JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 92, NO. B3, PAGES 2697-2712, MARCH 10, 1987

Seismic Waveform Modeling In Heterogeneous Media by Ray Perturbation Theory

VERONIQUE FARRA AND RAUL MADARIAGA

Laboratoire de Sismologie, Institut de Physique du Globe, Paris, France

We study the propagation of rays, paraxial rays, and Gaussian beams in a medium where slowness differs only slightly from that of a reference medium. Ray theory is developed using a Hamiltonian formalism that is independent of the coordinate system under consideration. Let us consider a ray in the unperturbed medium. The perturbation in slowness produces a change of the trajectory of this ray which may be calculated by means of canonical perturbation theory. We define paraxial rays as those rays that propagate in perturbed medium in the vicinity of the perturbed ray. The ray tracing equation for paraxial rays may be obtained by a linearization of the canonical ray equations. The linearized equations are then solved by a propagator method. With the help of the propagator we form beams, i. e. families of paraxial rays that depend on a single beam parameter. The results are very general and may be applied to a number of kinematic and dynamic ray tracing problems, like two-point ray tracing, Gaussian beams, wave front interpolation, etc. The perturbation methods are applied to the study of a few simple problems in which the unperturbed medium is homogeneous. First, we consider a two-dimensional spherical inclusion with a Gaussian slowness perturbation profile. Second, transmission and reflection problems are examined. We compare results for amplitude and travel time computed by exact and perturbed ray theory. The agreement is excellent and may be improved using an iterative procedure by which we change the reference unperturbed ray whenever the perturbation becomes large. Finally, we apply our technique to a three-dimensional problem: we calculate the amplitude perturbation and ray deflection produced by the velocity structure under the Mont Dore volcano (central France). Again a comparison shows excellent agreement between exact and perturbed ray theory.

INTRODUCTION

Exact solutions of wave propagation problems in three dimensions may be obtained only by numerical techniques like finite differences or finite elements. There are many problems, however, where approximate techniques like geometrical ray theory are applicable. So far most of the applications of ray theory have been limited to ray tracing because of some essential instabilities that appear in the numerical evaluation of ray amplitudes. These are mostly due to the fact that ray theory is an infinite frequency method. At finite frequencies many of these instabilities are smoothed out. Recently, several new methods have been proposed to avoid these instabilities in ray theory. For instance, *Chapman* [1978] proposed to use the WKB method to generate high-frequency seismograms in vertically heterogeneous media. *Chapman and Drummond* [1983] used the uniform Maslov method for the calculation of high-frequency fields in arbitrarily heterogeneous media. Even more recently, researchers in Leningrad and Prague proposed to use a superposition of Gaussian beams in order to generate high frequency synthetics [e.g., *Červený et al.*, 1982; *Červený and Pšenčik*, 1984; V. M. Babich et al., preprint, 1985]. All these methods allow relatively fast computation of complex seismograms without special care for caustics and other ray field singularities [*Nowak and Aki*, 1984; *Yomogida and Aki*, 1985; *George et al.*, 1986] . As shown by *Madariaga and Papadimitriou* [1985] and *Klimeš* [1984], Gaussian beam summation in vertically varying media is an extension of the WKB method with some complex initial conditions. The unifying concept here is the so-called paraxial ray theory that incorporates a series of apparently unrelated techniques like dynamic ray tracing, Gaussian beams, WKB ,ray bending, etc., in a common theoretical framework. We intend to show in this paper that all these problems are particular applications of the paraxial approximation, which is itself derived from first-order ray perturbation theory.

A related problem that frequently appears in practice is how will the ray field be affected by small changes in the velocity or slowness structure. This problem has already been properly formulated for surface wave tracing when the structure differs only slightly from a spherically averaged model. *Woodhouse and Wong* [1986] have recently showed how to calculate rays and amplitudes in this perturbed medium.

In this paper we will study non dispersive ray propagation in three-dimensional heterogeneous media. A general formalism based on Hamiltonian theory will be used so that our results are independent of the coordinate system. Once the general formalism is presented, we will discuss in detail the case of a three-dimensional homogeneous medium to which we add a small smooth velocity perturbation. Ray and amplitude perturbations will be compared with exact solutions. For this same problem we will show that Gaussian beam perturbations may be found by the same paraxial ray tracing used to evaluate geometrical spreading. An example of Gaussian beam calculation will be presented. Finally, we demonstrate a three-dimensional application to the calculation of amplitudes simultaneously with the travel time perturbations. This results may be used to include wave amplitudes in the inversion of seismic velocity by *Aki et al.* 's [1977] method.

RAY AND PARAXIAL RAY THEORY

Ray and ray perturbation theory will be presented in a coordinate independent form. We adopt for this pur-

Paper number 6B6079.
0148-0227/87/006B-6079$05.00

pose the Hamiltonian formalism of analytical mechanics. Let us consider the high-frequency asymptotic solution to the elastic wave equation in a laterally heterogeneous three-dimensional body. *Karal and Keller* [1959] showed that to zeroth order in frequency, P and S waves propagate independently and that the Fourier- transformed displacement field may be written in the form

$$u(q,\omega) = A(\omega)\sqrt{\frac{\rho_0\, c_0}{\rho\, c J(q,q_0)}}\exp[i\omega\theta(q,q_0)] \qquad (1)$$

where q_0 and q are the generalized coordinates of the source and receiver, respectively; ρ and ρ_0 are the densities at the receiver and the source, respectively; and c and c_0 are the corresponding wave speeds. $A(\omega)$ is a complex amplitude vector that depends on the polarity of the wave and the source excitation. For P waves, $c = \alpha$ (the P wave velocity) and A is tangent to the ray; for S waves, $c = \beta$ (the shear wave velocity) and A is a transverse vector tangent to the local wave front. In some of the examples discussed below we used an acoustic approximation for the P waves. In this case, the square rooted term in (1) is replaced by

$$\sqrt{\frac{c}{c_0 J}}$$

The travel time, $\theta(q,q_0)$, and geometrical spreading function, $J(q,q_0)$, are obtained by ray tracing in a three-dimensional perfectly elastic body. The rays can be found by solving the eikonal equation:

$$(\nabla\theta)^2 = c^{-2}(q) = u^2(q) \qquad (2)$$

where q denotes the generalized coordinates of a point inside the earth, c is the appropriate wave speed and u is the corresponding slowness. ∇ is the gradient operator in q coordinates, so that $\nabla\theta$ is normal to the wave front and represents the slowness vector.

Two different versions of the Hamiltonian formalism have been discussed in the seismological literature. The first is based on the use of an independent parameter to measure position along the ray. This parameter may be distance along the ray, travel time or the parameter discussed by *Burridge* [1976] and *Chapman and Drummond* [1982]. For instance, if we choose arc length s as the parameter, then the Hamiltonian is

$$H(q,p,s) = \frac{1}{2}\,u^{-1}(q)\,[p^2 - u^2(q)] = 0 \qquad (3)$$

where $p = \nabla\theta$ is the slowness vector. Use of this formalism leads to a system of six ray tracing equations, which must be integrated subject to the condition that $H = 0$. We will not discuss this approach any further since it will not be adopted in this paper.

The alternative formulation, which will be adopted here, is to use the eikonal equation (2) to eliminate one of the space variables. Since H in (3) is always zero, at least one of the space coordinates is cyclic [*Landau and Lifschitz*, 1981; *Goldstein*, 1981] and may be used as a parameter along the ray. Let σ denote this independent variable and q a two-dimensional vector containing the other two position variables. In general orthogonal curvilinear coordinates the slowness vector has the form:

$$\nabla\theta = \left(h_\sigma^{-1}\frac{\partial\theta}{\partial\sigma},\ h_1^{-1}\frac{\partial\theta}{\partial q_1},\ h_2^{-1}\frac{\partial\theta}{\partial q_2}\right)$$

where $h_\sigma(\sigma,q)$ and $h_i(\sigma,q)$ are the scale factors corresponding to each of the coordinates. Then we define for each of the two q_i a conjugate momentum

$$p_i = \frac{\partial\theta}{\partial q_i} \qquad (i \neq \sigma)$$

Let us remark that with this definition the p_i are not the components of the slowness vector except in the special cases when the scale factors $h_i = 1$.

The position q and momenta p define a four-dimensional phase space instead of the six-dimensional one for the classical Hamiltonian (3). A ray trajectory is defined as the curve described in this space by the canonical vector $y(\sigma) = [q(\sigma), p(\sigma)]$. Ray trajectories are parameterized by the independent coordinate σ. Finally, we define the reduced Hamiltonian:

$$H(q,p,\sigma) = -h_\sigma\sqrt{u^2(\sigma,q) - h_1^{-2}p_1^2 - h_2^{-2}p_2^2} \qquad (4)$$

Considering (2) and the definition of $\nabla\theta$, we observe that

$$H = -\partial\theta/\partial\sigma = -p_\sigma$$

Thus the reduced Hamiltonian H has a clear physical meaning: it is the momentum associated with the independent variable σ. With these definitions, the complete position coordinates are (σ, q_1, q_2), the conjugate momentum is $(-H, p_1, p_2)$, and the slowness vector is $(-h_\sigma^{-1}H, h_1^{-1}p_1, h_2^{-1}p_2)$.

Two systems of coordinates will be considered in this paper. First, the simplest, is the usual geographical system shown in Figure 1 in which the independent parameter σ is the vertical coordinate z that points into the earth. In this system the scale factors $h_z = h_i = 1$. The other coordinate system that will be used is the ray-centered system introduced by *Babich and Buldyrev* [1972] and shown in Figure 2. This system is built around a certain given unperturbed reference ray; σ is the distance along this ray, and q are orthogonal coordinates defined on a plane perpendicular to it. This is the system most frequently used to construct Gaussian beams (see, for example, *Červený et al.* [1982]). In this case, $h_\sigma = (1 - u^{-1}q\ \partial u/\partial q)$ and $h_q = 1$.

Equation (4) contains all the information that is needed to generate rays and calculate the wave fronts. Rays are obtained from (4) by applying Hamilton's method [*Červený et al.*, 1977; *Červený*, 1985] which yields the following canonical ray equations:

$$\dot{q} = \frac{\partial H}{\partial p} \qquad \dot{p} = -\frac{\partial H}{\partial q} \qquad (5)$$

where the dot indicates derivative with respect to the independent parameter σ. Equations (5) should be completed with appropriate initial or boundary conditions at $\sigma = \sigma_0$ and $\sigma = \sigma_1$, which will be discussed after we study slowness perturbations in next section.

