JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 93, NO. B6, PAGES 6585–6599, JUNE 10, 1988

Ray Tracing For Earthquake Location In Laterally Heterogeneous Media

JEAN VIRIEUX, VERONIQUE FARRA[1] AND RAUL MADARIAGA

Laboratoire de Sismologie
Université Paris 7 and Institut de Physique du Globe de Paris, France

We propose a technique for the fast solution of ray tracing in three-dimensional laterally heterogeneous media. Analytical solutions for both ray tracing and paraxial ray tracing for a medium with an arbitrary gradient of the square of slowness are used to develop a systematic procedure to construct exact solutions. Complex heterogeneous media are divided into tetrahedral finite elements inside which the square of the slowness has a simple linear distribution. Arbitrarily complex media may be studied by judicious choice of the elements. We develop appropriate boundary conditions for interfaces with zeroth or first-order velocity discontinuities. Two-point ray tracing is performed by a Newton method based on paraxial ray theory. With a few iterations the ray trajectories through the source and the observer are calculated with the same overall speed as with ray bending. The paraxial method has the additional advantage that it can be used to identify caustics and to separate travel time branches. We used the previously described method to replace the ray tracing algorithm of the HYPO71 program, probably the most commonly used earthquake location program. The ray tracing routine may be easily modified to adapt it to other earthquake location environments. With the modified HYPO71 program, we relocated the aftershock sequence of the November 23, 1980, Irpinia (Italy) earthquake. The introduction of a low-velocity zone proposed by different authors shifts aftershock locations to the NE. This horizontal displacement of hypocenters is not very sensitive to the thickness of the low-velocity zone.

INTRODUCTION

Earthquake location in three-dimensional laterally heterogeneous structures requires the accurate and fast solution of two-point ray tracing. Most standard ray tracing techniques are just too slow or inaccurate for routine use in earthquake location programs. For this reason, several alternative approximate methods have been proposed in which evaluation of ray trajectories is less accurate than that of travel times. Most of these techniques are based on ray bending, which is an iterative method for the calculation of travel times and trajectories that starts from a hypothetical trajectory and iterates it until Fermat's principle is satisfied. Recently, *Um and Thurber* [1987] and *Prothero et al.* [1987], have proposed to calculate only travel times by a sort of Galerkin method. These methods have the unsatisfactory feature that they converge to only one of the possible ray trajectories when there are multiple arrivals, a very frequent occurrence in three dimensions. In this paper we will explore a full ray tracing technique that is akin to finite elements in that the medium is divided into a set of tetrahedra with a simple slowness distribution law. In each tetrahedron, ray tracing is performed exactly.

The methods that will be proposed here draw extensively from recent work on ray methods for the calculation of synthetic seismograms. These are the Maslov method [*Maslov*, 1965; *Chapman and Drummond*, 1982; *Thomson and Chapman*, 1985] and Gaussian Beam Summation as proposed by *Popov* [1982] and *Červený et al.* [1982]. In exploring the foundations of these methods, *Madariaga* [1984], *Klimeš* [1984], *Babich et al.* [1985], and *Farra and Madariaga* [1987], showed that these techniques are in fact based on the use of paraxial ray theory. This is a method to calculate the propagation of rays and beams in the vicinity of a ray that is used as a reference. Paraxial methods were introduced in seismology for the calculation of geometrical spreading by *Popov and Pšenčik* [1978], who called their method dynamic ray tracing. *Červený and Pšenčik* [1984] and *Beydoun and Keho* [1986], among others, have used paraxial ray theory for the calculation of synthetics in two-dimensional media.

As shown by *Chapman* [1985] in Cartesian coordinates and by *Farra and Madariaga* [1987] in general coordinates, paraxial ray tracing is based on the application of perturbation theory to ray tracing. In order to fully exploit paraxial ray theory it is convenient to use Hamilton's methods. In this paper we use Hamiltonian techniques in order to develop an efficient three-dimensional ray tracing method that is suitable for use in a general earthquake location package. The most time consuming aspect of this problem is the need to perform two-point tracing, i.e., to find a ray that passes through both the source and the observer and to perturb the source position. We propose to divide the medium into a set of tetrahedral elements inside which the square of the slowness has a simpler linear distribution. Both ray tracing and paraxial ray tracing may be performed analytically in these elements so that ray tracing reduces to connecting analytical solutions at the sides of the tetrahedrons. A linear velocity gradient also leads to simple analytical ray tracing [*Chapman*, 1985] but we believe that our solutions are simpler to trace. Paraxial ray tracing is used in order to connect perturbations at the source to those at the observer. In this fashion we can move the source or the ray endpoints as required by the earthquake location programs.

Because of its wide applicability throughout the geophy-

[1] Also at Institut Français du Pétrole, Rueil-Malmaison, France.

Paper number 7B6080.
0148-0227/88/007B-6080$05.00

sics community we have chosen to modify the HYPO71 program so that it can locate earthquakes in a laterally heterogeneous medium. This program was originally written by *Lee and Lahr* [1975] for earthquake location in vertically layered media. With very little change in the preparation of data, our program allows for lateral heterogeneity. The method is used to relocate a group of selected aftershocks of the Irpinia earthquake of 1980 in southern Italy [*Deschamps and King*, 1984]. Following a proposition by P. Bernard and A. Zollo (The Irpinia (Italy) 1980 earthquake: Detailed analysis of a complet normal fault, submitted to *Journal of Geophysical Research*, 1988; hereafter referred to as BZ88), we introduced a shallow low-velocity zone to the NE of the main fault. The effect upon locations is quite important. For instance, the main aftershocks activity moves from the foot wall toward the hanging wall.

A Simple Hamiltonian Formulation Of Ray Tracing

Ray Tracing

Let us consider the scalar equation of wave motion in a heterogeneous medium :

$$\nabla^2\phi - \frac{1}{v^2}\frac{\partial^2\phi}{\partial t^2} = 0 \tag{1}$$

where $\phi(\mathbf{x},t)$ is some scalar wave field. As shown by *Červený et al.* [1977], ray tracing in elastic media does not differ substantially from that in acoustics, so that we will only discuss the simpler equation (1). When the velocity $v(\mathbf{x})$ is slowly varying inside the medium, we seek an asymptotic solution of the form :

$$\phi(\mathbf{x},t) = s(\omega)A(\mathbf{x})e^{-i\omega[t-\theta(\mathbf{x})]} \tag{2}$$

which is the approximation of geometrical optics. $A(\mathbf{x})$ is the amplitude, $\theta(\mathbf{x})$ the eikonal or travel time function, ω is the frequency, and $s(\omega)$ is the source time function. $A(\mathbf{x})$ is the first term of the expansion of amplitude $A(\mathbf{x},\omega)$ with respect to the inverse powers of ω (see *Červený* [1985], for example). Introducing this ansats in (1) and collecting terms of the same order in ω, we get

$$(\nabla\theta)^2 = u^2 = v^{-2} \tag{3}$$

known as the eikonal equation (u is slowness), and a transport equation

$$A\nabla\theta + 2\nabla A.\nabla\theta = 0 \tag{4}$$

for the amplitude.

In order to perform ray tracing we introduce the slowness vector $\mathbf{p} = \nabla\theta$, which is perpendicular to the surfaces of equal phase θ, or wave fronts. Let s be the arc length and $\mathbf{x}(s)$ the position along a ray. Rays are tangent to the slowness vector, so that this vector can be written as $\mathbf{p} = u\,d\mathbf{x}/ds$. Introducing, following *Thomson and Chapman* [1985], a ray parameter τ by the relationship $u\,d\tau = ds$ we get $\mathbf{p} = \dot{\mathbf{x}}$, where the dot denotes derivation with respect to τ. Parameter τ plays a role that is similar to time in analytical mechanics, but it does not have any particular physical meaning in ray theory. Introducing the Hamiltonian proposed by *Burridge* [1976],

$$H(\mathbf{x},\mathbf{p};\tau) = \frac{1}{2}[\mathbf{p}^2 - u^2(\mathbf{x})] \tag{5}$$

we observe that the eikonal equation implies that $H \equiv 0$ along a ray. Using Hamilton's canonical equations, we find the ray tracing equations

$$\begin{aligned}\dot{\mathbf{x}} &= \nabla_{\mathbf{p}}H = \mathbf{p}\\ \dot{\mathbf{p}} &= -\nabla_{\mathbf{x}}H = \frac{1}{2}\nabla_{\mathbf{x}}u^2 = u\nabla_{\mathbf{x}}u\end{aligned} \tag{6}$$

where $\nabla_{\mathbf{x}}$ or $\nabla_{\mathbf{p}}$ denote the gradient with respect to vectors $\mathbf{x}$ and $\mathbf{p}$, respectively. Ray tracing consists in solving the nonlinear system (6) for $\mathbf{x}(\tau)$, $\mathbf{p}(\tau)$ for given initial conditions (shooting) or boundary value conditions (the two-point boundary value problem).

It is important to note that the six equations of the system (6) are not really independent. In the first place, the modulus of the slowness should satisfy

$$\|\mathbf{p}\| = u \tag{7}$$

as implied by the eikonal equation. This reduces the system (6) by one equation. A further reduction of the system (6) is possible because the ray parameter τ may be determined implicitly from the position vector $\mathbf{x}(\tau)$. The proper way to reduce the system (6) is to use a "reduced" Hamiltonian (see *Farra and Madariaga*, 1987, for further details). This reduction of variables, which is very attractive for computation, is at the expense of a more complex structure of the Hamiltonian, and the associated phase space which is now curved and requires the introduction of scale factors. Because we are interested in obtaining analytical solutions, we will keep the simpler ray tracing system (6), but we should remember that the equations are not really independent.

Paraxial Ray Tracing

Suppose a ray has been traced and that we need to trace another one in its vicinity. Can we benefit from the already traced ray? The paraxial approximation provides a systematic approach to this problem [*Luneberg*, 1944; *Červený et al.*, 1982]. Around the already traced ray, called the central ray, we can trace other neighboring rays by means of time dependent perturbation theory as explained by *Farra and Madariaga* [1987].

