Finite-difference quasi-P traveltimes for anisotropic media

Jianliang Qian [*]*and William W. Symes* [†]

ABSTRACT

The first arrival quasi-P wave traveltime field in an anisotropic elastic solid solves a first order nonlinear partial differential equation, the qP eikonal equation. The difficulty in solving this eikonal equation by finite difference method is that for anisotropic media the ray (group) velocity direction is not the same as the direction of traveltime gradient so that the traveltime gradient can no longer serve as an indicator of the group velocity direction in extrapolating the traveltime field. However, establishing an explicit relation between the ray velocity vector and the phase velocity vector overcomes this difficulty. Furthermore, the solution of the **paraxial** qP eikonal equation, an evolution equation in depth, gives the first arrival traveltime along downward propagating rays. A second order upwind finite difference scheme solves this paraxial eikonal equation in O(N) floating point operations, where N is the number of grid points. Numerical experiments using 2-D and 3-D transversely isotropic models demonstrate the accuracy of the scheme.

Introduction

Traveltime computation plays a central role in many seismic data processing methods, such as Kirchhoff depth migration and tomographic velocity analysis. Since seismic wave propagation is anisotropic in many sedimentary rocks, maximal imaging resolution requires that traveltime computation honor anisotropy whenever it seriously affects data kinematics (Tsvankin and Thomsen, 1995; Anderson, 1996; Pratt and Chapman, 1992; Cherrett and Singh, 1998).

Contemporay traveltime computation methods fall roughly into two classes: ray tracing methods and finite difference eikonal solvers. Ray tracing and related meth-

[*]The Rice Inversion Project, Department of Computational and Applied Mathematics, Rice University, Houston TX 77251-1892 USA, email `jlqian@caam.rice.edu`

[†]The Rice Inversion Project, Department of Computational and Applied Mathematics, Rice University, Houston TX 77251-1892 USA, email `symes@caam.rice.edu`

ods for isotropic propagation extend without essential difficulty to anisotropic propagation, but they suffer from shadow zones and related interpolation problems (Cerveny, 1972; Shearer and Chapman, 1988; Qin and Schuster, 1993) which complicate robust implementations. Finite difference eikonal solvers compute approximate first arrival times directly on a prespecified grid (Reshef and Kosloff, 1986; Vidale, 1988; van Trier and Symes, 1991; Schneider et al., 1992; Schneider, 1995; Sethian and Popovici, 1999), involve rather simple data structures, and are easy to code efficiently. Extension of these methods to anisotropic wave propagation is not entirely straightforward, however; the methods cited here and others in the public literature describe only finite difference traveltime algorithms for isotropic media.

The finite difference eikonal solvers cited above depend on the fact that for isotropic media the ray velocity vector, i.e. the *group velocity*, has the same direction as the traveltime gradient, i.e. the *phase velocity*, so that we can use the traveltime gradient as a reliable indicator of energy flow (and thus causality) in extrapolating the traveltime field. However, this is no longer true for anisotropic media (Dellinger and Symes, 1997). The goal of the current paper is to establish a reliable indicator of qP ray velocity direction by formulating a relation between the group velocity direction and the phase velocity direction, and to present theoretical formulations, some implementation details, and illustrative applications of a finite difference method for traveltimes of first arriving quasi-compressional waves in heterogeneous anisotropic solids.

The mathematical foundation of the finite difference approach to traveltime computation is the observation of Lions (1982): the first arrival traveltime is a generalized solution of the eikonal equation. A generalized solution concept is necessary because the eikonal equation does not generally have solutions in the ordinary sense throughout its domain. However, generalized solutions are not unique because of the existence of several branches of traveltime in the presence of strong refraction. Lions and others abstractly characterized a particular generalized solution, which turns out to be the first arrival. Because of its stability with respect to the medium parameters, source location, etc., this generalized solution is computable by finite difference approximation.

The result of Lions (1982) pertains to isotropic media. The central hypothesis of this paper is that the first arrival quasi-compressional ("qP") traveltime is also a stable generalized solution, and therefore computable by suitable finite difference schemes. Also by analogy with the isotropic case, we expect so-called upwind schemes to be particularly successful. Dellinger and Symes (Dellinger, 1991; Dellinger and Symes, 1997) investigated this possibility but did not give full details of a workable algorithm. This paper presents a family of algorithms of Essentially NonOscillatory ("ENO") type (Osher and Sethian, 1988; Osher and Shu, 1991), applied to a depth-evolution ("paraxial") form of the eikonal equation. The computed solution gives an accurate approximate time at every point of a Cartesian grid, which is connected to the source by a first-arriving ray whose velocity vector makes an angle less than a prescribed angle with the vertical. A similar approach has proven quite successful for isotropic traveltime (and amplitude) computation for use in prestack modeling, migration, and inversion (Symes et al., 1994; El-Mageed et al., 1997; Qian et al., 1999). We expect similar applications for the algorithm presented here.

We first derive the eikonal equation for the quasi-compressional ("qP") wave, which is the fastest propagating body wave. [The kinematics of qP waves are considerably simpler than those of quasi-shear phases; an effective finite difference approach to the latter awaits resolution of difficult technical issues.] For downgoing qP waves, the eikonal equation can be transformed to an evolution equation in depth, which we call the paraxial eikonal equation. Definition of the paraxial eikonal depends on the relation of the aperture limitation, i.e. the indicator of the energy flow, defined in terms of the ray velocity vector and the traveltime gradient. We show how to formulate this relation and mention the simplifications which occur in the presence of transverse isotropy with vertical symmetry axis (VTI media). The VTI assumption permits us to give closed form expressions for all of the quantities appearing in the paraxial eikonal in terms of Thomsen's parameters (Thomsen, 1986), whereas in general these must be computed using numerical eigenvalue solvers. We use the ENO-Godunov family of finite difference schemes (Osher and Sethian, 1988; Osher and Shu, 1991) to build qP eikonal solvers of arbitrary order of accuracy, in two- or three-dimensions. Two and three dimensional VTI examples illustrate the accuracy of the ENO schemes as well as the effect of the paraxial (aperture-limiting) assumption. Application to (isotropic and anisotropic) Marmousi models (Alkhalifah, 1997) shows the ability of these schemes to reveal the kinematic effect of anisotropy in geologically complex models.

In one appendix, we give the construction of the paraxial eikonal equation. In another appendix, by using backward raytracing we give an elementary derivation of the basic (first order) Godunov method from which the other schemes derive.

The Paraxial Eikonal Equation for qP Waves

In velocity structures with mild lateral heterogeneity, most reflected wave energy propagates down to the target, then up to the surface. That is, the energy in such a wavefield propagates along downgoing rays: the x_3 ("z") component of the ray velocity vector remains positive from source to target. The traveltime along such downgoing rays increases with depth and should be the solution of an evolution system in depth. For isotropic propagation, Gray and May (1994) suggested modifying the eikonal equation in such a way that (i) the modified equation defines a depth evolution of traveltime, and (ii) solutions of the modified and original eikonal equation are identical at every point connected to the source by a first arriving ray making a angle with the vertical less than 90 degrees. The principal purpose of this section is to explain such a modification, resulting in a *paraxial eikonal equation*, for anisotropic propagation.

High frequency approximation to the elastic wave equation leads to the Christoffel equation (Musgrave, 1970, 84),

$$\sum_k(\sum_{i,l} a_{ijkl}p_ip_l - \delta_{jk})U_k \quad = \quad 0, \tag{1}$$

in which a_{ijkl} are the components of the elastic (Hooke) tensor divided by density, U_k is displacement vector for a particular asymptotic phase, $\mathbf{p} = \nabla\tau$ is the slowness vector, τ is the traveltime or phase of the mode, and δ_{jk} is the Kronecker delta, $\delta_{jk} = 1$ if $j = k$, $\delta_{jk} = 0$ otherwise. Note that all of these quantities depend on the spatial coordinate vector $\mathbf{x} = (x_1, x_2, x_3)$, though in this and some of the following displays this dependence has been suppressed for the sake of clarity. This equation has nontrivial solutions U_k only when

$$\det(\sum_{i,l} a_{ijkl}p_ip_l - \delta_{jk}) \quad = \quad 0. \tag{2}$$

The Christoffel matrix $\sum_{i,l} a_{ijkl}p_ip_l$ is positive definite and scales as $p^2 = \mathbf{p}\cdot\mathbf{p}$. Therefore each eigenvalue takes the form $v^2(\mathbf{x},\mathbf{p})p^2$, where v is a homogeneous function of degree zero in $\mathbf{p}$. The largest eigenvalue, denoted $v_{qP}^2(\mathbf{x},\mathbf{p})$, is simple (Fedorov, 1968, 95), hence depends smoothly on the components of the Christoffel matrix, i.e. on $\mathbf{p}$, and the Hooke tensor. The slowness vector $\mathbf{p}$ for which

$$S(\mathbf{x},\mathbf{p}) \equiv |\mathbf{p}|v_{qP}(\mathbf{x},\mathbf{p}) = 1, \tag{3}$$

form the **qP slowness surface**, and v_{qP} is the qP phase velocity. Note that these vectors solve the Christoffel equation (2).

The qP wave eikonal equation results from combining the slowness surface condition (3) with the slowness identity $\mathbf{p} = \nabla\tau$:

$$S(\mathbf{x},\nabla\tau) = 1. \tag{4}$$

The method of characteristics relates its solution τ of slowness surface equation (4) to the rays of geometrical optics, which are the solutions of the ordinary differential equations

$$\frac{d\mathbf{x}}{dt} \quad = \quad \nabla_{\mathbf{p}}S(\mathbf{x},\mathbf{p}), \tag{5}$$

$$\frac{d\mathbf{p}}{dt} \quad = \quad -\nabla_{\mathbf{x}}S(\mathbf{x},\mathbf{p}), \tag{6}$$

where we have used the homogeneity of the eigenvalue v_{qP} in $\mathbf{p}$, so that $\tau = t$ has dimension of time.

Downgoing rays correspond to the part of the slowness surface on which $\dfrac{dx_3}{dt} > 0$. A useful description of this part of the surface follows from its **convexity** (Musgrave, 1970, 92). In Appendix A, we show that

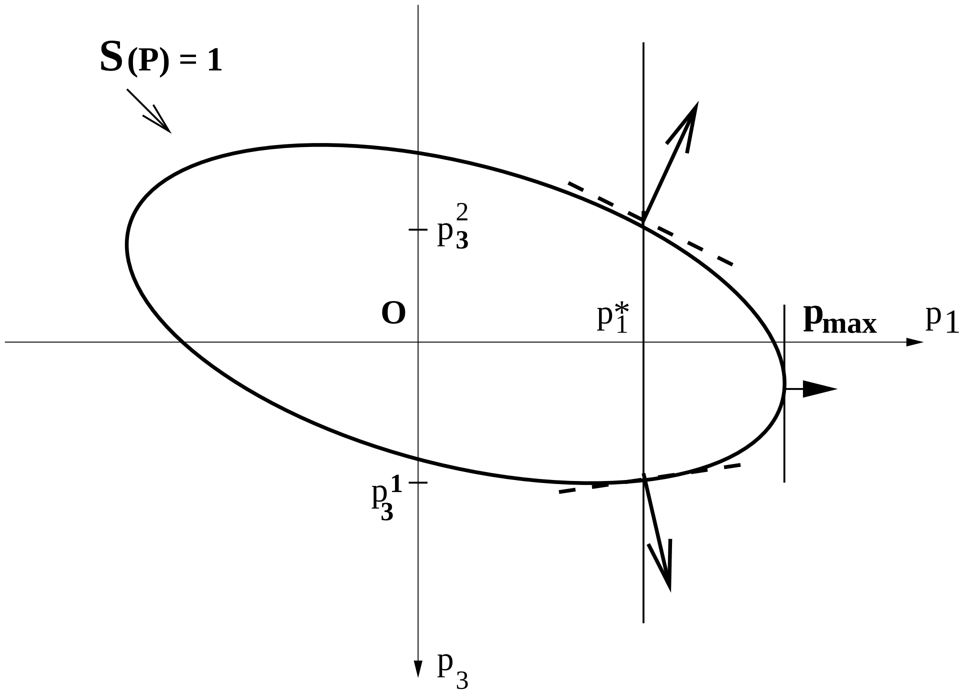


FIG. 1. The p_3 components of outward normals at the two intersections on the convex slowness surface have opposite signs.

