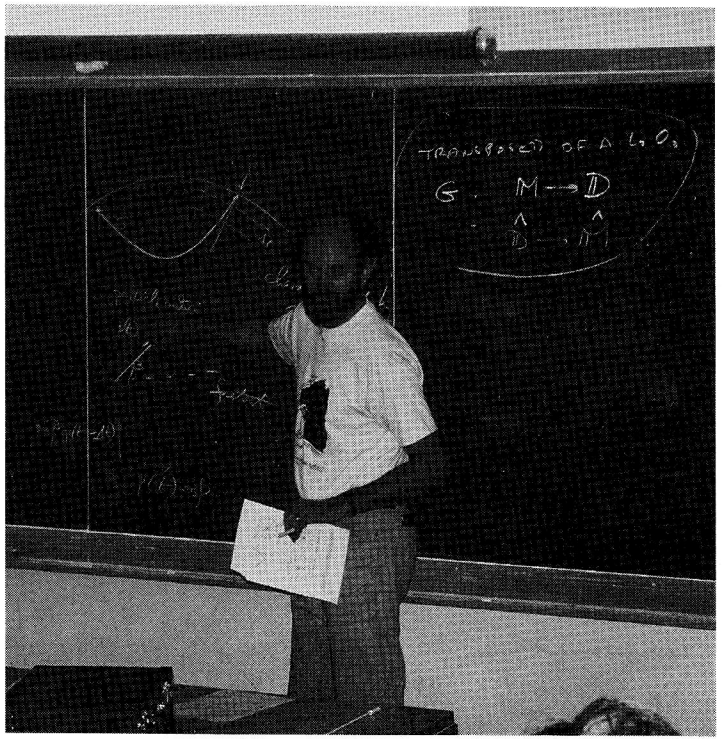
COURSE 2

THEORETICAL SEISMOLOGY

P.G. RICHARDS

Departement of Geological Sciences, and
Lamont–Doherty Geological Observatory
Columbia University
Palisades, New York, 10964, USA

Y. Desaubies, A. Tarantola and J. Zinn-Justin, eds.
Les Houches, Session L, 1988
Tomographie Océanographique et Géophysique /
Oceanographic and Geophysical Tomography



Contents

1. A review of underlying concepts

1.1. Introduction

The central exercise in seismology is understanding seismograms. Every generation of seismologists, for the last 100 years, has pushed for quantitative results, and for about the last 20 years this has meant that we try to interpret seismograms essentially by reproducing (i.e., modeling) the observed waveform shapes, as well as their times of arrival. For the resulting "synthetic seismograms" to be "successful" (i.e., more-or-less realistic), our model has to have a description of the seismic source; an appropriate theory of wave propagation; and of course a description of the Earth's internal structure. One can expect that in decades to come, all of these components of our modeling will continue to improve. Each generation has its own sense of what is meant by "realistic modeling", usually taking this to mean the latest improvement in capability.

These six lectures are principally concerned with the theory of seismic wave propagation. I shall discuss in broad terms the theory and practice of generating synthetic seismograms. The main idea is to work with the sum (or integral) of elementary solutions that more-or-less explicitly show the dependence on separate spatial and temporal variables. For example, the elementary solutions may be normal modes, or plane waves, or waves that are in part described by Bessel functions. The sum is taken over horizontal wavenumber (or ray parameter), and frequency. I shall concentrate on forward problems – that is, the problem of solving for seismic motions when source and structure are given, rather than on inverse problems, where the motion is to some extent known and the goal is to uncover properties of the structure, or, perhaps, of the source.

Seismic waves are observed across a great range of wavelengths, from thousands of kilometers down to about a meter – a factor of about a million. They are observed at distances ranging also from meters to thousands of kilometers from the causative source of the waves themselves. The medium in which the waves propagate – the Earth itself – is inhomogeneous on all scales, from grain-sized variability well below the wavelength of the shortest seismic wave, up to gross inhomogeneities like the difference between

mantle and cores, and continents and oceans – inhomogeneities thousands of kilometers across. Since the properties of a seismic wave depend broadly on the dominant wavelength of interest, on the distance from the source at which the solution is desired, and on the scale length of heterogeneity of Earth structure with which the wave has interacted, there is a truly vast number of different properties, whose effects are observed in seismograms, awaiting interpretation. No single method of generating synthetics is available. Rather, there are a large variety, that make different assumptions and approximations to assist in achievement of practical solutions. Every method, whether emphasizing a normal mode approach, or the interpretation of body waves, or whole classes of interaction of seismic waves with layered structures, has limitations. Although purely numerical methods are gaining in power, it may be expected that these too have restricted application, because foreseeable computers cannot handle the effects of 3-D varying Earth structure on waveforms that propagate more than a few hundred wavelengths.

Synthetic seismograms now do indeed quite often look remarkably like the data. Figures 1–4 give a short series of examples, and they represent a considerable achievement. They indicate in general terms that seismologists are now working with quite good models of the source; that they are using an appropriate theory of wave propagation; that reasonably accurate models of Earth structure are available; and that computational packages for generating sythetics are now effective and trustworthy.

However, these good results in figs. 1–4 must be understood as representing only a small part of the available types of seismic data. They emphasize a limited range of periods, and/or a limited range of time windows, and/or a limited range of source/receiver distances. In such cases we see that a deterministic approach to seismogram interpretation, in which one attempts to match observations "wiggle for wiggle," can succeed. There are other cases, notably in short period data, where synthetics are less sucessful.

These several examples in Figures 1–4 are representative of the best fits that can be achieved today, that is in 1988, between data and synthetics, for teleseismic waves. At these distances, the observed frequencies can often be modeled with a laterally homogeneous Earth model, or with lateral heterogeneities that may be handled as a perturbation. It is therefore natural in the theory of seismic wave propagation to emphasize methods for Earth models that vary in depth alone. The first five lectures in this series will take such an approach, and the sixth lecture will describe a procedure for handling waves in certain types of laterally varying structure, with dipping interfaces, where perturbation theory would not be adequate.

Before reviewing the concepts that lead to results such as those of figs. 1-

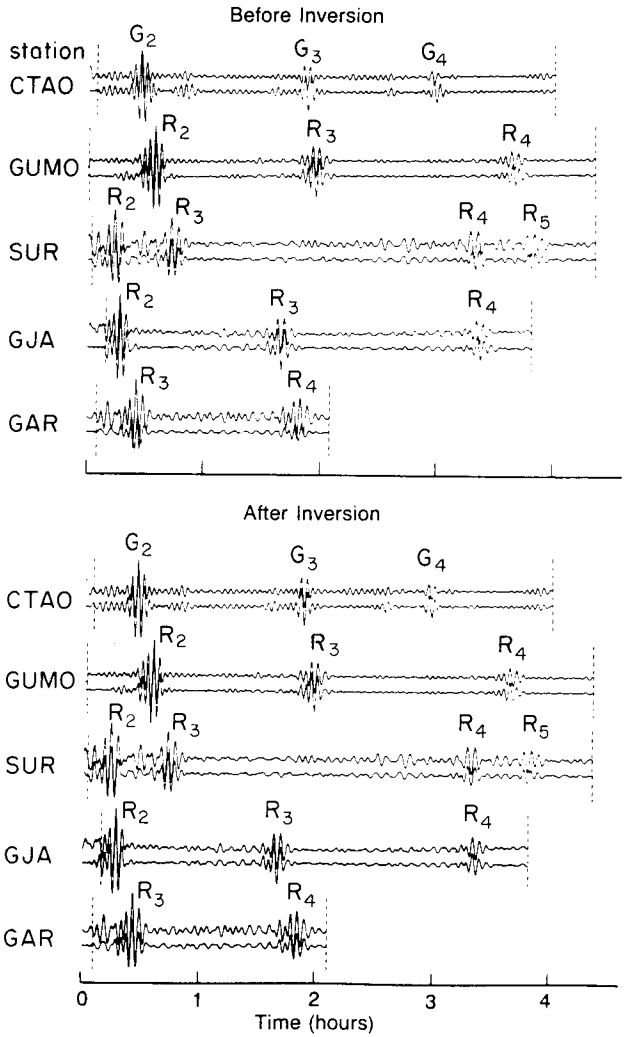
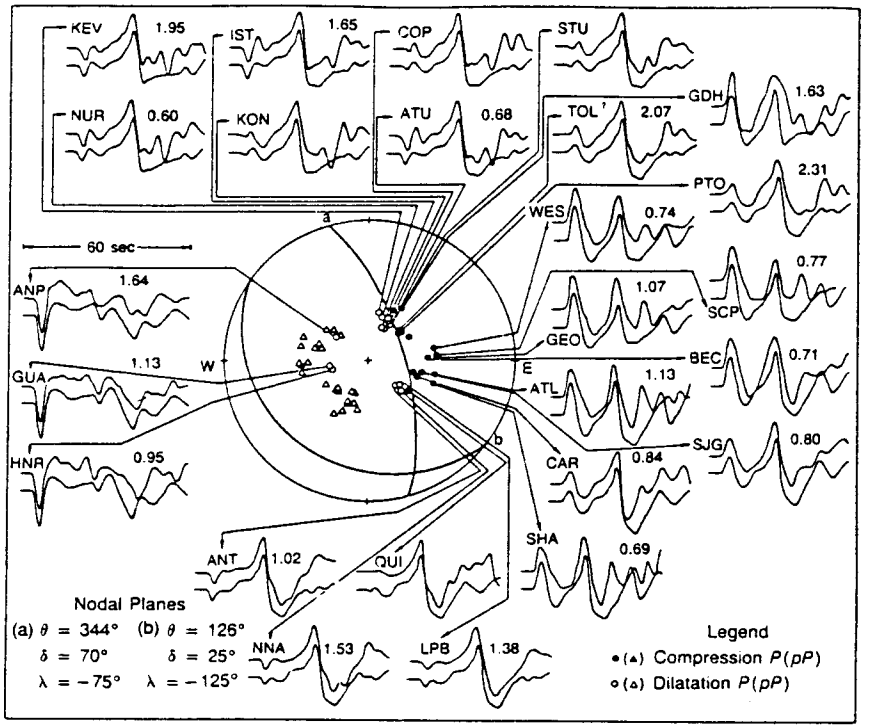
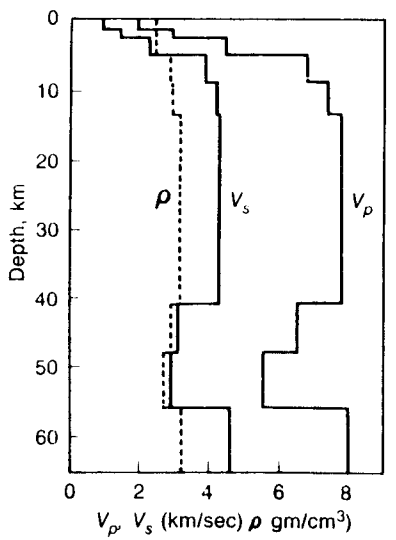


Fig. 1. A comparison of data (top trace in each pair) with synthetic seismograms for a shallow earthquake. The upper panel shows the result of a centroid moment tensor inversion in the Preliminary Reference Earth Model (PREM) of Dziewonski and Anderson (1981), and the lower panel shows the improved result after path corrections (controlled by lateral heterogeneity) have been determined. These synthetics are computed by summation of normal modes. The data is filtered to pass signals with periods longer than about a minute. From Woodhouse and Dziewonski (1984).





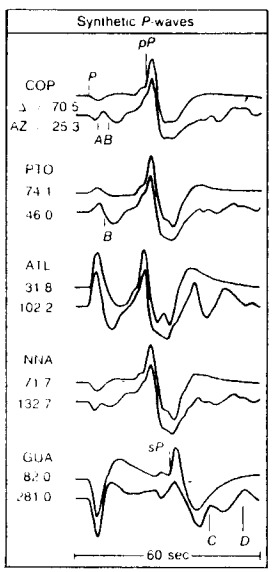
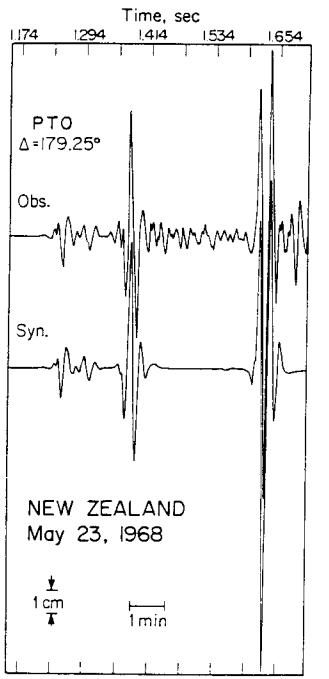


Fig. 2. A comparison between observed and computed P-waves for the 1965 April 29 Puget Sound, Washington, earthquake. The P waveforms were simultaneously modeled for source and structure parameters. Details of these waveforms, appearing between P and pP, were interpreted as reflections from interfaces above the source, associated with an inferred low velocity zone in the crust. The inferred crustal structure is shown on the lower left. On the lower right is a display to indicate the effect, on seismograms, of the low velocity zone. In each pair, the synthetic for a one-layer crust lies above, and the synthetic for a crust with low velocity zone lies below. Note that it is only a limited portion of the waveform that is affected l a fact which makes for difficulty in applying inverse methods. From Langston and Blum (1977). The synthetics are computed by geometrical ray theory.



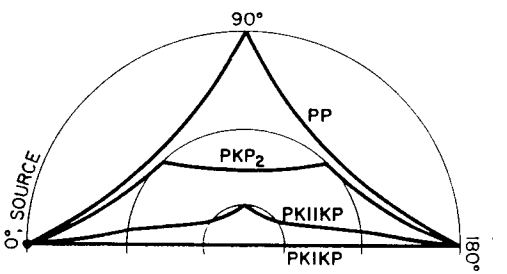
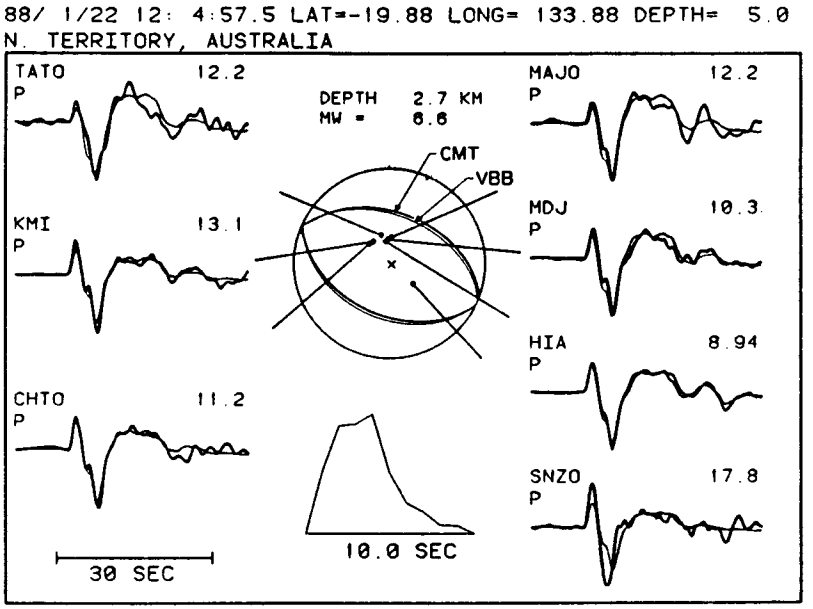


Fig. 3. A comparison between observed and computed waves that have interacted strongly with the Earth's cores. This earthquake occurred 1968 May 23 in New Zealand, and was observed almost at the antipode at a station in Portugal. The principal wave paths contributing in this special case are shown at the right. From Rial and Cormier (1981). Synthetics are computed by a method that makes allowance for the focussing effect that occurs near an antipode.

4, a comment is perhaps needed on future trends, and on whether the



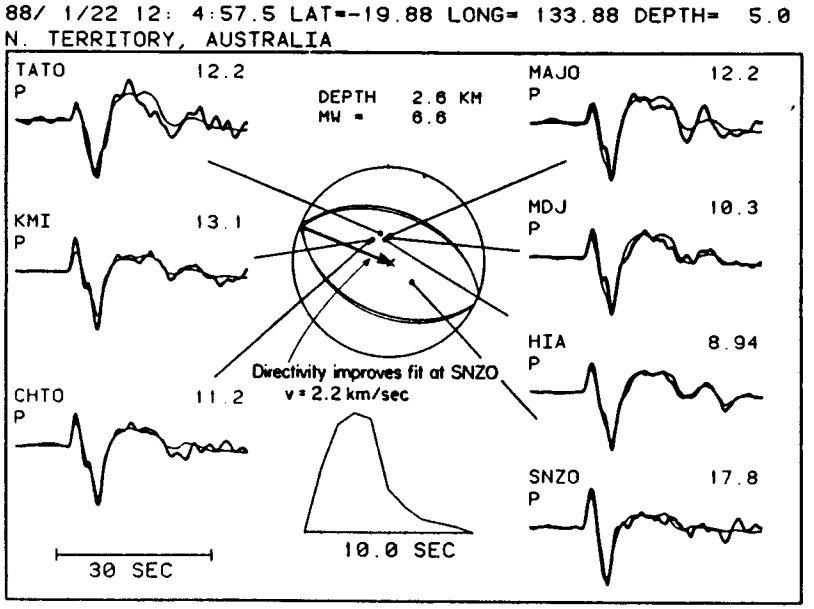


Fig. 4. An example of excellence of fit for a very shallow earthquake that occurred 1988 January 22 in Australia. The focal mechanism for an equivalent point source, and the shape of the pulse radiated from the focal sphere, were inverted for to obtain the upper panel. The data is the heavier line, and the synthetic is the lighter line. The number associated with each station is the maximum ground motion in microns. This is a thrusting mechanism, and though the fit is remarkably good at most stations, it is not so good at the station SNZO, for which the observed waveform is of relatively short duration. The analyst, Göran Ekström, found that the fit for this station could be improved, as shown in the lower panel, by modeling the earthquake source not as a point but as a line, on which the rupture propagated toward SNZO with a speed of 2.2 km/s. (Personal communication from Ekström, 1988.)

goodness of fit in this figure may be extended to a wider range of frequencies and source/receiver distances. The underlying question here, is whether a deterministic approach to waveform modeling can continue to provide satisfactory results. No doubt it will be possible to improve on current methods for generating synthetics – improvements that may accommodate some types of lateral heterogeneity, and that may provide for better ways to quantify attenuation. But in many cases the undersampling of seismic wave fields makes futile the attempt to explain every detail in the signal.

1.2. Key concepts in the theory of seismic wave propagation: Displacement, body force, traction, stress tensor

Under the small changes in stress and strain that typically accompany seismic motions in Earth materials, the simplifying assumption of a law of linear superposition is apparently accurate. That is, if two sets of seismic motions occur at the same time, each of which is separately understood, then the combined motion is simply the sum of the two.

By "motion" here, we mean *displacement*. In seismology it is natural to use the Lagrangian description of motion, in which the displacement $\boldsymbol{u}$, shown explicitly as a function of spatial position $\boldsymbol{x}$ and time t by writing $\boldsymbol{u}\ (\boldsymbol{x}\ , t)$, is defined as the vector distance of a particle at time t from the position $\boldsymbol{x}$ that it occupies at some reference time t_0. A seismogram, being a record of the motion of the part of the Earth to which the seismometer is attached, is directly a record of Lagrangian motion.

The forces acting upon the particles in a solid or fluid medium, due to physical processes outside the medium itself, and also the forces acting

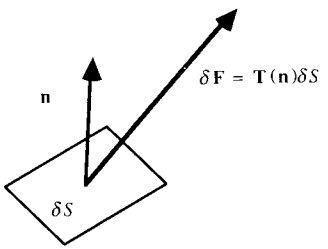


Fig. 5. The definition of the traction vector, $\boldsymbol{T}$ ($\boldsymbol{n}$), for an area element with normal $\boldsymbol{n}$.

between particles that are not adjacent, are called *body forces*. We use the notation $\boldsymbol{f}$ ($\boldsymbol{x}$, t) for the body force at time t acting per unit volume on the particle originally at position $\boldsymbol{x}$ at the reference time t_0.

To describe the internal forces acting mutually between adjacent particles within a continuum, we use the concepts of *traction* and *stress tensor*.

Traction is a vector, being the force acting per unit area across an internal surface within the continuum, and quantifying the contact force (per unit area) with which particles on one side of the surface act upon particles on the other side. For a given point of the internal surface S, traction is defined (see fig. 5) by considering an infinitesimal element of area δS, with normal $\boldsymbol{n}$ and centered on the point in question, and equating the force $\delta \boldsymbol{F}$ acting across the element to $\boldsymbol{T}\,\delta S$. The convention is adopted that $\delta \boldsymbol{F}$ (and hence $\boldsymbol{T}$) has the direction of force due to material on the side to which $\boldsymbol{n}$ points, acting upon material on the side from which $\boldsymbol{n}$ is pointing. The convention is easily remembered, if one recalls that "traction", in ordinary language, is the act of *pulling*.

We acknowledge the dependency of $\boldsymbol{T}$ on $\boldsymbol{n}$ by writing it out as $\boldsymbol{T}(\boldsymbol{n})$. In a fluid, $\boldsymbol{T}$ ($\boldsymbol{n}$) would be parallel to $\boldsymbol{n}$, and typically in the opposite direction. Thus, for a point in a fluid at pressure p, traction is related to pressure by $\boldsymbol{T}(\boldsymbol{n}) \cdot \boldsymbol{n} = -p$ for all orientations of the area element (i.e., all choices of the direction $\boldsymbol{n}$). The difference in sign is a reflection of the basic difference between the concepts of pressure ("pushing") and traction ("pulling"). In a solid, $\boldsymbol{T}$ ($\boldsymbol{n}$) will not in general be parallel to $\boldsymbol{n}$, because of the possible existence of shearing forces acting across the area element.

If the direction of $\boldsymbol{n}$ is reversed, then the associated traction vector is reversed, giving

$$\boldsymbol{T}(-\boldsymbol{n}) = -\boldsymbol{T}(\boldsymbol{n}). \tag{1.1}$$

This follows from Newton's law of action and reaction.

The traction $\boldsymbol{T}$ may be constructed for any particular choice of orientation of the area element by using the stress tensor.

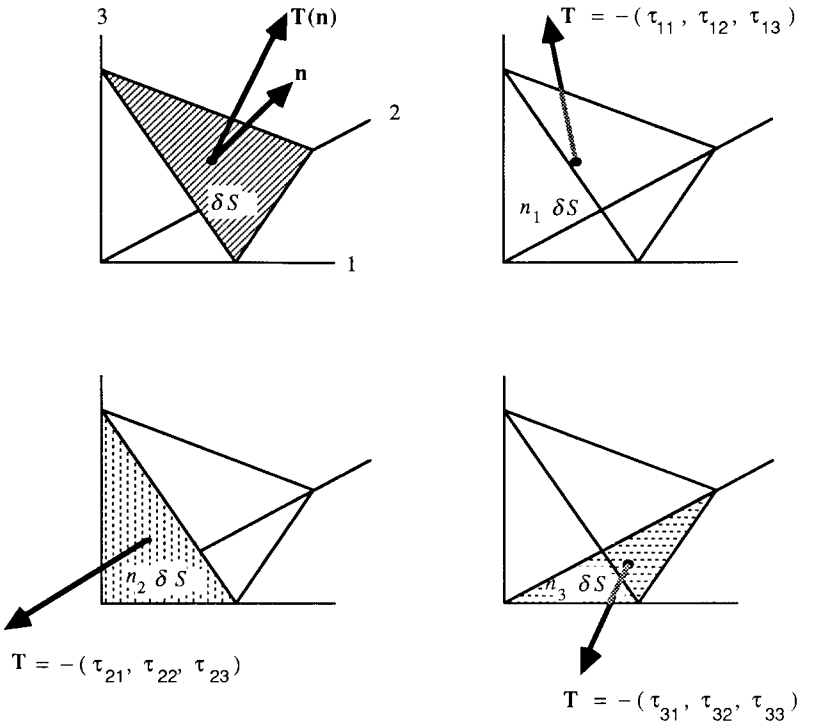


Fig. 6. The tractions acting on each face of a tetrahedron, three of whose sides are in the planes of a cartesian coordinate system. Thus, the oblique face is in the direction of interest, $\boldsymbol{n}$, and the traction on the other three faces is related to the components of the stress tensor.

The stress tensor, denoted by τ, is a cartesian second-order tensor having components τ_{kl}, defined as the l-th component of the traction acting across the plane normal to the k-th axis, due to material with greater x_k acting upon material with lesser x_k. If we consider the balance of forces acting on an elemental tetrahedron, with three of its faces in the planes of a cartesian coordinate system (see fig. 6), then

$$\begin{aligned}\boldsymbol{T}(\boldsymbol{n})\delta S - (\tau_{11}, \tau_{12}, \tau_{13}) n_1 \delta S \\ - (\tau_{21}, \tau_{22}, \tau_{23}) n_2 \delta S - (\tau_{31}, \tau_{32}, \tau_{33}) n_3 \delta S = (0, 0, 0).\end{aligned} \tag{1.2}$$

This result holds even if the elemental volume has finite acceleration, since the forces acting per unit volume are dominated by the forces acting per unit area.

The i-th component of eq. (1.2) is written, with the Einstein summation

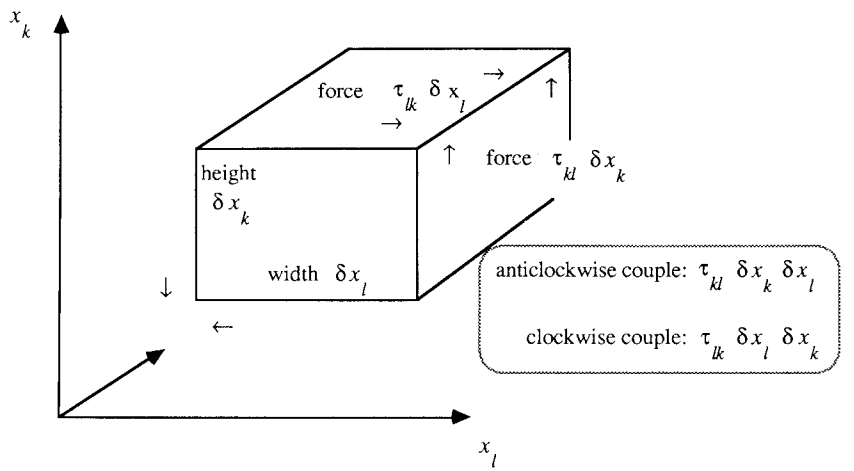


Fig. 7. Forces and couples, acting per unit length of a prism. (No summation over subscripts, in this figure. Forces and couples act in the directions indicated. It is possible to give an interpretation using the summation convention, reaching the same conclusion that the stress tensor must be symmetric.)

convention, as

$$T_i = \tau_{ji} n_j, \tag{1.3}$$

The nine symbols denoted by τ_{kl} ($k = 1$, 2 or 3; $l = 1$, 2 or 3) must have the symmetry

$$\tau_{lk} = \tau_{kl}. \tag{1.4}$$

since otherwise there would be a net couple acting to give angular momentum to an elemental prism with sides perpendicular to the x_k and x_l axes (see fig. 7).

