CWP-357
December 2000

Estimation of Multi-valued Green's Function by Dynamic Ray Tracing and True Amplitude Kirchhoff Inversion in 3-D Heterogeneous Media

Lan Wang

— PhD Thesis —
Mathematical and Computer Sciences

Center for Wave Phenomena
Colorado School of Mines
Golden, Colorado 80401
303/273-3557

ABSTRACT

Implementation and application of amplitude-preserving prestack Kirchhoff depth migration/inversion is of great interest for the detection of hydrocarbons (petroleum and natural gas) in the oil industry. 3-D prestack depth migration is considered as the ultimate imaging technique for complex geological structures. The analysis of amplitudes on recorded seismic data has contributed a great deal to exploration success and will become increasingly important in the near future. This thesis aims to be a contribution in that direction.

Ray-theoretic modeling requires accurate amplitude as well as phase both for forward modeling and Kirchhoff inversion, among other applications. There are no analytical solutions to the ray equations in realistic earth models; thus, we must use numerical solutions to solve problems of interest. For 3-D applications, it is a challenge to develop numerical modeling codes that require reasonable CPU time while achieving sufficient amplitude accuracy to be useful in applications. This thesis describes a ray-theoretical modeling method that produces all the quantities needed in the Kirchhoff inversion.

The concern of efficiency in 3-D application motivates to examine the use of analytical ray tracing techniques on tetrahedral grids. Here, a discussion of validity and accuracy of dynamic ray tracing on a tetrahedral model versus a Cartesian gridded model is included. For the case of linear sloth (slowness squared or inverse wavespeed squared), analytical solutions of the ray equations *do* exist, leading to an efficient combined numerical-analytical technique in tetrahedral models. However, the wavespeed model generated by this technique is not sufficiently smooth to produce numerically accurate amplitudes. Recent attempts to further smooth the physical model defeat the advantage of speed of the algorithm because the smoothness conditions across the faces of the tetrahedra generate a coupled system of equations of a size proportional to the number of tetrahedra in the global physical model. This is not practical in 3-D. Thus, a standard smoothed physical model on a Cartesian grid is likely to lead to a computer code of competitive CPU speed, when amplitude accuracy—dynamics—is of as much concern as traveltime accuracy—kinematics.

The application of smoothed velocity models presents a new challenge to the dynamic ray tracing, namely, how to incorporate the transmission coefficients of the "true interface" of the background model into the Greens functions. Here, a new technique was devised. In that technique, a transmission coefficient is introduced at each of the course grid steps of the Runge-Kutta method. Thus, in regions of rapid change in the smoothed model–representing the interfaces of the original model–the transmission

coefficient of the discontinuous model is built up cumulatively, through the transition region. In regions of more slowly varying velocity, the cumulative effect is nearly equal to a factor of unity. Numerical tests against analytical models indicate that this method provides adequately accurate transmission effects in the forward propagation direction needed for the ultimate Kirchhoff inversion program.

Another feature adopted in the forward modeling is the wavefront construction technique. In that method, the size of triangular plates connecting three nearby rays on the isochron (surface of constant traveltime) is used as an indicator of adequate density of rays. When at least one of the criteria for density of rays are violated, data at new points on the wavefront are interpolated into the family of rays and the wavefront construction continues. In this manner, the method does not require excessive density of rays at small traveltimes in order to maintain adequate density of rays at larger traveltimes. The resulting Green's functions are always defined on a fine enough wavefront grid to allow for acceptably accurate interpolation onto the Cartesian grid needed for the Kirchhoff inversion. The technique also permits multipathing (caustics) and computes amplitude propagation along each of the branches of the wavefront.

Sample implementations of the dynamic ray tracing algorithm are shown. Three examples are included, with each indicative of a situation that is common in applications. The accuracy of the smoothing procedure is tested by application on a horizontal-interface model with strong velocity contrast, where it is shown that the amplitude and traveltime adequately match an analytical solution of the discontinuous-model problem in the forward propagating direction. Also addressed is the problem of multi-pathing and caustics and we show that the method produces amplitude and phase on all arrivals at depth. The final example is taken from a section of the SEG/EAGE salt dome model. Here, the method properly characterizes the sparse penetration of rays into the salt dome, itself, as well as the transmission below.

Three-dimensional (3-D) prestack depth migration is an accurate model-consistent approach for imaging media with complex structures. The ability to localize processing to target zones gives Kirchhoff migration some CPU advantage over alternative 3-D prestack migration algorithms. Furthermore, other prestack depth migrations do not have demonstrated extensions to inversion. Thus, Kirchhoff prestack depth migration in 3-D is a relatively popular choice for heterogeneous media applications.

True amplitude Kirchhoff inversion is aimed at performing two tasks: to estimate angular dependent reflection coefficients, and to calculate the corresponding reflection angles from seismic data. The approach of Bleistein-Cohen is a high-frequency inverse method that combines the acoustic Kirchhoff theory with the WKBJ approximation and the multi-dimensional stationary phase method. Here, the weight functions in Kirchhoff inversion are recast as ray-theoretic expressions that are well-suited for an efficient computer implementation that uses the dynamic ray tracing. It was found

that the proposed method in the open literature for computing the Beylkin determinant for common offset data is based on an erroneous assumption about the relative orientation of two polarization directions at the source and receiver points. Polarization arises out of a method that reduces the 3D computations into 2D computations on locally plates orthogonal to the rays from the source and receiver to the output point. Thus, the author re-derived the reduction process necessary for the computation of the Beylkin determinant.

The true amplitude Kirchhoff inversion is applied to 3-D modeled data. The applications show that angle-dependent reflection coefficients are accurately estimated. The corresponding reflection angles are accurately estimated from seismic data by a technique that uses two simultaneous diffraction stacks.

TABLE OF CONTENTS

GLOSSARY OF MATHEMATICAL SYMBOLS

$\boldsymbol{x}, \boldsymbol{x}_0, \boldsymbol{y}$	subsurface position
$v(\boldsymbol{x})$	medium velocity
$\tau(\boldsymbol{x}, \boldsymbol{x}_0), t, \vartheta$	traveltime
$A(\boldsymbol{x}, \boldsymbol{x}_0)$	amplitude
σ	monotonic parameter along ray, with $d\tau = d\sigma/v^2$
(q_1, q_2, q_3)	ray-centered coordinates
$(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$	unit vectors in ray-centered coordinate system
$(\gamma_1, \gamma_2, \gamma_3)$	ray parameter coordinates
$\boldsymbol{p} = (p_1, p_2, p_3)$	slowness vector
J	Jacobian of transformation from $(\gamma_1, \gamma_2, \gamma_3)$ to (x_1, x_2, x_3)
Ω	ray
Σ_i	interface
$S, R, P_i, \tilde{P}_j$	points along the ray
$T(P_k)$	ray-theoretic transmission coefficient
$\mathcal{T}^c$	complete transmission coefficient
$\boldsymbol{Q}$	transform from (γ_1, γ_2) to (q_1, q_2)
$\boldsymbol{P}$	transform from (γ_1, γ_2) to (p_1, p_2)
$\hat{\boldsymbol{H}}$	matrix representation of $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ in Cartesian coordinates
$\boldsymbol{M}$	second derivatives of the traveltime field with respect to (q_1, q_2)
$\varphi(R, S)$	KMAH index from S to R
$q(\boldsymbol{x})$	quadratic sloth (slowness square) function
$f_k(\sigma)$	basis functions of ray trajectory in Laplace domain (see Section 2.3)
$g_k(\sigma)$	$\mathrm{d}f_k(\sigma)/\mathrm{d}\sigma$, basis functions of slowness vector in Laplace domain (see Section 2.3)
$\boldsymbol{x_s}$	source position
$\boldsymbol{x_g}$	receiver position
h	offset
$F(\omega)$	source spectrum, Fourier transform of $f(t)$
$c(\boldsymbol{x})$	background wave sound speed above the reflector
$\alpha(\boldsymbol{x})$	perturbation of slowness squared field
$\boldsymbol{\xi} = (\xi_1, \xi_2)$	parameters labeling source/receiver points
$\phi(\boldsymbol{x}, \boldsymbol{\xi})$	total traveltime from source to scattering point to receiver
$a(\boldsymbol{x}, \boldsymbol{\xi})$	product of WKBJ amplitude
$R(\boldsymbol{x})$	ray-theoretic reflection coefficient
$w(\boldsymbol{x}, \boldsymbol{\xi})$	spatial weighting factor
φ, μ	phase of complex amplitude
$g(\boldsymbol{x}, \boldsymbol{x}_0, \omega)$	WKBJ Green's function

$u_I(\boldsymbol{x_g}, \boldsymbol{x_s}, \omega)$	incident wavefield
$u_S(\boldsymbol{x_g}, \boldsymbol{x_s}, \omega)$	scattered wavefield
θ	source/receiver rays open angle
θ_s	angle between source ray and reflector normal
τ_s, τ_g	traveltime, source to reflector, receiver to reflector
$\beta(\boldsymbol{x})$	reflectivity function
$h(\boldsymbol{y}, \boldsymbol{\xi})$	Beylkin determinant
$\hat{\Gamma}_s^M$	derivatives of source position with respect to the surface parameters (ξ_1, ξ_2)
$\hat{\Gamma}_g^M$	derivatives of receiver position with respect to the surface parameters (ξ_1, ξ_2)
DRT	dynamic ray tracing
KRT	kinematic ray tracing
FD	finite-difference
PSDM	prestack depth migration
AVO	amplitude variation with offset
AVA	amplitude versus angle
ART	asymptotic ray theory
WF	wavefront
WFC	wavefront construction

ACKNOWLEDGMENTS

This work was partially supported by the Office of Naval Research, Applied Analysis Division (grant number N00014-91-J-1267) and Ocean Acoustics Division (grant number N00014-95-1-0508). This work was also partially funded by the sponsors of the Consortium Project on Seismic Inverse Methods for Complex Structures at the Center for Wave Phenomena, Colorado School of Mines.

Chapter 1

INTRODUCTION

The goal of seismic migration/inversion is to estimate the subsurface structure and medium parameters from observations at the upper surface. Such inversion processes are closely related to forward modeling. **Forward modeling** including scattering of waves uses the **wave equation** to synthesize the earth's response for a given set of model parameters. On the other hand, in the **inverse scattering problem**, the scattered wavefield is observed at the boundary surface of the earth, and the unknowns to be determined are the material parameters and their discontinuity surfaces—**reflectors**. In the field of seismic exploration, such an inverse scattering problem is called **migration** or **imaging** — an inversion operation involving rearrangement of seismic information elements so that reflections and diffractions are plotted at their true locations. In this thesis, I address both the forward and inverse scattering problems.

1.1 Estimation of Green's function by dynamic ray tracing

Accurate and efficient forward modeling is a key part in 3-D prestack migration and inversion. The forward modeling can be done in several ways, with the wave equation being the common starting point for all the methods. In finite-difference, finite-element or spectral methods (Kelly *et al.*, 1975; Mitchell & Griffiths, 1980; Farra, 1990), the wave equation is solved directly. This class of methods is computationally slow, but generally takes all wave modes into account. At the other extreme, "first arrival" methods such as the finite-difference solution of the eikonal equation (Vidale, 1988; Podvin & Lecomte, 1991; Sethian & Popovici, 1999) or shortest-path methods using graph theory (Saito, 1989; Klimes & Kvasnicka, 1994) have been introduced. However, these methods suffer from the disadvantage that they are restricted to the computation of first arrivals only and they produce unreliable amplitudes.

In **ray tracing** methods (Červený, 1987; Hubral, 1980; Červený, 1995; Virieux & Farra, 1991), new differential equations are derived from the wave equation under the high-frequency approximation. These methods are event oriented, which means that the wave mode to be computed (e.g., a primary reflection from a specific horizon) is an input parameter to the method. If just one (or a few) wave modes are wanted, the ray-

tracing methods are generally superior to the more comprehensive finite-difference, finite-element or spectral methods in terms of CPU efficiency. They are also superior to the first-arrival methods in terms of providing multiple arrivals, amplitude and phase. Ray-tracing methods provide a reasonable compromise between accuracy and computational efficiency among various seismic modeling methods. Thus, I apply the ray tracing in the estimation of all the wavefield model data needed for 3-D Kirchhoff inversion.

Estimation of the ray data can be carried out either by numerical solution of ray-tracing equations in general **smooth grid-based models** or by piecewise analytic solutions for certain simple velocity functions in **tetrahedral models**. In the method of analytic ray tracing, or cell ray tracing, the whole medium is divided into suitable cells (usually, tetrahedra in 3-D), in which the velocity can be approximated by simple functions that permit analytic ray solutions. The ray in the whole model is then obtained as a chain of analytically computed segments. On the other hand, in the method of numerical ray tracing, general smooth gridded models (possibly separated by smooth interfaces) are assumed. The estimation of the ray path, traveltime, amplitude and other ray data is based on numerically solving a system of first-order partial differential equations. For the 3-D Kirchhoff inversion, generating the Green's functions for all source/receiver points and all subsurface output points is computer-intensive, even when the computations are only carried out on a subset of those points with data generated at every grid point through interpolation. This leads to examming the possibility of applying analytic solutions in tetrahedral models for amplitude estimation. That method provides fast ray tracing, especially for linear velocity models in each tetrahedron. Here, the discussion of the two categories of ray tracing comes to the conclusion that the analytic approach in tetrahedral cells does not likely offer efficient algorithms in dynamic 3-D applications (Wang & Bleistein, 1998), where accurate amplitudes as well as traveltimes are required. Despite rejecting this method, I studied it in enough detail to re-derive a form of solution that is significantly simpler and computationally faster than the one proposed in the literature.

1.2 Wavefront construction on dynamic ray tracing

Where a fast solution of all ray contributions to a set of receivers from a shot point is desired, ray tracing is efficiently realized by the wavefront construction technique. Wavefront construction was first introduced by Vinje *et al.* (1993). Extensions and modifications on this technique include (Chilcoat & Hildebrand, 1995; Vinje *et al.*, 1996; Meng & Bleistein, 1997; Lucio *et al.*, 1996). In contrast to classical ray shooting, where single rays are propagated through the medium, wavefront construction is ray tracing using a wavefront oriented parameterization. Wavefronts consisting of rays arranged in a triangular network are propagated stepwise through the model. At each

time step, the differences in a number of parameters between each pair of rays on the wavefront are checked. New rays are interpolated whenever this difference between pairs of rays exceeds some predefined maximum value. A controlled sampling of the wavefront at all time steps is thus obtained. Ray data at receivers are given by linear interpolation between every two successive wavefronts. Multiple-event values can be obtained when different sequences of wavefronts pass the same receivers. The strength of the wavefront construction method is that it is robust and efficient. It is suitable for various kinds of ray-tracing methods and velocity distribution representation—either tetrahedral models or smooth grid-based models. In this thesis, I develop a new 3-D forward modeling module that uses physical model smoothing, dynamic ray tracing and wavefront construction.

1.3 Common-offset Kirchhoff inversion

Implementation and application of amplitude preserving prestack Kirchhoff depth migration/inversion is of great interest for the detection of hydrocarbons (petroleum and natural gas) in the oil industry. The use of 3-D migration is motivated strongly by 3-D complex model studies, and three-dimensional prestack depth migration (PSDM) is an accurate (model-consistent) approach for imaging media with complex structures. The ability to localize processing to target zones gives Kirchhoff migration some CPU advantage over alternative 3-D PSDM algorithms. Furthermore, other PSDM's do not have demonstrated extensions to inversion. Thus, Kirchhoff PSDM in 3-D is a relatively popular choice for heterogeneous media applications.

The pioneer works on incorporating the true amplitudes of arbitrary medium in the Kirchhoff diffraction summation are Newman (1975), Cohen and Bleistein (1979), Clayton and Stolt (1981), Tarantola (1984), Berkhout (1985), Cohen *et al.* (1986), Bleistein (1986; 1987) and Bleistein *et al.* (1987). The significance of these papers is that the inversion formulas give not only a reflector map but also an estimate of the angularly dependent reflection coefficient on the reflector. The approach of Bleistein-Cohen is a high-frequency inverse method that combines the acoustic Kirchhoff theory with the WKBJ approximation and the multi-dimensional stationary phase method. This approach provides a bridge between the classical migration techniques and more mathematically sophisticated high frequency techniques based on the generalized Radon transform (Beylkin, 1985; Miller *et al.*, 1987; Beylkin & Burridge, 1990).

The dependence of the Bleistein-Cohen method on the structure of the operator makes the approach significantly different from Kirchhoff migration in the interpretation of the output. The difference is in the spatial weighting which leads to an output proportional to the reflection coefficient in the inversion formula. In addition, based on the use of the Kirchhoff approximation as a starting point for deriving this

formula, Bleistein-Cohen (2000) call it **Kirchhoff inversion**. In the seismic exploration community, this is also called **amplitude preserving prestack Kirchhoff migration** or **true amplitude Kirchhoff migration**. Throughout this thesis, I will use the term Kirchhoff inversion and refer to the approach of Bleistein-Cohen as Bleistein-Cohen inversion. Many recent works on Kirchhoff migration/inversion for a hierarchy of more difficult background velocities and source/receiver configurations have been carried out; see Docherty (1987a; 1991), Dong *et al.* (1991), Schleicher *et al.* (1993), Hubral *et al.* (1991; 1996), Tygel *et al.* (1993; 1996), Hanitzsch (1995; 1997), Ou (1995), Xu (1996), Sun and Gajewski (1997; 1998), Gray (1997), Albertin *et al.* (1999). In this thesis, I focus on common-offset geometry and 3-D general background medium.

One difficulty in the implementation of Kirchhoff inversion for the common-offset configuration and general media is the estimation of the Beylkin determinant in the Bleistein/Cohen inversion formula (Bleistein *et al.*, 1987; Bleistein, 1987; Bleistein *et al.*, 2000). Červený and de Castro (1993) proposed a method for computing the Beylkin determinant by dynamic ray tracing. Hanitzsch (1995) has applied this approach to 2-D amplitude-preserving Kirchhoff migration/inversion and found valuable applications in amplitude variation with offset (AVO) and amplitude versus angle (AVA) analysis. Here, I modify Červený and de Castro's approximation, and apply it in 3-D heterogeneous structures. This method does not require the numerical measurements of any elementary area connected with the ray field or the determination of the derivatives of traveltimes, which play an important role in some other methods that have appeared in the recent literature, such as (Schleicher *et al.*, 1993; Sun & Gajewski, 1998).

1.4 Synthetic data sets

For common-offset Kirchhoff inversion, the computer process is similar to a traditional kinematic Kirchhoff migration algorithm. The challenges arise from the computation of the accurate amplitude 3-D WKBJ Green's function and the spatial weighting factor for such common offset data. Here, I use the dynamic ray tracing code to produce all the needed quantities. Two synthetic examples are tested by applying the common-offset Kirchhoff inversion algorithm. Positioning and amplitude accuracy are tested through these examples. Amplitudes in migrated sections are estimates of angle-dependent reflection coefficients, provided proper weights are applied in the diffraction stack migration to correct for amplitude loss due to geometrical spreading. The Bleistein-Cohen inversion provides those proper weights. The technique of double Kirchhoff inversion, that is, using two different weights, produces fairly accurate estimates of reflection angles from seismic data. Therefore, this technique is particularly well suited for an AVA analysis.

1.5 Content overview

Chapter 2 starts with the estimation of Green's functions, which play an important role in the inversion procedure. First I discuss the possibility of applying analytic solutions in tetrahedral models for amplitude estimation, and conclude that this approach cannot provide us accurate and reliable amplitudes. Instead, I develop an efficient and flexible 3-D forward modeling module that uses dynamic ray tracing and wavefront construction methods on a general smooth velocity field. Technique and implementation details are included in this chapter. Since I use a general smooth gridded velocity model which does not handle the transmission coefficients resulting from the layered models, I propose a method of re-introduction of transmission coefficients into the Green's function solution of the smoothed medium. A discussion of Maslov asymptotics at caustics is also included in the last section of this chapter.

Sample implementations of the dynamic ray-tracing algorithm are presented in Chapter 3. Three examples are included, each example being indicative of a situation that is common in applications.

Chapter 4 is devoted to the theory of true amplitude Kirchhoff inversion. The weights are recast into ray-theoretic expressions that are well suited for an efficient computer implementation that uses dynamic ray tracing. The proper weighting function for the common-offset configuration and 3-D acquisition dimension is given in this special form for computer implementation.

I show details of the implementation of the inversion theory in Chapter 5. The true amplitude Kirchhoff inversion accurately estimates angle-dependent reflection coefficients in modeled data. Examples of inversion on synthetic data are also shown in this chapter .

Finally, I show benchmark results in Chapter 6.

1.6 Software developed

For this research, I have developed some 3-D true amplitude Kirchhoff inversion algorithms and software. The algorithms and software include: 3-D dynamic ray tracing and 3-D inversion codes. Both include new results developed in this project and mentioned briefly above. The software has been tested using modeled and field data.

Chapter 2

ESTIMATION OF 3-D MULTI-VALUED GREEN'S FUNCTION BY RAY TRACING

2.1 Introduction

Accurate and efficient forward modeling is a key part of 3-D prestack migration and inversion, among other applications. In this chapter, I describe a ray-theoretical modeling method that uses smoothing of models, dynamic ray tracing and wavefront construction. All the quantities in the Kirchhoff inversion formula can be determined by that method.

Here, I address the problem of accurate and efficient determination of multi-valued 3-D maps for traveltimes and amplitudes, as well as other ray-related variables throughout the target zone from any shot and receiver position. The current interest in 3-D seismic imaging has considerably increased the importance of 3-D ray-tracing methods in wave-field computations. Among seismic modeling methods, ray-tracing methods provide a reasonable compromise between accuracy and computational efficiency. For the computation of traveltime, various methods have been described. Among those, the finite-differencing (FD) method, i.e., a FD-solver of the eikonal equation, has recently become widely used for calculation of "first arrival" traveltimes (Vidale, 1988). However, this method suffers from the disadvantages that it is restricted to the computation of first arrivals only and it produces unreliable amplitudes. Both are severe disadvantages for Kirchhoff-type inversion algorithms, such as the Bleistein/Cohen inversion (Bleistein *et al.*, 1987; Bleistein *et al.*, 2000), where the calculation of amplitudes is necessary to determine the weighting factor in heterogeneous media. Furthermore, in complex media, such as near salt domes and in sub-salt regions, later arrival traveltimes should be considered to obtain better image quality. Amplitudes can be used, among other things, to find most energetic arrivals.

Simultaneous computation of traveltimes and amplitudes is possible by dynamic ray tracing (DRT), which provides accurate multiple arrivals, amplitude and phase. Estimation of these ray data can be carried out either by numerical solution of ray-tracing equations in general smooth grid-based models or by piecewise analytic solutions for certain simple velocity functions in tetrahedral models. Among the choices of ray tracing procedures, the simplest and fastest solution of the ray tracing system is usually based on its analytic solution, wherever the complexity of the model allows one.

This is usually referred to as analytic ray tracing or cell ray tracing. Generally, the whole medium is divided into suitable cells (usually, tetrahedra in 3-D), in which the velocity can be approximated by simple functions that permit analytic ray solutions. The ray in the whole model is then obtained as a chain of analytically computed segments. The analytic ray tracing is usually performed for models in which either the velocity, $v(\boldsymbol{x})$, or $1/v(\boldsymbol{x})$, or $1/v^2(\boldsymbol{x})$, is a linear function of Cartesian coordinates. The simplest analytic solution for heterogeneous medium is that for constant gradient of squared slowness, also referred to as linear sloth media (Červený, 1987; Meng & Bleistein, 1997). However, this assumption leads to tetrahedral cells with artificial second-order discontinuities at their interfaces. As a consequence, this approach produces unreliable amplitude coefficients across the internal boundaries. Körnig (1995) proposed a method using quadratic sloth. In this approach, the squared slowness and its gradient with respect to spatial variables are continuous across each cell boundary. The analytic solutions for such a velocity function are determined by using the Laplace transform. Also, the computation of amplitude can be largely simplified by calculating the ray Jacobian directly from the analytical ray equations. However, determining the cell constants in quadratic sloth is rather difficult. The model design leads to a huge matrix inversion problem and is impractical in 3-D. Only 2-D implementation of the traveltime computation was carried out by Körnig (1995). I have concluded that the analytic approach in tetrahedral cells does not likely offer efficient algorithms in dynamic 3-D applications.

Traditionally, numerical DRT is performed by shooting a fan of rays from the source and extrapolating traveltimes and amplitude away from the rays into their nearby regions (Červený, 1987; Virieux & Farra, 1991; Sun & Biondi, 1995). The main disadvantage of the conventional shooting method is the lack of control of ray density in the search fan. Therefore, it is hard to reach a favorable compromise between efficiency and reliability, especially in complex 3-D models. It also produces shadow zones in areas of large velocity contrasts. The wavefront-construction traveltime-computation method (Vinje *et al.*, 1996) offers a solution to this problem by dynamically adding rays as needed. In this method, rays are maintained in a triangular network, and are traced stepwise in traveltime through the model. The wavefronts are then obtained automatically as a by-product of the ray tracing. In this chapter, the idea of wavefront construction is applied to 3-D models for estimation of both traveltime and amplitude coefficients. The dynamic interpolation of new rays assures that the wavefront is equipped with sufficient ray density at each computational step. Linear interpolation of traveltime with respect to the simulated wavefronts and linear interpolation of amplitude in terms of tube cross-sectional area are performed at grid points that fall into the sub-volume formed by every two successive wavefronts. A grid point can be passed by different sequences of wavefronts and, thereby, multi-valued arrivals can be detected and recorded. In this manner, all the grid points in the model are equipped with accurate—perhaps multi-valued—traveltimes and amplitudes.

In the following sections, I first discuss the possibility of applying analytic solutions in tetrahedral models for amplitude estimation and explain why I rejected this approach. Then, I propose a smooth gridded-model representation for the purpose of computational efficiency. Thereafter, I address some important issues in numerical DRT, such as interpolation of new rays and estimation of parameters at grid points. I also propose a method of re-introduction of transmission coefficients into the Green's function solution of the smoothed medium. Finally, I include a discussion of Maslov asymptotics at caustics in the last section.

2.2 Dynamic ray tracing

This section is a review of the dynamic ray-tracing theory, based on Červený (1987; 1995). The acoustic wave solution we seek is in the form,

$$A(\boldsymbol{x}, \boldsymbol{x}_0) e^{i\omega\tau(\boldsymbol{x},\boldsymbol{x}_0)}, \tag{2.2.1}$$

the leading order of an infinite series in powers of $1/i\omega$. Here, $\tau(\boldsymbol{x}, \boldsymbol{x}_0)$ and $A(\boldsymbol{x}, \boldsymbol{x}_0)$ are the traveltime and amplitude at $\boldsymbol{x} = (x_1, x_2, x_3)$ corresponding to the source $\boldsymbol{x}_0 = (x_{10}, x_{20}, x_{30})$. They are solutions of eikonal equation,

$$\nabla\tau(\boldsymbol{x}, \boldsymbol{x}_0)^2 = \frac{1}{v(\boldsymbol{x})}, \tag{2.2.2}$$

and transport equation,

$$2\nabla A \cdot \nabla\tau + A\nabla^2\tau = 0, \tag{2.2.3}$$

respectively. In equation (2.2.2), $v(\boldsymbol{x})$ is the velocity at $\boldsymbol{x}$.

2.2.1 Ray-centered coordinate system and the ray-parameter coordinate system

I begin by introducing two coordinate systems involved in DRT — the ray-centered coordinate system and the ray-parameter coordinate system. The **ray-centered coordinate system**, denoted by (q_1, q_2, q_3), is a curvilinear orthogonal coordinate system associated with any selected ray Ω (see Figure 2.1). One coordinate, say q_3, corresponds to any monotonic parameter along the ray, such as the arc length s, the traveltime τ or the parameter σ, with $d\tau = d\sigma/v^2$. Here, I take $q_3 = \tau$, the traveltime of the ray Ω away from the source. Thus, the traveltime, itself, is one of the coordinate axes in the ray-centered coordinate system. Coordinates q_1 and q_2 form a 2-D Cartesian coordinate system in the plane Σ perpendicular to Ω at $q_3 = \tau$, with the origin at Ω.

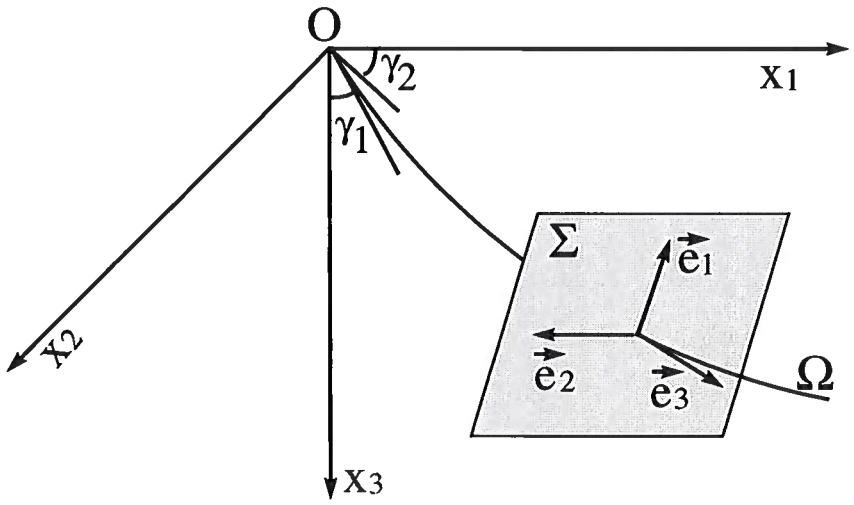

FIG. 2.1. The ray-centered coordinate system. $\boldsymbol{e}_3$ is tangent to the ray path Ω. $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ form a plane perpendicular to $\boldsymbol{e}_3$. γ_1 and γ_2 are the ray parameters that specify the ray, usually they are either the take-off angles (shown in this figure) or the slowness-vector components at the source.

The vector basis of the ray-centered coordinate system connected with Ω is formed at any arbitrary point $q_3 = \tau$ of ray Ω by a right-handed triplet of unit vectors $\boldsymbol{e}_1(\tau), \boldsymbol{e}_2(\tau), \boldsymbol{e}_3(\tau)$, as shown in Figure 2.1. The vectors $\boldsymbol{e}_i(\tau)$ can also be viewed as polarization vectors for isotropic media, when elastic wave propagation is considered. The unit vector $\boldsymbol{e}_3$ determines the direction of the displacement vector of P waves, which is always linearly polarized. Especially important are unit vectors $\boldsymbol{e}_1, \boldsymbol{e}_2$, since they determine the polarization of S waves, when we are dealing with vector solutions of the elastic wave equation.

The **ray parameter coordinates**, $(\gamma_1, \gamma_2, \gamma_3)$, are defined as follows: γ_1 and γ_2 are the ray parameters that specify the ray; usually they are either the take-off angles or the slowness-vector components at the source; γ_3 is any monotonic parameter, s, τ or σ, along the ray.

The Jacobian J of transformation from ray coordinates, $(\gamma_1, \gamma_2, \gamma_3)$, to the general Cartesian coordinates, (x_1, x_2, x_3), is an important factor in computation of the ray amplitude (Bleistein, 1984). The ray amplitude has the following form,

$$A(\boldsymbol{x}, \boldsymbol{x}_0) = \frac{const}{\sqrt{|J|}}, \tag{2.2.4}$$

with

$$J \equiv \det\left[\frac{\partial(x_1, x_2, x_3)}{\partial(\gamma_1, \gamma_2, \gamma_3)}\right]. \tag{2.2.5}$$

Here, J is the Jacobian, the constant is determined by the choice of $(\gamma_1, \gamma_2, \gamma_3)$.

2.2.2 Equations of dynamic ray tracing system

The rays are defined as the characteristics of the eikonal equation (Bleistein, 1984). That is, using the method of characteristics, one transforms the eikonal equation to the following six ray equations:

$$\begin{aligned} \frac{dx_i}{d\tau} &= v^2 p_i, \\ \frac{dp_i}{d\tau} &= -\frac{1}{v}\frac{\partial v}{\partial x_i}, \qquad i = 1, 2, 3. \end{aligned} \tag{2.2.6}$$

Here, $x_i(\tau)$ denotes the coordinates of position along the ray; $p_i(\tau)$ denotes the components of the slowness vector, τ denotes the traveltime along the ray, and $v(\boldsymbol{x})$ denotes the velocity. The system (2.2.6) is often referred as the **kinematic ray tracing (KRT) system**.

Differentiating the KRT system (2.2.6) with respect to the ray coordinates γ_i and applying the Taylor approximation up to second-order in q_i generates the **dynamic ray tracing system**. The DRT system can be expressed in many forms and in various coordinate systems. The simplest form of the DRT system is obtained in ray-centered coordinates connected with the ray Ω (Červený, 1987):

$$\begin{aligned} \frac{d\mathbf{Q}}{d\tau} &= v^2 \mathbf{P}, \\ \frac{d\mathbf{P}}{d\tau} &= -\frac{1}{v}\mathbf{V}\mathbf{Q}, \end{aligned} \tag{2.2.7}$$

where $\mathbf{Q}$, $\mathbf{P}$ and $\mathbf{V}$ are 2×2 matrices defined by

$$\begin{aligned} Q_{ij} &= \frac{\partial q_i}{\partial \gamma_j}, \qquad P_{ij} = \frac{\partial p_i}{\partial \gamma_j}, \qquad i, j = 1, 2, \\ V_{ij} &= \left.\frac{\partial^2 v(\boldsymbol{x})}{\partial q_i \partial q_j}\right|_{q_1 = q_2 = 0} = H_{kl} v_{,kl} H_{lj}, \qquad i, j = 1, 2, \\ H_{kl} &= \mathbf{i}_k \cdot \boldsymbol{e}_l, \qquad k, l = 1, 2, 3. \end{aligned} \tag{2.2.8}$$

Here, $\mathbf{i}_k$ are basis vectors in a general Cartesian coordinates (x_1, x_2, x_3) and H is the transform matrix from (q_1, q_2, q_3) to (x_1, x_2, x_3); an element H_{kl} represents the kth Cartesian component of basis vector $\boldsymbol{e}_l$. Elements V_{ij} of the matrix V are the second derivatives of velocity v with respect to q_i, and equivalent to the second derivatives

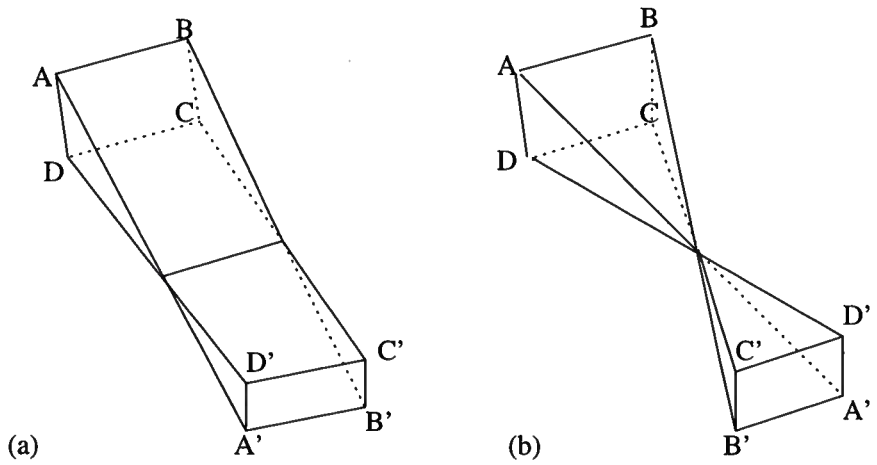

FIG. 2.2. Two types of caustic points along ray propagation. (a) At a caustic point of the first order, the ray tube shrinks into an arc, perpendicular to the direction of propagation. (b) At a caustic point of the second order (focus point), the ray tube shrinks to a point.

with respect to x_i under transformation H. Therefore, dynamic ray tracing (2.2.7) requires continuity of the velocity field up to second derivatives.[1]

The matrix $\mathbf{Q}$ is a transformation matrix from the ray parameters γ_1, γ_2 to the ray-centered coordinate q_1, q_2. Its determinant measures the ray Jacobian (2.2.5), and is also called the **geometrical spreading** factor, i.e.,

$$J = \frac{\mathrm{d}q_3}{\mathrm{d}\gamma_3} \det \boldsymbol{Q}. \tag{2.2.9}$$

The matrix $\mathbf{P}$ is a transformation matrix from the ray parameters γ_1, γ_2 to the slowness-vector component in the ray-centered coordinate system. The KRT system (2.2.6) computes the first derivatives of the traveltime field, while the DRT system (2.2.7) relates the second derivatives by the relationship,

$$\mathbf{M}(\tau) = \mathbf{P}\mathbf{Q}^{-1}. \tag{2.2.10}$$

where $\mathbf{M}$ is 2×2 matrix $\mathbf{M}$ of second derivatives of the traveltime field with respect to the ray-centered coordinate q_1, q_2, $M_{i,j} = \partial^2\tau/\partial q_i \partial q_j$.