- for each $\mathbf{x}$ and horizontal slowness vector (p_1, p_2), there are at most two choices of p_3 for which $\mathbf{p} = (p_1, p_2, p_3)$ solves slowness surface equation (4);
- when two distinct solutions exist, only one satisfies (Figure 1)

$$\frac{dx_3}{dt} = \frac{\partial S(\mathbf{x}, \mathbf{p})}{\partial p_3}(\mathbf{x}, \mathbf{p}) > \mathbf{0}.$$

This choice defines p_3 as a function of $\mathbf{x}, p_1, p_2$:

$$p_3 = H(\mathbf{x}, p_1, p_2), \tag{7}$$

which is also a partial differential equation for τ.

The characteristics (rays) of eikonal equation (7) are downgoing, so they can be parameterized by $x_3 = z$ and satisfy

$$\frac{dx_i}{dx_3} = -\frac{\partial H}{\partial p_i}, \ i = 1, 2; \tag{8}$$

$$\frac{d\tau}{dx_3} = H - p_1\frac{\partial H}{\partial p_1} - p_2\frac{\partial H}{\partial p_2}. \tag{9}$$

The ray group velocity is

$$v_g = \frac{\sqrt{\left(\frac{\partial H}{\partial p_1}\right)^2 + \left(\frac{\partial H}{\partial p_2}\right)^2 + 1}}{H(p_1, p_2) - \frac{\partial H}{\partial p_1}p_1 - \frac{\partial H}{\partial p_2}p_2}, \tag{10}$$

where

$$\frac{\partial H}{\partial p_1} = -\frac{a_{1jkl}p_l g_j g_k}{a_{3jkl}p_l g_j g_k}, \tag{11}$$

$$\frac{\partial H}{\partial p_2} = -\frac{a_{2jkl}p_l g_j g_k}{a_{3jkl}p_l g_j g_k}, \tag{12}$$

with g_j the eigenvector corresponding to v_{qP} and Einstein summation convention assumed.

However neither the partial differential equation (7) nor its rays are well-defined for all p_1, p_2. To remedy this defect, we introduce another family of Hamiltonian functions H_Δ, each of which is identical to H along "safely downgoing" rays, and defined everywhere in phase space. It is convenient to parametrize the horizontal variables (p_1, p_2) by polar coordinates: $(p_1, p_2) = (p\cos\phi, p\sin\phi)$. For each planar angle ϕ, the family of planes perpendicular to $(\cos\phi, \sin\phi, 0)$ is tangent to the slowness surface at exactly one point $(p_{\max}(\phi)\cos\phi, p_{\max}(\phi)\sin\phi, p_3(\phi))$. Choose $0 < \Delta < 1$, for $\phi, p \le (1-\Delta)p_{\max}(\phi)$, set $H_\Delta(p\cos\phi, p\sin\phi) = H(p\cos\phi, p\sin\phi)$; this is the unique root of $S = 1$ with $p_3 > p_3(\phi)$. For $p > (1-\Delta)p_{\max}(\phi)$, set $H_\Delta(p\cos\phi, p\sin\phi) = H((1-\Delta)p_{\max}(\phi)\cos\phi, (1-\Delta)p_{\max}(\phi)\sin\phi)$. By construction, rays of H_Δ are rays of H, and hence of the anisotropic elastic model, so long as their horizontal slowness components $(p\cos\phi, p\sin\phi)$ satisfy $p \le (1-\Delta)p_{\max}(\phi)$.

By convexity (Appendix A), the ray angle ψ with the vertical axis stays safely away from 90 degrees:

$$\begin{aligned}\tan\psi &= \frac{\sqrt{\left(\frac{\partial x_1}{\partial\tau}\right)^2 + \left(\frac{\partial x_2}{\partial\tau}\right)^2}}{\frac{\partial x_3}{\partial\tau}} \\ &= \sqrt{\left(\frac{\partial H_\Delta}{\partial p_1}\right)^2 + \left(\frac{\partial H_\Delta}{\partial p_2}\right)^2} \le O(\frac{1}{\Delta}).\end{aligned} \tag{13}$$

Numerical algorithms for the **paraxial eikonal equation**

$$\frac{\partial\tau}{\partial x_3} = H_\Delta\left(\mathbf{x}, \frac{\partial\tau}{\partial x_1}, \frac{\partial\tau}{\partial x_2}\right), \tag{14}$$

require explicit bounds on $\tan\psi$. Because of the convexity of the slowness surface, we need only look for these at the boundary of the region where $H_\Delta = H$, i.e. where $p = (1-\Delta)p_{\max}(\phi)$. Since H_Δ is radially constant outside of this region, the required bound is simply the maximum length of the 2-vector $\left(\frac{\partial H_\Delta}{\partial p_1}, \frac{\partial H_\Delta}{\partial p_2}\right)$ over the set $\{p = (1-\Delta)p_{\max}(\phi)\}$.

For the remainder of this paper, we put aside any further development of the algorithm for general anisotropy and pass to consideration of a special case, in which most of the details of the paraxial Hamiltonian construction are much simpler.

Application: VTI Media

The elastic modulus matrix for transversely isotropic media with vertical symmetry axis(VTI) has five independent components among twelve nonzero components (Musgrave, 1970; Thomsen, 1986). A closed form solution exists in this case for the eigenvalue problem (2). Consequently, the slowness surface for qP wave can be simplified as

$$\tilde{S}(p_1,p_2,p_3) \;\equiv\; p_3^2 - \frac{2c}{-b+\sqrt{b^2-4ac}} = 0. \tag{15}$$

where

$$a \;\equiv\; \alpha_0^2\beta_0^2, \tag{16}$$

$$b \;\equiv\; 2\alpha_0^2\beta_0^2(1+\delta+(\epsilon-\delta)\frac{\alpha_0^2}{\beta_0^2})(p_1^2+p_2^2)-\alpha_0^2-\beta_0^2, \tag{17}$$

$$c \;\equiv\; ((1+2\epsilon)\alpha_0^2(p_1^2+p_2^2)-1)(\beta_0^2(p_1^2+p_2^2)-1), \tag{18}$$

with Thomsen's parameters (Thomsen, 1986) (Anderson, 1996, 19): α_0 and β_0 are the vertical sound speeds for qP and qS waves, and ϵ and δ are two measures of anisotropy.

Note that $\tilde{S}$ is in the form different from S, but they are equivalent in the sense that they characterize the same qP slowness surface. Clearly,

$$\frac{\partial \tilde{S}}{\partial p_3} \;=\; 2p_3 > 0 \tag{19}$$

if

$$p_3 \;=\; \sqrt{\frac{2c}{-b+\sqrt{b^2-4ac}}} > 0. \tag{20}$$

So the qP eikonal equation for VTI is

$$p_3 \;=\; H\left(p_1,p_2\right) = \sqrt{\frac{2c}{-b+\sqrt{b^2-4ac}}}. \tag{21}$$

To obtain the paraxial Hamiltonian, we need consider only the case of radial angle $\phi = 0$ in the horizontal slowness space because of the rotational symmetry about the vertical axis. First note that for VTI media, we have

$$\sqrt{p_1^2+p_2^2} \;=\; \frac{\sin\theta}{v_{qP}(\theta)}, \tag{22}$$

$$p_3 \;=\; \frac{\cos\theta}{v_{qP}(\theta)} \tag{23}$$

where θ is the phase angle with the vertical direction, $v_{qP}(\theta)$ is the qP phase velocity for phase angle θ. Secondly, to make $\frac{\partial S}{\partial p_3} = 2p_3 = 0$, p_3 has to be zero, so the phase angle $\theta = \frac{\pi}{2}$. It follows that

$$p_{\max}(0) = \left(\sqrt{p_1^2 + p_2^2}\right)_{\max} = \frac{1}{v_{qP}(\frac{\pi}{2})}. \tag{24}$$

Thus

$$H_\Delta(p_1, p_2) = \begin{cases} H(p_1, p_2),\ p = \sqrt{p_1^2 + p_2^2} \le \dfrac{(1-\Delta)}{v_{qP}(\frac{\pi}{2})}; \\ H\left(\dfrac{(1-\Delta)\cos\phi}{v_{qP}(\frac{\pi}{2})}, \dfrac{(1-\Delta)\sin\phi}{v_{qP}(\frac{\pi}{2})}\right), \text{ else.} \end{cases} \tag{25}$$

Next we want to simplify the above Hamiltonian. For $0 < \Delta < 1$, we apply Intermediate Value Theorem to function $f(\theta) = \frac{\sin\theta}{v_{qP}(\theta)}$ in interval $[0, \frac{\pi}{2}]$, there exists $0 < \theta_{\max} < \frac{\pi}{2}$ such that

$$\frac{\sin\theta_{\max}}{v_{qP}(\theta_{\max})} = \frac{(1-\Delta)}{v_{qP}(\frac{\pi}{2})}. \tag{26}$$

Substitute the above relation into paraxial Hamiltonian (25) and combine two branch statements into one, and we have

$$H_{\theta_{\max}}(p_1, p_2) = \sqrt{\max\left(\frac{2c}{-b + \sqrt{b^2 - 4ac}}, \frac{\cos^2(\theta_{\max})}{v_{qP}^2(\theta_{\max})}\right)}, \tag{27}$$

where we have used the subscript $\theta_{\max}$ to emphasize the dependence of the Hamiltonian on phase angle $\theta_{\max}$. We call $\theta_{\max}$ the maximum phase angle. This Hamiltonian implies that we can set up a maximum phase angle to obtain a paraxial Hamiltonian in VTI.