Let us denote the position of the central ray $\mathbf{x}_c(\tau)$ and its slowness vector $\mathbf{p}_c(\tau)$. The position of the paraxial ray and its slowness vector are given by

$$\mathbf{x}(\tau) = \mathbf{x}_c(\tau) + \delta\mathbf{x}(\tau) \qquad \mathbf{p}(\tau) = \mathbf{p}_c(\tau) + \delta\mathbf{p}(\tau) \tag{8}$$

The perturbations of position and slowness vector $\mathbf{x}$ and $\mathbf{p}$ verify the paraxial ray tracing equations

$$\begin{bmatrix}\delta\dot{\mathbf{x}}\\ \delta\dot{\mathbf{p}}\end{bmatrix} = \begin{bmatrix}\nabla_{\mathbf{p}}\nabla_{\mathbf{x}}H & \nabla^2_{\mathbf{p}}H\\ -\nabla^2_{\mathbf{x}}H & -\nabla_{\mathbf{x}}\nabla_{\mathbf{p}}H\end{bmatrix}\begin{bmatrix}\delta\mathbf{x}\\ \delta\mathbf{p}\end{bmatrix} \tag{9}$$

where H and its derivatives are computed on the central ray at τ. This equation is obtained by first-order perturbation of the full ray tracing system (6). For our choice of the Hamiltonian (5), we deduce the simple linear system

$$\begin{bmatrix}\delta\dot{\mathbf{x}}\\ \delta\dot{\mathbf{p}}\end{bmatrix} = \begin{bmatrix}0 & \mathbf{I}\\ \mathbf{V} & 0\end{bmatrix}\begin{bmatrix}\delta\mathbf{x}\\ \delta\mathbf{p}\end{bmatrix} \tag{10}$$

where **I** is the identity matrix and **V** is a matrix of second order partial derivatives of the square of slowness defined by

$$V_{ij} = \frac{1}{2}\frac{\partial^2 u}{\partial x_i \partial x_j} \tag{11}$$

The structure of system (10) is the same as that obtained by *Thomson and Chapman* [1985] in Cartesian coordinates. The paraxial ray tracing system (10) is linear, so that any of its solutions may be written in terms of a propagator **P** :

$$\begin{bmatrix}\delta\mathbf{x}\\ \delta\mathbf{p}\end{bmatrix} = \mathbf{P}\begin{bmatrix}\delta\mathbf{x}_c\\ \delta\mathbf{p}_c\end{bmatrix} = \begin{bmatrix}Q_1 & Q_2\\ P_1 & P_2\end{bmatrix}\begin{bmatrix}\delta\mathbf{x}_c\\ \delta\mathbf{p}_c\end{bmatrix} \tag{12}$$

where we have adopted the now classical notation for the submatrices of the paraxial ray propagator [see *Červený*, 1985]. The initial ray vector is noted $[\delta\mathbf{x}_c, \delta\mathbf{p}_c]$. The propagator **P** may be built from a linear combination of elementary trajectories.

Just as with the full nonlinear ray tracing system (6), the six equations of the paraxial system (10) are not really independent; in other words, not every solution of the system (10) represents a paraxial ray trajectory. As with the full ray tracing system, a first additional condition that system (10) has to satisfy comes from the perturbation of the eikonal equation or Hamiltonian. To first order, when position and slowness vector are perturbed as in (8), the perturbation of the Hamiltonian (5) is given by

$$\delta H = \mathbf{p}_c \cdot \delta\mathbf{p} - u\nabla_{\mathbf{x}} u \cdot \delta\mathbf{x} = 0 \tag{13}$$

The perturbation of the Hamiltonian must be zero in order that a solution of (10) is a legitimate paraxial ray.

In order to find a valid solution of the system (10) we impose the additional condition that the paraxial rays belong to a certain family of rays (a beam) centered around the reference ray. Without this condition the paraxial rays would cross each other in random ways. This condition is then equivalent to that of continuity of the wave front. Following a notation introduced by *Popov* [1982], we require that for every value of τ

$$\delta\mathbf{p}(\tau) = \mathbf{M}(\tau)\delta\mathbf{x}(\tau) \tag{14}$$

where $\mathbf{M}(\tau)$ is a 3x3 matrix. With this relationship we can write the second-order expansion of the travel time around the central ray in the form

$$\theta(\mathbf{x},\tau) = \theta(\mathbf{x}_c) + \mathbf{p}_c \cdot \delta\mathbf{x} + \frac{1}{2}\delta\mathbf{x}^t\mathbf{M}(\tau)\delta\mathbf{x} \tag{15}$$

Condition (14) is a linear relation between the components of slowness and position perturbation vectors for a given value of τ. The matrices **M** for successive values of τ may be obtained from their initial value at the source

$$\mathbf{M}(0) = \epsilon^{-1} \tag{16}$$

where ϵ is a matrix that determines the initial shape of the ray beam (see *Červený* [1985] and *Farra and Madariaga* [1987] for further discussion). For the point sources considered in this paper, $\epsilon = \mathbf{0}$ the null matrix. With the conditions (13) and (14), the rank of the linear system (10) is reduced to four in a three-dimensional medium. The trajectory of a paraxial ray belonging to a certain ray pencil or beam depends only on four initial conditions.

Analytical Ray Tracing In A Simple Medium

We want to divide the medium into a set of finite elements with simple velocity distributions. After exploring different distributions of the velocity, the slowness or the square of the slowness, we concluded that a constant gradient of the square of the slowness gives the simplest analytical solutions both for ray tracing and paraxial ray tracing.

Consider the following linear distribution of the square of slowness:

$$u^2 = u_0^2 + \gamma \cdot \mathbf{x} = u_0^2 + \gamma_i \cdot x_i \tag{17}$$

where γ defines the gradient and the index i equal to 1 for x, 2 for y (noted also x_2) and 3 for z (noted also x_3). In a two-dimensional medium, the index 2 can be omitted. In order to perform ray tracing we replace the slowness distribution (17) into the ray tracing equations (6):

$$\frac{d\mathbf{x}}{d\tau} = \mathbf{p} \qquad \frac{d\mathbf{p}}{d\tau} = \frac{1}{2}\gamma \tag{18}$$

Solving (18), we find the simple expressions

$$\begin{aligned}\mathbf{p} &= \frac{1}{2}\gamma\tau + \mathbf{p}_0\\ \mathbf{x} &= \frac{1}{4}\gamma\tau^2 + \mathbf{p}_0\tau + \mathbf{x}_0\end{aligned} \tag{19}$$

where $\mathbf{x}_0$ and $\mathbf{p}_0$ are the initial position and slowness vector of the ray. As expected, (19) yields a parametric equation for the ray that depends only on its initial position and slowness vector. The parameter τ may be eliminated in order to find the actual equation of the ray, but it is preferable to use the parametric expression (19) in the computer implementation.

Finally, we calculate travel times. Recalling the relation $ds = u d\tau$, we find

$$\frac{d\theta}{d\tau} = u\frac{ds}{d\tau} = u^2 = \|\mathbf{p}\|^2 \tag{20}$$

which may be easily integrated using (19). The total travel time is obtained in the compact form

$$\theta = \mathbf{p}_0^2\tau + \frac{1}{2}\gamma \cdot \mathbf{p}_0\tau^2 + \frac{1}{12}\gamma^2\tau^3 \tag{21}$$

an expression which is equivalent to that obtained by Červený [1987].

Let us now solve the paraxial ray tracing. Because the square of the slowness has only linear terms, the matrix of second derivatives **V** defined in (11) is zero, and the solution of system (10) is straightforward

$$\delta\mathbf{x}(\tau) = \delta\mathbf{p}_0\tau + \delta\mathbf{x}_0 \qquad \delta\mathbf{p} = \delta\mathbf{p}_0 \tag{22}$$

In this medium, matrix $\mathbf{M}(\tau)$ has the simple expression

$$\mathbf{M}(\tau) = (\tau + \epsilon)^{-1} \tag{23}$$

so that travel times in the vicinity of the central ray are calculated immediately from (15). The solution of (22) has to satisfy the additional condition (13) which for the slowness distribution (17) becomes

$$\mathbf{p}_c \cdot \delta\mathbf{p} = \frac{1}{2}\gamma \cdot \delta\mathbf{x} \tag{24}$$

Thus there are only four independent perturbation parameters to choose.

Both ray tracing and paraxial ray tracing in media with constant gradient of the square of the slowness are thus extremely simple in Cartesian coordinates, and this is the reason we will use it in our ray tracing method.

Internal Boundary Conditions For Ray Tracing

The presence of interfaces with zero-order discontinuity of the square of the slowness, as well as boundaries with first-order discontinuity between the elements of the discrete medium (see the next section), requires the introduction of appropriate boundary conditions for ray and paraxial ray tracing. The phase matching method [*Deschamps*,1972; *Červený*,1985] may be used to obtain the boundary conditions. These may be obtained expanding the travel time around the central ray as in (15) and then imposing the continuity of the wave fronts across the boundary. The zeroth order term of the phase expansion around the central ray implies the continuity of the travel time along the central ray. The first-order term in $\delta\mathbf{x}$ gives Snell's law for the central ray, while the continuity of the second-order term yields the boundary condition for the paraxial ray vector $[\delta\mathbf{x}\ \delta\mathbf{p}]$. Alternatively, this boundary condition may be obtained by perturbation of Snell's law [*Chapman*,1985; *Farra*,1987]. In the following we present the formulation of *Deschamps* [1972] because it has a very simple geometrical interpretation. In order to avoid complex notation, the two-dimensional case will be described, but the three-dimensional case can be simply deduced because we are working in Cartesian coordinates.

Let us consider an already traced central ray that intersects an interface at point O. In the vicinity of this point O, the interface can be approximated by the quadratic equation

$$\mathbf{r} = \ell\,\mathbf{e} - \frac{1}{2}C\,\ell^2\,\mathbf{n} \tag{25}$$

where ℓ is the distance from the point O along the tangent to the interface, $\mathbf{r}$ is a point on the interface, $\mathbf{e}$ is the unit tangent vector, and $\mathbf{n}$ is the unit vector along the normal, as shown by Figure 1. A local Cartesian coordinate frame is defined at point O with axes X along the tangent and Z along the normal. These coordinates will be written with capital letters in order to distinguish them from the global Cartesian reference frame.

We consider now a paraxial ray of the incident ray. Its perturbation in position and slowness vector expressed in local coordinates is $[\delta X, \delta Z, \delta p_X, \delta p_Z]$, while that of the corresponding transmitted/reflected paraxial ray is $[\delta X_t, \delta Z_t, \delta p_{t_X}, \delta p_{t_Z}]$. The problem is to find the latter in terms of the incident paraxial ray coordinates. Following Figure 1, let the paraxial ray intersect the interface at a point R. Its projection into the tangent to the interface is M, as

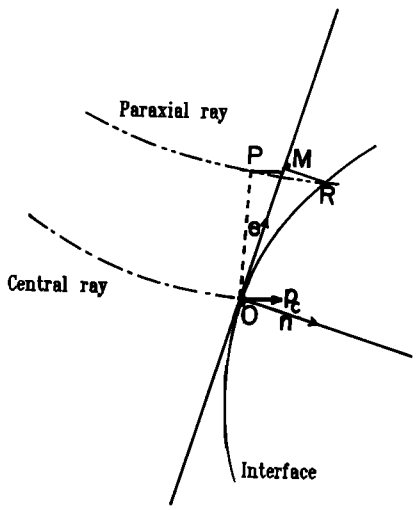

Fig. 1. Geometry of the linear transformation at the curved interface between two heterogeneous media. The central ray intersects the interface at O, while the paraxial ray arrives at R. A local coordinate system with base vectors $\mathbf{n}$, $\mathbf{e}$ is defined at the interface.

shown in Figure 1. To first order, this point M is also the projection of P, defined by OP equal to $\delta\mathbf{x}$, parallel to the slowness vector $\mathbf{p}$ at point O. This particular projection is the first-order continuation of (6) at point P. The local Cartesian coordinates of M are

$$\begin{aligned} \delta X + p_X(O)\,\delta\tau &= \delta X_t = \ell \\ \delta Z + p_Z(O)\,\delta\tau &= \delta Z_t = 0 \end{aligned} \tag{26}$$

where $\delta\tau$ is the increase or decrease of the ray parameter τ between points P and M. The new initial conditions for the transmitted/reflected paraxial ray are obtained from (26). In (26) we set $\delta Z_t = 0$ so that vector $\delta\mathbf{x}_t$ is taken along the tangent to the interface. Eliminating $\delta\tau$ from the two equations in (26), we find $\delta X_t = \ell$ in terms of δX and δZ.