2.2.3 At caustics

At caustics, the ray Jacobian (2.2.5), or the determinant of the matrix **Q** vanishes. In 3-D structures, there are two kinds of caustics depending on the dimension of the range[2] of the matrix. The ray tube may shrink to a caustic surface (envelope of rays) that is perpendicular to the direction of propagation (**a caustic point of the first order**); or the ray tube may shrink to a point (**a caustic point of the second order**), see Figure 2.2. In passing through the caustic point of the first order, the ray Jacobian J changes sign and the argument of $J^{1/2}$ takes on the phase shift $\pm\pi/2$. Similarly, in passing through the caustic point of the second order, the phase shift is $\pm\pi$. The **phase shift** due to caustics is cumulative; if we pass through several caustic points along the ray, the total phase shift is the sum of the individual phase shifts. This is often referred as the **KMAH index**, which was introduced by Ziolkowski and Dechamps (1980), acknowledging the work by Keller (1958), Maslov (1964), Arnold (1967) and Hörmander (1971) on this problem. The signs of the phase shifts are determined by the signs of the analytical signal. In general, the traveltime is decreased by the number of KMAH index times $\pi/2\omega$ (Kravtsov & Orlov, 1980).

2.3 Discussion on analytic ray tracing

Typically realistic velocity models seldom allow general, analytic solutions of the ray-tracing system. However, analytic ray tracing plays an important role in wave-field computation. In particular, analytic solutions are valuable in the cell approach, in which the whole model is subdivided into a set of tetrahedral cells with simple velocity functions within cells. The models allowing analytic solutions are usually those with either the velocity, or slowness, or squared slowness being a linear function of Cartesian coordinates. As mentioned above, this constraint does not provide enough smoothness at all boundaries for amplitude estimation. In this section, I discuss the analytic solutions for **quadratic sloth** media.

[1]It is the numerical sampling across tetrahedral interfaces that causes amplitude instability when the second derivative is not continuous.

[2]The range of a matrix is the vector space of all nonzero vectors that result from the product of the vector with this vector.

2.3.1 Analytic ray tracing in a quadratic sloth medium

The quadratic sloth distribution, denoted by q, is defined as a quadratic function in space,

$$\begin{aligned} q(x_1, x_2, x_3) &= \frac{1}{v^2(x_1, x_2, x_3)} \\ &= A + 2B_i x_i + C_{ij} x_i x_j. \end{aligned} \tag{2.3.11}$$

Here and throughout this thesis, repeated indices in a term indicate summation (summation convention). The analytical solutions of raypaths for this distribution were found by Körnig (1995) using the Laplace transform to the eikonal equation (2.2.2). In the Laplace domain, the ray coordinates, $X_i(s)$, are ratios of polynomials of sixth and seventh order, respectively, in s, the Laplace variable corresponding to σ, the ray tracing integral variable with $d\sigma = v^2 d\tau$. The expressions for the ray trajectories, $x_i(\sigma)$, can be obtained explicitly by inverse Laplace transform of $X_i(s)$ using partial fraction expansions of the Laplace transform. Depending on the distribution of eigenvalues of C_{ij}, the solutions of $x_i(\sigma)$ are generalized into seven different forms. For each case, the ray trajectories are in the general form

$$x_i(\sigma) = w_{ik} f_k(\sigma), \quad i = 1, 2, 3, \quad k = 1, 2, ..., 7. \tag{2.3.12}$$

Here, the w_{ik}'s are weighting factors, which are functions of x_{i0}, p_{i0}, B_j and C_{ij}, with x_{i0} and p_{i0} being the initial position and slowness components; $f_k(\sigma)$ are the basis functions corresponding to the inverse transform of the partial fraction expansions of the Laplace transform. One of them is unity; the others are either low-order polynomials in σ, or trigonometric, or hyperbolic functions.

Now I propose an alternative to Körnig's (1995) approach to calculate the amplitudes along rays. Notice that the $f_k(\sigma)$'s in (2.3.12) are functions depending only on σ, and w_{ik}'s can be expressed in terms of x_{i0}, p_{i0}, and constants B_j and C_{ij}. Therefore, if we choose $\gamma_1 = p_{10}, \gamma_2 = p_{20}$ and $\gamma_3 = \sigma$ in (2.2.5), the ray Jacobian J can be calculated analytically,

$$\begin{aligned} J &= \det \begin{bmatrix} \frac{\partial w_{ik}}{\partial p_{10}} f_k(\sigma) \\ \frac{\partial w_{ik}}{\partial p_{20}} f_k(\sigma) \\ w_{ik} g_k(\sigma) \end{bmatrix} \\ &= \det\{J_{ijk} j_k(\sigma)\}, \end{aligned} \tag{2.3.13}$$

with summation over the repeated index, k, and

$$j_k(\sigma) = \begin{cases} f_k(\sigma), & i = 1, 2 \\ g_k(\sigma), & i = 3. \end{cases} \tag{2.3.14}$$

In the above equations, $g_k(\sigma) = \mathrm{d}f_k(\sigma)/\mathrm{d}\sigma, k = 1, ..., 7$. The coefficients J_{ijk} can be expressed for all seven cases in (2.3.12). Notice that expression (2.3.13) has taken the place of the DRT system (2.2.7); and it is less computationally costly than solving the eight integrations in small time steps.

2.3.2 Why not use analytic ray tracing?

This approach of making the quadratic sloth assumption in tetrahedral models has eliminated the smoothing procedure across the internal interfaces. However, the constants A, B_i and C_{ij} in (2.3.11) are usually not known in advance for a given physical model. They have to be determined from the discrete model. The assumption of quadratic sloth is equivalent to the continuity of both $q(\boldsymbol{x})$ and the gradient of $q(\boldsymbol{x})$ across the internal cell faces. Therefore, the 10 constants A, B_i and C_{ij} in one tetrahedron cannot be determined from the velocity values at its four apices only, but also depend on the values in the neighboring cells. Such a model design problem for all the tetrahedral cells leads to a huge matrix inversion problem. If the whole model is divided into N cells, the size of the coupled system of equations is proportional to $10N$, making the computation very time consuming. Furthermore, this inverse system is not always solvable, or has solutions in a least squares sense, at best. This is impractical in 3-D and considerably limits the applicability of this approach.

From the above discussion, we see that although the assumption of quadratic sloth in tetrahedral models provides accurate amplitudes, this assumption leads to a difficult and inefficient numerical problem for determining the cell constants. This problem exists in all extensions to quadratic physical models, not just to the quadratic sloth model. We conclude that analytic ray tracing in tetrahedral models is not likely to give us an efficient module for dynamics, although it has its applications in traveltime calculations.

2.4 Wavefront construction on smooth gridded models

2.4.1 Model representation

The smoothness of the velocity model representation is critical to the calculation of amplitudes. The integration of the DRT system (2.2.7) requires continuity of the

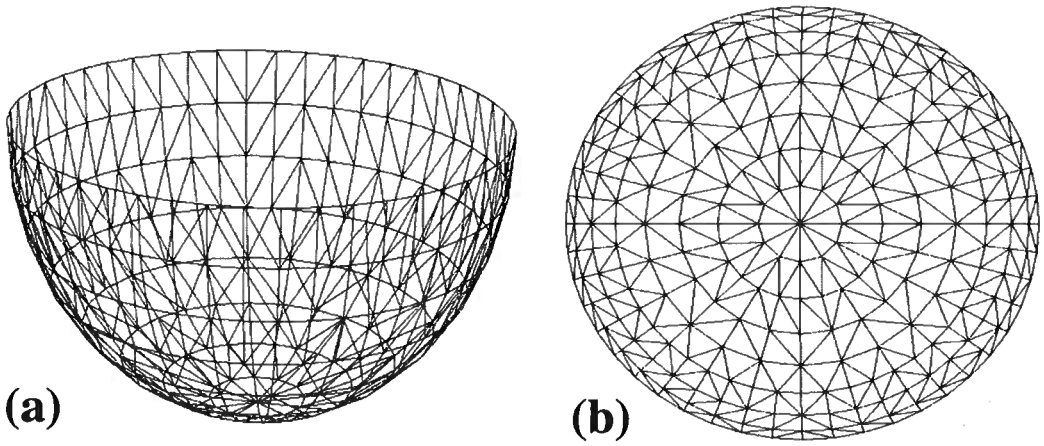

FIG. 2.3. The wavefront is represented as a triangular network. Each triangle is determined by its three neighboring ray endpoints on the wavefront. (a) a 3-D view, and (b) a planar view of the wavefront at a fixed time in a homogeneous medium.

velocity field up to the second derivatives. Here, in order to guarantee the velocity distribution is smooth enough for traveltime and amplitude calculation, I apply a smoothing filter to the whole velocity model. Many ray tracing procedures (Farra, 1990) involve a type of spline interpolation for the evaluation of velocities at arbitrary points. Spline interpolation, however, is a time consuming procedure. Here, we define the velocity model on a fine grid (about three or four grid points per shortest significant model wavelength) and pre-calculate its first and second derivatives at all grid points by finite differences of second order. Then, for the evaluation of the velocities and their derivatives at arbitrary points we use linear interpolation. For the smooth models defined on fine grids, the difference between this linear representation and a spline representation of the model is negligible.

When the considered model contains discontinuous velocities, however, a smoothing procedure must be applied to guarantee that the velocities vary smoothly. Thus, for the sake of computational efficiency, the interface conditions are eliminated here by using proper smoothing and a densely sampled grid model. Below, we study implications of this smoothing.

2.4.2 Wavefront construction

Here and below, I focus on numerically solving (2.2.6) and (2.2.7) for smooth gridded models. I will apply the technique of wavefront construction (WFC) to both kinematics and dynamics.

In the wavefront construction method, a relatively small number of rays are shot ini-

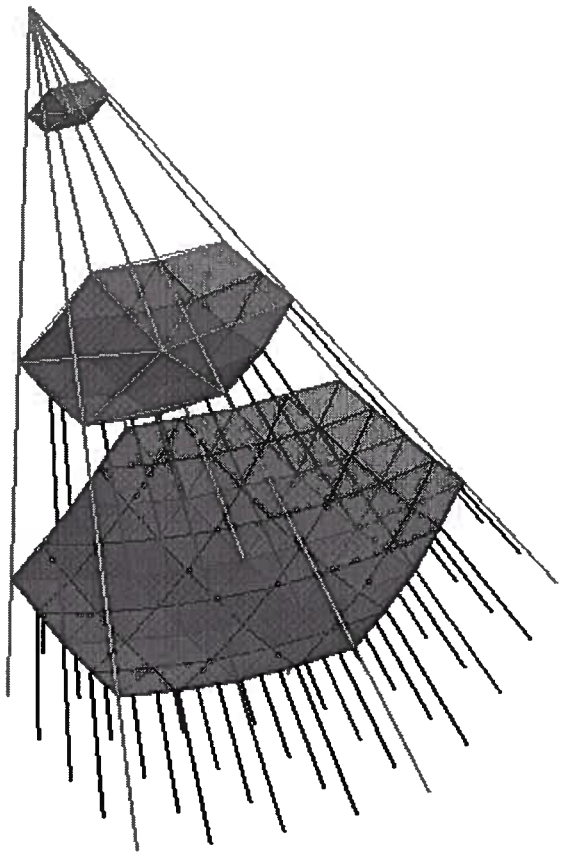

FIG. 2.4. The ray tracing algorithm has criteria to check and keep the size of each triangle and the propagation-direction derivation of every two adjacent rays within some pre-defined maximum values. When these ray-coverage conditions are violated, new rays are added such as the darker rays in the second wavefront and the darkest rays in the third wavefront. The original triangles are subdivided into new triangles.

tially. They differ from each other by the two take-off angles, and are extrapolated into the zone of interest by solving (2.2.6) and (2.2.7) numerically with appropriate initial conditions. Required accuracy of traveltime and amplitude can be approached by various standard numerical procedures, such as Runge-Kutta or predictor-corrector, for example.

At any computational step, the wavefront is obtained as a by-product of the ray tracing. Figure 2.3 shows one wavefront. The wavefront is represented by triangular plates that connect every three neighboring ray endpoints on the WF. The nearby rays in 3-D are then defined and organized by such a triangular mesh consisting of the internal ordering of connecting endpoints in each triangle and adjacent triangle(s) to each of its sides. The processes of checking, interpolating new rays and estimating grid point parameters (described below) are all performed within such a triangular network. Rays are added and the original triangle is subdivided into new triangles when certain criteria, restricting the size of the triangular plates, are violated. Figure 2.4 is an illustration of dynamically adding rays along ray tracing. Three wavefronts are displayed in this figure. A set of initial rays (lightest) pass through the first wavefront and no new rays are added. At the second wavefront, the darker rays are introduced and the original triangles are subdivided into new triangles, because the original triangles failed one of our sizing criteria at this wavefront. The lightest and darker rays with their attendant triangles are propagated to the third wavefront. There, again, new triangles are added along with the darkest rays; and so on. In this manner, the wavefront always has sufficient ray density without *a priori* estimation of the number of rays needed, or by imposing an excessive ray density on initiation. For complex 3-D velocity models, the wavefront surface may be very complicated, folding in on itself at some parts, for example; however, no tears or holes in the interior of the surface are allowed. In this sense, the wavefront is complete. On the other hand, a grid point can be passed by different sequences of wavefronts; multi-valued arrivals can then be estimated and distinguished by their initial take-off angles.

Figures 2.5 through 2.7 show ray tracing in a linear sloth model; that is, the velocity increases with depth with constant gradient of squared slowness. Figure 2.5 shows some in-plane rays. Double-arrivals are observed. The two co-located arrivals differ from each other by their take-off angles and traveltimes. Rays generated by using wavefront construction are shown in Figure 2.6. New rays are added at different computation steps whenever the ray coverage becomes too sparse. Figure 2.7 is one wavefront recorded at traveltime 0.8 s. The outer part of wavefront is generated by the up-going rays with large take-off angles, and the inner part of wavefront is formed by the down-going rays with small take-off angles. The edge of the wavefront sits in the caustic surface. Because of the double-arrivals, the wavefront is folded on itself in some places. This phenomenon is common in complex velocity models.

The most attractive advantage of the WFC method is that it is more efficient than the conventional ray tracing method because the number of rays is increased only

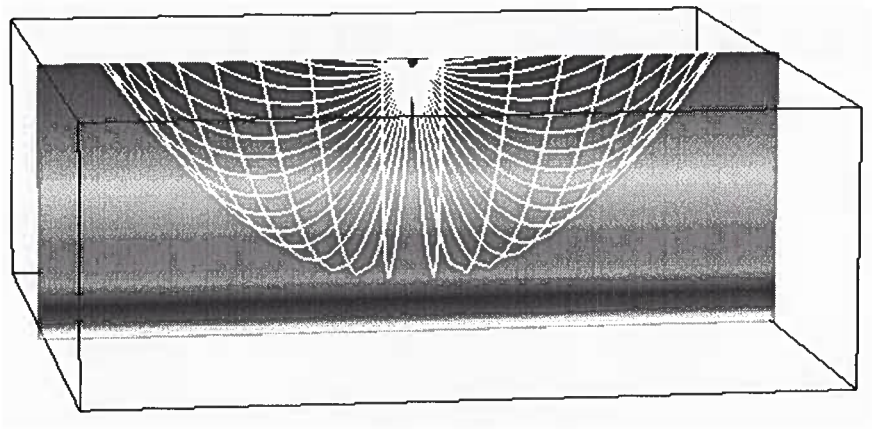

FIG. 2.5. Rays in a linear sloth model. Double-arrivals are observable. The dimension in this figure is 2×2 km in horizontal directions and 2 km in depth direction.

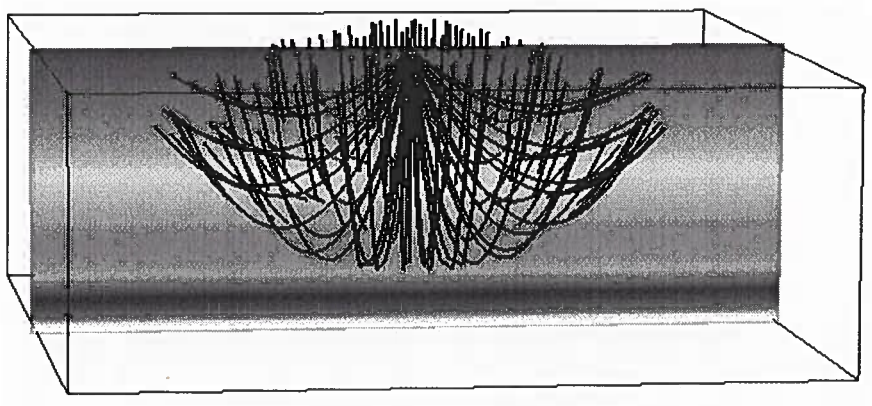

FIG. 2.6. Rays generated by wavefront construction in the same linear sloth model. The dimension in this figure is 2×2 km in horizontal directions and 2 km in depth direction.

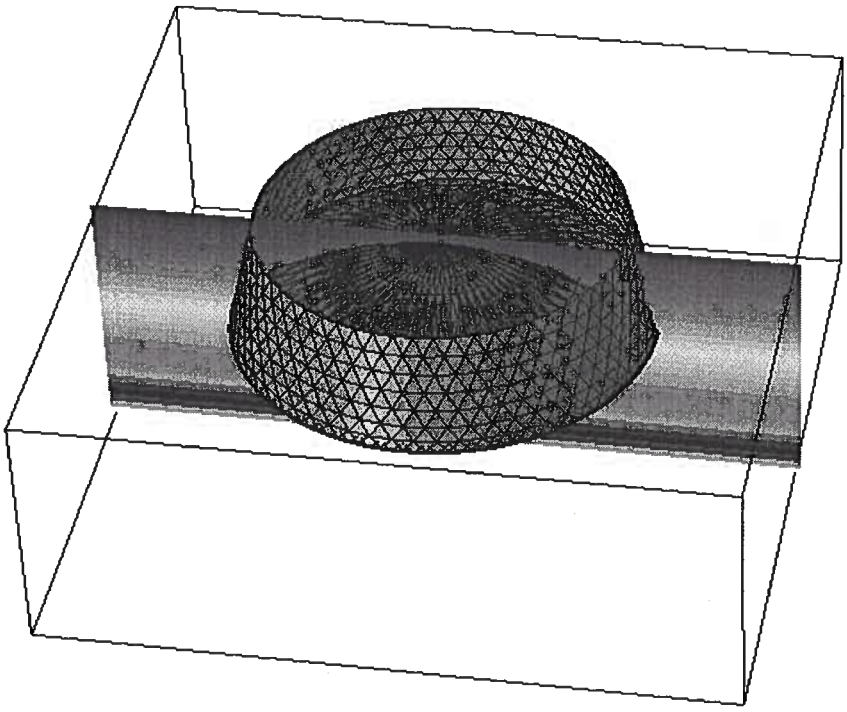

FIG. 2.7. Wavefront in the same linear sloth model. The dimension in this figure is 2×2 km in horizontal directions and 2 km in depth direction.

as needed to keep the density of rays approximately constant. In addition, WFC gives better ray coverage, especially in areas of large geometrical spreading where conventional ray tracing may give no arrivals. Furthermore, compared to FD-solvers, the WFC method is not restricted to the calculation of first arrivals only. Amplitude and other ray theoretical quantities are also available. Thus, it meets the requirements for accurate modeling of amplitude as well as phase, a requirement for inversion as opposed to migration.

2.4.3 Ray interpolation

The wavefront construction method is largely dependent on the procedure of interpolation of the wavefront at each step. New ray endpoints must be added along the simulated wavefront and must have the propagation direction that the rays would have had if they had been shot from the source. This section addresses an algorithm for accomplishing these requirements. This implementation was derived from a 2-D interpolation provided by D. Hanson (private communication).

Rays diverge and wavefronts expand throughout the wave-field. When new ray endpoints are needed to maintain a certain ray density on the wavefront, the whole

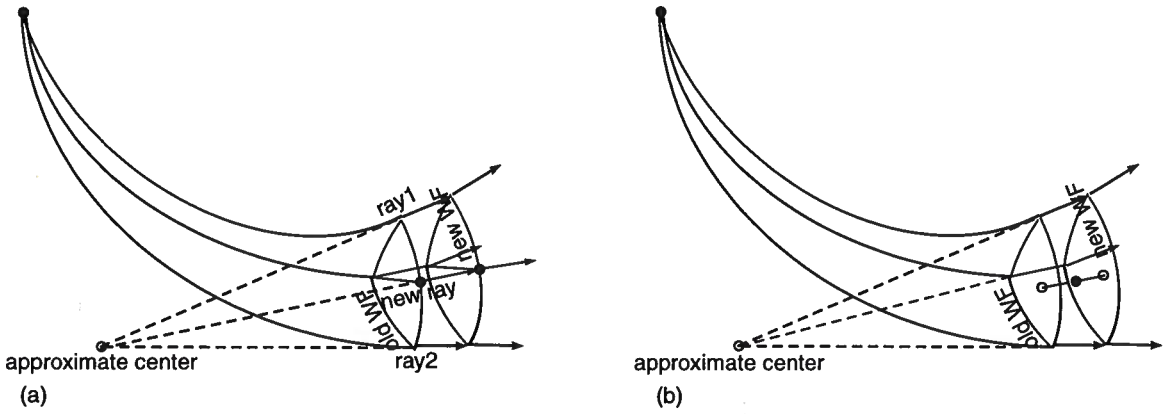

FIG. 2.8. Interpolation of new rays and estimation of grid parameters are performed in a ray tube. (a) Interpolation of a single new ray. The two ray endpoints and their propagation directions form two straight lines in 3-D. An "approximate" center can be defined as the midpoint of the line segment that connects the two straight lines at their points of shortest distance. This approximate center, along with the two ray ends at the old WF, form a fan and a circular curve connecting the two ray end points. The new ray position is then found along the dividing direction from the approximate center and at its intersection with the circular curve. Then the interpolated ray is traced from the old to the new WF. (b) Simple ray cell with an interior grid point. Ray data are estimated with respect to the two simulated WFs.

triangular network will have to be reorganized. The criterion for this interpolation can be that the area of triangles must not exceed a pre-specified limit, or that the angle deviation of the slowness vectors of two adjacent rays cannot be too large. New rays are always added in between the pairs of existing rays in order to meet the criteria for size and angle difference of triangular plates on the wavefront. Figure 2.8(a) illustrates the interpolation of a single new ray. The two ray endpoints and their propagation directions form two straight lines in 3-D. An "approximate" center can be defined as the midpoint of the line segment that connects the two straight lines at their points of shortest distance. This approximate center, along with the two ray ends at the old WF, form a fan and a circular curve connecting the two ray end points. The new ray position is then found along the dividing direction from the approximate center, and at its intersection with the circular curve. Other parameters along the new ray are interpolated linearly.

There are other alternative methods of interpolating new rays, such as the parameterization of a wavefront by a third-order polynomial (Vinje *et al.*, 1996). However, they require special treatment in the vicinity of caustic points, since only rays belonging to the same phase must be used to determine quantities of the new ray. The method described here does not have such a requirement. Moreover, we do not use the curvature of the wavefront obtained from DRT for the interpolation to keep the

problems of KRT and DRT separate. I have found this method stable.

2.4.4 Grid interpolation

Another interpolation procedure in DRT is the estimation of ray data at the output Cartesian-grid points. This is our final goal of the ray-tracing algorithm. The grid interpolation is performed within ray tubes, which are prism-shaped bodies bounded by three rays and the triangles that connect them on the two WFs. First the grid points falling into (or close to) each cell are found. Then, the traveltime can be estimated at each grid point in a way similar to the interpolation procedure for new rays (see Figure 2.8(b)). The approximate center is determined by the three rays with ray endpoints on any of the two WFs, and the distances from each grid point to the approximate center and to the simulated wavefront are calculated. The traveltime at the grid point is then recorded as $t_0 + d/v$, with t_0 the time at the wavefront, d the distance of the grid point away from the wavefront, and v the velocity at the grid point.

The above procedure is not suitable for interpolation of amplitude because the isochrons surface is usually not the iso-amplitude surface. Here, we apply linear interpolation for amplitudes, which is based on the assumption that the amplitudes vary only slowly; otherwise, the validity conditions of the underlying asymptotic theory would be violated. Since the ray Jacobian—the determinant of Q in (2.2.7)—is proportional to the cross sectional area of the ray tube, we interpolate the ray Jacobian linearly with respect to the triangle areas on the two wavefronts.

Another parameter that requires interpolation at every grid point is the initial shooting direction, i.e., the initial take-off angles of the ray that would reach the grid point if the ray actually had been traced. This parameter is stored in order to distinguish between arrivals, because two arrivals at a grid point cannot have almost equal take-off angles.

2.5 Transmission coefficients

In seismic processing, two types of velocity models are commonly used: models that consist of (often homogeneous) layers, and heterogeneous models without interfaces. The first type of velocity model is usually constructed by layer stripping methods. The second type of model usually results from methods that perform a global optimization to determine the velocity. Accurate amplitude requires velocity models to be smooth.

Velocity models of the second type do not contain interfaces, but for velocity models of the first type, transmission loss at interfaces can have significant influence on Green's

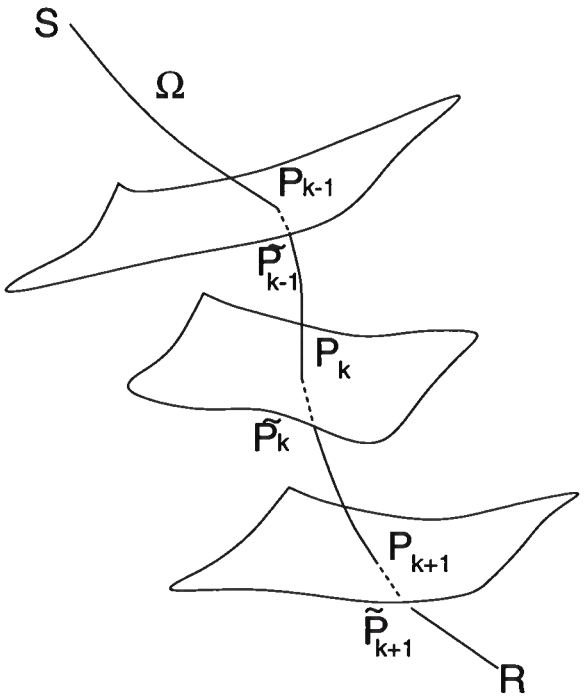

FIG. 2.9. Ray Ω from S to R strikes several interfaces.

function amplitudes. In the previous sections, the dynamic ray tracing is performed on a general smooth gridded model, even where there were real interfaces. The transmission coefficient has been taken to be unity, because the smooth solution does not take account the transmission coefficient. When a ray strikes a real interface, a reflection or transmission coefficient should be included in the Green's function. In this section, we discuss the re-introduction of transmission coefficients into the Green's function solution of the smoothed medium.

In smooth media, the amplitude determined by the transport equation is continuous, and we can estimate it using equations (2.2.4) through (2.2.7). Let us consider a ray Ω and two points S and R situated on Ω. Then the transport equation leads to the following relationship between the amplitudes at S and R if there is no transmission loss:

$$A(R) = \left[\frac{v(R)\det \boldsymbol{Q}(S)}{v(S)\det \boldsymbol{Q}(R)}\right]^{\frac{1}{2}} \exp[-i\frac{\pi}{2}\varphi(R,S)]A(S), \tag{2.5.15}$$

with $\varphi(R,S)$ the KMAH index from S to R.

Assume that the ray Ω strikes n interfaces, $\Sigma_1,\ldots,\Sigma_n$, at points, $P_1,\ldots,P_n$, between the points S and R (see Figure 2.9). $\tilde{P}_k$ $(k=1,\ldots,n)$ denote the points coinciding with P_k but corresponding to the transmitted branch of the ray. Therefore, the ray Ω

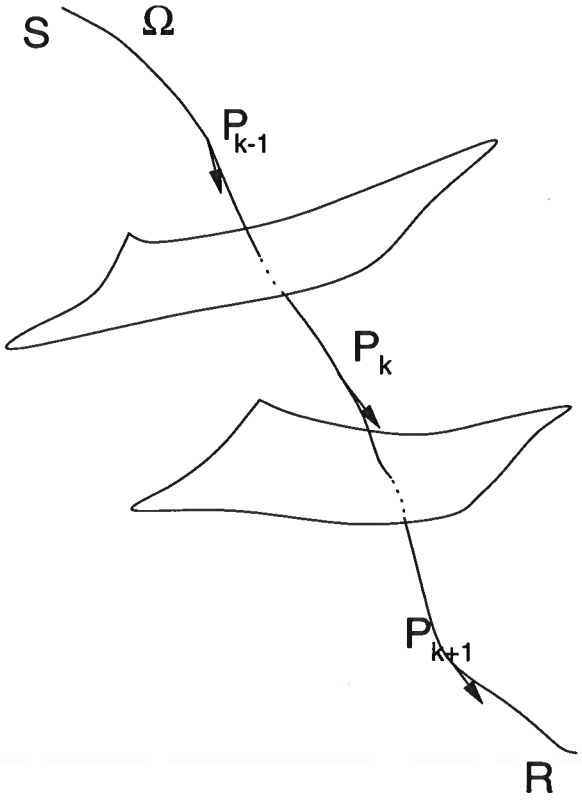

FIG. 2.10. Each ray segment hits an interface. The transmission coefficient is computed at each ray endpoint.

from S to R is divided into $n+1$ segments, with $\tilde{P}_0 = S$ and $P_{n+1} = R$. By connecting all the ray segments along S and R, the continuation relation for the amplitudes of acoustic waves becomes,

$$A(R) = \left[\frac{v(R)\det\boldsymbol{Q}(S)}{v(S)\det\boldsymbol{Q}(R)}\right]^{\frac{1}{2}} \mathcal{T}^c \exp[-i\frac{\pi}{2}\varphi(R,S)]A(S), \tag{2.5.16}$$

where

$$\begin{aligned}
\varphi(R,S) &= \sum_{k=1}^{n+1} \varphi(P_k, \tilde{P}_{k-1}), \\
\mathcal{T}^c &= \prod_{k=1}^{n} \mathcal{T}(P_k) = \prod_{k=1}^{n} T(P_k) \left[\frac{v(P_k)\cos i(\tilde{P}_k)}{v(\tilde{P}_k)\cos i(P_k)}\right]^{\frac{1}{2}}, \\
T(P_k) &= \frac{2v(P_{k+1})\cos i(P_k)}{v(P_{k+1})\cos i(P_k) + v(P_{k-1})\cos i(\tilde{P}_k)}.
\end{aligned} \tag{2.5.17}$$

Here, $i(P_k)$ is the acute angle of incidence and $i(\tilde{P}_k)$ the acute angle of transmission. $T(P_k)$ is the standard plane wave transmission coefficient at P_k. $\mathcal{T}^c$ is called the complete transmission coefficient. It is equal to the product of the normalized transmission coefficients at all points of incidence between S and R.

Comparing equations (2.5.15) and (2.5.16), we notice that the difference between the smooth amplitude solution and the one in layered structures is the factor $\mathcal{T}^c$. In the ray-tracing algorithm, a smooth gridded velocity model is assumed, so the location and shape of the interfaces are not known. Here, I propose to modify the amplitude at each endpoint along ray tracing. Suppose $P_1, \ldots, P_m$ are m ray end points along the ray between S and R, which are predicted by the dynamic ray-tracing method discussed in this chapter. Factors needed in (2.5.17) are the cosines of the incident and refracted angles, while the velocities are already interpolated in the ray-tracing procedure. Now assume that each ray segment hits an interface as shown in Figure 2.10. For the calculation of the cosines, I use the following approximation,

$$\begin{aligned}
\nabla'_{\boldsymbol{x}} v(P_k) &= (\nabla_{\boldsymbol{x}} v(P_{k-1}) + \nabla_{\boldsymbol{x}} v(P_k))/2, \\
\cos i(P_k) &= \boldsymbol{p}(P_{k-1}) \cdot \nabla'_{\boldsymbol{x}} v(P_k), \\
\cos i(\tilde{P}_k) &= \boldsymbol{p}(P_k) \cdot \nabla'_{\boldsymbol{x}} v(P_k),
\end{aligned} \tag{2.5.18}$$

where $\nabla_{\boldsymbol{x}} v(P_{k-1})$ and $\nabla_{\boldsymbol{x}} v(P_k)$ are the gradients of the velocity at P_{k-1} and P_k, which point to the direction of the normal to the interface that passes through the points P_{k-1} and P_k; $\boldsymbol{p}(P_{k-1})$ and $\boldsymbol{p}(P_k)$ are the slowness vectors at points P_{k-1} and

P_k, respectively.

Approximation (2.5.18) estimates angularly dependent transmission coefficients at arbitrary incident and transmission directions. The extra computation cost of this method is basically two inner products at each point along the ray. In the next chapter, an example of applying this method is given for a model with strong velocity contrast. In regions away from interfaces, the cumulative effect including this transmission coefficient is a factor near unity. In regions of rapid variation, the cumulative effect yields a good approximate of the transmission coefficient.