Finally we have a paraxial eikonal equation for qP wave in VTI,

$$p_3 = H_{\theta_{\max}}(p_1, p_2) = \sqrt{\max\left(\frac{2c}{-b + \sqrt{b^2 - 4ac}}, \frac{\cos^2(\theta_{\max})}{v_{qP}^2(\theta_{\max})}\right)}. \tag{28}$$

The two derivatives needed in the numerical algorithms satisfy that

$$\frac{\partial H}{\partial p_1} = -\frac{p_1(2\alpha_0^2(1+2\epsilon)\beta_0^2(p_1^2 + p_2^2) + Ap_3^2 - \alpha_0^2(1+2\epsilon) - \beta_0^2)}{p_3(2\alpha_0^2\beta_0^2 p_3^2 + A(p_1^2 + p_2^2) - \alpha_0^2 - \beta_0^2)}, \tag{29}$$

$$\frac{\partial H}{\partial p_2} = -\frac{p_2(2\alpha_0^2(1+2\epsilon)\beta_0^2(p_1^2 + p_2^2) + Ap_3^2 - \alpha_0^2(1+2\epsilon) - \beta_0^2)}{p_3(2\alpha_0^2\beta_0^2 p_3^2 + A(p_1^2 + p_2^2) - \alpha_0^2 - \beta_0^2)}, \tag{30}$$

with

$$A = 2\alpha_0^2(\alpha_0^2(\epsilon - \delta) + \beta_0^2(1+\delta)). \tag{31}$$

ENO schemes for Paraxial Eikonal Equations

Equation (14) is a nonlinear first-order PDE for traveltime τ. However traveltime is not unique: when the elastic parameters vary with position, in general many rays pass over at least some points in the subsurface. More than one traveltime may be assigned to each such point. One choice of unique traveltime for each subsurface point is the **least time** ("first arrival time"). It turns out that this first arrival time field is a **solution** of the eikonal equation (in a generalized sense). In fact it is the only possible single-valued choice of traveltime field uniquely and stably determined by the data (source position and elastic parameters). These assertions were established some time ago (Lions, 1982), and we surmise that they apply as well to anisotropic problems with a convex slowness surface.

To compute the first arrival traveltime field by a grid-based finite difference scheme, we derive a first order upwind scheme from ray tracing, rather than directly from the eikonal equation, demonstrated in Appendix B. Because our use of ray tracing in the derivation inherently honors causality, the resulting difference scheme is upwind, a term more or less synonymous with causality in this setting.

The first order upwind scheme is hopelessly inefficient; moreover the traveltime field produced by it is so inaccurate as to prevent convergent computation of associated fields such as take-off angles, amplitudes, etc. So we use the first order scheme as a building block to design high order schemes. To increase the order of convergence, we employ higher order ENO(Essentially NonOscillatory) refinements introduced by Osher and Sethian (Osher and Sethian, 1988) and Osher and Shu (Osher and Shu, 1991).

ENO schemes belong to a family of upwind schemes, which are attractive for the three reasons: (1) stable schemes of arbitrarily high order of accuracy exist, permitting accurate solutions on coarse grids; (2) versions exist in any dimension – in particular, three dimensional analogues are available; (3) ENO schemes are easy to program.

The principal feature of the ENO scheme is high order interpolation with adaptive stencil which tries to avoid high gradient regions whenever possible. To match with this high order discretization in x_1, x_2 direction, we use high order TVD (total variation diminishing) Runge-Kutta scheme to guarantee nonlinear stability and high order accuracy in $x_3 = z$ direction.

Given mesh sizes Δx_1, Δx_2 and Δx_3, we denote $\tau_{m,k}^n$ as the numerical approximation to the viscosity solution $\tau(x_1^m, x_2^k, x_3^n)$ of equation (14) at the grid point (x_1^m, x_2^k, x_3^n). Define the backward $(-)$ and forward $(+)$ first order difference quotient approximations to the left and right derivatives of $\tau(x_1, x_2, x_3)$ at the location

(x_1^m, x_2^k, x_3^n) with respect to x_1 and x_2 as

$$D^{\pm}_{x_1}\tau^n_{m,k} = \pm\frac{\tau^n_{m\pm 1,k} - \tau^n_{m,k}}{\Delta x_1}, \qquad D^{\pm}_{x_2}\tau^n_{m,k} = \pm\frac{\tau^n_{m,k\pm 1} - \tau^n_{m,k}}{\Delta x_2}. \tag{32}$$

The second order ENO refinements to $\frac{\partial \tau}{\partial x_1}$ (Osher and Sethian, 1988) are

$$D^{+,2}_{x_1}\tau = D^{+}_{x_1}\tau - \frac{1}{2}\Delta x_1 m(D^{+}_{x_1}D^{+}_{x_1}\tau, D^{-}_{x_1}D^{+}_{x_1}\tau), \tag{33}$$

$$D^{-,2}_{x_1}\tau = D^{-}_{x_1}\tau + \frac{1}{2}\Delta x_1 m(D^{-}_{x_1}D^{-}_{x_1}\tau, D^{-}_{x_1}D^{+}_{x_1}\tau) \tag{34}$$

with

$$m(x,y) = \min(\max(x,0), \max(y,0)) + \max(\min(x,0), \min(y,0)). \tag{35}$$

ENO refinements for $\frac{\partial \tau}{\partial x_2}$ are defined similarly.

So a second order ENO Runge-Kutta scheme for equation(14) can be formulated as

$$\tau^{n+\frac{1}{2}}_{m,k} = \tau^n_{m,k} + \Delta x_3^{cfl} H_\Delta \left((\widehat{\frac{\partial \tau}{\partial x_1}})^n_{m,k}, (\widehat{\frac{\partial \tau}{\partial x_2}})^n_{m,k} \right), \tag{36}$$

$$\tau^{n+1}_{m,k} = \frac{1}{2}\left(\tau^n_{m,k} + \tau^{n+\frac{1}{2}}_{m,k} + \Delta x_3^{cfl} H_\Delta \left((\widehat{\frac{\partial \tau}{\partial x_1}})^{n+\frac{1}{2}}_{m,k}, (\widehat{\frac{\partial \tau}{\partial x_2}})^{n+\frac{1}{2}}_{m,k} \right) \right), \tag{37}$$

where

$$(\widehat{\frac{\partial \tau}{\partial x_1}})^n_{m,k} = \mathrm{maxmod}(\max(D^{-,2}_{x_1}\tau^n_m, 0), \min(D^{+,2}_{x_1}\tau^n_m, 0)), \tag{38}$$

$$(\widehat{\frac{\partial \tau}{\partial x_2}})^n_{m,k} = \mathrm{maxmod}(\max(D^{-,2}_{x_2}\tau^n_m, 0), \min(D^{+,2}_{x_2}\tau^n_m, 0)), \tag{39}$$

with maxmod returning the larger value in modulus; ${\Delta x_3}^{cfl}$ is CFL step,

$${\Delta x_3}^{cfl}\left(\max_{p_1,p_2} \sqrt{\left(\frac{\partial H_\Delta}{\partial p_1}\right)^2 + \left(\frac{\partial H_\Delta}{\partial p_2}\right)^2} \right) \le \frac{\Delta x_1 \Delta x_2}{\sqrt{{\Delta x_1}^2 + {\Delta x_2}^2}}, \tag{40}$$

with the maximum taken over the relevant range of p_1 and p_2.

Numerical Examples

To give some idea of the accuracy obtainable with the difference schemes outlined in the preceding section, we demonstrate the scheme on smooth VTI models and the (isotropic and anisotropic) Marmousi model; that is, we solve the paraxial eikonal equation (28) by using a second order ENO scheme. All examples are assumed to be of constant density.

First we have to address the traveltime initialization. Due to the singularity of the traveltime field at the source which will lead to the contamination of global numerical accuracy, to initialize the traveltime we have to use some special techniques, such as the adaptive grid method (Qian et al., 1999), the local uniform mesh refinement (Kim and Cook, 1998). However, here we assume a homogeneous layer near the source (which is supposed to be easy to assign an accurate traveltime) and start the finite difference scheme some distance away from the source; namely, we use a nonlinear iteration method to compute the group velocity and initialize directly the traveltime at every grid point on a surface away from the source. This initialization technique, called the analytic method, is detailed in Appendix C.

2-D VTI

The first example occupies the rectangle $\{-0.5 \text{ km} \leq x_1 \leq 0.5 \text{ km}, 0 \leq x_3 \leq 1$ km $\}$; The source is located at $x_1 = 0.0$ km, $x_3 = 0.0$ km. The four Thomsen's parameters of homogeneous Green River Shale are

$$\begin{aligned} \alpha_0 &= 3.330\text{km/s}, \\ \beta_0 &= 1.768\text{km/s}, \\ \epsilon &= 0.195, \\ \delta &= -0.220. \end{aligned}$$

Because the parameter δ has the same magnitude as ϵ, the near-vertical anisotropic response is dominated by δ (Thomsen, 1986), so we use this example to demonstrate not only the accuracy of the second order scheme, but also the capability of the scheme in capturing the anisotropy of wave propagation. The initial data depth is at $x_3 = 0.24$ km; that is, the initial data for the finite difference scheme is given at this depth. The maximum phase angle is taken as $\theta_{\max} = 80$ degrees. The results are shown on Table 1, where Δx_1 is the x_1 direction grid sampling, Δx_3 is the x_3 direction grid sampling (taken as $\Delta x_3 = 0.01$ km), the maximum absolute error Abs.Err and maximum relative error Rel.Err are measured at bottom $x_3 = 1$ km, α is the estimated convergence order,

$$\begin{aligned} \text{Abs.Err}(\tau, \Delta x_1) &= \max |\tau_{ana} - \tau_{fd}^{\Delta x_1}|, \\ \text{Rel.Err}(\tau, \Delta x_1) &= \frac{\max |\tau_{ana} - \tau_{fd}^{\Delta x_1}|}{\max |\tau_{ana}|} \end{aligned}$$

Table 1. Convergence order of the scheme

Δx_1	Abs.Err($\tau, \Delta x_1$)(s)	Rel.Err($\tau, \Delta x_1$)	α
0.08	7.3754e-04	0.00168	
0.04	2.1380e-04	6.1296e-04	1.45
0.02	5.5932e-05	1.6035e-04	1.93
0.01	1.4162e-05	4.0602e-05	1.98
0.005	3.5643e-06	1.0218e-05	1.99

and

$$\alpha = \frac{1}{\log 2} \log\left(\frac{\text{Rel.Err}(\tau, 2\Delta x_1)}{\text{Rel.Err}(\tau, \Delta x_1)}\right),$$

where τ_{ana} denotes the traveltime from the analytic method and τ_{fd} the traveltime from the finite difference scheme.

From this table, both the absolute error and relative error are decreased four times as Δx_1 is halved, and the accuracy order α is going to 2 as Δx_1 goes to zero, so this scheme is of second order accuracy. When $\Delta x_1 = \Delta x_3 = 0.01$ km, the maximum absolute traveltime error at bottom is less than $\frac{1}{50}$ ms.

In the second example, the model occupies the rectangle $\{-0.5 \text{ km} \leq x_1 \leq 0.5$ km, $0 \leq x_3 \leq 1$ km $\}$, the source is located at $x_1 = 0.0$ km, $x_3 = 0.0$ km. The four elastic parameters are

$$\begin{aligned}
\alpha_0 &= \sqrt{11.0889 + x_1 + x_3}, \\
\beta_0 &= \sqrt{3.1329 + 0.5x_1 + 0.5x_3}, \\
\epsilon &= \frac{4.3247 + x_1 + x_3}{2(11.0889 + x_1 + x_3)}, \\
\delta &= \frac{(4.9477 + x_1 + x_3)^2 - (7.9560 + 0.5x_1 + 0.5x_3)^2}{2(11.0889 + x_1 + x_3)(7.9560 + 0.5x_1 + 0.5x_3)},
\end{aligned}$$

which are perturbations to Thomsen's parameters of homogeneous Green River Shale. To simulate the VTI media, these parameters necessarily satisfy the stress-strain coefficient inequalities in (Berryman, 1979). We design this example to test the capability of the scheme in dealing with both the vertical and lateral variations.

