
----> ITERATION #   3, (CALL# 31):   UPDATED PARAM. VECTOR: m(1 -> M)
    -0.17289   0.03761   0.76698   0.01846   0.35851   0.09197
     3.45705   4.11089   4.44552

     GAMMA**2 =    0.          E =   14.81712    RER (%)=    70.39
----> ITERATION #   4, (CALL# 41):   UPDATED PARAM. VECTOR: m(1 -> M)
    -0.17644   0.03794   0.76955   0.01271   0.35417   0.09261
     3.46617   4.05701   4.46057
     GAMMA**2 =    0.          E =   10.99367    RER (%)=    78.03
----> ITERATION #   5, (CALL# 51):   UPDATED PARAM. VECTOR: m(1 -> M)
    -0.16744   0.03846   0.76984   0.02495   0.35200   0.09230
     3.48518   4.03480   4.46185
     GAMMA**2 =    0.          E =    8.05435    RER (%)=    83.91
----> ITERATION #   6, (CALL# 61):   UPDATED PARAM. VECTOR: m(1 -> M)
    -0.16544   0.03865   0.77108   0.03308   0.34868   0.09370
     3.49271   4.00706   4.46815
     GAMMA**2 =    0.          E =    7.53079    RER (%)=    84.95
----> ITERATION #   7, (CALL# 71):   UPDATED PARAM. VECTOR: m(1 -> M)
    -0.15737   0.03890   0.77435   0.04018   0.35010   0.09366
     3.49657   4.00552   4.47120
     GAMMA**2 =    0.          E =    6.51910    RER (%)=    86.97
----> ITERATION #   8, (CALL# 81):   UPDATED PARAM. VECTOR: m(1 -> M)
    -0.15452   0.03894   0.77441   0.05096   0.34316   0.09275
     3.49660   3.99490   4.47664
     GAMMA**2 =    0.          E =    2.46324    RER (%)=    95.08
----> ITERATION #   9, (CALL# 91):   UPDATED PARAM. VECTOR: m(1 -> M)
    -0.14587   0.03810   0.77445   0.04844   0.34355   0.09300
     3.49547   3.99395   4.48241
     GAMMA**2 =    0.          E =    2.07989    RER (%)=    95.84
----> ITERATION #  10, (CALL#101):   UPDATED PARAM. VECTOR: m(1 -> M)
    -0.14313   0.03897   0.77437   0.05075   0.33860   0.09095
     3.49826   3.98878   4.48675
     GAMMA**2 =    0.          E =    1.88817    RER (%)=    96.23
----> ITERATION #  11, (CALL#111):   UPDATED PARAM. VECTOR: m(1 -> M)
    -0.13644   0.03937   0.77298   0.05801   0.33525   0.09099
     3.49608   3.99608   4.48493
     GAMMA**2 =    0.          E =    1.78092    RER (%)=    96.44
----> ITERATION #  12, (CALL#121):   UPDATED PARAM. VECTOR: m(1 -> M)
    -0.13432   0.03928   0.77074   0.06855   0.33592   0.09210
     3.49639   3.99828   4.48526
     GAMMA**2 =    0.          E =    1.56761    RER (%)=    96.87
----> ITERATION #  13, (CALL#131):   UPDATED PARAM. VECTOR: m(1 -> M)
    -0.13077   0.03852   0.77017   0.06739   0.33962   0.09161
     3.49629   4.00337   4.48779
     GAMMA**2 =    0.          E =    1.42116    RER (%)=    97.16
----> ITERATION #  14, (CALL#141):   UPDATED PARAM. VECTOR: m(1 -> M)
    -0.13079   0.03823   0.77018   0.07321   0.33648   0.09099
     3.49647   3.99979   4.48867
     GAMMA**2 =    0.          E =    1.45488    RER (%)=    97.09
```

Table 3. Information on inversion of Model B: Initial and updated parameters.

```
TOTAL NUMBER OF PARAMETERS:              M =     9
TOTAL NUMBER OF RESIDUALS:               N =  1096
LMDIF INVERSION EXIT CONDITION:     INFO =     2
TOTAL NUMBER OF FORWARD MODELS  CALLS =   146
TOTAL NUMBER OF ITERATIONS:              k =    15
FINAL LEVENBERG-MARQUARDT P: GAMMA**2 =     0.53075e+01
DATA ENERGY:                      E(DATA) =    52.98539
FINAL RESIDUAL ENERGY:                   E =     1.42116
ESTIMATED DATA VARIANCE:       SIGMA**2 =     0.13074e-02
FINAL RESID. ENERGY REDUCTION: RER(%) =     97.16
MODEL 1 SCALAR WEIGHT:                   w1 =     2.00000

 ESTIMATED PARAMETER VECTOR:  m(1 -> M)

    -0.13077    0.03852    0.77017    0.06739    0.33962    0.09161
     3.49629    4.00337    4.48779

PARAMETER STD.DEV. VECTOR: sqrt[Vm(1 -> M)]

    0.00004    0.00001    0.00140    0.00045    0.00004    0.00003
    0.00009    0.00011    0.00048

PARAM. COVARIANCE:Vm
                1                2                3                4                5                6                7                8
                9

  1   0.129685e-08    0.685881e-13   -0.670030e-10   -0.956364e-11   -0.666003e-12    0.216593e-11    0.150666e-10    0.751214e-11
      0.111784e-09

  2   0.685881e-13    0.420055e-10   -0.732763e-12   -0.402002e-12    0.347293e-12    0.498105e-12   -0.376007e-12    0.164663e-12
     -0.616762e-11

  3  -0.670030e-10   -0.732763e-12    0.197074e-05   -0.183667e-08   -0.107292e-10   -0.238159e-10    0.375515e-10   -0.177475e-10
     -0.001868e-10

  4  -0.956364e-11   -0.402002e-12   -0.183667e-08    0.204411e-06   -0.578758e-11   -0.920398e-11    0.241896e-11   -0.753736e-11
     -0.122204e-10

  5  -0.666003e-12    0.347293e-12   -0.107292e-10   -0.578758e-11    0.107065e-08   -0.551889e-11    0.758572e-11    0.582883e-11
      0.147193e-09

  6   0.216593e-11    0.498105e-12   -0.238159e-10   -0.920398e-11   -0.551889e-11    0.779545e-09   -0.238353e-11   -0.649258e-12
     -0.113551e-09

  7   0.150666e-10   -0.376007e-12    0.375515e-10    0.241896e-11    0.758572e-11   -0.238353e-11    0.776335e-08   -0.897225e-10
     -0.116262e-08

  8   0.751214e-11    0.164663e-12   -0.177475e-10   -0.753736e-11    0.582883e-11   -0.649258e-12   -0.897225e-10    0.119250e-07
     -0.726129e-09

  9   0.111784e-09   -0.616762e-11   -0.001868e-10   -0.122204e-10    0.147193e-09   -0.113551e-09   -0.116262e-08   -0.726129e-09
      0.233793e-06