It follows from eqs. (1.3) and (1.4) that

$$T_i = \tau_{ij} n_j, \tag{1.5}$$

and hence τ_{ij} must indeed be the components of a second order cartesian tensor, since it relates the components of two vectors.

The wave equation, strain, Hooke's Law, and more general equations of state . We can now give the relationship between acceleration, body force and stress tensor, at a point within a solid continuum.

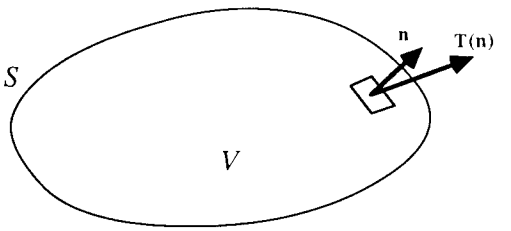


Fig. 8. A material volume V of the continuum, with surface S.

Equating the rate of change of momentum of particles constituting V to the forces acting on these particles,

$$\frac{\partial}{\partial t} \iiint_V \rho \frac{\partial \boldsymbol{u}}{\partial t}\, \mathrm{d}V = \iiint_V \mathrm{f}\, \mathrm{d}V + \iint_S \boldsymbol{T}(\boldsymbol{n})\, \mathrm{d}S.$$

The left-hand side can be written as

$$\iiint_V \rho \frac{\partial^2 \boldsymbol{u}}{\partial t^2}\, \mathrm{d}V,$$

since the particle mass $\rho\, \mathrm{d}V$ is constant in time. The second term of the right-hand side, together with eq. (1.5) and the Gauss divergence theorem, may be written as the vector whose i-th component is

$$\iiint_V \tau_{ij,j}\, \mathrm{d}V.$$

It follows, by continuity of the integrand of an arbitrary volume integral, that

$$\rho \ddot{u}_i = f_i + \tau_{ij,j}, \tag{1.6}$$

in which $\ddot{u}_i$ is $\partial^2 u_i / \partial t^2$. This is perhaps the single most important form of the equation of motion of particles in a solid continuum. It can be interpreted as an exact equation even if the displacements are large (see Aki and Richards, 1980, p.18). It makes no assumption about linearity, nor about the equation of state relating stress and deformation of the material. It simply relates the acceleration directly to body force and stress gradient.

To analyse the distortion occurring within a solid, we use the *strain tensor*. The particle initially at $\boldsymbol{x}$ moves to the position $\boldsymbol{x} + \boldsymbol{u}$, and the

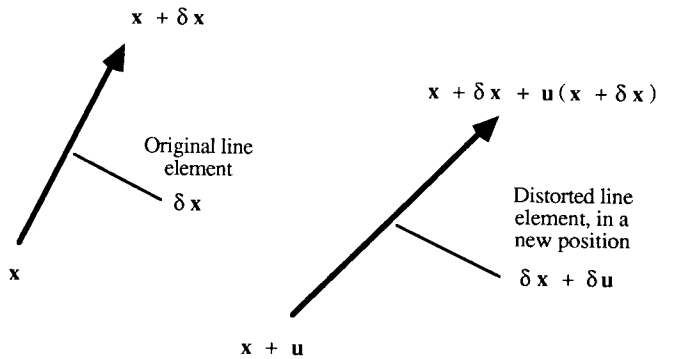


Fig. 9. An original line element moves to a new position. The neighborhood of $\boldsymbol{x}$ is partly translated, partly rotated as a rigid body, and also distorted.

relation $\boldsymbol{u} = \boldsymbol{u}\,(\boldsymbol{x}\,)$ describes the displacement field, becoming $\boldsymbol{u} = \boldsymbol{u}\,(\boldsymbol{x}\,, t)$ in general if the displacement continues to change. If $\boldsymbol{u}$ is a constant, independent of $\boldsymbol{x}$, the particles originally in the vicinity of $\boldsymbol{x}$ move all together, and there is merely a translation to the new position, with no distortion. If $\boldsymbol{u}$ is the result merely of a rigid rotation, then again there is no distortion of the neighborhood of $\boldsymbol{x}$. To quantify true distortions, it is necessary to study $\boldsymbol{u}\,(\boldsymbol{x} + \delta\boldsymbol{x}\,)$ (i.e., the displacement of the particle originally at $\boldsymbol{x} + \delta\boldsymbol{x}$) and see how the line element originally given by δx has changed length, as a function of the direction of $\delta\boldsymbol{x}$, after the displacement described by $\boldsymbol{u}$ takes place (see fig. 9).

The line element $\delta\boldsymbol{x}$ becomes $\delta\boldsymbol{x} + \delta\boldsymbol{u}$, which may be written down as the difference between the two endpoints of the line element in its new position, giving

$$\delta\boldsymbol{x} + \delta\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{x} + \delta\boldsymbol{x}) + \boldsymbol{x} + \delta\boldsymbol{x} - \boldsymbol{u}(\boldsymbol{x}) - \boldsymbol{x},$$

and the change $\delta\boldsymbol{u}$ in this element is therefore, to first order,

$$\delta\boldsymbol{u} = (\delta\boldsymbol{x} \cdot \nabla)\boldsymbol{u}.$$

The new length is (to first order, i.e., neglecting $\delta\boldsymbol{u} \cdot \delta\boldsymbol{u}$)

$$\begin{aligned} |\delta\boldsymbol{x} + \delta\boldsymbol{u}| &= \sqrt{\delta\boldsymbol{x} \cdot \delta\boldsymbol{x} + 2\delta\boldsymbol{u} \cdot \delta\boldsymbol{x}} \\ &= \sqrt{\delta x_i \delta x_i + 2\delta x_i \delta x_j u_{i,j}} \\ &= \sqrt{\delta x_i \delta x_i + 2\delta x_i \delta x_j e_{ij}}, \end{aligned} \tag{1.7}$$

where $e_{ij} = (u_{i,j} + u_{j,i})/2$ is the definition of the (i, j) component of the *strain tensor*. It follows that the extensional strain (i.e., the extension per unit original length), of the line element originally in the ν direction (ν is the unit vector given by $\delta\boldsymbol{x}/\,|\delta\boldsymbol{x}|\,$), is $e_{ij}\nu_i\nu_j$.

The i-th component of the change $\delta\boldsymbol{u}$ in the line element is given by

$$\delta u_i = (\delta u_i/\delta x_j)\,\delta x_j = e_{ij}\delta x_j + (1/over2)\,(u_{i,j} - u_{j,i})\,\delta x_j. \tag{1.8}$$

If spatial gradients of the displacement field are small, the last term of eq. (1.8) can be recognized as an infinitesimal rotation of amount $(\mathrm{curl}\,\boldsymbol{u})/2$, because of the identities

$$(u_{i,j} - u_{j,i})\,\delta x_j \equiv \varepsilon_{ijk}\varepsilon_{jlm}u_{m,l}\delta x_k \equiv (\mathrm{curl}\,\boldsymbol{u} \times \delta\boldsymbol{x})_i.$$

Putting these results together, the line element undergoes a displacement characterized by three terms: a translation of amount $\boldsymbol{u}$; a rigid body rotation of amount $(\mathrm{curl}\,\boldsymbol{u})/2$; and a displacement due to distortion (true strain) whose i-th component is $e_{ij}\delta x_j$.

More than 300 years ago, Robert Hooke reported the results of his experiments on springs. He gave the puzzle *ceiiinosssttuv*, which re arranges to the Latin phrase *ut tensio, sic vis*, translating to *as the extension, so the force*, i.e., the extension of a spring is proportional to the tension. Note that Hooke's Law is an experimental result, from which one is led to define the behavior of an ideal linearly elastic string. The modern generalization of Hooke's Law is that each component of the stress tensor is a linear combination of all components of the strain tensor, i.e., that there exist real constants c_{ijkl} such that

$$\tau_{ij} = c_{ijpq}e_{pq}. \tag{1.9}$$

The c_{ijkl} are constants, in that they are independent of strain. However, they can vary with position. They are the 81 components of a fourth order cartesian tensor, which can be shown to have the symmetries

$$c_{jipq} = c_{ijqp} = c_{pqij} = c_{ijpq}.$$

If the material whose stress–strain relation is (1.9) is isotropic, then the tensor c must be an isotropic tensor. The only possible fourth order tensor, with the symmetries noted above, is one with the form

$$c_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}), \tag{1.10}$$

in which there are now only two elastic constants, the Lamé constants λ and μ.

Just as an ideal gas has its pressure determined once its volume is given, and the relationship is different under isothermal or adiabatic conditions, we can regard eq. (1.9) as the equation of state for an ideal (i.e., linearly elastic) solid. It is usually used in geophysics with the assumption of adiabatic conditions, because when the solid material of the Earth is stressed and strained by the passage of a seismic wave, there is not enough time for diffusive equilibration of temperature to occur.

To say that the components of $\boldsymbol{c}$ are *real* (e.g., that λ and μ are real in an isotropic solid), is to say that stress and strain change in phase, with no hysteresis. If a material that is cycled in stress, with frequency ω, exhibits hysteresis, the associated phase delay of strain with respect to stress may be quantified by working with components of $\boldsymbol{c}$ that are *complex*, possibly with different values at different frequency. Net work is done, on the solid, in each complete cycle. In the time domain, this type of intrinsic friction (associated with attenuation of a seismic wave, as it propagates) can be quantified by generalizing Hooke's Law, (1.9), to a linear relation between stress, stress rate, strain, and strain rate. (We take this up again, in section 1.3, in a discussion of attenuation.)

The wave equation for a linear elastic solid is obtained by substituting (1.9) into (1.6), giving a vector equation, the i-th component of which is

$$\begin{aligned} \rho\ddot{u}_i &= f_i + (c_{ijpq}e_{pq})_{,j} \\ &= f_i + (c_{ijpq}u_{p,q})_{,j}. \end{aligned} \tag{1.11}$$

For an isotropic linear elastic solid we can use (1.10) and obtain

$$\rho\ddot{u}_i = f_i + (\lambda u_{k,k})_{,i} + (\mu u_{i,j})_{,j} + (\mu u_{j,i})_{,j}. \tag{1.12}$$

In a homogeneous part of such a solid, λ and μ are independent of position, and the result can be further simplified:

$$\rho\ddot{\boldsymbol{u}} = \boldsymbol{f} + (\lambda+\mu)\,\nabla(\nabla\cdot\boldsymbol{u}) + \mu\nabla^2\boldsymbol{u},$$

or

$$\rho\ddot{\boldsymbol{u}} = \boldsymbol{f} + (\lambda+\mu)\,\nabla(\nabla\cdot\boldsymbol{u}) - \mu\nabla\times(\nabla\times\boldsymbol{u}). \tag{1.13}$$

The Green function, reciprocity, and representations of the solution to the elastic wave equation . In the theory of scalar potentials for gravity fields, one can develop solutions to Poisson's equation,

$$\nabla^2\phi = 4\pi\gamma\rho,$$

by noting the special solution, called the Green function, $G(\boldsymbol{x} : \xi)$, which satisfies

$$\nabla^2 G = \delta(\boldsymbol{x} - \xi)\ .$$

Solving for ϕ as a function of position $\boldsymbol{x}$, due to a density distribution $\rho = \rho$ ($\boldsymbol{x}$), is simple once the "source term" in Poisson's equation is re-written as an integral over the sources which generate G. Thus, from writing the Poisson equation itself as

$$\nabla^2\phi = 4\pi\gamma \iiint_V \rho(\xi)\ \delta\ (\boldsymbol{x} - \xi)\ \mathrm{d}V\ (\xi)\ ,$$

where the integral is taken over the volume V in which the density distribution is non-zero, it follows that

$$\phi(\boldsymbol{x}) = 4\pi\gamma \iiint_V \rho(\xi) G(\boldsymbol{x};\xi)\,\mathrm{d}V(\xi)\ . \tag{1.14}$$

In this case, the specific form for G can be given as

$$G(\boldsymbol{x};\xi) = (-1/4\pi)/|\boldsymbol{x} - \xi|\,. \tag{1.15}$$

The solution (1.14) for ϕ is essentially a sum (a volume integral, in this case) over the potential field of individual point masses, weighted by the density distribution within V. In the present example, because an explicit expression is available for the Green function, an explicit volume integral is available for evaluating the potential ϕ.

All of these concepts generalize to solutions of the wave equation for displacements within solids, (1.6), provided only that the equation is linear, allowing for solutions obtained by superposition of elementary (Green function) solutions. Such linearity is found, in (1.11) (1.13). However, although the concepts used above to solve Poisson's equation can be applied to solve problems in linear elasticity, the complicating details tend to obscure this underlying simplicity. Complications arise from the time-dependence and vector nature of the displacement field and from directionalities at source and receiver, and from the application of surface boundary conditions. The Green function becomes a tensor (or, as some prefer to call it, a dyadic), and can be evaluated explicitly and exactly (as in eq. (1.15) for the scalar potential problem) only in the simplest of elasticity problems. This section concludes with a summary of the main results associated with the Green function in elasticity theory. Our subsequent work will often entail explicit

but approximate evaluations of the Green function, and their use to obtain synthetic seismograms meeting some desired standard of accuracy.

We shall use $G_{in}(\boldsymbol{x}\ , t\ ;\ \xi,\ \tau)$ to denote the i-th component of displacement at $\boldsymbol{x}$ at time t, due to a unit impulse applied as a body force at time τ on the particle at position ξ. From eq. (1.11) it follows that $\boldsymbol{G}$ satisfies

$$\rho\frac{\partial^2}{\partial t^2}G_{in} = \delta_{in}\delta(\boldsymbol{x}-\xi)\,\delta(t-\tau) + \frac{\partial}{\partial x_j}(c_{ijpq}\frac{\partial}{\partial x_q}G_{pn})\ . \tag{1.16}$$

However, this does not yet specify $\boldsymbol{G}$ uniquely. Equation (1.16) is valid throughout the volume V of the solid of interest, provided there is continuity of material properties. Across internal interfaces, where material properties (e.g., ρ, λ and μ) are discontinuous, we commonly require that displacements and tractions are continuous. To give a unique specification of $\boldsymbol{G}$, it is necessary to state the boundary conditions on the surface, S, of V. In practice, several different choices may be made, each useful in different circumstances (see Aki and Richards, 1980, sections 2.4, 2.5, and 3.1 for discussion). For example, S may be *rigid* ($\boldsymbol{u} = \mathbf{o}$ on S), or *traction free* ($\boldsymbol{T}\ (\boldsymbol{n}\ ) = 0$, on S, with $\boldsymbol{n}$ normal to S, often called the *stress free* condition).

$\boldsymbol{G}$ depends only on the time difference between t and τ :

$$\mathrm{G}(\boldsymbol{x},t;\xi,\tau) = \boldsymbol{G}(\boldsymbol{x},t-\tau;\xi,0)\ .$$

Also, there is an important reciprocity, an equality of certain displacement components if the source and receiver positions are interchanged. Expressed in terms of cartesian components, it is

$$G_{in}(\boldsymbol{x},t;\xi,0) = G_{ni}(\xi,\boldsymbol{t};\boldsymbol{x},0)\ .\ \text{(See also fig. 10)}. \tag{1.17}$$

Reciprocity is a fundamental property of the response of linear systems. Note that it is the quantity for which the wave equation itself is written, that has reciprocity when source and receiver are interchanged. Reciprocity is not claimed for derived non linear fields, such as energy. fig. 11 shows a situation in which the reciprocity is not obvious.

Just as eq. (1.14) is a way of representing the solution of Poisson's equation, so too we can write down representations of elastic displacement solutions of equations such as (1.11)-(1.13). If the elastic material is unbounded, and motion from initial rest is due solely to a body force active

G_{in} (x, t; ξ, 0) = G_{ni} (ξ, t; x, 0) . (See also Figure 1.7.)

(1.17)

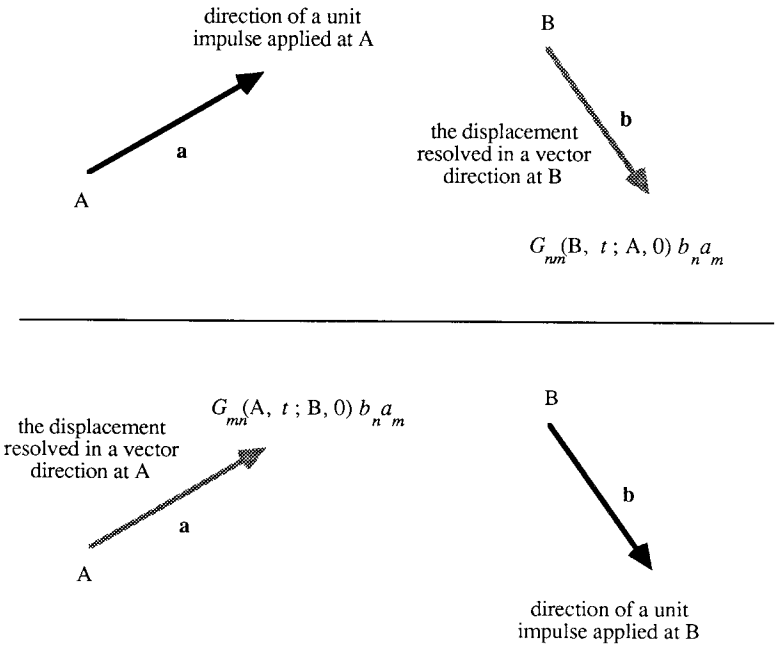


Fig. 10. An illustration of reciprocity. $\boldsymbol{a}$ and $\boldsymbol{b}$ are unit vectors. A unit impulse applied at A in direction $\boldsymbol{a}$ produces at B a displacement whose component in the $\boldsymbol{b}$ direction is the same as the component of displacement at A in the $\boldsymbol{a}$ direction if the source is instead a unit impulse applied at B in the $\boldsymbol{b}$ direction. The first situation is shown in the upper part of the figure, the second situation in the lower. The Einstein summation convention, of course, is applied.

only within region V after time $t = 0$, then by steps similar to the derivation of (1.14) we can show

$$u_i(\boldsymbol{x},t) = \int_0^\infty \mathrm{d}\tau \iiint_V f_p(\xi,\tau)G_{ip}(\boldsymbol{x},t-\tau;\xi,0)\,\mathrm{d}V(\xi). \tag{1.18}$$

This result is also valid for the interior of a solid occupying the volume V, if its surface S is one on which displacement or traction is zero (for purposes of defining both $\boldsymbol{u}$ and $\boldsymbol{G}$). In this representation, the complications present in evaluating the Green function are suppressed simply by writing the fundamental solution notionally as $\boldsymbol{G}\ (\boldsymbol{x}\ , t-\tau\ ;\ \xi,\ 0)$.

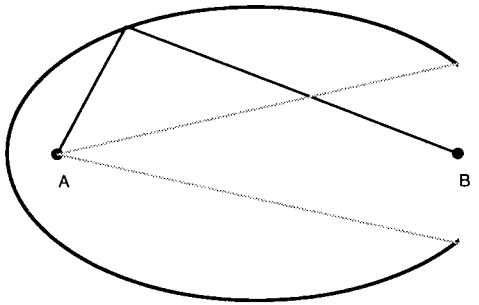


Fig. 11. Another example of reciprocity, for equal isotropic point sources of pressure placed at foci of an elliptical reflector. Think of this as sound waves propagating within a homogeneous fluid. Part of the ellipse has been cut away, so that half of the energy radiated from the source at B will escape without reflection, whereas a much lesser portion of the energy from the source at A escapes. Nevertheless, the singularity in the pressure at A due to the source at B equals the singularity at B due to the source at A. (See articles by Eisner, 1983, and Dahlen and Odom, 1984.)

For modeling the earthquake source, that is as a source of seismic waves, theories of linear wave propagation can still be used if non-linearities of motion in the source region itself are somehow excluded. This can usefully be done if faulting is regarded as slip across an internal surface, and the slip is treated as a discontinuity in the displacement field. The finite source region is then an area – the fault surface itself – rather than a volume as in eq. (1.19). The classical force equivalent for planar fault slip in an isotropic medium is a double couple distribution over the fault, as was first shown for a finite fault by Maruyama in 1963. More generally, for slip across a possible non planar fault surface A, with normal ν, in an anisotropic medium, the source has force equivalents given by couples described by

$$[u_i]\nu_j c_{ijpq}.$$

Here, $[u_i]$ is the i- component of the displacement discontinuity, taken across A in the ν direction. This force equivalent was first discovered by Burridge and Knopoff in 1964 (see also Aki and Richards, 1980, chapter 3). The couple has force in the p-direction, and arm in the q-direction, and, as described further in section 1.2, such a source at $(\xi\ ,\ \tau\ )$ generates displacement at $(\boldsymbol{x}\ ,\ t)$ in an amount given by

$$\frac{\partial}{\partial \xi_q} G_{np}(\boldsymbol{x}, t; \xi, \tau).$$

It follows that a representation of the seismic motions, set up by time-dependent slip across an internal fault surface, and propagated throughout the Earth, is

$$u_n(\boldsymbol{x}, t) = \int_0^\infty \mathrm{d}\tau \iint_A [u_i]\nu_j c_{ijpq} \frac{\partial}{\partial \xi_q} G_{np}(\boldsymbol{x}, t; \xi, \tau)\, \mathrm{d}A(\xi). \qquad (1.19)$$

For the waves set up by an explosion or an air gun, i.e., a point source, the volume integral in (1.18) reduces just to an evaluation at the source point. The same is true for an earthquake, if one is interested only in seismic waves with wavelength long enough that the source is effectively a point. In all these cases, the source is quantified by a particular choice of body force that is active at only one location. This is how figs. 1–3 and the upper part of fig. 4 were generated. For short wavelengths radiated by an earthquake, it may be necessary to recognize that the source region is of finite spatial extent. This is what is needed to interpret the signal at SNZO satisfactorily, in the lower part of fig. 4. In this way, seismologists are beginning to be able to quantify the spatial extent of faulting, and the direction of rupture, in earthquakes that may be located thousands of kilometers away.

2. Some important solutions of the wave equation

2.1. Introduction

This lecture is a commentary on special solutions of the second-order wave equation, for displacement in very simple types of elastic media. It is principally a review of classical material, in that many of the key ideas are over a hundred years old. Yet it is important to be reminded of these ideas, because they have given us the concepts, and even the language, needed to discuss and to understand the seismic waves propagating in more complex structures.

For example, one might wish in general to understand the way in which seismic waves are scattered by inhomogeneities. It is soon clear that there are two quite different sets of broad ideas that lead to different computational procedures for quantifying such scattering. On the one hand, we may regard the inhomogeneities of interest as having a boundary that is a surface across which there is a discontinuous change in material properties. The wave equation (e.g., 1.12) is not valid on this surface, and it is natural then to develop the study of scattering in terms of boundary conditions and reflection coefficients. On the other hand, the inhomogeneity

may instead be associated with material properties that change continuously, albeit with regions where the spatial rate of change is anomalously high. The wave equation is then valid everywhere, and must intrinsically be capable of quantifying "reflection" from regions of high gradient in elastic constants and density.

The point here is that an informed discussion, of which framework for handling scattering is more appropriate for giving numerical results in a given situation, requires first an understanding of the basic terminology, which was developed for wave propagation in very simple examples.

This lecture describes how elastic waves in homogeneous media can be studied via potentials, and gives the famous classical solution for the displacement field set up by a uni-directional force, suddenly applied within the medium, and acting at a single point. The moment tensor is described, for point sources of importance in geophysics, and the exact displacement solution for homogeneous media is approximated and generalized to give the geometrical ray theory solution for waves in a smoothly varying inhomogeneous medium.

2.2. Elastic waves in homogeneous media

Potentials . For a homogeneous medium, with wave propagation governed by eqs. (1.13), the solution for any displacement can be written in the form

$$\boldsymbol{u} = \nabla\phi + \nabla \times \psi, \tag{2.1}$$

in which potentials ϕ and ψ satisfy second-order wave equations,

$$\begin{aligned} \frac{1}{\alpha^2}\frac{\partial^2\phi}{\partial t^2} &= \nabla^2\phi + \frac{\Phi}{\lambda+2\mu}, \\ \frac{1}{\beta^2}\frac{\partial^2\psi}{\partial t^2} &= \nabla^2\psi + \mathrm{i}\frac{\boldsymbol{\Psi}}{\mu}, \end{aligned} \tag{2.2}$$

having source terms [the last term in each of eqs. (2.2)] that are given by potentials for the body force such that

$$\mathrm{f} = \nabla\Phi + \nabla \times \boldsymbol{\Psi}. \tag{2.3}$$

In these equations, α is the P-wave speed, $\sqrt{[(\lambda+2\mu)/\rho]}$, and β is the S-wave speed, $\sqrt{(\mu/\rho)}$. The vector displacement potential, ψ, and the vector body force potential, $\boldsymbol{\Psi}$, both have zero divergence.

It is not difficult to show that if potentials satisfying (2.2) and (2.3) exist, then the displacement generated by (2.1) does satisfy the equation of motion, (1.13). What is much less obvious, but more important and still true, is that for a displacement $\boldsymbol{u}$ satisfying the wave equation with body force $\boldsymbol{f}$, potentials do exist with the properties (2.2) and (2.3). This is known as Lamé's theorem (Aki and Richards, 1980, pp 68–69), and for over a hundred years it has been useful for solving elasticity problems for homogeneous media because it allows a treatment separately of the P-wave and S-wave motions (coupled, perhaps, by boundary conditions).

