INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

H4.SMR/782-14

Second Workshop on Three-Dimensional Modelling of Seismic Waves Generation, Propagation and their Inversion

7 - 18 November 1994

Waveform Synthesis by Ray Theoretical Methods

R. Madariaga

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

WAVEFORM SYNTHESIS BY RAY THEORETICAL METHODS

Ray theory has played a fundamental role in the interpretation and modeling of seismograms for many applications in seismology. Its use in applied geophysics has been limited to the calculation of travel times. The main reason has been the difficulty to handle singularities in the amplitude of the ray field due to caustics, small discontinuities, etc. A major development has occured in the last ten years with the simultaneous introduction of two methods that have greatly enhanced the possibilities of ray theory. First, dynamic ray tracing or, more appropriately, paraxial ray tracing, a method that permits to calculate not only a ray but a whole beam of rays propagating in its vicinity. Although initially derived from a parabolic approximation to the wave equation in ray centered coordinates, this method is in fact based on classical time-dependent perturbation theory. Paraxial ray tarcing can be derived in any coordinate system in which we can write the eikonal equation in a separable form. The second important development was the introduction of spectral methods in which the wave field is decomposed into a set or basis of independent beams, each of these beams is then propagated independently and finally they are summed at the observer. In the WKB and Maslov methods the beams are chosen as a set of "Snell waves" or initially flat beams. In the Gaussian beam summation method, the base functions are a set of gaussian beams which are propagated using the paraxial approximation. These notes will describe the fundamental aspects of these techniques without getting into excessive algebraic detail. Simple examples will be chosen.

1. INTRODUCTION

The purpose of these notes is to present in a relatively simple form several techniques to construct synthetic seismograms in the high frequency regime. The simplest and most widely used of the high frequency methods is classical geometrical ray theory, which is the basis of most practical methods for the modeling of seismograms and the inversion of travel times in seismology and applied geophysics. A comprehensive discussion of several aspects of ray theory may be found in the books of Cerveny et al (1977) and Bleistein (1984), and in the notes by Burridge (1976). Many programs that perform ray tracing have been written in order to calculate travel times and synthetic seismograms based on ray theory. One of the most difficult problems in generating synthetics was the calculation of geometrical spreading. A major advance in this field was made by Popov and Psencik (1978) who proposed a new technique for the calculation of geometrical spreading. This method is usually called dynamic ray tracing. Later work demonstrated that dynamic ray tracing was in fact a subset of what is called paraxial ray theory in optics, i.e. the study of rays propagating in the vicinity of another ray. This method is new to seismology but it has been used in optics for several decades (see, e.g. Deschamps, 1972). In fact, as demonstrated by Farra and Madariaga (1987), paraxial ray theory, dynamic ray tracing, Gaussian beams and many other problems in ray theory can all be derived from ray perturbation theory as described in all Classical Mechanics books (e.g. Landau and Lifschitz, 1981; Goldstein, 1981). Recognition of this similarity between ray theoretical methods and classical analytical mechanics dates back to last century and was the major contribution of Hamilton to modern physics. With the identification of paraxial ray theory as perturbation theory a wealth of powerful theoretical methods becomes available. For instance, the propagation of paraxial rays across curved interfaces reduces to a simple canonical transformation problem as shown by Farra and Madariaga (1987). A few simple applications of Hamiltonian methods to seismic ray theory will be presented in these notes.

Classical ray theory presents a number of practical problems due to the presence of singularities

of the ray field (caustics and focal points), and numerical instabilities due to small scale perturbations in the velocity model. In the last ten years or so, a number of methods based on the spectral decomposition of the wavefield at the source have been proposed in order to alleviate some of these problems. Among these methods the best known are WKB for vertically stratified velocity models (Chapman, 1978), Maslov or asymptotic Fourier transforms (Chapman and Drummond, 1972) and Gaussian Beams (Popov, 1982 and Cerveny et al, 1982). In the later part of this paper we will discuss these various techniques in the simple case of a homogeneous medium and we will show that they all have a common theoretical background in ray theory.

The basic methods that will be used in these notes are ray theory and ray perturbation theory. Ray theory is used to trace rays and determine travel times for a given set of initial conditions, for instance a point source. Ray perturbation theory is used to evaluate ray amplitudes and to iteratively solve two-point ray tracing problems. The most important perturbations in this context are perturbations of initial and final values of position and slowness. Perturbation theory provides a method to calculate the trajectories of paraxial rays that propagate in the vicinity of a reference ray. The study of the divergence and convergence of paraxial rays provides a method for the calculation of ray amplitudes and travel time extrapolation for different conditions at the source: plane waves, Snell-waves, point sources, Gaussian beams and other wavefront configurations that are needed in waveform synthesis and inversion. Paraxial rays provide also a natural way to interpolate rays so as to solve the two point ray tracing problem: to find a ray that passes through a given source and receiver. Thus we find a common background to such apparently disparate techniques as Gaussian beam summation, ray bending, and the WKB method.

2. CLASSICAL RAY THEORY

Ray theory is based on an ansatz or hypothesis about the form of the solution of the elastic field. As proposed by Babich (1956) and Karal and Keller (1959) we look for elastic waves of the form:

$$\mathbf{u}(\mathbf{x},\omega) = \mathbf{A}(\mathbf{x},\mathbf{x}_0,\omega)\sqrt{\frac{\rho_0 c_0}{\rho c J(\mathbf{x},\mathbf{x}_0)}}e^{i\omega\theta(\mathbf{x},\mathbf{x}_0)} \tag{1}$$

where $\mathbf{u}(\mathbf{x},\omega)$ is the Fourier transformed displacement at point $\mathbf{x}$ in the elastic medium and ω is the circular frequency. ρ and c are the density and wave velocity. This expression is valid both for P and S waves. For the former $c = \alpha$, the P-wave velocity, while for the latter $c = \beta$, the shear wave velocity. $\theta(\mathbf{x},\mathbf{x}_0)$ is diversely known as the eikonal, phase or travel time function. In (1) we have explicitely introduced the position of the source $\mathbf{x}_0$ as well as the density ρ_0 and velocityc_0 at this source. The parameter $J(\mathbf{x},\mathbf{x}_0)$ appearing under the square root is the ray Jacobian or geometrical spreading of the wavefronts; it will be defined later in the paper. In many applications J may be negative or complex so that the proper branch of the square root of J in (1) should be chosen. The vector amplitude $\mathbf{A}(\mathbf{x},\mathbf{x}_0,\omega)$ is a complex function of$\mathbf{x}$.

From a strict theoretical point of view $\mathbf{x}_0$ should not appear in (1), but since we know that solutions propagate along rays we introduce the source and geometrical spreading from the beginning. A more detailed formal development of (1) may be found in Cerveny (1985) or in the references cited in that paper.

In expression (1) there is no approximation since $\mathbf{A}$ is a general function of $\mathbf{x}$ and ω. In order to obtain the ray theoretical approximation we expand the vector amplitude into a series of inverse

powers of ω:

$$\mathbf{A}(\mathbf{x},\mathbf{x}_0,\omega) = s(\omega)\sum_{i=0}^{\infty}\mathbf{A}_i(\mathbf{x},\mathbf{x}_0)\omega^{-i} \tag{2}$$

and retain only the first few terms. In practice, however, only the lowest order term in the series is actually used. In this case, vector $\mathbf{A}_0$ contains the polarization of the wave, the radiation pattern of the source and, if appropriate, a product of reflection and transmission coefficients. For P waves the polarization is along the ray direction, while S-waves are polarized on a plane that is tangent to the wavefront. In (2), $s(\omega)$ is the source wavelet which contains information about the development of rupture for natural earthquakes, or the source time function for explosive sources. Using only the first term in the series (2), (1) reduces to

$$\mathbf{u}(\mathbf{x},\omega) = \mathbf{A}_0(\mathbf{x},\mathbf{x}_0)\sqrt{\frac{\rho_0 c_0}{\rho c J(\mathbf{x},\mathbf{x}_0)}}s(\omega)e^{i\omega\theta(\mathbf{x},\mathbf{x}_0)} \tag{3}$$

In some applications of ray theory, such as Gaussian beams, or in the presence of attenuation θ, J and$\mathbf{A}_0$ may all be complex.

Equation (3) is an approximation to the wave equation valid only at high frequencies when the higher order terms in (2) may be ignored. The ray theoretical approxiamtion assumes that $\mathbf{A}_0, J$ andθ are slowly varying functions of space; the only rapidly varying term in (3) being the exponential. This form of the soluton simplifies the calculation of seismograms in a substantial way. It is in fact simple to do the inverse Fourier transform of (1) in order to obtain the time-domain version of (3). Since $\mathbf{u}(\mathbf{x},t)$ is a real function, the Inverse Fourier Transform has the following form:

$$f(t) = \frac{1}{\pi}\mathrm{Re}\left[\int_0^{\infty} f(\omega)e^{-i\omega t}d\omega\right] \tag{4}$$

where Re denotes the Real part of the complex function in brackets. The inverse of (3) is straightforward:

$$\mathbf{u}(\mathbf{x},t) = \frac{1}{\pi}s(t) * \mathrm{Re}\left[\mathbf{A}_0\sqrt{\frac{\rho_0 c_0}{\rho c J}}\frac{1}{i(t-\theta-i\Delta t)}\right] \tag{5}$$

here Δt is a very small real quantity that is used to pull a possible pole at $t = \theta$ out of the Real t axis. Once the real part is evaluated Δt can be made to tend to zero, in which case (5) may be written in terms of generalized functions or distributions. In the computer implementation of (5) it is preferable to keep a small positive Δt in order to stabilize the numerical evaluation of this expression near the pole at $t = \theta$. As proposed by Madariaga and Papadimitriou (1985), a convenient value for Δt is the time step used in the discrete evaluation of $\mathbf{u}$. Comparing with (3), the inclusion of this small imaginary part is equivalent to multiplying the frequency domain expression (3) by $\tilde{f}(\omega) = exp(-\omega\Delta t)$ for $\omega > 0$. The time domain transform of this function is:

$$f(t) = \frac{\Delta t}{t^2 + (\Delta t)^2} \tag{6}$$

Thus, adding the small imaginary part $i\Delta t$ to θ is equivalent to a convolution of the time domain displacement $\mathbf{u}$ with the function (6). If Δt is equal to the time step, the effect of (6) is practically negligible bellow one half the Nyquist frequency of the signal.