Once the ray trajectories $[\sigma, q(\sigma), p(\sigma)]$ have been found, the travel time is found by direct integration along the trajectory:

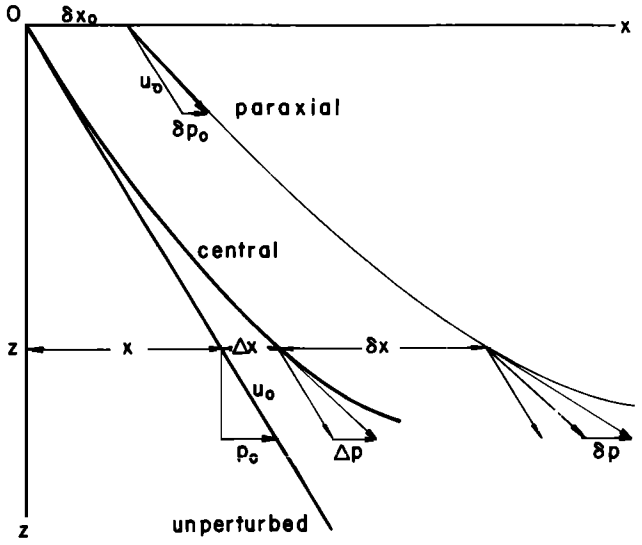

Fig. 1. Geometry of ray and paraxial ray parameters in a two-dimensional geographical coordinate system. Position is defined by $(\sigma, q) = (z, x)$, and the conjugate momentum p_x is simply the horizontal component of the slowness vector. The vertical slowness $p_z = -H$ is the Hamiltonian. The perturbed central ray is specified by the pertubation Δx with respect to the unperturbed central ray. A paraxial ray is specified by δx measured with respect to the perturbed reference ray.

$$\theta(\sigma, q, \sigma_0, q_0) = \int_{\sigma_0}^{\sigma} L[s, q(s), \dot{q}(s)]\, ds \tag{6}$$

where

$$L(\sigma, q, \dot{q}) = p\,\dot{q} - H = u(\sigma, q)\sqrt{h_\sigma^2 + h_1^2 \dot{q}_1^2 + h_2^2 \dot{q}_2^2}$$

is the Lagrangian function.

Geometrical spreading J, as well as Gaussian beams will be obtained from the so-called paraxial ray approximation, which is in fact a particular application of perturbation theory. Paraxial rays (see, for example, *Deschamps* [1972]) are those rays that propagate in the vicinity of a certain reference ray. They differ from this reference or central ray by small changes (perturbations) in position and slowness. Let us assume that we have already traced a reference ray in the medium with unperturbed slowness distribution $u_0(\sigma, q)$. We will denote with $y_0(\sigma) = [q_0(\sigma), p_0(\sigma)]$ the canonical vector trajectory of this reference ray. Paraxial rays are given by

$$q(\sigma) = q_0(\sigma) + \delta q(\sigma) \qquad p(\sigma) = p_0(\sigma) + \delta p(\sigma) \tag{7}$$

where $\delta y = [\delta q, \delta p]$ denote the perturbations in canonical vector (position and slowness). These perturbations are due to small changes $\delta q(\sigma_0)$ and $\delta p(\sigma_0)$ in the initial values of q and p. In order to trace paraxial rays we insert (7) into (5), and developping to first order, we find the following linearized system:

$$\begin{aligned} \delta\dot{q} &= \frac{\partial^2 H}{\partial p \partial q}\,\delta q + \frac{\partial^2 H}{\partial p^2}\,\delta p \\ \delta\dot{p} &= -\frac{\partial^2 H}{\partial q^2}\,\delta q - \frac{\partial^2 H}{\partial p \partial q}\,\delta p \end{aligned} \tag{8}$$

where all the derivatives of the Hamiltonian are calculated on the reference ray. Solutions to this linear system may be found by standard propagator techniques [*Gilbert and Backus*, 1966]. Let us first simplify the notation in (8) to

$$\delta\dot{y} = A_0(\sigma)\,\delta y \tag{9}$$

where δy denotes the perturbation of the canonical vector in phase space. A_0 is a matrix containing second-order partial derivatives of H, that appear in (8). Then, given the initial value $\delta y(\sigma_0)$, the subsequent evolution of the canonical vector in phase space is given by

$$\delta y(\sigma) = P_0(\sigma, \sigma_0)\,\delta y(\sigma_0) \tag{10}$$

where $P_0(\sigma, \sigma_0)$ is the propagator matrix, which is a solution of (9) with initial conditions $P_0(\sigma_0, \sigma_0) = I$, where I is the identity matrix.

When it is written in ray-centered coordinates (see Figure 2), the linear system (8) reduces to the so-called dynamic ray tracing system derived by *Cervený and Pšenčík* [1983] for the calculation of the Jacobian J.

Slowness Perturbation

Let us consider now a smooth perturbation of the model such that the slowness is slightly changed from u_0 to $u = u_0 + \Delta u$. Capital Δ will be used to denote perturbations due to the structure, and low case δ will be used for paraxial perturbations. The perturbation in slowness produces a corresponding perturbation of the Hamiltonian:

$$H(q, p, \sigma) = H_0(q, p, \sigma) + \Delta H(q, p, \sigma)$$

where $\Delta H = \partial H/\partial u\ \Delta u$ and H_0 is the Hamiltonian (4) for the reference slowness u_0 . In principle, ray tracing may be done by replacing the new Hamiltonian in the canonical equations (5) and solving this nonlinear system. To first order in Δu it is possible to linearize these equations considering only rays that deviate slightly from the reference

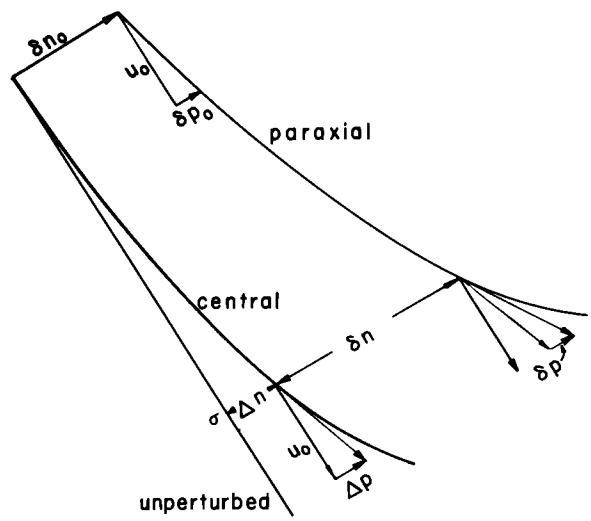

Fig. 2. Geometry of ray and paraxial ray parameters in a two-dimensional ray centered coordinate system. Position is defined by $(\sigma, q = n)$, and the conjugate momentum p_n is the components of slowness perpendicular to the central ray. The perturbed ray is specified by Δn measured with respext to the unperturbed central ray. A paraxial ray is specified by δn measured with respect to the perturbed central ray.

one. Introducing the perturbed canonical variables $q = q_0 + \Delta q$ and $p = p_0 + \Delta p$, we get:

$$\begin{aligned} \Delta\dot{q} &= \frac{\partial^2 H_0}{\partial q \partial p}\Delta q + \frac{\partial^2 H_0}{\partial p^2}\Delta p + \frac{\partial \Delta H}{\partial p} \\ \Delta\dot{p} &= -\frac{\partial^2 H_0}{\partial q^2}\Delta q - \frac{\partial^2 H_0}{\partial q \partial p}\Delta p - \frac{\partial \Delta H}{\partial q} \end{aligned} \tag{11}$$

where all the partial derivatives are calculated on the reference ray. Equations (11) form a linear system which has the same form as that of paraxial rays (9), except for the "source term" derived from ΔH. This term contains the gradient of the slowness perturbation, so that the perturbed rays behave like paraxial rays in the original medium which are continuously perturbed by the gradient of the slowness perturbation. We rewrite (11) in the simpler form

$$\Delta\dot{y}(\sigma) = A_0(\sigma)\,\Delta y(\sigma) + \Delta B(\sigma) \tag{12}$$

where Δy is the perturbation of the canonical vector y in phase space. The vector ΔB contains the Hamiltonian perturbations that act as sources in the linear system (11). The solution of (12) may again be found by standard propagator theory [*Gilbert and Backus*, 1966]:

$$\Delta y(\sigma) = P_0(\sigma,\sigma_0)\,\Delta y(\sigma_0) + \int_{\sigma_0}^{\sigma} P_0(\sigma,\tau)\,\Delta B(\tau)\,d\tau \tag{13}$$

where $\Delta y(\sigma_0)$ is the initial perturbation. If we are interested in a perturbation of the reference ray such that the initial conditions are conserved, then $\Delta y(\sigma_0) = 0$. In most of the applications in this paper we will be interested in the two-point boundary value problem so that the perturbed ray passes through both the source and the observer. In this case the boundary conditions for the solution of (13) are $\Delta q(\sigma_0) = 0$ and $\Delta q(\sigma_1) = 0$, where σ_1 is the end-point of the ray.

Finally, we may now look for the paraxial rays in the perturbed medium. These are rays that propagate in the vicinity of the perturbed ray $y = y_0 + \Delta y$. Careful analysis of the perturbation of the matrix A_0 in (9) leads to the following linear system of equations for the paraxial rays:

$$\delta\dot{y} = A(\sigma)\,\delta y \tag{14}$$

where δy is the perturbation measured from the perturbed reference ray. In (14) the matrix $A(\sigma) = A_0(\sigma) + \Delta A(\sigma)$, where $\Delta A = \Delta A_1 + \Delta A_2$, and

$$\Delta A_1 = \begin{bmatrix} \dfrac{\partial^2 \Delta H}{\partial q \partial p} & \dfrac{\partial^2 \Delta H}{\partial p^2} \\ -\dfrac{\partial^2 \Delta H}{\partial q^2} & -\dfrac{\partial^2 \Delta H}{\partial q \partial p} \end{bmatrix}$$

is a term due to perturbations in the Hamiltonian, while

$$\Delta A_2 = \left[\Delta q \frac{\partial}{\partial q} + \Delta p \frac{\partial}{\partial p}\right] \begin{bmatrix} \dfrac{\partial^2 H_0}{\partial q \partial p} & \dfrac{\partial^2 H_0}{\partial p^2} \\ -\dfrac{\partial^2 H_0}{\partial q^2} & -\dfrac{\partial^2 H_0}{\partial q \partial p} \end{bmatrix}$$

is due to the perturbation Δy of the reference central ray. All the derivatives are calculated on the original unperturbed ray y_0. Comparing with (10), we observe that the paraxials of the perturbed reference ray are given by

$$\delta y(\sigma) = P(\sigma,\sigma_0)\,\delta y(\sigma_0) \tag{15}$$

where $P(\sigma,\sigma_0)$ is the propagator of (14) which we may calculate by standard perturbation theory. To first order in the slowness perturbation, the Born approximation to P is

$$P(\sigma,\sigma_0) = P_0(\sigma,\sigma_0) + \int_{\sigma_0}^{\sigma} P_0(\sigma,\tau)\,\Delta A(\tau) P_0(\tau,\sigma_0)\,d\tau \tag{16}$$

Thus the paraxial rays $\delta y(\sigma)$ in the perturbed medium are given by an expression of the same form as that for the paraxials in the original medium (10), except that the propagator P has been perturbed as in (16). Let us remark again that $\delta y(\sigma)$ in (15) describes the trajectory of the paraxial ray referred to the perturbed trajectory of the central ray. Thus the total perturbation of a paraxial ray measured from the original unperturbed reference ray is $\delta y_T = \Delta y + \delta y$. This total perturbation is of little practical utility and will not be used in the following. We have purposely avoided using a separate notation for paraxial rays in the original medium and those in the perturbed one. This will simplify notation in next section.