Note that once the body force potentials are known, solutions of the wave eqs. (2.2) are readily available. For example, the first of these equations is usefully approached by seeing that

$$G(\boldsymbol{x},t) = \frac{1}{4\pi|\boldsymbol{x}|}\delta\left(t - \frac{|\boldsymbol{x}|}{v}\right), \tag{2.4}$$

is a solution of

$$\frac{1}{v^2}\frac{\partial^2 G}{\partial t^2} = \nabla^2 G + \delta(\boldsymbol{x})\,\delta(t). \tag{2.5}$$

This is a scalar Green function, with which it is easy to obtain the solution for displacement potential ϕ, given the force potential Φ.

The use of potentials in studies of elastic waves is but one example of many, of a classical technique that is now being superceded by other methods more directly suited to giving numerical solutions. Thus, the analytic approach to solving for waves in plane-layered media can be based more directly on matrix methods of working with physical components of displacement and stress, in each layer, rather than relying on potentials. And in cases of more general 2D and 3D inhomogeneity, the wave equation may be approximated by finite differencing, and then "integrated" directly to give a displacement solution. Nevertheless, potentials are often well suited to the solution of classical problems in homogeneous media, and the following solution is a famous example.

Stokes' solution . For an infinite homogeneous elastic solid, the solution for the displacement due to a force applied at a particular fixed point, and in a fixed direction, is known exactly. It was first published by Stokes in 1849, and is important in allowing wave properties to be described explicitly, in ways that led to an introduction of terminology that is still used today to describe solutions in more complicated media.

To sketch a derivation of Stokes' solution, it is helpful to consider an explicit source description. Thus, let us take the point of application of the

force as the origin of cartesian coordinates. Take the direction of application as the x_1 direction. Take the time-varying magnitude of the force as $F(t)$. Then, the body force $\boldsymbol{f} = \boldsymbol{f}\ (\boldsymbol{x}\ ,\ t)$ is given as:

$$f_i = \delta_{1i}\delta\,(\mathrm{x})\,F\,(t)\,.$$

One can then obtain the resulting displacement field by breaking the problem down into steps along lines suggested by Lamé's theorem. These obvious steps are:

(i) Find the body force potentials Φ and $\boldsymbol{\Psi}$ (with $\nabla\cdot\boldsymbol{\Psi} = 0$) that generate $\boldsymbol{f}$ as in (2.3).

(ii) Solve the separate wave eqs. (2.2) with these body force potentials as source terms, to obtain the displacement potentials ϕ and ψ (with $\nabla\cdot\psi = 0$) explicitly.

(iii) Generate the derived displacement field, $\boldsymbol{u} = \boldsymbol{u}\ (\boldsymbol{x}\ ,\ t)$, via (2.1).

The outcome can be found in this way as a sum of 3 terms:

$$\begin{aligned} u_i(\boldsymbol{x},t) =& \frac{1}{4\pi\rho\alpha^2}\gamma_1\gamma_i\frac{1}{R}F(t-\frac{R}{\alpha}) \\ &- \frac{1}{4\pi\rho\beta^2}(\gamma_1\gamma_i-\delta_{1i})\frac{1}{R}F(t-\frac{R}{\beta}) \\ &+ \frac{1}{4\pi\rho}(3\gamma_1\gamma_i-\delta_{1i})\frac{1}{R^3}\int_{R/\alpha}^{\mathrm{R}/\beta}\tau' F(t-\tau')\,\mathrm{d}\tau'. \end{aligned} \tag{2.6}$$

Here, γ is a unit vector in the direction from source to receiver (aligned along a straight line ray, for a homogeneous medium), and R is the source–receiver distance; $R = |\boldsymbol{x}\ |$. Details of the derivation are laid out in Aki & Richards (1980, pp70–73), and from this solution it is simple to state explicitly the form of the Green function itself, $\boldsymbol{G}\ (\boldsymbol{x}\ ,\ t\ ;\ \xi,\ \ \tau)$, in this simplest of media (isotropic, homogeneous, whole space). It is:

$$\begin{aligned} G_{ij}(\boldsymbol{x},t;\xi,\tau) =& \frac{1}{4\pi\rho\alpha^2}\gamma_i\gamma_j\frac{1}{R}\delta(t-\tau-\frac{R}{\alpha}) \\ &- \frac{1}{4\pi\rho\beta^2}(\gamma_i\gamma_j-\delta_{ij})\frac{1}{R}\delta(t-\tau-\frac{R}{\beta}) \\ &+ \frac{1}{4\pi\rho}(3\gamma_i\gamma_j-\delta_{ij})\frac{1}{R^3}\int_{R/\alpha}^{\mathrm{R}/\beta}\tau'\delta(t-\tau-\tau')\,\mathrm{d}\tau', \end{aligned} \tag{2.7}$$

with $R = |\boldsymbol{x} - \xi|$.

Note that the dependence on t and τ is only via the difference, $t-\tau$. So, it suffices to work with $\boldsymbol{G}\ (\boldsymbol{x}\ ,\ t;\ \xi,\ 0)$ and replace t by $t-\tau$ to evaluate $\boldsymbol{G}$

$(\boldsymbol{x}\ ,\ t;\ \xi,\ \tau)$. Each of the three terms in eq. (2.7) can then be described, in detail, as follows:

$$\text{First term of } G_{ij}(\boldsymbol{x},t;\xi,0) = \frac{1}{4\pi\rho\alpha^2}\gamma_i\gamma_j\frac{1}{R}\delta(t-\frac{R}{\alpha}). \tag{2.8}$$

This is a displacement that can naturally be termed the "far-field" P-wave. It is a vector displacement (components given by $i = 1,\ 2,\ 3$) that is a *longitudinal* particle motion (because it is a vector, parallel, at the receiver, to γ). It propagates *without change of shape* at speed α (because of the dependence on time purely via a dependence on $t - R/\alpha$). Its amplitude falls off with distance like $1/R$, and, as a spherically spreading wave, this can naturally be termed an amplitude dependence that is governed by *geometrical spreading*. There is also a dependence on directionality (i.e., of direction from source to receiver) specified by the factor γ_j, and this dependence gives us the *radiation pattern*.

Note that each of the above four properties, associated with phrases given in italics, arises from discussion of a separate factor in (2.8). Labeling this a "far-field" wave, becomes apparent in a comparison with the third term. But let us next identify the key properties of the second term of eq. (2.7):

$$\text{Second term of } G_{ij}(\boldsymbol{x},t;\xi,0) = -\frac{1}{4\pi\rho\beta^2}(\gamma_i\gamma_j-\delta_{ij})\frac{1}{R}\delta(t-\frac{R}{\beta}), \tag{2.9}$$

which is naturally labeled the "far-field S-wave". From the separate factors in this expression we can see that, corresponding to the 4 properties identified above for far-field P-wave, we now have a vector wave displacement which is *transverse* (because $(\gamma_i\gamma_j - \delta_{ij})\gamma_i = 0$); which propagates *without change of shape* at speed β, having amplitudes governed again by *geometrical spreading*; and with a *radiation pattern* governed by the dependence of $(\gamma_i\gamma_j - \delta_{ij})$ on γ_j.

By comparing the amplitude of corresponding terms in (2.8) and (2.9), it follows in general that the S-wave displacement will be larger than P-wave displacement, by a ratio of about α^2 to β^2. (A closer look at this ratio requires taking an average over the radiation pattern.)

$$\begin{aligned} \text{Third term of } G_{ij}(\boldsymbol{x},t;\xi,0) =& \frac{1}{4\pi\rho}(3\gamma_i\gamma_j-\delta_{ij})\frac{1}{R^3} \\ &\times\int_{R/\alpha}^{R/\beta}\tau'\delta(t-\tau')\,\mathrm{d}\tau'. \end{aligned} \tag{2.10}$$

This term is part of the exact expression for G_{ij}, but if in some sense R is large enough, it is a *smaller* displacement than those of (2.8) and (2.9)

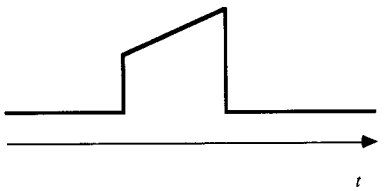


Fig. 12. The time-dependence of the near-field term in the Green function.

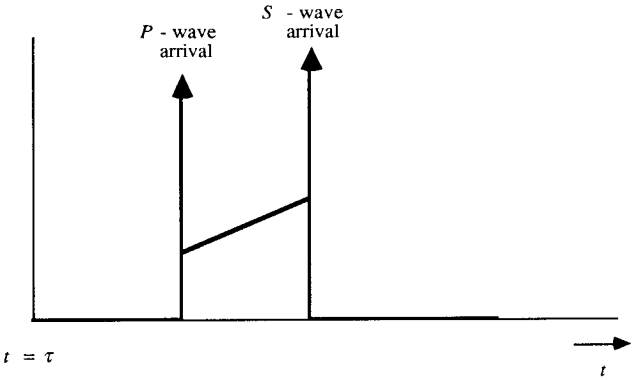


Fig. 13. The time-dependence of the Green function, shown here, has a δ-function pulse-shape at the P-wave and S-wave arrival times. From eq. (2.10), the ramp (i.e., the near-field term) has amplitude proportional to $1/R^2$ and duration proportional to R. Time-dependence is upon $t-\tau$, i.e., only upon time after the impulse is applied at the source.

because of the factor $1/R^3$ here, compared to factors of $1/R$ in each of the previous terms. It is this fact, that naturally leads to (2.8) and (2.9) being called far-field waves, and (2.10) a near-field wave. The near-field wave is neither longitudinal nor transverse (this follows from the factor $3\gamma_i\gamma_j - \delta_{ij}$). Neither is it purely P or S motion. To appreciate the time of arrival of this near-field wave, note that the dependence of

$$\int_{R/\alpha}^{R/\beta} \tau' \delta(t-\tau')\,\mathrm{d}\tau',$$

on time t is easily seen to be zero for $t \le R/\alpha$ and $t \ge R/\beta$; and has value t for $R/\alpha \le t \le R/\beta$. Thus, it is a pulse shape that is present only between P-wave and S-wave arrival times, and is then a ramp that extrapolates back to the origin at $t=0$ (fig. 12):

The time-dependence of $G_{ij}(\boldsymbol{x}, t : \xi, \tau)$, in total, is shown in fig. 13.

Note that the *area* under each of the three terms is proportional in each case to $1/R$. For a general force $\boldsymbol{F}(t)$ applied at $\boldsymbol{x} = \xi$, the resulting

displacement can be obtained by considering each term in the integrand of $\boldsymbol{F}(t) = \int \boldsymbol{F}(\tau)\delta(t-\tau)\,\mathrm{d}\tau$. It is:

$$\begin{aligned} u_i(\boldsymbol{x}, t) &= \int_{-\infty}^{\infty} F_j(\tau) G_{ij}(\boldsymbol{x}, t-\tau; \xi, 0)\,\mathrm{d}\tau \\ &= F_j * G_{ij} \end{aligned} \tag{2.11}$$

(using $*$ for temporal convolution; and note also the implied sum over subscript j). Therefore, the question of the relative weight of "far field" and "near field" terms will depend on the duration of $\boldsymbol{F}(t)$, compared to the time between P and S arrivals.

If, as is usually the case in seismic data, the source duration is significantly less than the time between P and S arrivals, then far field terms alone are dominant. But, if the duration of $\boldsymbol{F}(t)$ is comparable to the $S-P$ time (as will be the case for sources sufficiently near the receiver), then all three terms in (2.7) have comparable weight.

This concludes the present review of Stokes' solution. It will be apparent that, though straightforward, the basic solution for waves in elastodynamics is significiantly more complicated than that for the corresponding scalar problem, eq. (2.5), solved by (2.4), for which the geometric equation solution happens to be exact, and for which there is no radiation pattern.

The moment tensor . Having presented the properties of G_{ij}, we must now recognize that the point source force, $\mathrm{F}(t)$, giving displacement $\boldsymbol{u}$ via (2.11), is rarely a good description of the sources that naturally arise in seismology. Note, for example, that a point source, given as $\boldsymbol{F}(t)$, will in general add both linear momentum and angular momentum to the medium as a whole, and neither of these effects arises for earthquakes or for explosions contained within the medium. Instead, for example, we could accept seismic sources quantified in terms of *vector dipoles* rather than by vectors, and we would then find that these vector dipoles are needed in certain combinations as shown in fig. 14 for explosions and tension cracks and earthquakes.

However, there is no necessity in general to restrict ourselves to vector dipoles such that the 2 forces are aligned with the direction specified by the 2 points of application, as in each of the vector dipoles of fig. 14.

In general, to describe a force system that will not add linear momentum to the medium, we shall want to work with pairs of equal but opposite forces applied at two points that together specify a direction (called the "arm" direction) that can be different from that of the forces. In general, there are nine possibilities, as shown in fig. 15, from which to build up a description of point sources in seismology, and the total effect of these nine possibilities

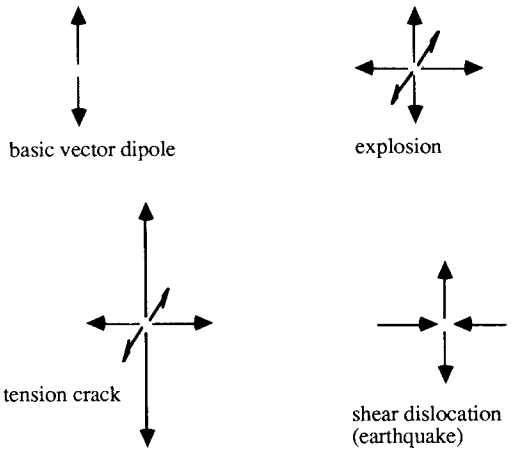


Fig. 14. Force equivalents for different types of seismic source. The basic vector dipole may be used to build up the force equivalent of an explosion (three mutually perpendicular vector dipoles of equal amplitude), of a tension crack (one dipole has larger amplitude than the other two), and of a shear dislocation (only two vector dipoles, one with outward forces and one inward).

must obey the constraint that no net angular momentum is added to the medium.

Such force systems will generate vector displacement fields that can be derived quantitatively in exactly the same way, in principle, that one derives the scalar potential field of a dipole from knowledge of the field due to a single charge. Thus, for a point charge q at $\boldsymbol{x} = \xi$, the scalar electrical potential is $q/\mid \boldsymbol{x} - \xi\mid$ at position $\boldsymbol{x}$. The field due to a dipole with strength M and orientation l (a unit vector) is obtained from the field due to two point sources, $+q$ at $\boldsymbol{x} = \xi$ and $-q$ at $\boldsymbol{x} = \xi - \epsilon \mathrm{l}$, and taking the limit as $q \to \infty$ and $\varepsilon \to 0$ such that $q\varepsilon = M$.

For the limiting value we then have

$$\frac{q}{|\boldsymbol{x} - \xi|} - \frac{q}{|\boldsymbol{x} - (\xi - \varepsilon \boldsymbol{l})|} \to M(\boldsymbol{l} \cdot \nabla)\frac{\mathbf{1}}{|\boldsymbol{x} - \xi|},$$

where the spatial differentiation is with repect to the source coordinates, ξ. In this case, we can naturally describe the source as a vector, $\boldsymbol{M} = M\mathrm{l}$, with resulting field given by

$$M_i \frac{\partial G(\boldsymbol{x}; \xi)}{\partial \xi_i}, \tag{2.12}$$

in which, here, G is just $1/|\boldsymbol{x} - \xi|$.

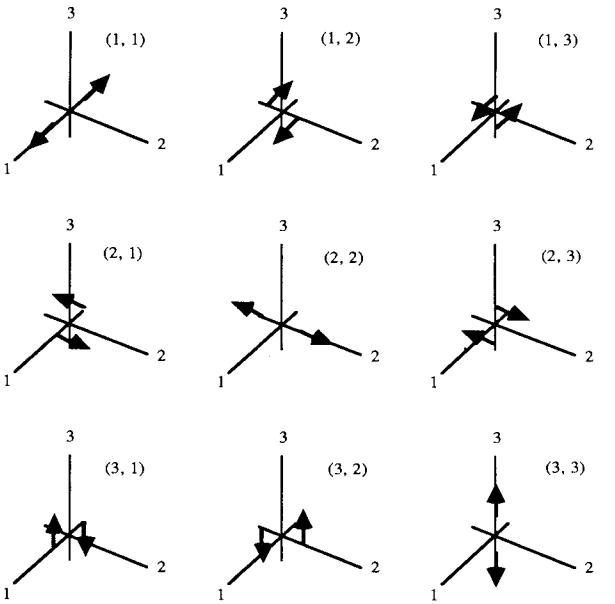


Fig. 15. The nine possible couples/dipoles, when force direction and arm direction may differ.

At last we are in a position to state the way in which vector displacements of importance in seismology are derived from knowledge of $G_{ij}(\boldsymbol{x}\ , t; \xi, 0)$, and, at the same time, give a quantitative description of the source strength. For a pair of forces in the p-direction ($\pm F_p$), applied with arm in the q-direction, we characterize the moment as M_{pq} (in the limit as $F_p \to \infty$ and arm length $\to 0$). M_{pq} is in general a function of time. The displacement is then simple to write down, by analogy with (2.12). It is

$$u_i(\boldsymbol{x}, t) = \int_{-\infty}^{\infty} M_{pq}(\tau)\frac{\partial G_{ip}(\boldsymbol{x}, t - \tau; \xi, 0)}{\partial \xi_q}\,\mathrm{d}\tau = M_{pq} * \frac{\partial G_{ip}}{\partial \xi_q}. \tag{2.13}$$

Thus, knowledge of the Green function, G_{ij}, is still vital. But it gives the displacement solution, for sources of interest in seismology, only after differentiation with respect to source coordinates. Note that (2.12) and (2.13) are valid in general once $\boldsymbol{G}$ is known – that is, these results apply to any elastic medium, not just for a medium that is homogeneous. The strength of the source is characterized by the nine components of a tensor. $\boldsymbol{M}$, called the *moment tensor*. It is simple to show that conservation of angular momentum requires $\boldsymbol{M}$ to be symmetrical. (Aki and Richards, 1980, pp 42 53) so that there are only 6 independent components. For an

explosion (3 equal and mutually orthogonal vector dipoles – see fig. 14), $\boldsymbol{M}$ is isotropic: $M_{ij} = M_I \delta_{ij}$ for some scalar M_I.

For an earthquake, characterized by a displacement discontinuity, $[\boldsymbol{u}\,]$, across a fault plane with normal ν and area A, in rocks with rigidity μ, we have from eq. (1.19), evaluated in an isotropic medium for a point source, that

$$M_{ij} = \mu A\{[u_i]\nu_j + [u_j]\nu_i\}. \tag{2.14}$$

Since $\boldsymbol{s} \cdot \nu = 0$, it follows that this moment tensor has three principal moments given by $\pm M_0$ and zero, where the so-called *seismic moment*, M_0, equals $\mu A|[\boldsymbol{u}\,]|$. Note that the above use of $\boldsymbol{M}$ characterizes a point source – i.e., one for which the radiated wavelengths are much longer than the linear dimensions of the area A. If one is dealing with seismic data for which wavelengths are comparable with fault length, one faces both a complication (in that the point source approximation is inadequate) and perhaps an opportunity (in that this finiteness of the source may possibly be determined from the data). Allowing for a finite area of faulting, we use

$$u_i(\boldsymbol{x}, t) = \int\int \mu\{[u_p(\xi, t)]\nu_q + [u_q(\xi, t)]\nu_p\}^* \frac{\partial G_{ip}(\boldsymbol{x}, t; \xi, 0)}{\partial \xi_q}\, \mathrm{d}A. \tag{2.15}$$

The synthetic seismogram of fig. 4, that improves the fit to the data for station SNZO, was computed via the source finiteness allowed for in (2.15).

Finally, in this review of the way point sources may be quantified in terms of a moment tensor, $\boldsymbol{M}$, note that $\boldsymbol{M}$ is essentially a kinematic, rather than a dynamic quantity. That is, $\boldsymbol{M}$ is determined by the net offset over the fault plane (see 2.14), rather than by the dynamic forces actually acting in the source region.

2.3. Geometrical ray theory

For a homogeneous isotropic whole space, we can perform the operation (2.13) on Stokes' solution for G_{ij}, eq. (2.7), and in this way obtain an exact and explicit solution in such a simple medium. The most important terms, for most applications, are the far-field P- and S-waves, and these arise from spatial differentiation of the delta functions in (2.8) and (2.9), since these are the most rapidly varying spatial factors. Noting, for example, that

$$\frac{\partial}{\partial \xi_k}\delta(t - \frac{R}{\alpha}) = \frac{1}{\alpha}\frac{\partial(-R)}{\partial \xi_k}\dot{\delta}(t - \frac{R}{\alpha}),$$

where an overdot denotes a time derivative, and that

$$\frac{\partial R}{\partial \xi_k} = -\gamma_k,$$

it follows that the far-field P-wave displacement is:

$$\begin{aligned} u_i^{\mathrm{Far-field}P}(\boldsymbol{x}, t) &= \frac{1}{4\pi\rho\alpha^3 R}\gamma_i\gamma_j\gamma_k M_{jk}(t) * \dot{\delta}(t - \frac{R}{\alpha}) \\ &= \frac{1}{4\pi\rho\alpha^3 R}\gamma_i\gamma_j\gamma_k \dot{M}_{jk}(t - \frac{R}{\alpha}). \end{aligned} \tag{2.16}$$

Similarly for the S-wave, we find

$$u_i^{\mathrm{Far-field}S}(\boldsymbol{x}, t) = \frac{1}{4\pi\rho\beta^3 R}(\delta_{ij} - \gamma_i\gamma_j)\gamma_k \dot{M}_{jk}(t - \frac{R}{\beta}). \tag{2.17}$$

The question then arises, of how to generalize these approximate solutions, valid for a homogeneous medium with a point source at ξ, to a medium in which α and β vary continuously with position. We can do this, by generalizing each of the factors in, for example, the far-field P-wave (2.16), using the physical interpretation of these factors we first gave for the wave solution (2.8). The interpretation we shall use is that based on geometrical ray theory, using concepts that first appeared in the physical sciences in the context of geometrical optics.

Thus, we now think of the signal as propagating along rays, with position along the ray specified by travel time evaluated as an integral,

$$T^P(\boldsymbol{x}, \xi) = \int_{\text{Ray from } \xi \text{ to } \boldsymbol{x}} \frac{\gamma \cdot \mathrm{d}\boldsymbol{s}}{\alpha}, \tag{2.18}$$

with $\mathrm{d}\boldsymbol{s}$ an element of ray length, γ a unit vector along the ray, and α a function of position. The ray path is in general curved, the actual path being one that makes travel time T^P in (2.18) an extremal with respect to path perturbations.

The factor γ_i in (2.16) is to be evaluated at the receiver, i.e., $\gamma_i = \gamma_i(\boldsymbol{x}\,)$, correctly describing the fact that a P-wave (in the far-field approximation) is a longitudinal motion.

The factors γ_j and γ_k in (2.16) are to be evaluated at ξ because they are to describe the radiation pattern at the source.

The factor $1/R$ must be replaced by a description of the geometrical spreading, and this is done in fig. 16, in terms of a function $R(\boldsymbol{x}\ ,\ \xi)$.

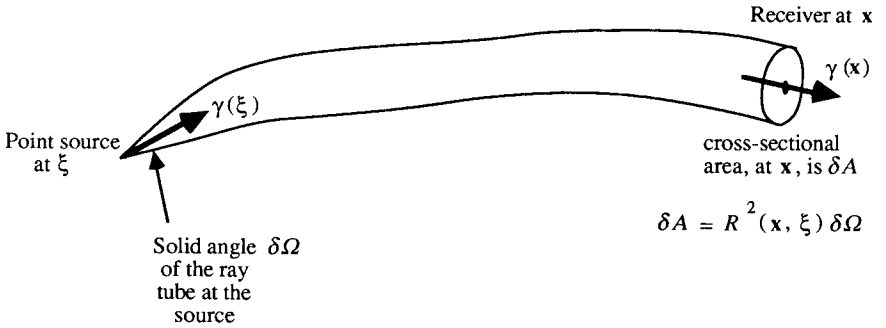


Fig. 16. Sketch to show the definition of the geometrical spreading function, $R(\boldsymbol{x}\ ,\ \xi)$. In a homogeneous medium, R reduces to the distance $|\boldsymbol{x} - \xi|$.

such that the cross-sectional area of a ray tube, at $\boldsymbol{x}$, that departs from ξ centered on direction $\gamma(\xi)$ within solid angle $\delta\Omega$, is $R^2(\boldsymbol{x}\ ,\ \xi)\delta\Omega$.

It follows that (2.16) generalizes in a smoothly varying medium to a displacement proportional to:

$$\frac{1}{4\pi R(\boldsymbol{x},\xi)}\gamma_i(\boldsymbol{x})\gamma_j(\xi)\gamma_k(\xi)\dot{M}_{jk}(t - T^{\mathrm{P}}),$$

and all that remains is a correct generalization of the remaining factor in (2.16), which is:

$$\frac{1}{\rho\alpha^3}.$$

This last step can be taken, once it is recognized that ray theory is based on the approximation that the energy propagating down a ray tube is conserved. The energy flux carried by a P-wave, being a sum of strain and kinetic energy fluxes, is

$$\rho\alpha|\dot{\boldsymbol{u}}|^2,$$

across an area element perpendicular to the P-wave ray (see Aki and Richards, 1980, pp126–127). Therefore,

$$\rho(\boldsymbol{x})\alpha(\boldsymbol{x})|\dot{\boldsymbol{u}}|^2R^2(\boldsymbol{x},\xi),$$

is constant along a ray tube. Taking the square root of this result, we find for the velocity and hence the displacement,

$$|\dot{\boldsymbol{u}}| \propto \frac{1}{\sqrt{\rho(\boldsymbol{x})\alpha(\boldsymbol{x})}}\frac{1}{R(\boldsymbol{x},\xi)},$$

as $\boldsymbol{x}$ varies along the ray tube. The last factor here, the geometrical spreading, we have already recognized. The earlier factors, however, tell us how to apportion

$$\frac{1}{\rho\alpha^3},$$

between values at ξ and $\boldsymbol{x}$. The end result, at last, is the expression

$$u_i^{\mathrm{Far-field}P}(\boldsymbol{x},t) = \frac{1}{4\pi\sqrt{\rho(\boldsymbol{x})\alpha(\boldsymbol{x})\rho(\xi)\alpha(\xi)}\,\alpha(\xi)^2R(\boldsymbol{x},\xi)} \times \gamma_i(\boldsymbol{x})\gamma_j(\xi)\gamma_k(\xi)\dot{M}_{jk}(t - T^{\mathrm{P}}). \tag{2.19}$$

This formula has wide application in seismology, since it is the basis upon which P-wave waveform data can often be interpreted, and the focal mechanism obtained. It gives the direction of particle motion, $\gamma_i(\boldsymbol{x})$; the radiation pattern, the way that $\gamma_j(\xi)\gamma_k(\xi)\dot{M}_{jk}$ varies with $\gamma(\xi)$; the pulse shape, a sum over $\dot{M}_{jk}(t - T^{\mathrm{P}})$; the travel time, T^P; the geometrical spreading, $1/R(\boldsymbol{x},\xi)$; and an absolute overall factor depending on ρ and α at source and receiver. As a "far-field" expression, (2.19) becomes increasingly accurate at higher frequencies (or, equivalently, at times nearer the arrival time of a propagating wavefront).