When the travel time $\theta(\mathbf{x},\mathbf{x}_0)$ is real we can evaluate (5) in a more familiar form in terms of the source time function $s(t)$ and its Hilbert transform. Letting $\Delta t \to 0$, we get

$$\mathbf{u}(\mathbf{x},t) = \mathrm{Re}\left[\mathbf{A}_0\sqrt{\frac{\rho_0 c_0}{\rho c J(\mathbf{x},\mathbf{x}_0)}}\right]s[t-\theta(\mathbf{x},\mathbf{x}_0)] + \mathrm{Im}\left[\mathbf{A}_0\sqrt{\frac{\rho_0 c_0}{\rho c J(\mathbf{x},\mathbf{x}_0)}}\right]s^*[t-\theta(\mathbf{x},\mathbf{x}_0)] \tag{7}$$

where $s^*(t)$ is the Hilbert transform of $s(t)$:

$$s^*(t) = \frac{1}{\pi}\text{P.V.}\int_{-\infty}^{\infty}\frac{s(\tau)}{\tau - t}d\tau = \frac{1}{\pi}Im\int_0^{\infty} s(\omega)e^{-i\omega t}d\omega \tag{8}$$

P.V. denotes the principal value of the integral. The last Fourier integral provides the most practical way of computing the Hilbert transform of $s(t)$. It is easy to see that in most computer calculations it is preferable to use (5) to calculate time-domain seismograms.

Let us return now to the ray theoretical expression (3). The main property of ray theoretical seismograms is the clear separation between the kinematics of rays and wavefronts represented here by θ , and the amplitudes and waveforms controlled by $J, \mathbf{A}_0$ and $s(t)$. At high frequencies the quantities outside the exponential vary slowly with position, while the exponential term varies very fast because ω is large. This simplicity of ray solutions comes from the neglect of the interaction of the waveform with the heterogeneities of the propagation medium. There is no scattering along the ray trajectory: rays are bent and deviated by the structure but energy is conserved along ray tubes. As shown by (7), in classical ray theory (θ is real) the only effect of propagation upon waveforms is an eventual Hilbert transformation (phase shift) of the signal. Unfortunately, the limits of validity of this approximation are difficult to evaluate and, except in simple cases, there is no general method for determining the validity of ray theory. A recent discussion of the limit of applicability of ray theory may be found in Ben Menahem and Beydoun (1986).

3. RAY TRACING

In order to calculate the different terms that make up the ray theoretical ansatz (3) we substitute it in the elastodynamic equations:

$$\rho(\mathbf{x})\omega^2\mathbf{u}(\mathbf{x},\omega) = \text{Div}\sigma(\mathbf{x},\omega) \tag{9}$$

where $\rho(\mathbf{x})$ is the density and $\sigma(\mathbf{x},\omega)$ is the stress tensor which is related to strain by the Lamé parameters $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$:

$$\sigma(\mathbf{x},\omega) = \lambda(\mathbf{x})\nabla\cdot\mathbf{u}(\mathbf{x},\omega)\mathbf{I} + 2\mu(\mathbf{x})\epsilon(\mathbf{x},\omega) \tag{10}$$

where $\mathbf{I}$ is the identity matrix and ϵ is the strain tensor:

$$\epsilon = 1/2[\nabla\mathbf{u} + (\nabla\mathbf{u})^T] \tag{11}$$

and the superscript T denotes transposition.

After collecting terms of the same order in ω one finds two sets of independent solutions (see, e.g. Cerveny et al., 1977). From the highest order terms in ω one finds for P waves:

$$(\nabla\theta)^2 = \alpha^{-1} \qquad \mathbf{A}_0\times(\nabla\theta) = 0 \tag{12}$$

and for S waves:

$$(\nabla\theta)^2 = \beta^{-1} \qquad \mathbf{A}_0\cdot(\nabla\theta) = 0 \tag{13}$$

From the next order term in ω, we get the equations for the Jacobian J, which will be discussed later in the text.

In order to interprete these equations we define as in Figure 1, a wavefront as the surface $\theta(\mathbf{x}) =$ constant. Then the vector

$$\mathbf{p} = \nabla\theta$$

is perpendicular to the wavefront and its length is the slowness of the P wave, α^{-1}, or of the S-wave, β^{-1}. $\mathbf{p}$ is simply the slowness vector. The right hand side equations in (12) and (13) define the polarization of vector $\mathbf{A}_0$. This is parallel to $\mathbf{p}$ for P-waves, and perpendicular to it for Swaves.

The left-hand side of equations (12) and (13) may be rewritten in the standard form:

$$(\nabla\theta)^2 = u^2 \tag{14}$$

where $u = c-1$ stands for the slowness α^{-1} of P-waves or β^{-1} of S-waves, and c for the corresponding wave velocity. This is a first-order non-linear equation for the travel time θ , that is usually called the eikonal equation, from eikon, image in greek. The standard method to solve it is the method of characteristics as developped, for instance, by Courant and Hilbert (1966). In ray theory the characteristics are called rays. Let us define the wavefronts as surfaces of equal travel time in $\mathbf{x}$ space: $\theta(\mathbf{x}, \mathbf{x}_0) = t =$ constant. The characteristics of the eikonal equations, or rays, are defined as the trajectories that are orthogonal to the wavefronts. The set of rays and wavefronts depend on the initial conditions for θ . Given an initial wavefront $t_0 = \theta_0(\mathbf{x}, \mathbf{x}_0)$, the rays and succesive wavefronts may be calculated by ray tracing.

In order to obtain the ray tracing equations we introduce the ray coordinate s, a parameter that measures position along the ray. There are many choices for this parameter, for instance it may be the curvilinear distance along the ray, the travel time θ itself, or other discussed by Cerveny (1985). In the following we will use curvilinear distance s as the ray parameter. The ray tracing equations may be easily rewritten for any of the other ray parameters using the Hamiltonian formulation to be discussed later in the paper. From figure 1, we remark that

$$\mathbf{p} = \nabla\theta = u(\mathbf{x})\frac{d\mathbf{x}}{ds} \tag{15}$$

so that the slowness is parallel to the local ray tangent $d\mathbf{x}/ds$. This is the first ray tracing equation, the other one may be obtained taking the gradient of the eikonal equation (14). We write the ray tracing system in the following way:

$$\begin{aligned} \frac{d\mathbf{x}}{ds} &= u^{-1}\,\mathbf{p} = c\mathbf{p}_0 \\ \frac{d\mathbf{p}}{ds} &= \nabla u \end{aligned} \tag{16}$$

The latter equation is closely related to ray curvature. It shows that rays deviate from a straight trajectory because of the gradient of the slowness. The ray curvature is actually given by:

$$\kappa = u^{-1}\,\mathbf{n}\cdot\nabla u = -c^{-1}\,\mathbf{n}\cdot\nabla c \tag{17}$$

where $\mathbf{n}$ is the unit normal to the ray.

Solution of (16) requires the specification of initial or boundary conditions. The simplest problem is to specify the initial position $\mathbf{x}(s_0)$, and slowness vector $\mathbf{p}(s_0)$ for each ray on some initial surface. The direction of the initial slowness vector is arbitrary but its norm has to satisfy the eikonal equation

(14). For a point source, for instance,$\mathbf{x}(s_0)$ is the same for all the rays, while $\mathbf{p}(s_0)$ changes from ray to ray. Once the initial conditions are specified the ray tracing system (16) can be integrated numerically, for instance, by the Runge Kutta method. There exist also a few models of slowness or velocity distribution for which equations (16) may be integrated analytically. Let us remark that the system of six ray tracing equations (16) has to be solved together with the eikonal equation (14) so that in fact only 5 of the equations are independent. In practice, when (16) is being solved numerically, the eikonal equation (14) can be used as a consistency check. Other ways to reduce the system based on coordinate transformations will be discussed in section 5. Once the rays have been traced, the travel time $\theta(\mathbf{x}, \mathbf{x}_0)$ may be calculated by direct integration of

$$\frac{d\theta}{ds} = u \tag{18}$$

along each ray.

Solution of the initial value ray tracing problem is relatively straightforward. In most seismological applications, however, the usual problem is to trace a ray that passes through two points $\mathbf{x}_0$ and $\mathbf{x}_1$. In this case, one has to find the initial value of the slowness $\mathbf{p}_0$ for the ray that satisfies the two boundary conditions. This problem may be solved by iterative methods using the paraxial ray tracing techniques to be discussed later in this paper.

Given appropriate initial conditions, the set of rays and wavefronts is uniquely determined in those regions of space that are illuminated by the initial data. Because the ray tracing system (16) is non-linear the ray field may present singularities. In order to understand these problems and to determine geometrical spreading we remark that the set of rays and wavefronts form a curvilinear coordinate system. As shown in Figure 2 we introduce orthogonal curvilinear coordinates γ_1 an γ_2 on the wavefronts in addition to the ray coordinate s. Each pair (γ_1, γ_2) defines a ray. The curvilinear coordinate set (s, γ_1, γ_2) defined in this form is usually called the ray coordinate system. Any point P in the region illuminated by rays may be defined by its ray coordinates. In this coordinate system the volume element is

$$dV = dxdydz = J(\mathbf{x}, \mathbf{x}_0) ds d\gamma_1 d\gamma_2 \tag{19}$$

where J is the Jacobian of the transformation from cartesian to ray coordinates. Since ds is a curvilinear abscissa along the ray the cross section dS (see fig. 2) of a beam of rays defined by the four rays with coordinates $\gamma_1, \gamma_2, \gamma_1 + d\gamma_1$ and $\gamma_2 + d\gamma_2$ is given by

$$dS = J(\mathbf{x}, \mathbf{x}_0) d\gamma_1 d\gamma_2 \tag{20}$$

so that in fact J is a measure of the variation of the cross section of this beam. J is usually called geometrical spreading, because it measures the spreading of the wavefront around the ray (γ_1, γ_2).

We can now explain the presence of $J^{-1/2}$ in the expression for ray theoretical seismograms (3). Elastic energy flow across a wavefront element of cross section dS, see fig. 2, is:

$$dE = 1/2\rho c ||\dot{u}||^2 dS = 1/2\rho c ||\dot{u}||^2 J d\gamma_1 d\gamma_2$$

Since in the ray approximation energy flows along a beam of rays without lateral scattering, the energy flux

$$F = dE/d\gamma_1 d\gamma_2 = \rho c J ||\dot{u}||^2$$

is conserved. Therefore the amplitudes of the ray theoretical velocity and displacement field are necessarily of the form:

$$\|u\| = \Phi\sqrt{\frac{\rho_0 c_0}{\rho c J)}}$$

where ρ_0 and c_0 have been introduced for convenience and Φ is a source excitation factor to be computed from the solution of a certain canonical problem.

As mentioned earlier, the transformation to ray coordinates may be singular. Near these singularities the usual expressions of ray theory as given by (3) fail and other methods, like WKB or Gaussian beam summation, have to be used. The most common singularity is a caustic which appears when $J \rightarrow 0$. An example of simple caustic formation in a two-dimensional reflexion seismogram is shown in Figure 3. The medium is homogeneous but the reflector is curved, rays deflected by the interface cross each other forming a cusp. We observe that the set of rays is tangent to two curves or caustics. Inside the area delimited by the two caustics three rays arrive at each observer, while only one ray reaches observers outside the caustics. The geometry of caustics may be described by catastrophe theory as shown, for instance, by Nye (1985).