2.6 Maslov asymptotics

Normal asymptotic ray theory (ART) gives the solution for wave equation in the spatial domain. Its disadvantage is that the discontinuities or smooth heterogeneities in the model can lead to regions where the rays form envelopes of varying degrees of complexity—caustics— where straightforward ray theory breaks down. It is important to note that the actual solution does not have these singularities; the problem is that the simple structure, $A\exp\{i\omega\tau\}$, is not valid in these special regions.

Maslov's method (Maslov, 1964; Chapman & Drummond, 1982; Kendall & Thomson, 1993; de Hoop & Brandsberg-Dahl, 1998) provides an alternative representation of the asymptotic wavefield that does not break down near caustics. In this approach, the ray trajectory, which is normally considered in the spatial domain, is viewed more generally in the six-dimensional phase space consisting of position and slowness. The solution in the physical space is a *projection* of the solution in six independent variables. However, the asymptotic solution for other mixed coordinates—other projections—in phase space can also be found. For instance, transform methods find the solution in the mixed domain where the horizontal slowness replaces one spatial coordinate. The reason for considering asymptotic solutions in the different mixed domains is that, in general, the singularities in different domains do not coincide. Thus, a nonsingular asymptotic solution in some alternative three-dimensional subset of the six variables in phase space can always be found. Maslov asymptotic theory extended this idea to heterogeneous media, and the asymptotic solution in a mixed domain (position and slowness) is obtained by a canonical transformation from the spatial domain. The method is useful because the singularities in the mixed and spatial domains are at different locations, and Maslov theory provides a uniform result by combining the solutions in the different domains.

2.6.1 1-D Maslov – first-order caustics

Equation (2.2.6) defines the ray trajectory in six dimensions $\boldsymbol{x} \times \boldsymbol{p}$. The Maslov solution is in the form of a Radon transform that introduces one or more slowness coordinates as parameters. This is done by applying the partial Legendre transformation.

At points of first-order caustics in $\boldsymbol{x}$ space domain, the cross-section of a ray tube is reduced to a line, and the KMAH index increases by unity or a phase retardation of $\pi/2$, which corresponds to one eigenvalue vanishing in (2.2.5). Here I assume the first-order caustic is not parallel to the x_1-axis (otherwise, we can transform whichever of the original coordinates is most nearly normal to the caustic surface and the discussion still holds). In this case, (p_1, x_2, x_3) constitute the proper coordinates on the Lagrangian manifold over a region Ω_I that will provide a nonsingular solution. Let X_1 denote the point in the plane (x_2, x_3) = constant where the wave front has (projected) slowness p_1. The mapping $p_1 \rightarrow X_1$ represents a change of coordinates on the Lagrangian manifold. This can be achieved by the following partial Legendre transform ϕ of τ,

$$\phi(p_1, x_2, x_3; \boldsymbol{x}') = \tau\left[X_1(p_1, x_2, x_3), x_2, x_3; \boldsymbol{x}'\right] - p_1 X_1(p_1, x_2, x_3). \tag{2.6.19}$$

Here, $\boldsymbol{x}'$ denotes any reference point along the ray Ω. Differentiating equation (2.6.19) with respect to p_1, and taking into account that $\partial_{X_1}\tau = p_1$, it follows that

$$\partial_{p_1}\phi = -X_1 \text{ for } x_2, x_3 \text{ fixed} \quad (\partial_{x_2}\phi = p_2, \partial_{x_3}\phi = p_3 \text{ for } p_1 \text{ fixed}). \tag{2.6.20}$$

The function ϕ in turn defines a hypersurface consisting of points (p_1, x_2, x_3, t) satisfying $t - \phi(p_1, x_2, x_3; \boldsymbol{x}') = 0$. Notationally, we describe this hypersurface by

$$\{(p_1, x_2, x_3, t) \mid t - \phi(p_1, x_2, x_3; \boldsymbol{x}') = 0\}.$$

The Legendre transformation can be re-applied to ϕ to yield

$$\tau(X_1, x_2, x_3; \boldsymbol{x}') = \phi(p_1(X_1, x_2, x_3), x_2, x_3; \boldsymbol{x}') + X_1 p_1(X_1, x_2, x_3),$$

The original traveltime function is recovered, and it follows that

$$\partial_{X_1}\tau = p_1 \text{ for } x_2, x_3 \text{ fixed} \quad (\partial_{x_2}\tau = p_2, \partial_{x_3}\tau = p_3 \text{ for } x_I \text{ fixed}). \tag{2.6.21}$$

Now we construct the phase function in the (p_1, x_2, x_3) domain,

$$\begin{aligned} \vartheta(p_1, \boldsymbol{x}; \boldsymbol{x}') &= \phi(p_1, x_2, x_3; \boldsymbol{x}') + p_1 x_1 \\ &= \tau(X_1(p_1, x_2, x_3), x_2, x_3; \boldsymbol{x}') + p_1(x_1 - X_1(p_1, x_2, x_3)) \,. \end{aligned} \tag{2.6.22}$$

The above expression is the arrival time at x_1 of a 'p_1 wave' reaching X_1 at time $\tau(X_1, x_2, x_3; \boldsymbol{x}')$. At $x_1 = X_1(p_1, x_2, x_3)$ this phase is stationary, i.e.,

$$\begin{aligned}(\partial_{p_1}\vartheta)(p_1, \boldsymbol{x}; \boldsymbol{x}') &= \frac{\partial\tau}{\partial X_1}\frac{\partial X_1}{\partial p_1} + (x_1 - X_1) - p_1\frac{\partial X_1}{\partial p_1} \\ &= p_1\frac{\partial X_1}{\partial p_1} + x_1 - X_1 - p_1\frac{\partial X_1}{\partial p_1} \\ &= x_1 - X_1(p_1, x_2, x_3) = 0. \end{aligned} \tag{2.6.23}$$

The stationary phase argument leads to

$$\vartheta((p_{\boldsymbol{x}})_1, \boldsymbol{x}; \boldsymbol{x}') = \tau(\boldsymbol{x}; \boldsymbol{x}'), \quad \text{and} \quad \nabla_{\boldsymbol{x}}\vartheta((p_{\boldsymbol{x}})_1, \boldsymbol{x}; \boldsymbol{x}') = \boldsymbol{p}_{\boldsymbol{x}}, \tag{2.6.24}$$

on the bicharacteristics.

The ('one-way') Maslov asymptotic representation on a given region, Ω_1, of the Lagrangian manifold (p_1, x_2, x_3) is given by

$$U(\boldsymbol{x}, \boldsymbol{x}', \omega) \sim \left(\frac{i\omega}{2\pi}\right)^{\frac{1}{2}} \times \int \mathrm{d}p_1\, B(p_1, x_2, x_3; \boldsymbol{x}') \exp[i\omega\,\vartheta(p_1, \boldsymbol{x}; \boldsymbol{x}')], \tag{2.6.25}$$

where $B(p_1, x_2, x_3; \boldsymbol{x}')$ is the weighting constant to be determined. We can consider the WKBJ seismogram as being constructed by decomposing the point source into p_1 waves, propagating each wave using ART, and recombining the p_1 wave at the receiver.

Away from caustics, I apply stationary phase (Bleistein, 1984) to (2.6.25), and obtain,

$$U(\boldsymbol{x}, \boldsymbol{x}', \omega) \sim \frac{B(p_1, x_2, x_3)}{|\partial_{p_1} X_1|^{\frac{1}{2}}} \exp\left[i\mathrm{sgn}(\omega)\frac{\pi}{4}[1 - \mathrm{sgn}(\partial_{p_1} X_1)]\right] \exp[i\omega\tau(\boldsymbol{x}, \boldsymbol{x}')] \tag{2.6.26}$$

where $|\partial_{p_1} X_1| \neq 0$. Expression (2.6.26) should agree with the following ART result:

$$A \exp\left[i\omega\tau(\boldsymbol{x}, \boldsymbol{x}') - i\frac{\pi}{2}\mathrm{sgn}(\omega)\varphi(\boldsymbol{x}, \boldsymbol{x}')\right].$$

Here, $\varphi(\boldsymbol{x}, \boldsymbol{x}')$ is the KMAH index from $\boldsymbol{x}'$ to $\boldsymbol{x}$. This implies

$$B(p_1, x_2, x_3) = \frac{v(\boldsymbol{x})}{\left|\frac{\partial(p_1, x_2, x_3)}{\partial(q_1, q_2, \tau)}\right|^{\frac{1}{2}}}, \tag{2.6.27}$$

and

$$-i\text{sgn}(\omega)\frac{\pi}{4}\left[2\varphi(\boldsymbol{x};\boldsymbol{x}') - (1-\text{sgn}(\partial_{p_1}X_1))\right] = 0. \tag{2.6.28}$$

Equation (2.6.28) implies that the first-order caustic occurs at $|\partial X_1/\partial p_1| = 0$, and $\partial X_1/\partial p_1$ changes signs when passing through a first-order caustic.

At the caustics, the normal ART, $A\exp[i\omega\tau(\boldsymbol{x},\boldsymbol{x}') - i\pi\text{sgn}(\omega)\varphi(\boldsymbol{x},\boldsymbol{x}')/2]$, breaks down, but (2.6.25) is still valid. Applying a higher-order stationary phase analysis to (2.6.25) and supposing $|\partial^2 X_1/\partial p_1^2| \neq 0$, we obtain,

$$U(\boldsymbol{x},\boldsymbol{x}',\omega) \sim O\left(|\omega|^{-\frac{1}{3}}|\omega|^{\frac{1}{2}}\right) = O\left(|\omega|^{\frac{1}{6}}\right). \tag{2.6.29}$$

This expression means that the asymptotic Green's function near caustics is larger than the solution away from caustics, through the higher power in ω. In the inversion formulas (see next chapter), the two Green's functions appear in the denominator. Therefore, when substituting the above expression into the inversion formula, the contribution at $|\partial X_1/\partial p_1| = 0$ is of lower order in ω than those points with $|\partial X_1/\partial p_1| \neq 0$. This suggests to us that we can ignore the contribution at the caustic points in the migration/inversion process. If, at the caustic points, $|\partial^2 X_1/\partial p_1^2| = 0$, we can apply higher-order stationary analysis at those points, which leads to even lower-order terms in inversion formula.

2.6.2 2-D Maslov – second-order caustics

At a point of second-order caustics in $\boldsymbol{x}$-space, the cross-section of a ray tube shrinks to a point. Two eigenvalues in (2.2.5) vanish, and the KMAH index increases by two with a phase shift of π in the stationary phase analysis. In this case, the caustic points are also singularities in 1-D Maslov domain (p_1, x_2, x_3). That is, the second-order caustics and 1-D Maslov caustics coincide. Hence, expression (2.6.25) cannot give a useful amplitude. We need to consider the 2-D Maslov mapping $(x_1, x_2, x_3) \to (p_1, p_2, x_3)$. Here we assume the second-order caustics are not parallel to the (x_1, x_2) plane (otherwise a proper coordinate rotation can be applied). The phase function in the (p_1, p_2, x_3) domain can be constructed by the partial Lengedre transform,

$$\begin{aligned}\vartheta(p_1,p_2,\boldsymbol{x};\boldsymbol{x}') &= \tau(x_1,x_2,x_3;\boldsymbol{x}') \\ &\quad + p_1(x_1 - X_1(p_1,p_2,x_3)) + p_2(x_2 - X_2(p_1,p_2,x_3)).\end{aligned} \tag{2.6.30}$$

Note that at $x_1 = X_1(p_1,p_2,x_3), x_2 = X_2(p_1,p_2,x_3)$, this phase is stationary, and the stationary phase argument leads to

$$\vartheta((p_{\boldsymbol{x}})_1,(p_{\boldsymbol{x}})_2,\boldsymbol{x};\boldsymbol{x}') = \tau(\boldsymbol{x};\boldsymbol{x}'), \quad \text{and} \quad \nabla_{\boldsymbol{x}}\vartheta((p_{\boldsymbol{x}})_1,(p_{\boldsymbol{x}})_2,\boldsymbol{x};\boldsymbol{x}') = \boldsymbol{p}_{\boldsymbol{x}},$$

2-D Maslov has to be applied to remove the second-order caustics.

$$U(\boldsymbol{x}, \boldsymbol{x'}, \omega) \sim \left(\frac{i\omega}{2\pi}\right) \times \int \mathrm{d}^2 p \, B(p_1, p_2, x_3; \boldsymbol{x'}) \exp[i\omega \, \vartheta(p_1, p_2, \boldsymbol{x}; \boldsymbol{x'})], \quad (2.6.31)$$

We can apply multi-dimensional stationary phase analysis to (2.6.31) as we did in the previous subsection. The same conclusion follows. That is, the contribution at the caustics to modeling is higher order in ω, but to inversion it is lower order in ω.

Note that in regions in physical space where simple solutions of the form, $A \exp\{i\omega\tau\}$ exist, we can obtain them by applying the method of stationary phase to the Maslov representation. The transformation to the (p_1, x_2, x_3) domain (2.6.25) removes only one zero from the Jacobian, and the point at which the Jacobian has two zeros is still a singularity in the (p_1, x_2, x_3) domain. In order to obtain a nonsingular, asymptotic expansion at these points, a second-order transformation (2.6.31) is necessary. Third-order zeros of the Jacobian are not possible for the seismic system as they imply infinite velocity. Equations (2.6.25) and (2.6.31), which involve one and two slowness integrals, respectively, are computationally more expensive than the WKBJ calculation. Therefore, except at or near second-order caustics, Chapman (1982) preferred expression (2.6.25) for numerical calculations. Maslov (1964) developed a method to obtain a uniform asymptotic solution which is valid everywhere and eliminates the problems due to the caustics in any domain. The solution is obtained by simply blending together the various asymptotic solutions, $A \exp\{i\omega\tau\}$ and equations (2.6.25) and (2.6.31), with weighting functions. This is expensive.

Another disadvantage in the Maslov computation is the difficulty of distinguishing between physical caustics and '*pseudo-caustics*'— the singularity points in mixed phase domain (p_1, x_2, x_3) or (p_1, p_2, x_3) that are not coincident with the physical caustics. Theoretically, since rays do not cross in phase space $\boldsymbol{x} \times \boldsymbol{p}$, it is always possible to choose a projection for which they do not cross. However, in the numerical computation, singularities in different domains are often located close to each other, and it is hard to separate the pseudo-caustics from the physical caustics. This limits the application of the Maslov integral asymptotic solution.

On the other hand, carrying out the stationary phase analysis for the Maslov integral representation, we find that the asymptotic Green's function near caustics is larger than the solution away from caustics, largely by being of higher power in ω. Since in the inversion formulas, the two Green's functions appear in the denominator and, hence, the contribution of regions near caustics is lower order in this processing than is the contribution from regions away from caustics. This suggests to us that we can simply zero out the contribution from the neighborhood of caustic points in the migration/inversion process by tapering.

Chapter 3

EXAMPLES OF DYNAMIC RAY TRACING

In this chapter, I provide some sample implementations of the dynamic ray tracing algorithm discussed in the previous chapter. Three examples are included, each example is indicative of a situation that is common in forward modeling and migration/inversion.

3.1 Introduction

For the first test, I construct the wavefield for a model consisting of a single horizontal-interface, with a strong velocity contrast. The numerical output for the smoothed model is compared with analytic solutions for the underlying discontinuous model. The comparison aims to assess the accuracy of the transmitted amplitude, as well as of traveltime. This is the direction of most interest to us for application to Kirchhoff inversion. In this example, both of the numerical outputs with and without transmission coefficient implementation are compared with the corresponding analytic solutions.

The dynamic ray-tracing algorithm is also applied to a model consisting of a low-velocity lens inclusion in an otherwise vertically monotonic velocity distribution. This earth model produces rays that exhibit multi-pathing and different kinds of caustics. Here, the first arrivals at depth are not always the most energetic arrivals. The test demonstrates separation of the various arrivals by the computer code.

The final example is based on a section of the SEG/EAGE saltdome model. The output demonstrates the ability of this method to treat the extreme contrasts across the saltdome interface and into the subsalt region underneath the dome.

3.2 Example I — horizontal-interface model

The preconditioning (smoothing) of the velocity model needed in this method leads to a question of the accuracy of the traveltimes and amplitudes transmitted through a sharp interface. This first simple example tests that accuracy.

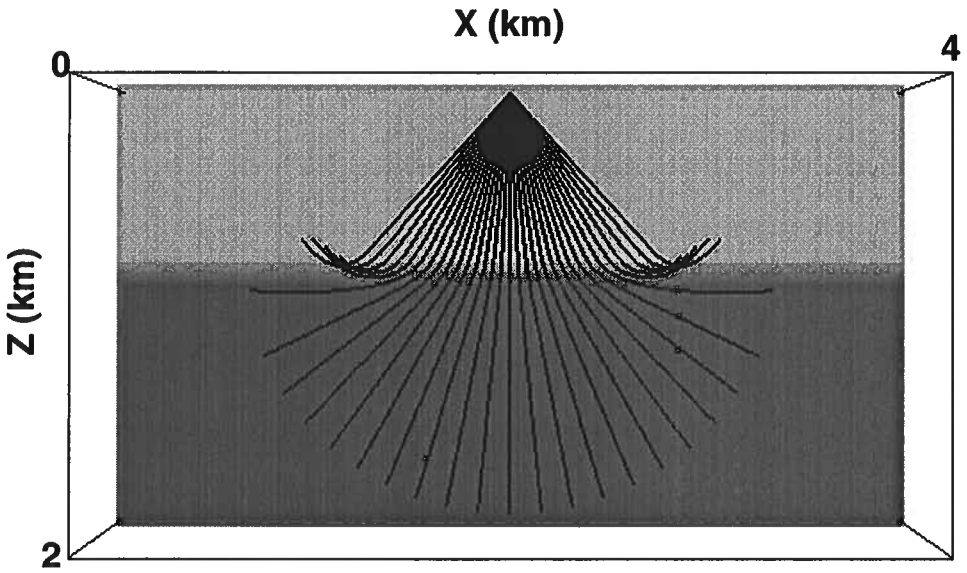

FIG. 3.1. Rays in horizontal-interface model of strong velocity variation. The velocities are 2 km/s in the upper layer and 4 km/s in the lower layer. The region of intermediate shading indicates the effective width of the zone of velocity smoothing across the interface.

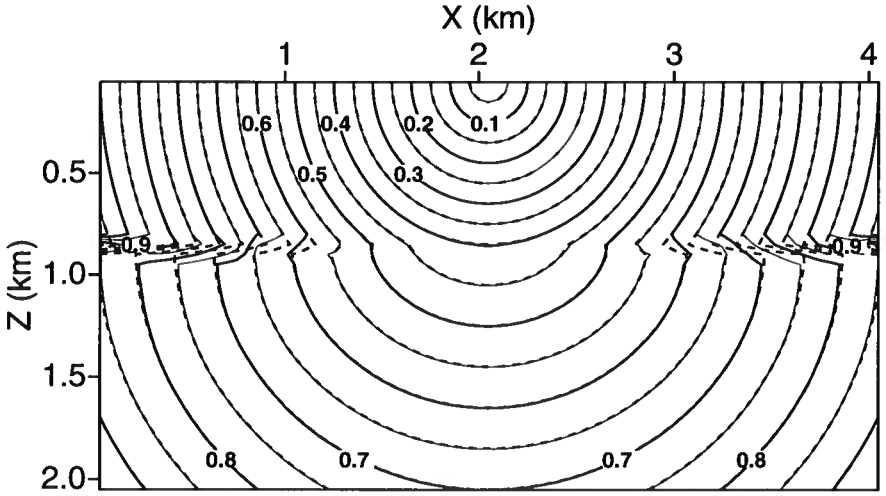

FIG. 3.2. Traveltime comparison for horizontal-interface model. The solid curves depict the in-plane wavefronts, or isochrons, computed from the smoothed model; the dashed curves depict the wavefronts for the analytic solutions.

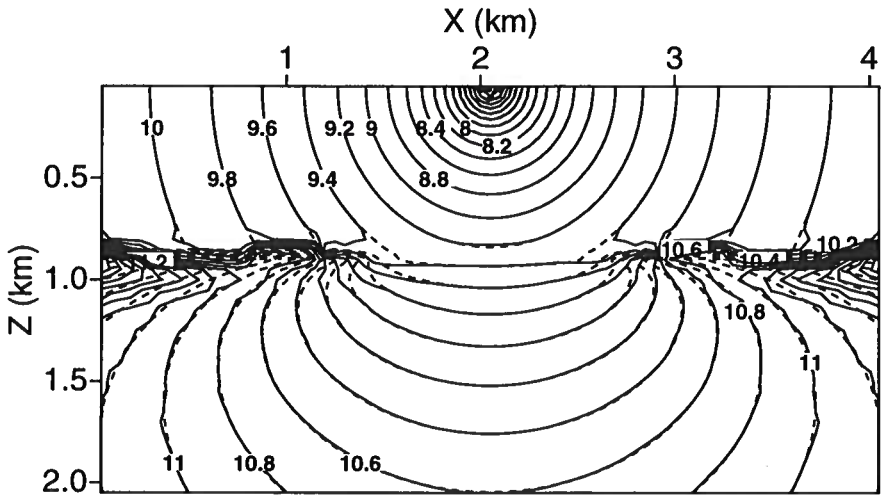

FIG. 3.3. Amplitude (-log) comparison for horizontal-interface model. The solid curves depict the computed level curves of amplitude from the smoothed model and the dashed curves depict the level curves for the analytic solutions, with transmission coefficient set equal to 1 throughout.

Inversion requires a smooth velocity function in propagating regions. Interface shapes must be described by smooth functions as well. Here, I choose an alternative treatment of interfaces. That is, I smooth velocities across those regions, as well, and then introduce an approximate transmission coefficient, as described in Section 2.5. The inversion then delineates the discontinuous surface where the background model was continuous. This ray-tracing method requires a velocity field that has continuous first and second derivatives everywhere. Thus, as the processing moves deeper into the subsurface, the earth model used for ray tracing and the earth model determined by inversion in the overburden disagree. Nonetheless, inversion requires that the transmitted wavefield remain reasonably close to the wavefield of a discontinuous overburden. The example chosen here provides a first test of the agreement of the transmitted wavefield produced numerically by using a smoothed velocity model and an analytical wavefield produced from the underlying discontinuous model.

The velocity model (see Figure 3.1) consists of two homogeneous layers, separated by a horizontal interface. The velocities are 2 km/s in the upper layer, and 4 km/s in the lower layer. Velocities in the region surrounding the interface between the two layers are smoothed by a smoothing operator based on damped least squares (Liu, 1994), so the interface is replaced by a smooth transition zone (see Figure 3.1). This smoothing is actually applied to slowness—inverse velocity. It has been our experience that this provides the most accurate traveltimes and amplitudes. Figure 3.1 shows rays traced in the smoothed velocity model. Rays propagate across the interface zone smoothly, with post-critical rays smoothly refracted upward after entering the

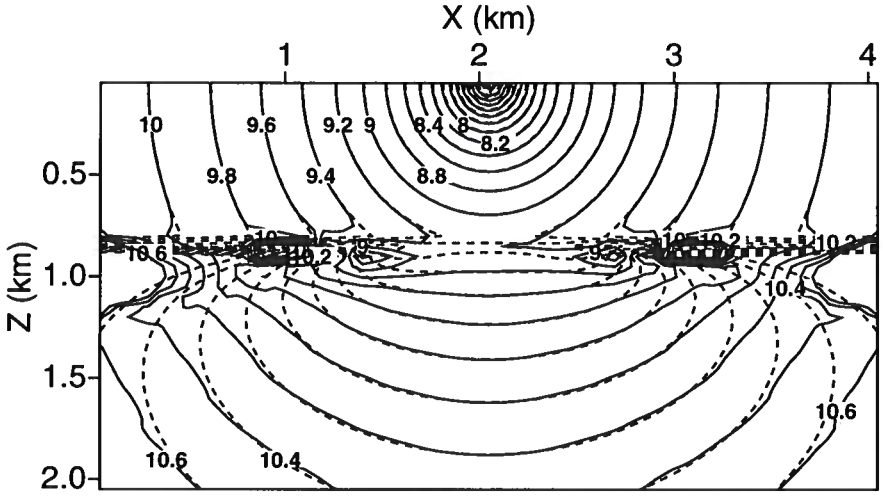

FIG. 3.4. Amplitude (-log) comparison for horizontal-interface model. The solid curves depict the computed level curves of amplitude from the smoothed model and the dashed curves depict the level curves for the analytic solutions, with the transmission coefficient implementation.

smoothed transition zone. The rays with near critical take-off angles refracted into the transition zone and formed a transition wavefield similar to a head wave.

Figure 3.2 shows the traveltime comparison. The dashed curves are the analytic in-plane wavefronts, and the solid curves are computed results. The two solutions are shown to be in good agreement in this figure. This indicates that the ray tracing method can produce accurate traveltimes even in the presence of large velocity contrast in the original model.

From equations (2.2.6) and (2.2.7), we can see that the computation traveltime requires smoothness of velocity gradient while the amplitude requires smoothness up to the second derivatives of the velocity. Therefore, the amplitude accuracy is of greater interest in this example. Solving equations (2.2.7) directly for the two constant-velocity layer model, we can get the following analytic amplitude solution for this model:

$$A(\boldsymbol{x}) = \begin{cases} \dfrac{1}{4\pi R_1}, & z < 0.8\text{km}, \\ \dfrac{T}{4\pi\sqrt{(R_1 + \dfrac{v_2}{v_1}R_2)(R_1 + \dfrac{v_2}{v_1}R_2\dfrac{\cos^2 a}{\cos^2 \tilde{a}})}}, & z \geq 0.8\text{km}, \end{cases} \tag{3.2.1}$$

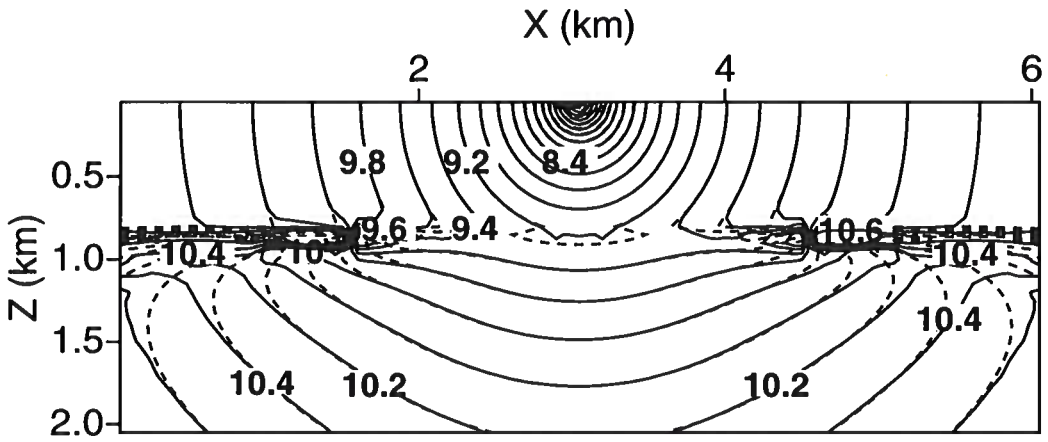

FIG. 3.5. Same as in Figure 3.4 but for a two-layer with velocities above and below the interface are 2 and 2.5 km/s, respectively.

with

$$T = \frac{2v_2 \cos a}{v_2 \cos a + v_1 \cos \tilde{a}}. \tag{3.2.2}$$

Here, a and $\tilde{a}$ are the incident and transmitted angles; v_1 and v_2 are the velocities above and below the interface; R_1 and R_2 are the lengths of the ray paths in the first and second layer with take-off angle a.

Figure 3.3 shows the amplitude comparison, with the dashed curves showing the level curves of amplitude for the analytic solution, and the solid curves used for the computed results. In this comparison, the transmission coefficient across the interface is set equal to one. Thus, I am only checking the agreement of the ray Jacobians for the smooth and the discontinuous velocity models. The only difference in these two solutions is at the near-critical region. We expect this difference arising from the replacement of the sharp interface by the smooth transition zone.

Figure 3.4 shows the amplitude comparison with the numerical calculation of the transmission coefficient included in the numerical solution and the analytical transmission coefficient (3.2.2) included in the analytical solution. Once again, the two solutions agree with each other at the incident wavefield and the transmitted wavefield with small transmission angle. However, the agreement of the two solutions is degraded largely at large transmission angles. Seen from Figure 3.4, the disagreement of the numerical estimate of the transmission coefficients with the solution tends to have an angular dependence. Since the model picked here has strong velocity contrast—the normal reflection coefficient is 0.3, a small range of pre-critical incident angles span to a large range of refracted angle. The instability of the transmission coefficient

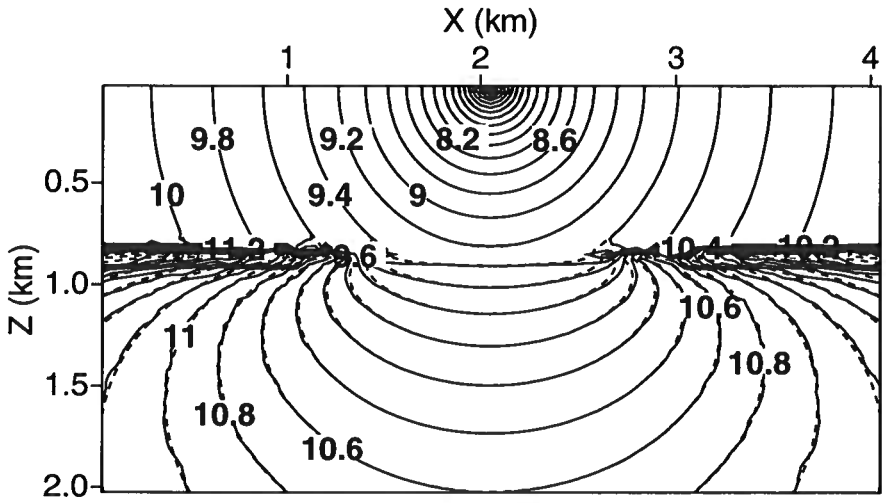

FIG. 3.6. Amplitude (-log) comparison for horizontal-interface model. The solid curves depict the computed level curves of amplitude from the smoothed model and the dashed curves depict the level curves for the analytic solutions, with the transmission coefficient set to 1.

computation may be caused by the rapid changes in the transmitted angles. Therefore, the accuracy checking is carried out on a flat-interface model with a relatively smaller velocity contrast, 2 km/s above the interface and 2.5 km/s below. Figure 3.5 shows the amplitude comparison on such a velocity model. Notice that the numerical estimate of the amplitude agrees with the analytic solution with a wider refracted angle range. For the accuracy check with error less than or equal to 1%, the dip angle is about 15° wider. This indicates that our transmission coefficient computation does depend on the rapidness of velocity variation.

Amplitude accuracy shown in Figures 3.4 and 3.5 is not favorable at large transmitted angles. Further tests are carried out on checking with the following parameters: (1) smoothing length, (2) velocity sampling size, (3) ray tracing step size (in traveltime) and (4) ray coverage density. The smoothing size (the number of velocity samples in the smoothing widow) is tied to the gradient of the velocity through the smoothing region. Too larger a smoothing length gives less accuracy in the transmitted angle and therefore less accurate amplitude. Too small a smoothing length will cause irregular behavior of rays at near critical, because in this case the ray tracing validity condition is not met. The velocity sampling size also has influence on the amplitude accuracy, because the velocities and their derivatives at ray endpoints along the central ray are linearly interpolated from the velocities and their derivatives at the surrounding grid points. This grid size should be related to the gradient of the velocity. However, for finer velocity grid size, smaller ray tracing step size is needed for the transition zone. Finer ray coverage density is also need to make the output amplitude look

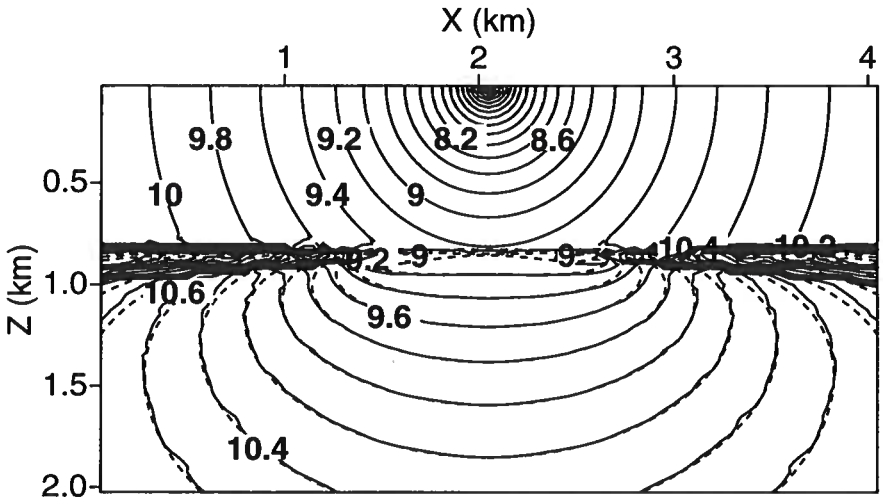

FIG. 3.7. Amplitude (-log) comparison for horizontal-interface model. The solid curves depict the computed level curves of amplitude from the smoothed model and the dashed curves depict the level curves for the analytic solutions, with the transmission coefficient implementation.

smoother for large transmitted angles. These four factor are related to each other, which bring certain difficulty in finding the best combination of them to provide best amplitude accuracy. In addition, finer ray tracing step size and denser ray coverage imply the longer computing time for the same size of target zone of the wavefield. Figures 3.6 and 3.7 show the comparison of computed amplitudes of finer mesh size and the analytic solutions, without and with the transmission coefficients included. In both figures, the numerical output has less (almost no) irregular amplitude above the interface and better agreement with the analytic solutions for the transmitted region.

Figures 3.2 through 3.7 indicate that even in the presence of a sharply discontinuous interface in the original velocity model, our ray tracing method provides accurate transmitted traveltime and amplitude. The traveltime has a better agreement with the analytic solution because it requires less smoothness in velocity than the amplitude computation. When the transmission coefficients are included, our numerical estimate is less accurate. However, with proper choices of the ray tracing parameters, the method can produce the numerical calculation of amplitude with fairly good accuracy.

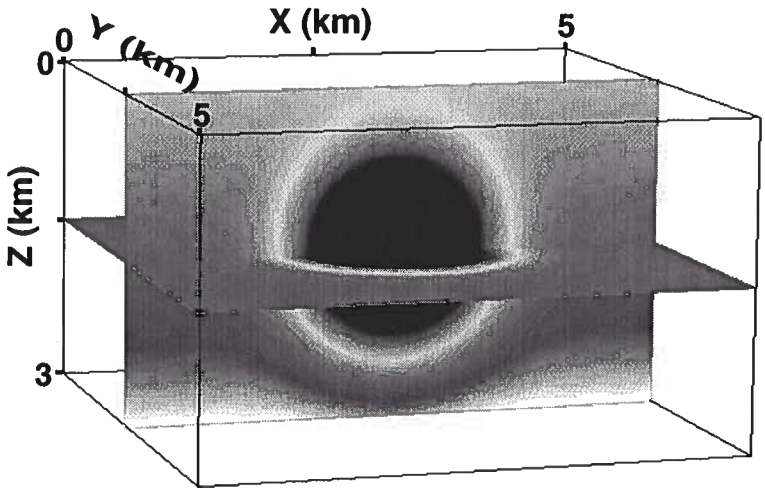

FIG. 3.8. Lens model. A spherical low-velocity lens is wedged into high-velocity surroundings.