The grid sampling is $\Delta x_1 = \Delta x_3 = 0.01$ km and the maximum phase angle is taken as $\theta_{\max} = 80$ degrees, Figure 3. Because of vertical and lateral variations of the example near the source, we can not use the initialization technique used in the first example; instead, we use horizontal qP sound speed ($= (1 + \epsilon)\alpha_0$) and an adaptive integration method (Stoer and Bulirsch, 1992) to obtain the traveltime at $x_3 = 0$ km. Figure 2-left shows the traveltime isochrons for the 2D model. Because the lateral variation is not symmetric with respect to the source, the isochrons could

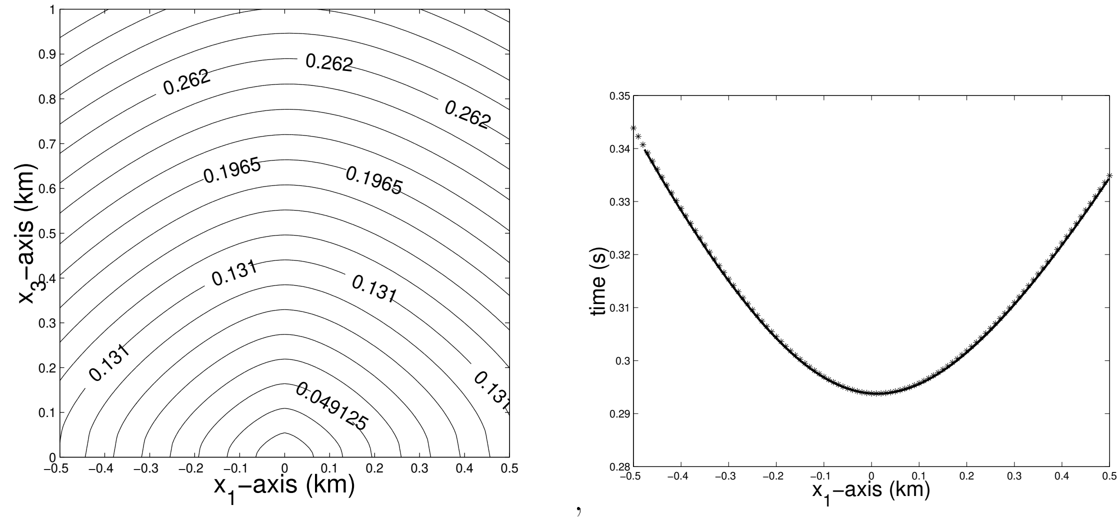


FIG. 2. Left: Traveltime isochrons for the 2-D model with vertical and lateral variations: anisotropic effects and lateral variations on the wave propagation are evident. Right: Traveltime comparison at $x_3 = 1.0$ km for the model with vertical and lateral variations: ENO traveltime (*) and Ray-tracing traveltime (-), they are consistent with each other.

not reach the same depth at $x_1 = 0.5$ km and $x_1 = -0.5$ km, the isochron of 0.262 s, for example. Also, from the figure we can see that the limitation on the maximum phase angle (artificial plane wave approximation) plays a role when the phase angle is near 90 degrees. Figure 2-right shows the calibration result for the traveltime at bottom $x_3 = 1.0$ km from the ENO scheme and the ray tracing method (Cerveny, 1972). Because the raytracing method (Cerveny, 1972) can not give us the traveltime at grid point directly and we can not guarantee that every traced ray will pass through the bottom, we have to use a linear interpolation method to extract the traveltime for some points (not necessarily grid points) from the ray tracing solutions. The ray tracing method uses the traveltime as a running parameter along the ray; i.e. (x_1, x_3, p_1, p_3) are parameterized by traveltime τ. When a ray passes through the bottom $x_3 = 1$ km, we pick out the nearest two points on the ray which embrace the bottom, then approximate by linear interpolation the x_1 coordinate and traveltime of the intersection point between the ray and the bottom. In the computation, the ray tracing method traces rays from phase angle -80 degrees to 80 degrees with a sampling interval of 3.14 degrees. Because the ray fan from ray tracing is rarefying quickly away from the source, most of the intersection points are not at the grid point, so it is not appropriate to assess the accuracy of ENO scheme and ray tracing method this way. Nevertheless, they are perfectly consistent with each other, Figure 2-right. Also, the effect of the lateral variations of the model is evident on the traveltime curves at the bottom.

3-D VTI

The first 3-D VTI example is the homogeneous Green River shale, and the four Thomsen parameters are the same as in the 2D homogeneous case. The model occupies the cube $\{-0.5 \text{ km} \leq x_1 \leq 0.5 \text{ km}, -0.5 \text{ km} \leq x_2 \leq 0.5 \text{ km}, 0 \leq x_3 \leq 1.0$ km $\}$, and the source is located at the center of the surface $x_3 = 0$.

The grid sampling is $\Delta x_1 = \Delta x_2 = \Delta x_3 = 0.02$ km, and the maximum phase angle is taken as $\theta_{\max} = 65$ degrees. We assume the initial homogeneous layer is at 0.1 km; that is, we start the finite difference scheme at 0.1 km. To initialize the traveltime at 0.1 km, we use the analytic method (Appendix E). Figure 3-left shows the isochrons at the bottom $x_3 = 1$ km, and the isochrons are circles because of the transverse isotropy. Figure 3-right shows the vertical traveltime profile at $x_2 = 0.3$ km, and the anisotropic effect is evident. Figure 4 shows the traveltime comparison at gridline $x_2 = 0.2$ km and $x_3 = 1.0$ km between the ENO traveltime and the analytic traveltime, and the maximum absolute error is less than 0.19 ms.

The second 3-D VTI example has the same geometry as the first model. The four Thomsen's parameters are

$$\begin{aligned}
\alpha_0 &= \sqrt{11.0889 + x_3}, \\
\beta_0 &= \sqrt{3.1329 + 0.5x_3}, \\
\epsilon &= \frac{4.3247 + x_3}{2(11.0889 + x_3)}, \\
\delta &= \frac{(4.9477 + x_3)^2 - (7.9560 + 0.5x_3)^2}{2(11.0889 + x_3)(7.9560 + 0.5x_3)},
\end{aligned}$$

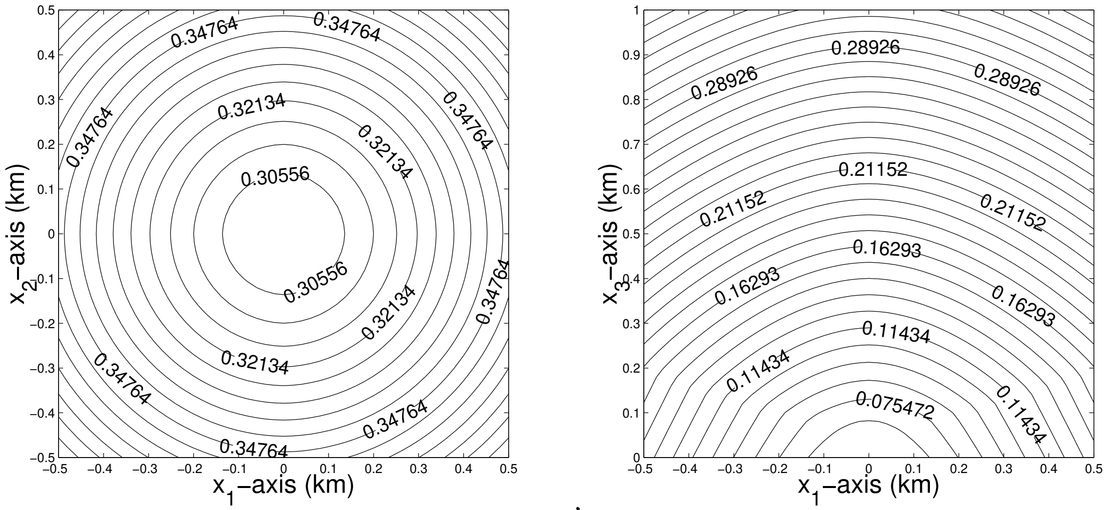
For this model, we use the discretization parameters as used in the homegeneous model. Because of the vertical variation near the source, we start the finite difference scheme at $x_3 = 0$ km, and the initial traveltime is assigned by using the horizontal qP sound speed, and the results are shown in Figure 5.

These two 3-D examples show that the scheme can be used to compute the first arrival traveltime field of 3-D smooth VTI models.

Isotropic and anisotropic Marmousi model

Because the Marmousi model is well known as a complex velocity model, we test the ENO scheme to see its robustness and stability on the original isotropic and the anisotropic Marmousi model created by Alkhalifah (Alkhalifah, 1997). We take portions of these two models to test our method; that is, the sampling domain is $\{4.5 \text{ km} \leq x_1 \leq 7.5 \text{ km}, 0 \leq x_3 \leq 2.9875 \text{ km}\}$, and the lateral and depth samples are 241 and 240, respectively, with the sampling intervals equal to 12.5 m.

In the computation, the maximum phase angle is taken as $\theta_{\max} = 75$ degrees. The source is located at $x_1 = 6.0$ km, $x_3 = 0.0$ km. For the isotropic Marmousi model, we initialize the traveltime at $x_3 = 0$ km by using the constant surface velocity. For



,

FIG. 3. 3-D homogeneous VTI model. The source is located at $x_1 = x_2 = x_3 = 0.0$. Left: Horizontal traveltime slice at $x_3 = 1.0$ km, isochrons are circles because of the transverse isotropy; Right: Vertical traveltime slice at $x_2 = 0.30$ km.