Thus, it is commonly found that (2.19) is accurate for teleseismic P-waves recorded in the Earth within the distance range $30° \leq \Delta \leq 90°$ and within the band 0.01–5 Hz. Note that in practice, to do adequate modelling of body waves from earthquakes, as for example in figs. 2 and 4, we also need to modify eq. (2.19) to account for the effect of interfaces such as the Moho and the Earth's free surface. These effects are quantified with transmission and reflection coefficients introduced as additional factors into (2.19). For example, these interfaces can convert S-waves (incident from the source region) into P-waves that arrive only a few seconds after the direct P-wave at teleseismic distances. Thus, observed pulse shapes of P must often be regarded as a superposition of P, pP, and sP pulse shapes, each modeled by expressions like (2.19). For *shallow* earthquakes, these pulse shapes are all superimposed.

An example of observed pulse shapes for which modelling with eq. (2.19) is now easy, is shown in fig. 17. In this case, the pulse shapes are particularly simple because these are the far-field P-waves (P, and pP) of a *deep* earthquake, and so the reflections from the free surface above the source arrive sufficiently late for the pulses to be well separated in time.

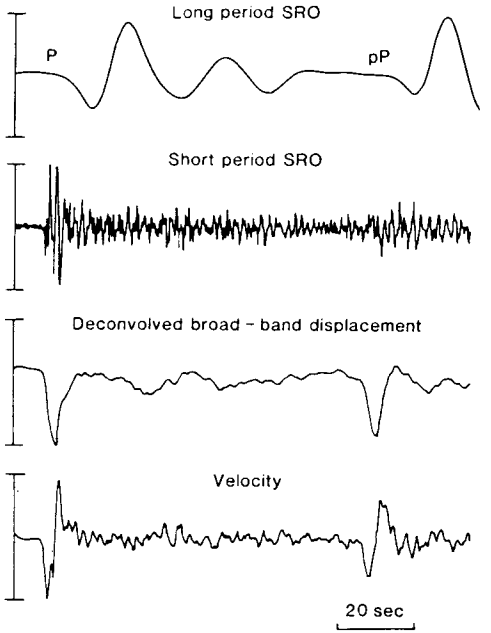


Fig. 17. Remarkably simple pulse shapes are commonly observed from deep earthquakes. Note that most seismometers recording teleseisms do so in narrow frequency bands. To derive a broadband signal, it may then be necessary to combine information from two different frequency bands, as done here for the Seismic Research Observatory (SRO) at MAIO. The top two traces are digitally recorded vertical seismograms from a deep event (377 km) at a distance of 63.4°. The third trace is constructed from these two frequency bands, after deconvolving the instrument responses. The pulse widths and arrival times of P and $\dot{p}P$ become much clearer than in the original records. (From Choy and Boatwright, 1981.) From data processing such as this, and the importance of details in the broadband pulse shapes, there is a strong trend in modern seismology to deploy sensors that can directly give broadband displacement or velocity signals. The bottom trace is the corresponding velocity record.

The short pulse durations are what would be expected from eq. (2.19) for a dependence of moment on time in which the final fault offset is acquired suddenly. In such cases, $\boldsymbol{M}\ (t)$ is roughly proportional to a Heaviside step function, and the time derivatives appearing in eq. (2.19) give impulsive pulse shapes. However, pulse broadening, as is observed, is a result not only of the finite duration of motion at the source, but also of attenuation in the Earth – as we shall see in section 1.3

3. Attenuation, from intrinsic friction or scattering

3.1. Introduction

The seismic waves set up by earthquakes and explosions propagate away from the source, and eventually fade away to become unobservable against the background of noise. By the term "fade away", we are describing the outcome of several different processes. For example, the waves typically become more and more spread out as they propagate further, so their amplitudes decay in a fashion that is more or less described by a geometrical spreading factor. Amplitudes are also in general attenuated by two types of physical process, which are the main subject of this lecture.

Attenuation may occur as the result of a transfer from kinetic and strain energy of motion, to the energy of heating or to the energy required to change the defect structure of the material within which the seismic wave is propagating. This is a true dissipation of the energy of motion, and it may occur in a medium that is homogeneous. Sometimes, the underlying cause of this dissipatation is referred to as "intrinsic friction". Stress in this case is not at all times proportional to strain – because there is typically some hysteresis – so such a medium is not strictly elastic. We say it is *anelastic*. For small motions, however, it may still be linear. Typically, the major effect on a seismic wave is the loss of signal at high frequencies.

Another physical process that will lead macroscopically to decreasing amplitudes is that of scattering of elastic waves by small-scale heterogeneities. The energy of motion is not lost, but becomes less and less coherent the further a wave propagates through the heterogeneities. Superficially, the main effect on a body wave that originated at the source as an impulsive signal is to produce an initial pulse shape that broadens, and that may have a loss of high frequencies. There is also what seismologists call a "coda", oscillations that continue to be present for some time after the first arrival has come and gone. Thus, for the first-arriving pulse shape, effects of scattering can be similar to the effects of anelasticity, though the underlying mechanisms are completely different.

To explain observed attenuation of seismic waves, it is often of interest in geophysics to apportion the cause between these two physical mechanisms, because of the information this could give about the mechanical state and lithology of the structures in which the waves have propagated. Sections below describe the main features of how these types of physical process achieve their attenuative effects.

3.2. Attenuation, due to anelasticity

Consider a very simple wave in an *elastic* medium:

$$e^{i(kx-\omega t)}, \quad \text{(wavenumber } k = \text{frequency/velocity} = \omega/v). \tag{3.1}$$

This is a wave propagating in the positive x-direction. It has unit amplitude at the origin, and this amplitude is preserved as the wave propagates, because in an elastic medium it does not attenuate.

In contrast, a simple wave in an a *anelastic* medium can be described by

$$e^{i(Kx-\omega t)}, \tag{3.2}$$

where $K = k + i\alpha$, that is, the wavenumber becomes complex.

Separating this into phase and amplitude factors (corresponding to real and imaginary parts of the complex wavenumber K), we have

$$e^{i(kx-\omega t)}e^{-\alpha x}. \tag{3.3}$$

The spatial decay rate, α, is often described in seismology (and other physical sciences dealing with wave propagation) by a dimensionless number, Q, which can be introduced via

$$\alpha = \frac{\omega}{2vQ}. \tag{3.4}$$

Seismic waves are observed to propagate with an attenuation that corresponds to only a weak dependence of Q on frequency ω. In fact, many observational seismologists routinely assume Q is a constant, independent of frequency. We shall return to this point below, but note here from (3.4) and (3.3) that, in a constant Q medium, high frequencies are preferentially filtered out of the spectrum.

There are several other ways to introduce Q. For example, the energy ΔE lost to intrinsic friction after propagating one wavelength is related to initial energy E, and to Q, via

$$\frac{1}{Q} = \frac{\Delta E}{2\pi E}. \tag{3.5}$$

Associated with this loss is a hysteresis or phase delay between applied stress and resultant strain, so that whereas these quantities are related to each other via a real constant M in an elastic medium, stress $= M \times$ strain,

Fig. 18. Two pulses, one more attenuated than the other.

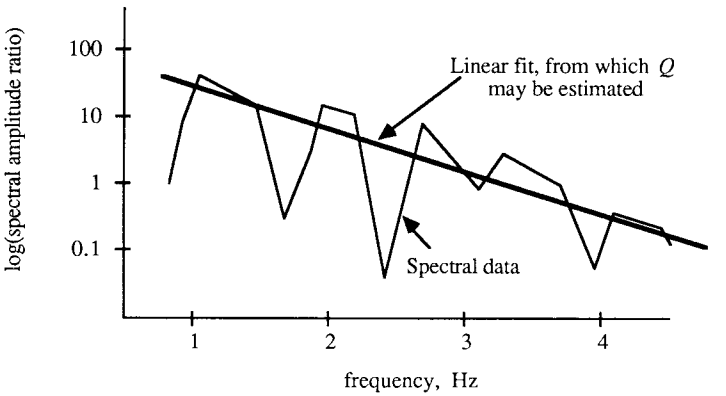


Fig. 19. Schematic straight-line fit of log spectral amplitudes, for constant Q.

M becomes complex (and, in general, frequency dependent) in an anelastic medium. From the one-dimensional equation of motion, $\rho\ddot{u} = \partial\tau/\partial x$, and the corresponding stress-strain relation, $\tau = M\epsilon$, we obtain for the wave (3.2) the relation

$$\rho\omega^2 = K^2 M(\omega). \tag{3.6}$$

O'Connell and Budiansky (1978) advocate the definition

$$\frac{1}{Q} \equiv \frac{-\mathrm{Im}\{M\}}{\mathrm{Re}\{M\}}. \tag{3.7}$$

The relation (3.4) then follows approximately (being accurate if $k \gg \alpha$, or, equivalently, if $Q \gg 1$).

One of the most basic ways to measure Q is to record a pulse at different locations a distance X apart (see fig. 18).

Forming the spectral amplitude ratio, and interpreting it as

$$e^{-\alpha X} = e^{-\omega X/2vQ}, \tag{3.8}$$

a log-linear plot will result in a straight line with negative slope if indeed Q is constant (see fig. 19).

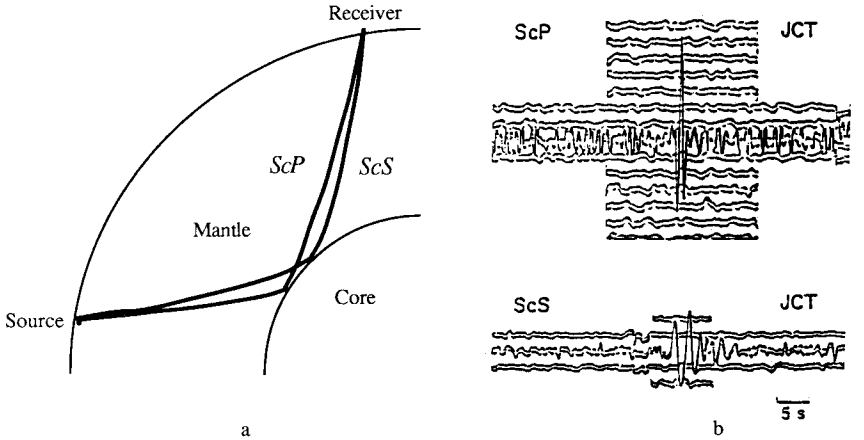


Fig. 20. (a) Ray paths of ScS and ScP. (b) Short-period records for Junction, Texas, for the deep South American earthquake of 1967 March 27. ScS contains signal up to about 1.5 Hz. (From Burdick, 1985.)

Many experiments have been designed along these lines in seismology, to take spectral ratios for a determination of Q. For example, fig. 20a shows the paths taken by rays ScS and ScP. These rays differ very little in their directions of departure from a deep earthquake to a fixed station, as both rays propagate steeply through the upper mantle. An example of these waves arriving at Junction, Texas, is shown in fig. 20b, and the ScS wave is observed to have less high-frequency signal content than does ScP, due partly to the larger number of wavelengths in ScS, and partly to the fact that Q for S is smaller than Q for P. Note that we can write (3.3) and (3.4) as

$$e^{i\omega(T-t)}e^{-\omega T/2Q}, \tag{3.9}$$

where T is travel time. The spectral ratio between ScS and ScP can then be interpreted as

$$\exp\left(-\frac{\omega}{2}\left[\frac{T^S}{Q_S}-\frac{T^P}{Q_P}\right]\right),$$

where T^P and Q_P refer to P-waves, averaged for one passage through the whole mantle, and T^S and Q_S refer to S-waves. The observability of P and S, at around 1Hz, after propagating several minutes in the Earth, then translates to values of at least a few hundred for Q_P and Q_S.

The need for dispersion: velocity $v = v(\omega)$. Let us assume a pulse shape $p = p(x,t)$ has the Fourier transform $(t \to \omega)$ given by e^{iKx} with

$$K = \frac{\omega}{v} + i\alpha = \frac{\omega}{v}\left(1 + \frac{i}{2Q}\right). \tag{3.10}$$

Then $p = p(x,t)$ is a delta function, $\delta(t)$, at $x = 0$.

It is a simple exercise to carry out the inverse transform for $p = p(x,t)$ if indeed Q is constant. The outcome is

$$p(x,t) = \frac{1}{\pi}\left[\frac{\frac{x}{2vQ}}{\left(\frac{x}{2vQ}\right)^2 + \left(\frac{x}{v} - t\right)^2}\right]. \tag{3.11}$$

But this is a totally unacceptable result, since (3.11) indicates that after propagation the pulse has non-zero values not only prior to the time $t = x/v$, but even prior to the time $t = 0$.

The property we wish to reproduce is that seismic signals, which we are analyzing within a framework of linear superposition of different frequency components, are *causal*. If the amplitude spectrum of a causal signal is specified then there are constraints on the behavior of the phase spectrum that require us to allow for a frequency-dependent phase-velocity, $v = v(\omega)$.

To be specific, let us suppose that the pulse $p(x,t)$ is actually zero, at x, for all times $t < x/v_\infty$. Then define

$$F(\omega) \equiv e^{i\left(K - \frac{\omega x}{v_\infty}\right)} = p(x,\omega)\, e^{\frac{-i\omega x}{v_\infty}} = \int_{\frac{x}{v_\infty}}^{\infty} p(x,t)\, e^{i\omega\left(t - \frac{x}{v_\infty}\right)}\, dt. \tag{3.12}$$

It follows that $F(\omega)$ is an analytic function of ω everywhere in the upper half of the ω-plane (because this integral and all its derivatives converges everywhere that ω has a non-negative imaginary part). For such a function, the real and imaginary parts are Hilbert transform pairs. In fact, it can be shown (Aki and Richards, 1980, pp. 173–174) that $\log F(\omega)$ too is analytic everywhere in the upper half of the ω-plane. The Hilbert transform relation is then available as

$$\frac{\omega}{v(\omega)} = k(\omega) = \frac{\omega}{v_\infty} + H\{\alpha(\omega)\}, \quad \text{and}$$
$$\alpha(\omega) = -H\{\frac{\omega}{v(\omega)} - \frac{\omega}{v_\infty}\}. \tag{3.13}$$

The Hilbert transform of a function $f(x)$ can be written as

$$H\{f(x)\} = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{f(y)\,dy}{y-x}.$$

It can be regarded as the outcome of a 90° phase shift of all the Fourier components of $f(x)$.

Many choices of $\alpha(\omega)$ and $v(\omega)$ have been proposed that satisfy the constraints (3.13) arising from causality. For example, if we take

$$\alpha(\omega) = \frac{\alpha_0\omega}{1+\alpha_1\omega}, \tag{3.14}$$

in which α_0 and α_1 are constants, then this can model a constant Q medium at frequencies for which $|\alpha_1\omega| \ll 1$. Azimi et al. (1968) pointed out that the Hilbert transform of (3.14) is

$$H\{\alpha(\omega)\} = \frac{2\alpha_0\omega}{\pi(1-\alpha^2{}_1\omega^2)}\ln\left(\frac{1}{\alpha_1\omega}\right), \tag{3.15}$$

and therefore the phase velocity is effectively

$$\frac{1}{v(\omega)} = \frac{1}{v_\infty} + \frac{2\alpha_0}{\pi}\ln\left(\frac{1}{\alpha_1\omega}\right).$$

This allows a comparison of the phase velocity at two different frequencies, and we find

$$\frac{v(\omega)}{v(\omega_0)} = 1 + \frac{1}{\pi Q}\ln\left(\frac{\omega}{\omega_0}\right). \tag{3.16}$$

Such a result is important in geophysics, because it shows that phase velocities (which potentially can be measured very accurately, by timing a wave over sufficiently great distances) are affected by attenuation to *first order* in $1/Q$, if Q is constant over the bandwidth of observations.

Figure 21 shows a comparison of two "constant Q" pulse shapes – one computed via (3.11), ignoring causality, the other by allowing for the logarithmic delay of lower frequencies, using the dispersion rule given in (3.16).

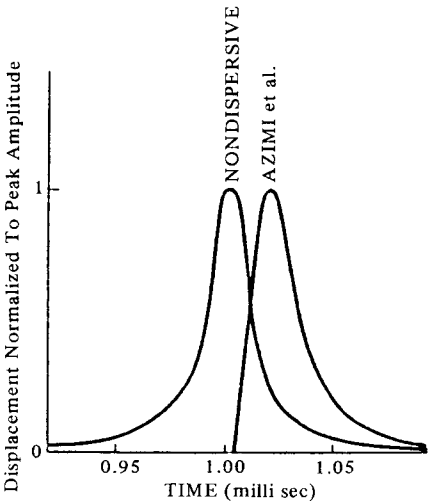


Fig. 21. Comparison of pulse shapes, dispersive and non dispersive, after propagation a distance 5 meters (v = 5 km/s, Q = 50. After Gladwin and Stacey, 1974).

Note, for the Azimi et al. pulse shape, the asymmetric shape with its steep onset and relatively slow decay.

An attenuation–dispersion pair, when Q is slightly dependent on frequency. If Q is constant, then in practice it is clear that causality requires a slight velocity dispersion, and we have obtained in eq. (3.16) the frequency dependence of this dispersion. For data of narrow bandwidth, typical of most seismometers and recording systems prior to the mid-1970's, Q does appear to be constant. However, the increasing deployment of broad-band instrumentation, and increasing efforts to model waveforms, has led to the general recognition that Q increases slightly with frequency. This conclusion was implied in earlier work, previous to the deployment of broad band systems, in that seismologists who studied the Earth at periods longer than a minute typically reported Q values as 200 or less for S waves propagating through the whole mantle, whereas at periods of a few seconds, Q values above 500 were typically reported (see also fig. 22, below, for P-waves).

A useful attenuation law with which to investigate this frequency dependence is that which results from taking α, the spatial decay rate as in eq. (3.3), proportional to some power of frequency. Thus, if α is proportional to ω^s, it follows that

$$Q(\omega) = Q_0 \left(\frac{\omega}{\omega_0} \right)^{1-s}, \tag{3.17}$$

giving a slight rise of Q with ω if s is somewhat less than the value 1. The Hilbert transform of α can be found (Strick, 1970), and from (3.13) the dispersion is described by

$$\frac{1}{v(\omega)} = \frac{1}{v_\infty} + \frac{\alpha(\omega)}{|\omega|} \tan \frac{s\pi}{2}. \tag{3.18}$$

Note that the singularity of $\tan(s\pi/2)$ at $s = 1$ indicates there can be a significant phase delay of the lower frequencies, if Q is almost constant. However, in practice one is not interested in the phase delay with respect to infinite frequency. The more relevant form of the dispersion relation is one which brings out the dispersion with respect to a reference frequency that may itself be part of the observed bandwidth. Thus, we can use

$$\frac{v(\omega)}{v(\omega_0)} = 1 + \frac{1}{2} \left(\frac{1}{Q_0} - \frac{1}{Q(\omega)} \right) \tan \frac{s\pi}{2}. \tag{3.19}$$

3.3. Comparison of pulse shapes with different dispersion

It is of interest to compare the pulse shapes that result when teleseismic P waves are modelled with different assumptions about Q. We shall find that the outcome of a Q with power-law frequency dependence is different from what might be expected from experience with constant-Q attenuation.

Thus, in fig. 22 is shown the depth-dependence of Q for P-waves in two Earth models. SL8 is typical of Q estimates derived from long period surface waves and normal modes, and AFL is more typical of Q estimates based on teleseismically observed high-frequency body waves.

In fig. 23 is shown the P-wave pulse shape from an impulsive source in Earth model SL8, for which Q is constant and quite low in value. Note the characteristic features of this pulse: high frequencies have been selectively removed, and the pulse shape is broadened with an asymmetric rise and fall. The attenuation–dispersion pair is based on eq. (3.16).

In fig. 24 is shown the same pulse, but for Earth model AFL. Q is again constant, but is significantly higher than the value for SL8.

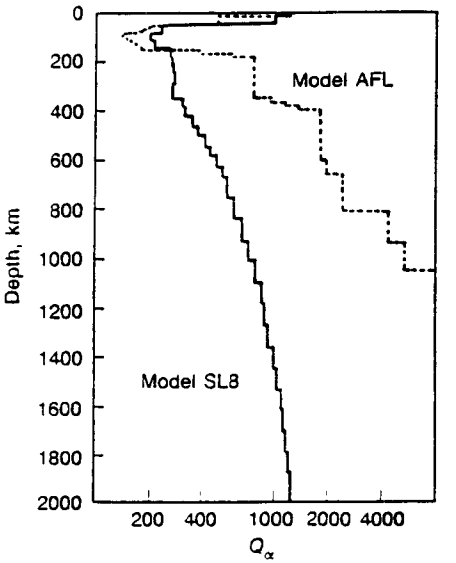


Fig. 22. Two models (SL8 and AFL) of Q as a function of depth. (From Choy and Boatwright, 1981.)

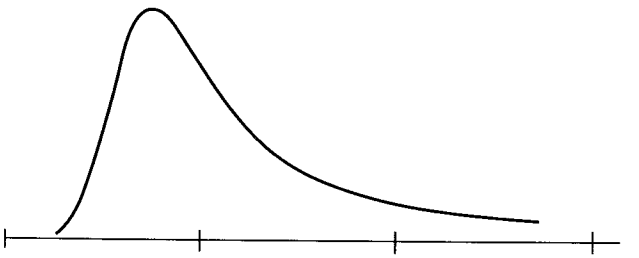


Fig. 23. The P-wave pulse shape at 48° in Earth model SL8, for an impulsive shallow source. Constant Q, and logarithmic dispersion as in eq. (3.16). Tick marks are at 1s intervals.

Since both AFL and SL8 are data-driven anelastic Earth models, the first based on relatively high frequency signals (about 2 Hz) and the second based on relatively low frequencies (about 0.005 Hz), it is clearly unsatisfactory that each model is presented in terms of a frequency-independent Q structure. That is, Q is constant and has low value in SL8; Q is constant but with a high value in AFL. Various high-frequency phenomena displayed in seismic data are not present in SL8, and low frequency phenomena in seismic data are not present in AFL.

It is therefore appropriate to consider some kind of interpolation for Q across the frequency range whose extremes are represented by the frequency

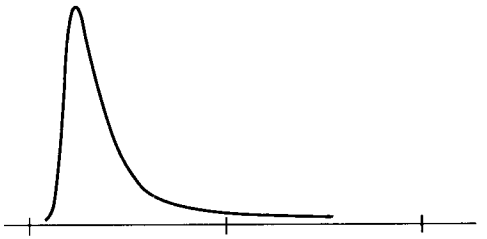


Fig. 24. The P-wave pulse shape at 48° in Earth model AFL, from a shallow impulsive source. Tick marks, 1s. Again, dispersion is governed by eq. (3.16), but with a high Q value.

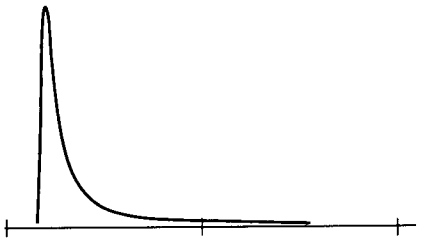


Fig. 25. The P-wave at 48° in an Earth model with power law Q, having approximately the Q of SL8 at frequencies for which the attenuation in SL8 was determined, and the Q of AFL at frequencies for which the attenuation of AFL was determined. Impulsive shallow source. Tick marks, at 1s intervals.

values typical of the data bases underlying these two Earth models. We can do this by assuming a Q that has a power-law dependence on frequency, as in eq. (3.17). Causality rules and the evaluation of Hilbert transforms then lead to eq. (3.19) as the dispersion rule.

Consideration of the difference in frequency ranges over which SL8 and AFL were generated, and the values of the two separate but constant Q values in each model, leads one to a value of about 0.7 for the constant s that appears in eq. (3.17). Thus, Q has about a 0.3 power-law dependence on frequency, to fit the two kinds of data base underlying these models.

The outcome of computing a pulse shape based on this $Q(\omega)$ might reasonably be expected to be a pulse somewhat intermediate between those of figs. 23 and 24, since we have merely found a way to interpolate for Q across a broad band of frequencies that is representative of SL8 at one end, and AFL at the other. But the actual outcome is different. It is shown in fig. 25, and one sees a pulse shape that has an even shorter rise time than does that of fig. 24 for AFL.

The reason for the shorter rise time is that at frequencies even higher than those which charaterize the data base for AFL, eq. (3.17) gives a higher value of Q than the constant Q value of AFL itself. Since it is the highest frequencies that determine the rise time, this latter feature will, in the frequency-dependent Q model, be reflective of the assumptions built into (3.17) concerning the very-high frequency dependence of Q. In this case, Q is unbounded as frequency rises, which presumably is why the rise time is so short in fig. 25.

We therefore reach two conclusions. First, that a constant Q attenuation model is unsatisfactory . (This has long been known: such a model does not explain observations across the observed range of frequencies.) Second, that its replacement is not obvious. Artificial features, associated with attenuation laws applied outside the range of frequencies for which there is a good data base, can in the time domain be drawn into the appearance of a pulse shape computed within the frequency band at which data is acquired. Note that because of the Hilbert transform rules, the phase delay at a finite frequency within the band of observations is fixed once one has determined the attenuation spectrum at all frequencies. The Hilbert transform of the log of the amplitude spectrum converges poorly if Q is nearly constant, resulting in effects (on phase delay, and therefore pulse shape) within the seismic bandwidth that may be artifacts of assumptions about the attenuation at absurdly high frequencies – for example the frequency-dependence of Q given in (3.17), applied onto gigahertz and higher.