 PARAMETER CORRELATION COEFFICIENTS

r( 1, 2)= 0.000 r( 1, 3)=-0.001 r( 1, 4)=-0.001 r( 1, 5)=-0.000 r( 1, 6)= 0.002 r( 1, 7)= 0.005
r( 1, 8)= 0.002 r( 1, 9)= 0.006 r( 2, 3)=-0.000 r( 2, 4)=-0.000 r( 2, 5)= 0.001 r( 2, 6)= 0.003
r( 2, 7)=-0.001 r( 2, 8)= 0.000 r( 2, 9)=-0.002 r( 3, 4)=-0.003 r( 3, 5)=-0.000 r( 3, 6)=-0.001
r( 3, 7)= 0.000 r( 3, 8)=-0.000 r( 3, 9)=-0.000 r( 4, 5)=-0.000 r( 4, 6)=-0.001 r( 4, 7)= 0.000
r( 4, 8)=-0.000 r( 4, 9)=-0.000 r( 5, 6)=-0.005 r( 5, 7)= 0.002 r( 5, 8)= 0.001 r( 5, 9)= 0.007
r( 6, 7)=-0.001 r( 6, 8)=-0.000 r( 6, 9)=-0.000 r( 7, 8)=-0.009 r( 7, 9)=-0.027 r( 8, 9)=-0.014
```

Table 3. Information on inversion of Model B: Final parameters, error information.

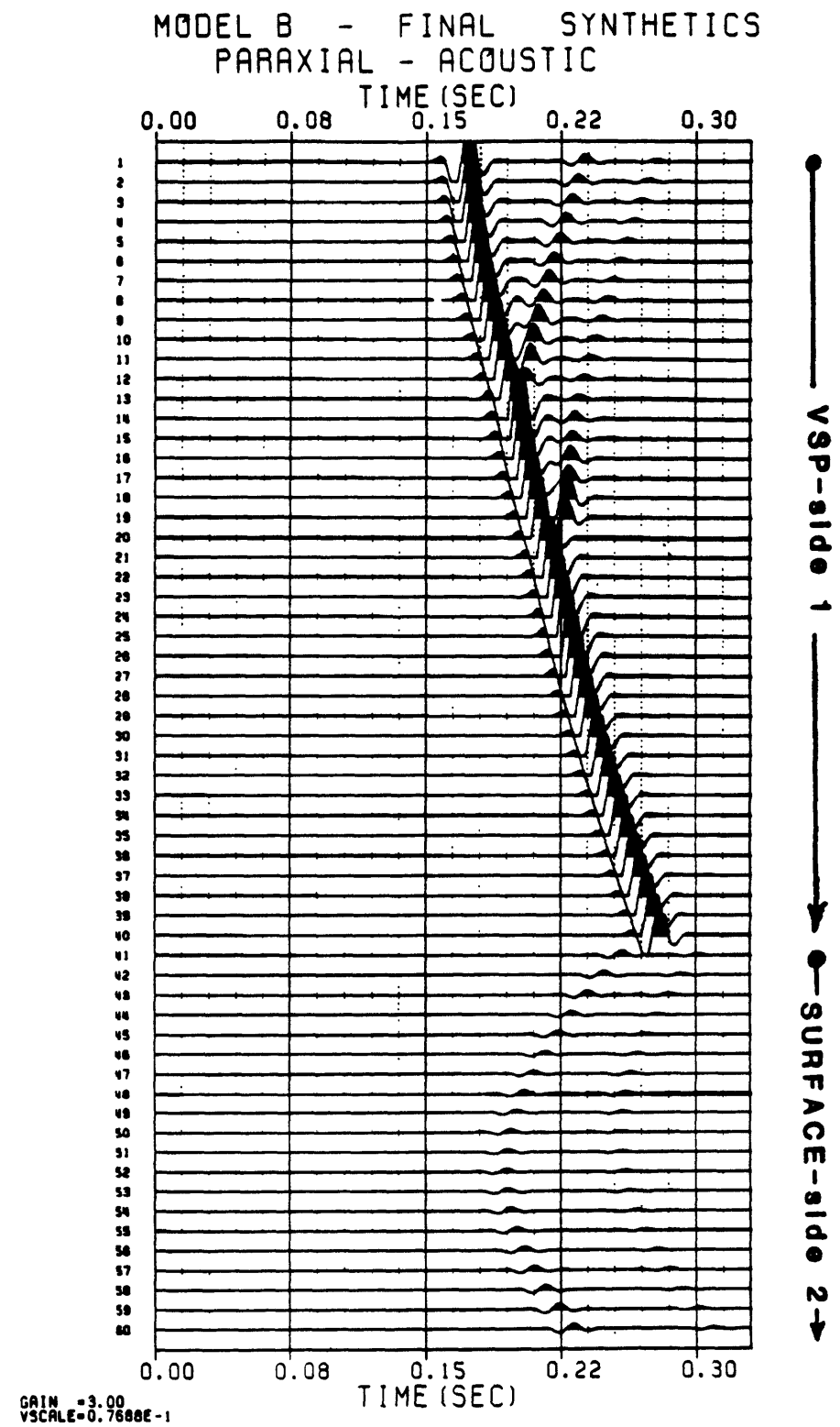


Figure 20. Paraxial synthetics after inversion, with figure 18 data.

Inversion of Field data

We shall present here a simple application of the described inversion for field data. An experimental group shoot for reservoir delineation was conducted by the Compagnie Génerale de Géophysique, with the scientific assistance of the Massachusetts Institute of Technology, during the fall of 1983. The site is in Manistee Co., Michigan. The experiment objective was to image a Niagaran reef, with data from vertical seismic profiling collected in St. Burch well, surface reflection, and logging tools. Site geology, field operations and results of the experiment are presented elsewhere. Dynamic ray tracing modeling was very helpful in planning and designing the experiment (Mellen, 1984). The set of experiments included several P and S wave offset VSP's with vibrators, a full waveform sonic recorded with Elf-Aquitaine's EVA system, a whole set of well logs, and 3D-CMP surface reflection survey.

An interpreted cross section of the region, prior to the experiment, is shown if figure 21. Part of the VSP data collected with a P wave source offset of 358 feet (109 meters), is shown in figure 22. Traces have been calibrated in amplitude with monitor phones situated close to the source.

Our aim, here, is to invert data within the reef area, for interface dip angles, and P velocities. This would estimate how far from the borehole the reef is, and possibly, determine how good the fit of a local part of the reef with flat interfaces is. Velocities will be constrained by a priori information resulting from full-waveform sonic estimates, and from a 1D-VSP full-waveform inversion (Stewart, 1983) done by Blackway (1985). The VSP inversions yield estimates of quality factors that are on the order of 60-100 for P waves, and separates up and down going waves. Interface locations on the borehole are precisely determined from combined sets of well logs.

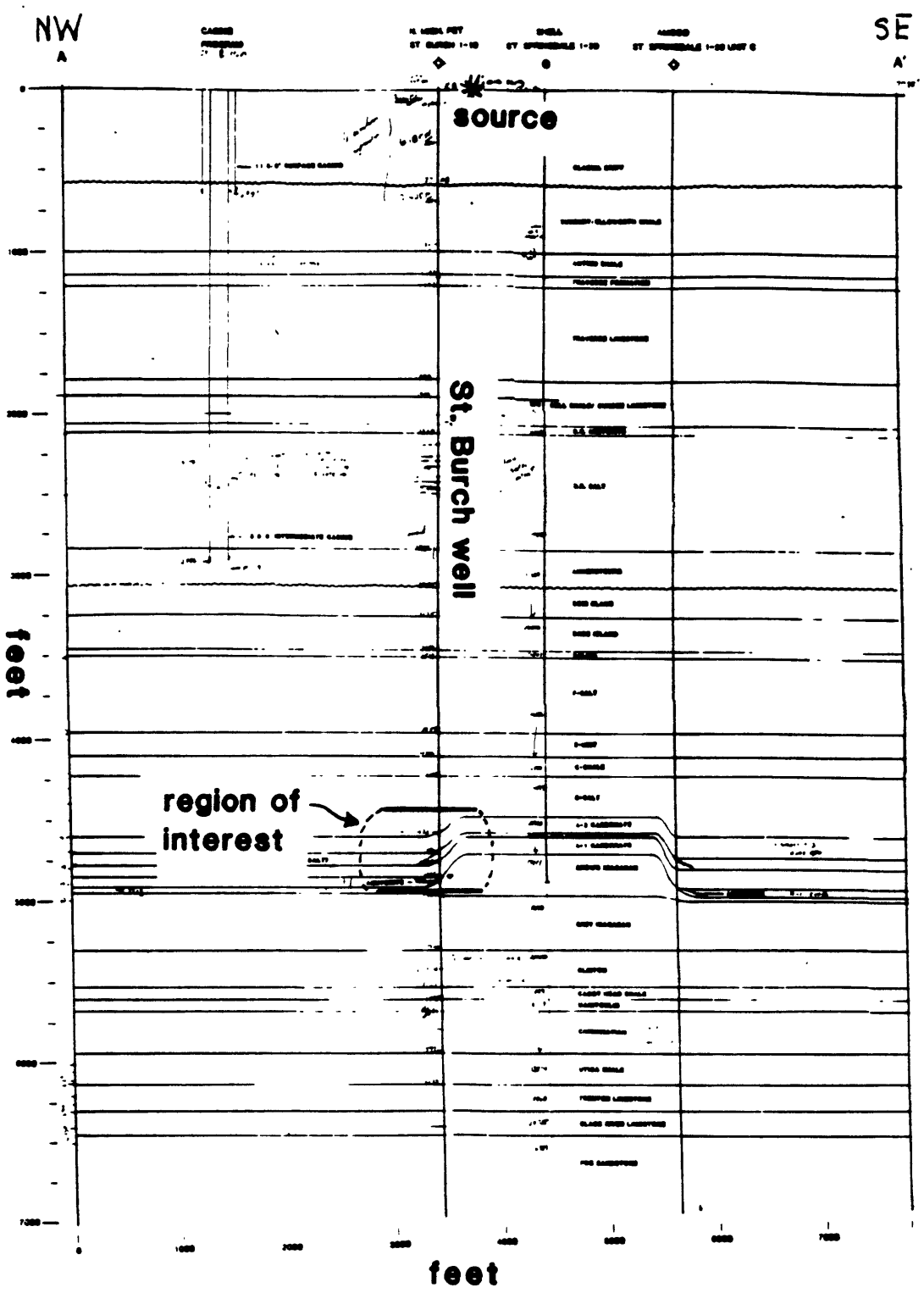


Figure 21. Interpreted cross section of St. Burch region, prior to the experiment (courtesy of West Bay Exploration Co.).

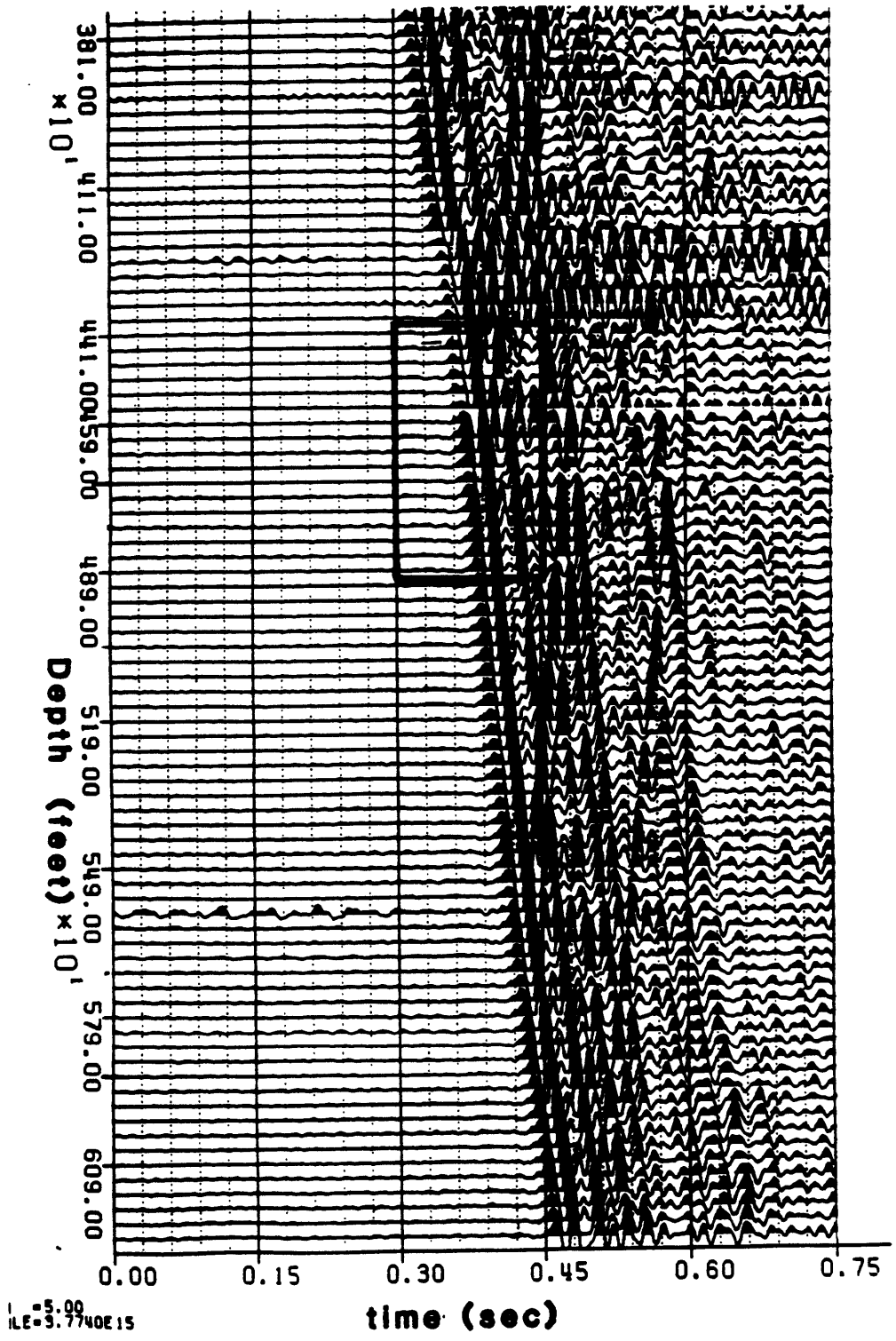


Figure 22. Portion of the VSP St. Burch field data, vertical component.

The depth interval that is of interest lies between 4410 feet (1344 meters) and 4890 feet (1490 meters). The receiver depth interval is 30 feet (9.144 meters). The time sampling interval is of 2 $msec$, and the time window is from 0.3 sec to 0.45 sec. The box in figure 22 corresponds to the set of data that will be inverted, which is expanded in figure 23. The sixth trace is somewhat "ringy"; we shall therefore not take it into account. The total number of traces is 16. The downgoing wave at depth 1490 meters, is given by the 1D-VSP inversion. This is the source wavelet with time function displayed in figure 24a, and its corresponding spectrum in figure 24b. The chosen frequency range is from 10 to 100 Hz, with 5 Hz sampling interval.