Finally, let us return to the expression (18) for the computation of travel times θ along a ray. Actually, travel times may be calculated not just by integration along rays but by integration along any curve that may be convenient for the problem at hand. In fact since $\mathbf{p} = \nabla\theta$, for any curve Γ joining two points $\mathbf{x}_1$ and $\mathbf{x}_2$ in the region illuminated by rays we get:

$$\theta_2 = \theta_1 + \int_\Gamma \mathbf{p}.d\mathbf{x} \tag{21}$$

where θ_1 and θ_2 are the wavefronts passing through points $\mathbf{x}_1$ and $\mathbf{x}_2$, respectively. The curvilinear integral (21) is path independent in the regions where the ray field is regular. As long as no caustics are crossed the path Γ may be chosen arbitrarily. This curvilinear integral will be used later to calculate wavefront approximations for paraxial rays and beams.

4. VARIATIONAL FORMULATION

The ray tracing problem has been posed so far in its differential form. The problem may also be posed in a variational form which may be used to develop alternative methods of solution of these equations, to introduce perturbation theory, to calculate wave-fronts, etc. The starting point for this formulation is Fermat's principle which may be stated in the following form: among all trajectories joining two fixed points $\mathbf{x}_0$ and $\mathbf{x}_1$, a ray is the trajectory for which the travel time is stationary. We write this condition in the form:

$$\delta\theta(\mathbf{x}_0, \mathbf{x}_1) = \delta \int_0^1 u(\mathbf{x}) \|\frac{d\mathbf{x}}{ds}\| ds = 0 \tag{22}$$

where s is as before the curvilinear distance along the ray. Let us note that since the ray tracing problems are highly nonlinear several rays may satisfy the variational condition (22). The rays that render (22) stationary may be found by standard techniques of the calculus of variations. The solution satisfies the following system of Euler equations:

$$\frac{d}{ds}(u\frac{d\mathbf{x}}{ds}) - \nabla u = 0. \tag{23}$$

which, taking into acount (15) is seen to be equivalent to the ray tracing equations (16). This result demonstrates the equivalence between the variational formulation based on Fermat's principle and the standard ray tracing system derived from the eikonal equation.

Let us note that the variational principle (22) looks for an extremal trajectory without any constraint upon s. The total distance s from $\mathbf{x}_0$ to $\mathbf{x}_1$ is allowed to change so that the time is extremal. Fermat's principle corresponds to the so-called Maupertius principle of minimum reduced action in analytical mechanics. A more powerful variational principle is that of Hamilton that states that the functional (22) is to be minimized under the constraint that the value of s at the initial point $\mathbf{x}_0$ and at the final point $\mathbf{x}_1$ remain constant. The perturbations to the functional that we consider are virtual so that they do not satisfy the condition that $||d\mathbf{x}/ds|| = 1$. Thus s has to be considered as a truly independent parameter. Under this conditions we rewrite $||d\mathbf{x}||$ in the form:

$$||d\mathbf{x}|| = \sqrt{(ds^2 - (||d\mathbf{x}||^2 - ds^2)}$$

where $|d\mathbf{x}|^2 - ds^2$ is the variation of the squared length vector. For small variation we can approximate:

$$||d\mathbf{x}|| \simeq 1/2ds(1 + ||\dot{\mathbf{x}}||^2)$$

where $\dot{\mathbf{x}} = d\mathbf{x}/ds$. We can write Hamilton's principle in the standard form:

$$\delta\theta = 0 = \delta \int_0^1 L ds \tag{24}$$

where the Lagrangian function is:

$$L(\mathbf{x}, \dot{\mathbf{x}}) = 1/2u(\mathbf{x})(1 + ||\dot{\mathbf{x}}||^2)$$

On the true ray the variation is equal to zero and $||\dot{\mathbf{x}}|| = 1$, so that the travel time function $\theta = \int u ds$ obtained from Fermat's principle (22) and for Hamilton's principle (24) are the same. Otherwise, these two principles are quite different.

We can now introduce the Hamiltonian formulation of the ray tracing problem. For that purpose we remark that the ray tracing equations define a ray by a couple of variables $\mathbf{x}(s), \mathbf{p}(s)$. $\mathbf{x}(s)$ describes the ray in configuration space, the physical space where rays are being traced. Rays are more naturally described in a six dimensional phase space, where the variables are$\mathbf{x}$ and $\mathbf{p}$. Rays are trajectories in this space, where they have a number of well known properties. For instance, two distinct ray trajectories never cross each other, because two rays that pass through the same point in phase space are identical. In order to obtain the Hamiltonian we introduce the Legendre transformation:

$$H(\mathbf{x}, \mathbf{p}) = \mathbf{p}\dot{\mathbf{x}} - L(\mathbf{x}, \mathbf{p})$$

where the derivative $\dot{\mathbf{x}}$ is replaced by the generalized momentum $\mathbf{p}$ using the classical relationship $\mathbf{p} = \partial L/\partial\dot{\mathbf{x}} = u(\mathbf{x})\dot{\mathbf{x}}$. Comparing with the definition of slowness in (15) we see that the generalized momentum associated with the coordinate $\mathbf{x}$ is the local slowness vector. Thus we find the Hamiltonian:

$$H(\mathbf{x}, \mathbf{p}) = 1/2u^{-1}(\mathbf{x})[\mathbf{p}^2 - u^2(\mathbf{x})] \tag{25}$$

The term in brackets is just the eikonal equation (14) so that the Hamiltonian is constant and equal to zero for all ray trajectories. Finally, the ray equations (16) may be found in a straightforward manner

from Hamilton's canonical equations (Landau, Lifschits, 1981 or Goldstein, 1981). These equations state that the trajectories that satisfy the variational principle (24) are given by the following set of canonical equations:

$$\frac{dq}{ds} = \frac{\partial H}{\partial p} \qquad \frac{dp}{ds} = -\frac{\partial H}{\partial q} \tag{26}$$

here, following conventional notations in analytical mechanics, q stands for the vector $\mathbf{x}, p$ for the vector $\mathbf{p}$ and partial derivatives are to be interpreted as gradients in phase space. It may be easily verified inserting the Hamiltonian (25) in the canonical equations (26) that these equations are just a short-hand notation for the ray tracing equations (16).

5. REDUCED RAY TRACING SYSTEMS

We remarked above that the standard ray tracing system (16) contains only five independent equations because $\mathbf{p}$ has to satisfy the eikonal equation $\|\mathbf{p}\|^2 = u^2$. In addition to this, the ray parameter s is not really independent and may also be eliminated taking any of the q or p coordinates as the independent variable. The new ray tracing system will contain only four equations . The best way to implement this reduction of the ray tracing equations is to use the Hamiltonian formulation. In fact, if we redefine the Hamiltonian in any convenient way that satisfies the eikonal equation we can new ray tracing equations using the canonical equations (26). We will present two examples of reduction for the two most frequently used coordinate systems: cartesian coordinates in this section, and ray centered coordinates in the following.

In most ray tracing problems in the Earth it is natural to consider ray tracing as a function of depth. This is always done in the case of vertically varying media (see e.g. Bullen and Bolt, 1986) where the ray tracing equations reduce to only 2 equations because the ray is contained in a plane through the source and the observer. In this section we will consider that the vertical coordinate z is used as an independent parameter. In this case we define a *reduced Hamiltonian*:

$$H_r(q, p, z) = -p_z = -\sqrt{u^2(q, z) - p^2} \tag{27}$$

where q stands for the two dimensional position vector $\mathbf{q} = (x, y)$ and p for the two-dimensional conjugate momentum $\mathbf{p} = (p_x, p_y)$. The reduced Hamiltonian (27) was obtained solving for p_z from the original Hamiltonian (25). Thus, the complete slowness vector (p_x, p_y, p_z) still satisfies the eikonal equation. The reduced Hamiltonian (27) contains the same information as (25), but it depends on only four variables. This apparent simplicity is offset by the fact that H_r depends on the independent variable z. Inserting (27) in the canonical ray tracing equations (26) we find explicitly:

$$\begin{aligned}
\frac{dx}{dz} &= \frac{p_x}{p_z} \\
\frac{dy}{dz} &= \frac{p_y}{p_z} \\
\frac{dp_x}{dz} &= \frac{\partial u}{\partial x}\frac{1}{p_z} \\
\frac{dp_y}{dz} &= \frac{\partial u}{\partial y}\frac{1}{p_z}
\end{aligned} \tag{28}$$

where p_z is given by (27). This form of the ray tracing system is entirel equivalent to (15). One may at first sight think that it is easier to solve (28) in order to trace rays in cartesian coordinates.

This is not necessarily so because solutions to (28) may be multiple-valued. This is typically the case in seismology where rays penetrate to a maximum depth and then return to the surface. At the maximum depth, is not defined, or equivalently $p_z = 0$, so that special care must be taken to integrate (28) across this point.

6. RAY CENTERED COORDINATES.

In the course of the study of rays in a multimirrored resonator Popov (1969), see also Babich and Buldyrev (1972), introduced an orthogonal curvilinear coordinate system centered around a reference curve. Later, Popov and Psencik (1978), Psencik (1979), Cerveny and Hron (1980) among many others, found that this coordinate system centered around a reference ray was very convenient to calculate geometrical spreading. The procedure to calculate geometrical spreading by this method is usually called dynamical ray tracing. The ray centered coordinate system was also found to be usefull for the development of paraxial ray theory and one of its main applications: the calculation of Gaussian beams (see, e.g. Cerveny, Klimes and Psencik, 1984; Madariaga, 1984; Klimes, 1984; Cerveny, 1985). Recently, Farra and Madariaga (1987) showed that this coordinate system is also very convenient for the calculation of ray perturbations and the modeling of slightly heterogeneous media. The ray centered coordinate system will be discussed in some detail in the following.

Refering to figure 4, we consider a curve parameterized by the curvilinear abcissa s. Around this curve we generate an orthogonal coordinate system $(s, \mathbf{q})$ where $\mathbf{q}$ is the position vector on a plane orthogonal to the ray. This coordinate system is regular in the vicinity of any curve with finite curvature. The system becomes double valued once the distance to the reference curve is greater or equal than the radius of curvature of the reference curve. The system is chosen so that it is cartesian in the plane (q_1, q_2) normal to the curve. In this system of coordinates, the slowness vector $\mathbf{p}$ is given by:

$$\mathbf{p} = \nabla\theta = h_s \frac{\partial\theta}{\partial s}\mathbf{t} + \frac{\partial\theta}{\partial q_1}\mathbf{e}_1 + \frac{\partial\theta}{\partial q_2}\mathbf{e}_2 \tag{29}$$

where the $\mathbf{e}_i$ are the unit vectors in the plane perpendicular to the curve, $\mathbf{t}$ is the tangent to the curve, and h_s is the scale factor of the curvilinear coordinate s. The scale factor h_s takes into acount the curvature κ of the reference curve, and is given by:

$$h_s(s, \mathbf{q}) = 1 - \kappa(s)\mathbf{n} \cdot \mathbf{q}$$

where $\mathbf{n}$ is the normal to the curve.