A good research plan to resolve this issue is to see if high-quality broadband pulse shape data can be acquired, for example from body-wave data of deep earthquakes observed teleseismically. An example of such simple broadband signals was shown previously in fig. 17. It may be possible to obtain the attenuation–dispersion pair directly from this type of data. A first attempt in this direction has been made by Choy and Cormier (1986). It is difficult to achieve results, because of the need to remove source effects, and effects of heterogeneity of structure at depth (e.g., effects of the downgoing slab, which may be expected to be present as a cause of the earthquake activity itself). However, observational programs are the natural basis for developing an attenuation–dispersion pair that is appropriate in computation of synthetics. Such programs will eventually narrow the range of current trade offs between source and propagation phenomena.

3.4. Attenuation, due to scattering

The P-wave velocity in the Earth's crust has fluctuations on all spatial scales. For example, fig. 26 shows a profile (i.e. a plot of velocity as a function of depth) obtained by measuring the acoustic velocity in the walls of a borehole at different depths. Also shown, is an idealization that approx-

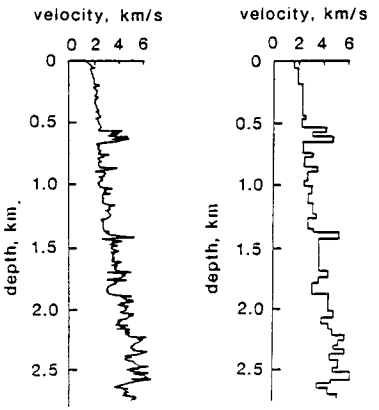


Fig. 26. A typical acoustic log from a borehole, and an idealization. From Richards and Menke (1983).

imates this measured velocity profile with over 40 different layers. Though this is far more complicated an idealization than is usually used in crustal models for which synthetic seismograms are computed, the approximation does not begin to approach the detailed complexity of the real crust, even if we assume plane layering.

We have seen in section 1.2 how heterogeneity on large spatial scales affects seismic body waves. Within a slowly varying medium, we found that ray paths are curved and signal amplitudes are focussed or defocussed as ray tubes shrink or expand. This section is instead concerned with the effects, upon seismic waves, of a rapidly varying medium. Note that by "slowly varying" or "rapidly varying", we refer to a comparison of the spatial scale of the heterogeneity with the spatial scale set by the wavelength. In this section, we are interest in the effects of fine-scale structure.

If a seismic wave propagates within a medium such as shown in fig. 26, then scattering occurs from each heterogeneity. If the wave propagates sufficiently far, than we may expect that eventually the wave will become scattered to such a degree that its identity will be lost. The question of interest, is how does this multiple scattering makes its effect apparent upon the pulse shape of a through-going wave?

An answer is given in fig. 27, which shows a medium with 500 layers whose velocities are selected at random from a gaussian distribution (mean velocity, 3.0 km/s, standard deviation 0.25 km/s; density is assumed constant). Also shown is what happens to a pulse shape, originally an impulse, after transmission vertically across a few hundred layers in such a medium with a somewhat greater standard deviation (0.5 km/s). The method of computation for waves in layered media is briefly reviewed in a later section

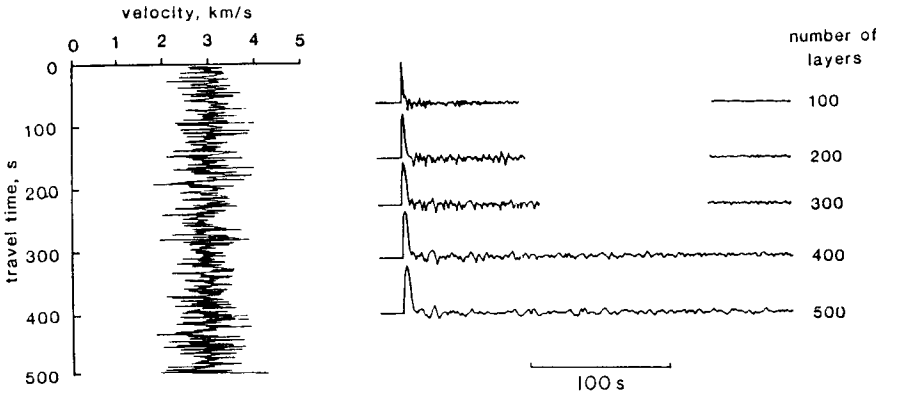


Fig. 27. The profile of 500 layers drawn from a gaussian distribution, and the transmission response of a similar medium, after passage across a few hundred such layers. Note the significant coda and pulse broadening. From Richards and Menke (1983).

(1.5), but the outcome is shown here in this figure because the effect on the propagating pulse shape is seen broadly to be similar to what happens in an anelastic medium. The "spike" of the initial impulse becomes an ever wider pulse, and, in contrast to what is seen in an anelastic medium, there is the development of a "coda" of scattered waves that arrive after the main signal.

Because the medium under discussion in fig. 26 is elastic, the high frequency signal components are not lost. Rather, they are more effectively scattered, and arrive later in the coda, as shown in fig. 28.

The exercise of computing synthetics such as these, can lead to a change in one's whole mind set about the Earth's internal structure and the nature of body-wave pulse propagation. Thus, subsequent to the very first non-zero arriving amplitude in the pulses of fig. 27, the signals are composed of forward-scattered multiples. Whereas ray theory for smoothly-varying media can be presented as a set of approximations that become simpler and more and more accurate at ever-higher frequencies, as discussed in connection with eq. (2.19), we find this approach is of little use in assessing scattering and coda, which are such prominent features of short-period seismograms. To demonstrate that indeed high-frequency components can arrive later in seismograms, fig. 29 shows a seismogram obtained in Norway for a teleseismic Soviet underground nuclear explosion, after the recorded signal has been passed through a sequence of narrow-band filters.

For many purposes in geophysics, it is not necessary to reach any conclusion about the mechanisms by which seismic waves are attenuated. Thus,

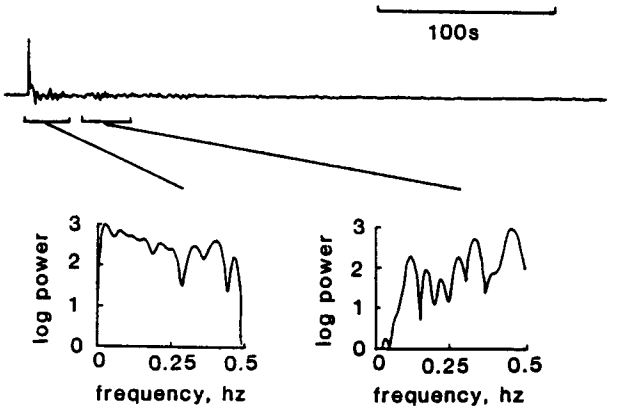


Fig. 28. Transmission response of upper 100 layers of log used in fig. 27; and the power spectra of the initial portion of the pulse, and of a later window within the coda. From Richards and Menke (1983).

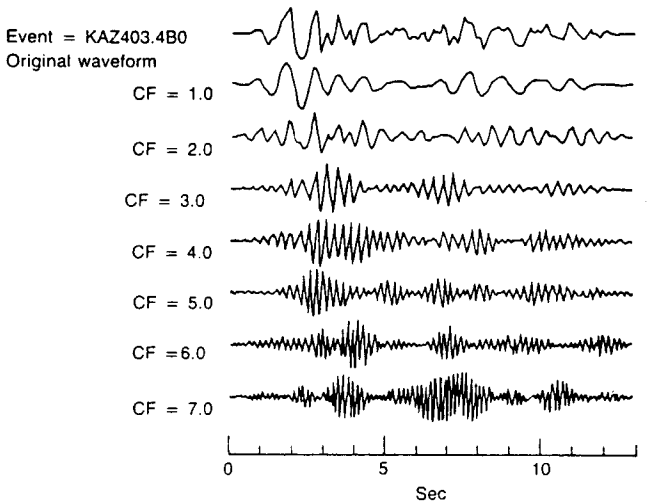


Fig. 29. Band-passed filtered seismograms, generated from the same original recording of vertical ground motion in Norway, using zero-phase shift gaussian filters with center frequencies of 1, 2, 3, 4, 5, 6, and 7 Hz. Each filter has a half-width maximum of 0.5 Hz. Note the later arrival of higher frequency signals, which may be a common feature of multiply-scattered seismic signals. From McLaughlin and Anderson (1985).

to read and interpret travel times, or even to study the systematic variability of source sizes, it is usually enough to work within the framework of wave propagation for smoothly varying media, perhaps with an allowance

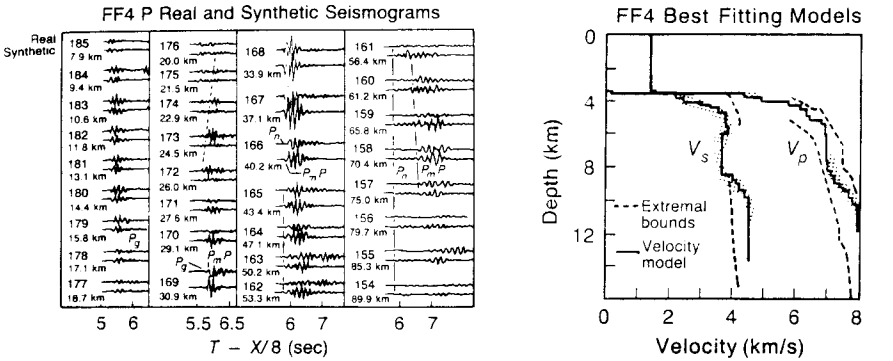


Fig. 30. Comparison between observed hydrophone records and synthetics, computed for plane layered media using the oceanic crustal model as shown. From Spudich and Orcutt (1980).

for phenomenological effects of attenuation upon amplitude spectra. However, to interpret not only the pulse shapes of seismic body waves, but also the statistics of observed coda, the question of what physical mechanisms are responsible for attenuation is an essential element, and an active research area, of modern seismology.

4. Seismic waves in media with plane parallel layering

4.1. Introduction

Sections 1.4 and 1.5 describe a way to represent, exactly, the seismic waves that propagate away from a point source in a restricted type of inhomogeneous medium – namely, one that is built up from homogeneous layers with plane parallel boundaries. We also discuss the case of a continuous variation of properties with depth. The representation entails thinking of the wavefield, set up by the point source, in terms of its frequency and wavenumber components. Such components of the wavefield of interest, evaluated at fixed frequency and fixed wavenumber in a cartesian coordinate system, are nothing but plane waves. So, these two sections are essentially about plane waves, and their use to represent point sources generating waves in layered media.

The initial rationale for this approach to seismic wave interpretation is that, for many practical purposes, the seismic waves observed to propagate in the Earth's crust and uppermost mantle behave in ways that can be effectively modelled in terms of the synthetics within a plane layered structure. An example is given in fig. 30.

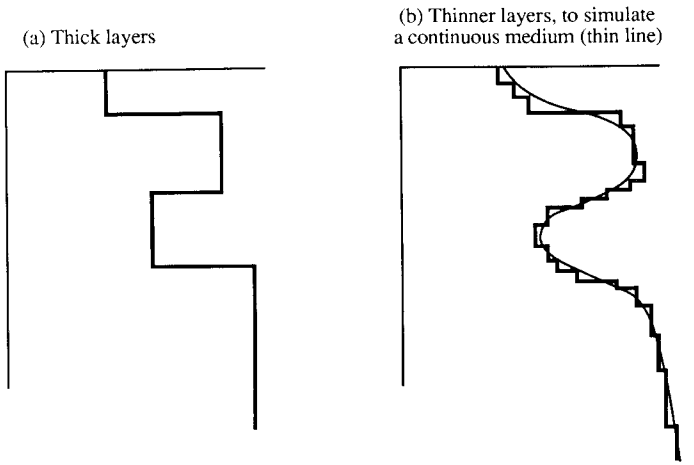


Fig. 31. The use of layered media, to build up depth dependent structures.

Once the capability has been developed, to compute seismic wave solutions in such plane layered media, one finds broadly that there are three direct ways in which it may be used. First, it can be used with just a few layers, so that individual layers and the interfaces between them have a recognizable identity in terms of units within the Earth. For example, we might have one homogeneous layer to represent the crust, then a discontinuity, the "Moho", below which are a couple of layers and a half space, representing lithosphere, low velocity zone, and the rest of the mantle (see fig. 31a). Second, plane layered modelling can be used with large numbers (perhaps hundreds) of layers, in order to reproduce the effect (on seismic waves) of Earth structure that varies smoothly with depth (see fig. 31b). In this way, one can quantify the reflection from regions of high velocity gradient, as well as from specific discontinuities. The third way in which the solution for waves in plane homogeneous layering is useful, is to quantify the cumulative effects of fine structure, on a spatial scale that may be much smaller than a wavelength. In this case it may be necessary to work with many thousands of layers, modelling random fluctuations of velocity with depth, and the choice of velocity in each layer is made on a statistical basis. For example, the choice might be drawn at random from a gaussian distribution, perhaps superimposed upon a mean velocity that varies smoothly with depth. We have described this use in the previous section (1.3), to quantify the apparent attenuation which may occur due to scattering from fine layering in an elastic medium.

We must recognize, in all of the above uses of plane-layered modelling, that the approach is fundamentally limited because of its inability to incorporate, directly, any wave propagation effects associated with lateral variation of structure in the Earth. The Earth has 3D spatial variability, and plane-layering can model the variability in one spatial direction alone. However, in practice, one finds that synthetics in 1D structures provide an important point of reference for interpreting high quality data, because variability of structure is usually strongest in the depth direction. Also, the algorithms that produce such synthetics, once thoroughly understood, can in some cases be adapted to handle lateral variability of structure. An example of this is given in section (1.6).

In what follows, only the broad features and principles of working with plane homogeneous layering are reviewed. It is not intended that these notes supply complete derivations of the main results. The literature on wave propagation in plane layered media is extensive, and continues to grow. Rather, the intent here is to identify key ideas. Thus, we begin with the plane wave representation for the outgoing spherical wave from a point source in a homogeneous medium. We outline some of the many ways in which this integral representation can be evaluated – analytically and/or numerically. In particular, we show the importance of the stationary phase approximation, and its use to obtain the results that might be predicted by geometrical ray theory. If the medium has one or more interfaces, we find representations for the scattered field, and (in section 1.5) sketch a matrix method for describing the physical components of displacement and stress. In sections (1.4) and (1.5), we briefly describe a generalized ray theory, showing in particular how body waves may be identified and quantified in the representation, and commenting on head waves and surface waves, and ways to incorporate the effects of anelasticity.

4.2. The plane wave/cylindrical wave components of a simple spherical wave

For the 3D wave equation

$$\frac{1}{v^2}\frac{\partial^2 W}{\partial t^2} = \nabla^2 W + \frac{\delta(\boldsymbol{x})\delta(t)}{v^2}, \tag{4.1}$$

valid in an infinite homogeneous medium, the exact solution, outgoing at infinity, is

$$W(\boldsymbol{x}, t) = \frac{1}{4\pi v^2 |\boldsymbol{x}|}\delta\left(t - \frac{|\boldsymbol{x}|}{v}\right). \tag{4.2}$$

We can write this wave as a sum of its frequency components (i.e., an integral in this case), in the form

$$W(\boldsymbol{x},t) = \frac{1}{8\pi^2 v^2 |\boldsymbol{x}|} \int_{-\infty}^{\infty} \mathrm{d}\omega \exp\left[\mathrm{i}\omega\left(\frac{|\boldsymbol{x}|}{v} - t\right)\right]. \tag{4.3}$$

Note here the constant amplitude spectrum, as expected for a delta function.

In a similar fashion, we can go on to write the solution in terms of a sum of its wavenumber components too. If we carry out the complete triple transform from $\boldsymbol{x}$ to $\boldsymbol{k}$, as well as from t to ω, using the convention

$$W(\boldsymbol{k},\omega) = \int_{-\infty}^{\infty} \mathrm{d}t \int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{d}y \int_{-\infty}^{\infty} \mathrm{d}z W(\boldsymbol{x},t) \times \exp[-\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)], \tag{4.4}$$

where $\boldsymbol{k} = (k_x, k_y, k_z)$, then it follows from (4.1) that

$$W(\boldsymbol{k},\omega) = \frac{1}{(v^2\boldsymbol{k}\cdot\boldsymbol{k} - \omega^2)}. \tag{4.5}$$

It is a simple exercise to carry out the inverse transform over k_z (see, for example, Aki and Richards, 1980, pp195–197), and thus to obtain

$$W(\boldsymbol{x},t) = \frac{\mathrm{i}}{16\pi^3 v^2} \int_{-\infty}^{\infty} \mathrm{d}\omega \int_{-\infty}^{\infty} \mathrm{d}k_x \int_{-\infty}^{\infty} \mathrm{d}k_y \times \frac{\exp[\mathrm{i}(k_x x + k_y y + \sqrt{\frac{\omega^2}{v^2} - k_x^2 - k_y^2}\,|z| - \omega t)]}{\sqrt{\frac{\omega^2}{v^2} - k_x^2 - k_y^2}}. \tag{4.6}$$

Just as (4.3) is a sum over the frequency components involved in describing the time dependence of the solution, so (4.6) is a sum over also the two horizontal wave number components, (k_x, k_y), of the underlying spatial dependence of the spherical wave. The integrand of (4.6), regarded as a function of $(\boldsymbol{x}, t)$, is just a steady-state plane wave with amplitude given as a function of (k_x, k_y, ω). To be more specific in making this identification of a plane wave, it is useful to appreciate that a plane wave is in general a function of space and time in the combination $\boldsymbol{x}.\boldsymbol{s} - t$. Any such function has constant values over planes perpendicular to the vector $\boldsymbol{s}$. The wave propagates in the $\boldsymbol{s}$ direction, and $\boldsymbol{s}$ is called the *slowness vector*. Its amplitude is the reciprocal of the wave speed. The dependence of the

integrand of (4.6) upon $\boldsymbol{x}.\boldsymbol{s} - t$ can be identified, once we divide k_x and k_y by frequency ω. Thus,

$$W(\boldsymbol{x},t) = \frac{1}{16\pi^3 v^2} \int_{-\infty}^{\infty} \mathrm{i}\omega \,\mathrm{d}\omega \int_{-\infty}^{\infty} \mathrm{d}p_x \int_{-\infty}^{\infty} \mathrm{d}p_y \times \frac{\exp[\mathrm{i}\omega(p_x x + p_y y + \xi|z| - t)]}{\xi}. \tag{4.7}$$

where the slowness vector $\boldsymbol{s}$ has horizontal components $(p_x, p_y) = (k_x/\omega, k_y/\omega)$ and a vertical component

$$\xi = \sqrt{\frac{1}{v^2} - p_x^2 - p_y^2}. \tag{4.8}$$

Thus, the slowness $\boldsymbol{s}$ is $(p_x, p_y, \pm\xi)$, with $+\xi$ or $-\xi$ according as $z > 0$ or $z < 0$.

Another useful way to think of the representation (4.7) is in terms of the excitation of latent waves, available in the medium. Thus, the plane wave $\exp[\mathrm{i}\omega(\boldsymbol{x}.\boldsymbol{s} - t)]$ is a solution of the homogeneous wave equation (i.e., the wave equation with no source term). It is a wave that can exist with any amplitude and with any value of (ω, p_x, p_y), provided the remaining component of the slowness, the vertical component, is such as to provide a solution to the wave equation. This condition reduces simply to the requirement that $|\boldsymbol{s}| = 1/v$. Representation (4.7) then gives the amount by which this plane wave, specified by (ω, p_x, p_y), is excited by the particular source under consideration.

The representation (4.7) is in the form that typically arises for waves in layered media. For example, there is a phase factor in the integrand, pertinent to the wave of interest at the receiver. For media more complicated than the present simplest of examples, there will be extra factors in the integrand that are determined solely by the vertical structure of the medium and not by the receiver or source positions.

Starting from expressions such as (4.7), there are many different ways to evaluate the triple integrals. For example, one approach is simply to carry out this integration numerically, applying a taper at high values of $\pm\omega$, and $\pm\, p_x$, and $\pm\, p_y$. Another approach is to see where the integrand has stationary phase, and to let the position of these stationary points, which are fixed values of (p_x, p_y), independent of frequency, guide one to a numerical or analytic procedure. The frequency integration can either be done first, or last.

Another approach to evaluation of the integrals in (4.7), which we next develop, is to recognize that integration over the (p_x, p_y) plane can be

carried out with respect to the variables p and ϕ' introduced via $p_x = p\cos\phi'$ and $p_y = p\sin\phi'$. This is a natural way to proceed, if we recognize that the solution has axial symmetry about a vertical line through the source, and so the waves depend only on $r = \sqrt{(x^2+y^2)}$ rather than on x and y independently. This approach allows a reduction from two slowness integrals to a single integral over p. We find below that this gives a result that is very convenient for interpreting the solution physically, permitting generalization to a medium with many layers and interfaces.

Thus, let us also introduce the azimuthal angle ϕ such that $x = r\cos\phi$ and $y = r\sin\phi$. Then

$$p_x x + p_y y = pr\cos(\phi - \phi') \quad \text{and} \quad \xi = \sqrt{\frac{1}{v^2} - p^2},$$

and we have:

$$W(\boldsymbol{x}, t) = \frac{1}{16\pi^3 v^2}\int_{-\infty}^{\infty} \mathrm{i}\omega\,\mathrm{d}\omega \int_0^{\infty} \mathrm{d}p \frac{p}{\xi}\mathrm{e}^{\mathrm{j}\omega\xi|z|} I(\omega pr),$$

where

$$I(\omega pr) = \int_0^{2\pi} \mathrm{d}\phi' \mathrm{e}^{[\mathrm{i}\omega pr\cos(\phi-\phi')]}. \tag{4.9}$$

From tables of integrals we can find that

$$I(\omega pr) = 2\pi J_0(\omega pr)$$

and at this stage we introduce an approximation, valid if ωpr is much greater than 1, namely that

$$J_0(\omega pr) \sim \sqrt{\frac{2}{\pi\omega pr}}\cos(\omega pr - \frac{\pi}{4}).$$

Hence,

$$I(\omega pr) \sim \sqrt{\frac{8\pi}{\omega pr}}\cos(\omega pr - \frac{\pi}{4}).$$

[This approximation for $I(\omega pr)$ could have been obtained more directly by carrying out the standard stationary phase analysis for the integral for I in (4.9), near the two points $\phi' = \phi$ and $\phi' = \phi + \pi$.]

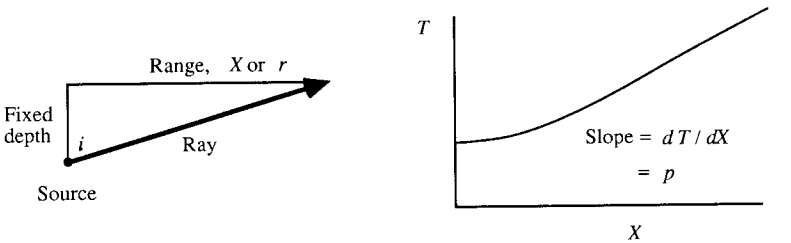


Fig. 32. Ray geometry (left), and travel-time curve (right),for the direct wave in a homogeneous medium.

Changing the range of the p integration in (4.9), we now have

$$W(\boldsymbol{x}, t) = \frac{1}{16\pi^3 v^2\sqrt{r}}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\pi/4}\sqrt{2\pi\omega}\,\mathrm{d}\omega \times \int_{-\infty}^{\infty} \mathrm{d}p \frac{\sqrt{p}}{\xi}\mathrm{e}^{\mathrm{i}\omega(pr+\xi|z|-t)}. \tag{4.10}$$

This is an approximation, the integrand here being accurate only if $|\omega pr| \gg 1$.

Interpretation of this representation of the wavefield can be given a physical basis by recognizing that the slowness has cartesian components

$$\boldsymbol{s} = \left(\frac{\sin i}{v}\cos\phi, \frac{\sin i}{v}\sin\phi, \frac{\cos i}{v}\right),$$

where i is the angle between the vertical and the ray direction. Therefore, the variable p in (4.10) is the so-called ray parameter,

$$p = \frac{\sin i}{v}.$$

When our knowledge of the wavefield is derived principally from the measurements made by, for example, hydrophones and seismometers at specific locations, these measurements are commonly first interpreted in terms of travel time curves, in which T, the travel time, is plotted against X(or r), the horizontal range, as in fig. 32. Note too that in representations such as (4.10) we are working with cylindrical waves, for which the exact form is given in terms of Bessel functions.

The tangent to the travel-time curve intercepts the T-axis at the so-called *delay time*, τ, where

$$\tau = T - pX. \tag{4.11}$$

Another natural name for τ is the intercept time. In the present example,

$$\tau = T - pX = \frac{zv}{\cos i} - \frac{\sin i}{v} z \tan i = \xi z,$$

which is the time taken for the phase of the plane wave, with horizontal slowness p, to travel a vertical distance z.

Representation (4.10) is now in the form

$$W(\boldsymbol{x}, t) = \int \mathrm{e}^{\mathrm{i}\pi/4} \sqrt{2\pi\omega}\, \mathrm{d}\omega \int_{-\infty}^{\infty} f(p) \mathrm{e}^{\mathrm{i}\omega[pr+\tau(p)-t]}\, \mathrm{d}p, \tag{4.12}$$

which is a quite general form, one that can often be found for body-wave ray paths of interest in seismology. To quantify the effects of a point source with a time dependence different from that implied by eq. (4.1), an extra function of frequency (the *source spectrum*) is needed in the frequency integral here. Another such function (the *instrument response*) can also be inserted to quantify the effect of a particular sensor and recording system, making $W(\boldsymbol{x}, t)$ a *synthetic seismogram.*

Before discussing some of the ways in which the double integral of (4.12) can be evaluated, let us consider another example in which the contribution associated with particular rays can be written in the form (4.12).

4.3. The plane wave/cylindrical wave components of a simple scattered wave

Thus, consider the effect of a plane interface between two homogeneous half spaces, in contact on the surface $z = 0$. If the source is at $z = (0, 0, -c)$ in the medium with wave speed v_1 (the region $z \leq 0$), then for receivers in the upper half space (see fig. 33) there can be both a direct wave and a reflected wave; and in the lower half space ($z \geq 0$, wave speed v_2) there can be a transmitted wave. The interaction of a plane wave and a planar interface between two half spaces is well understood in terms of reflected and transmitted plane waves, with amplitudes described by reflection and transmission coefficients, R and T, that are determined from satisfying boundary conditions at the interface.