The elastic model considered has 4 layers, with 3 flat interfaces (omitting boundary interfaces). Starting from the top, fixed depths of each interface, on the borehole wall, are 4570 feet (1393 m) for the first, 4700 feet (1433 m) for the second, and 4755 feet (1449 m) for the third. Within the range of observation, with a quality factor for P waves on the order of 80, for a 50 Hz signal and a velocity on the order of 6 km/sec, there is a maximum of 5% error between the elastic amplitude and the near-elastic amplitude. The upper layer parameters are that of the layer at the first receiver. The model calibration to the data is achieved by matching both, travel time and maximum amplitude of the down going wave (Fig. 24) on the first trace. This approximation is justified by the fact that we are far away from the source (about 1350 meters), and that the observation range is small (less that 150 meters). Ten ray codes are considered to reconstruct the full waveform; direct P, primary P reflections, primary P to S reflected converted waves, and direct P to S conversion at each interface.

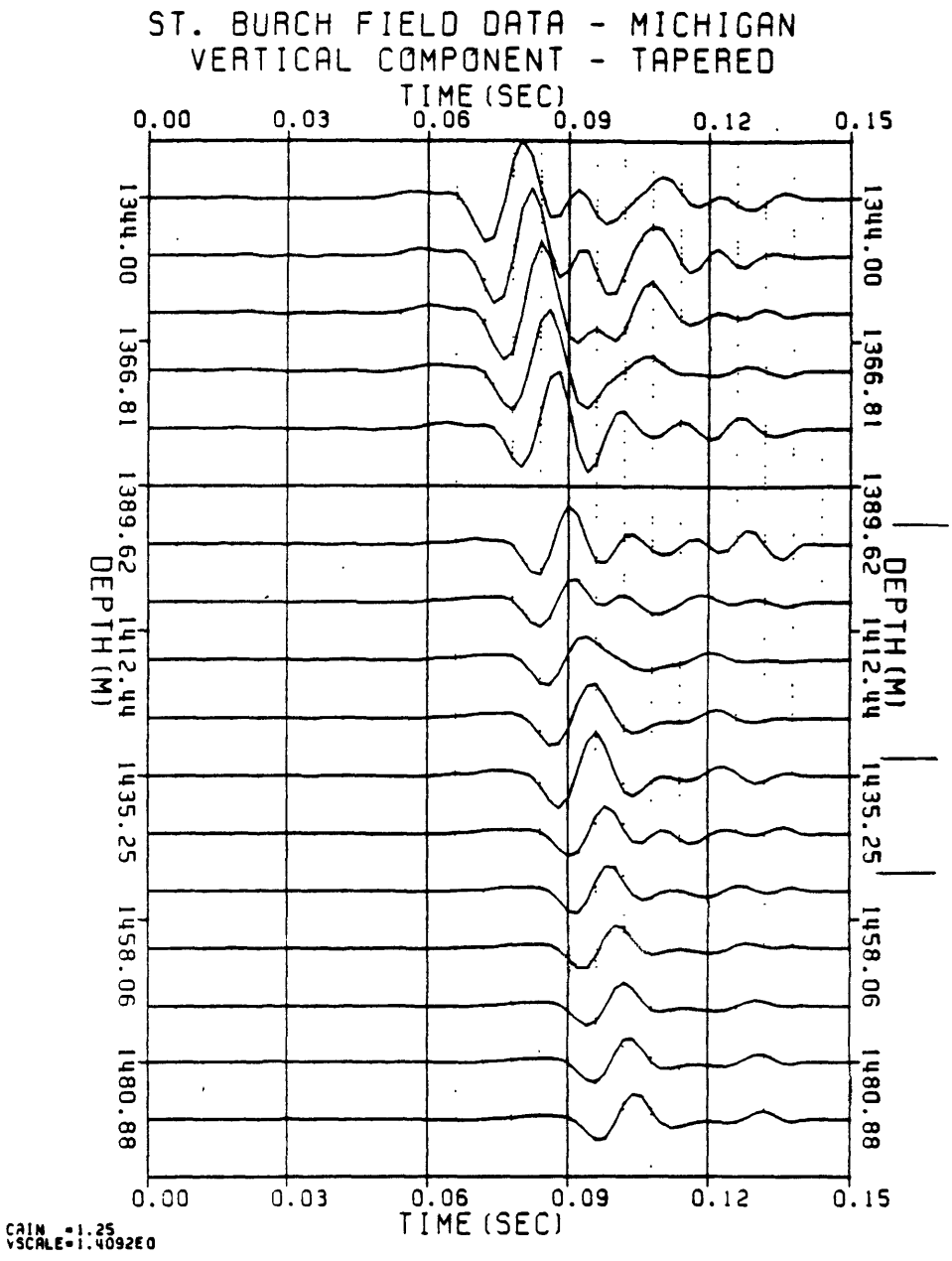


Figure 23. Enlarged section of the field data boxed in figure 22.

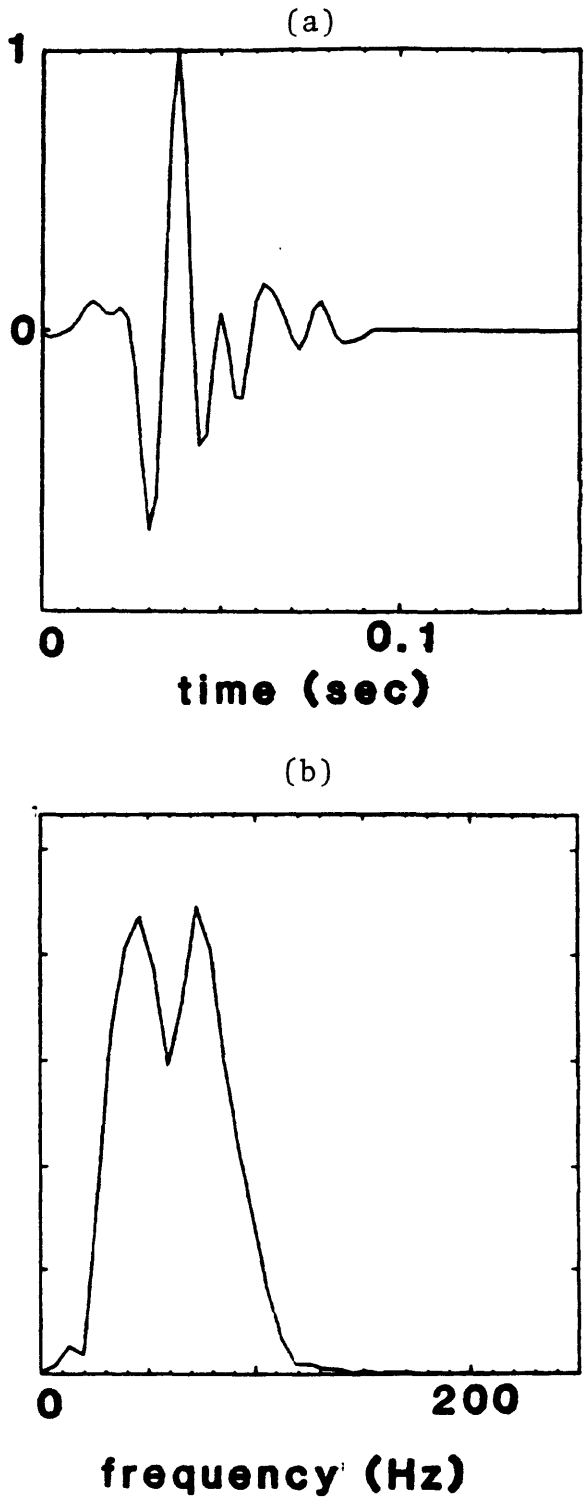


Figure 24. Source signal used in the inversion: (a) in time; (b) amplitude spectrum.

The parameters are the following, $[\vec{m}]^T = (\Phi_2 , \Phi_3 , \Phi_4 , \alpha_2 , \alpha_3 , \alpha_4)$. Initial values on velocities and densities are determined via full-wave sonic logging, 1D-VSP inversion, and the gravimetry survey (figures 25a and 25b). Values are shown in Table 4. In the inversion, at each iteration, S wave velocities will vary such that the P to S velocity ratio remains constant (i.e. constant Poisson ratio at each layer).

The standard deviation on velocities is taken to be about 10% of their values. This imposes constraints on velocity determination. Densities will remain constant. Initial values for interface dip angles are given by the full-wave sonic interpretation of the section (Paternoster, 1985). The initial angles will all be at -15^o.

Results of the inversion are shown in Table 4, with final seismograms in figure 26. The synthetics are satisfactory, considering the assumptions we have made, and the fact that we are fitting a flat interface to a local region of the reef that might not be flat. Some differences arise in the middle part of the section, which is thought to be due to local interface or medium changes, that are not handled by the model. The final local reconstruction of the model is presented in figure 27. Estimated parameters are the following,

$[\hat{m}]^T = (-6^o , -15^o , -9^o , 6.822 , 4.696 , 7.1505)$,

and all final parameters are presented in Table 4, along with estimated variances. The middle interface, is not well constrained by the data. Inversion with different initial values, yielded differences of Φ_3 that are on the order of 5 to 10 degrees. The seismograms for the same final model but with horizontal interfaces is shown in figure 28. These can be overlaid to figures 23 and 26. The slight differences indicate the need for the model to have non-zero dip interfaces.

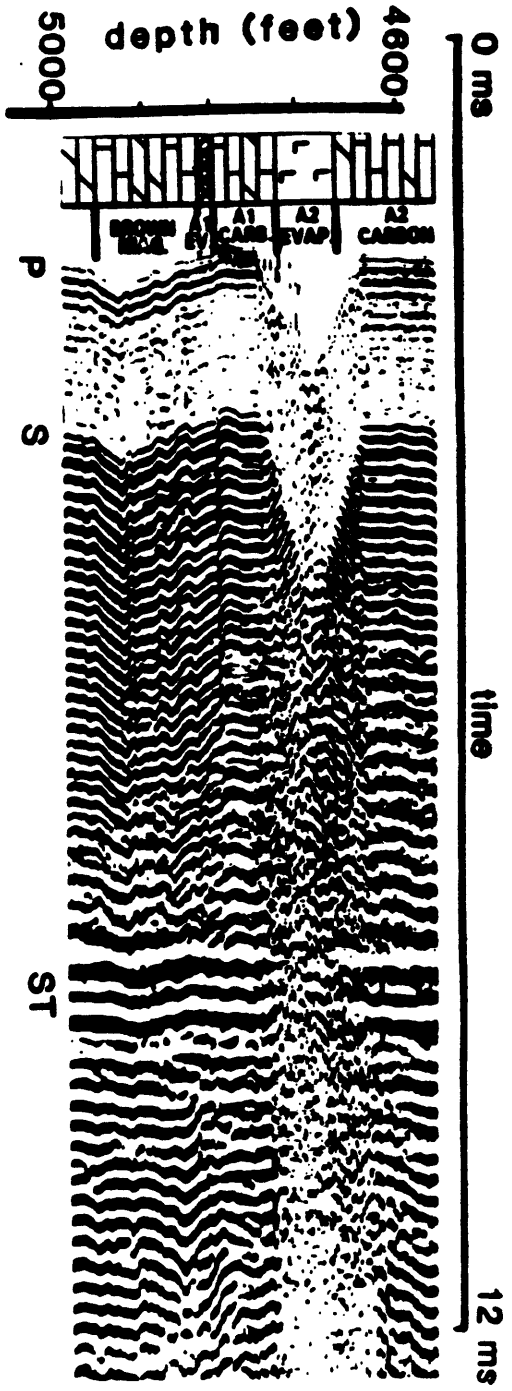


Figure 25a. Iso spacing section of Eva data of a portion of St. Burch well.

WELL LOGS

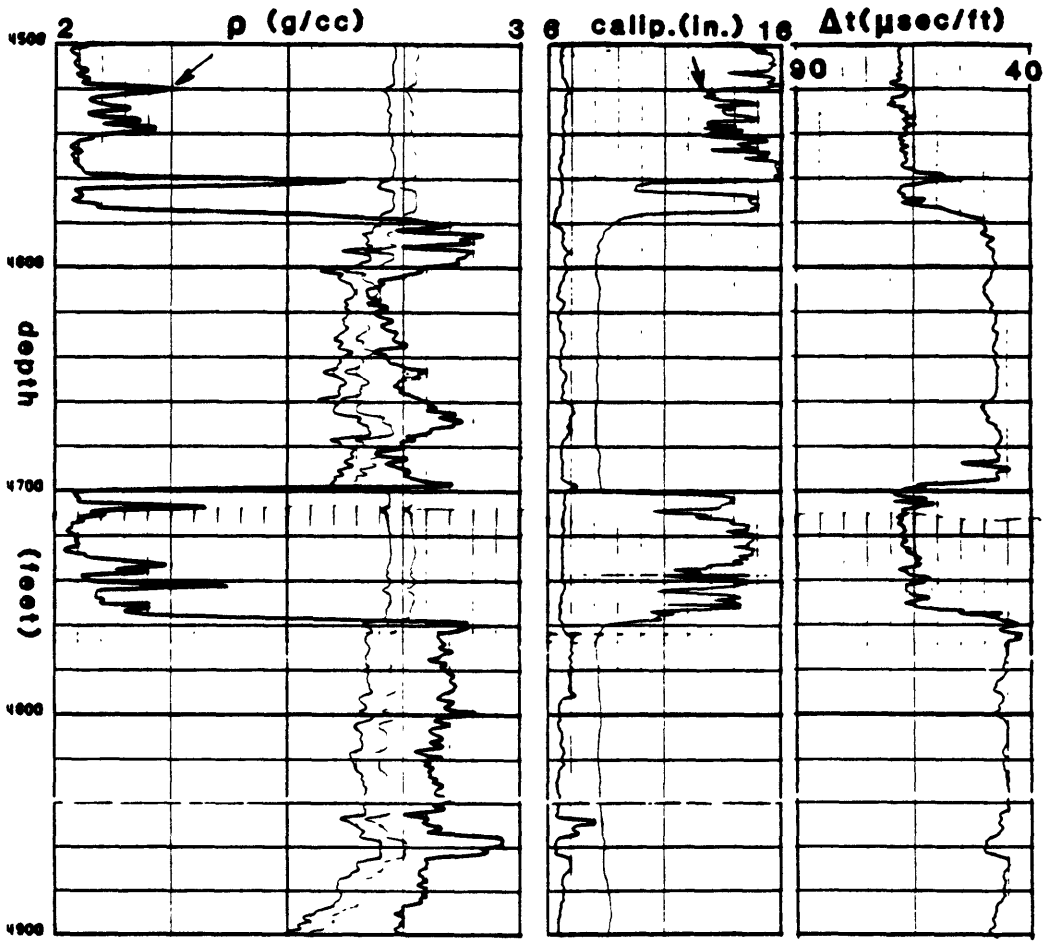


Figure 25b. Bulk density, caliper and sonic logs of a portion of St. Burch well.

The VSP inversion gives an apparent dip of a 3D interface, in the 2D plane spanned by the source and receivers. It is always smaller or equal than the true dip. These interface dip angles are consistent in absolute value with estimated values from full-wave sonic which predicts a $|\Phi_4|$ on the order of 15^o with a resolution of $\pm 10^o$. A similar approach done by Larrere (1985), estimates $|\Phi_4|$ to be positive but less than 10^o. Each of these methods give dip information of a fitted flat interface, for which the extent away from the borehole is very different. The full-waveform depth of investigation is on the order of a few meters, and the VSP range of investigation is on the order of 100 meters. This last value was determined geometrically, from the ray diagram. Considering these differences, the fit is satisfactory. However, a larger source offset would constrain better the problem, and should improve results.

INITIAL PARAMETERS FOR INVERSION OF FIELD DATA

LAYER #	$\alpha(km/sec)$	$\beta(km/sec)$	$\rho(g/cm^3)$
1	4.602	2.540	2.22
2	6.553	3.587	2.79
3	4.542	2.651	2.29
4	7.041	3.810	2.86

INTERFACE DIP ANGLES	ϕ
Interface 2	-15°
Interface 3	-15°
Interface 4	-15°