Let us introduce the components of the slowness vector in the plane perpendicular to the reference curve. From (29) we get:

$$p_1 = \frac{\partial\theta}{\partial q_1} \qquad p_2 = \frac{\partial\theta}{\partial q_2} \tag{30}$$

In the nomenclature of Hamiltonian theory, these two components of the slowness vector are the momenta conjugate to the position vectors q_1 and q_2. A four dimensional vector (q_1, q_2, p_1, p_2) completely defines the trajectory of a ray in the four-dimensional phase space of position and slowness. Following Popov and Psencik (1978, eq. 3.14) we define the reduced Hamiltonian in this phase space as :

$$H_s(\mathbf{q}, \mathbf{p}, s) = -h_s\sqrt{u^2(s, \mathbf{q}) - \mathbf{p}^2} \tag{31}$$

where as in the case of cartesian coordinates of the previous section, H_s is the negative component of the slowness along the curve, i. e., $p_s = -H_s$, as may be easily verified by reference to (29). Inserting this Hamiltonian in the canonical ray tracing equations (26) we find the ray tracing system in curve centered coordinates:

$$\begin{aligned} \frac{d\mathbf{q}}{ds} &= h_s \frac{\mathbf{p}}{\sqrt{u^2 - \|\mathbf{p}\|2}} \\ \frac{d\mathbf{p}}{ds} &= -h_s \frac{\nabla_T u}{\sqrt{u^2 - \|\mathbf{p}\|2}} + \nabla_T h_s \sqrt{u^2 - \|\mathbf{p}\|2} \end{aligned} \tag{32}$$

where ∇_T denotes the gradient on the (q_1, q_2) plane. Cerveny and Psencik (1979) derived an equivalent expression by a more complex method. As long as the trajectory $\mathbf{q}(s), \mathbf{p}(s)$ does not deviate beyond the evolute of the reference curve, the system (32) is regular and rays may be traced by numerical integration. As will be shown bellow the ray tracing system (32) is always linearized in order to perform paraxial ray tracing. I am not aware of any examples of ray tracing using the full non-linear system (32) although it may be convenient in the ray bending and continuity methods.

The ray tracing system (32) was derived for any reference curve. One particular curve of interest is a ray that has been previously traced by some numerical method. In this case we define the coordinate system around this reference ray as shown in Figure 4. Rays have some particular properties that simplify the calculation of h_s. The curvature of a ray is given by (17). This may be used to simplify the expression for the scale factor. Since $\mathbf{q} \cdot \mathbf{t} = 0$,

$$h_s = 1 - u^1 \mathbf{q} \cdot \nabla u,$$

an expression that was derived by Popov and Psencik (1978).

7. PARAXIAL RAY THEORY

The solution of two-point ray tracing, the calculation of Geometrical spreading, Gaussian beams, etc. become much simpler using the so-called paraxial ray theory, which is just a particular application of perturbation theory to the ray tracing equations. We define paraxial rays (see Figure 5) as those rays that propagate in the vicinity of another ray that is taken as a reference. Suppose that we have succeded in tracing a ray in a medium of slowness $u(\mathbf{x})$. We denote by $[\mathbf{x}_0(s), \mathbf{p}_0(s)]$ the trajectory as a function of s of this ray in phase space. In generalized coordinates this trajectory is written $y_0(s) = [q_0(s), p_0(s)]$, where q_0 represents the generalized position coordinates,p_0 the generalized slownesses, and y_0 is the canonical vector, a vector in phase space. Because of the generality of Hamilton's theory we will derive the paraxial ray equations in generalized coordinates. Results for particular coordinates systems may be determined by straightforward operations. A paraxial ray is described by a perturbation of the ray trajectory in phase space:

$$q(s) = q_0(s) + \delta q(s) \qquad p(s) = p_0(s) + \delta p(s) \tag{33}$$

Tracing paraxial rays consists in finding the canonical perturbation vector $[\delta q, \delta p]$ in phase space. These perturbations in the trajectory are due to small changes in the initial conditions of the ray at the initial point of the ray s_0. Let

$$\delta q(s_0) = \delta q_0 \qquad \text{and} \qquad \delta p(s_0) = \delta p_0 \tag{34}$$

be the perturbations in initial conditions. Inserting the perturbation vectors (33), into the canonical ray tracing equations (20) and developing to first order we find the following linear system for the calculation of paraxial rays:

$$\frac{d\delta q}{ds} = \frac{\partial^2 H}{\partial p \partial q}\delta q + \frac{\partial^2 H}{\partial p^2}\delta p \qquad (35)$$
$$\frac{d\delta p}{ds} = -\frac{\partial^2 H}{\partial q^2}\delta q - \frac{\partial^2 H}{\partial p \partial q}\delta p$$

where all the derivatives of the Hamiltonian are calculated on the reference ray.

In order to find the paraxial ray trajectories the linear system (35) has to be solved numerically together with the ray tracing system. Let us note however that since the system is linear all the paraxial rays in the neighborhood of a certain reference ray may be computed by simple linear operations. In order to see this we write the solution of (35) in the form of a propagator matrix (see, for instance, Gilbert and Backus , 1966):

$$\begin{pmatrix} \delta q \\ \delta p \end{pmatrix} = \mathcal{P}(s, s_0) \begin{pmatrix} \delta q_0 \\ \delta p_0 \end{pmatrix} \qquad (36)$$

where $\mathcal{P}(s, s_0)$ is the paraxial ray propagator from s to s_0 of the paraxial rays. This propagator has a number of very useful properties, the most important for us is that it may be easily inverted. Given then the initial perturbations of position and slowness at s_0, and knowing the propagator, the later position and slowness of the paraxial ray are entirely determined by (36). The paraxial approximation remains valid as long as the perturbation vector is small. The validity of this approximation is unfortunately very difficult to establish in the general case of a heterogeneous reference medium.

7.1 BEAMS

The elements of the propagator $\mathcal{P}$ have a relatively simple physical interpretation if we introduce the concept of a beam. We define a beam as a one parameter family of paraxial rays such that the perturbed initial conditions are related by:

$$[\delta q_0] = \epsilon[\delta p_0] \qquad (37)$$

where ϵ is a complex scalar that defines the shape of the beam. The scalar ϵ may be replaced by a constant complex matrix, but we will not discuss this possibility since we will not use it here.

Let us introduce the following partition of the propagator matrix:

$$\mathcal{P}(s, s_0) = \begin{pmatrix} Q_1 & P_1 \\ Q_2 & P_2 \end{pmatrix} \qquad (38)$$

where, depending on the number of dimensions, Q_i and P_i are submatrices or scalars. Inserting (37) and (38) into equation (36) we find the following solution for individual paraxial rays in a beam of parameter ϵ:

$$\delta q = (\epsilon Q_1 + Q_2)\delta p_0 = (Q_1 + \epsilon - 1 Q_2)\delta q_0 \qquad (39)\delta p = (\epsilon p_1 + P_2)\delta p_0 = (P_1 + \epsilon - 1 P_2)\delta q_0$$

so that, finally, we may write $\delta p = \mathbf{M}(s, s_0)\delta q$, where matrix $\mathbf{M}$ is

$$\mathbf{M}(s, s_0) = (\epsilon P_1 + P_2)(\epsilon Q_1 + Q_2)^{-1} \qquad (40)$$

This matrix is also the tensor of second order derivatives of the travel time function θ as will be demonstrated bellow. From (37) we observe that the initial value of this matrix is simply $\mathbf{M}(s_0, s_0) = \epsilon \mathbf{I}$, where $\mathbf{I}$ is the identity matrix.

The shape of a beam is controlled by ϵ. Two extreme values of ϵ give simple fundamental ray beams. For $\epsilon = 0$ we get a point source since all the rays in the beam leave from the same point in space q_0 with slightly different values of $p(s_0)$. The set of paraxial rays forms a solid angle at the source. This corresponds to the point source (see figure 6a). On the other hand, when $\epsilon = \infty$, $\delta p_0 = 0$ in (37), so that all the paraxial rays share the same initial slowness vector, but leave from slightly different positions $q(s_0) = q_0 + \delta q(s_0)$ in space. The geometry of the beam with $\epsilon = \infty$ depends on the coordinate system under consideration (Madariaga, 1984). For cartesian coordinates the paraxial rays form what is sometimes called a Snell-wave, i.e. as shown in Figure 6b a wave such that all the rays make a constant angle with respect to the (x, y) coordinate plane. In ray-centered coordinates, the $\epsilon = \infty$ beam corresponds, as shown in figure 6c, to an initially plane wave. As we mentionned before, complex values of ϵ are legitimate provided that the corresponding travel times satisfy causality conditions. This occurs for Im $\epsilon < 0$, in which case we get Gaussian beams as will be shown bellow.

Now that we have traced a beam and its paraxial rays we have to calculate their travel times for the paraxials. This is conveniently done using the curvilinear integral (21). We first write :

$$\theta[q(s_0), q(s_1)] = \theta_0[s_0, s_1] + \delta\theta_0 + \delta\theta_1 \tag{42}$$

where θ_0 is the travel time along the central ray from $q_0(s_0)$ to $q_0(s_1)$ and $\delta\theta_0$ and $\delta\theta_1$ are the travel time perturbations at the initial and final point, respectively. From (21) the end point perturbations are simply:

$$\delta\theta = \int_0^\delta q(\mathbf{p}_0 + \delta\mathbf{p}) \cdot d\delta\mathbf{q} \tag{43}$$

Replacing the relation $\delta\mathbf{p} = \mathbf{M}\delta\mathbf{q}$ in (43) and integrating we get:

$$\delta\theta = \mathbf{p}_0 \cdot \delta\mathbf{q} + 1/2\delta\mathbf{q}^T\mathbf{M}\delta\mathbf{q} \tag{44}$$

This is the second order Taylor expansion of the travel time θ around the central ray. From (44) we observe that $\mathbf{M} = \nabla_T\nabla_T\theta$, so that $\mathbf{M}$ is the tensor of second order derivatives of the travel time with respect to the (q_1, q_2) variables. The second order terms are needed in the generation of Gaussian beams and in paraxial approximations to the calculation of synthetic seismograms. In most of the following discussion we will consider point sources so that $\delta\theta_0$ in (42) will be zero.