For example, if we adapt (4.10) to give the incident wave for a source at $\boldsymbol{x} = (0, 0, -c)$ in medium one (wave speed v_1), with reflection back into this medium from an interface at $z = 0$, and transmission into medium two (wave speed v_2), we have:

$$W^{\text{Incident wave}}(\boldsymbol{x}, t) = \frac{1}{16\pi^3 v_1^2 \sqrt{r}} \int_{-\infty}^{\infty} e^{\mathrm{i}\pi/4} \sqrt{2\pi\omega}\, \mathrm{d}\omega \times \int_{-\infty}^{\infty} \mathrm{d}p \frac{\sqrt{p}}{\xi_1} \mathrm{e}^{\mathrm{i}\omega(pr+\xi_1|z-c|-t)}, \tag{4.13}$$

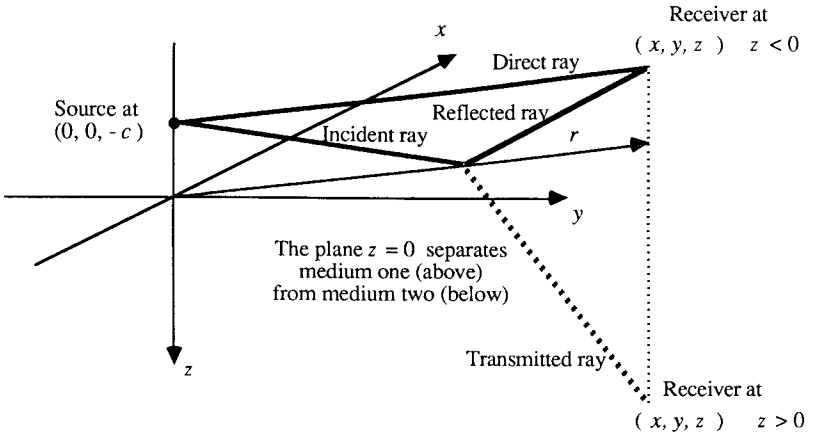


Fig. 33. Two half-spaces are in contact on the plane $z = 0$. 3D schematic, for four rays.

with

$$\xi_1 = \sqrt{\frac{1}{v_1^2} - p^2}.$$

This is an integral over cylindrical waves specified by their frequency ω and horizontal slowness p. It specifies both a direct wave, observed throughout the upper half-space, and the wave incident upon the interface. From this incident wave, a reflected wave is scattered upward, and a transmitted wave is scattered down into the lower half space.

The waves scattered by the interface must be

$$W^{\text{Reflected wave}}(\boldsymbol{x}, t) = \frac{1}{16\pi^3 v_1^2 \sqrt{r}} \int_{-\infty}^{\infty} e^{\mathrm{i}\pi/4} \sqrt{2\pi\omega}\, \mathrm{d}\omega \times \int_{-\infty}^{\infty} \mathrm{d}p \frac{\sqrt{p}}{\xi_1} R \mathrm{e}^{\mathrm{i}\omega(pr+\xi_1 c+\xi_1|z|-t)},$$

and

$$W^{\text{Transmitted wave}}(\boldsymbol{x}, t) = \frac{1}{16\pi^3 v_1^2 \sqrt{r}} \int_{-\infty}^{\infty} e^{\mathrm{i}\pi/4} \sqrt{2\pi\omega}\, \mathrm{d}\omega \int_{-\infty}^{\infty} \mathrm{d}p \frac{\sqrt{p}}{\xi_1} T \mathrm{e}^{\mathrm{i}\omega(pr+\xi_1 c+\xi_2 z-t)}. \tag{4.14}$$

Here, ξ_2 is the vertical slowness for the lower medium, and R and T are the cylindrical wave reflection and transmission coefficients. Fortunately,

R and T are also the plane wave coefficients, as can be seen by starting with (4.7) instead of (4.10) and finding that R and T depend on (p_x, p_y) only via the combination $p = \sqrt{(p_x^2 + p_y^2)}$. If W is the pressure, satisfying a scalar wave equation in a fluid, then from continuity of pressure and vertical displacement across $z = 0$ we find that

$$1 + R = T \quad \text{and} \quad \frac{\xi_1}{\rho_1}(1 - R) = \frac{\xi_2}{\rho_2}T, \quad \text{and hence}$$

$$R = \frac{\rho_2\xi_1 - \rho_1\xi_2}{\rho_2\xi_1 + \rho_1\xi_2} \quad \text{and} \quad T = \frac{2\rho_2\xi_1}{\rho_2\xi_1 + \rho_1\xi_2}. \tag{4.15}$$

We can now appreciate the main point of this examination of a simple problem in reflection and transmission, namely that, after substituting from (4.15) into (4.14), both reflected and transmitted waves have representations of the type given in (4.12). The phase factor, in each case, has a common term $\exp[\mathrm{i}\omega(pr - t),]$ this commonality being an expression of Snell's law (i.e., that the horizontal phase velocity is the same at every depth, in all the waves scattered by horizontal interfaces, for an incident wave of single fixed horizontal phase velocity). The phase factors also have a dependence upon vertical slownesses times vertical distance, accumulated along the path from source to receiver, and the resulting vertical time delay is in each case

$$\begin{aligned}\tau &= \int_{\mathrm{Ray}} \xi \,\mathrm{d}z = \int \frac{\cos i}{v}\,\mathrm{dz} = \int \frac{1 - \sin^2 i}{v\cos i}\,\mathrm{d}z \\ &= \int \frac{\mathrm{d}z}{v\cos i} - p\int \tan i \,\mathrm{d}z = T - pX.\end{aligned} \tag{4.16}$$

The generality of (4.12) is thus established. Even if many interfaces are crossed, via a ray that sometimes is reflected and sometimes transmitted, the principal phase factor due to propagation is always $\exp\{\mathrm{i}\omega[pr + \tau(p) - t]\}$. The factor $f(p)$ in (4.12) includes the product of plane wave coefficients appropriate for the interfaces crossed, and for the particular way in which energy is converted between wave types (for example, between downgoing and upgoing P and/or SV), as the wave of interest propagates along different segments of the total ray path. We refer to the wave representation (4.12) as the contribution from a *generalized ray*, and we next turn our attention to an evaluation of such double integrals.

4.4. Basic analysis of a generalized ray

The most important properties of generalized ray representations are determined by the position of stationary points, that is, values of p such that

$$\frac{\mathrm{d}}{\mathrm{d}p}[pr + \tau(p)] = 0.$$

But, another important property of the intercept time is that

$$\frac{\mathrm{d}\tau}{\mathrm{d}p} = \frac{\mathrm{d}T}{\mathrm{d}p} - p\frac{\mathrm{d}X}{\mathrm{d}p} - X = \frac{\mathrm{d}T}{\mathrm{d}p} - \frac{\mathrm{d}T}{\mathrm{d}X}\frac{\mathrm{d}X}{\mathrm{d}p} - X = -X = -X(p).$$

Hence, stationary points occur at p values such that

$$X(p) = r. \tag{4.17}$$

Solutions are thus p values such that a ray with ray parameter p fits between source and receiver. If p_0 is such a stationary point, we can make a Taylor series expansion for the phase and find that (4.12) is approximately

$$\begin{aligned}W(\boldsymbol{x}, t) = f(p_0) &\int e^{\mathrm{i}\pi/4}\sqrt{2\pi\omega}\,\mathrm{e}^{\mathrm{i}\omega[T_0 - t]}\,\mathrm{d}\omega \\ &\times \int_{-\infty}^{\infty} \exp\left[-\mathrm{i}\omega\frac{(p - p_0)^2}{2}\frac{\mathrm{d}X}{\mathrm{d}p}\right]\mathrm{d}p.\end{aligned} \tag{4.18}$$

where T_0 is the travel time along the ray. If X increases with p, as in fig. 32, so that $\mathrm{d}X/\,\mathrm{d}p$ is positive, this integral may be evaluated to give

$$W(\boldsymbol{x}, t) = \frac{f(p_0)}{\sqrt{\frac{\mathrm{d}X}{\mathrm{d}p}}}\int 2\pi\mathrm{e}^{\mathrm{i}\omega[T_0 - t]}\,\mathrm{d}\omega = 4\pi^2\frac{f(p_0)}{\sqrt{\frac{\mathrm{d}X}{\mathrm{d}p}}}\delta(t - T_0). \tag{4.19}$$

In the case of the direct ray, shown in fig. 33, it is a simple exercise to show that (4.19) is just our previous result, eq. (4.2) again. More generally, (4.19) gives the amplitude and arrival time expected on the basis of geometrical ray theory, with $\sqrt{(\,\mathrm{d}X/\,\mathrm{d}p)}$ being proportional to the geometrical spreading function, $R(\boldsymbol{x}, \xi)$, introduced in fig. 16.

The techniques by which the above wavefields are first written in terms of frequency and horizontal slowness, and then approximately evaluated after identifying points of stationary phase, would not be worthwhile if their outcome was merely another way to obtain the results predicted by

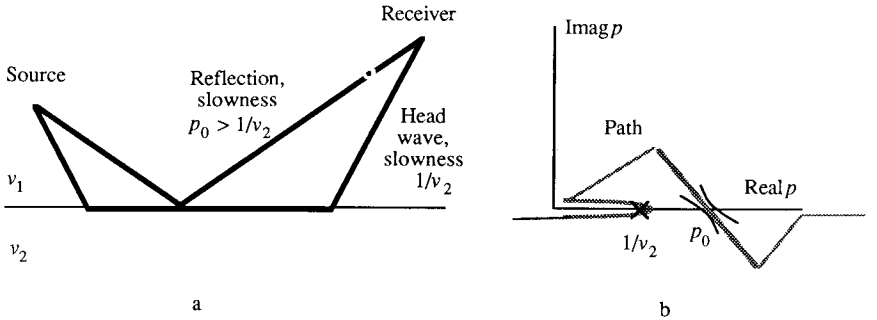


Fig. 34. (a) The ray paths of a reflected ray, and head wave. (b) Relative position of saddle and branch point, in the slowness plane. A path of integration, as shown, can numerically evaluate head wave and reflection contributions separately, these being respectively the integration taken over the part of the path around the branch point; and the integration taken over the segment of the path that goes across the saddle. Remaining parts of the path make negligible contribution.

geometrical ray theory. In fact, these techniques give much more. For example, if the two half-spaces shown in fig. 33 have a slower velocity in the upper half-space ($v_1 \leq v_2$), then a head wave can exist as shown in fig. 34.

It is associated with horizontal propagation at speed v_2 just below the interface, and therefore (by Snell's law) has horizontal slowness $p = 1/v_2$. The integral for the reflected wave in (4.14) includes the head wave contribution, and careful analysis shows how the head wave arrival may be made an identifiable part of this expression. The new contribution arises because the reflection coefficient R, a factor in the integrand, has a branch point at $p = 1/v_2$. The head wave may be evaluated after identifying first the path of steepest ascent and descent (crossing a "saddle point" in the complex p-plane at position p_0, see fig. 34b), associated with the wave that is truly reflected. This is a path on which integrands such as that of (4.18) have amplitudes given by a gaussian expression, $\exp[-(\omega/2)|p - p_0|^2 \, \mathrm{d}X/\,\mathrm{d}p]$. But the path of integration along which (4.14) is evaluated must meet the two constraints of being a path on which the integrand varies analytically, and being accessible from the original path (the real p-axis from $-\infty$ to $+\infty$) by path distortions that do not cross poles or branch points. The path sketched in fig. 34b is a suitable candidate (see also Aki and Richards, 1980, pp 207–214). Thus, in addition to the contribution from the saddle, one finds there is a contribution from having to go around the branch point. While this may seem an artificial way to quantify head wave properties, it reduces numerical evaluation to an integration over ranges of values of p that make sense from a physical point of view.

As well as the head wave, other waves guided by interfaces can exist. Thus, the structure may support one or more wave systems that have amplitudes which decay exponentially with distance from the interface, and these will typically be present only at discrete values of p. Examples are Rayleigh waves and Stoneley waves, and their discrete values of horizontal slowness are found to be poles of representations such as (4.14). Such interface waves may then be evaluated simply by the calculus of residues.

The extent to which one chooses to develop the relevant analysis, of head waves evaluated in terms of an integral around a branch pont, and interface waves evaluated in terms of residues, is essentially a matter of taste, since instead of obtaining detailed closed-form expressions corresponding to integrating around features in the complex p-plane, one may simply evaluate integrals such as (4.14) numerically, using just real values of p, perhaps with a taper at high and low values. However, such a numerical evaluation can be carried out much more efficiently and reliably if there is some understanding of the behavior of the relevant integrand, as a function of p.

4.5. Cagniard methods

For purposes of evaluation of multiple integrals in expressions such as (4.7) (a triple integral), and (4.10) and (4.12) (double integrals), it is sometimes possible to achieve the desired result without actually having to do any integration. Thus, a synthetic seismogram (a function of space and time) is achieved by playing the multiple integrals off against each other.

To give a sense of how this may be done, one can start, for example, with eq. (4.12) and see that

$$W(\boldsymbol{x}, \omega) = \mathrm{e}^{\mathrm{i}\pi/4} 2\pi\sqrt{2\pi\omega} \int_{-\infty}^{\infty} f(p) \mathrm{e}^{\mathrm{i}\omega[pr + \tau(p)]} \, \mathrm{d}p. \tag{4.20}$$

But we also know that the spectrum of $W(\boldsymbol{x}, t)$ is given by the forward transform,

$$W(\boldsymbol{x}, \omega) = \int_0^{\infty} W(\boldsymbol{x}, t) \mathrm{e}^{\mathrm{i}\omega t} \, \mathrm{d}t. \tag{4.21}$$

(There is no contribution from $t \leq 0$, because $W(\boldsymbol{x}, t)$ is zero at such times.)

Cagniard methods essentially entail manipulation of (4.20), without doing any integration, so that (4.20) is written in the form of (4.21). From a

uniqueness theorem, the two integrands can then be equated, and a closed form expression for $W(\boldsymbol{x}, t)$ is directly obtained. A key stage is that of equating exponents in the two integrands, so that

$$t = pr + \tau(p), \tag{4.22}$$

which then is used as an equation for p, regarded as a function of time t. As the dummy variable t of (4.21) is taken through all positive real values, from $t = 0$ to $t = \infty$, the solution $p = p(t)$ of (4.22) traces out a so-called Cagniard path in the complex p-plane.

Thus, recognizing in (4.20) that $W(\boldsymbol{x}, \omega)$ is a product of two different functions of frequency, it follows that $W(\boldsymbol{x}, t)$ is essentially a convolution of two functions of time.

The first, is a function of time whose spectrum is $\mathrm{e}^{\mathrm{i}\pi/4}\, 2\pi\sqrt{2\pi\omega}$. The second, is closely related to values of $f(p)\, \mathrm{d}p/\, \mathrm{d}t|_{p=p(t)}$.

A major application of this approach was made by Helmberger (1968), who showed how it could be done so quickly that it becomes feasible to carry out the procedure for many generalized rays, thus synthesizing the waves that propagate in multi-layered structures.

4.6. Incorporating anelasticity

If a representation for synthetics in an elastic medium has been obtained, such as (4.12), then the representation usually requires little change in application to the computation of synthetics for an anelastic medium. The change amounts to making a small perturbation in the Earth model, in which real elastic constants are replaced by complex-valued constants that correctly describe the anelastic relationships between stress and strain including the slight phase delay, and causality.

The perturbation can be characterized in terms of seismic velocities. Thus, where a velocity v appears in the elastic representation, its replacement by $v(1 - \mathrm{i}/2Q)$ in the anelastic representation will correctly change such expressions as

$$\exp \mathrm{i}\omega \left(\frac{x}{v} - t\right) \quad \text{to} \quad \exp \mathrm{i}\omega \left(\frac{x}{v} - t\right) \exp - \left(\frac{\omega x}{2vQ}\right).$$

[See eq.(3.3).] We are assuming $Q \gg 1$.)

However, though such a change can give correct amplitude spectra, it does not correctly enforce causality, for which a further perturbation is required, applied to the real-valued velocity.

To summarize: the substitution

$$v_{\text{elastic}} \to v_{\text{anelastic}} \left(1 - \frac{\mathrm{i}}{2Q}\right) \tag{4.23}$$

will give acausal results in the anelastic medium if $v_{\text{anelastic}}$ is independent of frequency. The same substitution, (4.23), will give causal results if $v_{\text{anelastic}}$ is frequency-dependent, and is chosen in conjunction with Q to form an attenuation–dispersion pair satisfying causality relations such as (3.13).

If the choice is made, not to impose causality, and to use (4.23) with Q as well as $v_{\text{anelastic}}$ independent of frequency, then the algorithm used in practice to obtain synthetics is typically not much more computationally intensive for anelastic than for elastic problems. But, if causality is imposed, for example in combining (4.23) and (3.16) so that

$$v_{\text{elastic}} \to v(\omega_0)[1 + \frac{1}{Q}\{\frac{1}{\pi} \ln(\frac{\omega}{\omega_0}) - \frac{\mathrm{i}}{2}\}], \tag{4.24}$$

then computation may require significantly longer execution times, since now the integrals for travel time, and intercept time, become not only complex valued, but frequency-dependent (O'Neill and Hill, 1979).

5. Matrix methods

5.1. Introduction

In the previous section, describing the use of plane waves or cylindrical waves to construct solutions for sources in depth-dependent structures, we worked with a scalar wavefield, namely the pressure in a fluid. But in seismology, we are typically interested in more complicated wavefields – for example, one or more components of displacement (or, sometimes, velocity, or acceleration, or strain). In this section we shall describe such wavefields and find that they can be built up from a sum of eigenvectors whose associated eigenvalues are related to vertical phase factors. This procedure explicitly brings out expressions for the physical quantities of interest (for example, components of displacement and traction that are continuous across interfaces). In this way, we develop a complete description of the solution of the governing wave equations – which we identify as a coupled first-order system. The fundamental solution is a matrix whose

columns are eigenvectors that characterize the physical properties of, for example, upgoing and downgoing P- and S-waves.

In a medium built up from horizontal layers with planar interfaces, the matrix solution can be used to include the internal multiple reflections, or used to characterize a particular generalized ray. In a medium with physical properties that vary continuously with depth, the fundamental matrix solution can be studied purely numerically, or by various asymptotic schemes.

We briefly describe some canonical problems for the generalized rays associated with an interface between two inhomogeneous layers, and list some of the many ways in which integrals over frequency and wavenumber are used to generate synthetic seismograms.

5.2. Use of column vectors, to describe a wave solution

The wave equation underlying section 1.4, concerning pressure in a homogeneous fluid, was very simple. It had plane wave scalar solutions, of type $\exp[\mathrm{i}\omega(\boldsymbol{x}\cdot\boldsymbol{s}-t)]$, with $|\boldsymbol{s}|=1/v$. It was simple, because in this case the phase factor is itself the wave solution. For seismic waves, the wave equation is of course more complicated, and the way in which we choose to describe the solution is not obvious. Seismic waves entail vector displacement, and there are associated tractions whose properties constrain the motion. We find that, in order to build up a systematic description of the important physical variables associated with a seismic wave solution, it is useful to develop a list of pertinent displacement and traction components, and to refer to it as a column vector of dependent variables. When we refer to the "wave solution" for an elastic motion in a solid, we mean this column vector.

For example, consider a downgoing P-wave in a homogeneous layer with density ρ, P-wave speed α, and S-wave speed β. If this is a plane wave propagating in the xz plane, then there is no displacement in the y direction. The slowness has cartesian components $s=(p,0,\xi)$, with the vertical slowness ξ given by:

$$\xi=\sqrt{\frac{1}{\alpha^2}-p^2},$$

and the phase of such a P-wave is $\exp[\mathrm{i}\omega(\boldsymbol{x}\cdot\boldsymbol{s}-t)]=\exp[\mathrm{i}\omega(px+\xi z-t)]$. Taking this as the scalar P-wave potential,we can evaluate any component of displacement and traction that we choose. The relevant components, for a systematic analysis of boundary conditions on horizontal interfaces in the

present P–SV problem for a layered elastic solid, are u_x, u_z, τ_{zx}, and τ_{zz}. Thus, we shall work in this case with the column vector

$$\boldsymbol{f}=\begin{pmatrix} u_x \\ u_z \\ \tau_{zx} \\ \tau_{zz} \end{pmatrix}. \tag{5.1}$$

From $\boldsymbol{u}=\nabla(P\text{-wave potential})$, together with an evaluation of the strain components, and Hooke's law for tractions in (5.1), the systematic description of physical variables in the downgoing plane P-wave is then given by

$$\boldsymbol{f}=\begin{pmatrix} u_x \\ u_z \\ \tau_{zx} \\ \tau_{zz} \end{pmatrix}=\begin{pmatrix} \alpha p \\ \alpha\xi \\ 2\mathrm{i}\omega\rho\alpha\beta^2 p\xi \\ \mathrm{i}\omega\rho\alpha(1-2\beta^2p^2) \end{pmatrix}\exp[\mathrm{i}\omega(px+\xi z-t)]. \tag{5.2}$$

This column vector has been normalized so that the P-wave displacement is a vector of unit amplitude, times the phase factor.

The associated downgoing SV-wave has the same horizontal phase factor, $\exp[\mathrm{i}\omega(px-t)]$, but a different vertical phase since now the overall phase of this plane wave is

$$\exp[\mathrm{i}\omega(\boldsymbol{x}\cdot\boldsymbol{s}-t)]=\exp[\mathrm{i}\omega(px+\eta z-t)] \quad \text{where} \quad \eta=\sqrt{\frac{1}{\beta^2}-p^2}.$$

Taking this phase factor as a potential, used to derive the SV-wave displacement via $\boldsymbol{u}=\nabla\times(0,SV-\text{wave potential},0)$, we can give explicitly the wave solution for the downgoing plane SV-wave. It is

$$\boldsymbol{f}=\begin{pmatrix} u_x \\ u_z \\ \tau_{zx} \\ \tau_{zz} \end{pmatrix}=\begin{pmatrix} \beta\eta \\ -\beta p \\ \mathrm{i}\omega\rho\beta(1-2\beta^2p^2) \\ -2\mathrm{i}\omega\rho\beta^3 p\eta \end{pmatrix}\exp[\mathrm{i}\omega(px+\eta z-t)]. \tag{5.3}$$

To complete the possibilities, for this P–SV example, we have also the upgoing waves:

$$\boldsymbol{f}=\begin{pmatrix} \alpha p \\ -\alpha\xi \\ -2\mathrm{i}\omega\rho\alpha\beta^2 p\xi \\ \mathrm{i}\omega\rho\alpha(1-2\beta^2p^2) \end{pmatrix}\exp[\mathrm{i}\omega(px-\xi z-t)], \quad \text{and}$$

$$\boldsymbol{f} = \begin{pmatrix} \beta\eta \\ \beta p \\ -\mathrm{i}\omega\rho\beta(1-2\beta^2p^2) \\ -2\mathrm{i}\omega\rho\beta^3p\eta \end{pmatrix} \exp[\mathrm{i}\omega(px - \eta z - t)]. \tag{5.4}$$

One of the main themes of this short course series is that a wave solution can often be written as a product of an amplitude factor, and a phase factor. Note that these explicit examples, of a column vector that describes the important physical components in a wave solution, also have this factorization. The factorization is useful, not just for analytical reasons, but for the consequences in a numerical evaluation. The wave field itself, in many situations in seismology, is typically observed only after propagation a distance of several tens of wavelengths (often, several hundreds of wavelengths), from source to receiver. A direct numerical evaluation of such fields would in some ways be like trying to track numerically the oscillating solution to $\mathrm{d}^2f/\mathrm{d}t^2 + \omega^2 f = 0$ across hundreds of wavelengths, instead of working with the known solutions, which are sums of $\sin\omega t$ and $\cos\omega t$. In the present examples, the phase factor has been obtained in closed form in each layer using ξz or ηz for the vertical delay time as in (5.2)–(5.4). More generally, it may be built up for a generalized ray using the intercept time τ, as in eq. (4.16), in a depth-dependent structure. Surprisingly, as we shall find in section 1.6, the factorization of a wave solution into an amplitude term, times a phase term in the form $\exp[\mathrm{i}\omega(pr + \tau - t)]$, can even be generalized to allow us to build up solutions for media that have three dimensional variability.

5.3. An eigenvector–eigenvalue approach to the wave solution

It is possible to find the differential equation satisfied by the column vector $\boldsymbol{f}$ of displacement/stress components. In general, for a linear elastodynamics problem, this equation has the form

$$\frac{\mathrm{d}}{\mathrm{d}z}\boldsymbol{f} = \mathbf{A}\boldsymbol{f}. \tag{5.5}$$

The matrix $\mathbf{A}$ is a function of frequency and horizontal wavenumber (or horizontal slowness) and material properties, and so is constant (independent of z) in a homogeneous layer.

For example, for the P–SV problem considered in the previous section, with $\boldsymbol{f}$ given by eq. (5.1), $\mathbf{A}$ is a 4 by 4 matrix. The coupled set of four first order equations, (5.5), includes both the wave equation, (1.13), and

Hooke's law, (1.9). A derivation is contained, for example, in Aki and Richards (1980, pp163–164 or p269). The matrix $\mathbf{A}$ is constant and is found to be

$$\mathbf{A} = \begin{pmatrix} 0 & -\mathrm{i}\omega p & \frac{1}{\mu} & 0 \\ \frac{-\mathrm{i}\omega p\lambda}{\lambda+2\mu} & 0 & 0 & \frac{1}{\lambda+2\mu} \\ \frac{4\omega^2p^2\mu(\lambda+\mu)}{\lambda+2\mu} - \rho\omega^2 & 0 & 0 & \frac{-\mathrm{i}\omega p\lambda}{\lambda+2\mu} \\ 0 & -\rho\omega^2 & -\mathrm{i}\omega p & 0 \end{pmatrix}. \tag{5.6}$$

Again and again, we have found in describing waves that it is natural to write the solution as a product of amplitude and phase factors. So here, we find that there is a natural factorization of solutions $\boldsymbol{f}$ of (5.5). To obtain this factorization, we first note that the eigenvalues of $\mathbf{A}$ are $\pm\mathrm{i}\omega\xi$ and $\pm\mathrm{i}\omega\eta$, where

$$\xi = \sqrt{\frac{1}{\alpha^2} - p^2}$$

and

$$\eta = \sqrt{\frac{1}{\beta^2} - p^2}.$$

In these expressions α and β are the P and S wave speeds.