Finding the boundary conditions for $\delta\mathbf{p}$ is slightly more involved. Following *Deschamps* [1972], we express the continuity of the differential phase $d\theta = \mathbf{p}(R) \cdot d\mathbf{r}$, across the interface. Linearizing (25), we get

$$d\mathbf{r} = d\ell\,\mathbf{e} - C\ell d\ell\,\mathbf{n} \tag{27}$$

The slowness vector at R is related to that at P by the first-order Taylor expansion

$$\mathbf{p}(R) \simeq \mathbf{p}(M) = \mathbf{p}(P) + \frac{\partial\mathbf{p}}{\partial\tau}(O)\,\delta\tau$$

so that using definition (8), we get

$$\mathbf{p}(R) = \mathbf{p}_c + \delta\mathbf{p} + \frac{\partial\mathbf{p}}{\partial\tau}(O)\delta\tau \tag{28}$$

Finally, the differential phase can be split in two terms

$$d\theta = d\ell\,\mathbf{p}_c \cdot \mathbf{e} + \chi d\ell \tag{29}$$

The first term is the variation of the travel time around point O of Figure (1), while the second includes effects from both the paraxial ray and the interface curvature. The latter is given by

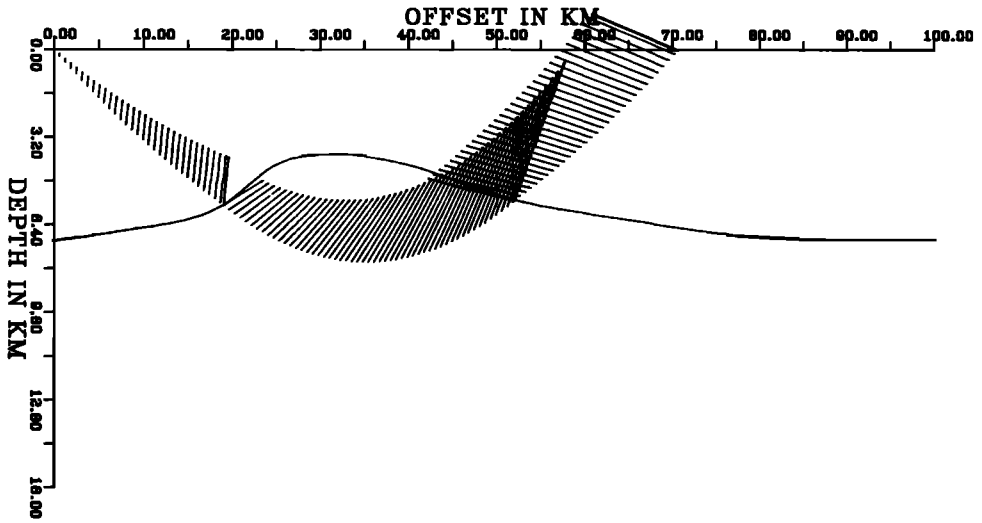

Fig. 2. Illustration of paraxial rays and their transformation at the crossing of an interface. A paraxial ray is obtained by drawing its vector $\delta\mathbf{x}$ from the central ray. The two gaps in the distribution of paraxial rays come from their transformations at the interface. A final transformation is also performed at the free surface in order to get a horizontal $\delta\mathbf{x}$.

$$\chi = \mathbf{e}\cdot\delta\mathbf{p} + \mathbf{e}\cdot\frac{\partial\mathbf{p}}{\partial\tau}(O)\,\delta\tau - C\ell\,\mathbf{p}_c\cdot\mathbf{n} \tag{30}$$

which using (26) may be written in local coordinates as

$$\chi = \delta p_X - C p_Z\,\delta X_t - \frac{u}{p_Z}\nabla_X u\,\delta Z \tag{31}$$

In homogeneous media it is easy to show that expression (30) reduces to that obtained by *Deschamps* [1972]. This expression is also equivalent to those obtained by *Červený* [1985] or by *Farra* [1987] by other means.

A similar expression may be obtained for the transmitted/reflected ray. The continuous quantity χ is given by

$$\chi = \delta p_{t_X} - C p_{t_Z}\,\delta X_t \tag{32}$$

Finally, let us impose condition (13) to the transmitted/reflected paraxial trajectory

$$dH_t = p_{t_X}\delta p_{t_X} + p_{t_Z}\delta p_{t_Z} - u_t\nabla_X u_t\,\delta X_t = 0 \tag{33}$$

where u_t is the slowness of the medium where the transmitted/reflected ray propagates.

Using (25) through (33), we find the following expression for the boundary conditions of a paraxial ray across the curved interface

$$\begin{aligned}
\delta X_t &:= \delta X - \frac{p_X}{p_Z}\delta Z \\
\delta Z_t &:= 0 \\
\delta p_{t_X} &:= \delta p_X + C(p_{t_Z} - p_Z)\delta X_t - \frac{u}{p_Z}\nabla_X u\,\delta Z \\
\delta p_{t_Z} &:= \frac{u_t}{p_{t_Z}}\nabla_X u_t\,\delta X_t - \frac{p_{t_X}}{p_{t_Z}}\,\delta p_{t_X}
\end{aligned} \tag{34}$$

The order of computation is very important in these expressions. This transformation $\mathbf{T}$ might be rewritten in matrix form in order to emphasize its linearity

$$\begin{bmatrix} \delta X_t \\ \delta Z_t \\ \delta p_{t_X} \\ \delta p_{t_Z} \end{bmatrix} = \begin{bmatrix} 1 & Q_{1_X Z} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ P_{1_X X} & P_{1_X Z} & 1 & 0 \\ P_{1_Z X} & P_{1_Z Z} & P_{2_Z X} & 0 \end{bmatrix} \begin{bmatrix} \delta X \\ \delta Z \\ \delta p_X \\ \delta p_X \end{bmatrix} \tag{35}$$

where the different elements of the matrix may be calculated from (34). Finally, using the rotation matrix $\mathbf{R}$ from global coordinates to local ones, we get the linear transformation for the transmitted/reflected paraxial ray $\mathbf{R}^{-1}\,\mathbf{T}\,\mathbf{R}$, which might be obtained easily for media without interfaces.

In order to test paraxial ray tracing, we considered two media with different velocity gradients and a curved interface between them defined by B-splines (see Figure 2). We perform ray tracing and paraxial ray tracing with a Runge-Kutta solver, and we represent the perturbation vector $\delta\mathbf{x}$ of one of the paraxial rays by straight lines drawn from the central ray. When the central ray hits the interface, we perform the local transformation defined by (34) and continue paraxial ray tracing starting with a perturbation vector parallel to the local tangent to the interface as explained above. This produces a jump in $\delta\mathbf{x}$ at every point of intersection of the central ray with the interface, as is clearly observed in Figure 2. The second jump in $\delta\mathbf{x}$ is particularly noticeable. The first jump at the interface is a forward extrapolation ($\delta\tau > 0$), in which the paraxials seem to leave a gap, while the second jump is a backward extrapolation ($\delta\tau < 0$) and the paraxials overlap. This is only a consequence of the particular choice we made to restart the tracing of the paraxial ray across the interface. Following the tip of the paraxial ray vectors one observes that this is continuous across the interface.

A Finite Element Approach To Ray Tracing In A Complex Medium

In order to take into account more complex velocity distributions, we divide the medium into elementary cells: triangles in two dimensions and tetrahedrons in three dimensions (see Figure 3). Inside each cell, a constant gradient of the square of slowness is assumed. Therefore the square of the slowness is continuous in the medium but presents first-order discontinuities at edges of each element. Inside each cell, ray tracing and paraxial ray tracing are performed analytically. Usually the most difficult numerical problem for ray tracing is to find the intersection between a ray and the internal

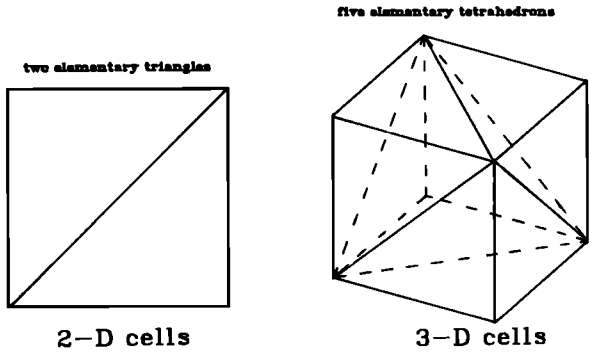

Fig. 3. The geometry of the elements used to discretize a heterogeneous medium. Two triangles are drawn inside a rectangle, while five tetrahedrons are defined inside a parallelogram.

boundaries. This is where the linear law of the square of slowness is particularly interesting: the intersection of a ray with a plane border may be calculated very easily. Let, for instance,

$$\mathbf{n} \cdot \mathbf{x} = q \tag{36}$$

be the equation of a straight line (in two-dimensions) or a plane (in three-dimensions) of unit normal $\mathbf{n}$. The intersection of the ray (19) with this edge may be found solving

$$\frac{1}{4}\mathbf{n} \cdot \gamma\, \tau^2 + \mathbf{n} \cdot \mathbf{p}_0\, \tau + \mathbf{n} \cdot \mathbf{x}_0 = q \tag{37}$$

for τ. This simple quadratic equation may be solved analytically. Similar simple expressions may be found for the intersection of a ray with boundaries described by piecewise parabolas or cubic splines. Once τ is found, it is an easy matter to obtain $\mathbf{x}(\tau)$ and $\theta(\tau)$ from the results of the section on analytical ray tracing. We observe that the parabola (19) intersects a plane boundary at two, one, or zero points depending on the discriminant of (37). It is this property that led us to prefer this parameterization of the media in order to perform analytical ray tracing inside finite elements of the medium. For elements with constant gradient of the velocity, rays are arcs of circles, whose intersection with a plane boundary is somewhat more difficult to solve than (37) [*Chapman*, 1985].