```
 INITIAL  PARAMETER VECTOR:  m0(1 -> M)
     -0.261000  -0.261800  -0.261800   6.553000   4.542000   7.041000
      GAMMA**2 =     0.          E =   43.14316    RER (%)=      0.

----> ITERATION #   1, (CALL#  8):    UPDATED PARAM. VECTOR: m(1 -> M)
     -0.120032  -0.26769  -0.13472   6.74642   4.68212   7.100716
      GAMMA**2 =     0.          E =   35.597005    RER (%)=     17.49
----> ITERATION #   2, (CALL# 15):    UPDATED PARAM. VECTOR: m(1 -> M)
     -0.070003  -0.31981  -0.15929   6.82438   4.66789   7.161004
      GAMMA**2 =     0.          E =   36.300793    RER (%)=     15.84
----> ITERATION #   3, (CALL# 25):    UPDATED PARAM. VECTOR: m(1 -> M)
     -0.110044  -0.27471  -0.13850   6.80247   4.69009   7.14857
      GAMMA**2 =     1.004721    E =   35.53677    RER (%)=     17.63
----> ITERATION #   4, (CALL# 33):    UPDATED PARAM. VECTOR: m(1 -> M)
     -0.105002  -0.27492  -0.14759   6.81238   4.69834   7.14246
      GAMMA**2 =     1.71183     E =   35.41612    RER (%)=     17.91
----> ITERATION #   5, (CALL# 40):    UPDATED PARAM. VECTOR: m(1 -> M)
     -0.100555  -0.27228  -0.15074   6.81837   4.69810   7.149300
      GAMMA**2 =     4.54530     E =   35.41001    RER (%)=     17.92
----> ITERATION #   6, (CALL# 47):    UPDATED PARAM. VECTOR: m(1 -> M)
     -0.100425  -0.27079  -0.15024   6.82007   4.69680   7.150200
      GAMMA**2 =    14.92362     E =   35.40081    RER (%)=     17.95
----> ITERATION #   7, (CALL# 54):    UPDATED PARAM. VECTOR: m(1 -> M)
     -0.100326  -0.27050  -0.15003   6.82070   4.69645   7.150048
      GAMMA**2 =    47.66368     E =   35.38346    RER (%)=     17.99
----> ITERATION #   8, (CALL# 61):    UPDATED PARAM. VECTOR: m(1 -> M)
     -0.100391  -0.27040  -0.14938   6.82173   4.69638   7.150096
      GAMMA**2 =    26.44896     E =   35.37545    RER (%)=     18.00
----> ITERATION #   9, (CALL# 68):    UPDATED PARAM. VECTOR: m(1 -> M)
     -0.100339  -0.27001  -0.14967   6.82353   4.69542   7.15216
      GAMMA**2 =    20.43200     E =   35.39267    RER (%)=     17.96
```

Table 4. Information on inversion of St. Burch well field data: Initial and updated parameters.