The basis of our factorization will be that a solution of (5.5) is given by any eigenvector of $\mathbf{A}$, multiplied by $\mathrm{e}^{\text{corresponding eigenvalue, times } z}$. This is easily checked, by seeing for such a product, when substituted into both left-hand and right-hand sides of (5.5), that both sides scale the solution up by the same amount (the eigenvalue). The eigenvalues, given above, are proportional to the upgoing and downgoing P-wave and S-wave vertical slownesses. So, we define a phase matrix as follows:

$$\boldsymbol{\Lambda} = \begin{pmatrix} e^{\mathrm{i}\omega\xi z} & 0 & 0 & 0 \\ 0 & e^{\mathrm{i}\omega\eta z} & 0 & 0 \\ 0 & 0 & e^{-\mathrm{i}\omega\xi z} & 0 \\ 0 & 0 & 0 & e^{-\mathrm{i}\omega\eta z} \end{pmatrix}. \tag{5.7}$$

The ordering here is: downgoing P, downgoing S, upgoing P, upgoing S. We also define a matrix $\mathbf{E}$ whose columns are the eigenvectors of $\mathbf{A}$,

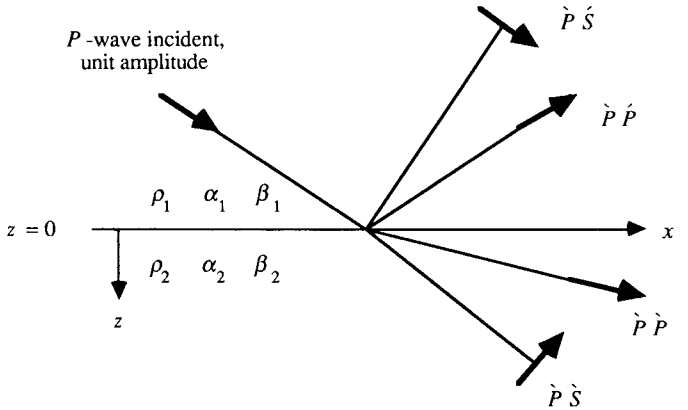


Fig. 35. The four scattered P–SV waves from a single interface, due to P incident from above. In this case, the four reflection and transmission coefficients are frequency independent.

arranged in the same order as that in which the corresponding eigenvalues appear in the phase matrix. It is

$$\mathbf{E} = \begin{pmatrix} \alpha p & \beta\eta & \alpha p & \beta\eta \\ \alpha\xi & -\beta p & -\alpha\xi & \beta p \\ 2i\omega\rho\alpha\beta^2 p\xi & i\omega\rho\beta(1-2\beta^2p^2) & -2i\omega\rho\alpha\beta^2 p\xi & -i\omega\rho\beta(1-2\beta^2p^2) \\ i\omega\rho\alpha(1-2\beta^2p^2) & -2i\omega\rho\beta^3 p\eta & i\omega\rho\alpha(1-2\beta^2p^2) & -2i\omega\rho\beta^3 p\eta \end{pmatrix} \tag{5.8}$$

The matrix formed by the product, $\mathbf{E}\boldsymbol{\Lambda}$, has columns which separately are solutions of eq. (5.5). Note too that these columns are just the column vectors we previously identified in (5.2)–(5.4) as giving the value of physical variables (displacement/stress components), associated with the separate P- and SV-waves.

It follows that the general P–SV solution is a linear combination of such columns, each of which is separately a solution of (5.5). This combination can be written as

$$\boldsymbol{f} = \mathbf{E}\Lambda w. \tag{5.9}$$

Here, the vector $\boldsymbol{w}$ is a column vector of constants, giving the weight of each of the basic P–SV waves present in the total solution $\boldsymbol{f}$. Because we have normalized the eigenvectors so that each has unit displacement, each of the elements of the weight vector gives a displacement amplitude directly.

The evaluation of reflection and transmission coefficients for a single interface provides an elementary example of the use of the eigenvector eigenvalue approach. Thus, in fig. 35 is shown a planar interface ($z = 0$) between two half-spaces. For the incident wave as shown, the four scattered waves are related to the weight vectors in each medium via

$$\boldsymbol{w}_1 = \begin{pmatrix} 1 \\ 0 \\ \dot{P}\acute{P} \\ \dot{P}\acute{S} \end{pmatrix}, \text{ and } \boldsymbol{w}_2 = \begin{pmatrix} \dot{P}\dot{P} \\ \dot{P}\dot{S} \\ 0 \\ 0 \end{pmatrix}. \tag{5.10}$$

But, since $\boldsymbol{f}$ is continuous across the interface, we know that

$$\boldsymbol{f}(0) = \mathbf{E}_1\boldsymbol{w}_1 = \mathbf{E}_2\boldsymbol{w}_2. \tag{5.11}$$

Here, $\mathbf{E}_1$ is the matrix of physical constants given by (5.8) with density and wave speeds appropriate for the upper medium; and $\mathbf{E}_2$ is evaluated from (5.8) for the lower medium. eqs. (5.10) and (5.11) result in four scalar equations for the four unknown scattering coefficients, and once found they can be used in expressions such as (4.14) to give the wavefield that propagates away from a point source. A detailed analysis can be given, of the body waves, head waves, and interface waves, along lines laid out in section 1.4. In general, these coefficients will have singularities at real values of p, corresponding to Stoneley waves that can propagate along the interface, and that may be evaluated in terms of polar residues of integrals such as (4.14).

5.4. Matrix methods for a medium composed of homogeneous layers

For a medium composed of many parallel layers, it is still true that $\mathrm{d}\boldsymbol{f}/\mathrm{d}z = \mathbf{A}\boldsymbol{f}$, but now the matrix $\mathbf{A}$ is piecewise constant as a function of depth, as is the matrix $\mathbf{E}$. We can still use the factorization $\boldsymbol{f} = \mathbf{E}\boldsymbol{\Lambda}\boldsymbol{w}$, and since $\boldsymbol{f}$ is continuous across interfaces but $\mathbf{E}$ is discontinuous, it follows that $\boldsymbol{w}$ must change from layer to layer. This makes sense on physical grounds, for the mix of upgoing and downgoing P- and SV-waves is different in each layer.

If the solution $\boldsymbol{f}$ is known on one interface, then it can be found at any other. With reference to fig. 36, we introduce the *propagator matrix*, $\mathbf{P}(z_n; z_0)$, relating the solutions $\boldsymbol{f}$ at levels z_0 and z_n via the equation

$$\boldsymbol{f}(z_n) = \mathbf{P}(z_n; z_0)\boldsymbol{f}(z_0). \tag{5.13}$$

The propagator can be built up as an explicit product of matrices, going down in the stack, alternately crossing layer and interface. Thus,

$$\boldsymbol{f}(z_1) = \mathbf{E}_1 \boldsymbol{\Lambda}_1(z_1)\boldsymbol{w}_1 \text{ and } \boldsymbol{f}(z_0) = \mathbf{E}_1 \boldsymbol{\Lambda}_1(z_0)\boldsymbol{w}_1,$$

giving

$$\boldsymbol{f}(z_1) = \mathbf{E}_1 \boldsymbol{\Lambda}_1(z_1)\boldsymbol{\Lambda}_1^{-1}(z_0)\mathbf{E}_1^{-1}\boldsymbol{f}(z_0),$$

and hence

$$\mathbf{P}(z_1; z_0) = \mathbf{E}_1 \boldsymbol{\Lambda}_1(z_1)\boldsymbol{\Lambda}_1^{-1}(z_0)\mathbf{E}_1^{-1}.$$

The propagator across a single layer, given here, can be found explicitly because $\mathbf{E}$ and Λ have explicit inverses. Continuing downward, we have

$$\boldsymbol{f}(z_2) = \mathbf{P}(z_2; z_1)\boldsymbol{f}(z_1) = \mathbf{P}(z_2; z_1)\mathbf{P}(z_1; z_0)\boldsymbol{f}(z_0),$$

and in general

$$\boldsymbol{f}(z_n) = \mathbf{P}(z_n; z_{n-1})\mathbf{P}(z_{n-1}; z_{n-2}) \ldots \mathbf{P}(z_1; z_0)\boldsymbol{f}(z_0),$$

so that

$$\mathbf{P}(z_n; z_0) = \prod_{i=0}^{n-1} \mathbf{P}(z_{n-i}; z_{n-i-1}). \tag{5.14}$$

Once the propagator is known, the scattering from a sequence of layers can be found. With the example of fig. 36, the four unknown scattering coefficients are determined from the four linear equations that result from the relation between $\boldsymbol{w}_{n+1}$ and $\boldsymbol{w}_0$. It is:

$$\begin{pmatrix} \grave{P}\grave{P} \\ \grave{P}\grave{S} \\ 0 \\ 0 \end{pmatrix} = \boldsymbol{w}_{n+1} = \boldsymbol{\Lambda}_{n+1}^{-1}\mathbf{E}_{n+1}^{-1}\boldsymbol{f}(z_n) = \boldsymbol{\Lambda}_{n+1}^{-1}\mathbf{E}_{n+1}^{-1}\mathbf{P}(z_n; z_0)\boldsymbol{f}(z_0)$$

$$= \boldsymbol{\Lambda}_{n+1}^{-1}\mathbf{E}_{n+1}^{-1}\mathbf{P}(z_n; z_0)\mathbf{E}_0\boldsymbol{\Lambda}_0(z_0) \begin{pmatrix} 1 \\ 0 \\ \grave{P}\acute{P} \\ \grave{P}\acute{S} \end{pmatrix}. \tag{5.15}$$

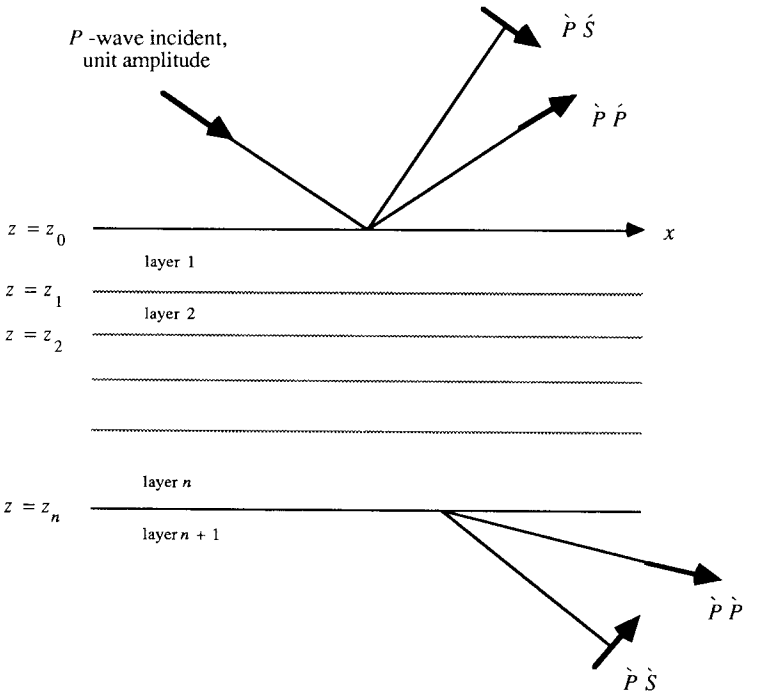


Fig. 36. The four scattered P SV waves from a sequence of interfaces, due to P incident from above. In this case, the four reflection and transmission coefficients are frequency dependent, since they include the effects of all the internal multiples, which interact differently at different frequencies.

Because these coefficients include all the internal multiples, the resulting synthetic seismogram in the time domain can be a series of long-lasting reverberations. Once the scattering coefficients have been obtained from (5.15), and stored for discrete values of ω and p (often, closely spaced values), they can be used in a direct integration such as given in (4.14). This procedure, in which scattering coefficients (R and T in eq. (4.14)) are obtained by matrix products such as in (5.15) and then integrated numerically over ω and p, is known in seismology as the reflectivity method. Reviews are given by Chapman and Orcutt (1985) and Müller (1985). Figure 37 shows an example of ocean-bottom seismometer data, consisting of significant reverberations, that is compared qualitatively with a synthetic seismogram computed by the reflectivity method. The display uses sonograms, which are power spectra calculated over a moving time window, in order to show the frequencies at which constructive and destructive interferences (resonances) are occurring.

Fig. 37. (upper) Vertical component acceleration and sonogram. Recorded by OBS Karen at 23°49′ S, 165°32′ W in 1983. (Personal communication from M. Hedlin and J.A. Orcutt.) (lower) Synthetic vertical acceleration and sonogram, obtained by the reflectivity method. The source is an explosion buried 30 meters at a range of 400 km. The structure consists of a low velocity surficial layer, in which the principal reverberations occur, and several other homogeneous layers over a half-space. (Personal communication from M. Hedlin and J.A. Orcutt.)

5.5. Matrix methods for continuously varying media

The equation $\mathrm{d}\boldsymbol{f}/\mathrm{d}z = \mathbf{A}\boldsymbol{f}$ still applies for a continuous variation of properties with depth, and the concept of a propagator is still relevant. Solutions for $\boldsymbol{f}$ can be found numerically and simply by a finite differencing appoach. Thus, starting from z_0, where $\boldsymbol{f}$ is presumed known, one can step down an amount Δz and use

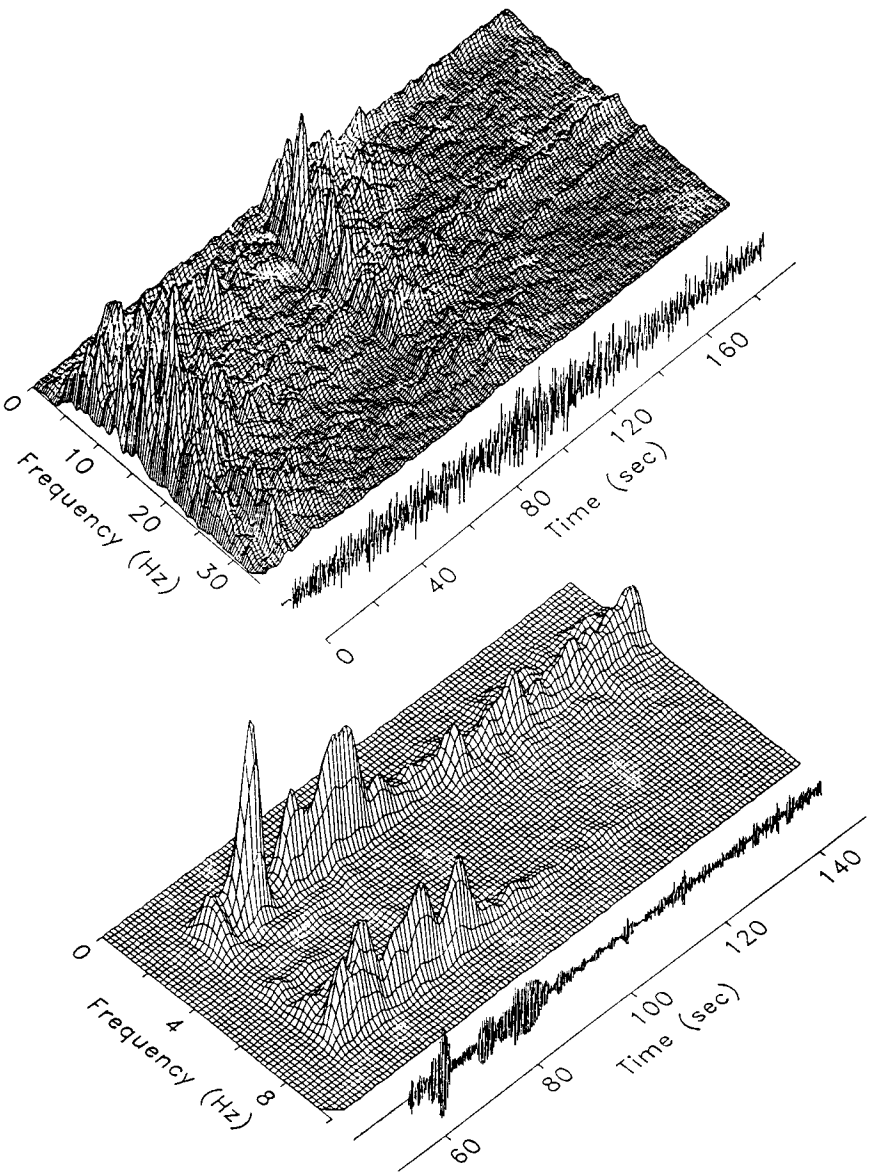
$$\boldsymbol{f}(z_0 + \Delta z) = \boldsymbol{f}(z_0) + \Delta z \frac{\mathrm{d}}{\mathrm{d}z}\boldsymbol{f} = \boldsymbol{f}(z_0) + \Delta z \mathbf{A}\mathrm{f} = (\boldsymbol{I} + \Delta z \mathbf{A})\boldsymbol{f}(z_0). \tag{5.16}$$

Repeatedly stepping down in this way, the solution at z_n can be constructed.

Alternatively, it is often possible to obtain an asymptotic approximation for the solution $\boldsymbol{f}$. Such approximations are useful, even if one is planning a purely numerical approach, because the behavior of the asymptotic solution indicates ways to implement numerical procedures accurately over propagation paths that are many wavelengths in extent. Thus, asymptotic methods help to determine the appropriate choice of Δz, which may be different at different depths and for different values of ω and p. What follows is an outline of the principal asymptotic results, for a medium in which velocities (both P and S) increase continuously with depth throughout an inhomogeneous layer. We continue to use the P SV problem as an illustrative example, i.e. eqs. (5.1), (5.5), and (5.6). The difficulty is now that the eigenvectors and eigenvalues of $\mathbf{A}$ are functions of depth. The eigenvalues, $\pm\mathrm{i}\omega\xi$ and $\pm\mathrm{i}\omega\eta$, where

$$\xi(p,z) = \sqrt{\frac{1}{\alpha(z)^2} - p^2} \quad \text{and} \quad \eta(p,z) = \sqrt{\frac{1}{\beta(z)^2} - p^2}, \tag{5.17}$$

can even change from imaginary to real values, as z increases with p held at a fixed real value. Such a depth is called a *turning point*. For a fixed



value of p, the turning point for S-waves is typically at much greater depth than that for P-waves.

Non-uniform asymptotics: WKBJ theory . The asymptotic solutions of a standard equation,

$$\frac{\mathrm{d}^2 f}{\mathrm{d}z^2} + \omega^2 \xi^2 f = 0,$$

with a single turning point, are

$$\frac{1}{\sqrt{\xi}} \exp \pm \mathrm{i}\omega\tau, \quad \text{where} \quad \tau = \int \xi \, \mathrm{d}z.$$

They are known as WKBJ solutions, acknowledging the work of Wentzel, Kramers, Brillouin, and Jeffreys. There is a similar solution for coupled differential equations such as (5.5).

Thus, using standard methods for solving coupled first-order linear differential equations, Richards (1971) showed that for continuous depth dependent media the solution of (5.5) can still be factorized as in eq. (5.9), $\boldsymbol{f} = \mathbf{E}\boldsymbol{\Lambda}\boldsymbol{w}$. The form of the solution is now

$$\boldsymbol{f} = \boldsymbol{f}(p, z, \omega) = \left[\mathbf{E}^{(0)} + \frac{\mathbf{E}^{(1)}}{\mathrm{i}\omega} + \frac{\mathbf{E}^{(2)}}{(\mathrm{i}\omega)^2} + \cdots \right]$$

$$\times \begin{pmatrix} \frac{\exp(+\mathrm{i}\omega\tau^P)}{\alpha\sqrt{\rho\xi}} & 0 & 0 & 0 \\ 0 & \frac{\exp(+\mathrm{i}\omega\tau^S)}{\beta\sqrt{\rho\eta}} & 0 & 0 \\ 0 & 0 & \frac{\exp(-\mathrm{i}\omega\tau^P)}{\alpha\sqrt{\rho\xi}} & 0 \\ 0 & 0 & 0 & \frac{\exp(-\mathrm{i}\omega\tau^S)}{\beta\sqrt{\rho\eta}} \end{pmatrix} \boldsymbol{w}. \quad (5.18)$$

Here $\tau^P = \int^z \xi \, \mathrm{d}z$ and $\tau^S = \int^z \eta \, \mathrm{d}z$, and the reference level, where these intercept times are zero, may be at any convenient depth. For many purposes, this reference level can be taken as the turning point. The matrix $\mathbf{E}$ of eigenvectors of $\mathbf{A}$ (see eq. (5.8)) is here replaced by a power series in descending powers of frequency. But for most purposes only the leading term, $\mathbf{E}^{(0)}$, is significant, and it is still given by eq. (5.8), though now with each of $\rho\ \alpha\ \beta\ \xi\ \eta$ depending on depth. Each column of $\mathbf{E}^{(0)}$ is therefore still an eigenvector, but now with a depth-dependent amplitude, and the amplitude factors appearing in the diagonal phase matrix in (5.18) are those which make the solution conserve energy flux across horizontal planes.

The framework in which the asymptotic form (5.18) is laid out is still one in which there are identifiable downgoing and upgoing P- and SV-waves, above the turning points, with a constant vector $\boldsymbol{w}$ giving the weight of each of these wave types in the inhomogeneous layer.

At high frequency, for smoothly varying media, use of (5.18) is accurate with just the leading term $\mathbf{E}^{(0)}$ (given by the local eigenvectors of $\mathbf{A}$). However, it is a non-uniform approximation in that it fails in the vicinity of turning points. The most serious effects appear, when $\boldsymbol{f}$ is to be evaluated at some fixed depth (where a boundary condition is to be applied), and when there is a need to work with values of p that make that fixed depth close to a turning point. Such problems arise, for example, in cases of near-grazing incidence of a ray upon a boundary. Though it is possible to develop a separate, non-uniform, expansion that is valid in the vicinity of turning points, it is more satisfactory to work with a uniform expansion, as we next discuss.

Uniform asymptotics: Langer theory . WKBJ asymptotic solutions fail in the vicinity of a turning point (where $\xi = 0$). However, defining

$$\tau = \tau(p, z) = \int_{\text{turning point}}^{z} \xi \, \mathrm{d}z,$$

the equation

$$\frac{\mathrm{d}^2 f}{\mathrm{d}z^2} + \omega^2 \xi^2 f = 0$$

has asymptotic solutions given in terms of Hankel functions of order one-third. The solutions are

$$\sqrt{\frac{\tau}{\xi}} H^{(1)}_{1/3}(\omega\tau) \quad \text{and} \quad \sqrt{\frac{\tau}{\xi}} H^{(2)}_{1/3}(\omega\tau).$$

Such uniformly asymptotic solutions, studied by Langer, and independently by many others, reproduce the properties of the WKBJ solutions away from the turning point, but remain valid even in the vicinity of $\xi = 0$.

For our present differential equation, (5.5), with turning points associated with both P- and S-wave motion, the WKBJ solutions given in (5.18) can be improved in a similar fashion.

For example, we can find the uniformly asymptotic approximation for $\boldsymbol{f} = \mathbf{E}\boldsymbol{\Lambda}\boldsymbol{w}$, where, as before, $\boldsymbol{w}$ is the vector of constants weighting the different linearly independent basic solutions in each layer. In this approximation, the matrix of phase factors becomes

$$\mathbf{\Lambda} = \begin{pmatrix} \sqrt{\frac{\tau^P}{\rho\alpha^2\xi}}H^{(1)}_{1/3}(\omega\tau^P) & 0 & 0 & 0 \\ 0 & \sqrt{\frac{\tau^S}{\rho\beta^2\eta}}H^{(1)}_{1/3}(\omega\tau^S) & 0 & 0 \\ 0 & 0 & \sqrt{\frac{\tau^P}{\rho\alpha^2\xi}}H^{(2)}_{1/3}(\omega\tau^P) & 0 \\ 0 & 0 & 0 & \sqrt{\frac{\tau^S}{\rho\beta^2\eta}}H^{(2)}_{1/3}(\omega\tau^S) \end{pmatrix} \tag{5.19}$$

The matrix $\mathbf{E}$, whose columns are eigenvectors of $\mathbf{A}$, becomes

$$\mathbf{E} = \begin{pmatrix} \alpha p & \beta\dot{\eta} & \alpha p & \beta\acute{\eta} \\ \alpha\dot{\xi} & -\beta p & -\alpha\acute{\xi} & \beta p \\ 2\mathrm{i}\omega\rho\alpha\beta^2\dot{\xi} & \mathrm{i}\omega\rho\beta(1-2\beta^2p^2) & -2\mathrm{i}\omega\rho\alpha\beta^2p\acute{\xi} & -\mathrm{i}\omega\rho\beta(1-2\beta^2p^2) \\ \mathrm{i}\omega\rho\alpha(1-2\beta^2p^2) & -2\mathrm{i}\omega\rho\beta^3p\dot{\eta} & \mathrm{i}\omega\rho\alpha(1-2\beta^2p^2) & -2\mathrm{i}\omega\rho\beta^3p\acute{\eta} \end{pmatrix}, \tag{5.20}$$

in which $\dot{\xi}$ and $\acute{\xi}$ are simply $\sqrt{\frac{1}{\alpha^2}-p^2} \equiv \xi$ above the turning point, but more generally

$$\dot{\xi} = \xi \mathrm{e}^{\mathrm{i}\pi/6}\frac{H^{(1)}_{2/3}(\omega\tau^P)}{H^{(1)}_{1/3}(\omega\tau^P)}, \quad \text{and}$$
$$\acute{\xi} = \xi \mathrm{e}^{-\mathrm{i}\pi/6}\frac{H^{(2)}_{2/3}(\omega\tau^P)}{H^{(2)}_{1/3}(\omega\tau^P)}, \tag{5.21}$$

with similar definitions for $\dot{\eta}$ and $\acute{\eta}$.

To see how (5.19) and (5.20) can be useful, it is helpful to identify the most important problems that arise in connection with seismic waves scattered at the horizontal interface between two inhomogeneous layers. Thus, for an interface where velocity above is faster than velocity below (but, with velocity increasing with depth in each inhomogeneous layer), a schematic is shown in fig. 38.