In order to illustrate the accuracy of our ray tracing method we consider a simple two-dimensional medium with a constant vertical gradient of the velocity. Rays in this medium are the well-known circular trajectories. Three of them are shown in Figure 4. We calculate ray tracing using our technique. First, the medium is divided into a mesh of regular rectangular elements as shown in Figure 4, which are in turn subdivided in triangles along the negative diagonal (Figure 3). In order to avoid cluttering the figures, only the rectangular cells are shown in the following figures. The velocity field was computed at each node of the mesh, and a linear gradient of the square of slowness inside every triangle was calculated. Taking the same initial conditions as for the three exact rays shown in Figure 4, we performed numerical ray tracing with our method. Instead of showing the result by continuous lines, the crosses in Figure 4 show the points where the numerical rays intersect the internal boundaries of the elements. Let us remark that the crosses are either on a vertical border, or on a horizontal border or on the negative diagonal. These are in fact the only points on the ray that are calculated in our numerical method. The agreement between these crosses and the exact rays is excellent. Using expression (21), the travel time is evaluated with a relative error that is lower than 10^{-3}.

Filling the three-dimensional space with arbitrary tetrahedra is a more difficult task than that of building a triangular mesh in two dimensions. We choose to define elementary parallelograms subdivided into five tetrahedrons: four right tetrahedrons and one internal element with the shape of a diamond (see Figure 3). Unfortunately, this structure is asymmetric with respect to the vertices of the parallelogram: four of the six vertices of the parallelogram coincide with the rectangular vertex of each tetrahedron. The other two vertices of the parallelogram define the axis of the internal diamond. There are thus three possible directions for the internal diamond. In order to avoid the introduction

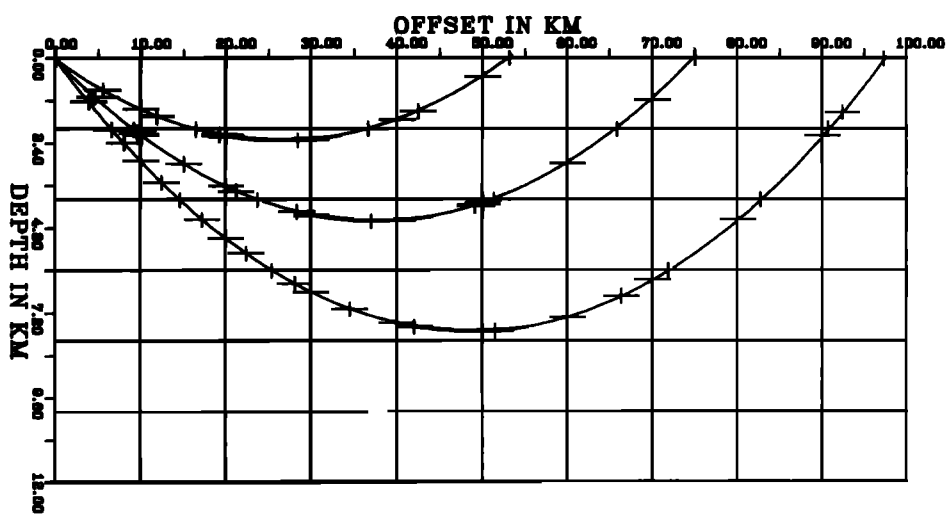

Fig. 4. A example of ray tracing in a triangular mesh. The medium has a velocity distribution, $v(z) = 3.00 + 0.02\ z\ km/s$. It is discretised using a triangular mesh, and rays are traced using our finite element technique. The continuous lines are the exact rays. The crosses indicate the points where the numerical rays intersect the sides of the elements.

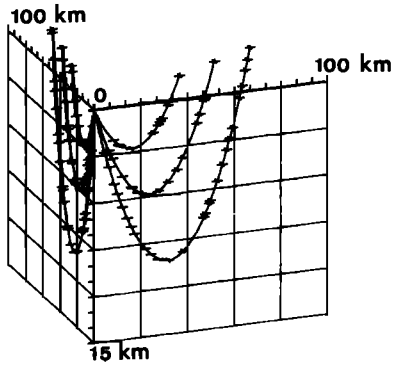

Fig. 5. Example of ray tracing in a three-dimensional medium with velocity distribution, $v(z) = 3.00 + 0.02\ z\ km/s$. The medium is discretized using tetrahedral finite elements. The continuous lines are the exact rays while the crosses indicate the points where the numerical ray intersects the sides of the tetrahedrons.

of numerical anisotropy we alternate the three diamonds in neighboring cells.

In order to test the three-dimensional routines, we consider again a medium with a vertical gradient of the velocity. This medium is divided into parallelograms and tetrahedra as described in the previous paragraph. In Figure 5 we show three exact circular trajectories on each of two vertical planes with different azimuth passing through the source at the origin of coordinates. We perform numerical ray tracing starting with the same initial slowness vectors as for the exact rays. The intersections of the numerical rays with the sides of the tetrahedrons are represented by horizontal crosses in Figure 5. They fit perfectly with the theoretical solution.

Elements with linear variation of the square of slowness present first-order discontinuities at the internal boundaries between elements. Paraxial ray tracing requires the application of the continuity conditions derived in the previous section (equation (34)). Let us assume that a paraxial ray has been traced up to a boundary and let $\delta\mathbf{x}$ be its perturbed position vector (8). In general there is no reason that $\delta\mathbf{x}$ lies along the element boundary that we are considering. In order to resume ray tracing across the boundary, we transform the ray vector using expression (34). Since the internal boundaries between elements are flat, the curvature of the interface C is zero, and (34) simplifies accordingly. Considering that there might be a large number of internal boundaries, this is not a minor advantage of the method that we propose. The transformed $\delta\mathbf{x}_t$ lies now along the boundary, and the propagation of the paraxial ray may be easily performed using (22).

In Figure 6 we present the result of paraxial ray tracing around each of the rays calculated numerically in the medium with constant vertical velocity gradient presented in Figure 4. The numerical mesh is not represented in Figure 6, where the central ray is drawn as well as $\delta\mathbf{x}$ at each intersection with internal boundaries between elements. Although the direction of $\delta\mathbf{x}$ changes abruptly at every boundary, one can easily follow the paraxial ray tube. In fact, three different ways of representing the ray tube are possible: by horizontal, vertical, and diagonal paraxial vectors $\delta\mathbf{x}$. These three representations are equivalent to first order, and they are also equivalent to the more usual paraxial ray tracing in ray centered coordinates [*Červený and Pšenčík*, 1984]. Using the formulation presented in this paper, we can easily move from one to the other as it fits the problem at hand. This flexibility in the choice of paraxial ray vectors greatly facilitates the finding of the paraxial ray passing through any point in the vicinity of the central ray. This facility to perturb rays without changing ray tracing equations is what convinced us that it was preferable to work with the full Cartesian formulation rather than reduced ray centered coordinates, as is more frequently done in the literature (see *Červený* [1985] for a review of paraxial ray theory in ray centered coordinates).

Two Point Ray Tracing

Two-point ray tracing is a very nonlinear problem. Let us note $\mathbf{x}_e$ the source position, $\mathbf{x}_s$ the station position and

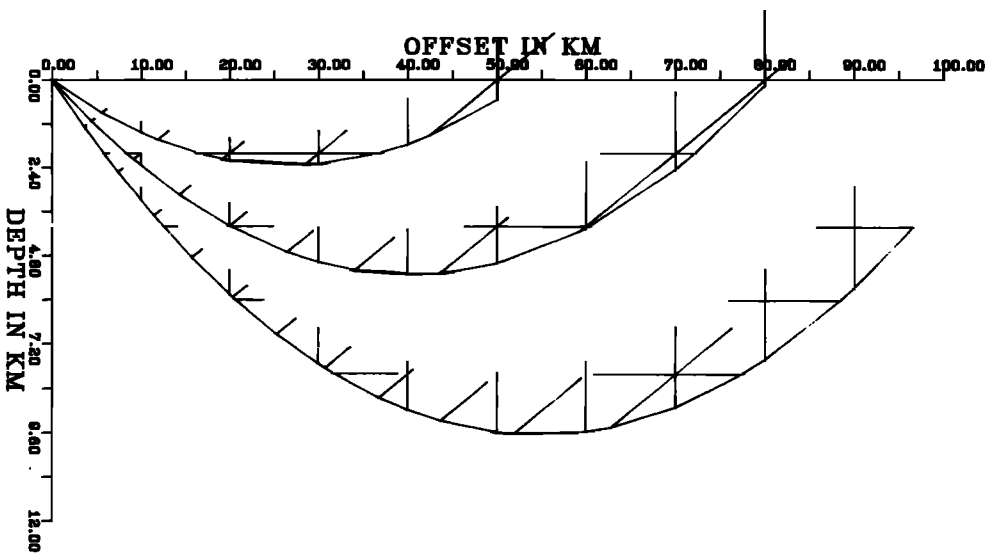

Fig. 6. A demonstration of paraxial ray propagation in a triangular mesh. In each cell, paraxial ray propagation is solved analytically. Because the velocity field has first-order discontinuities at the edges of the triangles, we must apply boundary conditions at each boundary. The vector $\delta\mathbf{x}$ is drawn at these boundaries, showing the pattern of the ray tube.

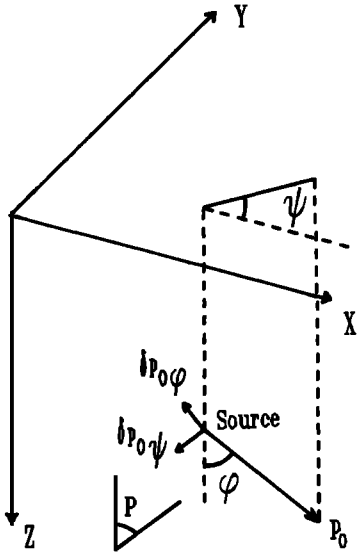

Fig. 7. Geometry of the shooting angles and two initial perturbations of the slowness vector used for paraxial ray tracing in a three-dimensional medium.

$\mathbf{p}_i$ an initial slowness vector. Two-point ray tracing may be formulated as follows: " find $\mathbf{p}_i$ such that $\mathbf{x}(\mathbf{p}_i, \mathbf{x}_c) = \mathbf{x}_s$."

In the following we shall use paraxial ray tracing to solve this problem. Let us assume that a ray has been traced that does not quite hit the station at $\mathbf{x}_s$. In order to improve the fit we linearize the ray tracing equations around the previous ray and we solve the approximate problem

$$\mathbf{x}(\mathbf{p}_i, \mathbf{x}_c) + \nabla_{\mathbf{p}_i}\mathbf{x}\,\Delta\mathbf{p}_i = \mathbf{x}_s \tag{38}$$

where $\Delta\mathbf{p}_i$ is the perturbation of the initial slowness vector we want to estimate. The only difficulty in (38) is estimating the gradient $\nabla_{\mathbf{p}_i}\mathbf{x}$, but this is calculated immediately from paraxial ray tracing. This technique has been called the paraxial ray method by previous workers [*Beydoun and Keho*, 1986; *Cormier and Beroza*, 1987]. This may be a little confusing, since paraxial ray theory can be used to build Gaussian beams, Snell waves, perturb the structure, etc.