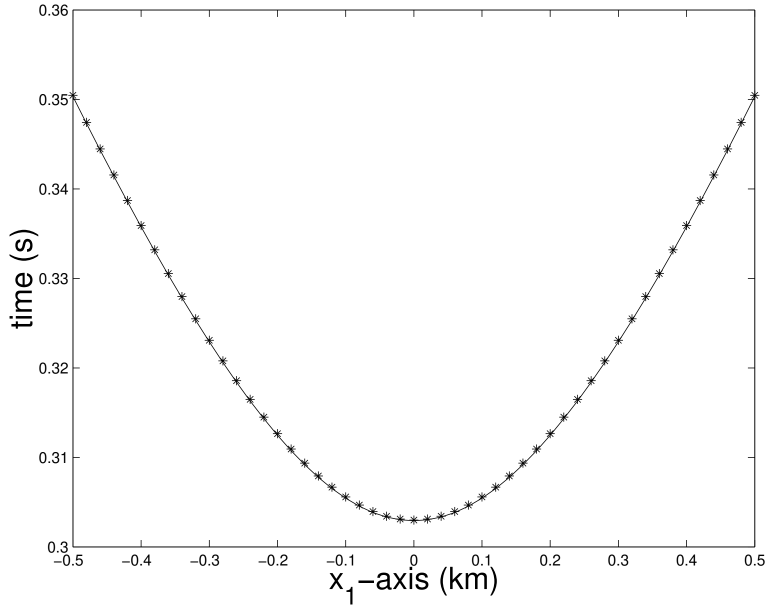
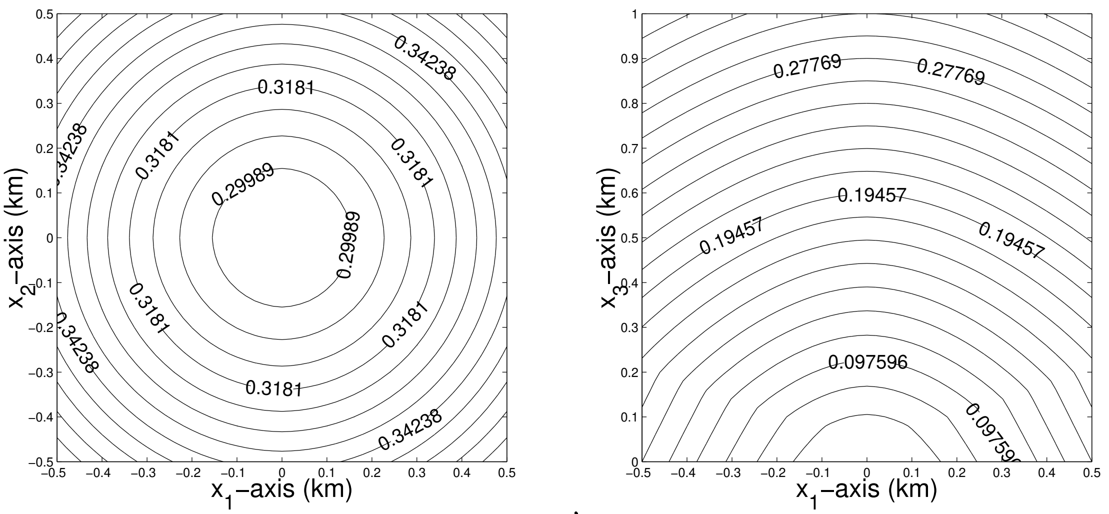


FIG. 4. Traveltime comparison at gridline $x_2 = 0.2$ km, $x_3 = 1.0$ km for 3-D homogeneous VTI model: ENO traveltime (*) and Analytic traveltime (-), ENO traveltime has almost the same accuracy as the analytic traveltime.



,

FIG. 5. 3-D VTI model with vertical variation. The source is located at $x_1 = x_2 = x_3 = 0.0$. Left: Horizontal traveltime slice at $x_3 = 1.0$ km, isochrons are circles because of the transverse isotropy; Right: Vertical traveltime slice at $x_2 = 0.30$ km.

the anisotropic Marmousi model, the traveltime is initialized by using the horizontal sound velocity. The final results are shown for windowed portions of the two models, corresponding to $\{5.4 \text{ km} \leq x_1 \leq 6.8 \text{ km}, 1.5 \text{ km} \leq x_3 \leq 2.9875 \text{ km} \}$, Figure 6. Because the vertical velocity in the anisotropic Marmousi model is the same as the isotropic Marmousi model, the portion of the wavefront corresponding to vertical wave propagation is similar (Alkhalifah, 1997), which we can see by subtracting the two traveltime fields, with results shown in Figure 7. However, there are still some differences between the two models, especially near the upper-right corner where η is larger than other places, and the maximum traveltime difference is 5.9 ms. Because we are using a second order finite difference scheme, we believe that these traveltime differences are due to the anisotropic effect rather than the numerical errors. These computations show that our scheme can deal with complex geological models.

CONCLUSIONS

We have formulated for the heterogeneous anisotropic solids the paraxial eikonal equation satisfied by the first arrival traveltime associated with the qP wave propagation. We have presented some implementation details for computing the paraxial Hamiltonian and illustrated the application to the transversely isotropic solid. We have derived a first order finite difference scheme for solving this paraxial eikonal equation. Because this first order scheme is based on backward raytracing, causality is automatically honored. We have constructed a second order scheme by using the first order scheme as a building block. Higher order schemes may be constructed by further use of this method (Qian et al., 1999). These schemes solve the qP paraxial eikonal equation in $O(N)$ floating point operations (where N is the number of grid

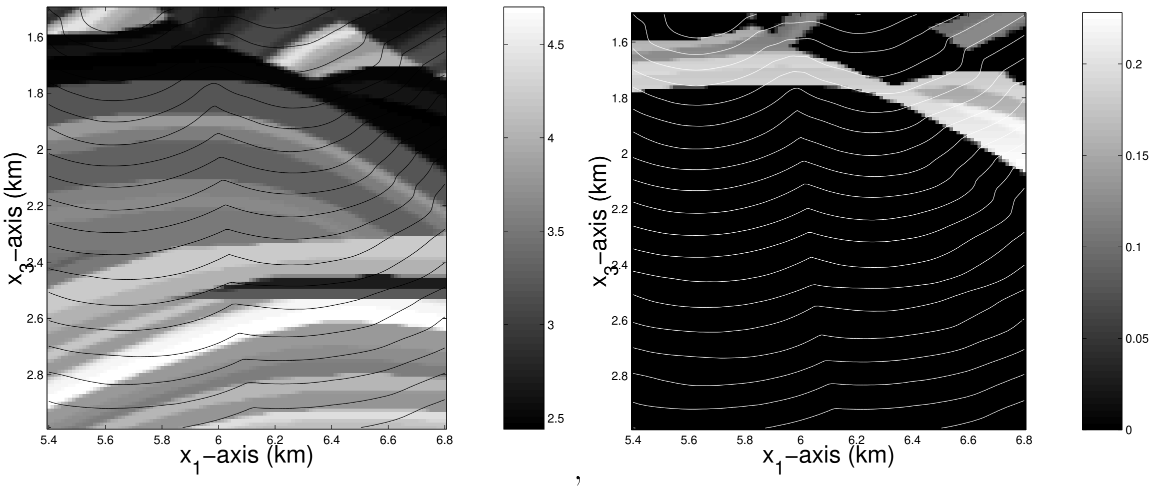


FIG. 6. Traveltime contours overlaying the model. The source is positioned at $x_1 = 6.0$ km, $x_3 = 0.0$ km. Left: Isotropic Marmousi velocity model; Right: Anisotropic Marmousi η model.

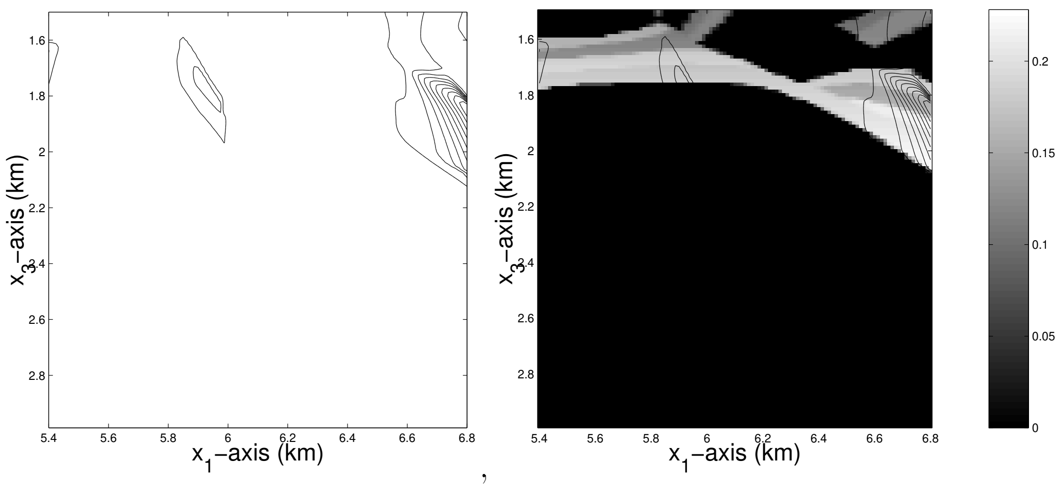


FIG. 7. Anisotropic effects on the wave propagation. Left: nonzero contours of traveltime differences between the two Marmousi models concentrates on the region where η is larger; Right: contours of traveltime differences overlaying η model. The maximum traveltime difference is 5.9ms.

points). Extensive numerical results have shown that our scheme is accurate, efficient and can deal with complex models and models with both lateral and vertical variations. The chief shortcoming is that the paraxial assumption permits us to compute traveltimes along downgoing rays only; for example, we do not compute the overturning wavefronts.

The scheme can be used in many geophysical applications requiring modeling anisotropic wave propagation such as 3-D Kirchhoff migration and modeling, tomography, and 3-D controlled illumination modeling.

ACKNOWLEDGEMENT

The authors performed part of the research reported here while guests of the Stanford Exploration Project; J.L. Qian and W.W. Symes thank SEP members and director Prof. Jon Claerbout, for their hospitality and many stimulating discussions. J.L. Qian thanks Dr. Joe Dellinger and Dr. S. Kim for many helpful discussions.

This work was partially supported by the National Science Foundation under grant numbers DMS 9627355, the Office of Naval Research under grant number N00014-96-1-0156, the U. S. Department of Energy under grant number DE-FG07-97 ER14827 and The Rice Inversion Project. TRIP Sponsors for 1999 are Amerada Hess, Amoco Research, Conoco Inc., Exxon Production Research Co., Geophysical Development Corporation, Landmark Graphics, Mobil, and Western Geophysical.

REFERENCES

Alkhalifah, T. (1997). Anisotropic Marmousi data set. In *SEP95*. Stanford University. (http://sepwww.stanford.edu/sep/tariq/Html_marmousi).

Anderson, J. E. (1996). *Imaging in Transversely Isotropic Media with a Vertical Symmetry Axis*. PhD thesis, Center for Wave Phenomena, Colorado School of Mines, Golden, CO80401-1887.

Berryman, J. G. (1979). Long-wave elastic anisotropy in transversely isotropic media. *Geophysics*, 44:896–917.

Cerveny, V. (1972). Seismic rays and ray intensities in inhomogeneous anisotropic media. *Geophys. J. R.Astr.Soc.*, 29:1–13.

Cherrett, A. and Singh, S. (1998). Traveltime and polarisation tomography in 3-D anisotropic media. In *68th SEG Annual Internat. Mtg. Expanded Abstracts*, pages 1859–1862. Soc. Expl. Geophys.

Dellinger, J. (1991). *Anisotropic Seismic Wave Propagation*. PhD thesis, Stanford University, Stanford, CA94305.

Dellinger, J. and Symes, W. W. (1997). Anisotropic finite-difference traveltimes using a Hamilton-Jacobi solver. In *67th SEG Annual Internat. Mtg. Expanded Abstracts*, pages 1786–1789. Soc. Expl. Geophys.

El-Mageed, M., Kim, S., and Symes, W. W. (1997). 3-D Kirchhoff migration using finite difference traveltimes and amplitudes. In *Annual Report. The Rice Inversion Project*. Rice University. (http://www.trip.caam.rice.edu/).

Fedorov, F. I. (1968). *Theory of Elastic waves in crystals*. Plenum press.