```
TOTAL NUMBER OF PARAMETERS:                M =     6
TOTAL NUMBER OF RESIDUALS:                 N =   426
LMDIF INVERSION EXIT CONDITION:         INFO =     1
TOTAL NUMBER OF FORWARD MODELS  CALLS =    70
TOTAL NUMBER OF ITERATIONS:                k =    10
FINAL LEVENBERG-MARQUARDT P: GAMMA**2 =     0.30675e+03
DATA ENERGY:                      E(DATA) =   200.04994
FINAL RESIDUAL ENERGY:                E =    35.37372
ESTIMATED DATA VARIANCE:      SIGMA**2 =     0.84223e-01
FINAL RESID. ENERGY REDUCTION: RER(%) =    18.01
MODEL 1 SCALAR WEIGHT:                w1 =     0.10000
```

ESTIMATED PARAMETER VECTOR: m(1 -> M)

-0.10387	-0.27038	-0.14940	6.82185	4.69631	7.15105

PARAMETER STD.DEV. VECTOR: sqrt[Vm(1 -> M)]

0.00064	0.00074	0.00066	0.00257	0.00296	0.00457

PARAM. COVARIANCE:Vm

	1	2	3	4	5	6
1	0.41497e-06	-0.38336e-09	-0.35897e-09	0.153528e-09	0.286618e-09	-0.481484e-10
2	-0.38336e-09	0.54589e-06	-0.941858e-10	0.837100e-09	0.482249e-09	-0.466350e-10
3	-0.35897e-09	-0.941858e-10	0.432242e-06	-0.111669e-08	-0.932598e-09	-0.109775e-09
4	0.153528e-09	0.837100e-09	-0.111669e-08	0.660400e-05	-0.683269e-08	-0.391238e-08
5	0.286618e-09	0.482249e-09	-0.932598e-09	-0.683269e-08	0.874001e-05	-0.135887e-08
6	-0.481484e-10	-0.466350e-10	-0.109775e-09	-0.391238e-08	-0.135887e-08	0.289191e-04

PARAMETER CORRELATION COEFFICIENTS

r(1, 2)=-0.001 r(1, 3)=-0.001 r(1, 4)= 0.000 r(1, 5)= 0.000 r(1, 6)=-0.000 r(2, 3)=-0.000
r(2, 4)= 0.000 r(2, 5)= 0.000 r(2, 6)=-0.000 r(3, 4)=-0.001 r(3, 5)=-0.000 r(3, 6)=-0.000
r(4, 5)=-0.001 r(4, 6)=-0.000 r(5, 6)=-0.000

Table 4. Information on inversion of St. Burch well field data: Error information.

FINAL PARAMETERS AFTER INVERSION

LAYER #	$\alpha(km/sec)$	$\beta(km/sec)$	$\rho(g/cm^3)$
1	4.602	2.540	2.22
2	6.822	3.734	2.79
3	4.696	2.741	2.29
4	7.151	3.870	2.86

INTERFACE DIP ANGLES	ϕ
Interface 2	−6°
Interface 3	−15°
Interface 4	−9°

Table 4. Information on inversion of St. Burch well field data: Final parameters.

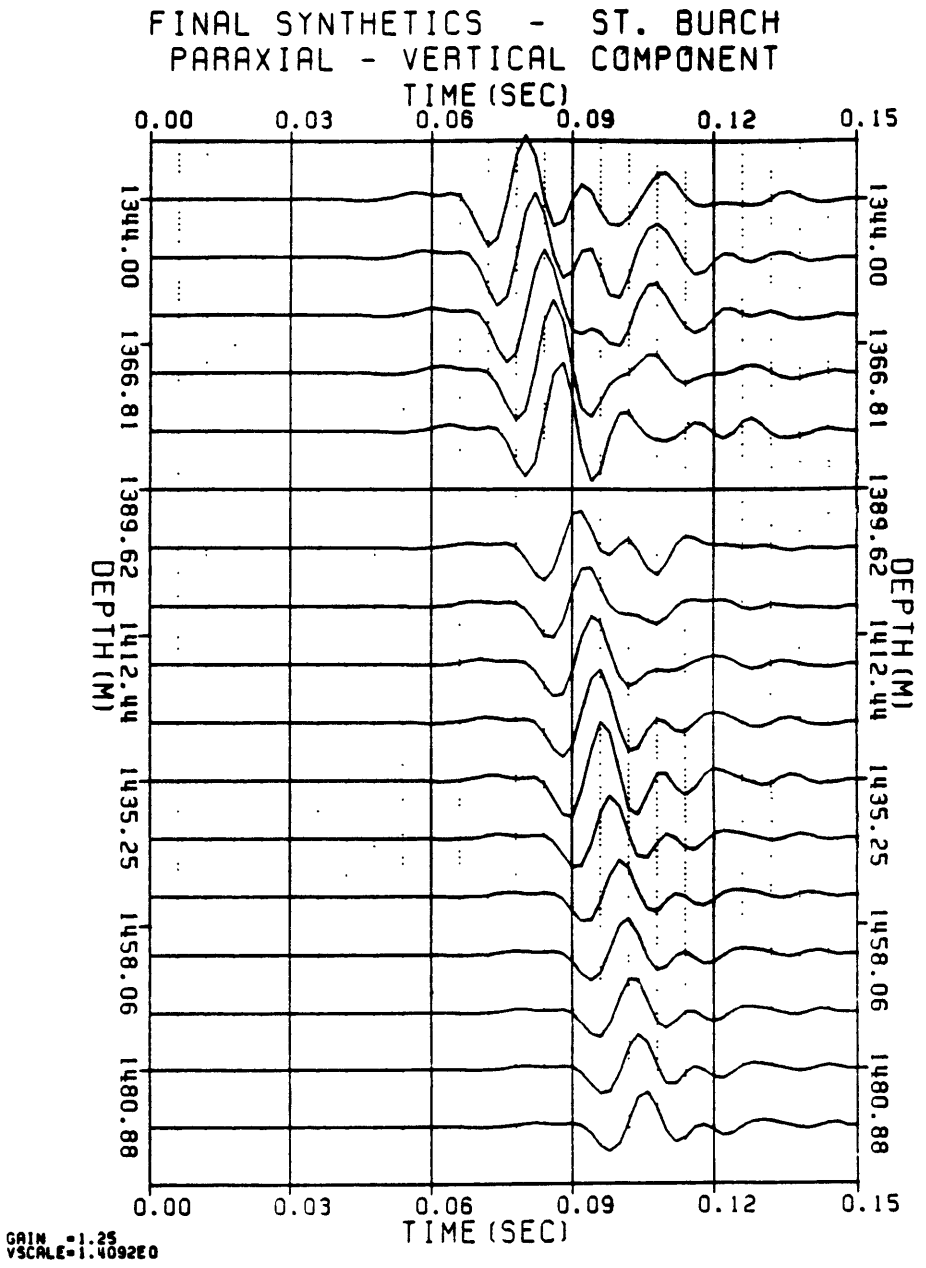


Figure 26. Paraxial synthetics after inversion, compare to figure 23.

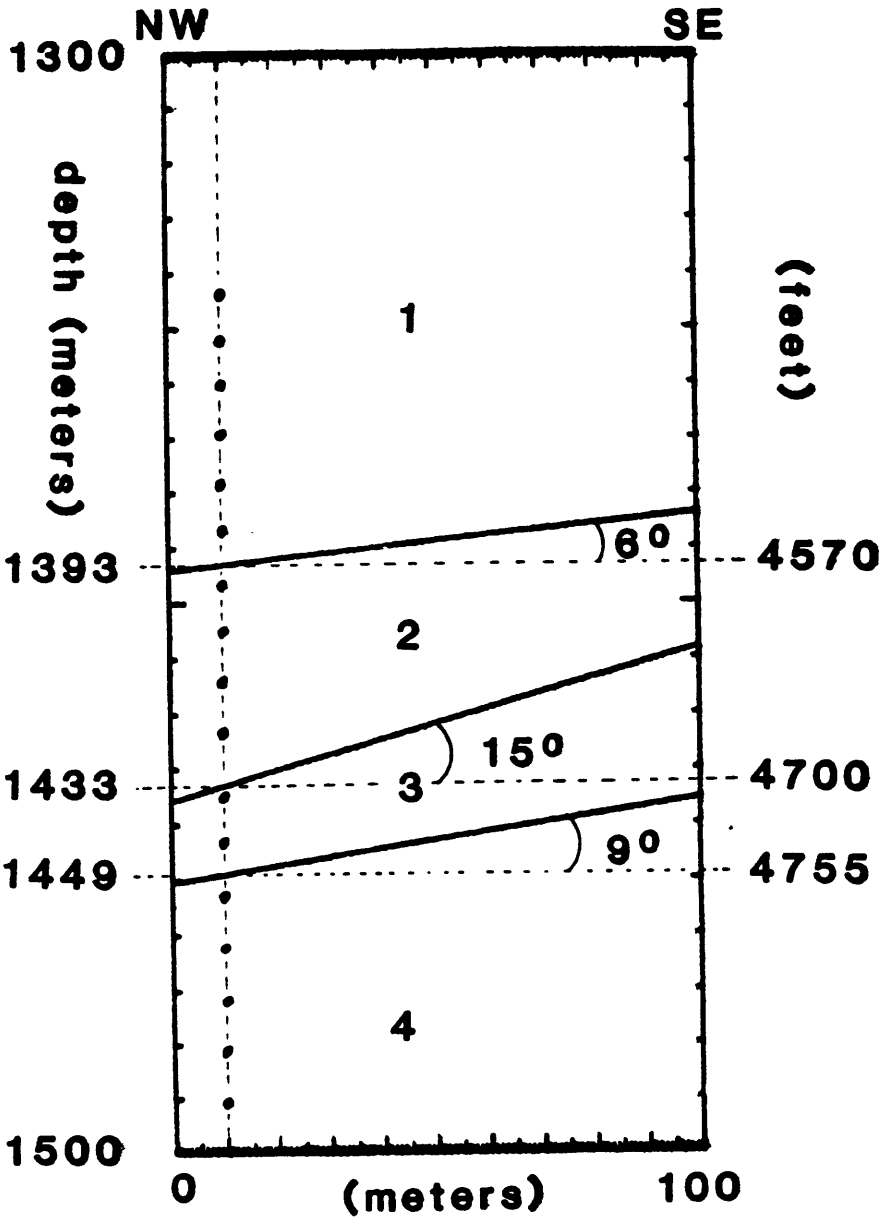


Figure 27. Local reconstruction of reef model in region boxed in figure 21.

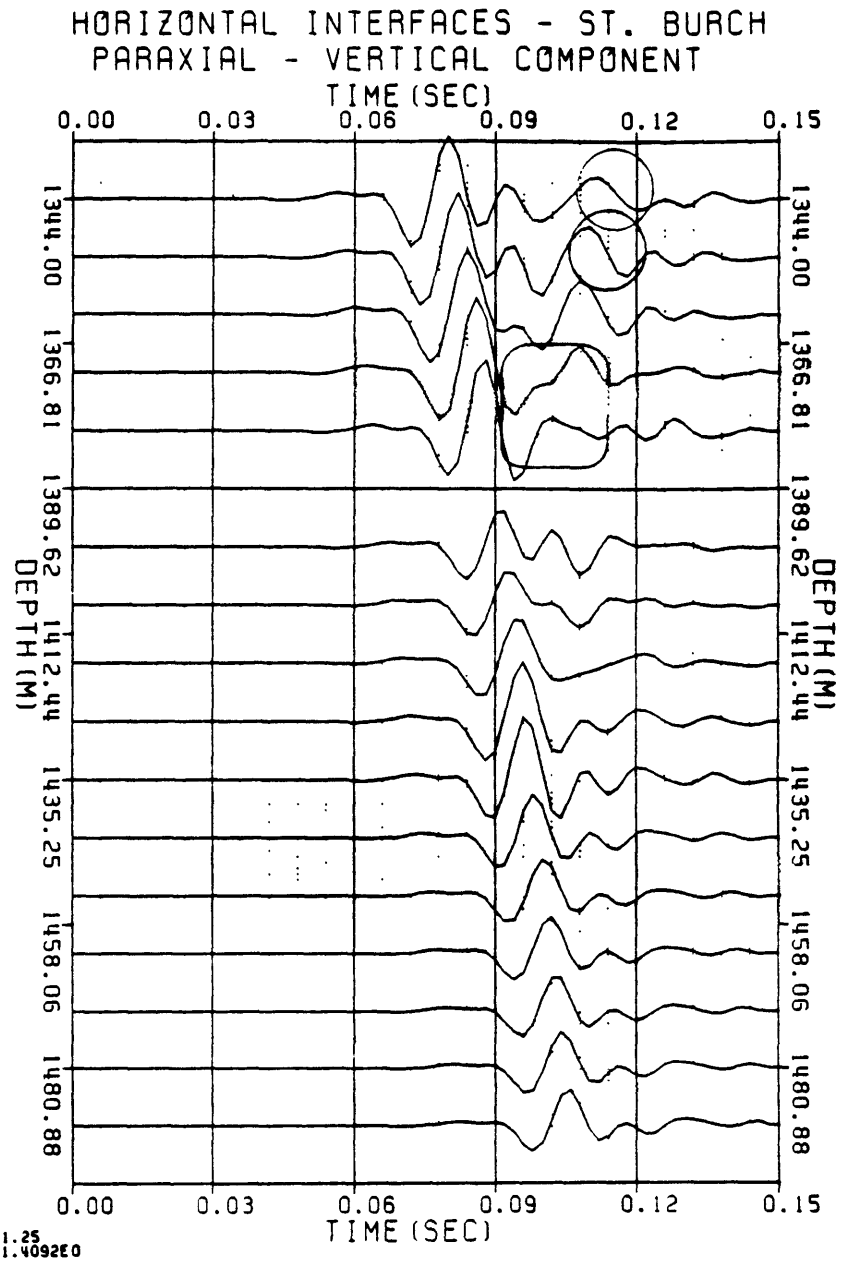


Figure 28. Paraxial synthetics of figure 27, with horizontal layers; compare to figures 23 and 26.

V.7. CONCLUSION

The non-linear least squares problem is presented in the context of tensor algebra. The residual energy is defined as one half the norm of the residual vector. The basis under which the data vectors are represented has its metrical tensor matrix equal to the inverse of the data covariance matrix. Linear operations on the residual energy (eg. high order derivatives) are straightforward. The Gauss-Newton method is rederived in this context.

Given a set of seismic data, resulting from combined seismic experiments with different types of prior information, the required estimates are layer and interface parameters. Full - wave inversion in the frequency domain is developed. The forward model uses the paraxial ray method.

Sensitivity analysis is investigated. Results show that velocity is a parameter with small variance but high non-uniqueness in its determination, whereas interface parameters, such as dipping angle, is more unique but has a large variance. The source shape has an important effect on residuals. Wavelet processing might be necessary if prior information is poor.

Finite difference data sets are inverted for interface parameters and layer velocities. For a given set of initial parameter estimates and prior information, examples show that parameters are estimated accurately for simple models. In complex media, the strategy could consists in inverting the data and use the velocity estimates, that are determined quite accurately, as initial velocity values for a second inversion. Initial interface parameters are the same as those in the first inversion. A simple example of field data inversion for structure and velocity information is presented. Results are compatible with that determined by full-wave sonic logging.

V.8. APPENDIX F: Definition and Review of Tensor analysis

A brief review on tensor analysis is presented. The scope is limited to an introduction of tensor formalism used in sections V.2, V.3 and V.4. Definitions are from Bass (1977), and Chambadal (1969). Basic properties and transformation laws can be found in Simmonds (1982), Morse and Feshbach (1953), Weyl (1952) and Ben-Menahem and Singh (1981).

Let E be a J dimensional discrete linear space on a commutative field K (commonly the set of real numbers R), and E* the dual of E, that is, the linear space of linear forms in K over E. E* is of dimension J and (E*)* = E. If p and q are two natural integers, the tensor of type (p,q) and order $(p+q)$ is an element of the linear space

$$\mathbf{T}^{p,q}(\mathbf{E}) = \mathbf{T}^{p}(\mathbf{E}) \otimes \mathbf{T}^{q}(\mathbf{E}^*), \tag{F.1}$$

where T(E) denotes the tensor algebra of E over K, and $\otimes$ is the tensor product of the two spaces $\mathbf{T}^{p}(\mathbf{E})$ and $\mathbf{T}^{q}(\mathbf{E}^*)$. The tensor is a $(p+q)$ linear form over $\mathbf{E}^{p} \times \mathbf{E}^{*q}$, and is p times contravariant and q times covariant. If p, q, r and s are four natural integers, the tensor product (or direct product) X of two tensors Z, in $\mathbf{T}^{p,q}(\mathbf{E})$, and W, in $\mathbf{T}^{r,s}(\mathbf{E})$, is defined as the function

$$(Z,X) \quad |\longrightarrow \quad X = Z \otimes W, \tag{F.2}$$

which is a bilinear form of $\mathbf{T}^{p,q}(\mathbf{E}) \times \mathbf{T}^{r,s}(\mathbf{E})$ in $\mathbf{T}^{p,q}(\mathbf{E}) \otimes \mathbf{T}^{r,s}(\mathbf{E})$. The resulting tensor X is of type $(p+r, q+s)$ and order $(p+q+r+s)$.

In the system of tensors, vectors are tensors of order 1, and scalars are tensors of order 0. Contravariant vectors (in E) are denoted with subscripts i.e. $\vec{u}_1, \vec{u}_2, \cdots$, whose components are denoted with superscripts, i.e. $x^1, x^2, \ldots, x^J$. Covariant vectors

(in E*) are denoted with superscripts, i.e. $\vec{v}^1, \vec{v}^2, \cdots$, whose components are denoted with subscripts $y_1, y_2, \ldots, y_J$.

Let $(\vec{a}^1, \vec{a}^2, \ldots, \vec{a}^J)$ be a basis of E*. If $\overset{\leftrightarrow}{Z}$ is a second order covariant tensor, it can be represented by

$$\overset{\leftrightarrow}{Z} = \sum_{i,j=1}^{J} Z_{ij}\ \vec{a}^i \otimes \vec{a}^j \ . \tag{F.3}$$

In other words, the Z_{ij} are the components of $\overset{\leftrightarrow}{Z}$ in the basis $\{\vec{a}^i \otimes \vec{a}^j\}$. A tensor can be identified as a linear operator, operating on vectors. In this sense, a second order tensor is called a dyadic, a third order tensor a triadic, etc. Elements of a dyadic such as $\overset{\leftrightarrow}{Z}$ in (F.3) are called dyads and are equal to $Z_{ij}\ \vec{a}^i \otimes \vec{a}^j$. The components Z_{ij} of $\overset{\leftrightarrow}{Z}$ can be represented by a matrix denoted $[\overset{\leftrightarrow}{Z}]$, linking, therefore, tensor algebra with linear operator algebra and matrix algebra. Sometimes, when no confusion arises, the matrix $[\overset{\leftrightarrow}{Z}]$ of a dyadic is simply denoted Z.

Henceforth, we define in E a scalar product (bilinear form) resulting in a Euclidian norm (quadratic form). If $(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_J)$ is a basis of E, the metrical structure underlying Euclidian space assign to every two vectors of E

$$\vec{\xi} = \sum_{i=1}^{J} \xi^i\ \vec{e}_i \qquad ; \qquad \vec{\zeta} = \sum_{j=1}^{J} \zeta^j\ \vec{e}_j, \tag{F.4}$$

a number independent of the basis considered (an invariant) and which is their scalar product

$$s(\vec{\xi},\vec{\zeta}) = (\vec{\xi} \bullet \vec{\zeta})_W = \sum_{i,j=1}^{J} W_{ij}\ \xi^i\ \zeta^j \ ; \qquad W_{ij} = \vec{e}_i \bullet \vec{e}_j \ . \tag{F.5}$$

The Euclidian (L_2) norm is defined by

$$s(\vec{\xi},\vec{\xi}) = \sum_{i,j=1}^{J} W_{ij}\, \xi^i\, \xi^j \,. \tag{F.6}$$

For two natural integers p and q different from zero, the contraction of a tensor Z of type (p,q) and order $(p+q)$ in $\mathbf{T}^{p,q}(\mathbf{E})$ is a function $C_{p,q}$ over $\mathbf{T}^{p-1,q-1}(\mathbf{E})$ such that if

$$Z = \vec{u}_1 \otimes \cdots \otimes \vec{u}_p \otimes \vec{v}^1 \otimes \cdots \vec{v}^q \,,$$

then

$$X = C_{p,q}(Z) = s(\vec{u}_p, \vec{v}^q)\, \vec{u}_1 \otimes \cdots \otimes \vec{u}_{p-1} \otimes \vec{v}^1 \otimes \cdots \vec{v}^{q-1} \,. \tag{F.7}$$

The contracted tensor X is of type $(p-1,\, q-1)$ and order $(p+q-2)$. In terms of tensor components, this operation can be represented by

$$X_{i_1, i_2 \ldots, i_p}^{j_1, j_2, \ldots, j_q} = \sum_{k=1}^{J} Z_{i_1, i_2 \ldots, i_{p-1}, k}^{j_1, j_2, \ldots, j_{q-1}, k} \,. \tag{F.8}$$

Some of the properties of tensors are presented. The generalization to more complex tensors is similar to that developed.

1) A dyadic $\ddot{Q}$ is said to be symmetric if $\ddot{Q}^T = \ddot{Q}$, where $\ddot{Q}^T$ is the transposed tensor of $\ddot{Q}$. For a (0,2) type dyadic, the matrix representation is $[\ddot{Q}]^T = [\ddot{Q}]$, and the component relation

$$Q_{ji} = Q_{ij} \,. \tag{F.9}$$

2) The direct product of two contravariant vectors given by (F.4), is a (2,0) dyadic $\ddot{Z}$ constructed following

$$\ddot{Z} = \vec{\xi} \otimes \vec{\zeta} = \sum_{i,j=1}^{J} \xi^i \zeta^j\, \vec{e}_i \otimes \vec{e}_j \,. \tag{F.9a}$$

The matrix representation is $[\ddot{Z}] = [\vec{\xi}]\,[\vec{\zeta}]^T$, and the component relation $Z^{ij} = \xi^i \zeta^j$.

3) Two tensors Z and Q are said to be inverse if

$$Z \bullet Q = Q \bullet Z = I, \tag{F.10}$$

where I is the unit tensor. We will use the notation Z^{-1} to denote the inverse of Z. In terms of components, if for example $\overset{\leftrightarrow}{Z}$ is of type (2,0), and $\overset{\leftrightarrow}{Q}$ of type (0,2) the component relation is

$$\sum_{k=1}^{J} Z_{ik}\, Q^{km} = \delta_i^m, \tag{F.11}$$