In order to solve the two-point ray tracing we need an initial guess for the ray from the source to the station. We obtain this guess by two-point ray tracing in a vertically stratified structure that is sufficiently close to the laterally heterogeneous one. The vertically layered approximation to the structure is not difficult to define, since in most earthquake location problems one starts using a stratified model and then perturbs it to take into account known or assumed lateral heterogeneities. From this initial ray tracing, we take the slowness vector at the source. Let us consider that this initial slowness vector makes an angle ϕ with the vertical axis and has an azimuth ψ with respect to the the horisontal Cartesian axis x as shown in Figure 7. By differentiating the slowness vector with respect to ϕ and ψ, we define two paraxial rays with initial conditions

$$\begin{aligned} \delta\mathbf{q}_{\phi 0} &= 0 \\ \delta\mathbf{p}_{\phi 0} &= (cos\psi cos\phi, sin\psi cos\phi, -sin\phi)u_0 \end{aligned} \tag{39a}$$

$$\begin{aligned} \delta\mathbf{q}_{\psi 0} &= 0 \\ \delta\mathbf{p}_{\psi 0} &= (-sin\psi sin\phi, cos\psi sin\phi, 0)u_0 \end{aligned} \tag{39b}$$

as shown in Figure 7, u_0 is the slowness at the source. The central ray intersects the free surface at the point (x_f, y_f) missing the station located at (x_s, y_s). Solving paraxial ray tracing with initial conditions (39a) and (39b), we obtain the paraxial vectors along the free surface measured from the exit point of the central ray. Let these two paraxial ray vectors be $\delta\mathbf{q}_\phi$ and $\delta\mathbf{q}_\psi$. Since paraxial rays form a linear system, any linear combination of the two independent paraxial trajectories (39a) and (39b) is also a paraxial ray (at least as long as the first-order approximation is valid). Combining these solutions, we can write (38) in the explicit form

$$\begin{aligned} x_s - x_f &= \delta q_{\phi x}\delta\phi + \delta q_{\psi x}\delta\psi \\ y_s - y_f &= \delta q_{\phi y}\delta\phi + \delta q_{\psi y}\delta\psi \end{aligned} \tag{40}$$

from which $\delta\phi$ and $\delta\psi$ may be solved, because all other quantities in (40) are known. Since the linear approximation (40) may not be used very far from the central ray, we iterate with the new initial slowness vector

$$\delta\mathbf{p}_0 = \delta\mathbf{p}_\phi\delta\phi + \delta\mathbf{p}_\psi\delta\psi \tag{41}$$

until we reach the station with the desired precision. As an example of two-point ray tracing, we show in Figure 8 the iteration procedure for a source at depth in the Irpinia area that will be studied in the following sections of this paper. For each station indicated by the labels at the exit point of the rays, we performed two-point ray tracing starting with an initial shooting angle estimated for a simple layered medium. The ray hits the free surface at some distance of the station; a small cross is drawn at this intersection point in free surface. We iterate the ray tracing with a new shooting angle obtained solving (40). The ray hits the free surface at a smaller distance to the station, as shown by a new cross. Starting again, we define a series of crosses at the free surface that approach closer and closer toward the station. When the station is reached within a certain precision, the final ray is traced in the medium.

The procedure just described might fail for several reasons. The most obvious one is that the initial point (x_f, y_f) may be too far from the station and the maximum number of

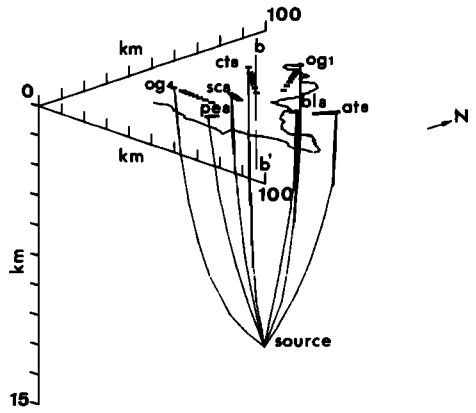

Fig. 8. An example of two-point tracing for earthquake location in the Irpinia area of central Italy. The crosses on the surface indicate the exit points of the succesive rays traced in the iterative solution of two-point ray tracing. The final result is indicated by continuous rays. The well-known leveling line transversing the area is plotted for reference, as well as the boundary bb' between the high-velocity and the low-velocity zone.

rays set by the user is reached before convergence. Another possibility is that at a certain iteration, the predicted shooting angles might be over-estimated, producing a ray which does not hit the free surface anymore. Or the discretization of the medium might be too rough and produce artificial shadow zones never reached by any ray. Numerical strategies can be set up to partially solve these problems: speeding up the convergence with numerical algorithms, fixing upper limits in the perturbation of shooting angles, and smoothing the medium by increasing the number of elements and by smoothing the square of the slowness. These problems will appear in any method that performs three-dimensional ray tracing because they are due to the instability of ray tracing in three-dimensional heterogeneous media. This is inherent to the nonlinear nature of the ray tracing problem.

Earthquake Location In Laterally Heterogeneous Media Using Hypo71

The earthquake location problem is an even more nonlinear problem than the two-point ray tracing. Let us note θ_s, the travel time observed at a station. Using the notation of the previous section, the problem can be posed as follows:" find $\mathbf{x}_e$ such that $\theta(\mathbf{x}_e, \mathbf{x}(\mathbf{p}_i, \mathbf{x}_e)) = \mathbf{x}_s) = \theta_s$," at all stations. This problem includes two-point ray tracing and might be linearised in the same way around a given source position $\mathbf{x}_e$.

$$\theta(\mathbf{x}_e, \mathbf{x}_s) + \nabla_{\mathbf{x}_e}\theta \, \Delta\mathbf{x}_e = \theta_s$$

where $\Delta\mathbf{x}_e$ is the perturbation of the source position we want to compute. Using the definition of $\mathbf{p}_0$ and the reciprocity theorem,

$$\mathbf{p}_0 = \nabla_{\mathbf{x}_e}\theta(\mathbf{x}_s, \mathbf{x}_e) = -\nabla_{\mathbf{x}_s}\theta(\mathbf{x}_e, \mathbf{x}_s) \tag{42}$$

we find

$$\theta(\mathbf{x}_e, \mathbf{x}_s) - \mathbf{p}_0 \Delta\mathbf{x}_e = \theta_s \tag{43}$$

Although a nonlinear search has been proposed recently for earthquake location [*Rabinowitz*, 1988], linearization is used by most programs.

We start by locating an earthquake in a predefined stratified medium by the standard HYPO71 program [*Lee and Lahr*, 1975]. The subroutine TRVDRV computes the travel time θ at station $\mathbf{x}_s$ and the gradient of the travel time with respect to source position $\mathbf{x}_e$. We use this location and the computed initial slowness vector as our initial guesses for starting two-point ray tracing in a laterally heterogeneous medium that is not too different from the vertically stratified medium used for the first guess. As previously explained, we define a mesh of tetrahedrons inside the medium; the square of the slowness is evaluated at the nodes of the mesh, and a constant gradient of the square of the slowness is calculated for each element of the mesh. A new travel time and new initial slowness vector are obtained for each station using the paraxial ray tracing method explained in the previous section. From the paraxial ray tracing for the last iteration for each source station ray path, we easily obtain the travel time gradient. In this form, we solve the same problem in a laterally inhomogeneous medium as the subroutine TRVDRV does in stratified media. The standard procedure of inversion in HYPO71 is reassumed, and a new location is obtained. The procedure may be started all over again from the current source position and the gradient (42) estimated by paraxial ray tracing. We may note in passing that the ray tracing procedure that we propose leads to a better correction for station elevation than the usual constant time corrections.

As long as the predefined stratified medium does not violate the first-order assumption of the paraxial ray (40), the new location is independent of the assumed stratified medium. In case of failure of two-point ray tracing for any of the reasons presented in the previous section, we use the following procedure: if the distance between the horizontal position of the central ray and the station is 5 times the required precision, we take the travel time of the paraxial ray. If not, we cancel the station for this iteration of the inversion procedure. When the station is missed three times, we attempt to shoot it from the nearest station reached. When too many stations are missed, we might end up with insufficient data for locating this particular earthquake.

The Irpinia Earthquake Of November 1980: Relocation of a Sequence of Aftershocks

The November 23,1980, Irpinia earthquake in southern Italy is one of the best studied earthquakes in the Mediterranean region [*Deschamps and King*, 1983, 1984; *Westaway and Jackson*, 1987; BZ88]. *Deschamps and King* [1984] remarked that the travel time residuals of aftershock location with the HYPO71 program, showed systematic positive time residuals for stations east of the fault and negative time residuals for stations west of it. Although the diffuse shape of the aftershock distribution observed for this event might reflect the complexity of the fault geometry (see BZ88 for discussion and further references), we decided to explore the possibility that it was due to lateral heterogeneity of the source region. The presence of a low-velocity zone in the NE part of the source has been suggested by several authors. Although its existence is well supported by geological and geophysical data, travel time inversion [*Del Pezzo et al.*, 1983] and accelerogram analysis (BZ88), its extension in depth (3 km up to 8 km), and the amplitude of the velocity perturbation are unknown. We adopt the velocity model (Table 1) derived by P. Bernard and A. Zollo from unpublished work done by M. Martini of Osservatorio Vesuviano. This velocity structure is similar to the one obtained by *Del Pezzo et al* [1983] from one-dimensional travel time inversion and therefore might still overestimate the velocities. The high-velocity area has a velocity close to the one

TABLE 1. *P* Wave Velocity Structure

Depth of nodes, km	High-Velocity Zone, km/s	Low-Velocity Zone, km/s
0.	4.50	2.27
3.	4.50	2.27
3.1	5.32	5.32
7.	6.03	6.03
10.0	6.28	6.28
20.0	6.54	6.54

The *S* wave velocity is deduced by dividing the *P* wave velocity by 1.79.