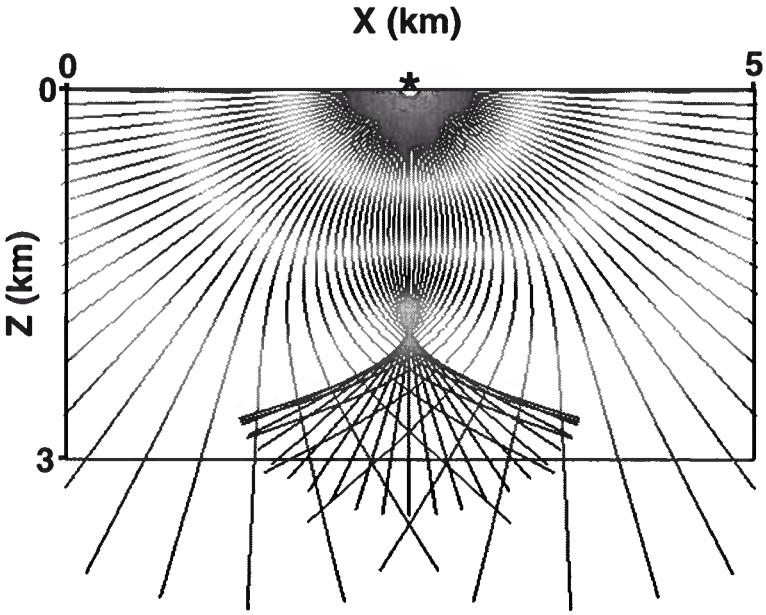

FIG. 3.9. Rays in the lens model. The color represents the traveltime along the ray.

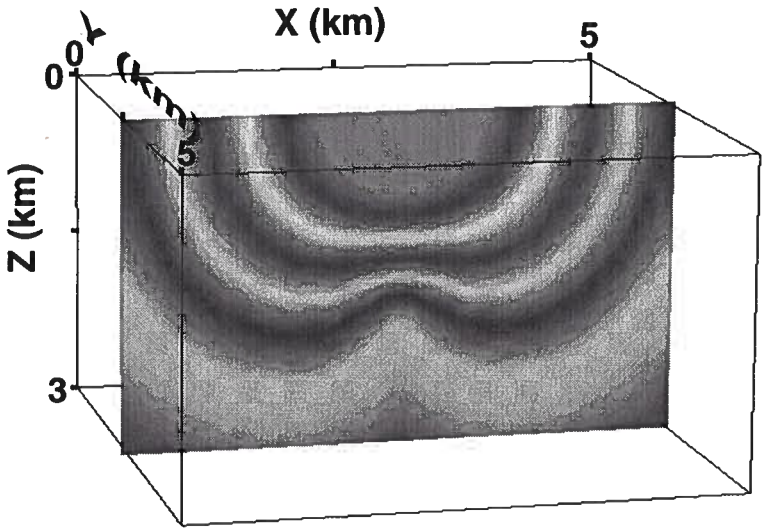

FIG. 3.10. Traveltime of first arrivals in the lens model.

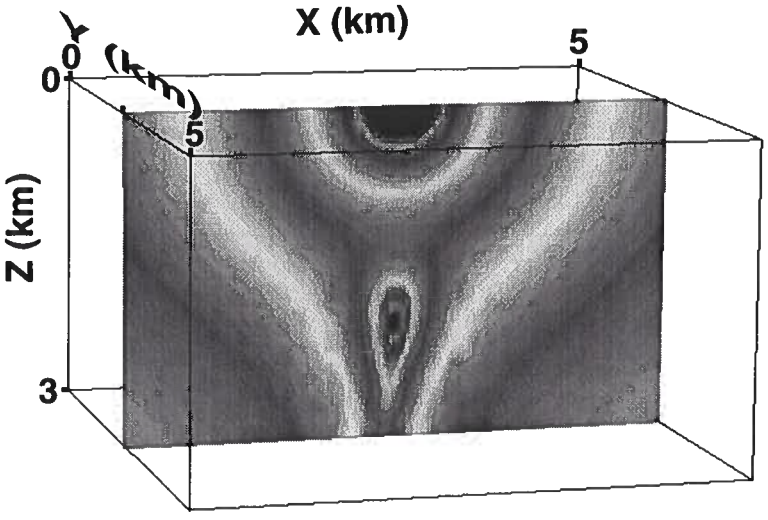

FIG. 3.11. Amplitude of first arrivals in the lens model.

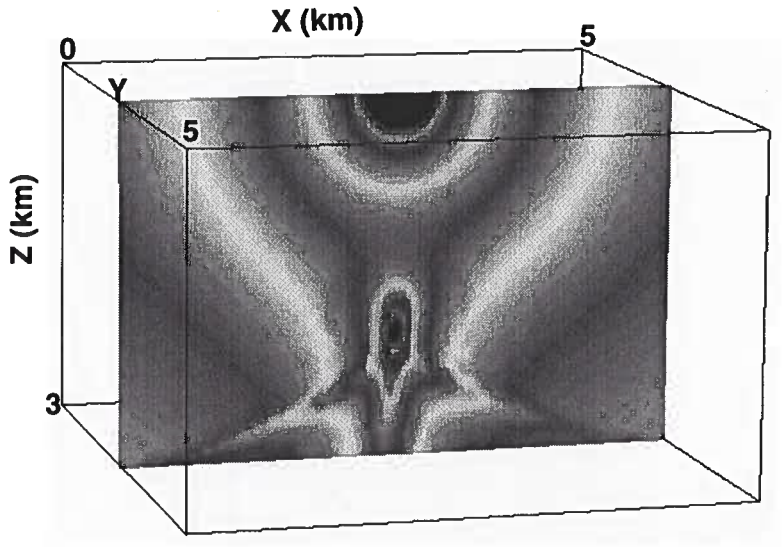

FIG. 3.12. Amplitude of most energetic arrivals in the lens model.

3.3 Example II — lens model

This example demonstrates the ability of the code to deal with multi-pathing and caustics. Multiple arrivals are a common phenomenon in heterogeneous media. One would expect a focusing of energy in the low-velocity region and a defocusing in the surrounding high-velocity region. In the presence of multi-pathing, later arrivals may carry more energy, and the amplitude coefficients exhibit phase shifts arising from passage through caustics. Therefore, it is important for the forward modeling algorithm to have the ability to find all the possible arrivals and amplitude coefficients. It should be remarked that the current code deals with only the more standard caustics of isotropic media. More exotic caustics produced by anisotropy do not lend themselves to adjustment via a simple KMAH index (de Hoop & Brandsberg-Dahl, 1998).

Figure 3.8 shows the velocity distribution of a lens model. A spherical-shape low velocity lens is included in an otherwise high velocity zone, with velocity monotonically increasing with depth. Figure 3.9 shows some in-plane rays, fired at the center of the upper surface. Due to the symmetry of the velocity model, only rays traced in one vertical slice are displayed. The rays bend, and multi-pathing occurs in the lower part of the model. The color represents the traveltime along the rays. Notice that, for equal traveltimes, the rays traveling outside the lens penetrate deeper than do the rays that pass through the lens; the velocity in the lens is lower than in the surrounding medium, leading to longer traveltimes at comparable depths.

Three kinds of caustics are observed in this example. The most visible caustic is the

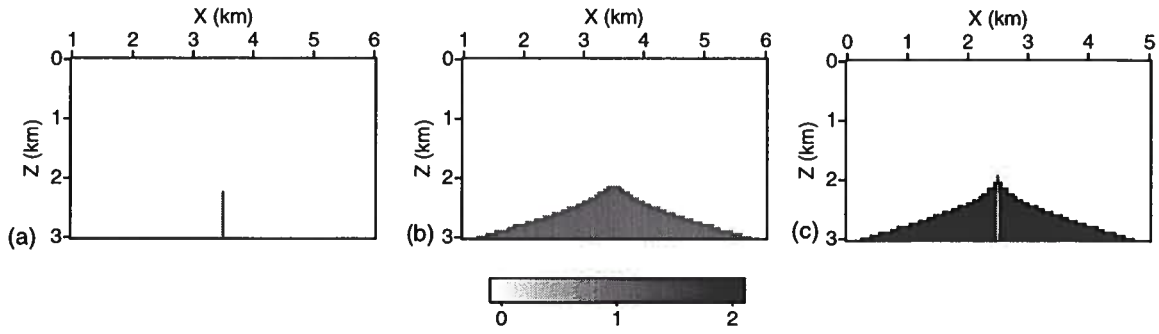

FIG. 3.13. KMAH index of (a) first arrivals; (b) second arrivals; (c) third arrivals.

inverted cusped structure just below 2 km. In the neighborhood of the cusp, itself, rays from each side of the central axis approach the caustic curve, but have not reached it yet and rays have already impinged on the caustic and moved on. As one moves out along the flank of the caustic, below it, there are only two nearby caustic rays, one that has not yet passed through the caustic and the other that has. We emphasize these caustic rays because ray theoretic models that characterize the wavefield here require representations in terms of Airy functions and their generalizations. In this development, we model the wavefield only before and after the caustic and not in the caustic transition zone. Thus there is the cusp of the caustic—the minimum depth of crossing rays—and the smooth flank of the cusped surface. This accounts for two caustics.

The third caustic is less apparent in this figure because it requires 3-D visualization for clearer depiction, but it can still be described in this figure. Below the cusped surface, we see rays crossing on the center vertical line. In 3-D, these are two rays of an entire sheet of rays of a surface of revolution. This is the third type of caustic. That is, rays with the same polar take-off angle and varying azimuthal take-off angle form a caustic on the axis. Varying the polar angle moves the caustic point along the axis.

For each of these caustics, the Jacobian used in amplitude calculations vanishes. Thus, the Jacobian matrix has at least one zero eigenvalue and corresponding eigen-vector that characterize the direction in which the cross-sectional area of a ray tube has shrunk to zero. At the cusp, two eigenvalues vanish, corresponding to variations with respect to the two initial take-off angles. Along the flank of the cusp, the principle direction of vanishing corresponds to the vanishing of the Jacobian with respect to variations in the polar angle, while at the on-axis caustic below the cusp, the vanishing is with respect to variations in the azimuthal angle.

Figure 3.10 shows the first-arrival traveltime. In the triplication region, the shapes of the wavefronts are those generated by connecting the ray endpoints that have been

traced from outside the lens. The pull-up of the wavefronts under the lens is due to the low-velocity in the lens.

Figure 3.11 shows the amplitude of first arrivals. The amplitudes shown in this plot are consistent with the ray coverage in Figure 3.9. The high intensity of amplitude in the region of the cusp characterizes the relative increase in the density of rays. The WKBJ wavefield modeling is not accurate near caustics where the computed amplitude approaches infinity. The code sets a threshold above which the amplitude is replaced by an upper bound to stabilize the code. The difference shown in the amplitude between first- and second- order caustics is caused by the averaging of the ray amplitude inside a ray tube of prescribed size. The actual caustic surfaces form a set of measure zero in 3-D. Thus, we anticipate that this approximation in caustic regions will not seriously distort the amplitude of the migration output, except perhaps, at image points that are also caustic points of the ray fields for the Green's functions from source and receiver. In fact, the entire theory employed for inversion breaks down in the neighborhood of such points, and even the imaging produced by migration alone is somewhat suspect.

Figure 3.12 shows the amplitude of most energetic arrivals. Figures 3.11 and 3.12 show the same amplitudes in the upper part of the figure, including the region on axis directly below the source point. These are regions in which the first arrival *is* the most energetic arrival. On the other hand, the amplitudes are quite different in the triplication region below the cusped caustic. The amplitudes of the most energetic arrivals in this region are formed by those rays that have traveled through the lens. As noted in the ray plot, Figure 3.9, these rays are slower. Hence, the first arrivals and the most energetic arrivals are predictably different here. The difference in position between the first arrival and the most energetic arrival can be 1 km or more.

Figure 3.13 shows the phase shifts in the triplication region, the area surrounded by the cusp surface. The first arrivals reach the receiver grid points from outside the lens (see Figure 3.9). Therefore, no phase shifts arise along the ray until the rays approach the axis, on which rays with the same polar take-off angle and varying azimuthal take-off angle form a caustic with phase shift of $\pi/2$. The second arrivals in time also come from rays outside the lens, but from the opposite side of lens with respect to the center axis. As a result, the second arrivals get a phase shift of $\pi/2$ on the axis and carry it further on to all the grid points in the triplication zone [Figure 3.13(b)]. The third arrivals, the latest arrivals in time, pass through the axis right near the cusp. At the cusp, itself, two eigenvalues vanish. The nearby rays first pass through the caustic on the axis corresponding the vanishing of the eigenvalue with eigenvector in the direction of azimuthal angle. The rays then pass through the caustic along the flanks leading out from the caustic. Here, the eigenvalue associated with the polar angle vanishes. The code is able to detect the two sign changes in the Jacobian, and results in a phase shift of π at each grid points [Figure 3.13(c)], except that the ray penetrating vertically has no shift; the program computes a phase shift

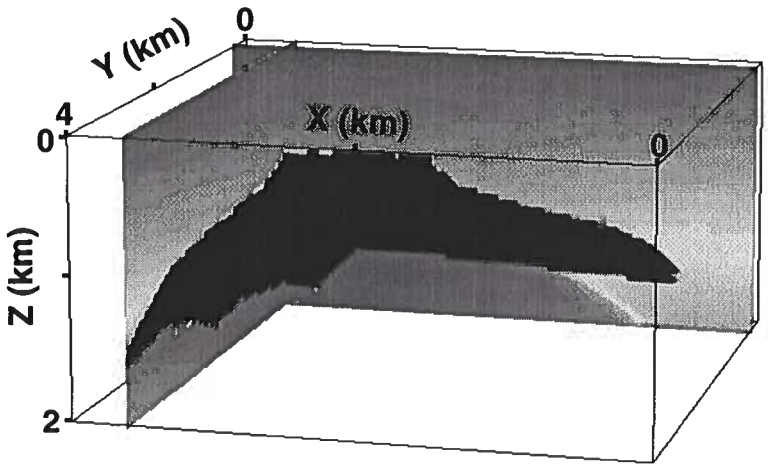

FIG. 3.14. SEG/EAGE salt model. The velocity in the salt (black) is 4482 m/s. 0-4 km in X runs right to left, and 0-4 km in Y runs back to front (GOCAD standard).

of $\pi/2$ at grid points below the source. Modeling of on-axis propagation, however, is highly suspect. The cusp is formed by a limit of rays with polar angle approaching zero. Yet, that limiting ray is associated with a limit of axis-crossings by rays forming the caustic flanks below and to the right and left. Thus the limiting ray, which would seem to have initial polar angle equal to zero crosses the axis only once—at the cusp—while the ray with initial polar angle zero remains on-axis throughout its propagation path.

3.4 Example III — SEG/EAGE model

The final example uses a section of the SEG/EAGE saltdome model, which is based on a typical US Gulf Coast salt structure. The complexity of the 3-D salt structure makes it a benchmark model used to evaluate various 3-D forward modeling and imaging algorithms. Figure 3.14 shows two orthogonal vertical slices of the velocity distribution. The velocity in the salt is 4482 m/s, shown in black, much higher than the velocities in the surrounding sediments. As a consequence, the salt body acts as a large secondary source that emanates waves from its surface, nearly in all directions. Such waves, typically, have low energy; however, due to their speed through the salt, they generate first arrivals in near-salt and sub-salt regions. These waves dominate the solution of first-arrival eikonal solvers. Some of these waves correspond to head-waves; others are just low-energy forward propagating parts of solutions. Using such a traveltime solution for imaging will result in a less than ideal image (Geoltrain & Brac, 1993).

Figure 3.15 shows one wavefront recorded at 0.9 s with the source located at the

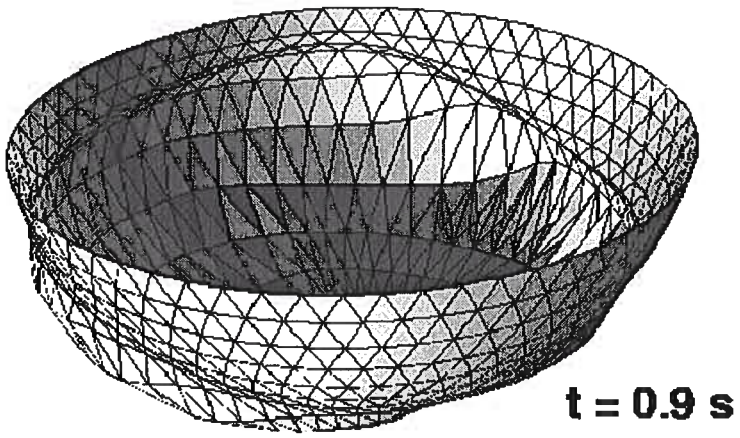

FIG. 3.15. One wavefront recorded at 0.9 s, with the source located at the center of the upper surface.

center of the upper surface. Its lower part changes from being spherical, due to the presence of the high-velocity salt. This structure causes portions of the wavefront to bulge and expand, especially those parts that have penetrated through the salt. The boundaries of this bulge are dominated by the pseudo-head-wave arrivals, which continuously connect the portions of the wavefront that penetrated through the salt with those that traveled only through the sediments. However, headwaves are low-energy arrivals that are not useful for imaging applications. Waves traveling directly from the source, without going through the salt, carry more energy than do rays that propagate through the salt.

Figure 3.16 is a ray plot, for source at the center of the upper surface. The front face of this figure shows a vertical slice at x = 3.5 km, which is 1.5 km from the source. It passes through the salt structure. The rays are shown in this plot, shot off with a uniform distribution of take-off angles. The feature of adding rays to maintain ray density was suppressed in this figure so that we could see the relative distribution of the set of rays around the salt structure, which are initially uniform in take-off angles. Notice that the density of the post-critical rays that do not enter the salt is much higher than that of the rays that enter the salt and emerge through the front face. This indicates that we should expect the amplitude in the salt to be lower than the amplitude just above the salt. Figure 3.18 is an amplitude plot on the same vertical slice as the front face of Figure 3.16. Here, the red and yellow colors indicate high amplitude and the purple color, low amplitude. The high- and low- amplitude zones agree with what we expected from the ray plot in Figure 3.16. We know from the previous discontinuous example that the amplitudes near turning are not as accurate as are the transmitted amplitudes, but they are close enough qualitatively to confirm the agreement we see between these two figures. Figure 3.17 shows the traveltime

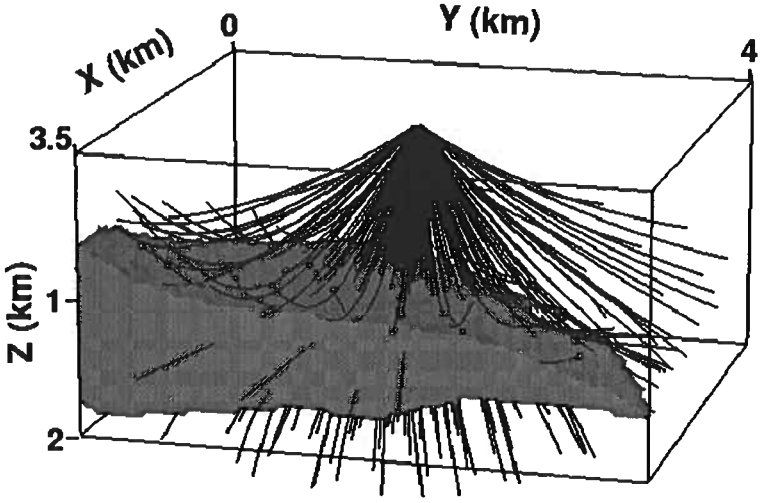

FIG. 3.16. Rays in the salt model. This model is rotated when compared to Figure 3.14. The front face is approximately the Y-Z plane that is shaded in the former figure.

on the same cross-section. The rapid changes of the wavefronts occur at the salt boundary.

3.5 Conclusions

We have demonstrated the implementation of a ray-theoretical forward modeling algorithm on three progressively more complex earth models. This method is based on dynamic ray tracing and wavefront construction in smoothed gridded earth models. The use of smoothing applied to the discontinuous model stabilizes the ray tracing while maintaining sufficient accuracy on transmitted traveltime and amplitude coefficients. The method of estimation transmission coefficients from smooth gridded model provide us an alternative treatment of interfaces. The method is capable of calculating all typical ray-tracing parameters such as traveltime, ray amplitude, slowness vector, etc. It can produce all the quantities involved in Kirchhoff inversion formula. The adaptability of the method was demonstrated by evaluating the wavefield in regions of extreme geometrical spreading, such as near a salt wall, in sub-salt regions, and in presence of a low velocity lens.

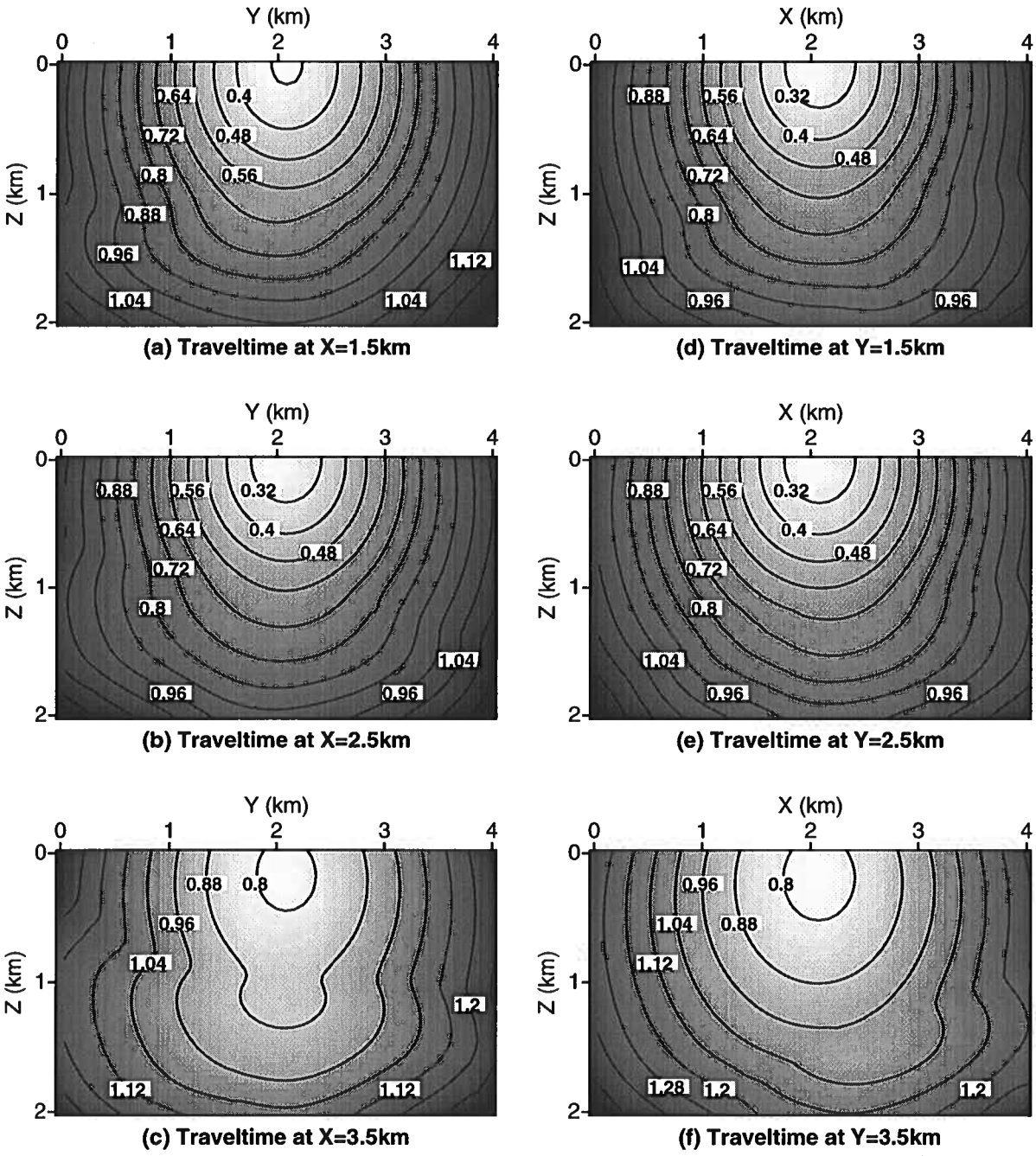

FIG. 3.17. Traveltime slices for the salt model.

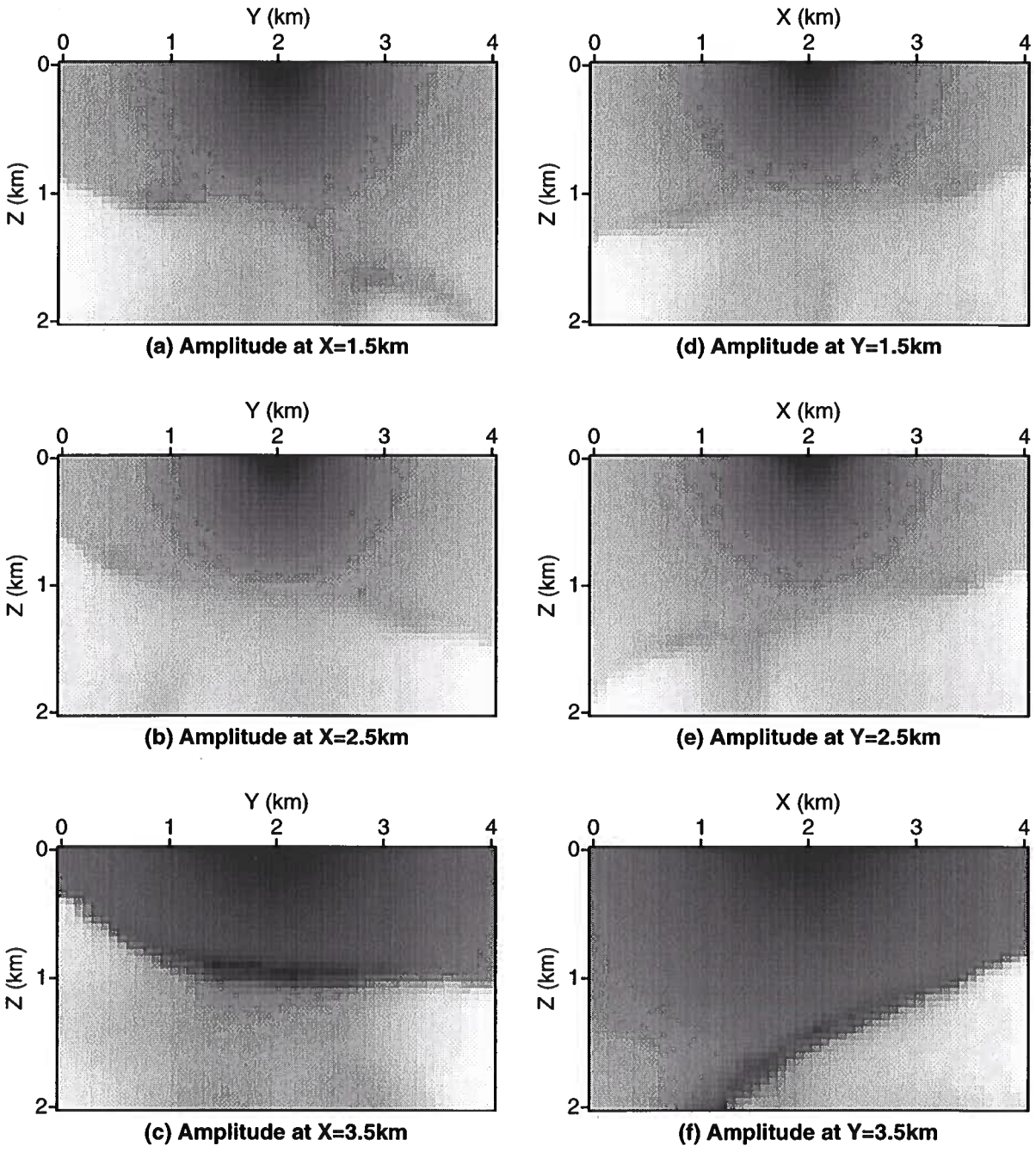

FIG. 3.18. Amplitude slices for the salt model.

Chapter 4

THEORY OF KIRCHHOFF INVERSION

In this chapter, I discuss the theory of inverting the linearized scattered field for the image of subsurface reflectors and medium parameter variations.

Some basic concepts of prestack migration/inversion are first provided in Section 4.1. Section 4.2 contains a short review of the approach of Bleistein and Cohen to Kirchhoff inversion and the technique of multiple Kirchhoff inversion. The latter offers a way to estimate the reflection angles on reflectors. In Section 4.3, I discuss the numerical estimation of the Beylkin determinant. The computation is realized by the ray-theoretical methods and can be carried out easily by the dynamic ray-tracing method discussed in Chapter 2. In Section 4.4, I derive the expression of the Beylkin determinant for a special case—zero-offset. In this case, the weighting factor can be largely simplified and reduced to a slowly varying factor.

4.1 Basic concepts

4.1.1 Diffraction stacking

Seismic migration is the process that maps the seismograms recorded on the Earth's surface into representations of its interior properties. Kirchhoff migration approaches this goal by linear inversion of the wave equation in terms of integral operators. A mathematical description of such a process can be written as

$$m(\boldsymbol{y}) = \int_A \mathrm{d}^2\xi\, w(\boldsymbol{\xi}, \boldsymbol{y}) \mathcal{F}[U(\boldsymbol{\xi}, \tau_D(\boldsymbol{\xi}, \boldsymbol{y}))]. \tag{4.1.1}$$

Here, A is the seismogram recording surface, parameterized by $\boldsymbol{\xi} = (\xi_1, \xi_2)$. The distribution of sources and receivers is specified by equations $\boldsymbol{x_s} = \boldsymbol{x_s}(\boldsymbol{\xi})$ and $\boldsymbol{x_g} = \boldsymbol{x_g}(\boldsymbol{\xi})$; U is the seismic data recorded along this surface; $\tau_D(\boldsymbol{\xi}, \boldsymbol{y})$ is the sum of traveltimes from source and receiver to the imaging point $\boldsymbol{y}$ (see Figure 4.1); $\tau_D(\boldsymbol{\xi}, \boldsymbol{y})$ is also called the **diffraction time surface**. The filter $\mathcal{F}$ denotes the time derivative in 3-D, or the Hilbert transform of the half-derivative in 2.5-D, of the seismic data U. $w(\boldsymbol{\xi}, \boldsymbol{y})$ is a spatial weighting function that contributes to the integral summation. Therefore, Kirchhoff-type migration/inversion algorithms can be viewed as **weighted**

diffraction summation (or **stacking**), which is performed for each grid point $\boldsymbol{y}$ in the target zone to be migrated. Schneider (1978), Berryhill (1979) and Berkhout (1985) provide extensive references to the Kirchhoff migration method. Descriptions of the Kirchhoff migration process from the geophysical point of view can be found in Yilmaz (1987) and Claerbout (1996).

The specific function $w(\boldsymbol{\xi}, \boldsymbol{y})$ distinguishes the output $m(\boldsymbol{y})$ among various migration/inversion algorithms. When $w(\boldsymbol{\xi}, \boldsymbol{y}) = 1$, or unweighted stacking, the stack $m(\boldsymbol{y})$ gives a structural image of the medium. On the other hand, when the weight function is carefully formulated, the weighted stacking can predict the medium parameters, in addition to a graphical image of the medium. This is referred as **Kirchhoff inversion** by Bleistein (2000). The weighting function in an inversion algorithm includes the model-consistent geometrical loss—**true amplitude**—and the **Beylkin determinant**, which accounts for any source-receiver configuration and parameterization in the background propagation speed. This last multiplier could not have been predicted by a migration approach to the high-frequency inverse problem. In this thesis, I study the importance and the properties of the weighting factors.

4.1.2 Linearity

The forward scattering problem can be formulated as a sum of two integrals. The first involves the perturbation as a linear factor of the integral. The second involves a product of the perturbation and the unknown scattered field at depth. It is argued that this product is at least quadratic in terms of perturbation, hence nonlinear in the perturbation. This nonlinear term can be ignored provided that perturbations of the medium parameters are small. This is the **Born approximation** [see Bleistein *et al.* (2000) for further discussion]. Therefore, the linearization in forward modeling means that we consider only the linear term of the forward scattering integral equation. This linear term physically corresponds to the **primary reflections**—energy that has been reflected only once and hence is not a multiple. This assumption implies that all other seismic events, such as multiple reflections, direct waves, refracted waves, should be viewed as noise and should be attenuated during preprocessing or properly interpolated in the output.

4.1.3 High frequency

The term "high-frequency data" means that length scales of interest in a medium must be "many" (in practice, at least three) times as large as the predominant inverse wavenumber propagating in the medium. When considering the resolution of narrowly spaced reflectors, the high-frequency condition translates into the familiar Rayleigh criterion for resolution (i.e., bed thickness $\geq$ wavelength/4). However, the theory does

not fail catastrophically if the high-frequency condition is not exactly met (Bleistein *et al.*, 2000).

Mathematically, the high-frequency approximation implies the use of **asymptotic methods** to create high-frequency formulations of the forward and inverse problems (Bleistein, 1984; Bleistein *et al.*, 2000). The Born approximation, as mentioned earlier, is a small-perturbation assumption that is compatible with the high-frequency assumption. The **WKBJ approximation** discussed in Chapter 2 is another fundamental tool in the forward modeling. WKBJ-approximate Green's functions are high-frequency solutions to the wave equation that assume that the properties of the medium are slowly varying over several wavelengths. The **Kirchhoff approximation** is an assumption relating high-frequency incident and scattered fields. The **method of stationary phase** is an important tool for finding approximate analytic solutions of Fourier-like integrals and for testing inversion formulas.

It should be noted that linearization and the high-frequency assumption are present in migration, just as they are in inversion. A "small-perturbation" assumption is hidden within the migration process. Furthermore, migration is based on the ideas of wavefronts, reflectors, and raypaths—all concepts of high-frequency wave theory.

4.1.4 True amplitude

The aim of a Kirchhoff inversion is to use amplitudes in seismic reflection data in order to invert for a physical parameter at the position of reflectors, e.g., the reflection coefficient. Amplitudes in seismic reflection data are usually influenced by a variety of factors (Sheriff, 1975). These include source and receiver factors (e.g., strength and coupling, geophone sensitivity), geometrical spreading, transmissions, reflections, attenuation, thin-layer tuning, scattering, and anisotropy, to name a few. The techniques presented below automatically correct the amplitude variations of geometrical spreading and recover values proportional to the angularly-dependent reflection coefficient, provided the uncorrected effects are negligible. These other effects can be corrected for, if they are quantitatively known.