GEOMETRICAL SPREADING OF A BEAM

Finally, in order to actually compute the ray amplitudes we need to calculate geometrical spreading J. Following Figure 7, we consider an elementary cross section of the beam at s_0 and from each point on this cross section we trace a paraxial ray. As the beam propagates, the paraxial rays may contract or dilate, so that the beam cross section at the end point s is a measure of the geometrical spreading. Consider as in Figure 7 two distinct paraxial rays denoted by their paraxial position vectors δq_1 and δq_2. At any point s along the beam, the cross section of the beam described by these paraxials and the central beam is:

$$dS(s) = (\delta q_1 \times \delta q_2) \cdot \mathbf{t} \tag{45}$$

where $\mathbf{t}$ is tangent to the ray. Let $dS(s_0)$ be the initial beam cross-section at position s_0 along the central ray (see fig. 5). We can follow the change in cross section as a function of s, tracing the paraxial rays by means of the paraxial ray equations (39). At the curvilinear abcissa s the paraxial position vectors δq_1 and δq_2 are given by:

$$\delta q_1(s) = (Q_1 + \epsilon^{-1}Q_2)\delta q_1(s_0) \qquad \delta q_2(s) = (Q_1 + \epsilon^{-1}Q_2)\delta q_2(s_0) \tag{46}$$

where as usual ϵ is the beam parameter. Replacing these vectors in (43) we can find the cross section dS at abcissa s:

$$dS(s) = \mathrm{Det}(Q_1 + \epsilon^{-1}Q_2)\cos\phi|\delta q_1(s_0)||\delta q_2(s_0)| \tag{47}$$

where ϕ is the angle between the normal to the plane defined by the vectors δq_1 and δq_2 and the tangent to the ray. Comparing (47) with the corresponding expression for the cross-section at the initial coordinate s_0, we get:

$$dS(s) = Det(Q_1 + \epsilon^{-1}Q_2)\frac{\cos\phi}{cos\phi_0}dS(s_0) \tag{48}$$

finally we use the very well known result that the geometrical spreading J is just the ratio of the cross section of the beam at s to the initial cross section:

$$J(s, s_0) = Det(Q_1 + \epsilon{-}1Q_2)\frac{cos\phi}{cos\phi_0} \tag{49}$$

This is a general expression for J that is independent of the coordinate system used to trace the rays.

Let us remark that the minors $\mathrm{Det}Q_1$ and $\mathrm{Det}Q_2$ of the propagator matrix $\mathcal{P}$ have a clear physical meaning. Det Q_1 measures the geometrical spreading of a plane or Snell wave because in this case $\epsilon = \infty$. For a point source, $\epsilon \to 0$, and J defined as in (49) becomes singular. For point sources we redefine J so that:

$$J = lim_{\epsilon\to 0}\frac{J}{\epsilon} = \mathrm{Det}Q_2\frac{cos\phi}{cos\phi_0} \tag{50}$$

This expression permits to calculate geometrical spreading for point sources from the propagator for paraxial rays. This method of calculating J is more stable than the more traditional one that consists in tracing several rays around the central ray, and calculating J by the ratio of cross sections.

Let us conclude this section by remarking that our derivation of the paraxial ray equations, the travel time perturbation and geometrical spreading are independent of the coordinate system used to trace the rays. It is equally valid for the original ray tracing system (16) as well as for the reduced ray tracing systems (28) in cartesian coordinates, or (32) in ray centered coordinates.

8. RAY CENTERED COORDINATES AND DYNAMIC RAY TRACING.

The expressions derived in the previous section for paraxial rays are much simpler in ray centered coordinates, which were the coordinates used by Popov and Psencik (1978) in their derivation of dynamic ray tracing. In ray centered coordinates the central ray is simply given by $\mathbf{q}_0 = \mathbf{p}_0 = 0$, since the reference ray is the origin of the coordinates $\mathbf{q}$ and$\mathbf{p}$. Thus in this coordinate system the paraxial ray coordinates (33) have the simple expressions $\mathbf{q}(s) = \delta\mathbf{q}(s)$ and $\mathbf{p}(s) = \delta\mathbf{p}(s)$. Ray tracing of paraxial rays may actually be performed exactly using the nonlinear system (33). In (33) there is no approximation in the region where the ray centered coordinate system is regular, i.e. inside a

region limited by the smallest radius of curvature of the ray. In the paraxial approximation, where $\delta\mathbf{q}$ and $\delta\mathbf{p}$ are considered to be first order perturbations, we use the system (35) which is obtained by linearization of (32). In ray centered coordinates $\partial^2 H/\partial q \partial p = 0$, so that (35) simplifies to:

$$\frac{d\delta\mathbf{q}}{ds} = u^{-1}\mathbf{I}\delta\mathbf{p} \qquad (51)\frac{d\delta\mathbf{p}}{ds} = -u^2\mathbf{V}\delta\mathbf{q}$$

where $\mathbf{V}$ is the matrix of second order derivatives of the velocity field with respect to the ray centered coordinates q_i:

$$\mathbf{V} = [V_i j] \qquad \text{with} \qquad V_i j = \frac{\partial^2 v}{\partial q_i \partial q_j} \tag{52}$$

This expression was originally derived by Popov and Psencik (1978) from the Ricatti equation for the wavefront curvature. Our derivation shows that (51) is just a particular application of the general paraxial ray approximation (35) to ray centered coordinates.

For a heterogeneous distribution of velocity, the numerical solution of equation (51) can be obtained simultaneously with the solution of the ray tracing equations (16). The only difficulty is in the calculation of the matrix of second order derivatives of the velocity in (52). This requires that velocities be interpolated with continuous second order derivatives. Local or global cubic splines are the preferred method for interpolating the velocity field with this condition.

The dynamic ray tracing or paraxial system in ray-centered coordinates may be integrated and the solution expressed as before in terms of a propagator matrix:

$$\mathcal{P} = \begin{pmatrix} Q_1 & P_1 \\ Q_2 & P_2 \end{pmatrix}$$

where the submatrices Q_1, P_1 are the propagators for "plane wave" initial conditions, while Q_2, P_2 are the propagators for a point source. Most of the other results obtained in the previous section apply with some slight modification to ray centered coordinates. The most important result is that the travel time of a paraxial ray is given by

$$\theta(s, \delta\mathbf{q}) = \theta(s, 0) + 1/2\delta\mathbf{q}^T\mathbf{M}\delta\mathbf{q} \tag{53}$$

where $\theta(s, 0)$ is the travel time along the central ray. Compared to the general expression (44) the linear term in $\delta\mathbf{q}$ has disappeared because in ray centered coordinates $\delta\mathbf{q}$ is by definition perpendicular to $\mathbf{p}_0$, which is parallel to the ray tangent. Looking at (53) we realize that the matrix $\mathbf{M}$ has a clear physical meaning: it is the Gaussian curvature matrix of the wave front $\theta(s, \delta q) =$ constant. Matrix $\mathbf{M}$ may be diagonalized in order to find the two principal radii of curvature of the wavefront. In two dimensions,M is scalar and is simply the curvature of the wavefront. The form (53) is a local paraboloidal approximation to the wavefront $\theta =$ constant, for this reason Cerveny, Popov and Psencik (1982) called it the parabolic approximation. In fact in order to derive the paraxial approximation they incorporated the parabolic approximation directly into the ray ansatz (1) and then derived the equations for δq and δp from the wave equation. In our opinion, the method used here is simpler and has a clearer physical meaning.

Finally, geometrical spreading J is given by either (46) or (47) with $\cos\phi = cos\phi_0 = 1$, since for ray centered coordinates the normal to the coordinate plane $\mathbf{q}$ is always parallel to the local tangent to the ray $\mathbf{t}$. Thus once the paraxial ray tracing propagator $\mathcal{P}$ has been calculated, travel times, geometrical spreading, etc are easily calculated in ray centered coordinates.

9. SEISMIC SOURCES

The last element that we need to calculate ray theoretical seismograms using the frequency domain expression (3) or its time domain counterpart (5) is the vector amplitude **A** and the source function $s(\omega)$. We write the excitation of P and S waves in the following form (Madariaga, 1982): For P waves:

$$\mathbf{A}_0 = \mathbf{t}\frac{1}{4\pi\rho_0\alpha_0^3} \qquad s(t) = M_{RR}(t) \tag{54}$$

For S-waves:

$$\mathbf{A}_0 = \mathbf{q}\frac{1}{4\pi\rho_0\alpha_0^3} \qquad s(t) = M_{Rq}(t) \tag{55}$$

Where $\mathbf{t}$ is the vector tangent to the ray at the observation point, and $\dot{\mathbf{M}}(t)$ is the moment rate tensor. M_{RR} is the radial component of the moment tensor in local spherical coordinates around the source. Similarly M_{Rq} is the tangent component of the moment tensor in spherical coordinates projected on a plane perpendicular to the take off direction of the ray. Let $\mathbf{t}_0$ be the tangent to the ray at the source, then the projection of **M** on the plane perpendicular to $\mathbf{t}_0$ is:

$$\mathbf{M}\cdot\mathbf{t}_0 = M_{RR}\mathbf{t}_0 + M_{Rq}\mathbf{q}_0$$

where $\mathbf{q}_0$ the unit vector in the direction of $\mathbf{t}_0 \times (\mathbf{M}\cdot\mathbf{t}_0)$. This unit vector points in the direction of the shear component of $\mathbf{M}\cdot\mathbf{t}_0$. With this notation for the components of the moment tensor used in (45), the source time functions may be rewritten as:

$$\dot{M}_{RR} = \mathbf{t}\cdot\mathbf{M}_0\cdot\mathbf{t}_0$$
$$\dot{M}_{Rq} = \mathbf{q}\cdot\mathbf{M}_0\cdot\mathbf{t}_0$$

Vector $\mathbf{q}$ is the polarisation vector for the S-wave, it may be obtained by propagation of the unit vector $\mathbf{q}_0$. The expressions in (54) permit to calculate ray amplitudes in any elastic medium where the ray tracing problem has been solved. The calculation of synthetic seismograms of far field body waves for the study of source processes are one example of the use of ray synthetics in practical applications. Calculation by ray theory is limited to rays that penetrate into the lower mantle, otherwise we would have to deal with triplications and caustics due to the upper mantle discontinuities. For this reason classical ray theory is limited to the calculation of synthetics in the range from 30° to 90°. At shorter distances WKB (Chapman, 1978) or Gaussian beam summation (Madariaga and Papadimitriou, 1985) provide a practical method to calculate synthetics. This kind of synthetic seismogram is widely used in order to study fault mechanisms and the distribution of asperities on the faults.

10. SPECTRAL METHODS.

The methods we have described so far are all based on the direct use of the ray theoretical expression (3). This expression becomes singular or unstable under many practical circumstances. For instance, as shown on Figure 8, in the vicinity of the caustic the geometrical spreading function J tends to zero and changes sign once the rays have crossed the caustic. Near a caustic, ray theory predicts infinite amplitude although we know that in reality amplitudes are finite around the caustic. This failure of ray theory comes from its inability to deal with finite frequency phenomena. Another problem, also illustrated in Figure 8 is the extreme sensitivity of ray theory to small local perturbations in the slowness field. This problem is due again to the assumption that frequency is infinite. At finite

frequencies, small discontinuities should only affect waves with a wavelength that is similar or shorter than the dominant wavelength of the diffracting object. At low frequencies, when the wavelength is longer than the characteristic size of the heterogeneity, the efect of the perturbation is to generate weak Rayleigh scattering, but the travel times, geometrical spreading, of the rays should not be severely affected. Spectral methods provide a partial solution to this problem without loosing the physical appeal and simplicity of ray theory.

In a spectral method, the field is not calculated directly from (3) but from a sum of beams of the form (3). We say that the elastic wave field is projected from configuration space into a base of beams with variable ray parameter. The best known of these methods is WKB for vertically heterogeneous media. In this method, originally developped in cartesian coordinates, the source is developped into a sum of Snell-waves (see figure 6b), then each Snell-wave is propagated independently and a seismogram is calculated summing all the propagated Snell waves. This process is illustrated in Figure 9. Since a seismogram is calculated evaluating a sum of beams, it is much easier to control the frequency contents of the final synthetic. Also, the problem with caustics is partially suppressed because the Snell waves usually behave regularly near caustics in configuration space. Unfortunately, caustics may also appear in the individual Snell waves and may render the calculation unstable. Appropriate combination of both classical ray theory and WKB seem to be the best answer to problems with the stability of the ray field.