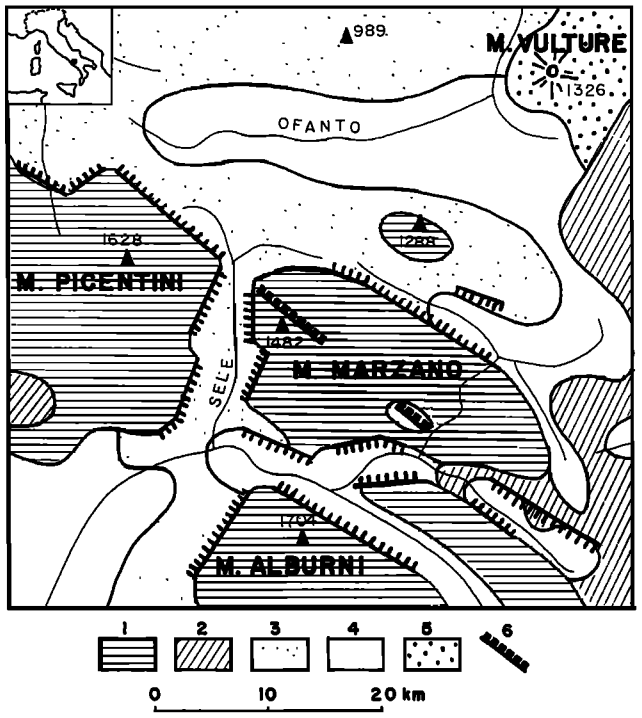

Fig. 9. Schematic geostructural map of the Irpinia area from P. Bernard and A. Zollo (BZ88). 1 Campano Lucanian carbonate unit (secondary); 2 Lagonegro unit (Oligo Miocene); 3 Flyschs of the internal unit (Miocene-Pliocene); 4 Plioquaternary deposists; 5 volcanic unit (quaternary); 6 reported fault trace for the earthquake of November 23, 1980. One might notice the contrast between the NE low-velocity zone and the SW high velocity zone. The SW zone has a somehow more complex pattern which is not considered in this study. Surface trace corresponds to the boundary between the two compartments.

determined by *Deschamps and King* [1984] from the best fitting model for aftershock location. The surface trace of the main active fault of this event was identified by *Westaway and Jackson* [1987]; it dipped toward the NorthEast with an angle of 60° , and had a NW-SE strike at the surface (Figure 9). The trace of the fault coincides with the vertical boundary between the two proposed compartments with different velocity.

Before studying real data, let us first look at a few synthetic examples in order to investigate the effect of the velocity distribution on the location of earthquakes. First, we computed synthetic travel times in the previously described heterogeneous medium for a set of events distributed on a T-shaped region near the center of the network. We then tried to relocate these events using the synthetic travel time data set. The results are shown in Figure 10. At the top is shown a perspective view of the hypocenters, in the center a plan view, and at the bottom a cross section. The larger dots in the leftmost plan view denote the stations used in the location. The second column shows results when the velocity distribution used to locate the aftershocks is the same as the heterogeneous medium used to compute synthetic travel times. The third column gives standard results of HYPO71 for the layered velocity distribution used by P. Bernard and A. Zollo (BZ88), while the fourth column shows the locations in the layered medium used by *Deschamps and King* [1984]. Using the same heterogeneous medium for locating earthquakes as for the synthetic travel time computation, gives a good precision of the horizontal locations. Use of a layered medium moves the epicenters to the SW whatever the stratified structure. For the heterogeneous medium, depth precision reduces rather quickly below 12 km. For stratified media, the error in depth determination increases dramatically, but the relative horizontal shape of the aftershock distribution is preserved (Figure 10, center).

What happens when we modify the low-velocity structure? Two numerical experiments were conducted. They are presented in Figure 11, which is similar to Figure 10 with the perspective view omitted. The first two columns of Figure 10 are the same as those of the preceding figure; they correspond to (*a*) the distribution of events we want to locate and (*b*) its relocation using the same heterogeneous medium as for the travel time calculation.

In the first experiment, we moved the southern part of the boundary between the two blocs 10 km to the NE, so that stations BL8 and BA8 were displaced from the low-velocity side to the high-velocity one. We recomputed synthetic travel times for this structure, and we relocated the events using the heterogeneous medium used in Figure 10. Results are presented in the third column of Figure 11. We find that the horizontal shape of the aftershock distribution is remarkably stable, without any lateral motion, while the depth is very sensitive to the difference in lateral structures. A more dramatic change is observed by increasing the thickness of the sedimentary bloc from 3 km to 5 km. In the fourth column of Figure 11, the large horizontal line of the T-shaped distribution of events is still well located, confirming the compensation of errors in a well-distributed set of stations. The hypocenters that form the NS part of the T-shaped distribution of events are very affected by the change in velocity model as shown in the last column of Figure 11. The events of this group are controlled by the two stations that were moved into the high-velocity zone.

At this point, two main conclusions can be drawn. In the central part of the network, the epicenter is almost independent of the vertical structure used for location, while the depth is not well constrained. When a lateral velocity anomaly exists, the hypocenter locations might be displaced by distance much larger than the estimated error given by HYPO71. This optimistic estimation of the error comes from the assumption of perfect knowledge of the velocity structure. Taking into account the lateral variation of the velocity structure in ray tracing reduces the horizontal displacement of the hypocenters, but vertical resolution remains poor. Adding normally distributed random errors to the data set will introduce scattering in the locations but not a systematic bias [*Pavlis and Hokanson*, 1985].

We analyzed aftershocks from December 7 to 15, 1980, using the data set obtained cooperatively by several French, Italian and British institutions [*Deschamps and King*, 1983]. The aftershock arrival time data come from *Deschamps and King* [1984], who estimated an accuracy of about 0.1 s for *P* wave arrival times. *S* wave arrival times were used in the locations with half the weighting of *P* waves in order to constrain the depth determination. A constant *P*-to-*S* wave velocity ratio of 1.79 was assumed. The distribution of large-magnitude aftershocks found by *Deschamps and King* [1984] defines an almost vertical trend in the central part of the fault. This locates a substantial part of the aftershock

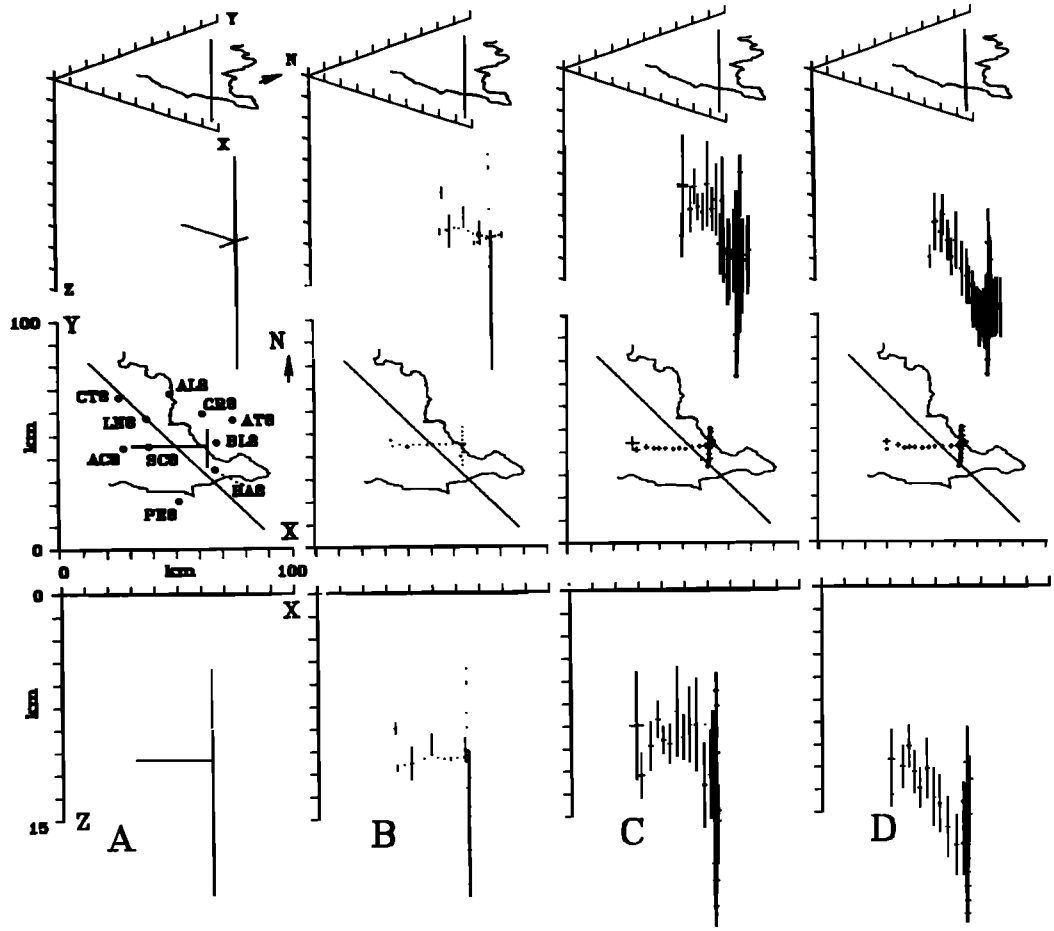

Fig. 10. A synthetic example where a T-shaped set of hypocenters is defined in a heterogenous medium. (*a*) Original data. (*b*) Locations obtained using the same heterogeneous medium as used to calculate synthetic travel times. (*c*) Using the stratified medium of P. Bernard and A. Zollo (BZ88). (*d*) Using the stratified medium of *Deschamps and King* [1984]. Three projections are shown in every column: stereo projection at the top, plan view projection at the center, and vertical projection at the top.

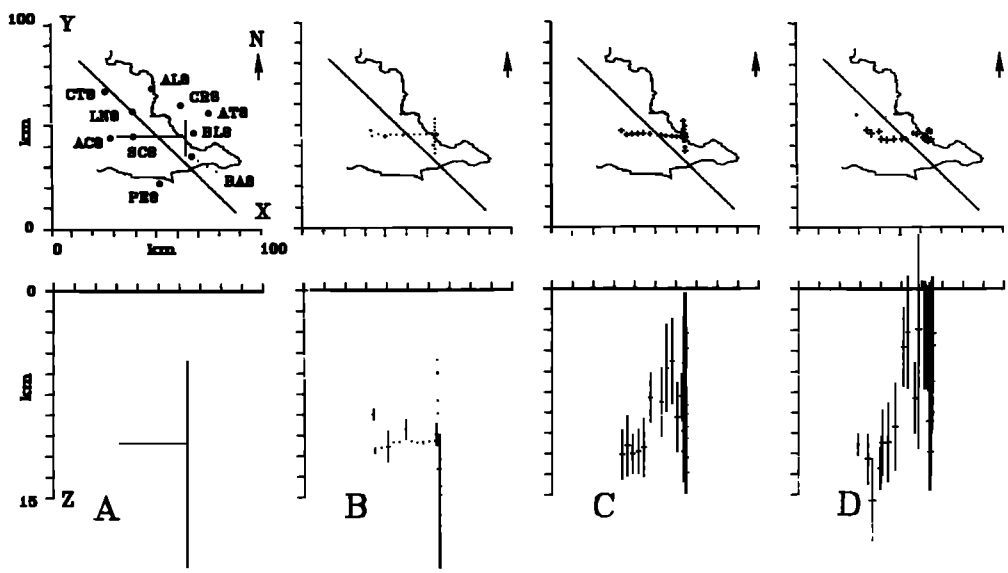

Fig. 11. Deformation of the T-shaped set of hypocenters shown in Figure 10 as the synthetic "real" medium differs more and more from the one used in the location program. (*a*) Initial geometry. (*b*) The earthquakes are relocated using the original medium. (*c*) The horizontal boundary is more complex for the "real" medium. (*d*) The thickness of the low-velocity bloc is increased from 3 km to 5 km for the previous "real" medium. In each column two projections are shown: plan view at the top and vertical projection at the bottom.