4.1.5 Migration versus inversion

The words "migration" and "inversion" are, in fact, used to describe related processes. To some people, seismic migration is strictly an imaging process through which the seismic data are mapped, or "migrated" into their proper location. To others, the migration and inversion processes derive maps of local reflectivity from seismic data. To still others, migration is a method of exploratory data analysis while inversion aims to make quantitative inferences—to characterize the entire set of models that fit the

data and are *a priori* reasonable— about the Earth's properties. It is now apparent that all accurate imaging methods can be viewed essentially as linearized inversions of the wave equation, whether in terms of Fourier integral or direct gradient-based optimization of a waveform misfit function. People in the exploration community who practice migration usually are not able to obtain or preserve the true amplitudes of the data, because of the sequence of processing steps that make up a migration (trace equalization, gaining, deconvolution, etc). Therefore, they use the traveltime as the only important parameter and image the reflectors so that the position and shape of reflectors are more correctly represented. However, there is no attempt to recover information about the material parameters of the subsurface in traditional migration procedures.

On the other hand, if we had true-amplitude data, then we could make quantitative statements about how spatial variations in reflector strength were related to changes in geological properties. The distinction here is the distinction between imaging reflectors, on the one hand, and doing a true inverse problem for the subsurface properties on the other.

The theories in Bleistein em et al. (2000) unify the approach of Kirchhoff-type migration/inversion algorithms, and provide the exact expression for the weighting factor needed to recover the medium parameters as well as an image of structure, subject to the assumed underlying model of forward propagation.

4.1.6 Other geophysical concepts

A **common-offset gather** consists of a collection of seismograms whose respective source-receiver separation is a constant value (**offset**) on all traces. In this thesis, I only consider flat datum surface. In this case, we can select one distinguished direction, say, ξ_1-direction, as the orientation of the common-offset vector. The source and receiver positions are then written as,

$$\boldsymbol{x_s} = \boldsymbol{x_s}(\xi_1 - |h|, \xi_2), \qquad \boldsymbol{x_g} = \boldsymbol{x_g}(\xi_1 + |h|, \xi_2), \tag{4.1.2}$$

with $|h|$ denoting the constant common-offset.

The common-offset gather is one of the source-receiver configurations used broadly in the oil industry. Other data collections include the shot record or common-shot (source) gather—a range of receivers from a single source—and common-receiver gather—a single receiver from a range of sources.

Band limiting and aperture limiting of data The bandlimited nature of the inverse problem can be caused by many factors. These include the frequency response of the seismic source and receiver, the degree of coupling between the source/receiver,

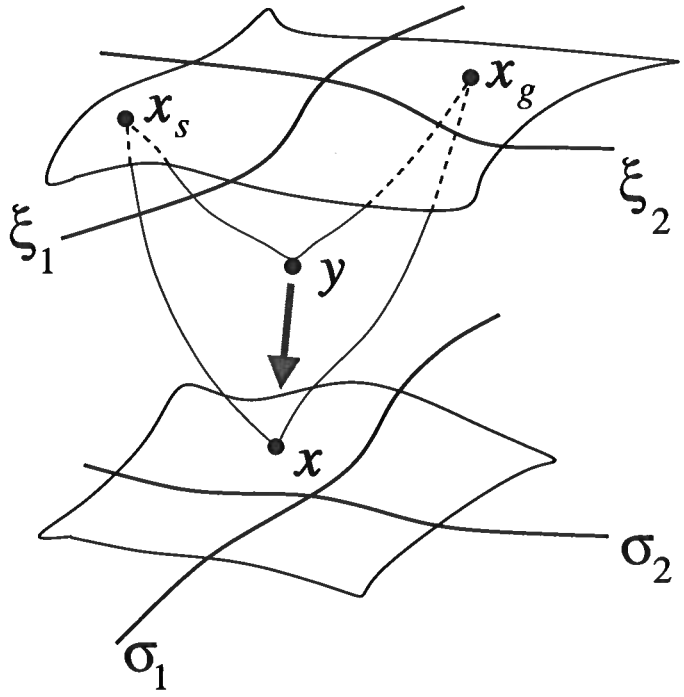

FIG. 4.1. Graphical representation of the "stationary triple" composed of the source position $\boldsymbol{x_s}$, receiver position $\boldsymbol{x_g}$, and reflection point $\boldsymbol{x}$ on the reflecting surface (Bleistein *et. al.*, 2000).

the propagating medium "earth filter," and anelastic filtering. Because of the limitations imposed by the high-frequency assumption, the ability to resolve structures of a particular length scale is limited by the frequency band that is available.

In addition to the bandlimiting in the frequency domain, the recording geometry introduces a limitation in the wavenumber domain. The aperture limiting may also degrade the resolution of the image created by inversion, and it gets worse as the opening angle between the incident ray and the reflected ray increases—which typically occurs as the offset between source and receiver increases.

4.2 The approach of Bleistein-Cohen

In this section, a brief review of the derivation of the the approach of Bleistein-Cohen is provided. The Bleistein-Cohen approach of Kirchhoff inversion (Bleistein *et al.*, 1987; Bleistein, 1987; Bleistein *et al.*, 2000), is based on WKBJ ray theory and high-frequency linearized inversion. The derivation of the theory involves the asymptotic ray theory, perturbation theory and the method of stationary phase. In the implementation, the method uses specific spatial weights in the diffraction stack migration,

and returns an angular dependent reflection coefficient, in addition to a background model-consistent reflector map of the structure interfaces. The computation time for such an approach is slightly greater than that needed for kinematic prestack depth migration, but it introduces a practical method for estimation of quantitative information on reflectors (namely, reflection coefficient or impedance contrast).

4.2.1 Review of the derivation

The Bleistein-Cohen inversion procedure starts from a linearized forward modeling formula and derives a pseudo-inversion of that formula. The linearized forward modeling formula derives from the acoustic Born theory with WKBJ approximation. It is written in terms of the unknown material parameters that describe the process of generating the seismic data. The inversion uses a multidimensional stationary-phase method to solve for the unknown material parameters.

In the Bleistein-Cohen theory, the unknown material parameter is the acoustic wavespeed perturbation, $\alpha(\boldsymbol{x})$, a measure of the difference between the background velocity $c(\boldsymbol{x})$ and the true velocity $v(\boldsymbol{x})$:

$$\alpha(\boldsymbol{x}) \equiv \frac{c^2(\boldsymbol{x})}{v^2(\boldsymbol{x})} - 1. \tag{4.2.3}$$

The Bleistein-Cohen inversion starts with the Born-approximate representation of the recorded scattered wavefield, u_S, in terms of the velocity perturbation $\alpha(\boldsymbol{x})$,

$$u_S(\boldsymbol{x_g}, \boldsymbol{x_s}, \omega) = \omega^2 \int_D \mathrm{d}^3x \, \frac{\alpha(\boldsymbol{x})}{c^2(\boldsymbol{x})} u_I(\boldsymbol{x_g}, \boldsymbol{x_s}, \omega) g(\boldsymbol{x_g}, \boldsymbol{x_s}, \omega). \tag{4.2.4}$$

Here, $u_I(\boldsymbol{x_g}, \boldsymbol{x_s}, \omega)$ is the incident wavefield and $g(\boldsymbol{x_g}, \boldsymbol{x_s}, \omega)$ is the free space Green's function corresponding to the point source at $\boldsymbol{x_s}$. In this forward modeling formula, a term that contains the product of the field u_S and the perturbation $\alpha(\boldsymbol{x})$ is ignored. This corresponds to the linearization of the forward scattering problem that is identified as the Born-approximation. It assumes moderate variation of the medium parameter; Bleistein *et al.* (2000) has shown how to substantially remove this restriction.

The next step is to invoke high-frequency asymptotics and apply the WKBJ (ray-theoretic) Green's function. The approximations take the form,

$$g(\boldsymbol{x}, \boldsymbol{x}_0, \omega) \sim A(\boldsymbol{x}, \boldsymbol{x}_0) e^{i\omega\tau(\boldsymbol{x}, \boldsymbol{x}_0)} e^{-i\pi \mathrm{sgn}(\omega)\varphi(\boldsymbol{x}, \boldsymbol{x}_0)/2}. \tag{4.2.5}$$

Here, $\tau(\boldsymbol{x}, \boldsymbol{x}_0)$ is the traveltime from $\boldsymbol{x}_0$ to $\boldsymbol{x}$ and $A(\boldsymbol{x}, \boldsymbol{x}_0)$ is the corresponding amplitude of the ray-theoretic Green function, derived by solving, respectively, the

eikonal equation (2.2.2) and the (first) transport equation (2.2.3); $\varphi(\boldsymbol{x}, \boldsymbol{x}_0)$ is the KMAH index from $\boldsymbol{x}_0$ to $\boldsymbol{x}$ due to the existence of caustics.

Substitution of the WKBJ Green's functions into equation (4.2.4) yields the scattered wave field at $\boldsymbol{x}_g$ for the point source at $\boldsymbol{x}_s$,

$$u_S(\boldsymbol{x}_g, \boldsymbol{x}_s, \omega) \approx \omega^2 F(\omega) \int \mathrm{d}^3 x \frac{\alpha(\boldsymbol{x})}{c^2(\boldsymbol{x})} a(\boldsymbol{x}, \boldsymbol{\xi}) \mu(\boldsymbol{x}, \boldsymbol{\xi}, \omega) e^{i\omega\phi(\boldsymbol{x}, \boldsymbol{\xi})}, \tag{4.2.6}$$

where

$$\begin{aligned} \phi(\boldsymbol{x}, \boldsymbol{\xi}) &= \tau(\boldsymbol{x}, \boldsymbol{x}_s(\boldsymbol{\xi})) + \tau(\boldsymbol{x}_g(\boldsymbol{\xi}), \boldsymbol{x}), \\ a(\boldsymbol{x}, \boldsymbol{\xi}) &= A(\boldsymbol{x}, \boldsymbol{x}_s(\boldsymbol{\xi})) A(\boldsymbol{x}_g(\boldsymbol{\xi}), \boldsymbol{x}), \\ \mu(\boldsymbol{x}, \boldsymbol{\xi}, \omega) &= e^{-i\pi \mathrm{sgn}(\omega)[\varphi(\boldsymbol{x}, \boldsymbol{x}_s) + \varphi(\boldsymbol{x}_g, \boldsymbol{x})]/2}. \end{aligned} \tag{4.2.7}$$

The exponential term $\mu(\boldsymbol{x}, \boldsymbol{\xi}, \omega)$ accounts for the phase action of caustics.

Bleistein *et al* (2000), Chapter 5, contains a detailed derivation of the inversion of (4.2.6). By diffraction theory, *all* points on the reflecting surface (not just a single specular point) act as Huygens sources that contribute to the field observed at a receiver. For a given source-receiver pair, and a given point on the reflector, the integrand is the product of source strengths times down-going wave times the up-going wave. Integrating over all the source/receiver pairs, and relating the down-going and up-going amplitudes, phases, and obliquity, and reflection coefficients, yields the inversion for reflectivity function

$$\begin{aligned} \beta(\boldsymbol{y}) \;=\; & \frac{1}{8\pi^3} \int \mathrm{d}^2 \xi \frac{|h(\boldsymbol{y}, \boldsymbol{\xi})|}{a(\boldsymbol{y}, \boldsymbol{\xi}) |\nabla_y \phi(\boldsymbol{y}, \boldsymbol{\xi})|} \int i\omega \, \mathrm{d}\omega \\ & \cdot e^{-i\omega\phi(\boldsymbol{y}, \boldsymbol{\xi})} \mu(\boldsymbol{y}, \boldsymbol{\xi}, \omega) u_S(\boldsymbol{x}_g, \boldsymbol{x}_s, \omega). \end{aligned} \tag{4.2.8}$$

Here,

$$h(\boldsymbol{y}, \boldsymbol{\xi}) = \det \begin{bmatrix} \nabla_y \phi(\boldsymbol{y}, \boldsymbol{\xi}) \\ \frac{\partial}{\partial \xi_1} \nabla_y \phi(\boldsymbol{y}, \boldsymbol{\xi}) \\ \frac{\partial}{\partial \xi_2} \nabla_y \phi(\boldsymbol{y}, \boldsymbol{\xi}) \end{bmatrix} = \det \begin{bmatrix} \boldsymbol{p}_s + \boldsymbol{p}_g \\ \frac{\partial}{\partial \xi_1} (\boldsymbol{p}_s + \boldsymbol{p}_g) \\ \frac{\partial}{\partial \xi_2} (\boldsymbol{p}_s + \boldsymbol{p}_g) \end{bmatrix} \tag{4.2.9}$$

is the Beylkin determinant. In the theoretical development, $\omega^2 |h(\boldsymbol{y}, \boldsymbol{\xi})|$ is the Jacobian of transformation between the variables $(\omega, \boldsymbol{\xi})$ and a local wavenumber, $\boldsymbol{k}$, defined by

$$\boldsymbol{k} = \omega \nabla \tau(\boldsymbol{y}, \boldsymbol{\xi}).$$

Equation (4.2.8) holds for quite arbitrary migration configurations because it contains the configuration-dependent parameter $\boldsymbol{\xi}$ and the Beylkin determinant, which reflects the influence of the configuration on the inversion. The reflectivity function $\beta(\boldsymbol{y})$ has been shown (Bleistein, 1987; Bleistein *et al.*, 2000) to consist of bandlimited delta functions having peak amplitudes occurring at the location of the reflector, S, with size scaled by the incidence angularly-dependent reflection coefficient. That is,

$$\beta_{\text{PEAK}} \sim R(\boldsymbol{y}, \theta_s) \frac{\cos\theta_s}{\pi c(\boldsymbol{y})} \int_{-\infty}^{\infty} F(\omega) \mathrm{d}\omega, \qquad \text{for } \boldsymbol{y} \text{ on } S, \tag{4.2.10}$$

where θ_s is half of the angle between the directions of rays to the source-receiver pair at reflection point $\boldsymbol{y}$.

4.2.2 Multiple Kirchhoff inversion

The structure of equation (4.2.8) allows us to alter the weighting factor and estimate reflection angles or the actual reflected rays. This can be done by simultaneously applying the diffraction stack with different weights on the same data and the same diffraction surfaces. Bleistein (1987; 1987) was the first to recognize this idea and to apply it to obtain reflection angles in addition to reflection coefficients. There, a 'double diffraction stack' technique is applied to estimate the reflection angle from two simultaneously performed diffraction stacks. Tygel *et al.* (1993) introduced the general name **multiple diffraction stack** for this type of technique. Hanitzsch (1995) applied this idea to 2.5-D amplitude-preserving Kirchhoff migration/inversion, and found valuable applications in AVO and AVA analysis.

Taking into account the relationship $|\nabla_y \phi(\boldsymbol{y}, \boldsymbol{\xi})| = 2\cos\theta_s / c(\boldsymbol{y})$, and introducing an extra divisor of $|\nabla_y \phi|$ in (4.2.8), Bleistein *et al.* (1987) found the following inversion operator,

$$\begin{aligned} \beta_1(\boldsymbol{y}) \;=\; & \frac{1}{8\pi^3} \int \mathrm{d}^2\xi \frac{|h(\boldsymbol{y}, \boldsymbol{\xi})|}{a(\boldsymbol{y}, \boldsymbol{\xi}) |\nabla_y \phi(\boldsymbol{y}, \boldsymbol{\xi})|^2} \\ & \int i\omega \, \mathrm{d}\omega \, e^{-i\omega\phi(\boldsymbol{y}, \boldsymbol{\xi})} \mu(\boldsymbol{y}, \boldsymbol{\xi}, \omega) u_S(\boldsymbol{x_g}, \boldsymbol{x_s}, \omega). \end{aligned} \tag{4.2.11}$$

$\beta_1(\boldsymbol{y})$ predicts the reflection coefficients at the reflector, S,

$$\beta_{1\text{PEAK}} \sim R(\boldsymbol{y}, \theta_s) \int_{-\infty}^{\infty} F(\omega) \mathrm{d}\omega, \qquad \text{for } \boldsymbol{y} \text{ on } S. \tag{4.2.12}$$

From the ratio of the outputs of $\beta(\boldsymbol{y})$ and $\beta_1(\boldsymbol{y})$, we can estimate $\cos\theta_s$ without ever having determined the specular source-receiver pair that produced the distinguished value θ_s. This, in turn, allows us to determine $R(\boldsymbol{y}, \theta_s)$ from the output. In Chapter 5,

my implementation of Kirchhoff inversion is actually based on (4.2.11).

In fact, instead of dividing the weight in $\beta(\boldsymbol{y})$ by $|\nabla_y \phi|$, we can multiply by $|\nabla_y \phi|$, which leads to,

$$\beta_2(\boldsymbol{y}) = \frac{1}{8\pi^3} \int \mathrm{d}^2\xi \frac{|h(\boldsymbol{y},\boldsymbol{\xi})|}{a(\boldsymbol{y},\boldsymbol{\xi})} \int i\omega \, \mathrm{d}\omega \, e^{-i\omega\phi(\boldsymbol{y},\boldsymbol{\xi})} \mu(\boldsymbol{y},\boldsymbol{\xi},\omega) u_S(\boldsymbol{x_g},\boldsymbol{x_s},\omega) \quad (4.2.13)$$

and

$$\beta_{2\mathrm{PEAK}} \sim R(\boldsymbol{y},\theta_s) \frac{c(\boldsymbol{y})}{4\pi \cos\theta_s} \int_{-\infty}^{\infty} F(\omega) \mathrm{d}\omega, \qquad \text{for } \boldsymbol{y} \text{ on } S. \quad (4.2.14)$$

The ratio of the two outputs, β_1 and β_2, can as easily be used to determine $\cos\theta_s$ as the ratio proposed above.

4.3 Beylkin determinant

One difficulty in the implementation of equations (4.2.8), (4.2.11), and (4.2.13) for general media and non-zero offset has been the estimation of Beylkin determinant h [equation (4.2.9)]. For the common-shot or common-receiver geometry, $\partial \boldsymbol{p_s}/\partial\xi_1 = \partial \boldsymbol{p_s}/\partial\xi_2 = 0$ (common-shot) or $\partial \boldsymbol{p_g}/\partial\xi_1 = \partial \boldsymbol{p_g}/\partial\xi_2 = 0$ (common-receiver), the Beylkin determinant can be simplified and computed efficiently either by direct numerical approximation of (4.2.9) or by rewriting (4.2.9) in terms of ray quantities. Implementations of common-shot/receiver inversion for 2-D, 2.5-D and 3-D have been carried out (Docherty, 1987a; Docherty, 1991; Dong *et al.*, 1991; Sun & Gajewski, 1997).

For the common-offset geometry, each element in the Beylkin matrix contains a sum of quantities from the source and the receiver. Bleistein *et al.* (2000) decompose the Beylkin determinant into eight terms, some of which can be recognized as being the Beylkin determinants for the common-shot and common-receiver geometries. Unfortunately, the remaining terms, which contain the cross products $\frac{\partial \boldsymbol{p_s}}{\partial\xi_1} \times \frac{\partial \boldsymbol{p_g}}{\partial\xi_2}$ and $\frac{\partial \boldsymbol{p_g}}{\partial\xi_1} \times \frac{\partial \boldsymbol{p_s}}{\partial\xi_2}$, are not easily simplified. Only solutions in simple background medium are found. Xu (1996) carried out common-offset inversion in a depth-dependent background medium. Lateral homogeneity allows him to write the traveltime, amplitude and weights as functions of offset and depth.

Červený and de Castro (1993) proposed a method for computing the Beylkin determinant by dynamic ray tracing. Appendix A contains a review and discussion of it. The derivation includes finding the mixed partial derivatives with respect to the distance vector from the two-point eikonal traveltime expression. The mixed derivatives are found to be written in terms of the ray-centered basis vectors and geometrical

spreading matrix $\boldsymbol{Q}$ [see (A.1.20) and (A.1.17)]. Then the Beylkin determinant can be written in terms of these matrices and the source-receiver configuration matrix by the chain rule. However, one assumption in the derivation of the key formula in (Červený & de Castro, 1993) leads to a computationally intensive implementation. In Appendix A.2, I point out the problem of that assumption and provide my own derivation of the correct expression of the Beylkin determinant. In summary, the Beylkin determinant (4.2.9) can be expressed as the following:

$$h(\boldsymbol{x},\boldsymbol{\xi}) = \det \begin{bmatrix} \boldsymbol{p_s}^T + \boldsymbol{p_g}^T \\ \hat{\boldsymbol{B}}_s^{MT} + \hat{\boldsymbol{B}}_g^{MT} \end{bmatrix}, \tag{4.3.15}$$

where

$$\begin{aligned} \hat{\boldsymbol{B}}_s^{MT} &\equiv \hat{\boldsymbol{\Gamma}}_s^{MT} \hat{\boldsymbol{H}}_s(\boldsymbol{x_s}) \hat{\boldsymbol{I}}^M \boldsymbol{Q}_2^{-1,T}(\boldsymbol{x_s},\boldsymbol{x}) \hat{\boldsymbol{I}}^{MT} \hat{\boldsymbol{H}}_s^T(\boldsymbol{x}), \\ \hat{\boldsymbol{B}}_g^{MT} &\equiv \hat{\boldsymbol{\Gamma}}_g^{MT} \hat{\boldsymbol{H}}_g(\boldsymbol{x_g}) \hat{\boldsymbol{I}}^M \boldsymbol{Q}_2^{-1,T}(\boldsymbol{x_g},\boldsymbol{x}) \hat{\boldsymbol{I}}^{MT} \hat{\boldsymbol{H}}_g^T(\boldsymbol{x}). \end{aligned} \tag{4.3.16}$$

Here, $\hat{\boldsymbol{\Gamma}}_s$ and $\hat{\boldsymbol{\Gamma}}_g$ denote constant configuration matrices (A.2.21). $\boldsymbol{Q}_{2,s}$ and $\boldsymbol{Q}_{2,g}$ are 2×2 point-source geometrical spreading matrices, which are quantities in the computation of the Green's function amplitude [see equation (2.2.7) in Chapter 2]. $\hat{\boldsymbol{H}}_s$ and $\hat{\boldsymbol{H}}_g$ contain the rotation and projection of rays centered in Cartesian coordinates at source and receiver [equation (2.2.8)]; $\hat{\boldsymbol{I}}^M$ is a 3×2 matrix [see equation (A.1.15)]; it is part of the 3×3 identity matrix. Detailed explanations of the definitions of the quantities are included in Appendix A and Section 2.2.

Equations (4.3.15) and (4.3.16) are completely general. They hold for arbitrary source-receiver configuration and arbitrary data acquisition surface. For specific source-receiver measurement configurations along a particular data surface, equations (4.3.15) and (4.3.16) can be simplified further. Appendix A.3 contains an expression for $\hat{\boldsymbol{\Gamma}}_s^M$ and $\hat{\boldsymbol{\Gamma}}_g^M$. For a planar data acquisition surface, and common-offset configuration, we have $\hat{\boldsymbol{\Gamma}}_s^M = \hat{\boldsymbol{\Gamma}}_g^M = \hat{\boldsymbol{I}}^M$. Then equation (4.3.16) becomes

$$\begin{aligned} \hat{\boldsymbol{B}}_s^{MT} &= \boldsymbol{H}_s(\boldsymbol{x_s}) \boldsymbol{Q}_{2,s}^{-1T}(\boldsymbol{x_s},\boldsymbol{x}) \hat{\boldsymbol{H}}_s^{MT}(\boldsymbol{x}) \\ \hat{\boldsymbol{B}}_g^{MT} &= \boldsymbol{H}_g(\boldsymbol{x_g}) \boldsymbol{Q}_{2,g}^{-1T}(\boldsymbol{x_g},\boldsymbol{x}) \hat{\boldsymbol{H}}_g^{MT}(\boldsymbol{x}). \end{aligned} \tag{4.3.17}$$

Equation (4.3.17) is the expression of the Beylkin determinant that will be implemented in next chapter. The advantage of this expression is that it can be computed directly from dynamic ray tracing. $\boldsymbol{Q}_{2,s}$ and $\boldsymbol{Q}_{2,g}$ are the geometrical-spreading matrices in the determination of the amplitudes. There is no need to introduce computation of new quantities. Thus, the computing times for the elements in Beylkin matrix bring

only marginally extra cost to the dynamic ray tracing. This method does not require the numerical measurements of any elementary area connected with the ray field, or the determination of the derivatives of traveltimes, which play an important role in some other methods (Schleicher *et al.*, 1993; Sun & Gajewski, 1997).

For a homogeneous medium, we have $\boldsymbol{Q}_{2,s}(\boldsymbol{x_s}, \boldsymbol{x}) = cr_s\boldsymbol{I}$, and $\boldsymbol{Q}_{2,g}(\boldsymbol{x_g}, \boldsymbol{x}) = cr_g\boldsymbol{I}$. So, for a plane datum surface and constant background wavespeed, the Beylkin determinant, h, can be considerably simplified to,

$$h = \frac{(1+\cos\theta)x_3}{c^3}\left[\left(\frac{1}{r_s}+\frac{1}{r_g}\right)\cdot\left(\frac{1}{r_s^2}+\frac{1}{r_g^2}\right)\right]. \tag{4.3.18}$$

Here, θ is the angle between the directions of rays to the source $\boldsymbol{x_s}$ and to the receiver $\boldsymbol{x_g}$ at point $\boldsymbol{x}$. Note that θ is twice the angle of θ_s in equations (4.2.10), (4.2.12) and (4.2.14). Equation (4.3.18) agrees with the corresponding result in Bleistein *et al.* (2000).

4.3.1 Comparison with weights of Schleicher, 1993

Schleicher *et al.* (1993) proposed an alternative expression for the weights in terms of spatial derivatives of traveltimes in a regular grid. In that approach, the weights can be computed by using Green's function computation methods performed in regular grids. It is suitable for planar configurations—shots and receivers on a surface, e.g., the Earth's surface. We should note that the weights in Schleicher *et al.* (1993) contain differences in the second derivatives of traveltimes. Green's function computation methods that compute only traveltimes (e.g., finite-difference solutions of the eikonal equation or shortest-path methods) are not suitable because of the numerical instability of the computation of the second derivatives. However, efficient methods that compute both traveltimes and their derivatives should, in principle, be well suited. The inversion formula in Schleicher *et al.* (1993) is the same as the Bleistein-Cohen inversion formula except for certain extensions. Schleicher *et al.* commented that the exponential term accounting for caustics was not considered in Bleistein *et al.* (1987), and they discussed the extension to isotropic elastic inversion. However, Bleistein-Cohen inversion can account for caustics as explicitly included in equations (4.2.6), (4.2.8), (4.2.11) and (4.2.13). Schleicher *et al.* also commented that the number of ray-branch caustics should be computed by dynamic ray tracing. That is, when caustics are under consideration, dynamic ray tracing is unavoidable. On the other hand, Sun and Gajewski (1998) provide a method of compute the ray-branch caustics from the second-order mixed-derivative matrix of the traveltime function. Nevertheless, among various forward modeling methods, dynamic ray tracing generally provides the better compromise between amplitude precision and computational efficiency. Thus, expressions (4.3.15) through (4.3.17) are more likely to give better

approximation than the second-order derivatives of traveltime in (Schleicher *et al.*, 1993).

4.4 Special case — zero-offset inversion

A number of simplifications occur when the offset between the source and receiver is zero. For this geometry

$$\hat{\boldsymbol{H}}_s(\boldsymbol{x}_s) = \hat{\boldsymbol{H}}_g(\boldsymbol{x}_g), \qquad \hat{\boldsymbol{H}}_s(\boldsymbol{x}) = \hat{\boldsymbol{H}}_g(\boldsymbol{x}), \qquad \boldsymbol{Q}_{2,s}(\boldsymbol{x}_s, \boldsymbol{x}) = \boldsymbol{Q}_{2,g}(\boldsymbol{x}_g, \boldsymbol{x}),$$

$$\boldsymbol{p}_s(\boldsymbol{x}) = \boldsymbol{p}_g(\boldsymbol{x}) = \frac{1}{c(\boldsymbol{x})}\boldsymbol{e}_3^s(\boldsymbol{x}), \qquad |\nabla_x \phi(\boldsymbol{x}, \boldsymbol{\xi})| = \frac{2\cos\theta_s}{c(\boldsymbol{x})}.$$

Therefore,

$$\begin{aligned} h &= \frac{8}{c(\boldsymbol{x})} \det \begin{bmatrix} \boldsymbol{e}_3^s(\boldsymbol{x}) \\ \boldsymbol{H}_s(\boldsymbol{x}_s)\boldsymbol{Q}_{2,s}^{-1T}(\boldsymbol{x}_s, \boldsymbol{x})\hat{\boldsymbol{H}}_s^{MT}(\boldsymbol{x}) \end{bmatrix} \\ &= \frac{8}{c(\boldsymbol{x})} \cdot \det(\boldsymbol{H}_s(\boldsymbol{x}_s)) \cdot \det(\boldsymbol{Q}_2^{-1T}(\boldsymbol{x}_s, \boldsymbol{x})). \end{aligned} \tag{4.4.19}$$

In this case, the inversion formula for $\beta_1(\boldsymbol{y})$ is

$$\beta_1(\boldsymbol{y}) = \frac{4}{\pi c(\boldsymbol{y})} \int \mathrm{d}^2\xi \cos a_s \int i\omega \, \mathrm{d}\omega \, e^{-i\omega\phi(\boldsymbol{y},\boldsymbol{\xi})} \mu(\boldsymbol{y}, \boldsymbol{\xi}, \omega) u_S(\boldsymbol{x}_g, \boldsymbol{x}_s, \omega), \tag{4.4.20}$$

with a_s being the initial take-off angle at the source(receiver).

Note that (4.4.20) holds for a general medium. The medium is not assumed to be a homogeneous. Equation (4.4.20) tells us that the characteristic part of the weight function for zero-offset inversion is just $\cos a_s$, a slowly varying factor ranging between zero and one.

4.5 Conclusions

The Beylkin determinant in Kirchhoff inversion is formulated in terms of ray-theoretic quantities. Thus, the weight function can be totally determined by dynamic ray tracing. The advantage of writing the Beylkin in terms of ray data is that it can be easily implemented by dynamic ray tracing. This method does not require the numerical measurements of any elementary area connected with the ray field, and/or the determination of the derivatives of traveltimes which play an important role in

some other methods that appeared in the recent issues of Geophysics.

Chapter 5

IMPLEMENTATION OF KIRCHHOFF INVERSION

In this chapter, I describe the computer processing for common-offset Kirchhoff inversion and I also carry out several examples to explore different characteristics.

5.1 True amplitude synthetic data

As described in the previous chapter, the Kirchhoff inversion formula is based on model-consistent forward modeling. The recorded seismogram $u_S(\boldsymbol{x_g}, \boldsymbol{x_s}, \omega)$ is assumed to be **true amplitude**, containing the effects of geometrical spreading and reflection/transmission loss due to the subsurface geological properties. The weighting factor in the inversion formula removes these effects as well as the source-receiver effects, and recovers the angular-dependent reflection coefficient. In this chapter, to implement and test the theory, I do forward modeling of data that contain the true amplitudes.

CSHOT (Docherty, 1987b) calculates true-amplitude shot data in 2.5-dimensional layered acoustic media. The source wavelet is zero phase. At a receiver, the source wavelet is scaled by an amplitude factor determined from spreading, reflection, and transmission. CSHOT also takes into account phase shifts due to post-critical reflections and caustics.

The ideal forward data has an infinite spectrum in the frequency domain. But in practice, the data can only be of finite bandwidth. With a zero-phase source wavelet, the reflection events are sinc-like symmetric wavelets with peaks at the location of the exact two-way traveltimes. After diffraction stacking with the proper weighting factor, I expect that the migrated/inverted section has wavelet peaks located at the true depths of the reflectors, scaled by the angular-dependent reflection coefficients.

The peak amplitudes of the synthetic data produced by CSHOT are governed by geometrical spreading, angular-dependent reflection and two-way transmission coefficients. Specifically, the peak amplitude for the primary reflection event from the *nth* reflector can be written as,

$$A_{peak,n} \sim \mathcal{A}_c \times R_n(\theta_n) \prod_{i=1}^{n-1} T_{twoway,i}(\theta_i) \times \frac{1}{G}, \tag{5.1.1}$$

Here G is the spreading factor due to the ray propagation. The scaling factor $\mathcal{A}_c$ is determined by parameters such as frequency spectrum of the source wavelet, the number of sampling points and wavelet length. The study of the exact expression for the $\mathcal{A}_c$ is out of the scope of this thesis. However, in each experiment the factor $\mathcal{A}_c$ is the same for all the reflection events and all the traces. In particular, it remains unchanged from the zero-offset experiment to the nonzero-offset experiment, and is independent of the physical velocity model. For example, in the first example in Section 5.3 (Figure 5.2 and 5.5), factor $\mathcal{A}_c$ equals to 54.678 for 701 samples per trace at 2-ms sample interval, 80-ms wavelet length and the source wavelet of 5~45 Hz in frequency spectrum with the lower bound spanning from 5 up to 15 Hz and the upper bound spanning from 35 down to 45 Hz.

In equations (4.2.8), (4.2.11), and (4.2.13), the input data first pass through a frequency filter $i\omega$, which is equivalent to a normal differentiation in the spatial domain. In this process, a scaling factor that results from the specific software applied. In fact, there is no need to look for the exact value of this scaling factor. Instead, I will seek the overall scaling factor that results from the combination of the *const* from the synthetic input data and the *const* from the frequency filter. This can be done by performing a zero-offset inversion to get the overall scaling factor, and then use this scaling factor in the nonzero-offset inversion to get an estimation of the reflection coefficients.

5.2 Inverse processing sequence

For the common-offset inversion formula in (4.2.11), the computer process is similar to a traditional kinematic Kirchhoff migration algorithm. It proceeds as follows:

1. *Preprocessing of the data.* The ω integration in (4.2.11) represents a filtering and inverse transform of the trace $u_S(\boldsymbol{x_s}, \boldsymbol{x_g}, \omega)$. The frequency filter $i\omega$ behaves asymptotically like a normal derivative operator in the spatial domain. Because this is independent of the ξ integration, the computation can be done as a preprocessing step.

2. *Computation of traveltime and amplitude tables for each source-receiver pair (midpoint).* The information needed in the computation of the weighting factor in (4.2.11) includes traveltime, amplitude, and Beylkin determinant at each grid point for each source/receiver point. For a 3-D nonzero common-offset inversion,

the Beylkin determinant is computed from a 3×3 matrix with each element containing a sum of quantities from the source and the receiver [see equations (4.3.15), (4.3.16) and (4.3.17)]. The nine elements in the Beylkin matrix for each source/receiver can be determined by eight quantities . These quantities can be computed by the dynamic ray tracer described in Chapter 4. Therefore, tables of traveltime, amplitude, and Beylkin matrix elements from each source and receiver are pre-computed and stored in files.

3. The aperture A, which consists of a 2-D grid of points in the $\boldsymbol{\xi}$-plane, and the target zone, V, of the image (the part within the background model for which the Kirchhoff inversion shall be performed) are determined.