Another method that has appeared relatively recently in the literature is the Gaussian beam summation method. Just as with WKB, the source is expanded into a series of Gaussian beams each of which is propagated independently. A synthetic is calculated by the summation or stacking of individually propagated Gaussian beams. The method is very similar to WKB, but presents the additional advantage that Gaussian beams do not have caustics and may be calculated everywhere. In the folowing we will briefly discuss these two methods.

11. THE WKB METHOD

The WKB method introduced by Chapman (1978) is probably the most widely used method for the calculation of high frequency synthetic seismograms. Derived originally only for vertically heterogeneous media, it was later extended to media with arbitrary heterogeneity (Chapman and Drummond, 1982). Its main limitation is that of ray theory: the heterogeneity of the medium should be smooth in comparison to the wavelength under consideration. For simplicity of the exposition we will develop the WKB method in a two dimensional medium, the three dimensional case is treated by Chapman (1978). The starting point for the WKB method is Weyl's integral in two dimensions or the Sommerfeld integral in three dimensions. In order to simplify the presentation we consider a line source in a homogeneous medium as a model problem. P wave radiation from this source is written in the form of a Green function:

$$\mathbf{g}(\mathbf{x},t) = \frac{1}{2\pi}\mathbf{A}_0(t^2 - \mathbf{x}^2/c^2)^{-1/2}\mathrm{H}(t - \|\mathbf{x}\|/c) \tag{57}$$

where $c = u-1$ is the constant velocity of the medium. $\mathbf{x}$ is the position of the observer with respect to the source, and the vector amplitude $\mathbf{A}_0$ is parallel to the radial vector $\mathbf{x}$, its scalar value depends on the source under consideration, and will be discussed later. The two dimensional Green function (57) has the classical inverse square-root singularity at the wave front. Its time Fourier transform is:

$$\mathbf{g}(\mathbf{x},\infty) = i/4\mathbf{A}_0 H_0^{(1)}(\omega\|\mathbf{x}\|/c) \tag{58}$$

where i is the imaginary unit. The Fourier transform in (58) is defined as in (4). The Green functions (57) and (58) are exact. In order to compare these solutions with ray theory we have to calculate their high frequency approximation. Using the symptotic expansion of the Hankel function, we get:

$$\mathbf{g}(\mathbf{x},\omega) = 1/4\pi\mathbf{A}_0 \left[\frac{2c}{||\mathbf{x}||}\right]^{1/2} \tilde{\lambda}(\omega) e^{i\omega\theta_2} \tag{59}$$

where the source function

$$\tilde{\lambda}(\omega) = \left[\frac{\pi}{\omega}\right]^{1/2} e^{i\pi/4} \tag{60}$$

is the Fourier transform of the inverse square root pulse:

$$\lambda(t) = t^{-1/2} H(t) \tag{61}$$

The travel time is

$$\theta_2 = \mathbf{x}/c \tag{62}$$

and comparing with (3) we observe that the geometrical spreading

$$J_2 = ||\mathbf{x}|| \tag{63}$$

In (61) and (62) a subindex 2 has been added to θ and J, in order to indicate that this is for two dimensional propagation. With these definitions we observe that (59) is in the form of the ray theoretical approximation (3) if in the latter we assume that the medium is homogeneous. The source function $s(\omega)$ in (2) has been replaced in (59) by $\tilde{\lambda}(\omega)$. We can calculate the time domain inverse of (59) using (7). Since all the amplitude terms in (59) are real we get:

$$\mathbf{g}(\mathbf{x},t) = \frac{1}{4\pi}\mathbf{A}_0 \left[\frac{2c}{J_2}\right]^{1/2} \lambda(t-\theta_2) \tag{64}$$

with θ_2 and J_2 given by (62) and (63).

This completes the demonstration that ray theory in a homogeneous medium is icentical to high frequency asymptotics. In quantum mechanics these high frequency methods are usually designated WKB approximations (see also Bleistein, 1984).

The spectral decomposition of (58) into plane waves is very well known and will be the starting point for our development of the WKB method. By means of Weyl's integral we can write (58) for $z > 0$ in the form:

$$\mathbf{u}(\mathbf{x},\omega) = \frac{i}{4\pi} \int_{-\infty}^{\infty} \mathbf{A}_0 e^{i\omega(px+qz)} dp/q \tag{65}$$

for $\omega > 0$. The integral representation (65) is in fact the Fourier transform with respect to x of the Green function (59). Equation (65) may be rewritten in order to show that it is in fact a superposition of plane waves. Let us define position and slowness as usual:

$$\mathbf{x} = (x,z) \qquad \text{and} \qquad \mathbf{p} = (p,q) \tag{66}$$

Because of the eikonal equation, $||\mathbf{p}|| = c-1$, so that q is not independent of p and may be written:

$$q = (c^{-2} - p^2)^{-1/2} \qquad \text{with} \qquad \mathrm{Im}(q) \geq 0 \tag{67}$$

Finally following the definitions in Figure 10, we can change the integration variable in (3) to the take off angle ϕ_0 defined by

$$p = \frac{sin\phi_0}{c} q = \frac{cos\phi_0}{c} \tag{68}$$

where the angle ϕ_0 should be taken along the countour $\mathcal{L}$ in the complex ϕ_0 plane shown in Figure 11. This contour is the mapping of Real p axis into the complex ϕ_0 plane. We then rewrite (65) in the form:

$$\mathbf{g}(\mathbf{x},\omega) = \frac{i}{4\pi} \int_{\mathcal{L}} \mathbf{A}_0 e^{i\omega\theta_2} d\phi_0 \tag{69}$$

where $\theta_2 = \mathbf{p} \cdot \mathbf{x}$ is the travel time from the origin to the current point $\mathbf{x}$ of a plane wave that leaves the origin in the direction $\mathbf{p}$. (69) is plainly a sum of plane waves of amplitude $i/4\pi\mathbf{A}_0$ and phase θ_2.

The plane wave decomposition provides an alternative way to calculate synthetic seismograms. A problem with it is the presence of inhomogeneous waves. The two vertical branches of $\mathcal{L}$ shown in figure 11, represent inhomogeneous plane waves that propagate horizontally along the $z = 0$ line and decrease exponentially with depth. These waves contribute to the near field of the Green function (57) and may be neglected as long as the observer is not too close to the x axis. In order to see this we remark that in the far field when $||\mathbf{x}|| >> 0$, the main contribution to the integral (69) comes from the vicinity of the stationary phase point:

$$\frac{\partial\theta_2(\phi_0)}{\partial\phi_0} = c^{-1}(x\cos\phi_0 - y\sin\phi_0) = 0 \tag{70}$$

Let us call ϕ_G the value of the angle ϕ_0 at the stationary point, solving (70) we find $\phi_G = \text{atan}(x/y)$. ϕ_G is just the take-off angle of the ray that joins the source to the observation point $\mathbf{x}$. Using the stationary phase approximation in the spectral integral (69), we obtain the ray theoretical solution (64). Thus a simple guess is that most of the contribution to a seismogram calculated using (69) will come from take off angles ϕ_0 close to ϕ_G. For this reason we restrict our integral to the real axis $-\pi/2 < \phi_0 < \pi/2$ and calculate the time domain transform of (69). Let us consider an individual plane wave in (69):

$$\mathbf{g}(\mathbf{x},\omega) = \frac{i}{4\pi}\mathbf{A}_0 e^{i\omega\theta_2(\phi_0)} \tag{71}$$

This plane wave may be transformed into the time domain using (5):

$$\mathbf{g}(\mathbf{x},t) = \frac{1}{4\pi^2}\mathbf{A}_0 \text{Re}\left[\frac{1}{t-\theta_2}\right] \tag{72}$$

We observe that the plane waves that are used to generate a synthetic seismogram by the sum (69), present a non causal behaviour. In fact the time signal in (72) is the Hilbert transform of a delta function.

Using the Fourier transform (72) in the sum (69) we get the WKB approximation to the Green function:

$$\mathbf{g}(\mathbf{x},t) = \frac{1}{4\pi^2}\text{Re}\left[\int_{\mathcal{L}} \frac{1}{t-\theta_2(\phi_0)} d\phi_0\right] \tag{73}$$

This way of calculating the synthetics, doing first the time Fourier transform and later the sum over the take off angle was introduced by Chapman (1984) and is the basis of the succes of the WKB and the Gaussian beam summation method. As long as the inverse Fourier transform (72) of the individual plane wave can be computed exactly, the plane wave sum can be easily computed. Let us

finally remark that evaluating the integral (73) is not straightforward because of the singularity that occurs when $t = \theta$.

This problem may be avoided by several different approximations described by Chapman (1984). In the following and by analogy with Gaussian beams we will use the method we proposed in (5), i.e. we add a small Imaginary part $i\Delta t$ to the plane wave phase θ . The resulting Green function, see Figure 11, is smoothed but the smoothing is perfectly compatible with the discretization path Δt. More serious is the problem with the cut off phases indicated also in figure 11. These phases that arrive before the geometrical onset of the Green function are due to the limitation of the integral (73) to a segment of the real axis. This problems comes from the fact that all plane waves contribute with the same amplitude to the WKB synthetic calculated using (73). Thus the arrest of the integration at a certain angle ϕ_1 produces a rather large effect on the integral. One of the possible ways to reduce the effect of these cut off phases is to filter the integral near the ends of the integration range, this may be done by simple tapering of the amplitudes $\mathbf{A}_0$ or by more sophisticated techniques. Another possibility is to replace the plane waves by Gaussian beams as will be shown in the next section.

The generation of the Green function (57) by the sum of plane waves is quite curious. The individual waves that are summed in (73) have their peak amplitude near $t = \theta_2$. Thus the contribution of each plane wave in the sum will be concentrated around θ_2. But θ_2 varies as a function of ϕ_0 in such a way that it it has maximum at $\phi_0 = \phi_G$, the stationary phase point already calculated above. Thus the maximum contribution of each plane wave occurs before the geometrical arrival time!

The WKB method can be easily extended to vertically heterogeneous media for which it was originally proposed by Chapman. It may also be applied in laterally heterogeneous media where Chapman prefers to call it the Maslov method, because in laterally heterogeneous media the plane wave decomposition (65) is only asymptotically valid. However, since the same asymptotic methods are used to write the propagation of the individual "plane" waves as to find the asymptotic Fourier transform it is preferable to present the Maslov-WKB method as a simple extension of the plane wave decomposition discussed above. This will be developped in the final version of the paper.