where δ_i^m is the Kronecker delta: $\delta_i^m = 1$ if $i=m$ and $\delta_i^m = 0$ if $i \neq m$. In particular, the basis $\{\vec{e}_i\}$ in **E**, has its the reciprocal basis $\{\vec{e}^j\}$ in **E***, such that $\vec{e}_i \bullet \vec{e}^j = \delta_i^j$. The scalar product (F.5) can be simply written if

$$\vec{\xi} = \sum_{i=1}^{J} \xi^i\, \vec{e}_i \qquad ; \qquad \vec{\zeta} = \sum_{j=1}^{J} \zeta_j\, \vec{e}^j. \tag{F.12}$$

Then,

$$s(\vec{\xi},\vec{\zeta}) = \sum_{i=\alpha}^{J} \xi_i\, \zeta^i, \tag{F.13}$$

since $W_i^j = \vec{e}_i \bullet \vec{e}^j = \delta_i^j$.

4) The following relation holds $(\vec{\xi} \otimes \vec{\zeta}) \bullet \vec{u} = \vec{\xi} \otimes (\vec{\zeta} \bullet \vec{u})$.

For example a dyadic $\overset{\leftrightarrow}{Z}$ of type (1,1) can be right-multiplied (dot product) with a vector $\vec{u}$ in **E** to form a vector $\vec{w}$ in **E** by contraction

$$\vec{w} = \overset{\leftrightarrow}{Z} \bullet \vec{u} = \sum_{i,j=1}^{J} \vec{e}_i\, Z_j^i\, u^j. \tag{F.14}$$

The matrix representation is $[\vec{w}] = [\overset{\leftrightarrow}{Z}][\vec{u}]$, and the component relation is

$$w^i = \sum_{j=1}^{J} Z_j^i\, u^j.$$

A dyadic $\overset{\leftrightarrow}{Z}$ of type (0,2) when left-multiplied by $\vec{\xi}$ yields a vector $\vec{v}$ in **E***

$$\vec{v} = \vec{\xi} \bullet \ddot{Z} = \sum_{i,j=1}^{J} \xi^i \, Z_{ij} \, \vec{e}^j \, . \qquad \text{(F.15)}$$

The matrix representation is $[\vec{v}]^T = [\vec{\xi}][\ddot{Z}]$, and the component relation is

$$w_j = \sum_{i=1}^{J} \xi^i \, Z_{ij} \, .$$

A dyadic $\ddot{Z}$ of type (1,1) when right-multiplied by a dyadic $\ddot{Q}$ of type (1,1) yields a dyadic $\ddot{X}$ of same type

$$\ddot{X} = \ddot{Z} \bullet \ddot{Q} = \sum_{i,k=1}^{J} \left(\sum_{j=1}^{J} Z_i^j \, Q_j^k \right) \vec{e}^i \otimes \vec{e}_k \, . \qquad \text{(F.16)}$$

The matrix representation is $[\ddot{X}] = [\ddot{Z} \bullet \ddot{Q}] = [\ddot{Z}][\ddot{Q}]$, and the component relation

$$X_i^k = \sum_{j=1}^{J} Z_i^j \, Q_j^k \, .$$

The dyadic $\ddot{Q} \bullet \ddot{Z}$ is not equal to $\ddot{X}$.

The scalar product product (F.5) is, then, a scalar contraction of the tensor of type (2,2) with components $W_{ik} \xi^l \zeta^m$. The second order tensor $\ddot{W}$ with components W_{ik} is called the fundamental (covariant) metrical tensor. $\ddot{W}$ and its inverse $\ddot{W}^{-1}$ are real and symmetric. A vector $\vec{u}_1$ expressed in E with $\ddot{W}$, has its associated vector $\vec{u}^1$ expressed in E* with $\ddot{W}^{-1}$, via the relation $\vec{u}^1 = \ddot{W} \bullet \vec{u}_1$, and conversely $\vec{u}_1 = \ddot{W}^{-1} \bullet \vec{u}^1$.

Given $\ddot{W}$, use of $W_i^k = \delta_i^k$ and relation (F.15) enables the scalar product (F.5) to be written in an equivalent form

$$s(\vec{\xi},\vec{\zeta}) = \vec{\xi} \bullet \ddot{W} \bullet \vec{\zeta} \, , \qquad \text{(F.17)}$$

for which the Euclidian norm $s(\vec{\xi},\vec{\xi})$ is used in equation (V.2.2). The matrix representation is $s = [\vec{\xi}]^T [\ddot{W}] [\vec{\zeta}]$, and the component relation is given by (F.5). Note that since $\ddot{W}$ is symmetric $\vec{\xi} \bullet \ddot{W} \bullet \vec{\zeta} = \vec{\zeta} \bullet \ddot{W} \bullet \vec{\xi}$.

5) The gradient of a vector $\vec{f}$ in E with respect to a vector $\vec{m}$ in E is a dyadic of type (1,1)

$$\nabla\vec{f} = \sum_{i,j=1}^{J} \frac{\partial}{\partial m^i} f^j \, \vec{e}^i \otimes \vec{e}_j + \sum_{i,j=1}^{J} f^j \, \vec{e}^i \otimes \frac{\partial \vec{e}_j}{\partial m^i} . \tag{F.18}$$

The gradient operator can be viewed as $\nabla = \sum_{i=1}^{J} \vec{e}^i \, \nabla_i$, with $\nabla_i = \partial / \partial m^i$.

We shall use the following tensor indentities derived from the preceding relations. $\ddot{A}$, $\ddot{B}$ are dyadics and $\vec{\xi}$ and $\vec{\zeta}$ are vectors.

$$\nabla(\ddot{A} \bullet \ddot{B} \bullet \vec{\xi}) = \nabla\ddot{A} \bullet \ddot{B} \bullet \vec{\xi} + (\nabla\ddot{B} \bullet \vec{\xi}) \bullet \ddot{A}^T + \nabla\vec{\xi} \bullet \ddot{B}^T \bullet \ddot{A}^T , \tag{F.19}$$

$$\nabla(\vec{\xi} \bullet \ddot{A} \bullet \vec{\zeta}) = \nabla\vec{\xi} \bullet \ddot{A} \bullet \vec{\zeta} + (\nabla\ddot{A} \bullet \vec{\xi}) \bullet \vec{\zeta} + \nabla\vec{\zeta} \bullet \ddot{A}^T \bullet \vec{\xi} . \tag{F.20}$$

In general, for a vector $\vec{\xi}$, and two arbitrary tensors A and B, the operation $A \bullet \vec{\xi} \bullet B$ is ambiguous since

$$(A \bullet \vec{\xi}) \bullet B \neq A \bullet (\vec{\xi} \bullet B) . \tag{F.21}$$

Therefore, if such an operation arises, we shall always put between parenthesis which dot product is computed first (see F.19 and F.20).

Since tensor elements are ordered (eg. F.3), we shall code this order following the alphabetical order of the tensor indices. For example a tensor with component G_i^j has its corresponding dyad $G_i^j \, \vec{e}^i \otimes \vec{e}_j$, whereas the tensor with component G_n^m has its corresponding dyad equal to $G_n^m \, \vec{e}_m \otimes \vec{e}^n$.

6) A change of coordinates from a f^j-system of dimension M to a m^k-system of same dimension, is defined by a transformation of the form

$$f^j = f^j(m^k) \quad ; \quad k=1,M . \tag{F.22}$$

Suppose that the components of a second order (0,2) tensor $\ddot{W}$ are known in the f^j-

system, then the corresponding components in the m^k-system, are (Simmonds, 1982)

$$\tilde{W}_{il} = \frac{\partial f^j}{\partial m^i} W_{jk} \frac{\partial f^k}{\partial m^l}, \tag{F.23}$$

or in tensor form,

$$\tilde{\vec{W}} = (\nabla \vec{f}) \bullet \vec{\vec{W}} \bullet (\nabla \vec{f})^T, \tag{F.24}$$

where the tilde denotes the tensor in the m^k-system, the differentiation being with respect to $\vec{m}$.

We shall assume that equation (F.24) remains valid when going from a subspace **U** in **E** with a f^j-system of dimension N to a subspace **V** in **E** with a m^k-system of dimension M, the transformation being of same form as (F.22).

V.9. APPENDIX G: Two examples of analytical interfaces

Each interface is approximated by a cubic spline interpolation scheme. The approximation is tested by super-imposing the exact interface with the approximate. The number of points required in the interpolation is chosen so that the fit is good and reliable. This requires the exact interface to be at least of class C^0: The interpolation is well behaved, provided the number of points is large enough, for elementary functions that are continuous. Two interface types are considered.

1) The simplest parametrization of an interface is a flat tilted interface. For interface i, the form of the function is

$$z_i = (x - x_f) \tan(\Phi) + z_f . \tag{G.1}$$

Inputs are a fixed point on the interface (x_f, z_f), determined by a priori information (eg. from a zero offset VSP). Other inputs could be a minimum and a maximum allowable

depths, z_{min} and z_{max}. These depths can be used to construct for example "smooth" pinchouts, wedges or faults. If $z_i \geq z_{max}$, then we set $z_i = (1-\varepsilon)z_{max}$. Similarly, if $z_i \leq z_{min}$, we would have $z_i = (1+\varepsilon)z_{min}$. Typical values of ε are on the order of 0.001.

The parameter, Φ, is the angle of the interface with respect to the x axis. The origin of angles is the x-axis (i.e. Φ=0 corresponds to a horizontal flat interface). An angle is positive if the rotation of the interface with respect to the y axis is toward the z axis (clockwise rotation). Figure 3a shows an example of such an interface.

2) The second type of interface models a reef. It is a continuous function constructed by matching five functions at their end points. Inputs are again a fixed point (x_f, z_f) at the interface, but off the reef. Five parameters describe the interface:

(1) $S = x$ translation with respect to x_f of the reef base

(2) $B = x$ length of large base of the reef

(3) $b = x$ length of small base of the reef

(4) $L = x$ length of reef transition region close to x_f

(5) $h =$ reef elevation with respect to z_f .

Set

$S = x_1 - x_f$

$B = x_4 - x_1$

$b = x_3 - x_2$

$L = x_2 - x_1$.

Then, for interface i, the function z_i is constructed following (Figure 3b)

$$
\begin{array}{ll}
x \le x_1 & z_i = z_f \\
x_1 < x < x_2 & z_i = z_f - \frac{h}{2}\left(1 - \cos[\pi \frac{x - x_1}{L}]\right) \\
x_2 \le x \le x_3 & z_i = z_f - h \\
x_3 < x < x_4 & z_i = z_f - \frac{h}{2}\left(1 - \cos[\pi \frac{x - x_3}{x_4 - x_3}]\right) \\
x \ge x_4 & z_i = z_f \ .
\end{array}
\qquad \text{(G.2)}
$$

Mathematical constraints are that $|B| > |L| + |b|$ and $L \neq 0$, and in addition, if $S < 0$, then $|B| > -S$. These constraints are imposed under the penalty model described in section V.3, with equations similar to (3.5-6).

Combination of two reef interfaces leads to more complicated geometry. For example, two reef interfaces with same parameters an inputs but opposite elevation, h, could model an inclusion. Or setting b larger than the model size would simulate a normal fault.

More realistic or complex analytical functions can be defined if necessary. For example combination of these two interfaces would yield a tilted reef, or a reef with different slopes on each side etc. A library of typical interface types should be available in order to adapt the inversion to general interfaces.

Coordinates of the fixed point on the interface could be introduced as constrained parameters. This would allow more flexibility and generality of the inversion, increasing however computation time.

V.10. APPENDIX H: Simple example on the Gauss-Newton approximation

We shall consider here the case where ε is a scalar function of a scalar parameter m. The same notation than the multi-dimensional problem will be taken, to conserve symbolically the more general case.

Assume that $d - \sum_{n=0}^{3} a_n m^n$, is a good approximation of the function $\varepsilon(m)$, in the range of interest, that is, not too far from the minimum, at m^*. The a_n are real coefficients. We could extend the summation to higher orders if this approximation is not sufficient without invalidating what follows. Therefore,

$$\varepsilon = d - \sum_{n=0}^{3} a_n m^n . \tag{H.1}$$

From (2.2) we have $E = \varepsilon\, \varepsilon / 2$, with W=1. Then from (4.5)