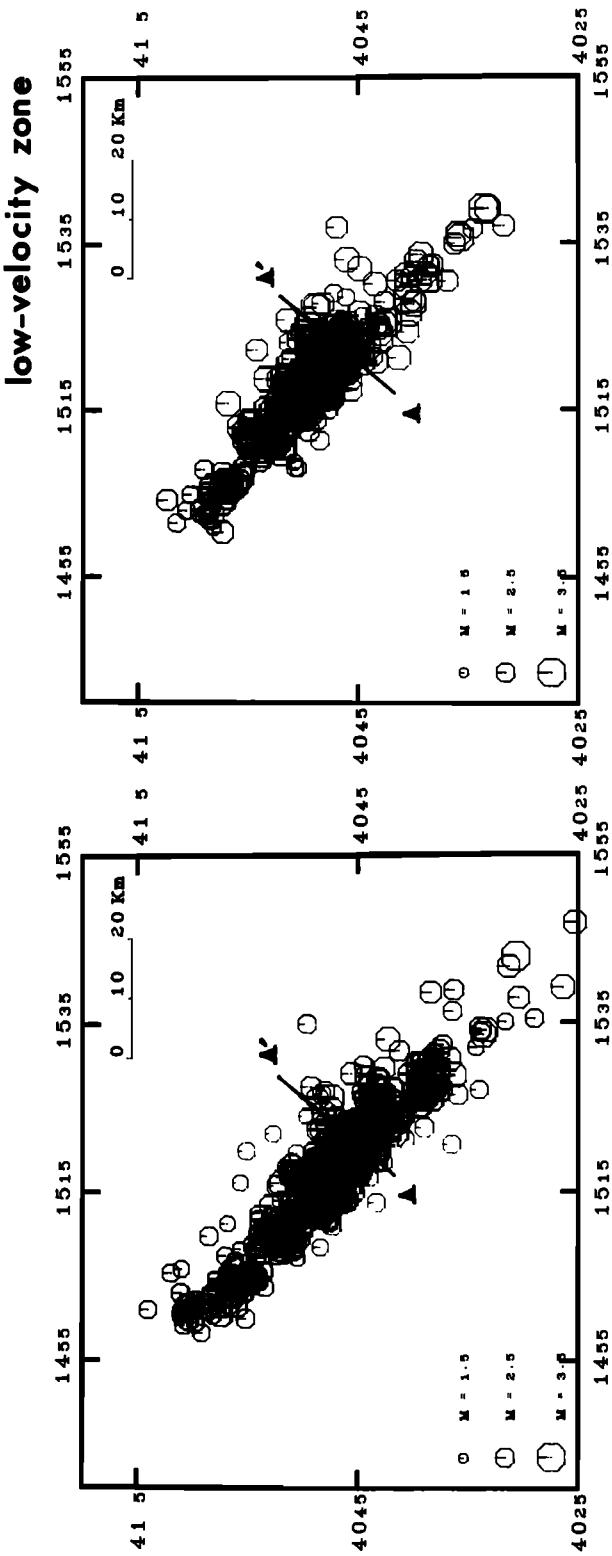

Fig. 12. Comparison between epicenter locations for the stratified medium of *Deschamps and King* [1984] on the left and for the heterogeneous medium described in Table 1 at the right. The pattern is very similar with a general drift of the hypocenters toward the NW.

activity in the footwall. Let us now study whether the introduction of the low-velocity zone to the NE of the fault reduces the dip of the aftershock trend, as suggested by P. Bernard and A. (BZ88).

Close to 450 earthquakes were relocated which had at least four P wave, two S wave readings, and nominal locating errors lower than 4 km. *Deschamps and King* [1984] obtained more than 500 locations using this criterion. Figure 12 presents their result and compares it to ours. The pattern of epicenters is nearly the same, with the expected translation toward the northeast. Three segments may be observed in the seismic activity, as already noted by *Deschamps and King* [1984]. Although the SE segment might be subdivided in two parts for the heterogeneous model, its position slightly outside the network does not justify such detail.

A cross section of 5 km on each side of line AA' in Figure 12 demonstrates the change in vertical distribution of hypocenters when the laterally heterogeneous model is used. The central zone around this line is the area best controlled by the seismic network; this is also the area where the two main events of this sequence occurred. Shallow earthquakes, although existing, are not shown in the cross section of Figure 13 because their location error is greater than 4 km. In any case, they are deeper than the location obtained with the standard version of HYPO71, because stations in the high-velocity compartment are reached by refracted rays, a result that is strongly dependent on the geometry of the low-velocity zone. Subvertical planes defined by aftershocks are still observed, but they are translated to the NE with respect to the surface fault traces. The more or less vertical plane of activity below the fault trace is not observed anymore. The main activity is concentrated now in the hanging wall as hypothesized by *Westaway and Jackson* [1987], who argued that the footwall of the fault is expected to return to equilibrium after the main shock, and thus to be less active.

The presence of the low-velocity zone does not generate any obvious structure with a dip of 60° corresponding to the main fault plane inferred for the main shock. We have no indication that other velocity distributions might change this. Taking into account the location of the surface fault traces, we may conclude from the hypocenter distribution shown on the right of Figure 13 that the fault plane dips to the NE. This is more coherent with the fault plane solution for the main shock than the vertical plane deduced from locations in a layered medium. However, the dip of the fault remains unknown, because of the poor depth resolution that depends on the details of the low-velocity zone to the NE of the fault.

It is interesting to note that two vertical parallel lines plunging down to 15 km are observed with a quiet area of 5-km width between them. This separation is larger than any location errors. This particular geometrical structure of the hypocenter distribution might be guessed in the standard HYPO71 results. Although verification with other velocity models is necessary, these features may be related to two of the main events in this area [*Westaway and Jackson*, 1987; BZ88].

Discussion

Classical ray tracing is usually considered too slow for routine use in earthquake location programs in laterally

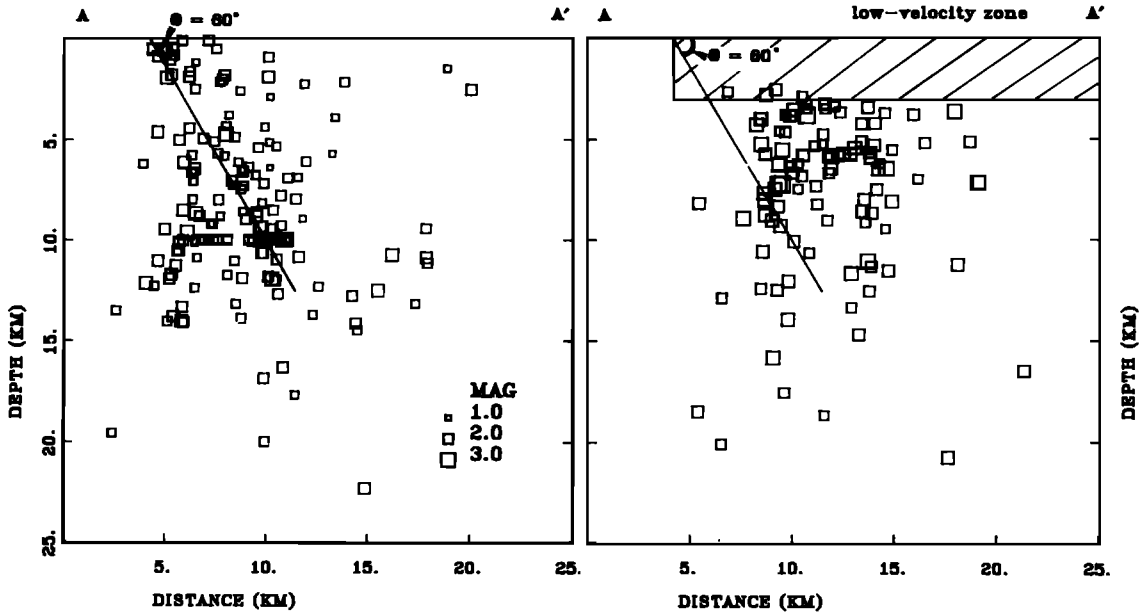

Fig. 13. A cross section of hypocenters along the direction AA' defined in Figure 12 near the central part of the active fault. Hypocenters within a box of 5-km width on each side of the profile were plotted in this figure. The figure on the left shows the locations in a stratified medium, and on the right the locations in the heterogeneous medium. The NW displacement of the hypocenters is clearly seen. Shallow aftershocks are removed from the right-hand figure because of their large error in depth. Two distinct planes are observed in the locations in the heterogeneous medium: their dip angle is not very well constrained, but their parallelism might be a real feature.

heterogeneous media. The lack of speed is due to the use of a finite difference approach for the integration of the ray tracing equations. In the classical Runge-Kutta technique the most time consumming routine in three dimensions is the determination of ray intersections with internal boundaries. In order to accelerate ray tracing, several authors [*Chapman*, 1985] have proposed a finite element approach in which the medium is divided into a set of elements with simple velocity distributions. This is the approach we adopted in this paper. After an extensive search we came up with a medium in which all the elements of ray tracing are calculated analytically, including intersections with parabolic or cubic boundaries. This is a medium with constant gradient of the square of slowness. Using the expressions derived in this paper, initial value ray tracing in three dimensions reduces to solving a number of polynomial equations for the value of the ray parameter τ at the boundaries. Since these equations may be solved exactly for polynomials up to degree 4, our method is much faster than Runge-Kutta ray tracing.

Two-point ray tracing is usually an order of magnitude more expensive than initial value ray tracing, except when very simple media are assumed for ray tracing [*Thurber*, 1983]. Again, the simple medium adopted inside the individual elements yields analytical expressions for paraxial ray tracing which can then be used very efficiently for the calculation of the gradients needed in Newton's method for the solution of two-point ray tracing. Again our method is fast enough for ray tracing to be used in solving the two-point ray tracing problem. In general, however, the shooting angle is not a single valued function of station position. There might be several rays that connect the station to the observer. Multipathing is a pervasive problem in three dimensional laterally heterogeneous media. If we wanted to find all the ray paths between a source and an observer, we would have to locate the separate branches where the travel time is single-valued. These different branches are bounded by caustics that describe very complex figures (umbilics, swallow tails, etc.) which can be described by catastrophe theory [*Nye*, 1985]. Unfortunately, locating the caustics in three dimensions is still an open problem and, certainly, a much more difficult one than ray tracing. For this reason, starting from a given initial slowness vector, the paraxial method might not converge to the ray with the true minimum travel time. Nevertheless, as many other authors have implicitly done, we will assume that lateral variations of the medium are such that starting from a ray traced in the initial vertically stratified medium will converge to the correct solution. Obviously this problem requires further work, especially if three dimensional ray tracing is used in tomography.