11. GAUSSIAN BEAMS

As we mentioned above, one of the main problems with the calculation of the plane wave sum (73) is the spurious cut off phases that appear when the integral is limited to a finite segment of ϕ_0. One succesful method for reducing the influence of these cut off phases is to use Gaussian beams. Before introducing Gaussian beams we must rewrite the plane wave (70) as a beam. Refering to figure 11 we note that this expression may be rewritten in terms of the distance s along the ray leaving the origin in the direction ϕ_0. This ray plays the role of the central ray in the paraxial theory presented in section 8. We can rewrite θ_2 in the simple form:

$$\theta_2 = s/c \tag{76}$$

which is the travel time of the plane wave that passes through $\mathbf{x}$. s is the distance along the central ray and plays the role of the independent parameter s in ray theory. The set of rays associated with this plane wave are parallel to the central axis. In the vicinity of the central ray we define the paraxial zone. In this zone the ray centered coordinates of a paraxial ray is simply $\delta q = \delta q_0$, $\delta p = \delta p_0 = 0$ wich states the trivial fact that paraxial rays are parallel to the central ray. With respect to the the

paraxial theory developped above we observe that the plane waves that are used in the WKB integral form what we called a plane beam. Since the medium is homogeneous the plane beam remains plane as it propagates away from the source.

In Gaussian beam summation we replace the plane beams of the WKB method by Gaussian beams. In order to construct Gaussian beams we slightly perturb the plane beam introducing a small complex part in p_0. Refering to (31) in section 7 we take:

$$\delta p_0 = \epsilon - 1 \delta q_0 = i\delta \delta q_0 \tag{77}$$

The small parameter δ produces a fanning of the plane plane wave. For real ϵ the plane wavefront deforms into a parabolic front and the paraxial rays tend to move away from the central ray. Let us demonstrate this: (77) gives the initial conditions for the tracing of the paraxial ray that leaves the source plane at a distance δq_0 from the origin (see Figure 12). We can then trace this paraxial ray with the help of the paraxial propagator (32). In a homogeneous medium this propagator takes the simple form:

$$\mathcal{P}(s, s_0) = \begin{pmatrix} 1 & c(s - s_0) \\ 0 & 1 \end{pmatrix} \tag{78}$$

so that the paraxial ray is given by (33):

$$\delta q = [1 + \epsilon^{-1} c(s - s_0)] \delta q_0 \qquad \text{and} \qquad \delta p = \epsilon^{-1} \delta q_0 = \delta p_0 \tag{79}$$

Thus the paraxial ray is a straight line as one expects for a homogeneous medium, but the direction of the ray differs from that of the central ray. Finally we can compute the travel time θ along the perturbed beam:

$$\theta_2 = \frac{s}{c} + \frac{1}{2} M \, \delta q^2 \tag{80}$$

where M the curvature of the wavefront is given by

$$M = \frac{\epsilon^{-1}}{1 + \epsilon^{-1} c(s - s_0)} = [\epsilon + c(s - s_0)]^{-1}$$

For $\epsilon > 0$ the radius of curvature M^{-1} grows with distance from the origin and is zero for a points $- s_0 = -\epsilon/c$. Thus as shown in figure 12 it appears as if the perturbed beam was coming from a source situated behind the origin at a distance $-\epsilon/c$. The paraxial approximation for the travel time replaces the spherical wavefront emanating from the focus at $x = -\epsilon/c$ by a parabolic wavefront. In this sense the paraxial ray theory is a parabolic approximation to the wave equation. It differs with respect to the parabolic approximation used applied geophysics (Claerbout, 1985) because it is also a high frequency approximation.

We can now introduce Gaussian beams. They are directly obtained from the paraxial approximation (79) replacing ϵ by a large imaginary value. The paraxial rays are complex and the travel time θ_2 is also complex. If we collect the results obtained so far we find that a Gaussian beam is given by:

$$\mathbf{g}(\mathbf{x}, \omega) = \frac{i}{4\pi} J^{-1/2} e^{i\omega\theta_R} e^{-\omega\theta_I} \tag{81}$$

where θ_R and θ_I denote the real and imaginary part of the travel time (80). For $\mathrm{Im}\, \epsilon < 0$, we can verify that $\theta_I > 0$ and therefore (81) presents a Gaussian amplitude decrease in the direction perpendicular to the central ray:

$$e^{-\omega\theta_I} = e^{-1/2 \omega \mathrm{Im}(M) \delta q^2}$$

so that the Imaginary part of the curvature of the wave front determines the attenuation of the beam away from the central ray.

The last thing we need to calculate (81) is to determine J, but this is given by (49) for a beam of arbitrary shape:

$$J = 1 + \epsilon^{-1} c(s - s_0) \tag{82}$$

Thus replacing (82) and (80) in (81) and taking $s_0 = 0$ we have a simple practical way for computing a Gaussian beam centered around a ray passing through the source situated at the origin of coordinates. The calculation of these Gaussian beams is not much more difficult than the calculation of a plane wave, the only additioanl difficulty is estimating the distance δq from the observation point to the central ray. This is not difficult to evaluate in the present problem but may be a serious problem in heterogeneous media. In that case it would be preferable to use cartesian coordinates for the calculation of Gausian beams. This has been discussed by Madariaga (1990).

We can now replace every plane wave in the WKB integral (73) by its corresponding Gaussian beam. Since this is being introduce arbitrarily here one could in principle take any suitable value of ϵ for each ray. In practice however, ϵ should be constant for all rays when evaluating the Gaussian beam sum. We get then

$$\mathbf{g}(\mathbf{x}, \infty) = \frac{i}{4\pi} \int \mathbf{A}_0 J^{-1/2} e^{i\omega\theta} d\phi_0 \tag{83}$$

where J and θ are the complex geometrical spreading and travel time determined above. Finally using (5) we find the time domain Gaussian beam sum:

$$\mathbf{g}(\mathbf{x}, t) = \frac{1}{4\pi} \mathrm{Re} \left[\int \mathbf{A}_0 J^{-1/2} \frac{1}{t - \theta} d\phi_0 \right] \tag{84}$$

Gaussian beams synthetics are evaluated discretizing this integral. If one uses the broadening parameter Δt introduced in (5) this discretization presents no stability problems. However in more complex situation one has to insure an appropriate density of rays in order to avoid interference problems in calculating it. An example of the calculation of a Green function is shown in Figure 13 where it is compared with the equivalent result obtained by plane wave summation (WKB).

Examples

A large number of examples of calculation of theoretical seismograms using classical ray theory or Gaussian beam summation have appeared in the literature. Among the numerous publications where the method has been used to generate realistic seismograms we can cite: Cerveny (1985), Cormier and Spudich (1984), Nowack and Aki (1985), Madariaga and Papadimitriou (1985), etc. These authors show that the Gaussian beam summation method can be used to solve many of the problems encountered when using straigtforward ray theory, without completely loosing the simplicity and physical appeal of rays. Comparison with finite difference calculations by George et al (1987) showed that the method works for caustics giving not only the right amplitudes in the illuminated zone but also on the shadow regions. The same authors showed that edge diffraction and head waves are difficult if not impossible to model with Gaussian beam summation unless special techniques are used. Recently, White et al. (1987) made a careful study of the conditions of validity of the sum of gaussian beams for several simple structural models. We refer the interested reader to these papers for further information. We conclude presenting with an example of a realistic structure with five discrete

reflectors presented in Figure 13. At the top of the Figure is a diagram of the model together with an example of the ray tracing that is needed to obtain good precision in the Gaussian beam summation. For the construction of the cross section at the bottom we have to calculate a series of reflected rays from each interface. The presence of small kinks and imperfections in the interfaces creates numerous caustics which would render impossible to calculate these profiles with classical ray theory. Gaussian beams on the other hand can smooth these imperfections yielding continuous reflections in the seismograms presented at the bottom. The calculation was carried out in a minicomputer and included several thousand rays.

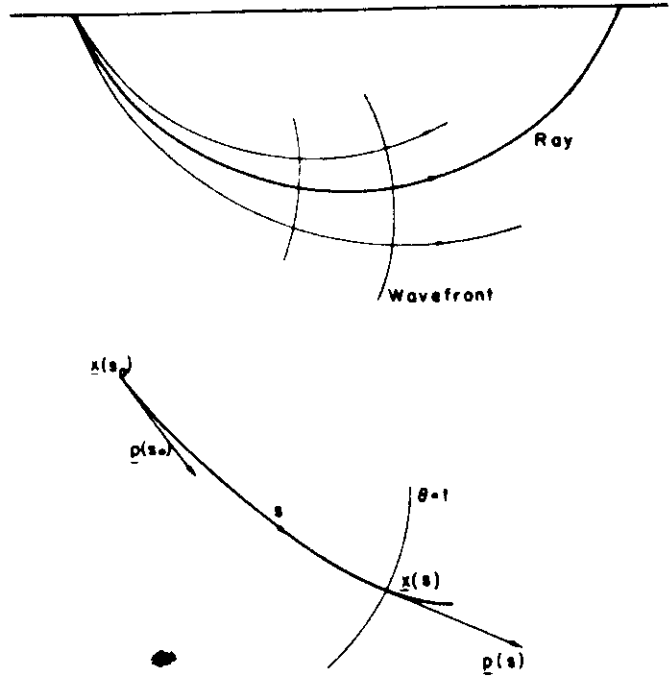

Fig. 1. Geometry of Rays and Wavefronts, p denotes the slowness vector.

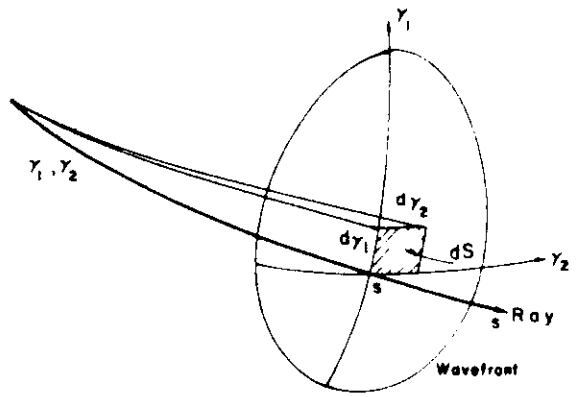

Fig. 2. Definition of the Ray coordinates (γ_1, γ_2) and the elementary surface area on a wavefront dS.

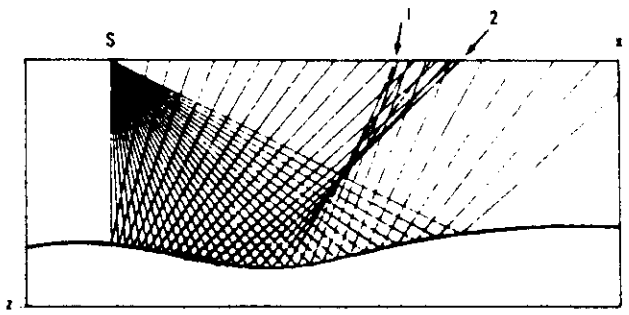

Fig. 3. Formation of a caustic in a reflection seismogram. The change in curvature of the reflector produces two caustics (1 and 2) in the reflected rays issued from a point source S.