4. *Summation over trace locations.* The input data after preprocessing Step 1, is read in trace by trace. For each trace, tables of the traveltime, amplitude and Beylkin matrix at the source and receiver location are input or interpolated from neighboring tables, if they are not already in the memory. For each trace, the following steps are repeated for every subsurface point $\boldsymbol{y}$ within target zone V that falls inside the aperture.

   (a) The diffraction traveltime is computed at each $\boldsymbol{y}$. It is the sum of traveltimes along both ray branches to the source and the receiver specified by source-receiver configuration. The diffraction traveltime at $\boldsymbol{y}$ will give a time index in the trace.

   (b) The weighting factor at each output grid point that falls inside the aperture is computed. Usually, the table spacing is sparser than the image spacing and the table grid points are not coincident with the output grid points. Linear interpolation is used to determine the weighting factor at each output grid point for that particular source point.

   (c) The data value in the trace is obtained by a diffraction traveltime index and then linearly interpolating between two adjacent samples on that trace. The contribution to the inversion sum from that particular trace location is obtained by multiplying this value with weights.

   Finally, summing these contributions over all the possible midpoints (traces) that meet certain anti-aliasing criteria, we obtain the angularly dependent reflection coefficient at the particular output point.

5.3 Example 1 — horizontal-interface model

The first example provides a simple test of positioning and amplitude accuracy. The model consists of four constant-velocity layers, as shown in Figure 5.1. Figure 5.2 shows the zero-offset model data produced by CSHOT. It consists of 200×200 traces,

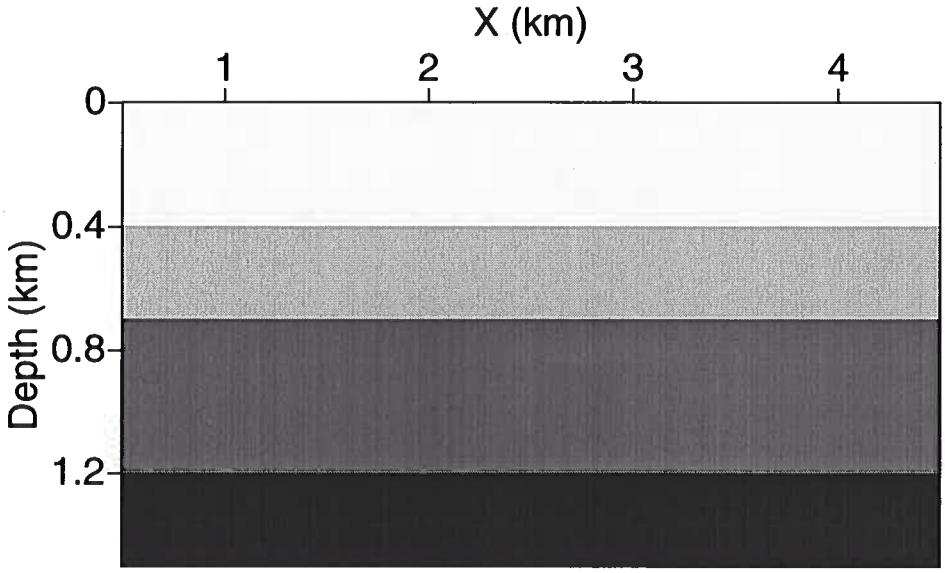

FIG. 5.1. Velocity model consists of four constant-velocity layers with three horizontal interfaces. From top to bottom, the layer velocities are 2, 3, 5 and 7.5 km/s.

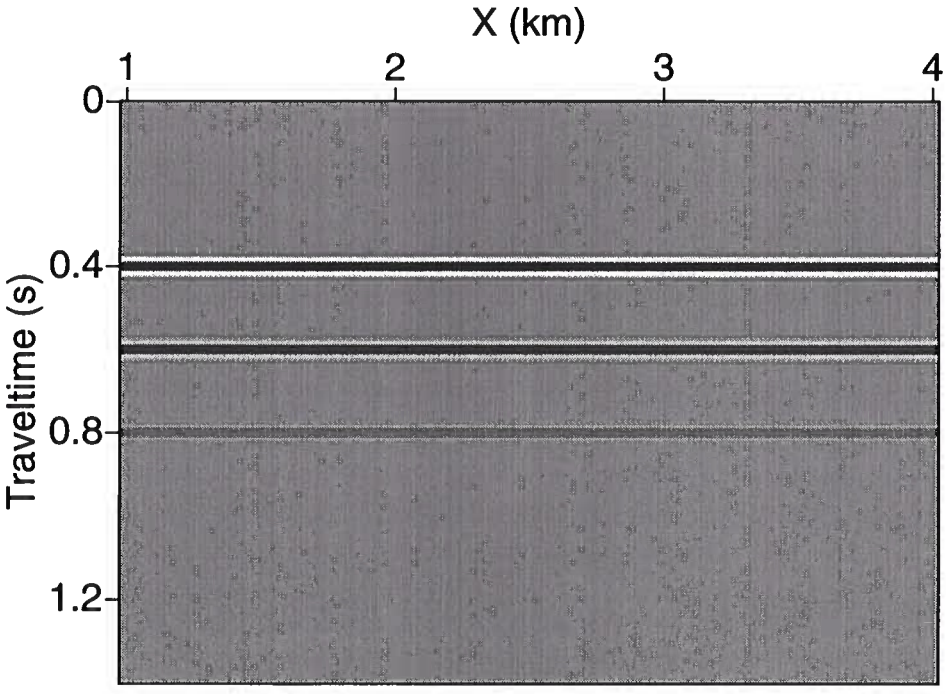

FIG. 5.2. Zero-offset synthetic data of horizontal-interface model.

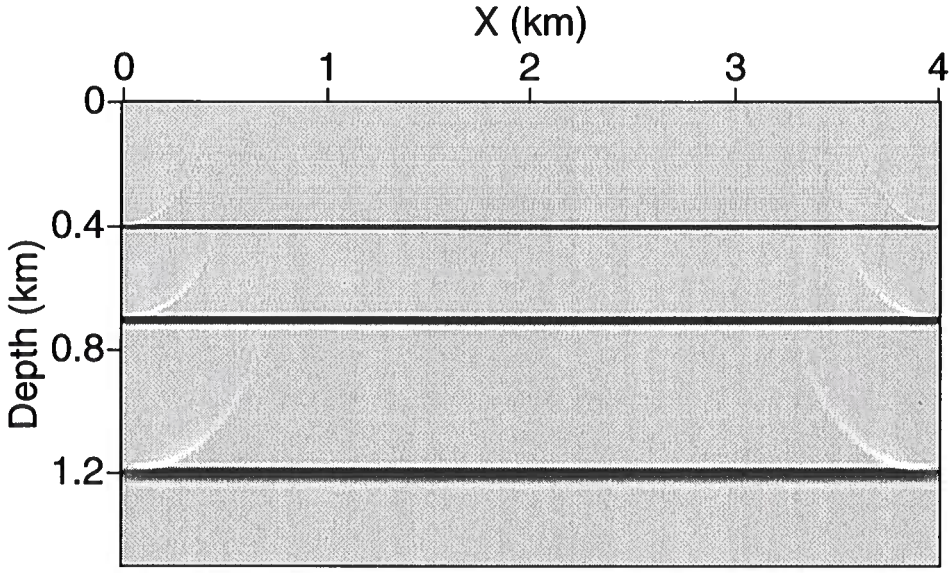

FIG. 5.3. Output β_1 for the zero-offset data.

with 701 samples per trace at 2-ms sample interval and 20-m trace interval in both x and y directions. Here, the synthetic data have amplitudes that are proportional to those influenced by geometrical spreading, reflection, and transmission.

Figure 5.3 shows the result for β_1—the local reflectivity function. The inversion results for R and the reflector depths are compared with the exact values in Table 5.1. The scaling factor is computed by taking the ratio of the exact and computed reflection coefficients. The actual scaling factor at the three interfaces is not important. The significant result is that they are close to each other. That is, the ratios of computed reflection coefficients for the three interfaces are close to the correct values. From the results in the table, we will use the average scaling factor to be 29.04 for the nonzero-offset inversion. This factor is a combination of the factor $\mathcal{A}_c$ in CSHOT amplitude (5.1.1) and a scaling factor from the frequency filter $i\omega$ in (4.2.11).

Interface	True d(epth) (km)	Inversion d (km)	True R_t	Inversion R_i	R_t/R_i
1	0.4	0.4	0.2	0.006879	29.07
2	0.7	0.7	0.25	0.008608	29.04
3	1.2	1.2	0.2	0.006891	29.02

Table 5.1 True and inversion results for reflection coefficients R.

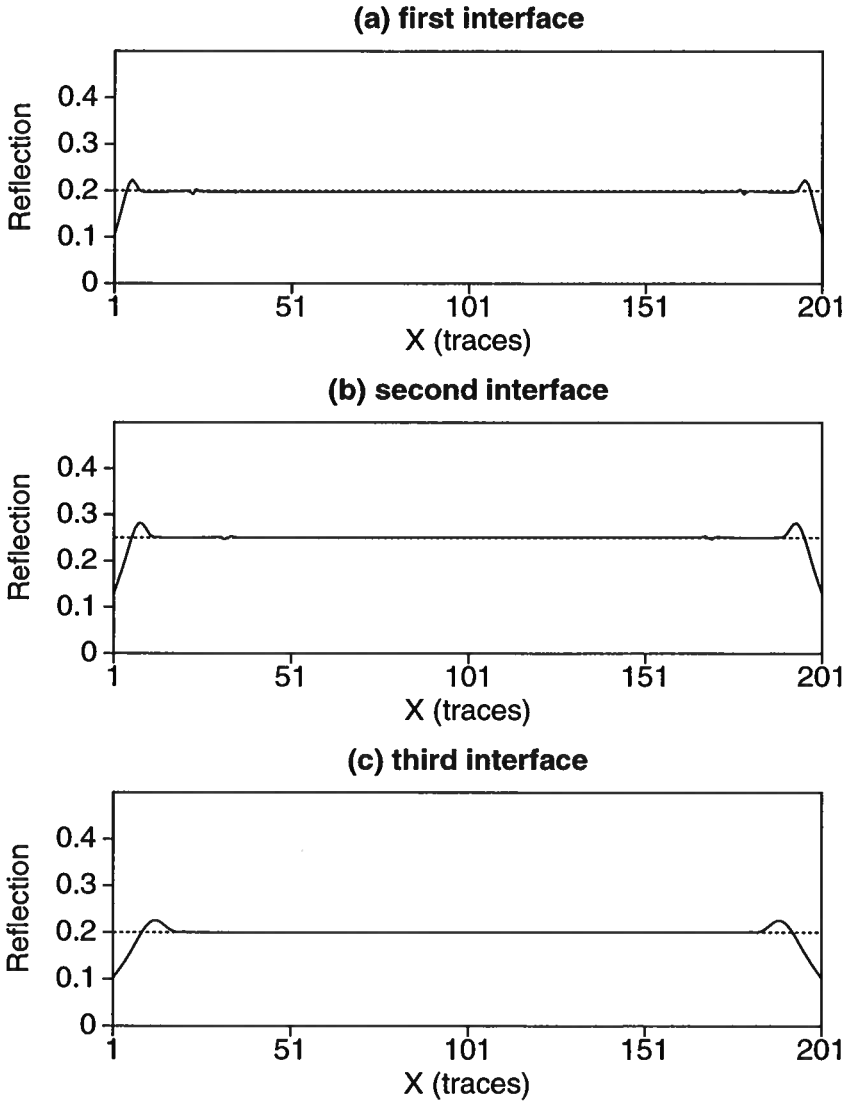

FIG. 5.4. Comparison between theoretical reflection coefficients (dashed) and peak amplitudes (solid) of zero-offset inversion.

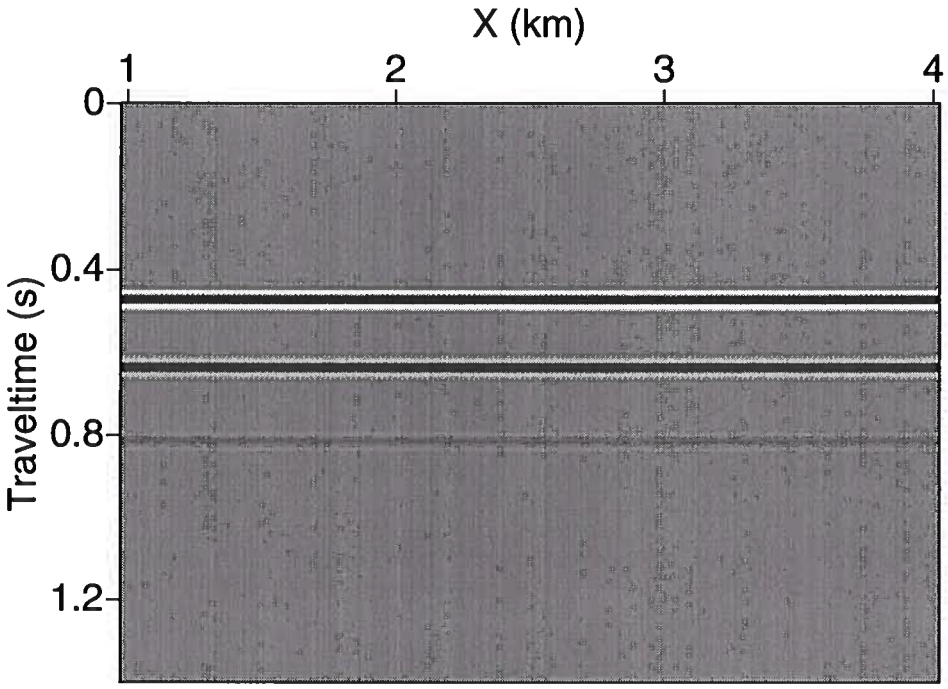

FIG. 5.5. Synthetic data with offset= 0.5 km.

In Figure 5.4 the amplitudes along each reflector are picked and compared to exact reflection coefficients. Note that close to the border, the amplitudes are influenced by the interference with the migration smiles due to the limited aperture. In the middle of the upper reflector, the amplitudes are slightly affected by a migration smile from the lower reflector. Since the smiles from deeper reflectors have larger radius, the degraded amplitudes move in somewhat on the second reflector compared with that on the first reflector. Apart from these artifacts, the match between exact and estimated values is good.

Figure 5.5 shows the synthetic model data with offset= 0.5 km produced by CSHOT for the same velocity distribution. The inversion produces two outputs: for β_1 and β_2. Figure 5.6 shows the result for β_1. The amplitudes in the two outputs should differ by the factor $\cos\theta$, where θ is the specular reflection angle. Tables 5.2 and 5.3 show comparisons of exact values and inversion results for R and θ where $\cos\theta$ is obtained by computing the ratio of R and $R\cos\theta$. Therefore, no scaling factor is applied in the computation of θ. Note in Tables 5.2 and 5.3 that the error in the calculation of $\cos\theta$ is much less than that of R and $R\cos\theta$. Because the processes for computing β_1 and β_2 are similar, their errors tend to have the same trend and their ratio is more accurate than is either of them, individually.

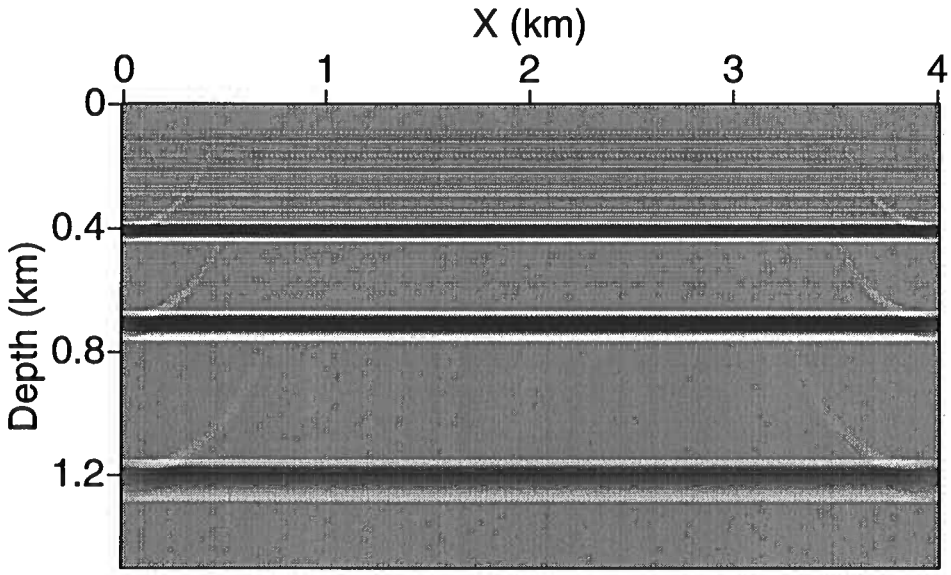

FIG. 5.6. Output β_1 for synthetic data with offset= 0.5 km.

Interface	True d(epth) (km)	Inversion d (km)	True R_t	Inversion R_i	Error (%)
1	0.4	0.4	0.354	0.3497	1.41
2	0.7	0.7	0.353	0.3468	1.70
3	1.2	1.2	0.229	0.2265	1.00

Table 5.2 True and inversion results for reflection coefficients R.

Interface	True θ (deg)	Inversion θ (deg)	Error (%)
1	64.0	63.98	0.03
2	48.67	48.92	0.51
3	33.56	33.62	0.18

Table 5.3 True and inversion results for specular angle θ.

Migration smiles are visible at both ends of the image section in Figures 5.3 and 5.6. They are caused by the limited aperture of the seismic data acquisition. In Figure 5.7 the amplitudes along each reflector are picked and compared to exact reflection coefficients. Similar as in the zero offset case, amplitude distortions due to migration smiles are observed close to the border. Again, apart from these artifacts, the match

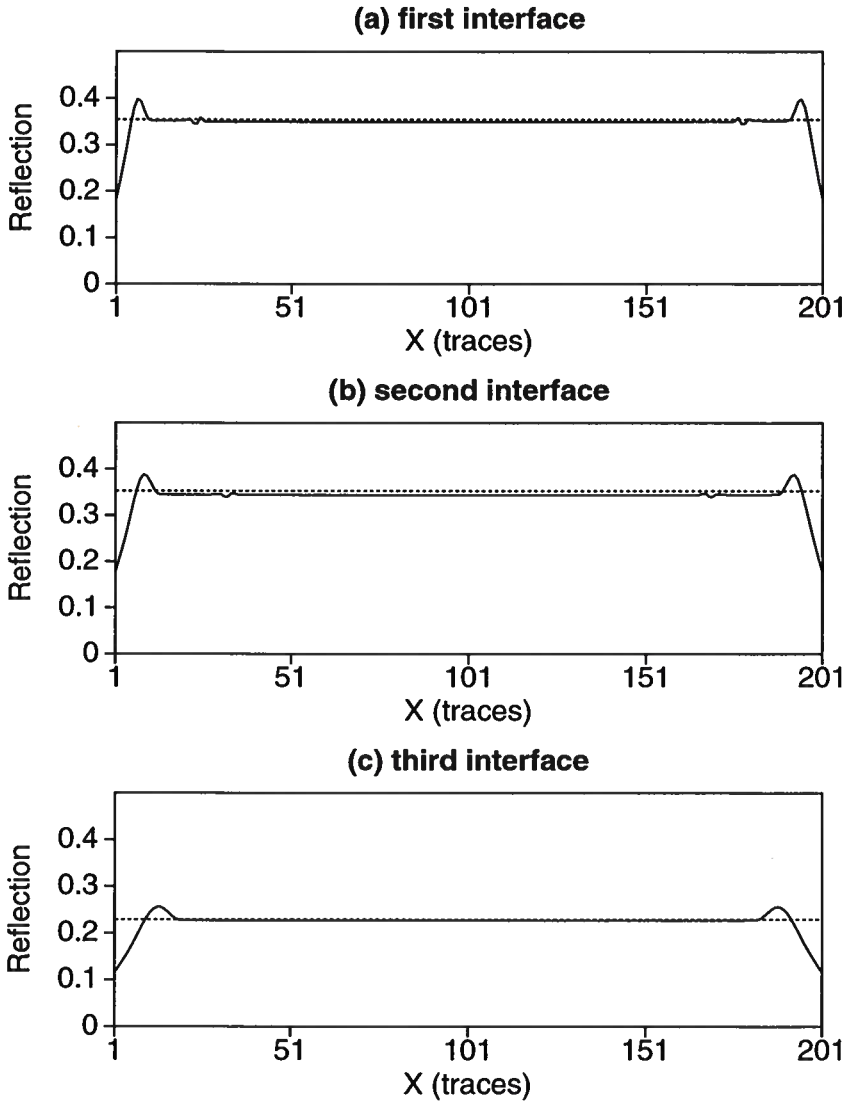

FIG. 5.7. Comparison between theoretical reflection coefficients (dashed) and peak amplitudes (solid) of common-offset inversion, (offset=0.5 km).

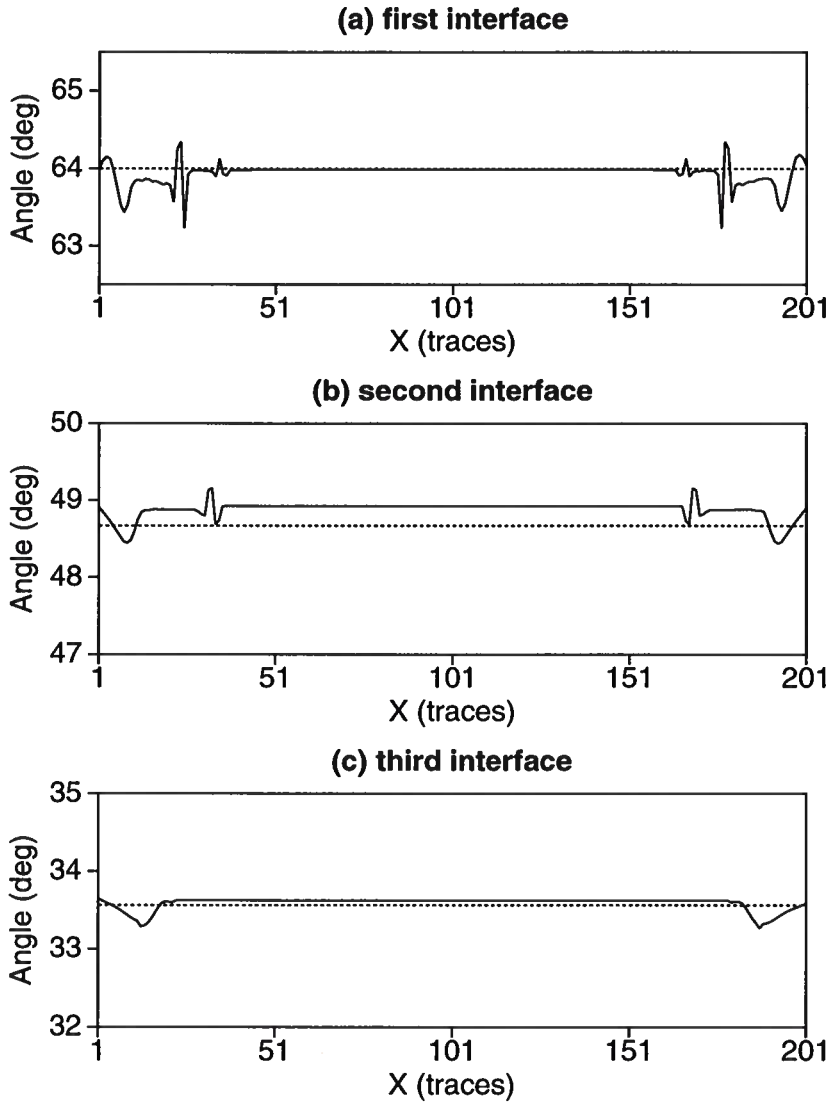

FIG. 5.8. Comparison between theoretical reflection angles (dashed) and the computed results (solid) from common-offset inversion.

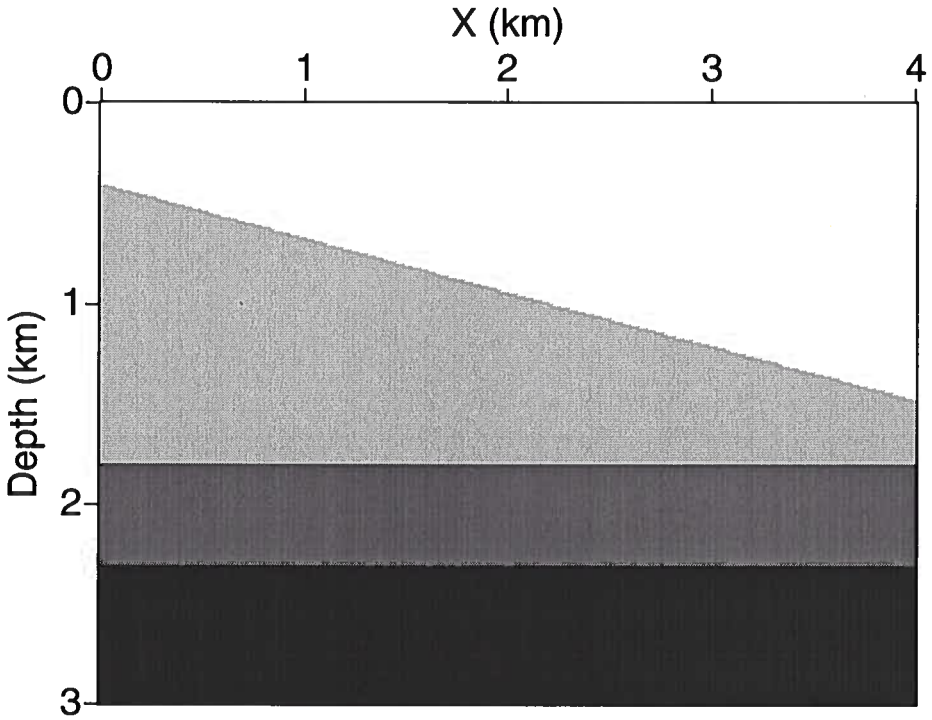

FIG. 5.9. Velocity model consists of four constant-velocity layers with one dipping reflector and two horizontal interfaces. From top to bottom, the layer velocities are 2, 3, 5 and 7.5 km/s.

between exact and estimated values is good. Figure 5.8 shows the comparison between the estimated reflection angles and their exact values.

5.4 Example II — dipping overburden

In this example, we test the positioning and amplitude accuracy for a dipping overburden model. The model consists of a dipping overburden and two horizontal interfaces, as shown in Figure 5.9. Once again I first perform a zero-offset inversion to find the scaling factor for model data. Figure 5.10 shows the zero-offset synthetic model data produced by CSHOT. It consists of 200 $\times$ 200 traces, with 1001 samples per trace at 2-ms sample interval and 20-m trace interval in both x (in-line) and y (off-line) directions. Because of the varying geometrical spreading, the amplitude for each reflection event is not uniform along any given reflection. Figure 5.11 shows the result for β_1, and Figure 5.12 shows the comparison of exact values and inversion results for R. The scaling factors obtained from the three interfaces range close to 29.1. I will use this number as the scaling factor for nonzero-offset experiments.

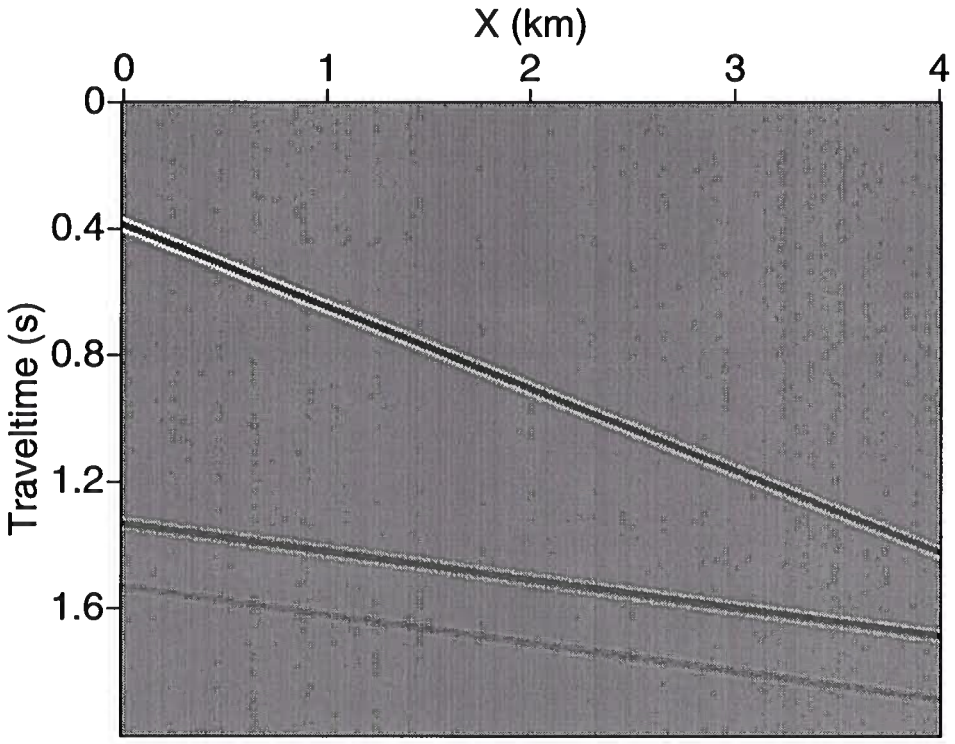

FIG. 5.10. Zero-offset synthetic data of dipping-overburden model

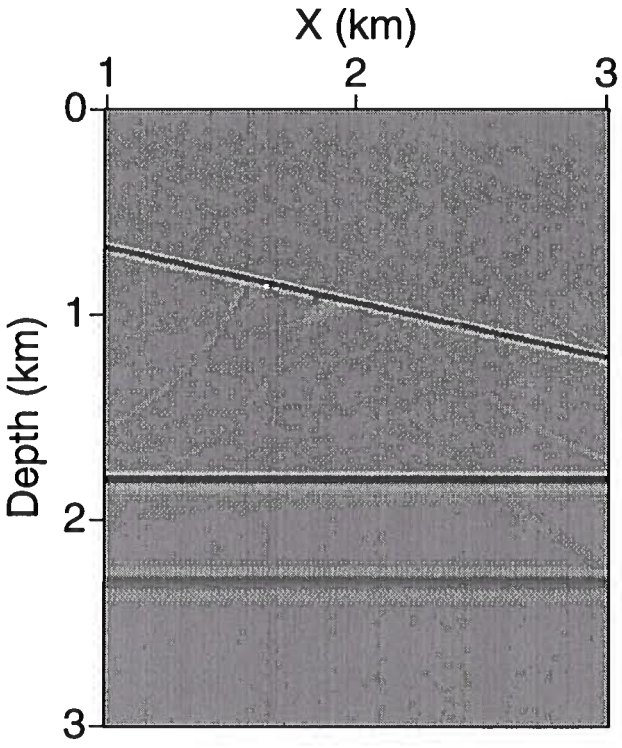

FIG. 5.11. Output β_1 for zero-offset data in Figure 5.10.

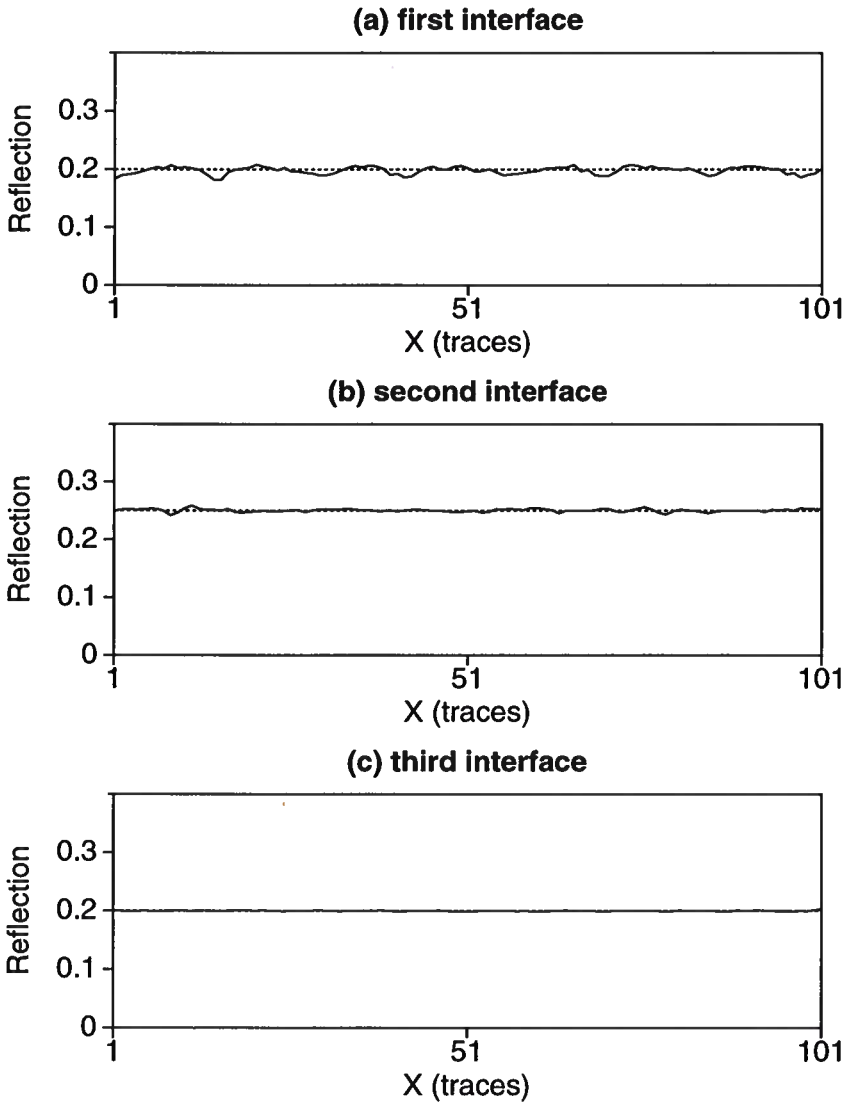

FIG. 5.12. Comparison between theoretical reflection coefficients (dashed) and peak amplitudes (solid) from zero-offset inversion of the data in Figure 5.10.