Initial Conditions for Paraxial Rays and Beams

Let us consider first the solution of the original nonlinear ray tracing equations (5). When $[q(\sigma_0), p(\sigma_0)]$ are given as initial conditions at σ_0 we obtain the standard ray tracing problem. On the other hand, if $q(\sigma_0)$ and $q(\sigma_1)$ are specified at σ_0 and σ_1, respectively, we obtain a two-point boundary value problem [*Keller*, 1968]. Of these two problems the first is simpler since it consists of a relatively straightforward integration of (5). The two-point problem is more difficult since one has to find by some iterative procedure the initial slowness $p(\sigma_0)$ of the ray that passes through $q(\sigma_1)$. One of the most powerfull methods to solve this problem is the ray bending method [*Julian and Gubbins*, 1977] which is closely related to ray perturbation theory.

In the following we will assume that a set of rays has been traced by some numerical method in the unperturbed medium and that we have also constructed the propagator $P_0(\sigma,\sigma_0)$. If we introduce a small perturbation in the slowness of the medium, the ray trajectories will be perturbed by a small amount Δy which can be calculated using (13). The new ray trajectories are then given by $y(\sigma) = y_0(\sigma) + \Delta y(\sigma)$. In order to calculate geometrical spreading J and to generate Gaussian beams, we construct a beam around each of these ray trajectories using the paraxial approximation. Let us remark that since the paraxial ray equations in unperturbed (10) and perturbed media (15) have the same form, the results derived in the following will be applicable to both the initial and perturbed media.

Depending on the form of the initial perturbation $\delta y(\sigma_0)$, we will distinguish two fundamental types of paraxial rays:

1. If $\delta p(\sigma_0) = 0$, we have "plane wave" initial conditions in the nomenclature introduced by *Cervený et al.* [1982]. In fact, keeping $\delta p(\sigma_0) = 0$ while changing $\delta q(\sigma_0)$ generates a

beam of rays which are initially parallel to the reference ray. In homogeneous media the rays propagate along straight lines so that this boundary condition generates a true plane wave solution. In heterogeneous medium the rays bend and the actual form of the beam will depend on the particular coordinate system.

2. If $\delta q(\sigma_0) = 0$, we obtain point source initial conditions since the coordinates of the starting point of the paraxial is kept fixed. As $\delta p(\sigma_0)$ changes, a pencil of paraxial rays emanating from the source point is formed.

Let us introduce the following partition of the propagator matrix:

$$P(\sigma,\sigma_0) = \begin{bmatrix} Q_1 & Q_2 \\ P_1 & P_2 \end{bmatrix}$$

so that the paraxial solution is written as

$$\begin{bmatrix} \delta q \\ \delta p \end{bmatrix} = \begin{bmatrix} Q_1 & Q_2 \\ P_1 & P_2 \end{bmatrix} \begin{bmatrix} \delta q_0 \\ \delta p_0 \end{bmatrix} \tag{17}$$

where subscript 1 denotes "plane" wave initial conditions and subscript 2 denotes point source conditions. This nomenclature has been chosen following that introduced by *Cervený et al.* [1982] in the special case of ray centered coordinates.

In order to calculate geometrical spreading and Gaussian beams we have to generate a beam around the reference ray. We define a beam as a set of paraxial rays such that the perturbed initial conditions satisfy the linear relationship:

$$[\delta q_0] = \begin{bmatrix} \epsilon_1 & 0 \\ 0 & \epsilon_2 \end{bmatrix} [\delta p_0] \tag{18}$$

where the ϵ_i are complex scalars that define the initial form of the beam. For point sources $\epsilon_1 = \epsilon_2 = 0$, while for "plane" initial conditions $\epsilon_1 = \epsilon_2 = \infty$. For Gaussian beams the ϵ are complex with a negative imaginary part [*Cervený et al.*, 1982; *Madariaga*, 1984]. Although the ϵ_i may be different, for simplicity we will assume that $\epsilon_1 = \epsilon_2 = \epsilon$. In that case, (17) gives:

$$\begin{aligned} \delta q &= (\epsilon Q_1 + Q_2)\delta p_0 = (Q_1 + \epsilon^{-1} Q_2)\delta q_0 \\ \delta p &= (\epsilon P_1 + P_2)\delta p_0 = (P_1 + \epsilon^{-1} P_2)\delta q_0 \end{aligned} \tag{19}$$

and the perturbed position and slowness are linearly related by $\delta p = M\,\delta q$, where

$$M(\sigma) = (\epsilon P_1 + P_2)(\epsilon Q_1 + Q_2)^{-1} \tag{20}$$

The matrix $M(\sigma)$ is related to the curvature of the wave front, but except for ray-centered coordinates, this relation is not simple. Let us note that the initial value of $M(\sigma_0) = \epsilon^{-1} I$.

We may now calculate the travel time θ for the paraxial rays in a beam. From (6) we find

$$\theta(\sigma, q + \delta q) = \theta(\sigma, q) + \int_0^{\delta q} (p + \delta p)\, d\delta q \tag{21}$$

where $[q(\sigma), p(\sigma)]$ is the perturbed canonical vector of the reference ray in phase space. To first order in the slowness perturbation, the travel time $\theta(\sigma, q)$ along the perturbed reference ray is

$$\theta(\sigma, q) = \theta_0(\sigma, q_0) + \int_{\sigma_0}^{\sigma} \Delta u(\sigma, q_0)\, d\sigma + p_0(\sigma)\, \Delta q(\sigma) \tag{22}$$

with the obvious notation that $\theta_0(\sigma, q_0)$ is the travel time along the unperturbed reference ray $[q_0(\sigma), p_0(\sigma)]$. We can use $\delta p = M\,\delta q$ in (21) to obtain

$$\theta(\sigma, q + \delta q) = \theta(\sigma, q) + p\,\delta q + \frac{1}{2}\delta q^{\dagger} M\,\delta q \tag{23}$$

where dagger denotes transposition. A similar expression for cartesian coordinates was obtained by *Madariaga and Papadimitriou* [1985].

We may now determine geometrical spreading. Let us consider again a beam around the reference ray $q(\sigma)$. At $\sigma = \sigma_0$, an elementary beam cross section is defined by

$$dS(\sigma_0) = \|\delta q_1 \times \delta q_2\| \cos\phi_0 \tag{24}$$

where δq_1 and δq_2 are two arbitrary but nonparallel paraxial ray vectors; ϕ_0 is the angle between the tangent to the perturbed ray and the unit vector perpendicular to δq_1 and δq_2. We now follow this elementary cross section along the reference ray. At a point σ the cross section $dS(\sigma)$ is again given by (24) where the δq_i are updated using (19) and ϕ_0 is replaced by ϕ, the local angle between the normal to the surface dS and the tangent to the perturbed ray. Thus

$$dS(\sigma) = \det(Q_1 + \epsilon^{-1} Q_2) \frac{\cos\phi}{\cos\phi_0}\, dS(\sigma_0)$$

and, since by definition the Jacobian $J = dS(\sigma)/dS(\sigma_0)$, we find

$$J(\sigma, \sigma_0) = \det(Q_1 + \epsilon^{-1} Q_2) \frac{\cos\phi}{\cos\phi_0} \tag{25}$$

This expression for geometrical spreading has to be modified for point sources since in this case $\epsilon \to 0$, so that J becomes singular. This difficulty is solved incorporating ϵ in the excitation vector $A(\omega)$ (see equation (1)). In this case, $J(\sigma, \sigma_0) = \det(Q_2) \cos\phi \cos^{-1}\phi_0$, which is the expression most commonly used in seismology.

Using (23) for θ and (25) for J in (1), we find the most general expression for a beam in the vicinity of a reference ray. These expressions are valid both in unperturbed and perturbed media. In the former case, Q_i and P_i are obtained partitioning the unperturbed propagator $P_0(\sigma, \sigma_0)$, in the latter from the perturbed propagator (16). Gaussian beams are a particular case of a beam in which ϵ is complex and its imaginary part is less than zero. Thus we can use the theory presented here to solve perturbation problems for rays and beams.

Homogeneous Reference Model

The general results obtained above take a very simple form when the reference medium is homogeneous. In the following we will discuss this particular case with some detail because the examples that will be presented later in the paper are perturbations with respect to a homogeneous reference medium. A reference ray in the unperturbed homogeneous medium is of course a straight line. The ray-centered coordinate system associated with this reference ray is Cartesian. We take as independent parameter σ, the abscissa

along the reference ray, and q and p are two-dimensional vectors defining position and slowness in plane perpendicular to the reference ray. We can write the reduced unperturbed Hamiltonian in this coordinate system as

$$H_0(p) = -\sqrt{u_0^2 - p^2}$$

which is independent of σ and q. The ray tracing equations (5) take the simple form

$$\dot{q} = \frac{p}{\sqrt{u_0^2 - p^2}} \qquad \dot{p} = 0 \tag{26}$$

Let us denote with $y_0(\sigma) = (q_0, p_0)$, the four-dimensional canonical vector representing a ray at the abscissa σ. The reference ray in the unperturbed medium is simply $y_0(\sigma) = 0$. The paraxial rays in the unperturbed medium are obtained by a small perturbation of the initial conditions, $\delta y(\sigma_0)$. Developing equations (26) up to first order in δy, we obtain the linear differential system,

$$\delta \dot{y} = A_0(\sigma)\, \delta y \tag{27}$$

with

$$A_0 = \begin{bmatrix} 0 & v_0 I \\ 0 & 0 \end{bmatrix}$$

where v_0 is the velocity in the reference medium and I is the identity matrix. The solution of (27) is

$$\delta y(\sigma) = P_0(\sigma, \sigma_0)\, \delta y(\sigma_0)$$

where P_0 is the propagator matrix given by

$$P_0(\sigma, \sigma_0) = \begin{bmatrix} I & v_0(\sigma - \sigma_0) I \\ 0 & I \end{bmatrix} \tag{28}$$

This is the same as used by *Cervený et al.* [1982] in their study of Gaussian beams in a homogeneous medium.

Let us consider now the perturbed slowness distribution $u(\sigma, q) = u_0 + \Delta u(\sigma, q)$, where the perturbation Δu is a smooth function which is assumed to have continuous second-order derivatives. In order to determine the perturbation of ray trajectories we expand the new reduced Hamiltonian, $H(\sigma, q, p) = -\sqrt{u^2(\sigma, q) - p^2}$, in a Taylor series around the reference ray in the homogeneous medium. Keeping terms up to first order in Δu,

$$H(\sigma, q, p) = H_0(p) + \Delta H(\sigma, q, p)$$

where

$$\Delta H(\sigma, q, p) = -\frac{u_0}{\sqrt{u_0^2 - p^2}}\, \Delta u(\sigma, q)$$

The perturbed rays are solutions of (12) which for a homogeneous reference medium becomes

$$\Delta \dot{y} = A_0(\sigma)\, \Delta y + \Delta B(\sigma) \tag{29}$$

with

$$\Delta B(\sigma) = \begin{bmatrix} 0 \\ \frac{\partial \Delta u}{\partial q} \end{bmatrix}$$

The perturbation $\Delta y(\sigma)$ of the reference ray is obtained solving (29) by the propagator technique

$$\Delta y(\sigma) = P_0(\sigma, \sigma_0)\, \Delta y(\sigma_0) + \int_{\sigma_0}^{\sigma} P_0(\sigma, \tau)\, \Delta B(\tau)\, d\tau \tag{30}$$

where P_0 is the propagator (28). The perturbation solution (30) may be used to solve a number of initial and boundary value problems. For instance, if we want to trace a perturbed ray with the same initial conditions as the reference ray, we would take $\Delta y(\sigma_0) = 0$. If we want to calculate the perturbed ray passing through the same source and observer as the unperturbed ray, then we take $\Delta q(\sigma_0) = \Delta q(\sigma) = 0$ and use (30) to calculate $\Delta p(\sigma_0)$.