Gray, S. and May, W. (1994). Kirchhoff migration using eikonal equation traveltimes. *Geophysics*, 59:810–817.

Kim, S. and Cook, R. (1998). Second-order finite difference traveltimes in three dimensions. In *Annual Report, The Rice Inversion Project*. Rice University. (http://www.trip.caam.rice.edu/).

Lions, P. L. (1982). *Generalized solutions of Hamilton-Jacobi equations*. Pitman Advanced Publishing Program.

Musgrave, M. J. P. (1970). *Crystal Acoustics*. Holden-Day.

Osher, S. and Shu, C. W. (1991). High-order Essentially NonOscillatory schemes for Hamilton-Jacobi equations. *SIAM J. Num. Anal.*, 28:907–922.

Osher, S. J. and Sethian, J. (1988). Fronts propagating with curvature dependent speed: algorithms based on Hamilton-Jacobi formulations. *J. Comput. Phys.*, 79:12–49.

Pratt, R. and Chapman, C. (1992). Traveltime tomography in anisotropic media -II: Application. *Geophys. J. Int.*, 109:20–37.

Qian, J. L., Belfi, C., and Symes, W. (1999). Adaptive finite difference method for traveltime and amplitude. In *69th SEG Annual Internat. Mtg. Expanded Abstracts*, pages 1763–1766. Soc. Expl. Geophys.

Qin, F. and Schuster, G. (1993). First-arrival traveltime calculation for anisotropic media. *Geophysics*, 58:1349–1358.

Reshef, M. and Kosloff, D. (1986). Migration of common shot gathers. *Geophysics*, 51:324–331.

Schneider, W. A. J. (1995). Robust and efficient upwind finite-difference traveltime calculations in three dimensions. *Geophysics*, 60:1108–1117.

Schneider, W. A. J., Ranzinger, K., Balch, A., and Kruse, C. (1992). A dynamic programming approach to first arrival traveltime computation in media with arbitrarily distributed velocities. *Geophysics*, 57:39–50.

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

Shearer, P. M. and Chapman, C. H. (1988). Ray tracing in anisotropic media with a linear gradient. *Geophysical Journal*, 94:575–580.

Stoer, J. and Bulirsch, R. (1992). *Introduction to numerical analysis.* Springer-Verlag, New York.

Symes, W. W., Versteeg, R., Sei, A., and Tran, R. H. (1994). Kirchhoff simulation, migration and inversion using finite difference traveltimes and amplitudes. In *Annual Report, The Rice Inversion Project.* Rice University. (http://www.trip.caam.rice.edu/).

Thomsen, L. (1986). Weak elastic anisotropy. *Geophysics*, 51:1954–1966.

Tsvankin, I. and Thomsen, L. (1995). Inversion of reflection traveltimes for transverse isotropy. *Geophysics*, 60:1095–1107.

van Trier, J. and Symes, W. W. (1991). Upwind finite-difference calculation of traveltimes. *Geophysics*, 56:812–821.

Vidale, J. (1988). Finite-difference calculation of travel times. *Bull., Seis. Soc. Am.*, 78:2062–2076.

Appendix A. SOME RESULTS FROM CONVEXITY

Assume the slowness surface S is C^1, bounded, closed and strictly convex, we prove the following results.

Theorem 1 For any given $(p_1^*, p_2^*) \in \{(p_1, p_2) : \exists p_3 \text{ s.t. } S(p_1, p_2, p_3) = 0\}$, draw the straight line passing through (p_1^*, p_2^*) and parallel to p_3 axis in the (p_1, p_2, p_3) space, then there exist at most two intersection points (p_1^*, p_2^*, p_3^1) and (p_1^*, p_2^*, p_3^2). The outward normals at the two intersection points have p_3 components of opposite signs.

Proof: Consider 2-D case, 3-D similar. By strict convexity of slowness surface, for p_1^* given, there are at most two intersection points. Without loss of generality, for p_1^* given, we can assume there exist p_3^1 and p_3^2 such that $S(p_1^*, p_3^1) = S(p_1^*, p_3^2) = 0$, Figure 1.

Assume $p_3^1 > p_3^2$, for $p_3 \in [p_3^2, p_3^1]$ define two functions

$$f(p_3) = \sup\{\alpha : S(\alpha, p_3) = 0, \alpha \geq p_1^*\}, \tag{A-1}$$

$$g(p_3) = \inf\{\alpha : S(\alpha, p_3) = 0, \alpha \leq p_1^*\}. \tag{A-2}$$

Both f and g are well defined and convex(or concave) by convexity of S, one of them must satisfy that

$$f(p_3^1) = p_1^* = f(p_3^2) \tag{A-3}$$

or

$$g(p_3^1) \;=\; p_1^* = g(p_3^2). \tag{A-4}$$

Without loss of generality, suppose f satisfies the above constraint, its graph corresponds to a section of slowness surface S.

By Rolle's mean theorem, there exists a $\xi \in (p_3^2, p_3^1)$ such that

$$\frac{df}{dp_3}(\xi) \;=\; 0. \tag{A-5}$$

However, $\frac{df}{dp_3}$ is strictly monotonic by strict convexity of f, so

$$\frac{df}{dp_3}(p_3^1)\frac{df}{dp_3}(p_3^2) \;<\; 0. \tag{A-6}$$

Assume $\frac{df}{dp_3}(p_3^1) < 0$ and $\frac{df}{dp_3}(p_3^2) > 0$, they define the tangents at those two points, Figure 1. It follows that the outward normal at (p_1^*, p_3^1) has acute angle with positive p_3 direction, so it has positive p_3 component. Similarly, the outward normal at (p_1^*, p_3^2) has negative p_3 component.

By **Theorem 1**, the following function is well defined.

Definition 1 For $(p_1, p_2) \in \{(p_1, p_2) : \exists p_3 \text{ s.t. } S(p_1, p_2, p_3) = 0\}$, define $p_3 = H(p_1, p_2)$ satisfying that

$$S(p_1, p_2, H(p_1, p_2)) = 0, \qquad \frac{\partial S}{\partial p_3}(p_1, p_2, H(p_1, p_2)) \geq 0. \tag{A-7}$$

Theorem 2 2D case. Suppose $\partial\Omega = \{(p_1, p_3) : S(p_1, p_3) = 0\}$ is a bounded closed set. Under the assumption of **Theorem 1**, the following two sets are both nonempty:

$$F \;=\; \{p_1^* : \exists p_3^1, p_3^2 \Rightarrow f(p_3^1) = p_1^* = f(p_3^2)\}, \tag{A-8}$$

$$G \;=\; \{p_1^* : \exists p_3^1, p_3^2 \Rightarrow g(p_3^1) = p_1^* = g(p_3^2)\}, \tag{A-9}$$

where p_1^* satisfies that there exist p_3^1 and p_3^2 such that $S(p_1^*, p_3^1) = S(p_1^*, p_3^2) = 0$, and f and g are defined in **Theorem 1**. Moreover, $\inf F \leq \sup G$.

Proof: Obviously, at least one of them is nonempty. Without loss of generality, assume F is nonempty. Suppose G is empty, it means that all the $p_1 \in \{p_1 : \exists p_3 \text{ s.t. } S(p_1, p_3) = 0\}$ is in F. Because F is a closed bounded set, $\beta = \inf F > -\infty$. So for $p_1^* = \beta$, $\exists p_3^1 and dp_3^2$ such that $\frac{df}{dp_3}$ is strictly monotonic in $[p_3^2, p_3^1]$, hence S can not be closed, which is a contradiction, so G is nonempty. Moreover, if $\inf F > \sup G$, S consists of two branches which are not connected, so it can not be a closed convex set. It follows that $\inf F \leq \sup G$.

Theorem 3 2D case. Under the assumption of **Theorem 1** and **Theorem 2**, among all (p_1, p_3) such that

$$S(p_1,p_3)=0, \qquad \frac{\partial S}{\partial p_3}(p_1,p_3)\geq 0, \tag{A-10}$$

there are $p_1^{\min}$, $p_1^{\max}$ such that

$$S(p_1^{\min},H(p_1^{\min}))=0, \qquad \frac{1}{|\nabla S|}\frac{\partial S}{\partial p_3}(p_1^{\min},H(p_1^{\min}))=0, \tag{A-11}$$

$$S(p_1^{\max},H(p_1^{\max}))=0, \qquad \frac{1}{|\nabla S|}\frac{\partial S}{\partial p_3}(p_1^{\max},H(p_1^{\max}))=0. \tag{A-12}$$

For $0<\Delta<1$ and $p_1\in[(1+\Delta)p_1^{\min},(1-\Delta)p_1^{\max}]$,

$$\frac{1}{|\nabla S|}\frac{\partial S}{\partial p_3}(p_1,H(p_1)) \;\;\geq\;\; O(\Delta)>0. \tag{A-13}$$

Proof By **Theorem 1** and $F\neq\emptyset$, for $\forall p_1^*\in F$, $\exists f$ such that $\frac{df}{dp_3}\leq 0$ and $\frac{df}{dp_3}$ is strictly decreasing in $[\xi, p_3^1]$, where $\frac{df}{dp_3}(\xi)=0$. Define $p_1^{\max}=f(\xi)$, then $\xi = H(p_1^{\max})$, so (A-12) holds. By strict monotonicity of function f in $[\xi,p_3^1]$, for $p_1=(1-\Delta)p_1^{\max}$, there exists $\eta\in(\xi,p_3^1)$ such that

$$f(\eta)=(1-\Delta)p_1^{\max} \quad , \quad \sigma=\frac{df}{dp_3}(\eta)<0. \tag{A-14}$$

Then the outward normal at $(f(\eta),\eta)$ is $(\frac{1}{\sqrt{1+\sigma^2}},\frac{|\sigma|}{\sqrt{1+\sigma^2}})$. Because $\frac{df}{dp_3}$ is strictly decreasing, $\frac{df}{dp_3}(p_3)\leq\frac{df}{dp_3}(\eta)<0$ if $p_3\geq\eta$. Furthermore, for $p_1\in F$ and $p_1\leq(1-\Delta)p_1^{\max}$,

$$\frac{1}{|\nabla S(p_1,H(p_1))|}\frac{\partial S}{\partial p_3}(p_1,H(p_1)) \;\;\geq\;\; O(\Delta)>0. \tag{A-15}$$

Similarly, by $G\neq\emptyset$, there exists $p_1^{\min}$ such that (A-11) holds. Also, for $p_1\in G$, $p_1\geq(1+\Delta)p_1^{\min}$,

$$\frac{1}{|\nabla S(p_1,H(p_1))|}\frac{\partial S}{\partial p_3}(p_1,H(p_1)) \;\;\geq\;\; O(\Delta)>0. \tag{A-16}$$

Finally, by $\inf F \leq \sup G$, for $p_1 \in [(1+\Delta)p_1^{\min}, (1-\Delta)p_1^{\max}]$ we have

$$\frac{1}{|\nabla S(p_1, H(p_1))|}\frac{\partial S}{\partial p_3}(p_1, H(p_1)) \;\; \geq \;\; O(\Delta) > 0. \qquad \text{(A-17)}$$

Definition 2 2D Hamiltonian H_Δ. Given $\Delta > 0$, $p_1^{\min}$, $p_1^{\max}$ defined as in **Theorem 3**, for $\forall p_1 \in R^1$, define function H_Δ:

$$H_\Delta(p_1) = \begin{cases} H(p_1) & \text{if } p_1 \in [(1+\Delta)p_1^{\min}, (1-\Delta)p_1^{\max}], \\ H\left(p_1^{\min}\right) & \text{else if } p_1 < (1+\Delta)p_1^{\min}, \\ H\left(p_1^{\max}\right) & \text{else if } p_1 > (1-\Delta)p_1^{\max}. \end{cases} \qquad \text{(A-18)}$$

Theorem 4 Suppose S is C^2. The Hamiltonian H_Δ is concave, C^2 in the interval $((1+\Delta)p_1^{\min}, (1-\Delta)p_1^{\max})$ and C^0 in R.