How is the paraxial method connected to other methods for solving the two-point ray tracing problem? Paraxial ray tracing can be interpreted in three ways. In the method adopted here a ray is traced in the medium with a given initial slowness vector. If this ray misses the observer, the initial conditions defined by (41) yield a new estimation of the initial slowness vector by first-order perturbation. The procedure can be iterated until the ray hits the receiver with a given precision. From another point of view, we might consider also that the source and the receiver are always connected by some trajectory that does not exactly satisfy the ray tracing equations (6). Considering that this trajectory is a paraxial ray of the real trajectory, we may "bend" the paraxial ray until it coincides with the exact ray joining the two endpoints. The equations for bending and paraxial ray tracing are the same as can be easily verified comparing our results to those of *Julian and Gubbins* [1977]. Finally, the paraxial ray can be considered as an exact ray in a very particular medium expanded quadratically around a certain

central ray. By continuous modification of this medium, we obtain the ray in the correct medium to a given precision [*Keller and Perozzi*, 1983]. Using perturbation theory [*Farra and Madariaga*, 1987], we can show that again in this case the equations that are actually solved are equivalent to (40). The equivalence between the three methods is not surprising, since they are all based on Newton's method for the iterative solution of nonlinear problems. The actual paths in parameter space followed by the three solutions are different, but at every iteration step they use a linearization of the ray equations (6) that may be solved by paraxial ray tracing. Paraxial ray tracing provides a simple but powerful method for the unification of these apparently different techniques.

Although many other iterative methods have been proposed in the literature, HYPO71 is still the most widely used program for earthquake location. For this reason we decided to maintain its procedures for source determination, and we adapted our paraxial ray tracing technique to replace one of the subroutines in that program. In our method we start locating the earthquake with classical HYPO71 using an approximate vertically layered medium. We use the ray trajectories in this medium as a starting point for iterative ray shooting and earthquake location.

Because we start by locating the earthquake in a stratified medium, total computer time includes the standard computer time of HYPO71. For vertically varying media, a maximum distance error of 2 km is enough to guarantee a precision of 5 ms in travel time for a total length of the ray around 50 km, when the quadratic interpolation of travel time (15) is used. For this geometry, we multiply the standard computer time of HYPO71 by a factor of 25, which is still acceptable for an interactive use. More complex structures, which are defined by the number of nodes of the tetrahedral structure, might increase the computer time dramatically, and a conservative approach is necessary for choosing input parameters of the program. For our simple application to the sequence of aftershocks of the Irpinia earthquake, an average 1 mn CPU time was necessary to locate a single event on a Prime 9950.

Conclusions

We proposed a simplified ray tracing algorithm for three dimensions that is fast enough to be used for the relocation of earthquakes in laterally heterogeneous structures. The use of tetrahedral elements with constant gradient of the square of slowness greatly simplifies ray tracing. Analytical results inside the tetrahedral elements may be used in order to accelerate ray tracing. The solutions inside the tetrahedrons may be easily connected across their walls by simple boundary conditions. Complex media with lateral variations of elastic wave velocity are modeled dividing the medium into a series of tetrahedrons with the simple velocity law proposed above. Ray tracing reduces in this case to the solution of a series of polynomial equations in one variable. Our technique accelerates ray tracing in three dimensions by more than an order of magnitude compared with classical Runge-Kutta techniques.

We apply ray and paraxial ray tracing in heterogeneous structures to the solution of the two point ray tracing problem: we accurately compute the travel time and its derivatives at different stations for a given source. By including this ray tracing in the HYPO71 program, we were able to locate earthquakes in more complex media than the layered structures for which HYPO71 was originally designed. The sequence of aftershocks of the Irpinia, central Italy, earthquake of November 23, 1980, is a good example of a region that needs the use of a lateral variation of the velocity structure because of a simple systematic pattern of time residuals. Hypocenter relocation with our method shows, as expected, a general displacement of the aftershocks toward the NE where there is a low-velocity zone. Unfortunately, depth resolution is still very poor and prevents any refinement on the fault plane activity.

Acknowledgments. We thank all the researchers that helped collect the Irpinia aftershock data set. We thank Pascal Bernard, Anne Deschamps and Roland Gaulon for numerous discussions on the Irpinia area, and Michael Hamburger for suggesting the synthetic experiment on the Irpinia network. Many thanks to Wafik Beydoun, who critically commented a first draft. This work has been supported by several grants from the Secteur TOAE du Centre National de la Recherche Scientifique. contribution IPG 1011.

References

Babich, V. M., I. A. Molotov, and M. M. Popov, Gaussian beams solutions concentrated around a line and their use (in Russian), Inst. of Radiotech. and Electron., Leningrad, 1985.

Beydoun, W. B., and T. H. Keho, The paraxial ray method: Fast computation of Green's functions for inverse scattering, paper presented at the 56th Annual International Meeting and Exposition, Soc. of Explor. Geophys., Washington D. C., 1986.

Burridge, R., *Some Mathematical Topics in Seismology*, Courant Institut of Mathematical Sciences, New York University, New York, 1976.

Červený V., I. Molotkov, and I. Pšenčík, *Ray Method in Seismology*, Karlova Universita, Prague, 1977.

Červený, V., M. M. Popov, and I. Pšenčík, Computation of wave fields in inhomogeneous media—Gaussian beam approach, *Geophys. J. R. Astron. Soc.*, *70*,109—128, 1982.

Červený, V., and I. Pšenčík, Gaussian beams in elastic two dimensional laterally varying layered structures, *Geophys. J. R. Astron. Soc.*, *78*, 65—91, 1984.

Červený, V., The application of ray tracing to the numerical modelling of seismic wave fields in complex structures, in *Handbook of Geophysical Exploration*, Sect. 1, *Seismic exploration*, vol. 15A, pp. 1—119, Geophysical Press, London, 1985.

Červený, V., Ray tracing algorithms in three dimensional laterally varying layered structures, in it Tomography in Seismology and Exploration Seismics, edited by G. Nolet, D. Reidel, Hingham, Mass., 1987.

Chapman, C. H., Ray theory and its extensions: WKBJ and Maslov seismograms, *J. Geophys.*,*58*, 27—43, 1985.

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

Cormier, V. F., and G. C. Beroza, Calculation of strong ground motion due to an extended earthquake source in laterally varying structure, *Bull. Seismol. Soc. Am.*, *77*, 1—13, 1987.

Del Pezzo, E., G. Iannaconne, M. Martini, and R. Scarpa, The 23rd November 1980 Southern Italy earthquake, *Bull. Seismol. Soc. Am.*, *73*, 187—200, 1983.

Deschamps, G.A., Ray techniques in electromagnetics, *Proc. IEEE*, *60*, 1022—1035, 1972.

Deschamps, A., and G. C. P. King, The Campania-Lucania (Southern Italy) earthquake of 23 November 1980, *Earth Planet. Sci. Lett.*, *62*, 296—304, 1983.

Deschamps, A., and G. C. P. King, Aftershocks of the Campania-Lucinia (Italy) earthquake of 23 November 1980, *Bull. Seismol. Soc. Am.*, *74*, 2483—2517, 1984.

Farra, V., Approche hamiltonienne de la théorie des rais : application à l'étude de perturbations, inversion des temps de

parcours en sismique réflexion, Thèse de l'Université Paris 7, France, 1987.

Farra, V., and R. Madariaga, Seismic waveform modeling in heterogeneous media by ray perturbation theory, *J. Geophys. Res.*, *92*, 2697—2712, 1987.

Julian, B. R., and D. Gubbins, Three dimensional seismic ray tracing, *J. Geophys.*, *43*, 95-114, 1977.

Keller, H. B., and D. J. Perozzi, Fast seismic ray tracing, *SIAM J. Appl. Math.*, *43*, 981—992, 1983.

Klimeš, L., The relation between Gaussian beams and Maslov asymptotic theory, *Stud. Geophys. Geod.*, *28*, 237—247, 1984.

Lee, W. H. K., and J. C. Lahr, HYPO71: a computer program for determining hypocenter, magnitude, and first motion pattern of local earthquakes, *U.S. Geol. Surv. Open file rep.*, *75-311*, 1975.

Luneberg, R. K., *Mathematical Theory of Optics*, Brown University, Providence, R. I., 1944.

Madariaga, R., Gaussian beam synthetic seismograms in a vertically varying medium, *Geophys. J. R. Astron. Soc.*, *79*, 589—612, 1984.

Maslov, V. P., *Perturbation theory and asymptotic methods* (in Russian), Moskov., Gos. Univ., Moscow, 1965. (translated into French by J. Lascoux and R. Sénéor, Dunod, Paris, 1972).

Nye, J. F., Caustics in seismology, *Geophys. J. R. Astron. Soc.*, *83*, 477—485, 1985.

Pavlis, G. L., and N. B. Hokanson, Separated earthquake location, *J. Geophys. Res.*, *90*, 12,777—12,789, 1985.

Popov, M. M., A new method of computation of wave fields using Gaussian beams, *Wave Motion*, *4*, 85—97, 1982.

Popov, M. M. and I. Pšenčík, Computation of ray amplitudes in inhomogeneous media with curved interfaces, *Stud. Geophys. Geod.*, *22*, 248—258, 1978.

Prothero, W. A., W. J. Taylor and J. A. Eickemeyer, A comparison of two fast 3-D raytracers, *Eos Trans. AGU*, *68*, 1373, 1987.

Rabinowitz, N., Microearthquake location by means of nonlinear simplex procedure, *Bull. Seismol. Soc. Am.*, *78*, 380—384, 1988.

Thomson, C. J., and C. H. Chapman, An introduction to Maslov's asymptotic method, *Geophys. J. R. Astron. Soc.*, *61*, 729—746, 1985.

Thurber, C. H., Earthquake locations and three dimensional crustal structure in the Coyote Lake area, central California, *J. Geophys. Res.*, *88*, 8226—8236, 1983.

Um, J., and C. H. Thurber, A fast algorithm for two-point seismic ray tracing, *Bull. Seismol. Soc. Am.*, *77*, 972—986, 1987.

Westaway, R., and J. Jackson, The earthquake of 1980 November 23 in Campania Basilicata (Southern Italy), *Geophys. J. R. Astron. Soc.*, *90*, 375—443, 1987.

V. Farra, R. Madariaga and J. Virieux, Laboratoire de Sismologie, Université Paris 7 and Institut de Physique du Globe, 4 Place Jussieu, Tour 14, 75252 Paris Cedex 05, France.

(Received September 1, 1987;
revised February 12, 1988;
accepted March 1, 1988.)