The paraxial rays with respect to the perturbed reference ray are given by (14), where $A(\sigma) = A_0(\sigma) + \Delta A(\sigma)$ with

$$\Delta A(\sigma) = \begin{bmatrix} 0 & -v_0^2 \Delta u I \\ \Delta U & 0 \end{bmatrix} \tag{31}$$

Here, $\Delta U(\sigma) = \partial^2 \Delta u / \partial q \partial q$ is the matrix of second-order partial derivatives of Δu in the q plane. As shown in (15), the perturbed paraxial rays are given by

$$\delta y(\sigma) = P(\sigma, \sigma_0)\, \delta y(\sigma_0) \tag{32}$$

where P is the perturbed propagator. To first order in Δu,

$$P(\sigma, \sigma_0) = P_0(\sigma, \sigma_0) + \int_{\sigma_0}^{\sigma} P_0(\sigma, \tau) \Delta A(\tau) P_0(\tau, \sigma_0)\, d\tau \tag{33}$$

where $\Delta A(\sigma)$ is given explicitly by (31) and P_0 by (28). Once the perturbed trajectory of the central ray $\Delta y(\sigma)$ and the propagator $P(\sigma, \sigma_0)$ have been calculated, the geometrical spreading may be calculated using (25) and the complex travel time θ of a Gaussian beam is calculated using (23). Thus we have all the elements to calculate the ray field in the perturbed medium.

A Simple Example: A Gaussian Spherical Inclusion

We now apply the method developed in the previous sections to a very simple two-dimensional acoustic wave propagation problem. As described in Figure 3, the velocity distribution consists of a homogeneous medium of velocity $v_0 = 3$ km/s containing a smooth circular low-velocity zone. The total velocity is

$$v = v_0 - \Delta v_p \exp\left[-\frac{1}{2}\left(\frac{d}{D_0}\right)^2\right]$$

where

$$d^2 = (x - 5)^2 + (z - 5)^2$$

The peak velocity perturbation is $\Delta v_p = 0.5$ km/s, and the effective radius of the low velocity perturbation is $D_0 = 1$ km. As shown in Figure 3, the main effect of the low-velocity zone is to focus the rays behind the perturbation. In Figure 3 we compare the rays traced using the perturbation method with the results of exact ray tracing. The exact rays shown with continuous line were obtained by numerical integration of the ray equations with the Runge Kutta method. The results of perturbation theory are shown by the dotted lines.

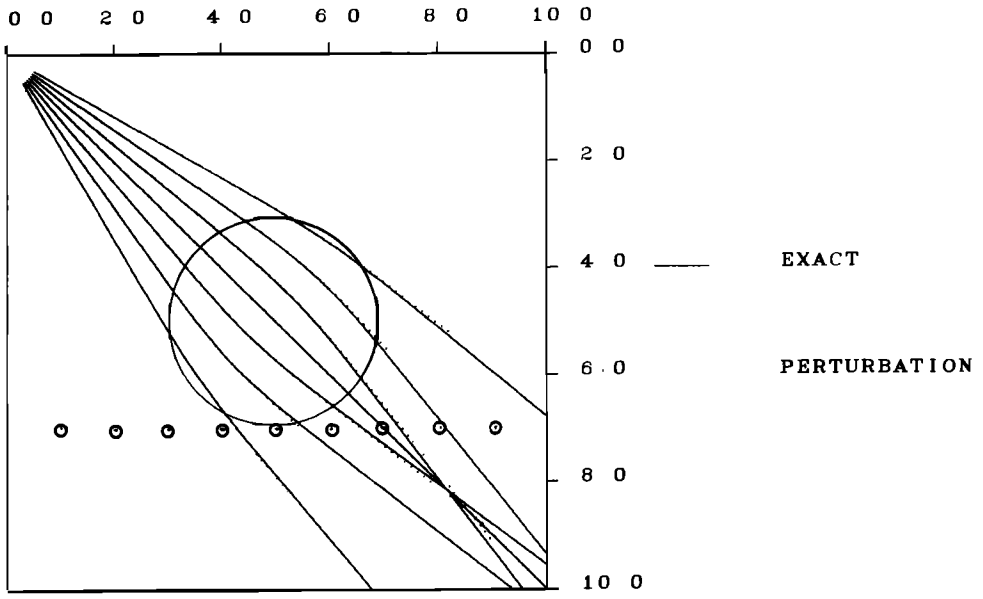

Fig. 3. Geometry of the spherical inclusion model and ray tracing through it. The perturbation in velocity is centered at (5,5) and has a Gaussian profile with effective radius 1 Km. Source is located at the origin. Exact rays (solid lines) traced by numerical integration of the ray equations may be compared with rays calculated by perturbation theory (dotted lines).

The error in the perturbation method is larger for the rays that travel near the center of the velocity perturbation. This is clearly seen in Figure 3, where the perturbed rays get farther and farther from their respective reference ray once they cross the perturbation.

Let us compare now synthetics calculated by exact ray theory and by the perturbation method. We consider an acoustic line source located at the origin. The ray field is given by

$$u(\sigma,\omega) = if(\omega)\sqrt{\frac{v(\sigma)}{v(\sigma_0)\,J(\sigma)}}\exp[i\omega\theta(\sigma,\sigma_0)] \qquad (34)$$

where $f(\omega)$ is the source time function defined below.

We compare in Figure 4 the synthetics calculated with the perturbation method with those calculated using exact ray tracing. The receivers are located on an horizontal line at a depth of 7 km (see Figure 3). For the calculation of exact seismograms we would have in principle to solve a series of two-point boundary value problems in order to trace rays from source to receiver. We used instead the paraxial technique described by *Cervený and Pšenčík* [1983] in order to interpolate between neighboring rays. All the calculations were carried out in two dimensions; however, in order to simulate a three-dimensional medium we chose a source time function of the following form

$$f(t) = \frac{d}{dt}\left[\frac{H(t)}{\sqrt{t}}\right]$$

Moreover, in order to avoid singularities in the time domain, the synthetics were smoothed with the sampling function

$$s(t) = \frac{\Delta t_1}{t^2 + \Delta t_1^2} \qquad (35)$$

where $\Delta t_1 = 0.02$ s. The perturbation in propagation times are of the order of 0.15 s, which is much greater than the source function width. The perturbation method gives these time perturbations with a maximum relative error of 2 % at the studied receivers. The amplitude of the signals are also very well modeled by the perturbation method even at the horizontal range of 7 km, where amplitude has doubled with respect to its value in the unperturbed medium. The results for travel time are quite good as we could expect from Fermat's principle [*Aki et al.*, 1977]. The results for amplitude are also quite encouraging at least for smooth perturbation like the one considered in this example.

We also applied the perturbation method to the computation of a single Gaussian beam traveling in the perturbed medium. We considered an acoustic source located at x = 5 km on the surface of the model (z = 0). We studied the Gaussian beam that leaves the source with a vertical initial direction. Because of the symmetry of the perturbation

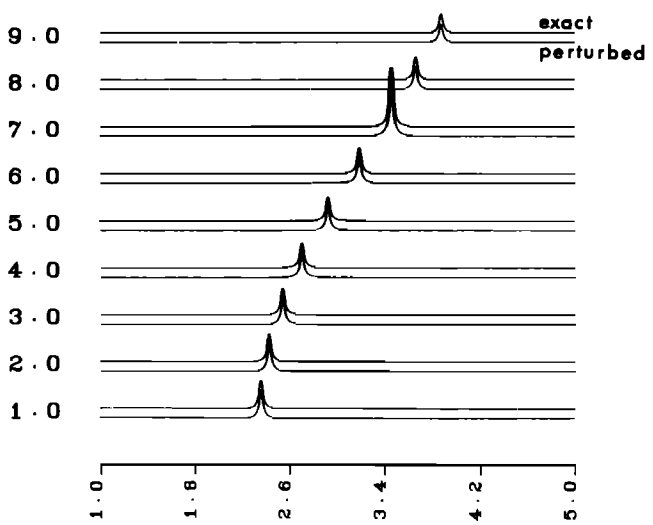

Fig. 4. Seismograms calculated at a series of receivers located at a depth of 7 km in the model of Figure 3.

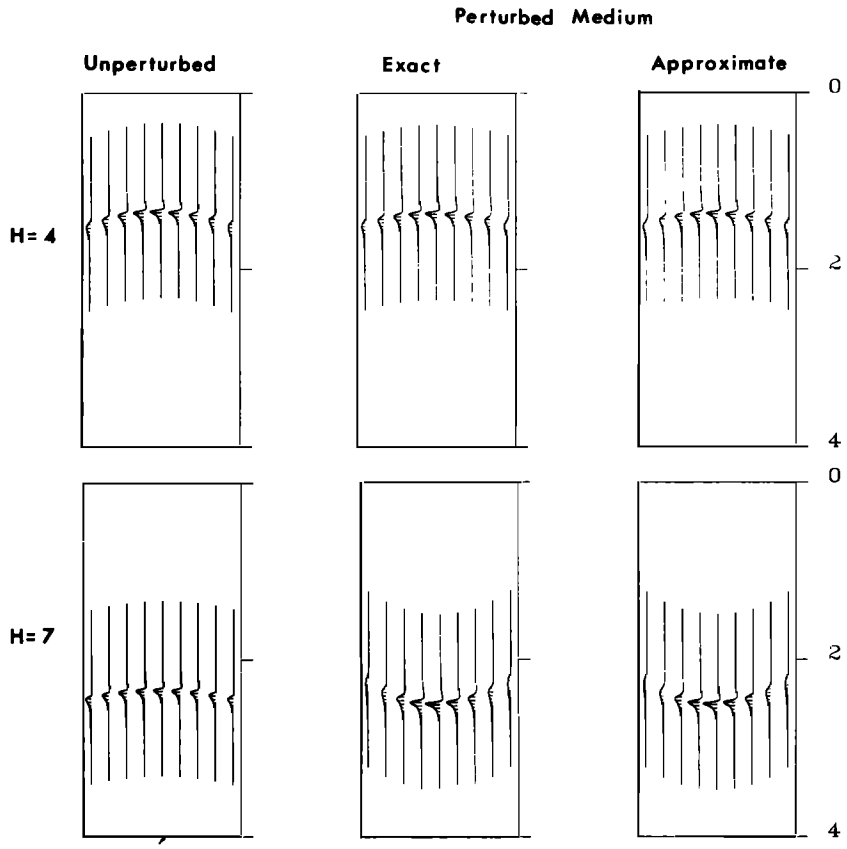

Fig. 5. Comparison of a Gaussian beam calculated by exact and approximate propagators. The central ray of the beam is vertical and the source is at (5,0) in the model of Figure 3. The top panel shows the Gaussian beams computed at depth of 4 km, while the bottom one shows the beams at 7 km depth. The traces are calculated at horizontal distance increments of 0.5 km and are centered around x = 5 km. The first column shows the beams in a homogeneous reference medium. The second column shows beams calculated exactly in the perturbed medium. The last column shows the results obtained for the perturbed model with perturbation theory.

about this line, the central ray remains straight in the perturbed medium, i. e. $\Delta y(\sigma) = 0$. The Gaussian beam is given in the time domain by

$$u(\sigma, q, t) = \sqrt{\frac{v(\sigma)}{v(\sigma_0)\, J(\sigma)}}\,\mathrm{Re}\left[\frac{1}{t - \theta(\sigma, q)}\right] \qquad (36)$$

Where J and θ are computed using (25) and (23), respectively. The complex beam parameter was chosen as $\epsilon = -i$. In order to eliminate the problem with the pole at $t = \theta$ when $\delta q = 0$, the synthetics were smoothed with the sampling function $s(t)$ defined in (35), in which $\Delta t_1 = 0.02$ s. As shown by *Madariaga and Papadimitriou* [1985], this is obtained by adding a small imaginary part, equal to $i\Delta t_1$, to θ.