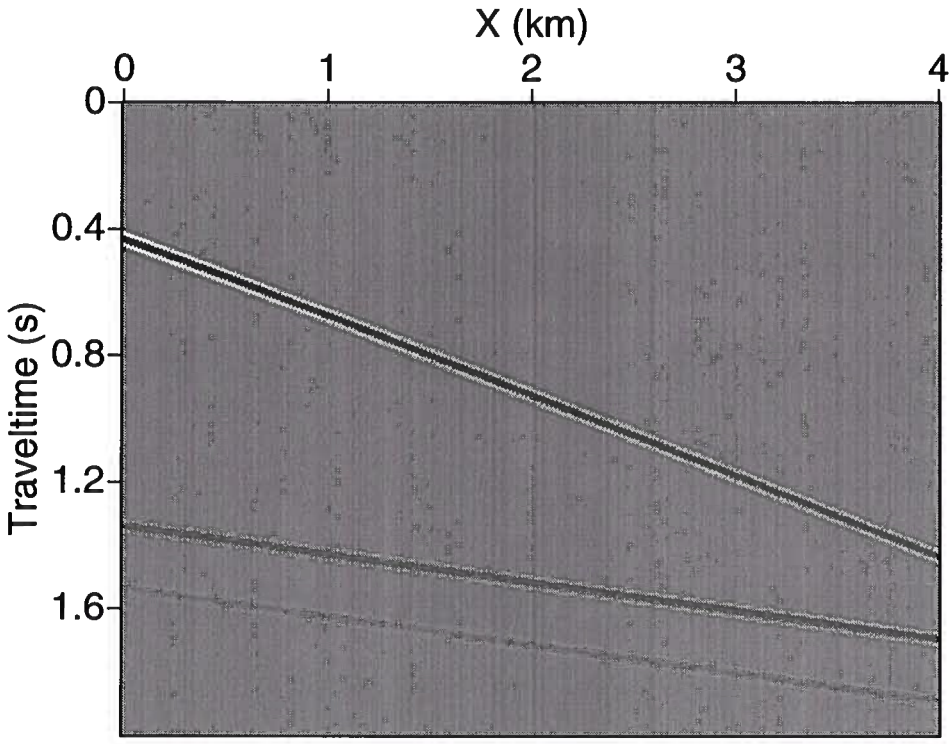

FIG. 5.13. Synthetic data with offset=0.4 km.

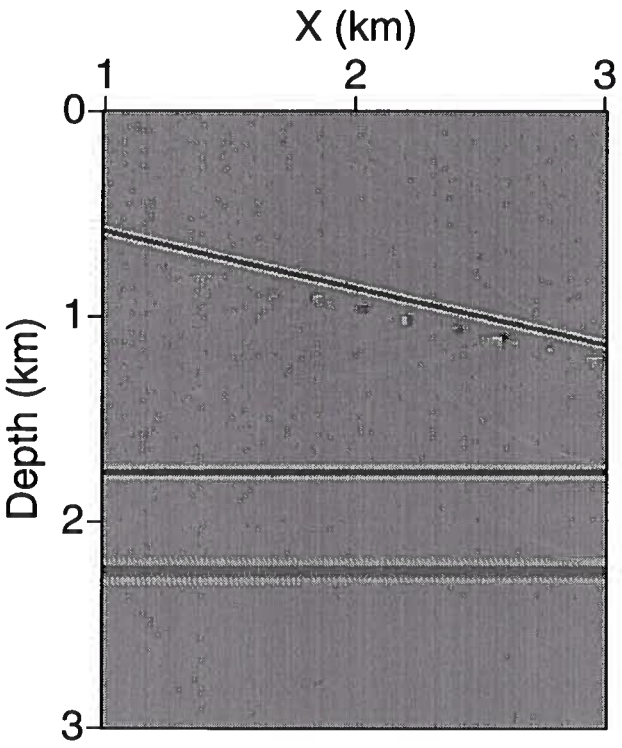

FIG. 5.14. Output β_1 for nonzero-offset data.

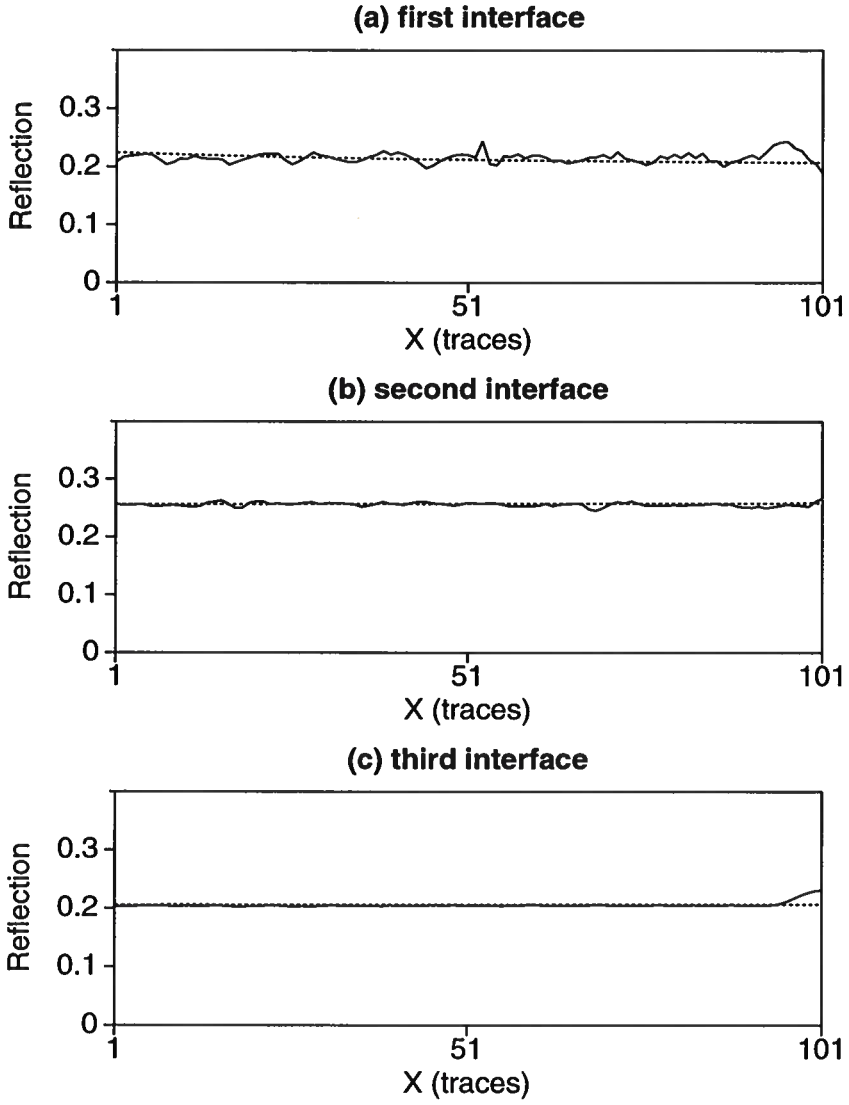

FIG. 5.15. Comparison between theoretical reflection coefficients (dashed) and peak amplitudes (solid) of common-offset inversion.

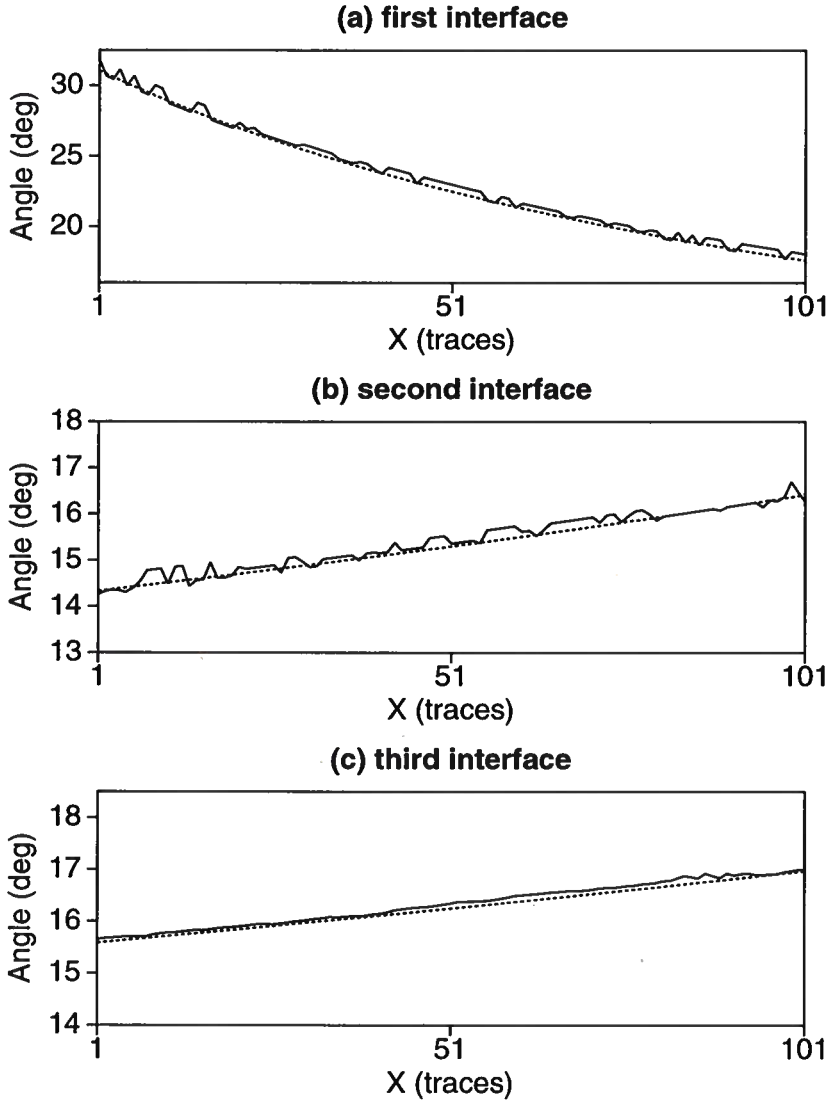

FIG. 5.16. Comparison between theoretical reflection angles (dashed) and the computed results (solid) from common-offset inversion.

Figure 5.13 shows the model data, produced by CSHOT for the same velocity distribution with offset=0.4 km. Figure 5.14 shows the result for β_1. In Figure 5.15, the amplitudes along each reflector are picked and compared to the exact reflection coefficients. For this model, I computed inversion sections for both β_1 and β_2, and computed $\cos\theta$ from the the ratio of R and $R\cos\theta$. Figure 5.16 shows a comparison of the estimated reflection angles along each reflector with the exact angles. Note that the errors in $\cos\theta$ are much smaller than those of R and $R\cos\theta$. Because the processes for computing β_1 and β_2 are similar, their errors tend to have the same trend and their ratio is more accurate than is either of them, individually.

Chapter 6

CONCLUSIONS AND FUTURE WORK

6.1 Summary of major contributions

In this thesis, I studied the ray-theoretic modeling and true amplitude Kirchhoff inversion. Accurate and efficient forward modeling was a key part of 3-D prestack migration/inversion, among other applications. True amplitude Kirchhoff inversion aims at estimating angular-dependent reflection coefficients, and to calculating the corresponding reflection angles from seismic data.

In Chapter 2, I developed a ray-theoretical modeling method that uses smoothing of the physical model, dynamic ray tracing and wavefront construction (WFC). Following are the main conclusions:

1. The use of smoothing applied to the discontinuous model stabilizes the ray tracing while maintaining sufficient accuracy on transmitted traveltime and amplitude coefficients for our inversion technique. The assumption of general smooth gridded model does not allow for transmission coefficients due to the large velocity contrast. A method of re-introducing the transmission coefficient to the smooth amplitude solution is proposed and demonstrated to produce accurate angular transmission coefficients and therefore provide accurate ray amplitude.

2. Analytic ray tracing in tetrahedral-based model has its applications in traveltime calculations. However, the tetrahedral-based approach produces unacceptable amplitudes due to the difficulty in efficiently obtaining accurate amplitudes across internal interfaces.

3. The method efficiently finds all the possible arrivals and is capable of calculating all typical ray-tracing parameters such as traveltime, ray amplitude, initial take-off angles and slowness vectors. It can produce all the quantities involved in Kirchhoff inversion formula.

4. The wavefront construction procedure based on proper interpolation of new rays makes dynamic ray tracing more efficient and results in a sufficiently dense and consistent ray coverage throughout the model, even in areas of large geometrical

spreading. When accurate amplitudes are required, I believe that this is a competitive method for development of ray-theoretic wavefields.

5. Maslov asymptotics offers a valid amplitude solution at caustics where the conventional asymptotic ray theory breaks down. The stationary-phase analysis of the Maslov integral representation shows that the asymptotic Green's function near caustics is larger than the solution away from caustics, largely by being of higher power in ω. This suggests us that we can simply zero out the contribution from the neighborhood of caustic points in the Kirchhoff inversion process.

The implementation of this ray-theoretical forward modeling algorithm is demonstrated in Chapter 3 on three progressively more complex earth models. The adaptability of the method was demonstrated by evaluating the wavefield in regions of extreme geometrical spreading, such as near a salt wall, in sub-salt regions, and in the presence of a low velocity lens.

The approach of Bleistein-Cohen to Kirchhoff inversion is reviewed in Chapter 4. It dictates weights needed in diffraction stack migration in order to account for amplitude loss in seismic data due to geometrical spreading and other problems. The weights are formulated in ray-theoretic notation, which is well suited to an efficient computer implementation that uses dynamic ray tracing.

The applications of true amplitude Kirchhoff inversion to synthetic models show that angle-dependent reflection coefficients are accurately estimated. The corresponding reflection angles are accurately estimated from seismic data by a technique that uses two simultaneous diffraction stacks. It can straightforwardly be implemented in the true amplitude Kirchhoff inversion.

6.2 Future work

The success of the 3-D prestack migration/inversion methods depends strongly on the computation of Green's functions. In this thesis, I have limited the forward modeling to a method that uses smoothing of physical model, dynamic ray tracing and wavefront construction. This choice of ray-tracing method is made on an "objective-oriented" basis. Nevertheless, the method should be compared with other ray-tracing algorithms such as ray tracing in layered structures. The latter contains interfaces and therefore provides a more straightforward treatment of the transmission coefficients.

The true amplitude Kirchhoff inversion, especially in 3-D, is still in an early development phase. More testing should be performed using the technique in Chapter 5.

There are several assumptions that preclude the direct application of the true amplitude Kirchhoff inversion in practice. For example, the seismic data are assumed

to contain only compressional primaries from a lossless, isotropic elastic earth. In practice, multiple reflections, mode conversions, attenuation caused by lossy material and amplitude, and moveout distortion caused by anisotropy, causes amplitudes to differ from those needed for the inversion theory in its current form. In addition, the high-frequency approximation fails to account for the presence of fine detail in the background velocity model. Removing these limitations requires further development. Nevertheless, the analysis of true-amplitude Kirchhoff inversion is of increasingly great interest due to its contribution to the exploration success.

REFERENCES

Albertin, U., Herman, J., Yingst, D., Bloor, R., Chang, W., Beasley, C., & Mobley, E. 1999. Aspects of true amplitude migration. *69th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, Tulsa*, 1358–1361.

Arnold, V. I. 1967. Characteristic classes entering in quantization conditions. *Funct. Anal. Appl.*, **1**, 1–13.

Berkhout, A. J. 1985. *Seismic migration: Imaging of acoustic energy by wavefield extrapolation A. Theoretical aspects.* Elsevier Science Publ., New York.

Berryhill, J. R. 1979. Wave-equation datuming. *Geophysics*, **44**, 1329–1344.

Beylkin, G. 1985. Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon Transform. *J. Math. Phys.*, **26**, 99–108.

Beylkin, G., & Burridge, R. 1990. Linearized inverse scattering problems in acoustics and elasticity. *Wave Motion*, **12**, 12–52.

Bleistein, N. 1984. *Mathematical Methods for Wave Phenomena.* Academic Press Inc., New York.

Bleistein, N. 1986. Two-and-one-half dimensional in-plane wave propagation. *Geophysical Prospecting*, **34**, 686–703.

Bleistein, N. 1987. On the imaging of reflectors in the earth. *Geophysics*, **52**, 931–942.

Bleistein, N., Cohen, J. K., & Hagin, F. G. 1987. Two and one-half dimensional born inversion with an arbitrary reference. *Geophysics*, **52**, 26–36.

Bleistein, N., Cohen, J. K., & Stockwell, Jr., J. W. 2000. *Mathematics of Multidimensional Seismic Migration, Imaging and Inversion.* Springer-Verlag, New York, to appear.

Chapman, C. H., & Drummond, R. 1982. Body-wave seismograms in inhomogeneous media using Maslov asymptotic theory. *Bull. Seism. Soc. Am.*, **72**, S277–S317.

Červený, V. 1987. Ray tracing algorithms in three-dimensional latrally varying layered struc tures. *Pages 99–133 of: Seismic Tomography.* G. Nolet (ed.).

Červený, V. 1995. *Seismic Waves Fields in Three-Dimensional Isotropic and Anisotropic Structures.* University of Trondheim, Trondheim, Norway, to appear as Seismic Ray Theory, Cambridge Univ Pr.

Červený, V., & de Castro, M. A. 1993. Application of dynamic ray tracing in the 3-D inversion of seismic-reflection data. *Geophysics Journal International*, **113**, 776–779.

Chilcoat, S. R., & Hildebrand, S. T. 1995. Wavefront construction in 3-D. *65th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, Tulsa*, 1255–1257.

Claerbout, J. F. 1996. *Imaging the earth's interior.* Blackwell Scientific Publications, Boston.

Clayton, R. W., & Stolt, R. H. 1981. A Born-WKBJ inversion method for acoustic reflection data. *Geophysics*, **46**, 1559–1567.

Cohen, J. K., & Bleistein, N. 1979. Velocity inversion procedure for acoustic waves. *Geophysics*, **44**, 1077–1087.

Cohen, J. K., Hagin, F. G., & Bleistein, N. 1986. Three-dimensional Born inversion with an arbitrary reference. *Geophysics*, **51**, 1552–1558.

de Hoop, M. V., & Brandsberg-Dahl, S. 1998. Maslov asymptotic extension of Generalized Radon Transform inversion in anisotropic elastic media: a Least-Squares approach. *Inverse Problems, in print.*

Docherty, P. 1987a. *Ray theoretical modeling, migration, and inversion in two-and-one-half-dimensional layered acoustic media.* Ph.D. thesis, Colorado School of Mines.

Docherty, P. 1987b. Two-and-one-half dimensional common shot modeling. *Tech. rept., CWP-50, Center for Wave Phenomena, Colorado School of Mines.*

Docherty, P. 1991. A brief comparison of some Kirchhoff integral formulas for migration. *Geophysics*, **56**, 1164–1169.

Dong, W., Emanuel, M. J., & Bleistein, N. 1991. A computer implementation of 2.5-D common-shot inversion. *Geophysics*, **56**, 1164–1169.

Farra, V. 1990. Amplitude computation in heterogeneous media by ray perturbation theory: a finite element approach. *Geophysics Journal International*, **103**, 341–354.

Geoltrain, S., & Brac, J. 1993. Can we image complex structure with first-arrival traveltime. *Geophysics*, **58**, 564–575.

Gray, S. 1997. True-amplitude seismic migration: a comparison of three approaches. *Geophysics*, **62**, 929–936.

Hanitzsch, C. 1995. *Amplitude preserving prestack Kirchhoff depth migration/inversion in laterally inhomogeneous media.* Ph.D. thesis, University Karlsruhe, Germany.

Hanitzsch, C. 1997. Comparison of weights in prestack amplitude-preserving Kirchhoff depth migration. *Geophysics*, **62**, 1812–1816.

Hörmander, L. 1971. Fourier integral operators. *Acta Math.*, **127**, 79–183.

Hubral, P. 1980. Wavefront curvatures in 3-D laterally inhomogeneous media with curved interfaces. *Geophysics*, **45**, 905–513.

Hubral, P., Tygel, M., & Zien, H. 1991. Three-dimensional, true-amplitude zero-offset migration. *Geophysics*, **56**, 18–26.

Hubral, P., Schleicher, J., & Tygel, M. 1996. A unified approach to 3-D seismic reflection imaging, Part I: Basic concepts. *Geophysics*, **61**, 742–758.

Keller, J. B. 1958. *A geometrical theory of diffraction. Calculus of variations and its applications.* McGraw, New York.

Kelly, K. R., Ward, R. W., Treitel, S., & Alford, R. M. 1975. Synthetic seismograms: a finite-difference approach. *Geophysics*, **41**, 2–27.

Kendall, J-M., & Thomson, C. J. 1993. Maslov ray summation, pseudo-caustics, Lagrangian equivalence and transient seismic waveforms. *Geophysics Journal International*, **113**, 186–214.

Klimes, L., & Kvasnicka, M. 1994. 3-D network ray tracing. *Geophysics Journal International*, **116**, 726–738.

Körnig, M. 1995. Cell ray tracing for smooth, isotropic media: a new concept based on a generalized analytic solution. *Geophysics Journal International*, **123**, 391–408.

Kravtsov, Yu. A., & Orlov, Yu. I. 1980. *Geometrical optics of inhomogenous media (in Russian).* Nauka, Moscow, translation to English by Springer, Berlin 1990.

Liu, Z. 1994. A velocity smoothing technique based on damped least squares. *Tech. rept., CWP-149, Center for Wave Phenomena, Colorado School of Mines.*

Lucio, P. S., Lambare, G., & Hanyga, A. 1996. 3D multivalued travel time and amplitude maps. *Pure and Applied Geophysics*, **148**, 449–479.

Maslov, V. P. 1964. *Theory of perturbations and asymptotic methods.* Izd. MGU, Moscow, USSR.

Meng, Z., & Bleistein, N. 1997. Wavefront contruction (WF) ray tracing in tetrahedral models – application to 3-D traveltime and ray path computations. *67th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, Tulsa*, 1734–1737.

Miller, D., Oristaglio, M., & Beylkin, G. 1987. A new slant on seismic imaging: Migration and integral geometry. *Geophysics*, **52**, 943–964.

Mitchell, A. R., & Griffiths, D. F. 1980. *The finite difference method in partial differential equations.* John Wiley, New York.

Newman, P. 1975. Amplitude and phase properties of a digital migration process. *37th Annual Internat. Mtg., Europ. Assoc. Expl. Geophys.*

Ou, B. 1995. 2.5-*D common offset inversion in triangulated background models of the Earth.* Master thesis, Colorado School of Mines.

Podvin, P., & Lecomte, I. 1991. Finite-difference computation of traveltimes in very contrasted velocity models: a massively parallel approach and its associated tools. *Geophysics Journal International*, **105**, 271–284.

Saito, H. 1989. Traveltimes and raypaths of first-arrival seismic waves: computation method based on Huygen's principle. *59th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, Tulsa*, 1024–1027.

Schleicher, J., Tygel, M., & Hubral, P. 1993. 3-D true-amplitude finite-offset migration. *Geophysics*, **58**, 1112–1126.

Schneider, W. A. 1978. Integral formulation for migration in two and three dimensions. *Geophysics*, **43**, 49–76.

Sethian, J. A., & Popovici, A. M. 1999. 3-D traveltime computation using the fast marching method. *Geophysics*, **64**, 516–523.

Sheriff, R. E. 1975. Factors afftecting seismic amplitudes. *Geophysical Prospecting*, **23**, 125–138.

Sun, J., & Gajewski, D. 1997. True-amplitude common-shot migration revisited. *Geophysics*, **62**, 1250–1259.

Sun, J., & Gajewski, D. 1998. On the computation of the true-amplitude weighting functions. *Geophysics*, **63**, 1648–1651.

Sun, Y. Clapp, R., & Biondi, B. 1995. Three dimensional dynamic ray tracing in complex geological structures. *Tech. rept. SEP-93, Standford University.*

Tarantola, A. 1984. Linearized inversion of seismic reflection data. *Geophysical Prospecting*, **32**, 998–1015.

Tygel, M., Hubral, P., & Hanitzsch, C. 1993. Multiple weights in diffraction stack migration. *Geophysics*, **58**, 1820–1830.

Tygel, M., Schleicher, J., & Hubral, P. 1996. A unified approach to 3-D seismic reflection imaging, Part II: Theory. *Geophysics*, **61**, 759–775.

Vidale, J. E. 1988. Finite-difference calculation of travel times. *Bull. Seismol. Soc. Am.*, **78**, 2062–2076.

Vinje, V., Iversen, E., & Gjöystdal, H. 1993. Traveltime and amplitude estimation using wavefront construction. *Geophysics*, **58**, 1157–1166.

Vinje, V., Iversen, E., Astebol, K., & Gjöystdal, H. 1996. Estimation of multivalued arrivals in $3D$ models using wavefront construction-Part I & II. *Geophysical Prospecting*, **44**, 819–858.

Virieux, J., & Farra, V. 1991. Ray tracing in 3-D complex isotropic media: An analysis of the problem. *Geophysics*, **56**, 2057–2069.

Wang, L., & Bleistein, N. 1998. $3D$ Multivalued Traveltime and Amplitude Maps. *68th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, Tulsa*, 1879–1882.

Xu, M. 1996. *3-D common offset inversion in depth dependent media and its parallel implementation.* Master thesis, Colorado School of Mines.

Yilmaz, Ö. 1987. *Seismic Data processing.* Society of Exploration Geophysicists, Tulsa.

Ziolkowski, R. W., & Deschamps, G. A. 1980. The Maslov method and the asymptotic Fourier transform: Caustic analysis. *Electromagnetic Laboratory Scientific Rep. No. 80-9, University of Illinois, Urbana-Champaign.*

Appendix A

BEYLKIN DETERMINANT EXPRESSION BASED ON DYNAMIC RAY TRACING

This appendix contains a review and discussion of the computation of the Beylkin determinant in Červený and de Castro (1993). Derivation details of some key results in that paper are included here in order to make this thesis self-contained on this subject.

In this appendix, I use several 2×2 and 3×3 matrices. To distinguish between them, I use the circumflex $\hat{}$ above the letter for 3×3 matrices, and boldface letters for 2×2 matrices and vectors. If the same letter is used for both 2×2 and 3×3 matrices, e.g., $\boldsymbol{M}$ and $\hat{\boldsymbol{M}}$, matrix $\boldsymbol{M}$ denotes the 2×2 left upper submatrix of $\hat{\boldsymbol{M}}$,

$$\hat{\boldsymbol{M}} = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix}, \qquad \boldsymbol{M} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}.$$

As discussed in Chapter 2, several coordinate systems are involved in ray tracing. We use the symbol (x) as a superscript to specify the vectors and matrices in **Cartesian coordinates**, and the letter itself denotes quantities in **ray-centered coordinates**. For example, $\boldsymbol{p}^{(x)}(\boldsymbol{x})$ denotes the slowness vector at point $\boldsymbol{x}$ in Cartesian coordinates, while $\boldsymbol{p}(\boldsymbol{x})$ denotes the slowness vector in ray-centered coordinates.

A.1 Two-point Eikonal and its mixed derivatives

We shall consider a central ray Ω and two points, S and R, on Ω. Let S' denote a point situated close to S, and R' close to R. We are interested in finding an expansion for the **two-point eikonal** traveltime from S' to R', $T(R', S')$, and its mixed partial derivatives. The latter plays an important role in the expression of the Beylkin determinant.

We start with the expression for the traveltime field in the vicinity of the central ray in the form of Taylor series, up to the quadratic terms, (Červený, 1995)

$$T(R') = T(R) + \boldsymbol{x}^T(R', R)\boldsymbol{p}^{(x)}(R) + \frac{1}{2}\boldsymbol{x}^T(R', R)\hat{\boldsymbol{M}}^{(x)}(R)\boldsymbol{x}(R', R). \qquad \text{(A.1.1)}$$

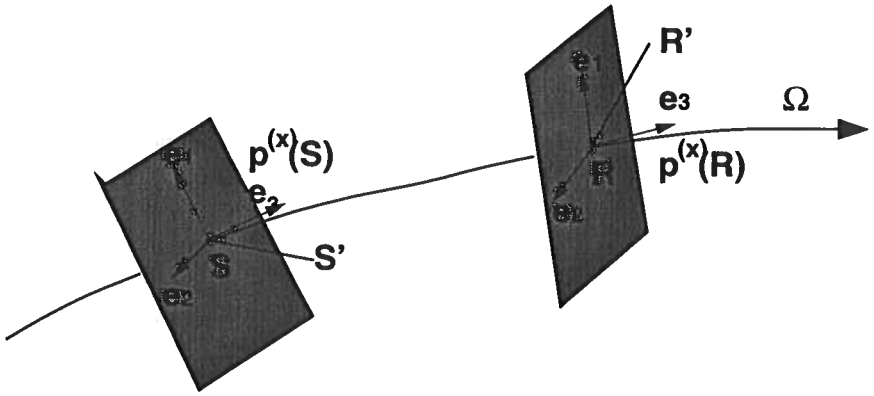

FIG. A.1. S and R are points situated on the central ray Ω. Point S' locate close to S and R' close to R.

Here, $\boldsymbol{p}^{(x)}(R)$ is the slowness vector at R in Cartesian coordinates, and $\hat{\boldsymbol{M}}^{(x)}$ is the matrix of second derivatives of the traveltime field, i.e.,

$$\begin{aligned}\hat{\boldsymbol{M}}^{(x)}(R) &= \left(\frac{\partial^2 T(\boldsymbol{x})}{\partial x_i \partial x_j}\right)_R \qquad i,j=1,2,3,\\ &= \hat{\boldsymbol{H}}(R)\hat{\boldsymbol{M}}(R)\hat{\boldsymbol{H}}^T(R), \end{aligned} \tag{A.1.2}$$

with $\hat{\boldsymbol{M}}(R)$ being the second derivatives of traveltimes with respect to the ray-centered parameters. These are quantities computed by ray-tracing (see Chapter 2). $\hat{\boldsymbol{H}}(R)$ is the transformation matrix from the ray-centered coordinates to the general Cartesian coordinates, $\hat{\boldsymbol{H}}(R) = (\boldsymbol{e}_1^T, \boldsymbol{e}_2^T, \boldsymbol{e}_3^T)$, with $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ being the orthogonal basis in the ray-centered system (see Figure 2.1).

Taking the derivatives of the paraxial traveltimes (A.1.1), we obtain the slowness vector at R',

$$\boldsymbol{p}^{(x)}(R') = \boldsymbol{p}^{(x)}(R) + \hat{\boldsymbol{M}}^{(x)}(R)\boldsymbol{x}(R',R). \tag{A.1.3}$$

Using equations (A.1.3) and (A.1.1), we can express the traveltime from S to R' in terms of the slowness vectors at R and R',

$$T(R') = T(R) + \boldsymbol{x}^T(R',R)\boldsymbol{p}^{(x)}(R) + \frac{1}{2}\boldsymbol{x}^T(R',R)[\boldsymbol{p}^{(x)}(R') - \boldsymbol{p}^{(x)}(R)]. \tag{A.1.4}$$

Similarly, the traveltime from S to S' is

$$T(S') = T(S) + \boldsymbol{x}^T(S',S)\boldsymbol{p}^{(x)}(S) + \frac{1}{2}\boldsymbol{x}^T(S',S)[\boldsymbol{p}^{(x)}(S') - \boldsymbol{p}^{(x)}(S)]. \tag{A.1.5}$$

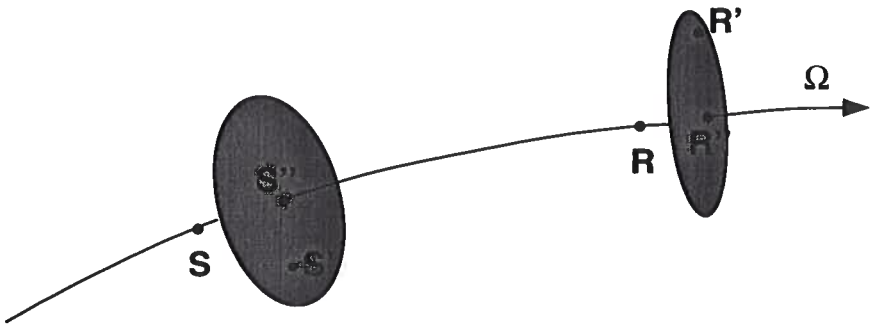

FIG. A.2. Points S' and R' are not necessarily in a plane perpendicular to Ω. Points S'' and R'' are the intersects of the ray Ω with the perpendicular planes to Ω that pass through S' and R', respectively.

Now we have found the quadratic approximation of the two-point eikonal traveltime from S' to R' in terms of $\boldsymbol{p}^{(x)}(R')$ and $\boldsymbol{p}^{(x)}(S')$,

$$\begin{aligned} T(R',S') &= T(R') - T(S') \\ &= T(R) - T(S) + \frac{1}{2}\boldsymbol{x}^T(R',R)[\boldsymbol{p}^{(x)}(R') + \boldsymbol{p}^{(x)}(R)] \\ &\quad -\frac{1}{2}\boldsymbol{x}^T(S',S)[\boldsymbol{p}^{(x)}(S') + \boldsymbol{p}^{(x)}(S)]. \end{aligned} \tag{A.1.6}$$

The next step is to express $\boldsymbol{p}^{(x)}(R')$ and $\boldsymbol{p}^{(x)}(S')$ analytically in terms of the displacements, $\boldsymbol{x}(R',R)$ and $\boldsymbol{x}(S',S)$. Recall that the dynamic ray tracing system can be written as an integral matrix system—the ray propagator matrix, along the central ray. The propagator matrix is non-singular along the whole ray Ω in smooth media; therefore, we can connect the solution of the dynamic ray-tracing system along different segments of the ray through the propagator matrix. Now, denote the projection of S' into the plane perpendicular to Ω at S by S'', and the projection of R' into the plane perpendicular to Ω by R''. Then the ray-centered components of the displacement and slowness vectors, $\boldsymbol{q}$ and $\boldsymbol{p}$, at S'' and R'' can be found in terms of the ray propagator matrix $\boldsymbol{\Pi}(R,S)$,

$$\begin{aligned} \begin{pmatrix} \boldsymbol{q}(R'') \\ \boldsymbol{p}(R'') \end{pmatrix} &= \boldsymbol{\Pi}(R,S) \begin{pmatrix} \boldsymbol{q}(S'') \\ \boldsymbol{p}(S'') \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{Q}_1(R,S) & \boldsymbol{Q}_2(R,S) \\ \boldsymbol{P}_1(R,S) & \boldsymbol{P}_2(R,S) \end{pmatrix} \begin{pmatrix} \boldsymbol{q}(S'') \\ \boldsymbol{p}(S'') \end{pmatrix}. \end{aligned} \tag{A.1.7}$$

In the 4×4 ray propagator matrix $\boldsymbol{\Pi}(R,S)$, $\boldsymbol{Q}_1$ and $\boldsymbol{P}_1$, are the solutions of the dynamic ray tracing system (2.2.7) for the initial condition $\boldsymbol{Q}(0) = \boldsymbol{I}, \boldsymbol{P}(0) = \boldsymbol{0}$ with

0 being the shorthand of initial point, which is often called the **telescopic point** or **plane wavefront** initial condition. $\boldsymbol{Q}_2$ and $\boldsymbol{P}_2$, are the solutions of the dynamic ray tracing system for the initial condition $\boldsymbol{Q}(0) = \boldsymbol{0}, \boldsymbol{P}(0) = \boldsymbol{I}$, which is referred to as the **point source** initial conditions. This initial condition is used in the computation in Chapter 3. The ray propagator plays a fundamental role in dynamic ray tracing. As soon as the ray propagator matrix is known, we can find the solution of the dynamic ray-tracing system analytically for any initial conditions, without repeating the dynamic ray tracing. For a more detailed explanation of the ray propagator matrix $\boldsymbol{\Pi}(R, S)$, see Červený (1995). In (A.1.7),

$$\begin{aligned} \boldsymbol{q}(R'') &= \begin{bmatrix} q_1(R'') \\ q_2(R'') \end{bmatrix}, \qquad \boldsymbol{q}(S'') = \begin{bmatrix} q_1(S'') \\ q_2(S'') \end{bmatrix}, \\ \boldsymbol{p}(R'') &= \begin{bmatrix} p_1(R'') \\ p_2(R'') \end{bmatrix}, \qquad \boldsymbol{p}(S'') = \begin{bmatrix} p_1(S'') \\ p_2(S'') \end{bmatrix}. \end{aligned} \tag{A.1.8}$$

correspond to the first two ray-centered components of the displacement and slowness vectors at S'' and R'', respectively. They are the only 2×1 column vectors appearing in this appendix.