$$G = G^T = -\nabla\varepsilon = \sum_{n=1}^{3} n\, a_n\, m^{n-1} . \tag{H.2}$$

Thus, (4.7) yields

$$E' = \nabla E = - G^T \varepsilon , \tag{H.3}$$

and from (4.8) and (4.9)

$$\nabla G^T = \sum_{n=2}^{3} (n-1) n\, a_n\, m^{n-2} , \tag{H.4}$$

$$H = E'' = \nabla\nabla E = G^T G - \nabla G^T \varepsilon . \tag{H.5}$$

The Gauss-Newton approximation assumes $E'' \approx G^T G$, or, equivalently that

$$| G^T G | \gg | \nabla G^T \varepsilon | . \tag{H.6}$$

In terms of series this condition yields

$$\left(\sum_{n=1}^{3} n\, a_n\, m^{n-1} \right)^2 \gg \left| \left(\sum_{n=2}^{3} (n-1) n\, a_n\, m^{n-2} \right) \left(d - \sum_{n=0}^{3} a_n\, m^n \right) \right| , \tag{H.7}$$

with explicit left hand term

$$a_1^2 + 4a_1a_2m + (4a_2^2 + 6a_1a_3)m^2 + 12a_2a_3m^3 + 9a_3^2m^4 , \qquad (H.8)$$

and right hand term

$$|(2a_2 + 6a_3m)(d - a_0 + a_1m + a_2m^2 + a_3m^3)| . \qquad (H.9)$$

In order for (H.6) to be verified one the two following must hold:

(1) the residual must be small, that is $(d - \sum a_n m^n) \rightarrow 0$; or

(2) the residual ε is close to a linear function of the parameter m; that is $a_0 \approx d$, and a_2 and a_3 are much smaller than a_1; this would imply that the function describing the residual energy E is close to a parabola.

Since E'' is the slope of E', this approximation could either underestimate or overestimate the slope of E', depending on the sign of the last term in (H.5), which would slow up the convergence for the zero finding of E' (see Figure 1). Typically, within some regularity conditions, the rate of convergence of Newton's method is quadratic (i.e. $|m_{k+1} - m^*|$ proportional to $|m_k - m^*|^2$), whereas the Gauss-Newton method has linear convergence (i.e. $|m_{k+1} - m^*|$ proportional to $|m_k - m^*|$).

VI. GENERAL CONCLUSION

... no proposition that we put, no axiomatic system, no formal language, is ever final.

- J. Bronowski

1. SUMMARY OF RESULTS

Physical parameters characterizing heterogeneous media have been defined. Conditions of validity of three asymptotic wave methods, (1) dynamic ray tracing, (2) Gaussian beams, and (3) paraxial ray method, were explicitly introduced and derived. Dynamic ray tracing must satisfy four conditions, and Gaussian beams three additional ones. The paraxial ray method is a hybrid method situated between standard ray and Gaussian beams. Amplitudes and phases of these methods are reliable, and can be used with confidence in regions where the validity conditions are satisfied. Green's function for an explosive point source in a medium with constant velocity gradient is then investigated. Validity conditions are tested numerically in this medium. Rate of breakdowns have been established. Dynamic ray tracing is more accurate than Gaussian beams. Both methods remain valid at ray turning points. The paraxial ray is a fast method for computing asymptotic Green's functions. Modeling examples show the robustness and flexibility of the method.

A full waveform nonlinear least squares inversion is then introduced. The forward model is generated by the paraxial ray method. The inversion is set up in the frequency domain, and handles prior information and constraints on parameters. Estimated parameters are layer and interface parameters (delineation). Combined VSP and surface

reflection data, multi-offset VSP or multi-component data, etc., can be simultaneously taken into account. Examples of inversion of VSP and surface reflection data, generated by finite difference method are presented. As an example of application, inversion of VSP field data from Michigan is performed, and medium estimates agree with other independent measures.

2. SUGGESTIONS FOR FURTHER RESEARCH

Numerical (wave and asymptotic wave) methods of solving wave propagation problems in heterogeneous media are being developed at a fast rate. Analytical solution for Green's functions in simple heterogeneous media are very scarce. Green's function in a medium with negative velocity gradient, or a layer over velocity gradient (presence of caustics), or a velocity gradient over layer (shadow zone), is within our reach, and should be investigated. This would allow a more extensive testing procedure of asymptotic wave methods in critical regions. Numerical wave methods would also benefit from these solutions. Green's function for a point force in medium containing a spherical inclusion would extend the testing range. This solution enables one to have direct control on interface curvature, and its interaction with waves. Quantitative estimation of the Gaussian beam generated head waves should be attempted with either exact solutions (Cagniard-de Hoop methods) or a calibrated and tested numerical wave method (for example the finite difference method used in this study). If head waves are correctly taken into account, one could consider hybrid method, where paraxial synthetics are super-imposed to GB-head-waves synthetics. Shear waves is the next step in the analysis, after acoustic (or SH) problems are solved. In more complicated media, validity conditions of asymptotic wave methods can be tested with numerical wave methods. Paraxial ray method can be used in imaging methods with variable

medium background, where a large number of approximate Green's function computations is required.

On the inversion with the paraxial ray method, it is possible to obtain analytic Jacobian matrix for homogeneous layers with smooth interfaces. Density and quality factors estimation is the next step in generalizing the inversion. More analytical interfaces must be available (anticline, overthrust). Addition of coherent noise in time would test further the inversion. Eventually, if the number of parameters is to high, we can attempt to solve sequentially, or by parts the general problem. The development of this inversion would help in the design of a seismic experiment. Given a localized region and its prior information, one would investigate some optimal ways of illuminating this region with waves, and of recording the information. This would improve the resolution and characterization of the region under consideration.

REFERENCES

Abramowitz, M. and I.A. Stegun; *Hanbook of Mathematical Functions*, Dover Publications Inc., New York, 1972.

Acton, F.S; *Numerical Methods That Work*, Harper & Row, Publishers, New York, 1970.

Alford, R.M. ,Kelly, K.R. and Boore, D.M; Accuracy of Finite-Difference Modelling of the Acoustic Wave Equation, *Geophysics, 39*, 834-842, 1974.

Aki, K., and P.G. Richards; *Quantitative Seismology: Theory and Methods (2 volumes)*, W.H. Freeman and Co., San Francisco, 1980.

Alverson, R.C., F.C. Gair and T.F. Hook; Uncoupled equations of motion in nonhomogeneous elastic media. *Bull. Seism. Soc. Amer. 53*, 1023-1030, 1963.

Bass, J; *Cours de Mathématiques (4 volumes)*, Masson Publishing Co., Paris, 1977 (in French).

Baker, B.B. and E.T. Copson; *The Mathematical Theory of Huygen's Principle*, Oxford University Press, 1939.

Bastiaans, M.J; Transport equations for the Wigner distribution function *Optical Acta 26*, 1265-1272, 1979.

Beck, J.V. and K.J. Arnold; *Parameter Estimation in Engineering and Science*, John Wiley & Sons, Inc, New York, 1977.

Bender C.M. and S.A. Orszag; *Advanced Mathematical Methods for Scientists and Engineers*, Mc Graw-Hill Book Co., 1978.

Ben-Menahem, A. and S.J. Singh; *Seismic Waves and Sources* Springer Verlag, 1981.

Blackway, C; One dimensional Full Waveform Inversion of VSP data, Earth Resources Laboratory Internal Report, Massachusetts Institute of Technology, 1985.

Born, M. and E. Wolf; *Principles of Optics*, Macmillan, New York, 1964.

Bouchon, M. and K. Aki; Discrete Wave Number Representation of Seismic-source Wave Fields, *Bull. Seism. Soc. Am, 67*, 259-277, 1977.

Box, M.J.; Bias in Nonlinear Estimation (with discussion), *J. R. Statist. Soc., B 32*, 171-201, 1971

Bremmer, H; The WKB approximation as the first term of a geometric-optical series. *Commun. Pure App. Math 4*, 105-115, 1951.

Carter, J.A. and L.N. Frazer; Accommodating lateral velocity changes in Kirchhoff migration by means of Fermat's principle. *Geophysics, 49*, 46-63, 1984.

References

Cerveny, V., I.A. Molokov and I. Psencik. *Ray methods in seismology*, Univerzita Karlova, Praha, 1977.

Cerveny, V. and F. Hron; The ray-series method and dynamic ray tracing for three-dimensional inhomogeneous media. *Bull. Seism. Soc. Amer., 70*, 47-77, 1980.

Cerveny, V., M.M. Popov and I. Psencik; Computation of wave fields in inhomogeneous media - Gaussian beam approach. *Geophys. J. R. astr. Soc., 70* 109-128, 1982.

Cerveny, V; The application of ray tracing to the propagation of shear waves in complex media, preprint, in book series Seismic Exploration (ed. Treitel and Helbig),1983[a]

Cerveny, V; Synthetic body wave seismograms for laterally varying layered structures by the Gaussian beam method, *Geophys. J. R. astr. Soc.*, 73, 389-426, 1983[b]

Chambadal, L; *Dictionnaire des Mathématiques Modernes*, Larousse Publishing Inc., 1969 (in French).

Chapman, C.H; A New Method for Computing Synthetic Seismograms, *Geophys. J. R. astr. Soc.*, 54, 481-518, 1978.

Chapman, C.H. and R. Drummond; Body-wave seismograms in inhomogeneous media using Maslov asymptotic theory, *Bull. seism. Soc. Am.*, 5277-5917, 1983.

Charnes, A., E.L. Frome and P.L. Yu; The equivalence of Generalized Least Squares and Maximum Likelihood Estimates in the Exponential Family, *Journal of the American Statistical Association*, 71, 169-171, 1976.

Claerbout, J.F; *Fundamentals of Geophysical Data Processing*, Mc Graw-Hill, New York, 1976.

Cormier, V.F. and P. Spudich; Amplification of Ground Motion and Waveform Complexity in Fault Zones: Examples from the San Andreas and Calaveras Faults, *Geophys. J. R. astr. Soc.*, 79, 135-152, 1984.

Cornbleet, S; Geometrical optics reviewed: A new light on an old subject, *Proc. IEEE, 71*, 471-502, 1983.

Corones, J; Bremmer series that correct parabolic approximations. *Jour. Math. Anlys. and Applic., 50*, 361-372, 1975.