We computed the displacement due to this Gaussian beam on two horizontal lines located at $z = 4$ km and $z = 7$ km. The receivers were distributed every 0.5 km, symmetrically with respect to the central ray. At the center in Figure 5 we show the results obtained for the perturbed medium using exact numerical solutions for the central ray and the propagator P. For reference, we show at left in Figure 5 the Gaussian beam in the unperturbed homogeneous medium. One can see the effect of the low-velocity zone on the amplitude and the curvature of the wave front which changes sign after crossing the perturbation: the concave wave front at 4 km becomes convex when it crosses the heterogeneity. Finally, at right in Figure 5 we show the Gaussian beams in the perturbed medium obtained with the perturbation method. The travel time and amplitude are practically identical with those of the center column, obtained with the exact method.

Reflection and Transmission of Perturbed Rays and Beams

The interaction of perturbed rays with an interface was not considered in the general formulation presented above. In the following we discuss this problem in the case in which the interface remains fixed. The more difficult problem of the perturbation of the shape of the interface itself will be considered elsewhere. Let M be the point of incidence of the perturbed ray on the interface. In order to propagate the transmitted and reflected rays away from the interface we have to change the reference unperturbed ray. We choose, as the new reference ray, the ray of the unperturbed medium that is tangent at M to the direction of emergence of the reflected or transmitted ray in the perturbed medium (Figure 6). The new reference ray is therefore connected at M with the incident perturbed ray by the Snell-Descartes law expressed in the perturbed medium,

$$\frac{\sin\phi_i}{c_i} = \frac{\sin\phi_t}{c_t}$$

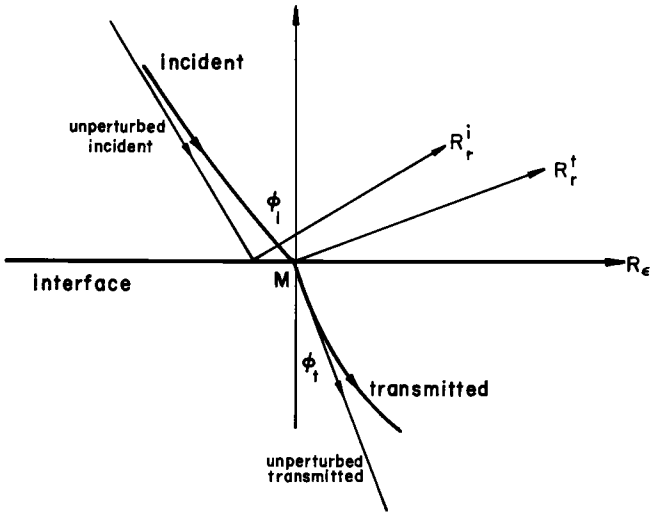

Fig. 6. Geometry of the interaction of a ray with an interface and the different coordinate systems used to continue the propagators across the interface. M is the point of incidence on the interface of the perturbed central ray.

where c_i and c_t are the perturbed velocities at M of the incident medium and of the medium in which the reflected or transmitted wave propagates. The initial canonical vector representing the perturbed transmitted or reflected ray in this new reference system is simply $\Delta y^t(\sigma) = 0$.

Complex reflection coefficients are used at each interface. The amplitude of the reflected-transmitted wave is calculated by multiplying the complex amplitude $A(\omega)$ in (1) by the reflection coefficient. Complex amplitudes are taken into account in the inverse Fourier transforms as described by *Madariaga and Papadimitriou* [1985].

The determination of initial conditions for the reflected or transmitted paraxial rays are somewhat more difficult to obtain. Let us introduce the following coordinate systems:

R_τ^i and R_τ^t are coordinate systems at M centered around the unperturbed incident and reflected-transmitted ray, respectively.

R_ϵ is the Cartesian coordinate system of the interface at M.

The δy^i and δy_ϵ^i are the canonical vectors representing an incident paraxial ray in R_τ^i and R_ϵ, respectively.

The δy^t and δy_ϵ^t are the vectors of the corresponding reflected or transmitted paraxial ray in R_τ^t and R_ϵ.

Continuity of the phase at the interface is equivalent to the continuity of the paraxial canonical vectors expressed in the R_ϵ system, i. e.,

$$\delta y_\epsilon^i = \delta y_\epsilon^t \tag{37}$$

We can now relate a canonical vector in the R_ϵ coordinates to that in ray centered coordinates R_τ^i or R_τ^t by canonical transformations as discussed in the Appendix. Denoting M_i and M_t the canonical transformation matrices from R_τ^i to R_ϵ and from R_ϵ to R_τ^t, respectively, we obtain the following relations:

$$\begin{aligned} \delta y_\epsilon^i &= M_i(R_\epsilon, R_\tau^i)\, \delta y^i \\ \delta y^t &= M_t(R_\tau^t, R_\epsilon)\, \delta y_\epsilon^t \end{aligned} \tag{38}$$

Finally, using the continuity condition (37), we obtain

$$\delta y^t = M_t M_i \quad \delta y^i \tag{39}$$

When there is no slowness perturbation, the matrix product $M_t M_i$ reduces to the transformation matrix obtained by *Cervený and Pšenčík* [1984] using a phase matching method. The canonical perturbation vector δy^t determined from (39) is used as the new initial condition to propagate the reflected-transmitted paraxial ray away from M.

As an example, we apply this method to the model described in Figure 7. The reference medium consists of three homogeneous layers with velocities of 3 km/s, 5 km/s and 6 km/s, respectively. We perturb this medium inserting a smooth circular high-velocity zone of the form

$$\Delta v = v_1 \exp\left[-\frac{1}{2}\left(\frac{d}{D_0}\right)^2\right]$$

where $v_1 = 1$ km/s, $d^2 = (x-5)^2 + (z-5)^2$, and $D_0 = 1$ km.

The acoustic source was located on the free surface as shown in Figure 7. We traced in the perturbed medium the rays reflected on the second interface by the perturbation method. These rays are shown in Figure 7. One can see that the rays are deflected by the perturbation and spread out as they cross it. The same rays calculated using the Runge Kutta method are undistinguishable from those calculated by perturbation.

We calculated synthetics at a number of receivers distributed on a vertical line, located 7 km away from the source as indicated by the vertical line on Figure 7. The source function was the derivative of Ricker's function:

$$s(t) = \exp\left[-\frac{1}{2}\left(\frac{t}{\Delta t_1}\right)^2\right]$$

where $\Delta t_1 = 0.02$ s. In Figures 8 and 9 we show the synthetics calculated by classical ray theory in the reference

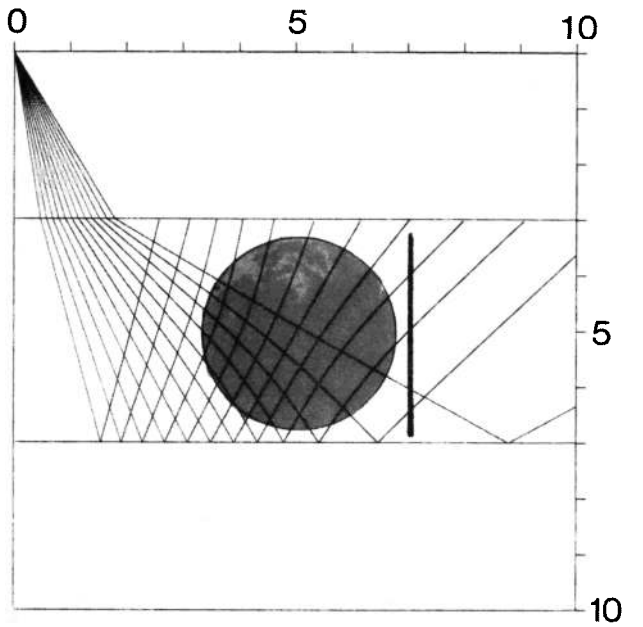

Fig. 7. Geometry of a simple layered structure with a spherical inclusion and rays traced with the perturbation method. The model is the same as that in Figure 3, except that the unperturbed medium is a layered structure. The source is located at the origin. Rays traced with the exact method are practically identical to those shown here.

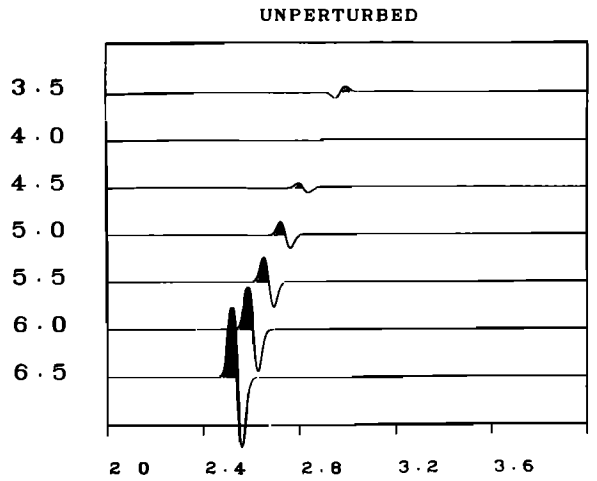

Fig. 8. Synthetic seismograms calculated on a vertical set of receivers located at a horizontal distance of 7 km from the origin in the model of Figure 7 without the circular high-velocity inclusion.

medium and in the perturbed medium, respectively. They were computed using (34) modified to take into account the reflection and transmission coefficients at the interfaces. One can see the important effect of the perturbation on the synthetics, in particular the vertical shift of the nodal (almost zero amplitude) trace. This is due to the interaction of the perturbed rays with the interface rather than to perturbation of geometrical spreading. The node in amplitude is due to the particular form of the reflection coefficient on the second interface. As the rays cross the high-velocity inclusion, they get deflected and the angle of incidence to the reflector decreases. Since we are close to critical angle, the reflection coefficient is very sensitive to these changes in incidence angle. At the bottom in Figure 9 we show the synthetics calculated by the perturbation method. The comparison with the results given by classical ray theory is excellent.

Iterative Perturbation

The perturbation theory that we presented above has the obvious limitation that the perturbed ray may not move too far away from the unperturbed one in phase space, i. e. Δy should be small. It is possible to reduce errors when Δy increases by changing the reference ray as soon as the perturbed ray ceases to verify the following criterion: $\Delta y < \Delta y_{max}$, where Δy_{max} is a predetermined maximum perturbation which should obviously depend on the acceptable error and the spatial scale of the velocity perturbation. The new reference ray is chosen as that ray of the unperturbed medium that is tangent to the perturbed ray at the point M where Δy becomes larger than Δy_{max}. The solution of the differential system (12) may be continued beyond M using $\Delta y = 0$ as new initial condition. Initial conditions for the calculation of the propagator P are obtained by canonical transformations connecting the old and the new coordinate systems. The change of reference ray is essential in order to maintain accuracy when the ray deviates too far from its unperturbed trajectory.