The solution of (A.1.7) for the slowness vectors at S'' and R'' in terms of $\boldsymbol{q}$ is

$$\begin{aligned} \boldsymbol{p}(S'') &= \boldsymbol{Q}_2^{-1}\boldsymbol{q}(R'') - \boldsymbol{Q}_2^{-1}\boldsymbol{Q}_1\boldsymbol{q}(S'') \\ \boldsymbol{p}(R'') &= \boldsymbol{P}_2\boldsymbol{Q}_2^{-1}\boldsymbol{q}(R'') + (\boldsymbol{P}_1 - \boldsymbol{P}_2\boldsymbol{Q}_2^{-1}\boldsymbol{Q}_1)\boldsymbol{q}(S''). \end{aligned} \tag{A.1.9}$$

Notice that the ray propagator $\boldsymbol{\Pi}(R, S)$ satisfies the symplectic property, i.e.,

$$\boldsymbol{\Pi}^T \begin{pmatrix} \boldsymbol{0} & \boldsymbol{I} \\ -\boldsymbol{I} & \boldsymbol{0} \end{pmatrix} \boldsymbol{\Pi} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{I} \\ -\boldsymbol{I} & \boldsymbol{0} \end{pmatrix}. \tag{A.1.10}$$

Substituting equation (A.1.7) into the above, it follows that,

$$\boldsymbol{Q}_2^T\boldsymbol{P}_1 - \boldsymbol{P}_2^T\boldsymbol{Q}_1 = -\boldsymbol{I}, \text{ and } \boldsymbol{Q}_2^T\boldsymbol{P}_2 = \boldsymbol{P}_2^T\boldsymbol{Q}_2.$$

This leads to the following expression,

$$\boldsymbol{P}_1 - \boldsymbol{P}_2\boldsymbol{Q}_2^{-1}\boldsymbol{Q}_1 = -\boldsymbol{Q}_2^{-1T}. \tag{A.1.11}$$

We shall denote by

$$\boldsymbol{M}(R, S) = \boldsymbol{P}_2(R, S)\boldsymbol{Q}_2^{-1}(R, S), \tag{A.1.12}$$

the matrix of the second derivatives of traveltime field at R, due to the point source

situated at S. Similarly, we define $\boldsymbol{M}(S,R) = \boldsymbol{P}_2(S,R)\boldsymbol{Q}_2^{-1}(S,R)$ which represents the matrix of the second derivatives of the traveltime field at S, due to the point source situated at R. Based on the chain rule of the ray propagator matrix $\boldsymbol{\Pi}$, it follows that

$$\boldsymbol{\Pi}(S,R) = \boldsymbol{\Pi}^{-1}(R,S),$$

which implies that

$$\boldsymbol{P}(S,R) = \boldsymbol{Q}_1^T(R,S), \text{ and } \boldsymbol{Q}(S,R) = -\boldsymbol{Q}_2^T(R,S).$$

The above two relations imply

$$\boldsymbol{M}(S,R) = -\boldsymbol{Q}_2^{-1}(R,S)\boldsymbol{Q}_1(R,S). \tag{A.1.13}$$

Substituting the results of equations [(A.1.11), (A.1.12), and (A.1.13)] into equation (A.1.9), we obtain

$$\begin{aligned} \boldsymbol{p}(S'') &= \boldsymbol{M}(S,R)\boldsymbol{q}(S'') + \boldsymbol{Q}_2^{-1}\boldsymbol{q}(R''), \\ \boldsymbol{p}(R'') &= -\boldsymbol{Q}_2^{-1T}(R,S)\boldsymbol{q}(S'') + \boldsymbol{M}(R,S)\boldsymbol{q}(R''). \end{aligned} \tag{A.1.14}$$

Now, introduce the decomposition of $\hat{\boldsymbol{M}}$,

$$\hat{\boldsymbol{M}} = \hat{\boldsymbol{I}}^M \boldsymbol{M} \hat{\boldsymbol{I}}^{MT} + \hat{\boldsymbol{M}}^+, \tag{A.1.15}$$

with

$$\hat{\boldsymbol{I}}^M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \hat{\boldsymbol{M}}^+ = \begin{pmatrix} 0 & 0 & M_{13} \\ 0 & 0 & M_{23} \\ M_{13} & M_{23} & M_{33} \end{pmatrix}.$$

Using the relation $\boldsymbol{q}(R'') = \hat{\boldsymbol{H}}^T(R)\boldsymbol{x}(R',R)$, and equations (A.1.2), (A.1.15), and (A.1.14), we can express the slowness vector $\boldsymbol{p}^{(x)}(R')$ in terms of the displacements $\boldsymbol{x}(R',R)$ and $\boldsymbol{x}(S',S)$,

$$\begin{aligned} \boldsymbol{p}^{(x)}(R') &= \boldsymbol{p}^{(x)}(R) + \hat{\boldsymbol{H}}(R)\hat{\boldsymbol{M}}(R)\hat{\boldsymbol{H}}^T(R)\boldsymbol{x}(R',R) \\ &= \boldsymbol{p}^{(x)}(R) + \hat{\boldsymbol{H}}(R)[\hat{\boldsymbol{I}}^M \boldsymbol{M}(R)\hat{\boldsymbol{I}}^{MT}\hat{\boldsymbol{H}}^T(R)\boldsymbol{x}(R',R) \\ &\quad + \hat{\boldsymbol{M}}^+(R)\hat{\boldsymbol{H}}^T(R)\boldsymbol{x}(R',R)] \end{aligned}$$

$$
\begin{aligned}
&= \boldsymbol{p}^{(x)}(R) + \hat{\boldsymbol{H}}(R)[\hat{\boldsymbol{I}}^{M}\boldsymbol{M}(R,S)\hat{\boldsymbol{I}}^{MT}\hat{\boldsymbol{H}}^{T}(R)\boldsymbol{x}(R',R) \\
&\quad -\hat{\boldsymbol{I}}^{M}\boldsymbol{Q}_2^{-1T}(R,S)\hat{\boldsymbol{I}}^{MT}\hat{\boldsymbol{H}}^{T}(S)\boldsymbol{x}(S',S) + \hat{\boldsymbol{M}}^{+}(R)\hat{\boldsymbol{H}}^{T}(R)\boldsymbol{x}(R',R)] \\
&= \boldsymbol{p}^{(x)}(R) + \hat{\boldsymbol{M}}^{(x)}(R,S)\boldsymbol{x}(R',R) - \hat{\boldsymbol{A}}^{(x)T}(R,S)\boldsymbol{x}(S',S). \qquad \text{(A.1.16)}
\end{aligned}
$$

Here,

$$
\hat{\boldsymbol{A}}^{(x)}(R,S) = \hat{\boldsymbol{H}}(S)\hat{\boldsymbol{I}}^{M}\boldsymbol{Q}_2^{-1}(R,S)\hat{\boldsymbol{I}}^{MT}\hat{\boldsymbol{H}}^{T}(R). \qquad \text{(A.1.17)}
$$

Similarly, we can obtain the following expression for $\boldsymbol{p}^{(x)}(S')$,

$$
\boldsymbol{p}^{(x)}(S') = \boldsymbol{p}^{(x)}(S) + \hat{\boldsymbol{M}}^{(x)}(S,R)\boldsymbol{x}(S',S) - \hat{\boldsymbol{A}}^{(x)T}(S,R)\boldsymbol{x}(R',R), \qquad \text{(A.1.18)}
$$

with $\hat{\boldsymbol{A}}^{(x)}(S,R) = -\hat{\boldsymbol{A}}^{(x)T}(R,S)$.

Taking into account relations (A.1.16) and (A.1.18), it is not difficult to obtain the two-point eikonal traveltime from S' to R',

$$
\begin{aligned}
T(R',S') &= T(R,S) + \frac{1}{2}\boldsymbol{x}^{T}(R',R)[\boldsymbol{p}^{(x)}(R') + \boldsymbol{p}^{(x)}(R)] \\
&\quad -\frac{1}{2}\boldsymbol{x}^{T}(S',S)[\boldsymbol{p}^{(x)}(S') + \boldsymbol{p}^{(x)}(S)] \\
&= T(R,S) \\
&\quad +\frac{1}{2}\boldsymbol{x}^{T}(R',R)\left[2\boldsymbol{p}^{(x)}(R) + \hat{\boldsymbol{M}}^{(x)}(R,S)\boldsymbol{x}(R',R) - \hat{\boldsymbol{A}}^{(x)T}(R,S)\boldsymbol{x}(S',S)\right] \\
&\quad -\frac{1}{2}\boldsymbol{x}^{T}(S',S)\left[2\boldsymbol{p}^{(x)}(S) + \hat{\boldsymbol{M}}^{(x)}(S,R)\boldsymbol{x}(S',S) - \hat{\boldsymbol{A}}^{(x)T}(S,R)\boldsymbol{x}(R',R)\right] \\
&= T(R,S) + \boldsymbol{x}^{T}(R',R)\boldsymbol{p}^{(x)}(R) - \boldsymbol{x}^{T}(S',S)\boldsymbol{p}^{(x)}(S) \\
&\quad +\frac{1}{2}\boldsymbol{x}^{T}(R',R)\hat{\boldsymbol{M}}^{(x)}(R,S)\boldsymbol{x}(R',R) - \frac{1}{2}\boldsymbol{x}^{T}(S',S)\hat{\boldsymbol{M}}^{(x)}(S,R)\boldsymbol{x}(S',S) \\
&\quad -\boldsymbol{x}^{T}(R',R)\hat{\boldsymbol{A}}^{(x)T}(R,S)\boldsymbol{x}(S',S). \qquad \text{(A.1.19)}
\end{aligned}
$$

With equation (A.1.19), we can determine the mixed partial derivative of the travel-

time,

$$\nabla^2_{\boldsymbol{x}(S',S),\boldsymbol{x}(R',R)} T(R',S') = -\hat{\boldsymbol{A}}^{(x)}(R,S). \tag{A.1.20}$$

This is an important equation for the determination of the Beylkin determinant. Notice that the kernel of $\hat{\boldsymbol{A}}^{(x)}(R,S)$, (A.1.17), is the geometrical-spreading matrix $\boldsymbol{Q}_2(R,S)$, which also contributes to the computation of the ray-theoretic amplitude. We shall see that $\boldsymbol{Q}_2(R,S)$ also plays an important role in the calculation of the Beylkin determinant.

A.2 Computation of the Beylkin determinant

Consider an arbitrary data surface S_0, parameterized by parameters $\boldsymbol{\xi} = (\xi_1, \xi_2)$. The source and receiver locations of the seismic experiment are specified by

$$\boldsymbol{x_s} = \boldsymbol{x_s}(\xi_1, \xi_2), \qquad \boldsymbol{x_g} = \boldsymbol{x_g}(\xi_1, \xi_2).$$

The following notations are introduced for the derivatives of source and receiver positions with respect to the surface parameters,

$$\hat{\boldsymbol{\Gamma}}^M_s = \begin{bmatrix} \frac{\partial x_{s1}}{\partial \xi_1} & \frac{\partial x_{s1}}{\partial \xi_2} \\ \frac{\partial x_{s2}}{\partial \xi_1} & \frac{\partial x_{s2}}{\partial \xi_2} \\ \frac{\partial x_{s3}}{\partial \xi_1} & \frac{\partial x_{s3}}{\partial \xi_2} \end{bmatrix}, \qquad \hat{\boldsymbol{\Gamma}}^M_g = \begin{bmatrix} \frac{\partial x_{g1}}{\partial \xi_1} & \frac{\partial x_{g1}}{\partial \xi_2} \\ \frac{\partial x_{g2}}{\partial \xi_1} & \frac{\partial x_{g2}}{\partial \xi_2} \\ \frac{\partial x_{g3}}{\partial \xi_1} & \frac{\partial x_{g3}}{\partial \xi_2} \end{bmatrix}. \tag{A.2.21}$$

Using equations (A.1.17) and (A.1.20), we can determine the quantities of the second and third rows in the Beylkin determinant $h(\boldsymbol{x}, \boldsymbol{\xi})$,

$$\nabla^2_{\boldsymbol{x},\boldsymbol{\xi}} T(\boldsymbol{x}, \boldsymbol{x_s}) = -\hat{\boldsymbol{H}}_s(\boldsymbol{x}) \hat{\boldsymbol{I}}^M \boldsymbol{Q}_2^{-1}(\boldsymbol{x_s}, \boldsymbol{x}) \hat{\boldsymbol{I}}^{MT} \hat{\boldsymbol{H}}^T_s(\boldsymbol{x_s}) \hat{\boldsymbol{\Gamma}}^M_s \equiv -\hat{\boldsymbol{B}}^M_s,$$

$$\nabla^2_{\boldsymbol{x},\boldsymbol{\xi}} T(\boldsymbol{x}, \boldsymbol{x_g}) = -\hat{\boldsymbol{H}}_g(\boldsymbol{x}) \hat{\boldsymbol{I}}^M \boldsymbol{Q}_2^{-1}(\boldsymbol{x_g}, \boldsymbol{x}) \hat{\boldsymbol{I}}^{MT} \hat{\boldsymbol{H}}^T_g(\boldsymbol{x_g}) \hat{\boldsymbol{\Gamma}}^M_g \equiv -\hat{\boldsymbol{B}}^M_g \tag{A.2.22}$$

Here, $\hat{\boldsymbol{H}}_s(\boldsymbol{x}) \equiv (\boldsymbol{e}^s_1(\boldsymbol{x}), \boldsymbol{e}^s_2(\boldsymbol{x}), \boldsymbol{e}^s_3(\boldsymbol{x}))$ and $\hat{\boldsymbol{H}}_g(\boldsymbol{x}) \equiv (\boldsymbol{e}^g_1(\boldsymbol{x}), \boldsymbol{e}^g_2(\boldsymbol{x}), \boldsymbol{e}^g_3(\boldsymbol{x}))$ denote the ray-centered basis vectors at point $\boldsymbol{x}$, from the source $\boldsymbol{x_s}$ and from the receiver $\boldsymbol{x_g}$, respectively. The expression for the Beylkin determinant $h(\boldsymbol{x}, \boldsymbol{\xi})$ then reads

$$h(\boldsymbol{x}, \boldsymbol{\xi}) = \det \begin{bmatrix} \boldsymbol{p_s}^{(x)T} + \boldsymbol{p_g}^{(x)T} \\ \hat{\boldsymbol{B}}^{MT}_s + \hat{\boldsymbol{B}}^{MT}_g \end{bmatrix}. \tag{A.2.23}$$

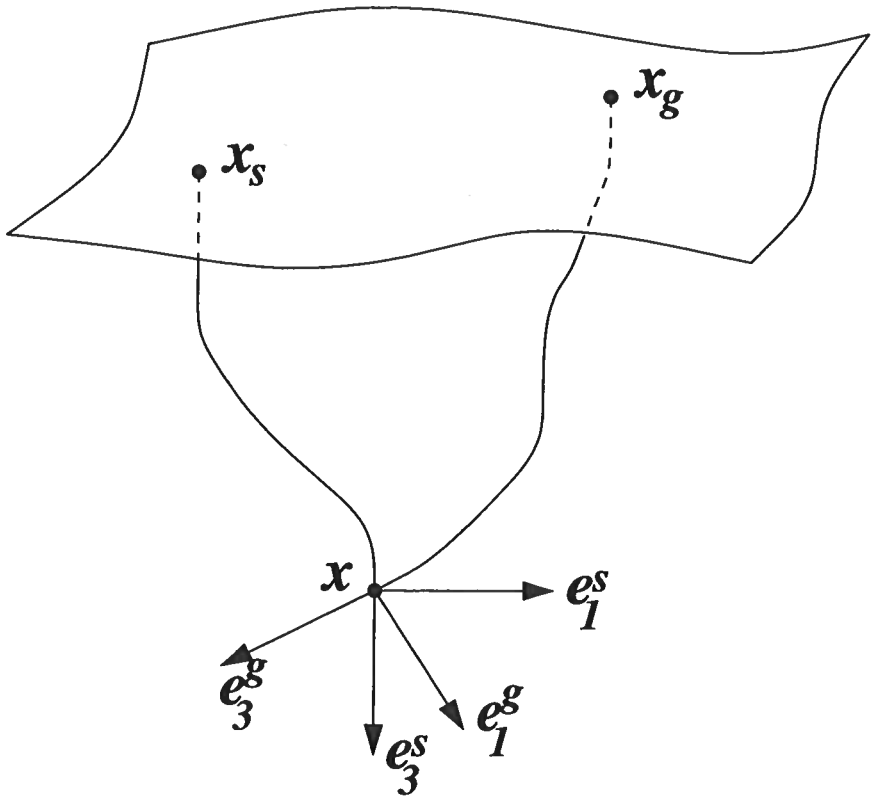

FIG. A.3. Ray-centered basis vectors at $\boldsymbol{x}$ from the source $\boldsymbol{x_s}$— $(\boldsymbol{e}_1^s(\boldsymbol{x}), \boldsymbol{e}_2^s(\boldsymbol{x}), \boldsymbol{e}_3^s(\boldsymbol{x}))$, and from the receiver $\boldsymbol{x_g}$—$(\boldsymbol{e}_1^g(\boldsymbol{x}), \boldsymbol{e}_2^g(\boldsymbol{x}), \boldsymbol{e}_3^g(\boldsymbol{x}))$. If we choose the same unit vectors for $\boldsymbol{e}_2^s(\boldsymbol{x})$ and $\boldsymbol{e}_2^g(\boldsymbol{x})$, perpendicular to the plane specified by $\boldsymbol{e}_3^s(\boldsymbol{x})$ and $\boldsymbol{e}_3^g(\boldsymbol{x})$—pointing out of the front of the page in this figure, then the four basis vectors $\boldsymbol{e}_1^s(\boldsymbol{x})$, $\boldsymbol{e}_3^s(\boldsymbol{x})$, $\boldsymbol{e}_1^g(\boldsymbol{x})$, $\boldsymbol{e}_3^g(\boldsymbol{x})$ are co-planar.

To simplify equation (A.2.23), Červený and de Castro (1993) made the further assumption that the unit vectors, $\boldsymbol{e}_2^s(\boldsymbol{x})$ and $\boldsymbol{e}_2^g(\boldsymbol{x})$, are the same. This common vector is chosen to be perpendicular to the plane specified by $\boldsymbol{e}_3^s(\boldsymbol{x})$ and $\boldsymbol{e}_3^g(\boldsymbol{x})$. Then the unit vectors $\boldsymbol{e}_1^s(\boldsymbol{x})$ and $\boldsymbol{e}_1^g(\boldsymbol{x})$ lie on that plane too. That is, the remaining basis vectors $\boldsymbol{e}_1^s(\boldsymbol{x}), \boldsymbol{e}_1^g(\boldsymbol{x}), \boldsymbol{e}_3^s(\boldsymbol{x}), \boldsymbol{e}_3^g(\boldsymbol{x})$ are co-planar. See Figure A.3. Let θ be the angle between the directions of rays to the source $\boldsymbol{x_s}$ and to the receiver $\boldsymbol{x_g}$ at point $\boldsymbol{x}$, i.e.,

$$\cos\theta = \boldsymbol{e}_3^{sT}(\boldsymbol{x}) \cdot \boldsymbol{e}_3^g(\boldsymbol{x}).$$

Then the transformation from $(\boldsymbol{e}_1^g(\boldsymbol{x}), \boldsymbol{e}_2^g(\boldsymbol{x}), \boldsymbol{e}_3^g(\boldsymbol{x}))$ to $(\boldsymbol{e}_1^s(\boldsymbol{x}), \boldsymbol{e}_2^s(\boldsymbol{x}), \boldsymbol{e}_3^s(\boldsymbol{x}))$ can be explicitly expressed as $\hat{\boldsymbol{R}}\hat{\boldsymbol{H}}_g(\boldsymbol{x}) = \hat{\boldsymbol{H}}_s(\boldsymbol{x})$, with $\hat{\boldsymbol{R}}$ being the rotation transform matrix,

$$\hat{\boldsymbol{R}} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}. \tag{A.2.24}$$

Hence, $\hat{\boldsymbol{B}}_s^{MT} + \hat{\boldsymbol{B}}_g^{MT}$ can be simplified as

$$\begin{aligned} &\hat{\boldsymbol{B}}_s^{MT} + \hat{\boldsymbol{B}}_g^{MT} \\ &= \left[\hat{\Gamma}_s^{MT}\hat{\boldsymbol{H}}_s(\boldsymbol{x_s})\hat{\boldsymbol{I}}^M\boldsymbol{Q}_{2,s}^{-1T} + \hat{\Gamma}_g^{MT}\hat{\boldsymbol{H}}_g(\boldsymbol{x_g})\hat{\boldsymbol{I}}^M\boldsymbol{Q}_{2,g}^{-1T}\right]\hat{\boldsymbol{I}}^{MT}(\hat{\boldsymbol{R}} + \hat{\boldsymbol{I}})\hat{\boldsymbol{H}}_g^T(\boldsymbol{x}). \end{aligned} \tag{A.2.25}$$

Now, let us re-write the first row of (A.2.23) in terms of $\hat{\boldsymbol{R}}$ and $\hat{\boldsymbol{H}}_g^T(\boldsymbol{x})$,

$$\begin{aligned} \boldsymbol{p}_s^{(x)T} + \boldsymbol{p}_s^{(x)T} &= \frac{1}{c(\boldsymbol{x})}[\boldsymbol{e}_3^s(\boldsymbol{x}) + \boldsymbol{e}_3^g(\boldsymbol{x})] \\ &= \frac{1}{c(\boldsymbol{x})}\left\{[0,0,1]\hat{\boldsymbol{H}}_s^T(\boldsymbol{x}) + [0,0,1]\hat{\boldsymbol{H}}_g^T(\boldsymbol{x})\right\} \end{aligned} \tag{A.2.26}$$

$$= \frac{1}{c(\boldsymbol{x})}[0,0,1](\hat{\boldsymbol{R}} + \hat{\boldsymbol{I}})\hat{\boldsymbol{H}}_g^T(\boldsymbol{x}). \tag{A.2.27}$$

Finally, the Beylkin determinant (A.2.23) can be written as

$$\begin{aligned} h(\boldsymbol{x},\boldsymbol{\xi}) &= \frac{1}{c(\boldsymbol{x})}\det\begin{bmatrix} [\hat{\Gamma}_s^{MT}\hat{\boldsymbol{H}}_s(\boldsymbol{x_s})\hat{\boldsymbol{I}}^M\boldsymbol{Q}_{2,s}^{-1T} + \hat{\Gamma}_g^{MT}\hat{\boldsymbol{H}}_g(\boldsymbol{x_g})\hat{\boldsymbol{I}}^M\boldsymbol{Q}_{2,g}^{-1T}]\hat{\boldsymbol{I}}^{MT} \\ 0 \qquad\qquad 0 \qquad\qquad 1 \end{bmatrix} \\ &\quad \cdot\det(\hat{\boldsymbol{R}} + \hat{\boldsymbol{I}}) \cdot \det\hat{\boldsymbol{H}}_g^T(\boldsymbol{x}) \\ &= \frac{4}{c(\boldsymbol{x})}(1 + \cos\theta)\det\boldsymbol{D}, \end{aligned} \tag{A.2.28}$$

where, the 2×2 matrix $\mathbf{D}$ is given by,

$$\boldsymbol{D} = \hat{\boldsymbol{\Gamma}}_s^{MT} \hat{\boldsymbol{H}}_s(\boldsymbol{x}_s) \hat{\boldsymbol{I}}^M \boldsymbol{Q}_{2,s}^{-1T} + \hat{\boldsymbol{\Gamma}}_g^{MT} \hat{\boldsymbol{H}}_g(\boldsymbol{x}_g) \hat{\boldsymbol{I}}^M \boldsymbol{Q}_{2,g}^{-1T}. \tag{A.2.29}$$

Equation (A.2.28) is one of the key results in Červený and de Castro (1993). However, its derivation depends on the assumption of the coincidence of $\boldsymbol{e}_2^s(\boldsymbol{x})$ and $\boldsymbol{e}_2^g(\boldsymbol{x})$. If we think in terms of the two normal planes to the rays from $\boldsymbol{x}_s$ and $\boldsymbol{x}_g$, then, these two vectors must lie along the intersection of those two planes. In general, if the rays have torsion, then the choices of $\boldsymbol{e}_2^s(\boldsymbol{x})$ and $\boldsymbol{e}_2^g(\boldsymbol{x})$ at each output point will arise from *different* initial values of the same two vectors. Therefore, if we start with particular initial values at $\boldsymbol{x}_s$ and $\boldsymbol{x}_g$, then we cannot guarantee the agreement of $\boldsymbol{e}_2^s(\boldsymbol{x})$ and $\boldsymbol{e}_2^g(\boldsymbol{x})$ at depth. Consequently, at each output point, we must introduce a rotation from the propagating vectors, $\boldsymbol{e}_1^s(\boldsymbol{x})$, $\boldsymbol{e}_2^s(\boldsymbol{x})$ and $\boldsymbol{e}_1^g(\boldsymbol{x})$, $\boldsymbol{e}_2^g(\boldsymbol{x})$, to new sets of vectors for which $\boldsymbol{e}_2^s(\boldsymbol{x})$ and $\boldsymbol{e}_2^g(\boldsymbol{x})$ are common. Note that the matrices $\hat{\boldsymbol{H}}_s$, $\hat{\boldsymbol{H}}_g$ at $\boldsymbol{x}$, $\boldsymbol{x}_s$ and $\boldsymbol{x}_g$ appearing in (A.2.25) have to be the matrices *after* rotation. This is an expensive computation. We can only avoid this extra computation in some cases when the rays have no torsion. Even in these cases, $\boldsymbol{e}_2(\boldsymbol{x})$ can vary in different ways on the rays from $\boldsymbol{x}_s$ and $\boldsymbol{x}_g$. The governing equation for the propagation of $\boldsymbol{e}_2^s(\boldsymbol{x})$ is (Červený, 1995, equation (4.1.14))

$$\mathrm{d}\frac{\boldsymbol{e}_2^s}{\mathrm{d}s} = -v^2 \left(\boldsymbol{e}_2^s \cdot \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}s} \right) \boldsymbol{p}. \tag{A.2.30}$$

The derivative, $\mathrm{d}\boldsymbol{p}/\mathrm{d}s$, is closely related to the curvature along the ray, so that this dot product can be nonzero, even when the torsion is zero. However, for a homogeneous medium, this derivative is zero and $\boldsymbol{e}_2^s(\boldsymbol{x})$ is constant on rays. Thus, if $\boldsymbol{e}_2^s(\boldsymbol{x})$ and $\boldsymbol{e}_2^g(\boldsymbol{x})$ are equal initially, they remain equal, all along the ray. Further, for a depth-dependent medium and for in-plane rays, we need only take the initial value of $\boldsymbol{e}_2^s$ to be in the out-of-plane direction. In this case, the dot product in (A.2.30) is zero and, again, $\boldsymbol{e}_2^s(\boldsymbol{x})$ and $\boldsymbol{e}_2^g(\boldsymbol{x})$ remain equal if they are taken to be the equal out-of-plane unit vectors, initially. These are a sparse set of cases with insufficient generality for the arbitrary heterogeneous media to which the method of this thesis is meant to be applied. Thus, we conclude that it is necessary to use equation (A.2.23) instead of equation (A.2.28).

A.3 Source-receiver measurement configurations

Equation (A.2.23) holds for an arbitrary source-receiver configuration and arbitrary data surface S_0. Now, we consider specific source-receiver measurement configurations along a particular data surface. Assume the data surface, S_0, is given by the relation, $x_3 = f(x_1, x_2)$. The positions of the receivers along data surface S_0 are then

determined by parametric equations,

$$\boldsymbol{x_g} = \begin{pmatrix} x_{g1}(\xi_1, \xi_2) \\ x_{g2}(\xi_1, \xi_2) \\ f(x_{g1}(\xi_1, \xi_2), x_{g2}(\xi_1, \xi_2)) \end{pmatrix}. \tag{A.3.31}$$

The position of the sources depends on the positions of the receivers. To simulate general source-receiver configurations, we use the following general expression for source positions,

$$\begin{aligned} \begin{pmatrix} x_{s1} \\ x_{s2} \end{pmatrix} &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_{g1}(\xi_1, \xi_2) \\ x_{g2}(\xi_1, \xi_2) \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \\ x_{s3} &= f(a_{1I} x_{gI} + b_1, a_{2I} x_{gI} + b_2). \end{aligned} \tag{A.3.32}$$

Here a_{IJ} and b_I $(I, J = 1, 2)$ are constants. Specific combinations of a_{IJ} and b_I can describe common-shot(receiver) and common-offset configurations. For a common-shot configuration,

$$a_{11} = a_{12} = a_{21} = a_{22} = 0, \qquad b_1 = x_{s1}, \qquad b_2 = x_{s2}.$$

For a common-offset configuration,

$$a_{11} = a_{22} = 1, \qquad a_{12} = a_{21} = 0, \qquad b_1 = \Delta x_1, \qquad b_2 = \Delta x_2.$$

Using relations (A.3.31-A.3.32), $\hat{\Gamma}_g^M$ and $\hat{\Gamma}_s^M$ can be written as,

$$\hat{\Gamma}_g^M = \begin{bmatrix} \frac{\partial x_{g1}}{\partial \xi_1} & \frac{\partial x_{g1}}{\partial \xi_2} \\ \frac{\partial x_{g2}}{\partial \xi_1} & \frac{\partial x_{g2}}{\partial \xi_2} \\ f_I^g q_{I1}^g & f_I^g q_{I2}^g \end{bmatrix}, \qquad \hat{\Gamma}_s^M = \begin{bmatrix} a_{1K} q_{K1}^g & a_{1K} q_{K2}^g \\ a_{2K} q_{K1}^g & a_{2K} q_{K2}^g \\ f_I^s a_{IK} q_{K1}^g & f_I^s a_{IK} q_{K2}^g \end{bmatrix},$$

$$I, K = 1, 2. \tag{A.3.33}$$

Here we have used the notation,

$$q_{IJ} = \frac{\partial x_{gI}}{\partial \xi_J}, \qquad I, J = 1, 2,$$

and

$$f_I^g = \left(\frac{\partial f(x_1, x_2)}{\partial x_I} \right)_{x_1 = x_{g1}, x_2 = x_{g2}}, \quad f_I^s = \left(\frac{\partial f(x_1, x_2)}{\partial x_I} \right)_{x_1 = x_{s1}, x_2 = x_{s2}}, \quad I = 1, 2.$$

From the above, it follows that $\hat{\Gamma}_g^M$ and $\hat{\Gamma}_s^M$ can be expressed in terms of four quantities, $(\partial x_{gI}/\partial \xi_J)$, $(I, J = 1, 2)$.

For a planar data surface, $x_3 = 0$, and $f_J^s = f_J^g = 0, I = 1, 2$. Let $x_1 = \xi_1$ and $x_2 = \xi_2$. Then, we have

$$\hat{\Gamma}_g^M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad \hat{\Gamma}_s^M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{A.3.34}$$

Notice that, in this case, the common offset Beylkin determinant is just,

$$h(\boldsymbol{x}, \boldsymbol{\xi}) \;=\; \det \begin{bmatrix} \boldsymbol{p_s}^{(x)T} + \boldsymbol{p_g}^{(x)T} \\ \boldsymbol{H}(\boldsymbol{x_s}) \boldsymbol{Q}_{2,s}^{-1T}(\boldsymbol{x_s}, \boldsymbol{x}) \hat{\boldsymbol{H}}_s^{MT}(\boldsymbol{x}) + \boldsymbol{H}(\boldsymbol{x_g}) \boldsymbol{Q}_{2,g}^{-1T}(\boldsymbol{x_g}, \boldsymbol{x}) \hat{\boldsymbol{H}}_g^{MT}(\boldsymbol{x}) \end{bmatrix}. \tag{A.3.35}$$

For a homogeneous medium, we have $\boldsymbol{Q}_{2,s}(\boldsymbol{x_s}, \boldsymbol{x}) = c r_s \boldsymbol{I}$, and $\boldsymbol{Q}_{2,g}(\boldsymbol{x_g}, \boldsymbol{x}) = c r_g \boldsymbol{I}$. The unit vectors $\boldsymbol{e}_1, \boldsymbol{e}_2$ remain constant along the ray in this case. For a plane datum surface in a constant background wavespeed medium, h can be considerably simplified to

$$h = \frac{(1 + \cos\theta) x_3}{c^3} \left[\left(\frac{1}{r_s} + \frac{1}{r_g} \right) \cdot \left(\frac{1}{r_s^2} + \frac{1}{r_g^2} \right) \right], \tag{A.3.36}$$

which agrees with the corresponding result in (Bleistein *et al.*, 2000).

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	18 Dec 00	Jan 91 - Oct 01, Performance Tech.

4. TITLE AND SUBTITLE Estimation of multi-valued Green's function by dynamic ray tracing and true amplitude Kirchhoff inversion in 3-D heterogeneous media

5. FUNDING NUMBERS
N00014-91-J-1267
N00014-95-1-0508

6. AUTHOR(S)
Lan Wang

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Center for Wave Phenomena
Colorado School of Mines
Golden, CO 80401

8. PERFORMING ORGANIZATION REPORT NUMBER
CWP-357

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Applied Analysis Division
800 No. Quincy Street
Arlington, VA 22217-5660

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
unclassified

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Implementation and application of amplitude-preserving pre-stack Kirchhoff depth migration/inversion is important for the detection of hydrocarbons (petroleum and natural gas) in the oil industry. 3-D pre-stack depth migration is the ultimate imaging technique for complex geological structures. The same method has application to the mapping of the seabed for analysis of signal propagation, such as sonar. The analysis of amplitudes on recorded seismic data has contributed a great deal to exploration success and will become increasingly important in the near future. This thesis aims to be a contribution in that direction.

Ray-theoretic modeling require accurate amplitude as well as phase both for forward modeling and Kirchhoff inversion, among other applications. There are no analytical solutions to the ray equations in realistic earth

(continued)

14. SUBJECT TERMS
pre-stack depth migration, ray-theoretic modeling, Kirchhoff inversion, 3-D heterogeneous media

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std 239-18
298-102

(continued)

models; thus, we must use numerical solutions to solve problems of interest. For 3-D codes that requrie reasonable CPU time while achieving sufficient amplitude accuracy to be useful in applications. This thesis describes a ray-theoretical modeling method that produces all the quantities needed in the Kirchhoff inversion.

This approach to the solution of the seismic inverse problem has experienced high acceptance in the oil exploration community.