Proof Only need to consider H_Δ in the interval $((1+\Delta)p_1^{\min}, (1-\Delta)p_1^{\max})$. By

$$S(p_1, p_3) \;\; = \;\; S(p_1, H(p_1)) = 0, \qquad \text{(A-19)}$$

take the derivative twice with respect to p_1 in the above,

$$\begin{pmatrix} 1 \\ \frac{\partial H}{\partial p_1} \end{pmatrix}^T \begin{pmatrix} \frac{\partial^2 S}{\partial p_1 \partial p_1} & \frac{\partial^2 S}{\partial p_1 \partial p_3} \\ \frac{\partial^2 S}{\partial p_1 \partial p_3} & \frac{\partial^2 S}{\partial p_3 \partial p_3} \end{pmatrix} \begin{pmatrix} 1 \\ \frac{\partial H}{\partial p_1} \end{pmatrix} + \frac{\partial S}{\partial p_3}\frac{\partial^2 H}{\partial p_1 \partial p_1} \;\; = \;\; 0. \qquad \text{(A-20)}$$

The Hessian matrix in the above is positive definite by S strictly convex and $\frac{\partial S}{\partial p_3} > 0$, so we have $\frac{\partial^2 H}{\partial p_1 \partial p_1} < 0$. It follows that H is concave in $((1+\Delta)p_1^{\min}, (1-\Delta)p_1^{\max})$. H_Δ is constant outside $((1+\Delta)p_1^{\min}, (1-\Delta)p_1^{\max})$, so it is concave.

Theorem 5 3D case. Given $\Delta > 0$, for any radial plane $p_2 = kp_1$ cutting through slowness surface S, there exist two intersection points $(p_1^{\min}, kp_1^{\min}) = (p_1^{\min}(k), kp_1^{\min}(k))$ and $(p_1^{\max}, kp_1^{\max}) = (p_1^{\max}(k), kp_1^{\max}(k))$ such that

$$S(p_1^{\min}, kp_1^{\min}, H(p_1^{\min}, kp_1^{\min})) \;\; = \;\; 0, \qquad \text{(A-21)}$$

$$\frac{1}{|\nabla S|}\frac{\partial S}{\partial p_3}(p_1^{\min}, kp_1^{\min}, H(p_1^{\min}, kp_1^{\min})) \;\; = \;\; 0, \qquad \text{(A-22)}$$

$$S(p_1^{\max}, kp_1^{\max}, H(p_1^{\max}, kp_1^{\max})) \;\; = \;\; 0, \qquad \text{(A-23)}$$

$$\frac{1}{|\nabla S|}\frac{\partial S}{\partial p_3}(p_1^{\max}, kp_1^{\max}, H(p_1^{\max}, kp_1^{\max})) \;\; = \;\; 0. \qquad \text{(A-24)}$$

For $p_1 \in [(1+\Delta)p_1^{\min}(k), (1-\Delta)p_1^{\max}(k)]$ and $p_2 = kp_1$,

$$\frac{1}{|\nabla S|}\frac{\partial S}{\partial p_3}(p_1, p_2, H(p_1, p_2)) \;\; \geq \;\; O(\Delta) > 0. \qquad \text{(A-25)}$$

Proof When $k = 0$ or $k = \infty$, the Theorem reduces to the 2D case, which is proved in Theorem 3. Otherwise we can rotate the axes of the coordinate by orthogonal transformation

$$M = \begin{pmatrix} \frac{1}{\sqrt{1+k^2}} & \frac{k}{\sqrt{1+k^2}} & 0 \\ -\frac{k}{\sqrt{1+k^2}} & \frac{1}{\sqrt{1+k^2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{A-26}$$

such that the new $op_1^{'}$ lies in the plane $p_2 = kp_1$ and $op_2^{'}$ is perpendicular to it. Thus we have reduced the 3D problem to a 2D problem in $op_1^{'}p_3$ coordinate plane. By Theorem 3, there exists $p_1^{'\min}$ and $p_1^{'\max}$ such that

$$S'(p_1^{'\min}, H'(p_1^{'\min})) = 0, \qquad \frac{1}{|\nabla S'|}\frac{\partial S'}{\partial p_3}(p_1^{'\min}, H'(p_1^{'\min})) = 0, \tag{A-27}$$

$$S'(p_1^{'\max}, H'(p_1^{'\max})) = 0, \qquad \frac{1}{|\nabla S'|}\frac{\partial S'}{\partial p_3}(p_1^{'\max}, H'(p_1^{'\max})) = 0, \tag{A-28}$$

where $S^{'}$ and $H^{'}$ are intersections of the slowness surface S and Hamiltonian H with $op_1'p_3$ plane (or radial plane $p_2 = kp_1$), respectively.

Perform the inverse transform to pull the above result back to original coordinates, we have an interval $[p_1^{\min}, p_1^{\max}]$, where $p_1^{\min} = \frac{1}{\sqrt{1+k^2}}p_1^{'\min}$ and $p_1^{\max} = \frac{1}{\sqrt{1+k^2}}p_1^{'\max}$, such that for $p_2 = kp_1$ **Theorem 5** holds.

It is convenient to parameterize the horizontal plane (p_1, p_2) by polar coordinates

$$(p_1, p_2) = (p\cos\phi, p\sin\phi), \tag{A-29}$$

we can summarize **Theorem 5** into a Corollary.

Corollary 1 For each planar angle ϕ, the family of planes perpendicular to $(\cos\phi, \sin\phi, 0)$ is tangent to the slowness surface at exactly one point which is $(p_{\max}(\phi)\cos\phi, p_{\max}(\phi)\sin\phi, p_3(\phi))$ and satisfies that

$$S(p_{\max}(\phi)\cos\phi, p_{\max}(\phi)\sin\phi, p_3(\phi)) = 0, \tag{A-30}$$

$$\frac{1}{|\nabla S|}\frac{\partial S}{\partial p_3}(p_{\max}(\phi)\cos\phi, p_{\max}(\phi)\sin\phi, p_3(\phi)) = 0. \tag{A-31}$$

For $0 < \Delta < 1$, $p \le (1-\Delta)p_{\max}(\phi)$ and $(p_1, p_2) = (p\cos\phi, p\sin\phi)$, we have

$$\frac{1}{|\nabla S|}\frac{\partial S}{\partial p_3}(p_1, p_2, H(p_1, p_2)) \ge O(\Delta) > 0. \tag{A-32}$$

Remark By method of characteristics, we can see that **Theorem 5** just puts a positive lower bound for the third component of normalized ray direction, namely, the group angle of the ray with the vertical axis never goes near 90 degrees.

Definition 3 3D Hamiltonian H_Δ. Given $\Delta > 0$, for any $(p_1, p_2) = (p\cos\phi, p\sin\phi)$, there exists $p_{\max}(\phi)$ as in **Corollary 1**, we define function H_Δ:

$$H_\Delta(p_1,p_2) = \begin{cases} H(p_1,p_2), \text{ if } p \le (1-\Delta)p_{\max}(\phi); \\ H((1-\Delta)p_{\max}(\phi)\cos\phi, (1-\Delta)p_{\max}(\phi)\sin\phi), \text{ else.} \end{cases} \tag{A-33}$$

Corollary 2 Under the assumption of **Corollary 1**, the inequality

$$\sqrt{\left(\frac{\partial H_\Delta}{\partial p_1}\right)^2 + \left(\frac{\partial H_\Delta}{\partial p_2}\right)^2} \le O(\frac{1}{\Delta}) \tag{A-34}$$

holds.

Proof By using ray equations (5), (8) and (9), we have

$$\frac{\partial S}{\partial p_1} = -\frac{\frac{\partial H}{\partial p_1}}{H - p_1\frac{\partial H}{\partial p_1} - p_2\frac{\partial H}{\partial p_2}}, \tag{A-35}$$

$$\frac{\partial S}{\partial p_2} = -\frac{\frac{\partial H}{\partial p_2}}{H - p_1\frac{\partial H}{\partial p_1} - p_2\frac{\partial H}{\partial p_2}}, \tag{A-36}$$

$$\frac{\partial S}{\partial p_3} = \frac{1}{H - p_1\frac{\partial H}{\partial p_1} - p_2\frac{\partial H}{\partial p_2}}. \tag{A-37}$$

The inequality follows by using **Corollary 1** .

Appendix B. A FIRST ORDER UPWIND SCHEME

To compute the first arrival traveltime field by a grid-based finite difference scheme, we derive a first order scheme from ray tracing rather than directly from the eikonal equation. Nonetheless, the end result will be recognizable as a difference approximation to the eikonal equation. Because our use of ray tracing in the derivation inherently honors causality, the resulting difference scheme is upwind, a term more or less synonymous with causality in this setting. We consider the two dimensional case to illustrate the idea, which means all out-the-plane components vanish in our equations formulated for 3-D media and all rays stay in the x_1x_3 plane. The three dimensional case is similar.

The first order finite difference approximation to the qP eikonal equation rests on several approximations; that is, we may approximate locally near a grid point

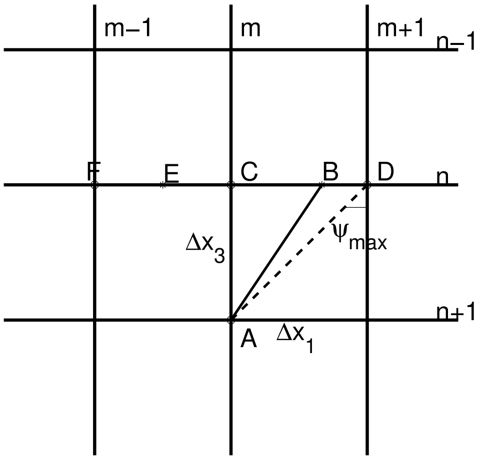


FIG. 8. Cartesian grid : a possible ray hitting point A

1. the group velocity v_g by a constant;
2. the rays by straight line segments;
3. the traveltime by piecewise linear functions of x_1 with nodes at the grid ponit.