A Three-Dimensional Example: Amplitude Modeling in the Mont Dore Volcano

In this section we use the perturbation method to trace rays and calculate amplitudes in a complex three-dimensional medium. The model is derived from the velocity structure under the Mont Dore volcano structure obtained by inversion of PmP arrival times by *Nercessian et al.* [1984]. Nercessian et al. used critically reflected *PmP* waves in order to find velocity perturbations inside a fixed volume under the volcano. Travel time residuals were inverted to three-dimensional velocity perturbations using an inverse method. The calculated velocity perturbations reach values of the order of 20 % which are in fact quite large. In Figure 10 we show a very simplified plane view and two cross sections through the structure inverted by *Nercessian et al.* [1984]. Further details about the structure may be found in their paper. In order to eliminate artificial velocity discontinuities at the boundary of the studied volume, we enlarged it with a boundary layer in which the velocity tends gradually to the reference velocity of the unperturbed medium.

The velocity of the medium is described by interpolating between the velocity defined on a regularly spaced grid of nodal points. We used the cubic splines proposed by *Thomson and Gubbins* [1982] in order to interpolate the velocity between grid points. These functions permit a local representation in which the velocity at one point is related to the parameters of the 64 nearest nodes. The main advantage of this interpolation is its computational speed: The kernels are independent of the values of the velocity at nodes, and consequently, they need to be computed only once. Moreover, instead of using the original procedure of *Thomson*

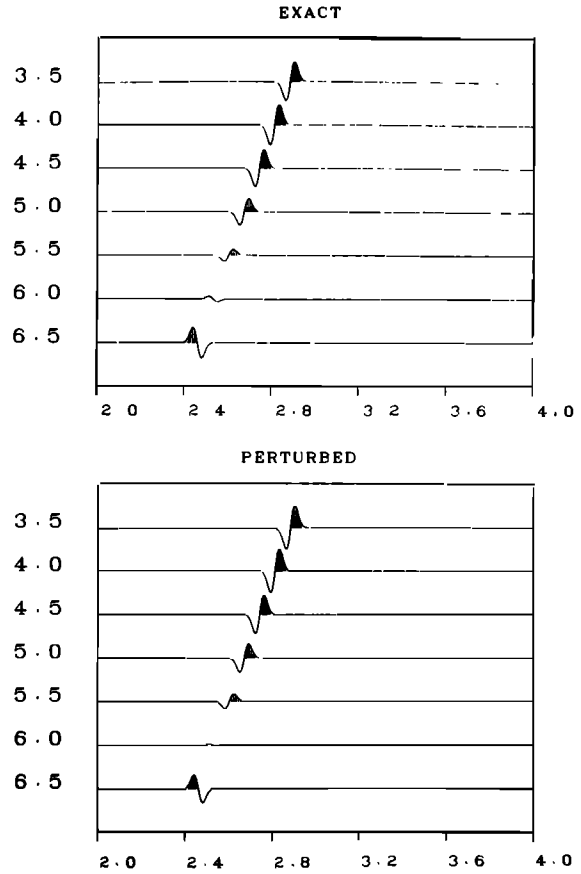

Fig. 9. Synthetic seismograms calculated on a vertical set of receivers located at a horizontal distance of 7 km from the origin in the model of Figure 7. At the top we show the seismograms calculated exactly by numerical integration, while at the bottom are those obtained by perturbation theory.

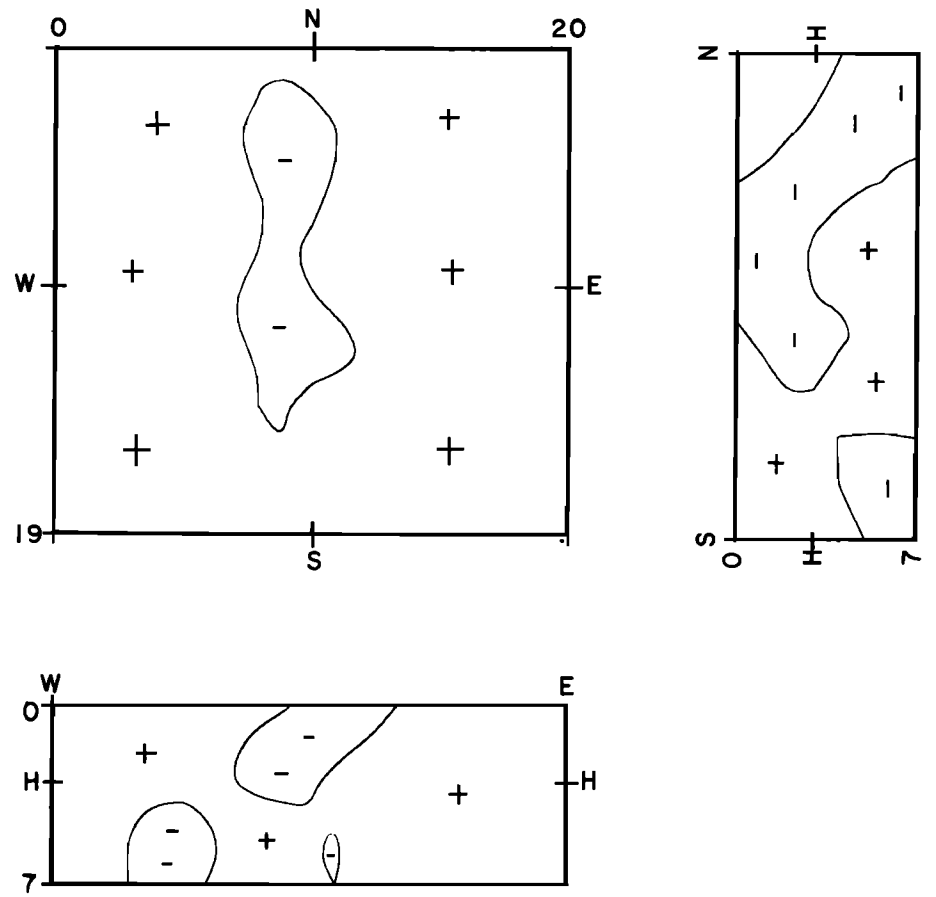

Fig. 10. Geometry of the velocity structure under the Mont Dore volcano as derived by *Nercessian et al.* [1982] from travel time inversion. On top right is a horizontal section at a depth of 4 km. The line indicates the zero perturbation line. Plus and minus signs indicate the sense of the perturbation of the velocity. To the right and bottom are two vertical cross sections in the NS and EW directions, respectively. More details of the model may be found in the original paper by Nercessian et al.

and Gubbins [1982], we evaluated the kernels and their first and second derivatives analytically. In this form we estimate velocity and its first and second derivatives which are continuous as required by ray theory.

The source was located 60 km south from the studied volume. We calculated the rays that reflect on the Moho and go through the perturbed volume. Figure 11 shows a perspective plot of a few rays selected among those that traverse the heterogeneous structure under the Mont Dore volcano. A severe distortion of the rays is observed. In order to give a more quantitative assessment of the effect of the structure we show on Figure 12 the intersection of a number of rays with several horizontal planes. In order to improve the visibility of these maps we connected with continuous lines the points corresponding to rays that leave the source on the same vertical plane (constant azimutal angle). These lines, which would be straight in a homogeneous medium, become distorted when the rays cross the different perturbation patches, especially in the low-velocity zone located near the center of the model. A caustic is formed in the vicinity of the maximum gradients that separate the low-velocity body from the higher-velocity rocks outside the old magma chamber. The reference ray in the unperturbed medium was changed every time the cartesian norm of the ray perturbation in configuration space $\|\Delta q\|$ failed to pass the criterion $\|\Delta q\| < 0.01$. The ray crossing maps obtained by a numerical integration of the ray equations in three dimensions are practically identical to those shown on Figure 12 and will not be shown here.

We then computed the amplitudes obtained at several points on the horizontal plane located at 2 km depth. In order to avoid interpolation the amplitudes were calculated along profiles defined by the crossing points of the rays indicated on Figure 13. In this figure there are fewer equal takeoff azimuth lines than in Figure 12, but we have plotted many more points along each profile. The amplitudes were calculated using the perturbed Jacobian J determined by

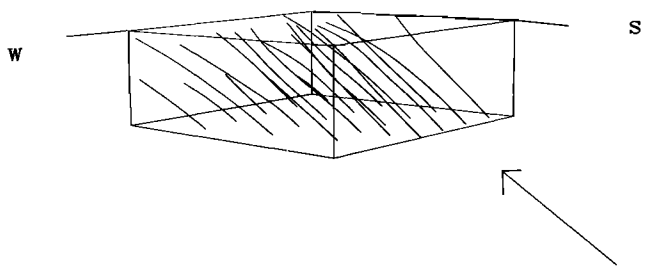

Fig. 11. Stereo view of selected rays traversing the structure of the Mont Dore volcano. The arrow indicates the direction of incidence of the rays.