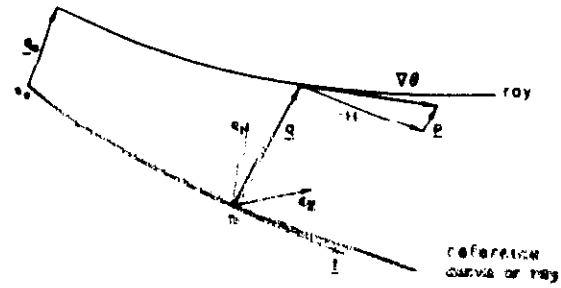

Fig. 4. Geometry of the curve centered coordinate system (s,q_1,q_2). Vector $\underline{q}$ is in the plane (q_1,q_2) perpendicular to the local ray tangent $\underline{t}$.

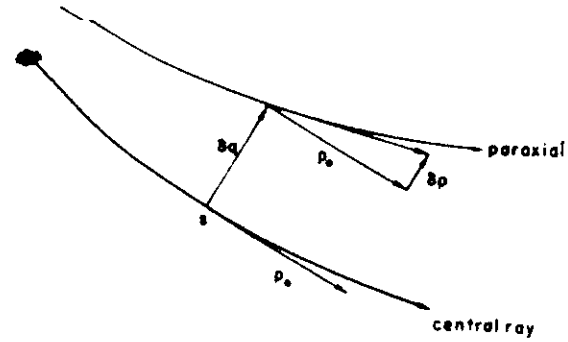

Fig. 5. Geometry of a paraxial ray propagating in the vicinity of a central ray. δq and δp are the perturbations in position and slowness in ray centered coordinates, respectively.

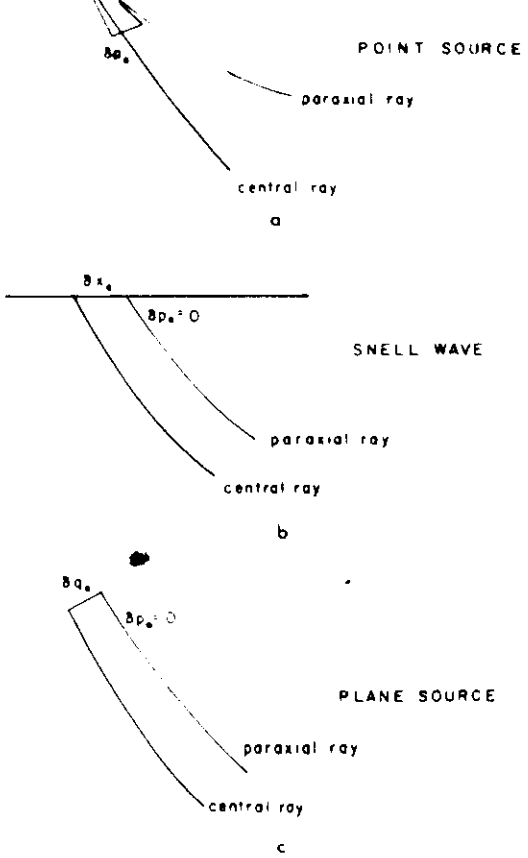

Fig. 6. The three basic geometries for the paraxial rays. On top, a, the point source. At the center the "Snell wave" where the paraxial leaves the reference plane with the same take off direction as the central ray. At the bottom, c, the plane source of Cerveny et al. (1982). Here the rays form a plane wave at the source.

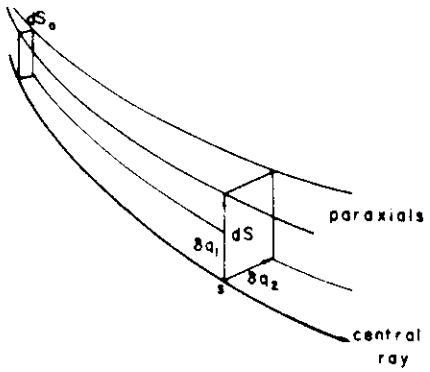

Fig. 7. Geometry of a beam for the calculation of geometrical spreading.

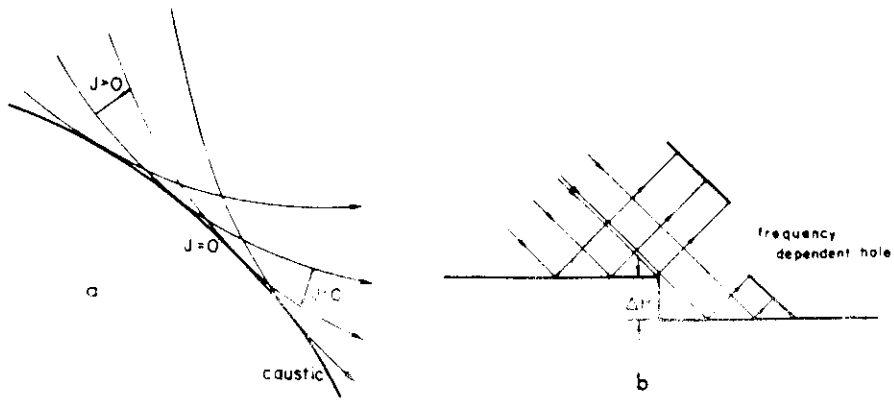

Fig. 8. Two examples of problems with ray theory. On top a caustic, at the bottom a small discontinuity in a reflector.

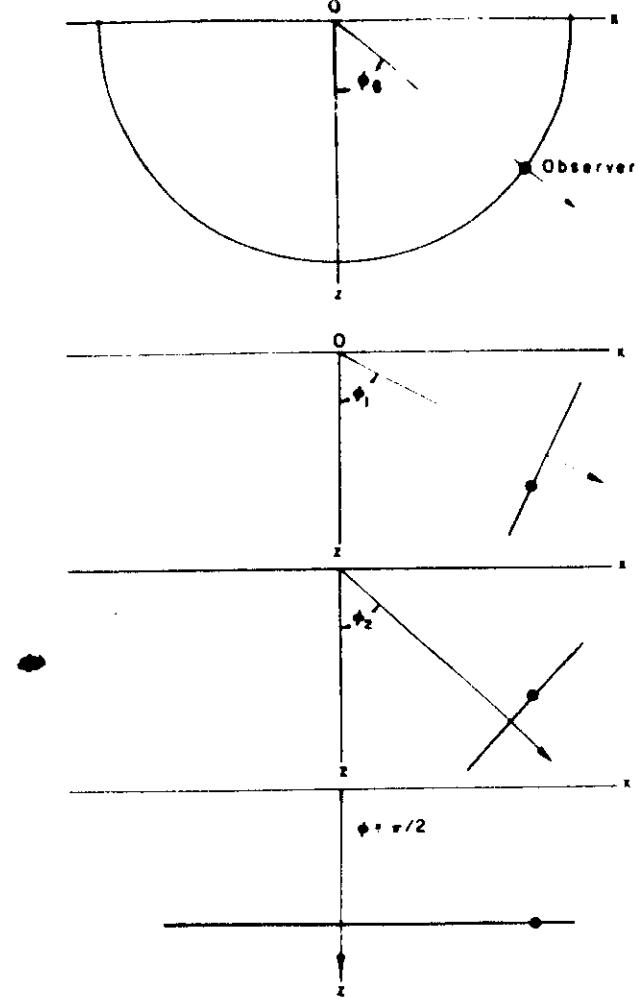

Fig. 9. Construction of a cylindrical wave by the summation of plane waves in the Weyl integral representation of the two-dimensional Green function.

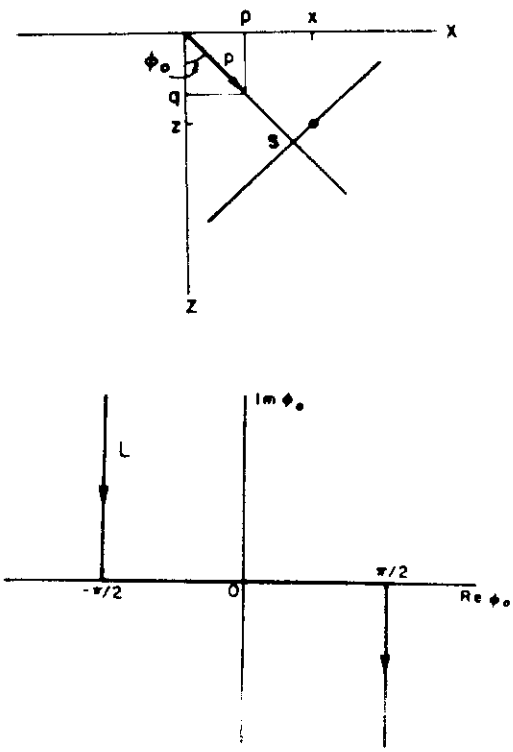

Fig. 11. Contour of integration in the complex ϕ_o-plan or the WKB or plane wave spectral decomposition of the two-dimensional Green function.

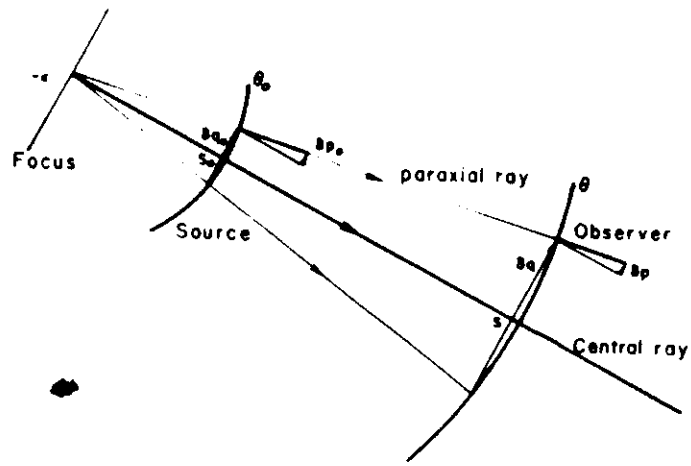

Fig. 12. Geometry of a parabolic beam. The focus does not coincide with the source.

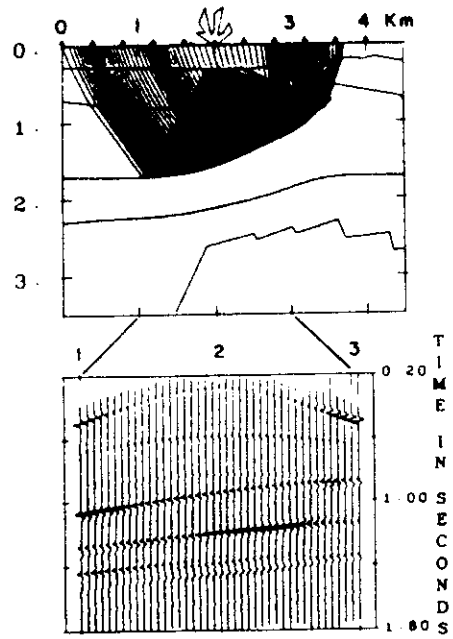

Fig. 13. Example of the calculation of a seismic reflection cross section by the Gaussian beam summation method. The model is a multilayered structure. All contributing reflections and multiples have been considered in the computation.