Suppose we are given a Cartesian grid $(m\Delta x_1, n\Delta x_3)$ in two-dimensional space, Figure 8. Let

$$\tau_m^n \approx \tau(m\Delta x_1, n\Delta x_3), \tag{B-1}$$

be the grid function approximating the traveltime. Near $(m\Delta x_1, n\Delta x_3)$, we define the approximation of derivative τ_{x_1} at $(x_1, n\Delta x_3)$ by

$$\tau_{x_1}^n \approx \begin{cases} D_{x_1}^- \tau_m^n \equiv \dfrac{\tau_m^n - \tau_{m-1}^n}{\Delta x_1} & \text{for } (m-1)\Delta x_1 < x_1 < m\Delta x_1 \ , \\ D_{x_1}^+ \tau_m^n \equiv \dfrac{\tau_{m+1}^n - \tau_m^n}{\Delta x_1} & \text{for } m\Delta x_1 < x_1 < (m+1)\Delta x_1, \end{cases} \tag{B-2}$$

and the group velocity by

$$v_g^{\pm} = \frac{\sqrt{\left(\frac{\partial H}{\partial p_1}(D_{x_1}^{\pm}\tau_m^n)\right)^2 + 1}}{H(D_{x_1}^{\pm}\tau_m^n) - \frac{\partial H}{\partial p_1}(D_{x_1}^{\pm}\tau_m^n)D_{x_1}^{\pm}\tau_m^n}. \tag{B-3}$$

A ray segment meeting the line $x_3 = n\Delta x_3$ at x_1 and passing through $(m\Delta x_1, (n+1)\Delta x_3)$ has length

$$L = \Delta x_3 \sqrt{1 + \left(\frac{x_1 - m\Delta x_1}{\Delta x_3}\right)^2}, \tag{B-4}$$

so that the time predicted at $(m\Delta x_1, (n+1)\Delta x_3)$ under the foregoing assumption is

$$\tau_m^{n+1} \approx \begin{cases} \tau_m^n + (x_1 - m\Delta x_1) D_{x_1}^+ \tau_m^n + \dfrac{L}{v_g^+} & \text{for } m\Delta x_1 \le x_1 \le (m+1)\Delta x_1 \ , \\ \tau_m^n + (x_1 - m\Delta x_1) D_{x_1}^- \tau_m^n + \dfrac{L}{v_g^-} & \text{for } (m-1)\Delta x_1 \le x_1 \le m\Delta x_1 . \end{cases} \tag{B-5}$$

Finally, the ray velocity vector should obey the ray equation (8),

$$-\frac{x_1 - m\Delta x_1}{\Delta x_3} = \frac{dx_1}{dx_3} \approx \begin{cases} -\dfrac{\partial H}{\partial p_1}(D_{x_1}^+ \tau_m^n) & \text{for } m\Delta x_1 \le x_1 \le (m+1)\Delta x_1, \\ -\dfrac{\partial H}{\partial p_1}(D_{x_1}^- \tau_m^n) & \text{for } (m-1)\Delta x_1 \le x_1 \le m\Delta x_1, \end{cases} \tag{B-6}$$

i.e.,

$$x_1 - m\Delta x_1 \approx \begin{cases} \Delta x_3 \dfrac{\partial H}{\partial p_1}(D_{x_1}^+ \tau_m^n) & \text{for } m\Delta x_1 \le x_1 \le (m+1)\Delta x_1 \ , \\ \Delta x_3 \dfrac{\partial H}{\partial p_1}(D_{x_1}^- \tau_m^n) & \text{for } (m-1)\Delta x_1 \le x_1 \le m\Delta x_1 . \end{cases} \tag{B-7}$$

Substituting expression (B-7) into (B-5) and carrying out the algebra, we arrive at

$$\tau_m^{n+1} = \begin{cases} \tau_m^n + \Delta x_3 H \left(D_{x_1}^+ \tau_m^n \right) & \text{for } m\Delta x_1 \le x_1 \le (m+1)\Delta x_1 \ , \\ \tau_m^n + \Delta x_3 H \left(D_{x_1}^- \tau_m^n \right) & \text{for } (m-1)\Delta x_1 \le x_1 \le m\Delta x_1 , \end{cases} \tag{B-8}$$

where x_1 is chosen to satisfy (B-6) if possible.

Next we have to examine the significance of the condition switching the branches of (B-8). If $m\Delta x_1 \le x_1 \le (m+1)\Delta x_1$; i.e., the ray slope is nonpositive, we see from (B-6) that $\dfrac{\partial H}{\partial p_1}(D_{x_1}^+ \tau_m^n) \ge 0$. Similarly, the second branch occurs when $\dfrac{\partial H}{\partial p_1}(D_{x_1}^- \tau_m^n) \le 0$. When neither of these occur; i.e., $\dfrac{\partial H}{\partial p_1}(D_{x_1}^+ \tau_m^n) \le 0 \le \dfrac{\partial H}{\partial p_1}(D_{x_1}^- \tau_m^n)$, rays fan away from $(m\Delta x_1, n\Delta x_3)$; that is, this point is a center of rarefaction. Then to good approximation, the ray entering $(m\Delta x_1, (n+1)\Delta x_3)$ is vertical; i.e., along it $\frac{\partial H}{\partial p_1} = 0$. Finally, if both conditions are satisfied; i.e., $\dfrac{\partial H}{\partial p_1}(D_{x_1}^+ \tau_m^n) \ge 0 \ge \dfrac{\partial H}{\partial p_1}(D_{x_1}^- \tau_m^n)$, then two rays converge on $(m\Delta x_1, (n+1)\Delta x_3)$, and we should select the least of the two times provided by (B-8).

Because the Hamiltonian H is concave, all four of these options can be combined in the simple formula

$$\tau_m^{n+1} = \tau_m^n + \Delta x_3 H \left(\mathrm{maxmod}(\max(D_{x_1}^- \tau_m^n, 0), \min(D_{x_1}^+ \tau_m^n, 0)) \right), \tag{B-9}$$

where the function maxmod returns the larger value in modulus.

Consulting Osher and Sethian (Osher and Sethian, 1988) and Osher and Shu (Osher and Shu, 1991), we see that the scheme just proposed is identical to the so-called Godunov first order scheme. The problem we have solved (in an approximate but a completely elementary way !) to obtain this scheme is the so-called Riemann problem for the eikonal equation, in which the initial traveltime data is piecewise linear.

There remains one further detail to take care of: the approximate ray might meet $x_3 = n\Delta x_3$ **outside** of the interval $(m-1)\Delta x_1 \leq x_1 \leq (m+1)\Delta x_1$. The difference scheme (B-9) would necessarily become unstable, as the numerical domain of dependence of $(m\Delta x_1, (n+1)\Delta x_3)$ (namely, the three points on $x_3 = n\Delta x_3$ in the scheme (B-9)) would no longer contain in its convex hull the continuum domain of dependence $\{x\}$. In order that this Courant-Friedriches-Lewy ("CFL") criterion be satisfied with surely, we limit the slope of ray:

$$\max\left\{\left|\frac{\partial H}{\partial p_1}\left(D^+_{x_1}\tau^n_m\right)\right|, \left|\frac{\partial H}{\partial p_1}\left(D^-_{x_1}\tau^n_m\right)\right|\right\} \quad \leq \quad \frac{\Delta x_1}{\Delta x_3}. \tag{B-10}$$

By definition of H_Δ (Appendix A), we just need to choose $\Delta > 0$ such that

$$\frac{\Delta x_1}{\Delta x_3} \quad \geq \quad O(\frac{1}{\Delta}), \tag{B-11}$$

and then we can find $p_{\max}$ needed in the defintion of H_Δ (see Definition 3 in Appendix A). So the resulting difference scheme is

$$\tau^{n+1}_m \quad = \quad \tau^n_m + \Delta x_3 H_\Delta\left(\mathrm{maxmod}(\max(D^-_{x_1}\tau^n_m, 0), \min(D^+_{x_1}\tau^n_m, 0))\right), \tag{B-12}$$

which is consistent with the **paraxial eikonal equation**

$$p_3 \quad = \quad H_\Delta\left(p_1\right). \tag{B-13}$$

Solutions of (B-13) are identical to solutions of the eikonal equation at points whose associated rays make angle $\leq \psi_g^{\max} = \arctan(\frac{\Delta x_1}{\Delta x_3})$ with the vertical ($x_1 =$ *constant*) throughout their length. (For point source problem, the associated ray is the first arriving ray connecting the subsurface point with the source point). Thus with this paraxial limitation, the scheme is suitable for computing traveltime of **downgoing ray fans**.

Appendix C. INITIALIZATION OF TRAVELTIME IN VTI

Assume the VTI medium is homogeneous, we only need to consider 2D case. Given source point (x_1^s, x_3^s) and target point (x_1^o, x_3^o), we try to find the traveltime between these two points. Suppose $x_1^o - x_1^s > 0$ and $x_3^o - x_3^s > 0$, the group angle ϕ

at (x_1^o, x_3^o) satisfies that $\tan\phi = \dfrac{x_3^o - x_3^s}{x_1^o - x_1^s}$ and $\phi = \phi(\theta)$ where θ is the phase angle at (x_1^o, x_3^o). We find the phase angle by a bisection iterative method, and we compute the group velocity from the raytracing equation. Finally the traveltime follows from the group velocity.

The simplified algorithm framework is as follows:

- Input: (x_1^s, x_3^s), (x_1^o, x_3^o) and $\alpha_0, \beta_0, \epsilon, \delta$.
- Set $x_1 = x_1^o - x_1^s$, $x_3 = x_3^o - x_3^s$, $\vec{n}_1 = (1,0)^T$, $\vec{n}_2 = (0,1)^T$;
- Compute $\vec{n} = \frac{1}{2}(\vec{n}_1 + \vec{n}_2)$; $\vec{n} = \dfrac{\vec{n}}{|\vec{n}|}$.
- do
  - Compute $(v_g^1, v_g^3)^T = \vec{F}(\vec{n}, \alpha_0, \beta_0, \epsilon, \delta)$.
  - if $\dfrac{v_g^3}{v_g^1} \geq \dfrac{x_3}{x_1}$, $\vec{n}_2 = \vec{n}$; else $\vec{n}_1 = \vec{n}$;
  - Compute $\vec{n} = \frac{1}{2}(\vec{n}_1 + \vec{n}_2)$; $\vec{n} = \dfrac{\vec{n}}{|\vec{n}|}$.
- until $(n_1(2) \geq n(2)$ or $n(2) \geq n_2(2))$
- Compute traveltime $t = \dfrac{\sqrt{x_1^2 + x_3^2}}{\sqrt{(v_g^1)^2 + (v_g^3)^2}}$.

In the above algorithm, we use the raytracing equation ($\vec{F}$ denotes the righthand side):

$$v_g^1 = p_1(2\alpha_0^2(1+2\epsilon)\beta_0^2 p_1^2 + A p_3^2 - \alpha_0^2(1+2\epsilon) - \beta_0^2)/D, \tag{C-1}$$

$$v_g^3 = p_3(2\alpha_0^2\beta_0^2 p_3^2 + A p_1^2 - \alpha_0^2 - \beta_0^2)/D, \tag{C-2}$$

with

$$A = 2\alpha_0^2(\alpha_0^2(\epsilon - \delta) + \beta_0^2(1+\delta)), \tag{C-3}$$
$$D = 2 - (\alpha_0^2(1+\epsilon) + \beta_0^2)p_1^2 - (\alpha_0^2 + \beta_0^2)p_3^2; \tag{C-4}$$

p_1, p_3 can be computed from the phase velocity and phase normal $\vec{n}$.