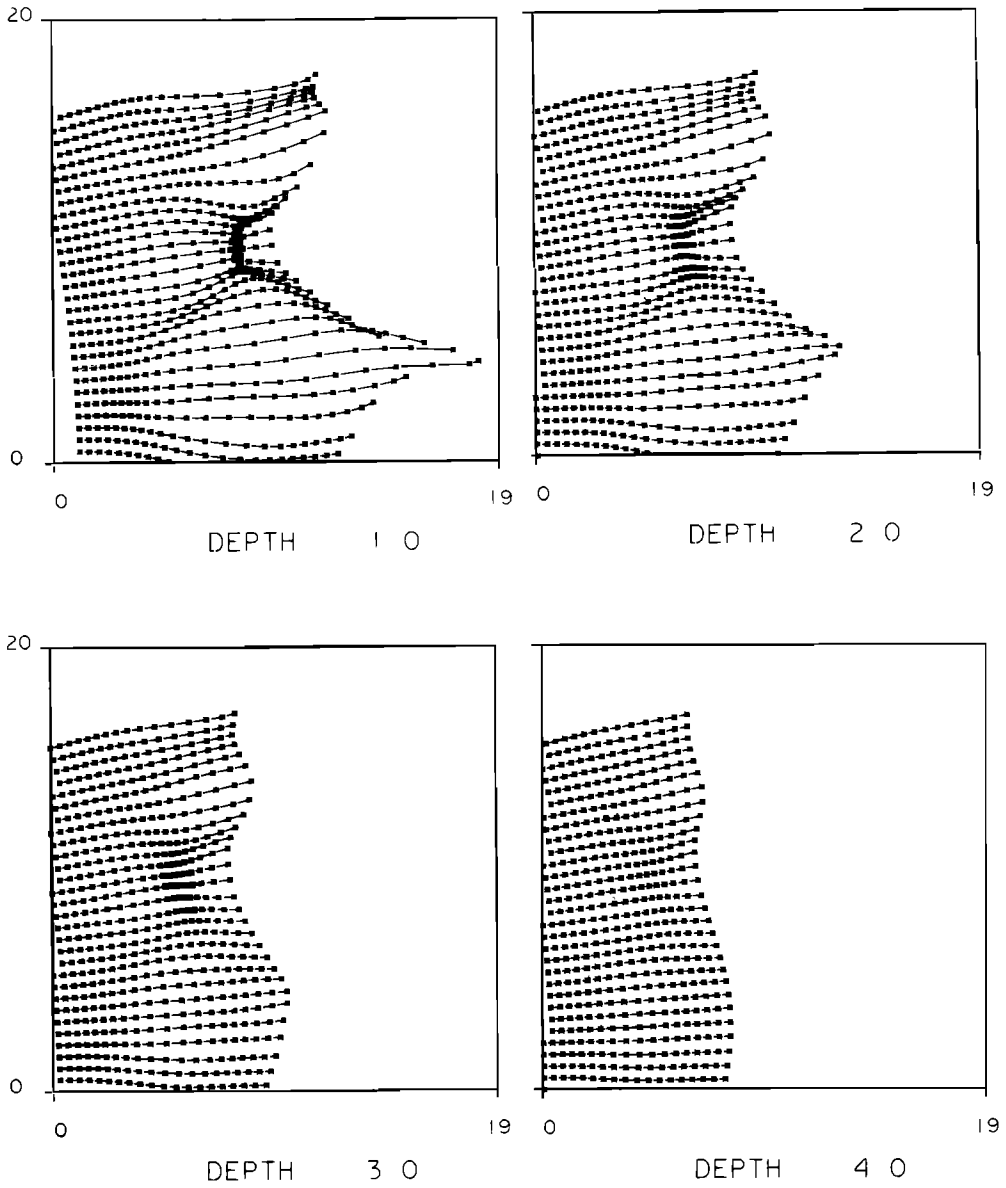

Fig. 12. Ray intersections with successive horizontal planes across the heterogeneous volume under the Mont Dore volcano. Each line in the cross sections joins rays that left the source with a constant azimuth. In a homogeneous medium these lines should be straight. We observe near the center, at shallow depths, the formation of caustics connected with the presence of a low-velocity zone.

means of the propagator (16). Since the propagator has an analytic expression along the segment of the ray path between the source and the perturbed volume, the total propagator was computed using the analytical results as initial values for the propagation inside the volume. In Figure 14 we show the amplitude profiles obtained along the ray crossing lines shown in Figure 13. On the left of Figure 14 we show exact results calculated by numerical integration of dynamic ray tracing equations, while on the right is the corresponding perturbation result. The amplitudes are very large around the edges of the low-velocity intrusion outlined in Figure 10. At the bottom of Figure 14 there are two loops in the amplitude curves which are due to the formation of a caustic. In the vicinity of this caustic the results obtained with exact or perturbed ray theory are not valid and the Gaussian beam summation should be used. The amplitudes are very well modeled by the perturbation method except near the caustics, where ray theory breaks down anyway.

Discussion

We have applied perturbation theory to the calculation of complete ray fields and synthetic seismograms in three-dimensional heterogeneous media. We addressed two principal problems. First, the calculation of paraxial rays, i. e., rays that propagate in the vicinity of another ray. Fast calculation of these rays is essential to compute geometrical spreading and to generate Gaussian beams. Second, we calculated the effect of a smooth perturbation of the slowness of the medium upon the reference rays and their paraxials. A number of important problems in seismic ray synthesis

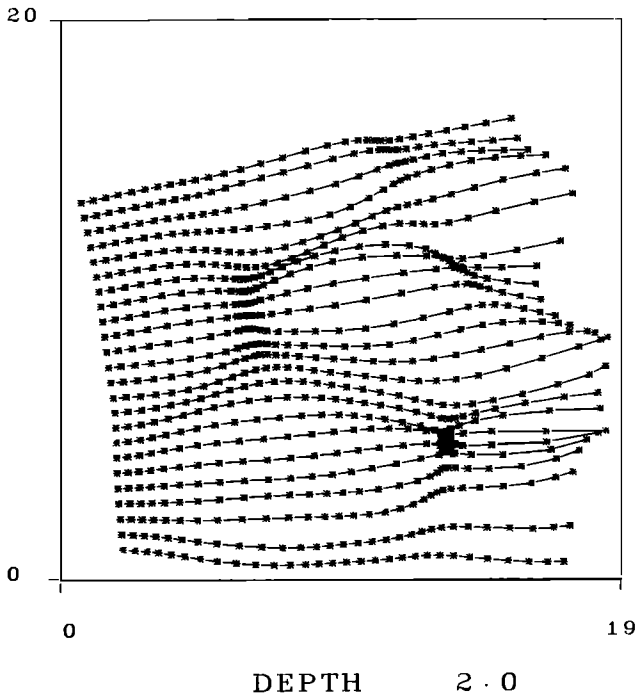

Fig. 13. Lines of ray crossing along which we calculated the amplitude profiles of Figure 14. These lines are a subset of those presented in Figure 12, but they are longer.

may be approached with the techniques developed in this paper.

In order to keep theory as independent as possible from the particular geometry under consideration, we used a Hamiltonian formalism in general orthogonal curvilinear coordinates. The Hamiltonian chosen here differs from that used most commonly in seismology (see, for example, *Chapman and Drummond* [1982] and *Cervený* [1985]) by the choice of one of the spatial coordinates as the independent parameter of the problem. In usual seismic applications an independent parameter is used along the rays. In the examples shown in this paper we used simple geometries, but our results can be easily extended to deal for instance with spherical geometry and the tracing of surface waves (see also *Woodhouse and Wong* [1986]).

Tracing of paraxial rays in general heterogeneous media and coordinates is one of our main results. As shown by *Cisternas et al.* [1984], *Madariaga and Papadimitriou* [1985] and V. M. Babich et al. (preprint 1985), Gaussian beams may be directly derived from ray theory without solving the complex parabolic equation introduced by *Cervený et al.* [1982]. This paper extends those results to arbitrary coordinate systems and to perturbed media. Furthermore, the matrices that allow to construct Gaussian beams are shown to be the propagators of the ray equations linearized around the reference ray. This provides a new insight into the construction of beams in general, and the close relationship between Gaussian beams, WKB, "plane wave" initial conditions and point sources. The fundamental concept here is that of a beam, which is a continuous set of paraxial rays that depend on a single (complex) number called the beam parameter. The two extreme values of the beam parameter define the "canonical" beams. A point source for $\epsilon = 0$ and, for $\epsilon = \infty$, an initially plane wave which corresponds very closely to the intuitive idea of a collimated beam. For other values of ϵ we get intermediate situations corresponding to as many different initial conditions. For complex ϵ we get Gaussian beams. Furthermore, the Hamiltonian formulation used here is very well suited for the propagation of rays and beams across interfaces. Also, when the perturbation becomes large, we can change the reference ray in the unperturbed medium. This allows to maintain an adequate precision for perturbation theory even as the perturbations become large.

In the applications considered in this paper we considered a relatively simple unperturbed background medium. This was either a uniform medium or vertically layered one. Since in these two types of media ray tracing is easily implemented, it is relatively straightforward to calculate the effect of smooth perturbations in the velocity of the medium. For more complex reference media the calculation of unperturbed rays becomes more difficult, and the application of perturbation theory becomes more complex because we have to estimate higher-order derivatives of the velocity distribution in the unperturbed medium.

The method proposed in this paper may be compared with the work published by a number of researchers for more specific applications. For instance, *Woodhouse and Wong* [1986] developed ray perturbation theory in a form which closely ressembles ours, but for a spherical shell. Other authors like *Thomson and Gubbins* [1982] calculated amplitudes by perturbation theory for the case of the structure under the NORSAR array in Norway. However, to our knowledge a systematic development of ray and beam perturbation theory proposing a common background to apparently unrelated problems like amplitude calculation and Gaussian beams had not been published.

Finally, we remark that the theory presented here gives the Fréchet differentials for travel time and amplitude in a explicit form, so that it may be used in the simultaneous inversion of amplitudes variations and time delays. It could also be easily adapted to the evaluation of differential seismograms which could then be used in the inversion of waveforms.

The examples shown in this paper are indicative of the possibilities of the perturbation technique. In all three cases ((1) perturbation of a homogeneous medium by a heterogeneous inclusion, (2) perturbation of a layered structure with a circular inclusion, and (3) a three-dimensional example of a laterally heterogeneous volcanic structure) the results of ray tracing, amplitude, and waveforms obtained by perturbation matched almost perfectly those determined with exact ray theory. The results are encouraging and show that a simple, fast, and efficient technique exists to make estimates of amplitudes in the presence of weak lateral heterogeneities. For larger heterogeneity the perturbation technique may be used as the basic step in an iterative technique. The algorithms that we developed are very efficient and permit iteration to change reference ray when perturbations become large.

Whenever caustics or other ray field singularities appear, we can resort to Gaussian beam summation to improve the results of ray theory. The same program used to calculate synthetics by ray theory may be used to calculate Gaussian beams.