Corones, J., B. De Facio and R.J. Krueger; Parabolic approximations to time-dependent elastic wave equation, *J. Math. Phys., 23*, 577-586, 1982.

De Boor, C; *A practical Guide to Splines*, Springer - Verlag, New York, 1980.

Dennis, J.E. and J.J. More; Quasi-Newton Methods, Motivation and Theory, *SIAM Review, 19* 46-89, 1977.

DeSanto, J.A; Relation between the solutions of the Helmholtz and parabolic equations

for sound propagation, *J. Acoust. Soc. Amer. 62*, 295-297, 1977.

Eisenhart, L.P; *A Treatise on the Differential Geometry of Curves and Surfaces*, Ginn and Comp., 1909.

Esmersoy, G; *Multidimensional Velocity Inversion for a Single Wide-Band Point Source*, Ph.D. Thesis, Massachusetts Institute of Technology, 1985.

Felsen, L.B; Complex-Source-Point Solutions of the Field Equations and their Relation to the Propagation and Scattering of Gaussian Beams, *Istituto Nazionale di Alta Matematica, Symposia Mathematica*, Volume 18, 39-56, 1976.

Fishman L. and J.J. McCoy; Derivation and Application of Extended Parabolic Wave Theories. I. The factorized Helmholtz Equation, *J. Math. Phys., 25*, 285-296, 1984.

Haus, H.A; *Waves and Fields in Optoelectronics*, Prentice Hall, 1983.

Hermann, R.B. and C.Y. Wang; A Comparison of Synthetic Seismograms, *Bull. Seism. Soc. Am., 75*, 41-56, 1985.

Hohmann, G; *Inverse Theory*, Course GG628 notes, University of Utah, 1979.

Hook, J.F; Separation of the vector wave equation of elasticity for certain types of inhomogeneous, isotropic media, *Acoust. Soc. Amer. 33*, 302-313, 1961.

Hook, J.F; Green's functions for axially symmetric elastic waves in unbonded inhomogeneous media having constant velocity gradients. *Jour. App. Mechanics, 293-298, 1962.*

Husdon, J.A; A Parabolic Approximation for Elastic Waves, *Wave Motion, 2*, 207-214, 1980.

Jackson, D.D. and M. Matsu'ura; A Bayesian Approach to Nonlinear Inversion, *J. Geophys. Res. 90*, 581-591, 1985.

Karal, K.B. and D.J. Perozzi; Fast Seismic Ray Tracing, *SIAM J. Appl., MATH., 43*, 981-992, 1983.

Karmanov, V.; *Programmation Mathématique*, Mir publications, Moscow, 1975 (in French).

Kelly, K.R., Ward, R.W., Treitel, S., and Alford, R.M; Synthetic Seismograms: A finite-difference approach, *Geophysics*, vol. 41, 2-27, 1976.

Kennett, B.L.N. and J.A. Harding; Is Ray Theory Adequate for Reflection Seismic Modelling Methods, *First Break, 3*, 9-14, 1985.

Klimes, L; The Relation Between Gaussian Bemas ans Maslov Asymptotic Theory, *Studia geophys. geod.*, 1984 (submitted).

References

Kravtsov, Y.A; Two new asymptotic methods in the theory of wave propagation in inhomogeneous media. *Soviet Physics - Acoustics 14*, 1-17, 1968.

Kravtsov, Y.A. and Z.I. Feizulin; Some consequences of the Huygens-Kirchhof principle for a smoothly nonuniform medium, *Izvestiya VUZ, Radiofizika 12*, 886-893, 1969. (Translated by Consultants Bureau, Plenum Publishing Corp. , 1972).

Kravtsov, Y.A. and Y.I. Orlov; Limits of applicability of the method of geometric optics and related problems. *Soviet Physics VSP 23*, 750-762, 1980.

Landers, T. and J.F. Claerbout; Numerical calculations of elastic waves in laterally inhomogeneous media. *J. Geophys. Res., 77*, 1476-1482, 1972.

Landers, T.E; Elastic scattering of plane P-waves at irregular boundaries in the lower crust. *Earthquake Notes 45*, 12-21, 1974.

Larrere, M; Etude de Evenements Pentes Observes sur les Enregistrements de la Sonde Acoustique EVA, Rapport de fin d'Etudes I.S.T., Univ. P.&.M. Curie, Paris, 1984 (in French).

Leontovitch, M.A. and V.A. Fock; Solution of the Problem of Propagation of Electromagnetic Waves Along the Earth's Surface by the Method of Parabolic Equation, *J. Phys. U.S.S.R., 10*, p. 13, 1946.

Lines, L.R. and S. Treitel; Tutorial Review of Least-Squares Inversion and its Application to Geophysical problems, *Geophys. Prosp., 32*, 159-186, 1984.

Lock, M.H.; Axially symmetric elastic waves in an unbounded inhomogeneous medium with exponentially varying properties. *Bull. Seism. Soc. Amer. 53*, 527-538, 1963.

Luenberger, D.G.; *Introduction to Linear and Nonlinear Programming*, Adison-Wesley Publishing Co. Inc., Massachusetts, 1973.

Luneburg, R.K.; *Mathematical Theory of Optics*, University of California Press, Berkeley, 1964.

Madariaga, R.; Gaussian beam synthetic seismograms in a vertically varying medium, *Geophys. J.R. Astron. Soc.*, 79, 589-612, 1984.

Madariaga, R. and Papadimitriou, P.; Gaussian beam modeling of upper mantle phases, Anales Geophysicae, 1984 (submitted).

Madden, T. and K. Aki; Inverse Problems, Course 12.53 notes, Massachusetts Institute of Technology, 1980.

Marfurt, K; Elastic Wave Equation Migration-Inversion, Ph.D. Thesis, Columbia University, 1978.

McCoy, J.J; A Parabolic Theory of Stress Wave Propagation Through Inhomogeneous

Linearly Elastic Solids, *J. Appl. Mech., 44*, 462-468, 1977.

Mellen, M.H; Offset Vertical Seismic Profiling: Two-Dimensional Forward Modeling with Asymptotic Ray Theory, S.M. Massachusetts Institute of Technology , 1984.

MINPACK-1 Fortran package, by J.J Moré , B.S. Garbow and K.E. Hillstrom, Argonne National Laboratory, Illinois, report ANL-80-74, 1980.

Morse, P.M. and H. Feshbach; *Methods of Theoretical Physics (2 Volumes)*, Mc Graw-Hill Book Co., 1953.

Morse, P.M. and K.U. Ingard; *Theoretical Acoustics*, Mc Graw-Hill Book Co., 1968.

McDaniel, S.T.; Parabolic approximations for underwater sound propagation. *J. Acoust. Soc. Amer. 58*, 1178-1185, 1975.

Nabelek, J.L; Determination of Earthquake Source Parameter from Inversion of Body Waves, Ph.D. Thesis, Massachusetts Institute of Technology, 1984.

Nagel, A. and E.M.Stein; Lectures on pseudo-differential operators. Princeton Univ. Press, Princeton, N.J., 1979.

Nazareth, L.; Some Recent Approaches to Solving Large Residuals Nonlinear Least Squares Problems, *SIAM Review, 20*, 1-11, 1980.

Nowack, R. and K. Aki; The 2-D Gaussian beam synthetic method: testing and application *J. Geophys. Res. 89*, 7797-7819, 1984

Oppenheim, A.V. and R.W. Schafer; *Digital Signal Processing*, Prentice-Hall Inc., New Jersey, 1975.

Palmer, D.R; Eikonal approximation and the parabolic equation, *J. Acoust. Soc. Amer. 60*, 343-354, 1976.

Palmer, D.R; A path-integral approach to the parabolic approximation, *J. Acoust. Soc. Amer. 66*, 862-871, 1979.

Paternoster, B; Effects of Formation Boundaries on Full-Waveform Acoustic Logs, S.M. Massachusetts Institute of Technology , 1985.

Pekeris, C.L; The Theory of Propagation of Sound in a Half Space of Variable Sound Velocity Under Conditions of Formation of a Shadow Zone, *J. Acoust. Soc. Amer. 18*, p. 295, 1946

Polyanskii, E.A; Relationship between the solutions of the Helmholtz and Schrödinger equations, *Soviet Phys. Acoust. 20*, p. 90, 1974.

Popov, M.M; A new method of computation of wave fields using Gaussian beams. *Wave-motion 4*, 85-97, 1982.

References

Prange, M.D; *A Numerically Stable Method for Computing High Frequency Sub-Surface Seismograms in Layered Media using the Discrete Wavenumber Method*, Earth Resources Laboratory Internal Report, Massachusetts Institute of Technology, 1985.

Psencik, I; Program for the computation of kinematic and dynamic parameters of seismic body waves in 2-D laterally inhomogeneous media with curved interfaces, in Programs for the Interpretation of Seismic Observations, Vol. 3, Nauka, Leningrad, 1983 (in Russian).

Roddier, F; *Distributions et Transformation de Fourier*, Mc Graw-Hill Inc., Paris, 1978 (in French).

Sandell, N.R. and J.H. Shapiro; Stochastic Processes and Applications, Course 6.432 notes, Massachusetts Institute of Technology, 1976.

Savarensky, E; *Seismic Waves*, Mir publications, Moscow, 1975.

Scott, P. and D. Helmberger; Applications of the Kirchhoff-Helmholtz integral to problems in seismology. *Geophys. Jour. astr. Soc. 72*, 237-254, 1983.

Simmonds, J.G; *A Brief on Tensor Analysis*, Springer - Verlag, New York, 1982.

Singh, S.J. and A. Ben-Menahem; Asymptotic theory of body waves in a radially heterogeneous earth. *Bull. Seism. Soc. Amer. 59*, 2039-2059, 1969.

Stephen, R.A; Finite-Difference Seismograms for Laterally Varying Marine Models, *Geophys. Jour. astr. Soc., 79*, 185-198, 1984.

Stewart, R.R; Vertical Seismic Profiling: The One Dimensional Forward and Inverse Problems, Ph.D. Thesis, Massachusetts Institute of Technology, 1983.

Tappert, F.D; The Parabolic Approximation Method, *Wave Propagation in Underwater Acoustics*, Lecture Notes in Physics No. 70, edited by J.B. Keller and J.S. Papadakis, Springer-Verlag, New York, 1977.

Thomson, C; Ray-Theoretical Amplitude Inversion for Laterally Varying Velocity Structure Below NORSAR, *Geophys. Jour. astr. Soc., 74*, 525-558, 1983.

Titchmarsh, E.C; *Introduction to the theory of Fourier integrals*, Oxford, 1967.

Toksoz, M.N. and D.J. Johnston; *Seismic Wave Attenuation*, Geophysics reprint series No. 2, Society of Exploration Geophysicists, 1981.

Van Trees, H.L; *Detection, Estimation, and Modulation Theory*, John Wiley and Sons, New York, 1968.

Vygodski, M; *Aide-mémoire de mathématiques supérieures*, Mir publications, Moscow, 1975 (in French).

References

Wales S.C. and J.J. McCoy; A comparison of Parabolic Wave Theories for Linearly Elastic Solids, *Wave Motion*, *5*, 99-113, 1983.

Weyl, H; *Space Time and Matter*, Dover Publications, Inc., 1952.

White, J.E. *Underground Sound*, Elsevier Pubications, Amsterdam, 1983.

Yechiam, Y. and J. Pearl; Asymptotic Properties of Discrete Unitary Transforms, *IEEE Trans. Pattern Analys. Machin. Intell.*, vol. PAMI-1, 366-371, 1979.

Yomogida, K; Gaussian Beams for Surface Waves in Laterally Slowly-Varying Media, *Geophys. J. R. astr. Soc., 1984* (submitted).

Yomogida, K. and K. Aki; Toatal Waveform Synthesis of Surface Waves in Laterally Heterogeneous Earth by Gaussian Beam Method *J. Geophys. Res., 1984 (submitted).*

Yosiyama, R; Elastic waves from a point in an isotropic heterogeneous sphere, Part 2, *Bull. Earthq. Res. Inst. 18*, 41-56, 1940.

THESIS EXAM COMMITTEE

Keiiti Aki (USC)

Vernon Cormier (ERL, MIT)

Gregory Duckworth (EAPS, MIT)

Bernard Levy (EE, MIT)

Theodore Madden (chairman)

M. Nafi Toksöz (advisor)