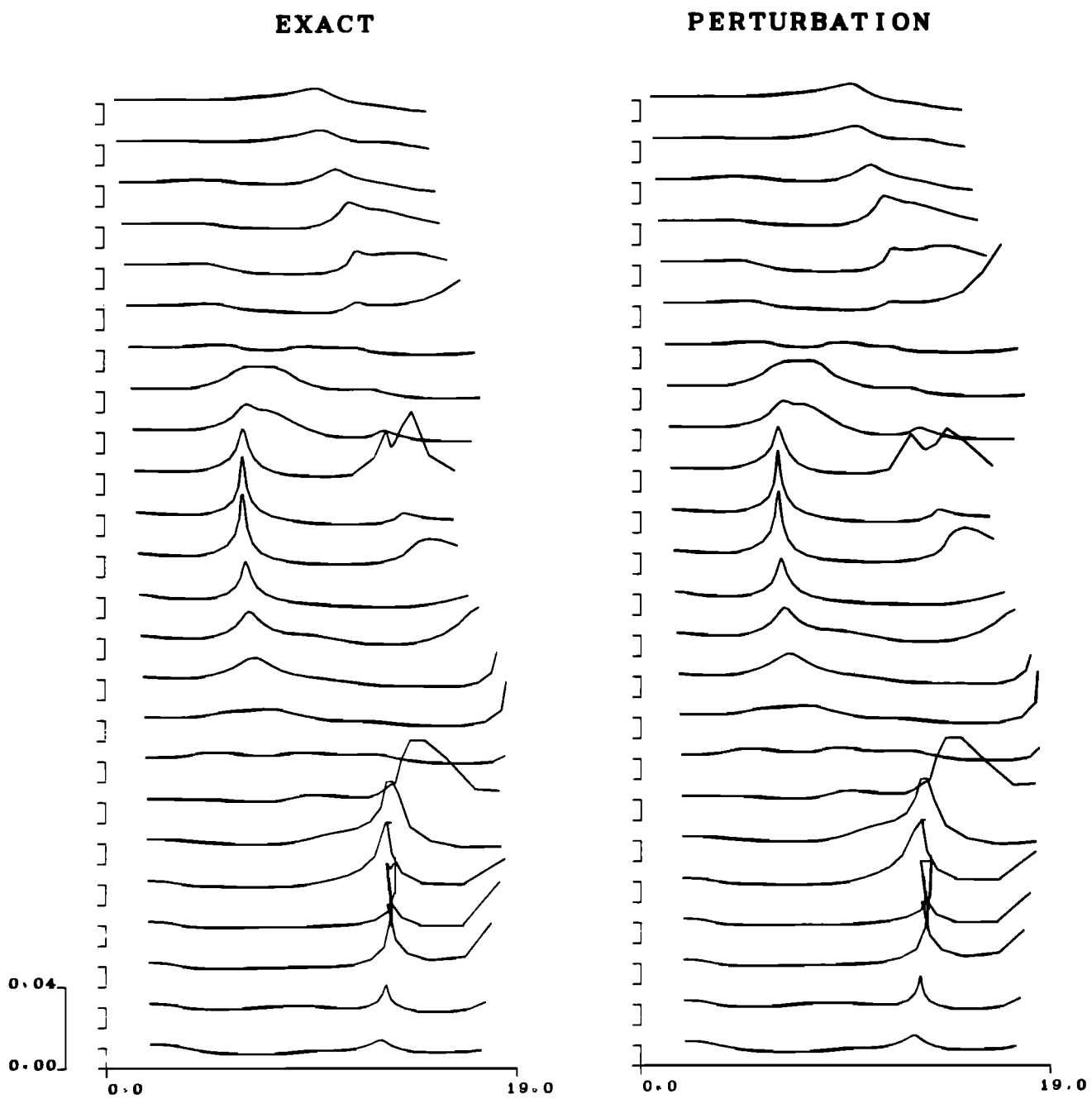

Fig. 14. Amplitude profiles along the lines shown in Figure 13. At left the exact results obtained by exact ray tracing; at right the results obtained with perturbation theory. The strong concentration of energy near the caustics is very clear in the loops near the bottom right of the profiles.

Appendix: Canonical Transformations

We present the canonical transformations from ray centered coordinates to curvilinear coordinates in a two-dimensional space, the following results being easily generalized for a three-dimensional medium. Let (σ_1, q_1) be the ray-centered coordinate system and (σ_2, q_2) the curvilinear system related to the interface. These two systems are connected by the following known transformation :

$$\begin{aligned} \sigma_2 &= F_\sigma(\sigma_1, q_1) \\ q_2 &= F_q(\sigma_1, q_1) \end{aligned} \tag{A1}$$

In both coordinates systems, $\sigma_i (i = 1, 2)$ denotes the independent variable, and H_i is the corresponding reduced Hamiltonian. The momentum conjugate to q_i is p_i, as discussed in the text. Then the travel time differential $d\theta$ for a small variation $(d\sigma_i, dq_i)$ is in either coordinate system (see (6)),

$$d\theta_i = p_i \, dq_i - H_i \, d\sigma_i$$

Imposing the condition that $d\theta_1 = d\theta_2$, the following linear relations are easily obtained from standard calculus:

$$p_1 = \frac{\partial F_q}{\partial q_1} p_2 - \frac{\partial F_\sigma}{\partial q_1} H_2$$
$$H_1 = -\frac{\partial F_q}{\partial \sigma_1} p_2 + \frac{\partial F_\sigma}{\partial \sigma_1} H_2 \qquad (A2)$$

Which may be easily inverted to

$$p_2 = \frac{1}{D}\left(\frac{\partial F_\sigma}{\partial \sigma_1} p_1 + \frac{\partial F_\sigma}{\partial q_1} H_1\right)$$
$$H_2 = \frac{1}{D}\left(\frac{\partial F_q}{\partial \sigma_1} p_1 + \frac{\partial F_q}{\partial q_1} H_1\right) \qquad (A3)$$

where

$$D = \left(\frac{\partial F_q}{\partial q_1}\frac{\partial F_\sigma}{\partial \sigma_1} - \frac{\partial F_q}{\partial \sigma_1}\frac{\partial F_\sigma}{\partial q_1}\right)$$

is the determinant of the transformation. Using the expression of H_1 as a function of q_1, p_1, and σ_1, we finally obtain a relation

$$p_2 = F_p(\sigma_1, q_1, p_1)$$

which, with the functions F_q and F_σ, defines completely the canonical transformation connecting the two phase spaces. These transformations are nonlinear because they are valid for arbitrarily large values of $y_i = (q_i, p_i)$. In order to find the linearized canonical transformations that are needed to connect the paraxial vectors $\delta y_1(\sigma_1) = [\delta q_1(\sigma_1), \delta p_1(\sigma_1)]$ and $\delta y_2(\sigma_2) = [\delta q_2(\sigma_2), \delta p_2(\sigma_2)]$, we use again perturbation theory.

As an example, we give the expression of the canonical transformation between the ray-centered coordinate system (R_r) and the Cartesian coordinate system (R_c) associated with a plane interface as shown in Figure 6. Let us define the position of a point with (σ, n) in the ray-centered coordinate system and with (z, x) in the Cartesian coordinate system defined in Figures 2 and 1, respectively. Then, we have the following nonlinear transformation equivalent to (A1) above:

$$z = -\sin\phi \quad n + \int_0^\sigma \cos\phi \, d\sigma$$
$$x = \quad \cos\phi \quad n + \int_0^\sigma \sin\phi \, d\sigma$$

from which we can calculate the following transformation for the conjugate momentum:

$$p_x = \cos\phi \; p_n + \sin\phi \sqrt{u^2(\sigma, n) - p_n^2}$$

where $\phi(\sigma)$ is the angle between the vertical axis and the tangent to the ray at σ. Since the central ray in ray-centered coordinates is given by $(n = 0, \; p_n = 0)$, these transformations may be linearized in the vicinity of this ray to

$$\delta z = -\sin\phi \; \delta n + \cos\phi \; \delta\sigma$$
$$\delta x = \quad \cos\phi \; \delta n + \sin\phi \; \delta\sigma$$
$$\delta p_x = \cos\phi \; \delta p_n + \sin\phi \; \frac{\partial u}{\partial n} \delta n + \frac{\partial p_x}{\partial \sigma} \delta\sigma \qquad (A4)$$

Let us note with $[\delta n(\sigma), \delta p_n(\sigma)]$ and $[\delta x(z), \delta p_x(z)]$ the canonical vectors of the paraxial rays with respect to the central ray. In order to rotate from ray-centered to Cartesian coordinates we impose $\delta z = 0$. Using this condition on the last two equations, we find

$$\delta x = \cos^{-1}\phi \; \delta n$$
$$\delta p_x = \cos\phi \; \delta p_n + \left(\frac{\kappa}{c}\sin\phi + \frac{dp_x}{d\sigma}\tan\phi\right) \delta n \qquad (A5)$$

where $\kappa = c \; (\partial u/\partial n)$ is the curvature of the central ray. In vertically heterogeneous media, p_x is constant and (A5) reduces to the expressions derived by *Madariaga* [1984]. The matrix M_i and M_t in (38) may be derived from (A5) by inspection.

Acknowledgments. Numerous discussions with J. Virieux and T. George are gratefully acknowledged. We also thank C. H. Chapman, I. Pšenčik, and two anonymous reviewers for very estimulating reviews. This research was supported by CNRS through the programs ATP Geophysique Appliquée and ATP Propagation d'ondes dans des milieux hétérogènes: application au génie parasismique. IPG contribution 933.

REFERENCES

Aki, K., A. Christofferson, and E. S. Husebye, Determination of the three-dimensional seismic structure of the lithosphere, *J. Geophys. Res., 82*,277–296, 1977.

Babich, V. M., and V. S. Buldyrev, *Asymptotic Methods in Short Wave diffraction Problems*, Nauka, Moscow, 1972.

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

Cervený, V., The application of ray tracing to the numerical modeling of seismic wavefields in complex structures, in *Handbook of Geophysical Exploration*, section 1, *Seismic Exploration*, vol. 15A, pp 1–119, Geophysical Press, London, 1985.

Cervený, V., and F. Hron, The ray series method and dynamic ray tracing system for three-dimensional inhomogeneous media, *Bull. Seismol. Soc. Am., 70*, 47–77, 1980.

Cervený, V., and I. Pšenčik, Gaussian beams and paraxial ray approximations in three-dimensional elastic inhomogeneous media, *J. Geophys., 53*, 1–15, 1983.

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

Cervený, V., I. A. Molotkov, and I. Pšenčik, *Ray Method in Seismology*, Karlova Univerzita, Prague, 1977.

Cervený, V., M. M. Popov, and I. Pšenčik, Computation of wave fields in inhomogeneous media-Gaussian beam approach, *Geophys. J. R. Astron. Soc., 70*, 109–128, 1982.

Cervený, V., L. Klimeš, and I. Pšenčik, Paraxial ray approximations in the computation of seismic wavefields inhomogeneous media, *Geophys. J. R. Astron. Soc., 79*, 89–104, 1984.

Chapman, C. H., A new method for computing synthetic seismograms, *Geophys. J. R. Astron. Soc., 54*, 481–513, 1978.

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

Cisternas, A., G. Jobert, and P. Compte, Classical theory of Gaussian beams (in Spanish), *Rev. Geofísica, 40*, 27–32, 1984.

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

George, T., J. Virieux and R. Madariaga, Seismic wave synthesis by Gaussian beam summation: a comparison with finite differences, *Geophysics*, in press 1987.

Gilbert, F., and G. E. Backus, Propagator matrices in elastic wave and vibration problems, *Geophysics, 31*, 326–333, 1966.

Goldstein, H., *Classical Mechanics*, Addison-Wesley, Reading, Massachusetts, 1981.

Julian, B. R., and D. Gubbins, Three-dimensional seismic ray tracing, *J. Geophys., 43*, 95–113, 1977.

Karal, F. C., Jr., and J. B. Keller, Elastic wave propagation in homogeneous and heterogeneous media, *J. Acoust. Soc. Am.*, *31*, 694–705, 1959.

Keller, H. B., Numerical Methods for Two-Point Boundary-Value Problems, Blaisdell, Waltham, Mass., 1968.

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

Landau, L., and E. Lifchitz, *Mecanique*, Mir, Moscow, 1981.

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

Madariaga, R., and P. Papadimitriou, Gaussian beam modelling of upper mantle phases, *Ann. Geophys.*, *6*, 799–812, 1985.

Nercessian, A., A. Hirn, and A. Tarantola, Three-dimensional seismic transmission prospecting of the Mont Dore volcano, France, *Geophys. J. R. Astron. Soc.*, *76*, 307–315, 1984.

Nowack, R., and K. Aki, The two-dimensional Gaussian beam synthetic method: testing and application, *J. Geophys. Res.*, *89*, 7797–7819, 1984.

Thomson, C. J., and D. Gubbins, Three-dimensional lithospheric modelling at Norsar : linearity of the method and amplitude variations from the anomalies, *Geophys. J. R. Astron. Soc.*, *71*, 1–36, 1982.

Woodhouse, J.H., and Y.K. Wong, Amplitude, phase and path anomalies of mantle waves, *Geophys. J. R. Astron. Soc*, in press, 1986.

Yomogida, K., and K. Aki, Waveform synthesis of surface waves in a laterally heterogeneous earth by the Gaussian beam method, *J. Geophys. Res.*, *90*, 7665–7688, 1985.

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

(Received May 20, 1986;
revised October 28, 1986;
accepted October 31, 1986.)