The question of whether a non-uniform or a uniform asymptotic approximation is appropriate is answered by checking to see if rays incident at near grazing angles are involved for the distance X at which seismograms are to be synthesized. For distances away from X_1 and X_2, the non-uniform asymptotics may be adequate, since the associated rays, to the extent that

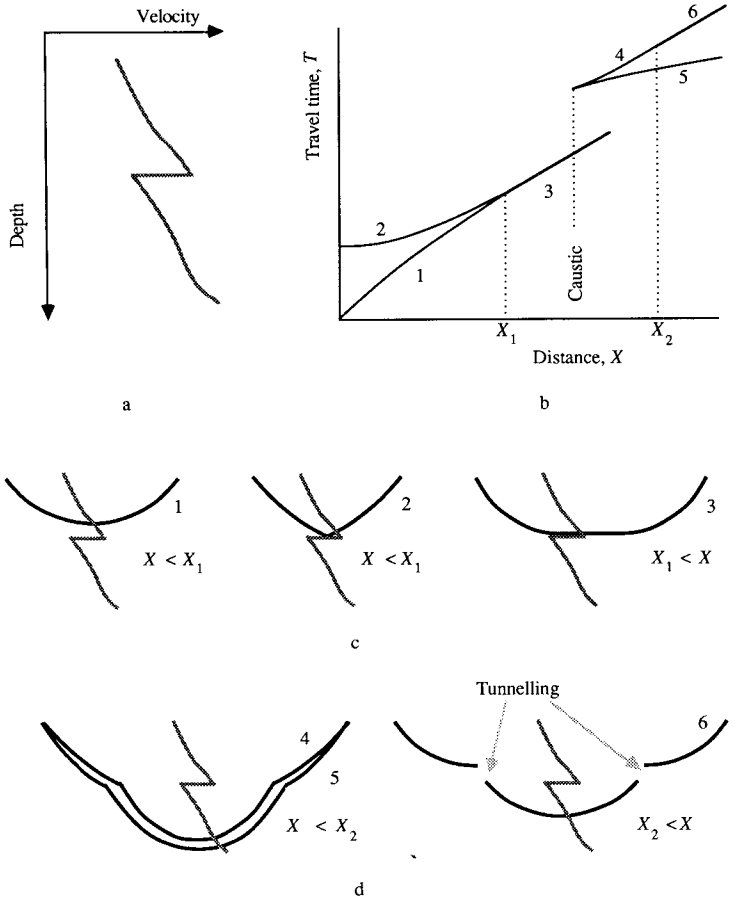


Fig. 38. Canonical problems, of rays and a single interface, when velocity just above the interface is faster than velocity below: (a) the profile; (b) the travel-time curve; (c) three rays that stay in the upper layer; (d) three rays that are transmitted into the lower layer. Grazing incidence is associated with rays that emerge near the distances X_1 and X_2. There is a caustic between these two distances, where rays 4 and 5 merge. Ray 3 is diffracted along the upper side of the interface. Ray 6 follows a path that entails "tunnelling" through the high velocities at the bottom of the upper layer.

they interact with the interface, do so at p-values (saddles of integrands such as (4.12)) which do not make the interface close to a turning point. But, a range of interesting wave propagation phenomena are found at distances near X_1 and X_2, which the uniform (Langer) approach is ideally equipped to handle. These include the way in which rays 1 and 2 (see

fig. 38bc) merge to become ray 3 at X_1; and the way in which energy can tunnel through the lowest levels of the upper medium, to reach the low velocity layer and thence travel to distances beyond X_2, by a path that geometric ray theory would not predict (see fig. 38d for ray 6).

For the situation, more common in geophysics, of an interface where velocity above is slower than velocity below, a somewhat more complex canonical problem emerges, as shown schematically in fig. 39. Again, provided the interaction of rays with interface occurs at angles away from grazing, a non uniform (WKBJ) approach to the analysis of each generalized ray is appropriate. At distances near X_2, the analysis may require a uniform (Langer) approach, but it is not significantly diferent from analysis of distances near X_1 in fig. 38. The problem that is more complex is the discussion of synthetics near the distance X_1 in fig. 39, for which an infinity of rays can be drawn, all with similar travel times. They are incident repeatedly upon the interface from below, with a reflection coefficient (near grazing) that typically is near unity, so that the contribution from higher-order multiples may not converge quickly. This is an example of the *whispering gallery* phenomenon, and, as usual in an observational science when something new is encountered, the complexity of this phenomenon is both a nuisance and an opportunity. It is a nuisance, because it is harder to find out how appropriate integrals like (4.12) can be obtained and evaluated numerically. But more important, it is an opportunity, because the interaction of the multiples that make up the whispering gallery results in a set of waveforms, observable at distances greater than X_1 and times just after that of ray 3, which are quite sensitive to the properties of the medium just below the interface. In particular, by selectively using waveforms observed in this distance range and time window, useful estimates may be made of the velocity gradient at the top of the layer just beneath the interface. Further details of this canonical problem, the whispering gallery, appear in Aki and Richards (1980, pp451–461).

5.6. Summary of basic choices for ω–k integration

For studying wave propagation in a laterally homogeneous structure, we have repeatedly used solutions in which the depth dependence, the horizontal dependence, and the time dependence appear via separate factors. Such basic solutions are typically specified by frequency ω and by horizontal wavenumber k, and we have seen how a subsequent integration over such solutions can give the wavefield due to a source that is active in only a limited region of space and time. However, because of various choices in the way integrands and integrations here can be approximated, there are

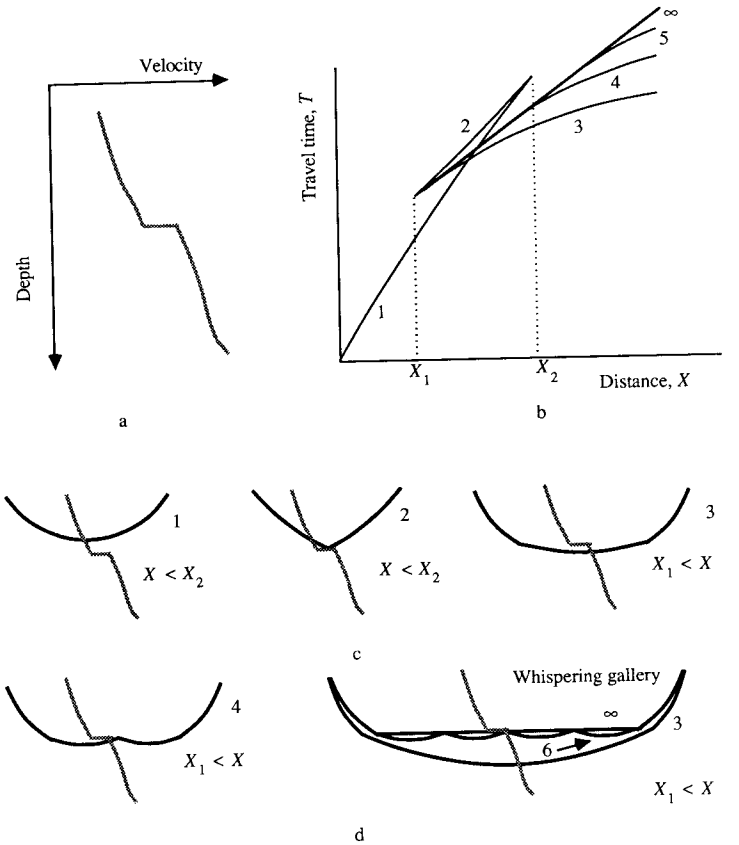


Fig. 39. Canonical problem of rays and a single interface, when velocity just above is slower than velocity below the interface: (a) the profile; (b) the travel-time curve, and numbered rays; (c) the most important three rays, which together form a "triplication", i.e. they are all three present between distances X_1 and X_2; (d) examples of rays that are reflected back downward below the interface. The sequence of rays 3, 4, 5, ... all merge to the same ray as distance X tends down to X_1 from above. For examples of how this canonical problem arises in seismic waves that interact with the Earth's cores, and how they are evaluated, see Aki and Richards (1980, pp447–461).

dozens of different ways in which the desired wavefield can be computed.

Thus, we have mentioned the reflectivity method for layered media earlier in this lecture, and pointed out that it has the merit of including the effect of all internal multiples in the layering. For this reason, such a method will often be favored, particularly if the synthetic is needed for a substantial time window – although there are choices here too, over the thickness of

layers needed to simulate a continuous velocity change, over the ranges of frequency and wavenumber needed in the integration, and over how finely the integrand should be sampled as a function of (ω, k).
However, in order to analyse isolated portions of a broadband seismogram the reflectivity method is not efficient. The main types of choice that can then arise are listed as follows:
Model choices
1. Many thin homogeneous layers, to simulate continuous depth-dependence of properties, or a few inhomogeneous layers, with relatively few interfaces?
2. Attenuation – constant Q or frequency-independent? Causal or acausal?

Integrand choices
The integrand will typically have many factors, but most of the effort goes into evaluation of those factors that depend directly on depth-dependence of the model.
1. Factorize $k = \omega p$? (Appropriate for analysing effects of rays. Less relevant for studying surface waves.)
2. Work with generalized rays, specifically identified for the problem being addressed, or use matrix methods for all multiples?
3. Use exact matrix methods, if working with many homogeneous layers (and then care is needed, to avoid round-off in extensive matrix products such as (5.14)); or use layer matrices with inhomogeneous layers, computed by asymptotics or by direct numerical integration of an arbitrary velocity profile?

Integration choices to obtain results in the $(\boldsymbol{x}, t)$ domain
1. Integrate over p first (or k), thus obtaining spectra as an intermediate step? (Such a procedure is called a *spectral method*. It is effective with generalized rays when a uniform asymptotic method is invoked, and/or if the medium has a frequency-dependent velocity profile, for example to handle attenuation accurately. There is still a choice, over whether to confine the integration to the real p-axis, or whether to use complex paths such as noted in fig. 34.)
2. Integrate over ω first (called a *slowness method*, very effective for elastic media: this too can be done with p integration real or complex)?

As noted in the extensive review of these issues by Chapman and Orcutt (1985), "the computation of synthetic seismograms is still a compromise between expense and accuracy." Of course, such compromises are made differently as computational facilities improve. But even if the complete synthetic were instantly and freely available for any depth-dependent structure, it would be important to know what part of the seismogram is influenced by what part of the structure – and such knowledge comes from an understanding of ω k and ω p representations, and figures such as 38 and 39.

6. Analytic study of seismic waves in 3D structures

6.1. Summary

Wave propagation in layered media has been studied fruitfully for several decades, with many applications in geophysics and other fields. The fundamental limitation of this type of modeling, in most cases, is its inability to handle lateral variations of structure. A recent discovery now indicates that analytical methods, developed and widely used for layers with plane parallel interfaces, can be extended to the case of plane interfaces that are not parallel.

Consequently, it appears possible to develop a generalized ray theory for wave propagation in a stack of wedge-shaped layers. Layer interfaces can have arbitrary dip and strike, permitting the study of waves in truly three dimensional structures. The methods that can be so generalized, from parallel layering to non-parallel, are those that derive from a fundamental and well known relationship between travel-time, range, and a sum over vertical slownesses (weighted by layer thickness). It is remarkable that this relationship, so familiar in purely depth-dependent structures, is virtually unchanged in the three-dimensional case.

This section outlines the newly generalized methods to study wave propagation associated with randomly dipping planar interfaces.

6.2. Introduction

As a contribution to methods for computing seismograms in three-dimensionally varying media, a generalized ray theory is outlined in this section for structures consisting of a stack of homogeneous layers that meet on planar interfaces of arbitrary dip and strike. A rapid procedure is then suggested for finding the rays, travel times, and geometrical spreading in such a model without the need to carry out ray tracing. The effects of intrinsic attenuation in each layer (or wedge) can easily be accommodated, and the method appears suitable for further generalizations that extend Cagniard–de Hoop theory, and other generalized ray theories such as disk ray theory (Wiggins, 1976; Chapman, 1976) to waves in layering with arbitrary dip and strike.

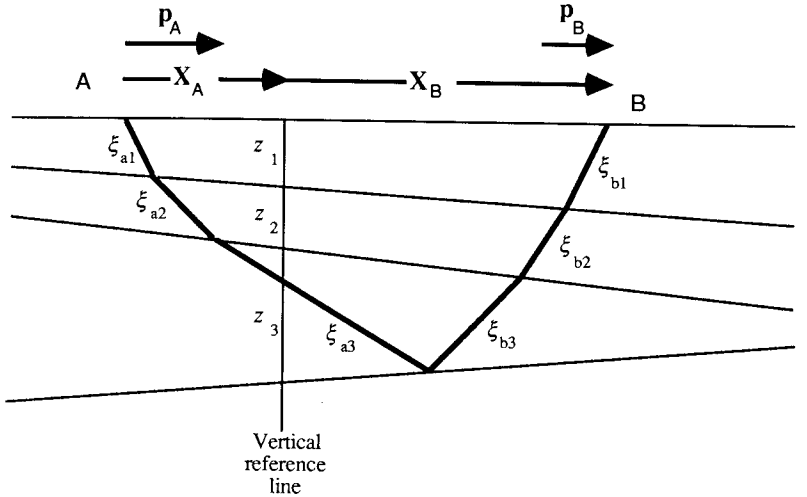


Fig. 40. "Layer thicknesses" are defined by intersection of interfaces with a vertical reference line. This is a 2D schematic of a 3D ray path from A to B which is composed of line segments that in general are not co-planar. Dipping interfaces are planar, but are not presumed to share a common strike. Notation is shown for vertical slownesses in each layer.

The basis for this work is a formula derived by Diebold (1987), giving the travel time along a ray path between two points in a stack of homogeneous layers with interfaces of arbitrary dip and strike. He proved that this time is

$$T = \boldsymbol{p}_{\mathrm{A}}.\boldsymbol{X}_{\mathrm{A}} + \boldsymbol{p}_{\mathrm{B}}.\boldsymbol{X}_{\mathrm{B}} + \sum_j (\xi_{\mathrm{a}j} + \xi_{\mathrm{b}j}) z_j \qquad (6.1)$$

where (see fig. 40): $\boldsymbol{p}_A$ is the horizontal slowness of the ray departing from a source at A; $\boldsymbol{X}_{\mathrm{A}}$ is the horizontal vector from A to a vertical reference line on which layer thicknesses z_j are defined; ξ_{aj} is the vertical slowness in the jth layer for the downward ray segments, and ξ_{bj} are the upward vertical slownesses (the summation is just for those layers traversed by the ray); and $\boldsymbol{p}_{\mathrm{B}}$ and $\boldsymbol{X}_{\mathrm{B}}$ are respectively the horizontal slowness and horizontal position vector from the vertical reference line, to a receiver at B.

Diebold's formula has long been known for the case of parallel layering, when it is trivial to prove from the travel time integral

$$T = \int_{\mathrm{ray}} \boldsymbol{s} \cdot \mathrm{d}\boldsymbol{l}.$$

The integrand here is slowness $\boldsymbol{s}$, and dl is an element of length along the ray. In cartesian coordinates with x horizontal in the plane of propagation, and z in the depth direction, components of slowness are s $= (p, \xi)$ with p constant along the ray (i.e., Snell's law); and dl $= (\mathrm{d}x, \mathrm{d}z)$. It follows for such one-dimensional structures that the travel time to the receiver at range X is

$$T = \int_{\mathrm{ray}} (p, \xi).(\mathrm{d}x, \mathrm{d}z) = pX + 2\sum_j \xi_j z_j = pX + \tau(p). \qquad (6.2)$$

The τ function here is the integral of vertical slowness ξ over the depth range spanned by the ray: τ is often called the intercept time, because from $\tau = T(p) - pX(p)$ it is the time intercepted on the time axis by the tangent to a travel-time curve (T vs. X).

A formula similar to (6.1) has been known for several years for two dimensional problems, i.e., for structures composed of homogeneous layers in contact on planes with arbitrary dip but all sharing a common strike direction, and with ray paths confined to the vertical plane perpendicular to that strike (Adachi, 1954; Johnson, 1976; Hong and Helmberger, 1977). It is remarkable, as Diebold (1987) showed, that (6.1) still applies even for three dimensional problems where the planar interfaces have arbitrary dip *and* strike. The total ray path is then composed of a series of straight line segments, with Snell's law applying locally on each interface between layers. The net result in 3D is a ray that is no longer confined to a single plane, let alone a vertical plane. The propagation path thus becomes much more difficult to draw, though conceptually there is little difficulty in tracing the ray from layer to layer, starting with a given "take off" direction (specified by $\boldsymbol{p}_{\mathrm{A}}$; or equivalently, by azimuth and magnitude of the horizontal phase velocity at the source). In eq. (6.1), the vertical slownesses differ, in general, for downward and upward ray segments in the same layer. The vertical reference line is fixed, but can initially be chosen anywhere (e.g., through the source, or at the receiver). Layer thicknesses z_j (see fig. 40) will differ, as will $\boldsymbol{X}_{\mathrm{A}}$ and $\boldsymbol{X}_{\mathrm{B}}$, for different choices of this reference line. But choices here do not influence the slowness components $\boldsymbol{p}_{\mathrm{A}}, \boldsymbol{p}_{\mathrm{B}}$, ξ_j.

Diebold (1987) gave a trigonometrical proof of eq. (6.1). The result is surprisingly simple, yet apparently had not been previously recognized for 3D applications. In the following section, his result is demonstrated from first principles with the intent of showing that a proof is hardly needed when the overall ray path is decomposed in a particular way; the result itself can simply be written down as self-evidently true. A procedure is

then outlined, for determining sequentially the vertical slownesses ξ_j along the ray path. This does not require ray tracing.

It is known that one of many possible uses of eq. (6.2) is the derivation of a generalized ray representation for parallel interfaces (including the effect of head waves if present: see earlier discussion of (4.12)). This section 1.6 will show that an outcome of Diebold's result is a generalized ray theory for media with planar dipping interfaces. In 3D problems this is likely, as a minimum, to give an effective procedure for computing ray paths, travel times and geometrical spreading. There is no need to iterate, to fit a ray between a particular source and receiver. If desired, attenuation (intrinsic dissipation) can readily be incorporated in the calculation of amplitudes. Other generalizations of the theory for parallel interfaces to 3D structures are likely to include Cagniard–de Hoop analysis of refractions (head waves) along dipping interfaces, and related extensions to WKBJ theory. The analytic method can to some extent be validated, by comparing its predictions (i.e. the synthetics) with those of geometrical ray theory for the simple problem of reflection from a single dipping interface.

6.3. The basic relation between travel time and slownesses

We here show that the 3D travel-time result can readily be written down even in complicated situations where layers may pinch out, where there is different layering beneath source and receiver, where $P-S$ conversions may occur, and when refractions and/or multiple legs are permitted in a particular layer. We then show explicitly how to account for changes in horizontal slowness along a ray path.

Proof of the basic travel-time formula . We begin with a simple problem, of one reflection from an interface as shown in fig. 41.

Source (at A) and receiver (at B) are on the same horizontal free surface, and the ray travels wholly within a homogeneous isotropic layer with seismic body wave speed v. In this case, both 3D slowness vectors $\boldsymbol{s}_\mathrm{A}$, $\boldsymbol{s}_\mathrm{B}$ have magnitude $1/v$. For general dip and strike, the plane of the two ray segments need not be vertical. To obtain the travel time along the ray, consider a plane wavefront moving perpendicular to the ray. The time to propagate from A to M_1 (see fig. 42) is $\boldsymbol{p}_\mathrm{A}.\boldsymbol{X}_\mathrm{A}$ because this is the time for the same wavefront to propagate in the horizontal plane along $\boldsymbol{X}_\mathrm{A}$ (i.e. to $\mathrm{X}_{\mathrm{ref}}$) with horizontal slowness $\boldsymbol{p}_\mathrm{A}$.

The time to propagate from $\mathrm{M}_1$to N_1 is $\xi_\mathrm{A} z$, because this is the time for the wavefront to propagate down the vertical reference line from $\mathrm{X}_{\mathrm{ref}}$ to Z with vertical slowness ξ_A. The time from $\mathrm{N}_1$to M_2 via R equals the time

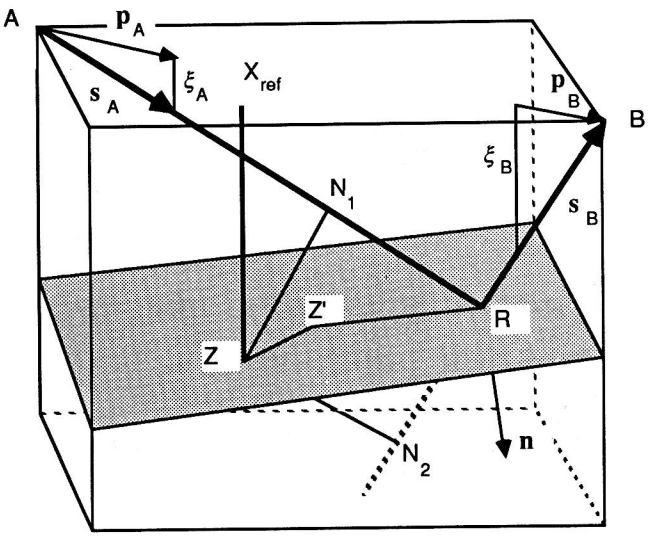


Fig. 41. A ray is shown, reflected at the point R on a planar interface. Incident and reflected slownesses are $\boldsymbol{s}_\mathrm{A}$ and $\boldsymbol{s}_\mathrm{B}$. These 3D vectors have horizontal components $\boldsymbol{p}_\mathrm{A}, \boldsymbol{p}_\mathrm{B}$ (2D vectors in the horizontal plane). Vertical slownesses are ξ_A and ξ_B (scalars). A vertical reference line, from a point X_{ref} in the horizontal surface, meets the dipping interface at the point Z. Lines $\mathrm{ZN}_1, \mathrm{ZN}_2$ are perpendiculars to the incident ray and to a continuation of the reflected ray. Lengths $\mathrm{N}_1\mathrm{R}, \mathrm{N}_2\mathrm{R}$ are equal. This can best be seen by dropping a perpendicular from Z to the plane containing $\boldsymbol{s}_\mathrm{A}, \boldsymbol{s}_\mathrm{B}$ and $\boldsymbol{n}$ (the normal to the dipping reflector). The foot of the perpendicular is Z′, and it too must lie in the dipping plane. Points $\mathrm{Z}', \mathrm{N}_1, \mathrm{R}, \mathrm{N}_2$, lie in the same plane, and equality of $\mathrm{N}_1\mathrm{R}$ and $\mathrm{N}_2\mathrm{R}$ follows from equality of angles of incidence and reflection $(\boldsymbol{s}_\mathrm{A}.\boldsymbol{n} = -\boldsymbol{s}_\mathrm{B}.\boldsymbol{n})$.

from $\mathrm{N}_2$to M_2 with speed v (from the equality of $\mathrm{N}_1\mathrm{R}$ and $\mathrm{N}_2\mathrm{R}$ claimed in the fig. 41 caption), which is $\xi_\mathrm{B} z$ because the wavefront is moving from Z up to $\mathrm{X}_{\mathrm{ref}}$ with vertical slowness ξ_B. Finally, the time from M_2 to B is $\boldsymbol{p}_\mathrm{B}.\boldsymbol{X}_\mathrm{B}$, as the wavefront propagates horizontally from $\mathrm{X}_{\mathrm{ref}}$ to B. It follows that

$$T = \boldsymbol{p}_\mathrm{A} \cdot \boldsymbol{X}_\mathrm{A} + \boldsymbol{p}_\mathrm{B} \cdot \boldsymbol{X}_\mathrm{B} + (\xi_\mathrm{A} + \xi_\mathrm{B}) z,$$

which is a simple example of eq. (6.1).

Next, consider two additional levels of complication as shown in fig. 43. Here, the source at A is not necessarily on the same horizontal level as the receiver at B. The downgoing ray is transmitted across another dipping interface, beneath a layer that pinches out so that it is not traversed by the eventual upgoing ray, reflected from a deeper dipping surface.

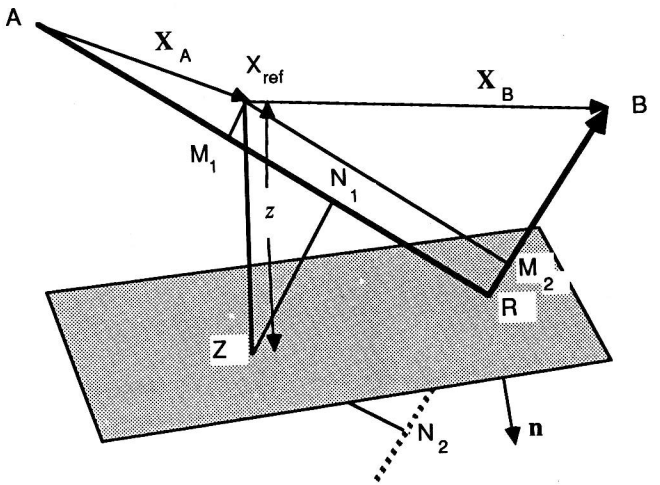


Fig. 42. Another view of the reflected ray shown in fig. 41. Lines M_1X_{ref}, N_1Z are perpendicular to the incident ray (AR). Line M_2X_{ref} is perpendicular to the reflected ray (RB).

Ray segments AR_1, R_1R_2 lie in the same plane as $\boldsymbol{n}_1$ (the normal to the dipping refractor). Z_1' is defined to lie in this plane, at the foot of the perpendicular from Z_1. It follows that Z_1' also lies in the refracting interface. To get the overall travel time from A to B consider again a plane wavefront moving along (and perpendicular to) the ray. The time from A to L_1 is $\boldsymbol{p}_A.\boldsymbol{X}_A$. The time from L_1 to M_1 is more interesting. It equals the time from L_1 to L_2 using the speed of the first ray segment, because the time from R_1 to L_2 at this speed equals the time from R_1 to M_1 at the speed of the refracted ray. This last equality is a direct consequence of Snell's law: the time from R_1 to L_2 for the incident ray equals the time for the wavefront to propagate obliquely from R_1 to Z_1'. All refracted or reflected wavefronts, including cases of conversion between P and S, are coupled on the interface and travel along it at the same speed. So, the time taken for oblique propagation at this speed, as any of these coupled wavefronts move from R_1 to Z_1', equals the time to travel directly along one of the rays, to the point (such as M_1) where the associated wavefront will then also lie on Z_1'. It remains to note that such a wavefront will then also lie on Z_1. For example, perpendiculars from Z_1' and Z_1, to the refracted ray R_1R_2, both meet that ray at M_1. The remainder of the path, from M_1 to R_2 and reflection to B, is similar to the discussion of fig. 42. We can get the overall travel time, from L_1 to M_2, by tracking on the vertical reference line the position of the wavefront moving perpendicular to the

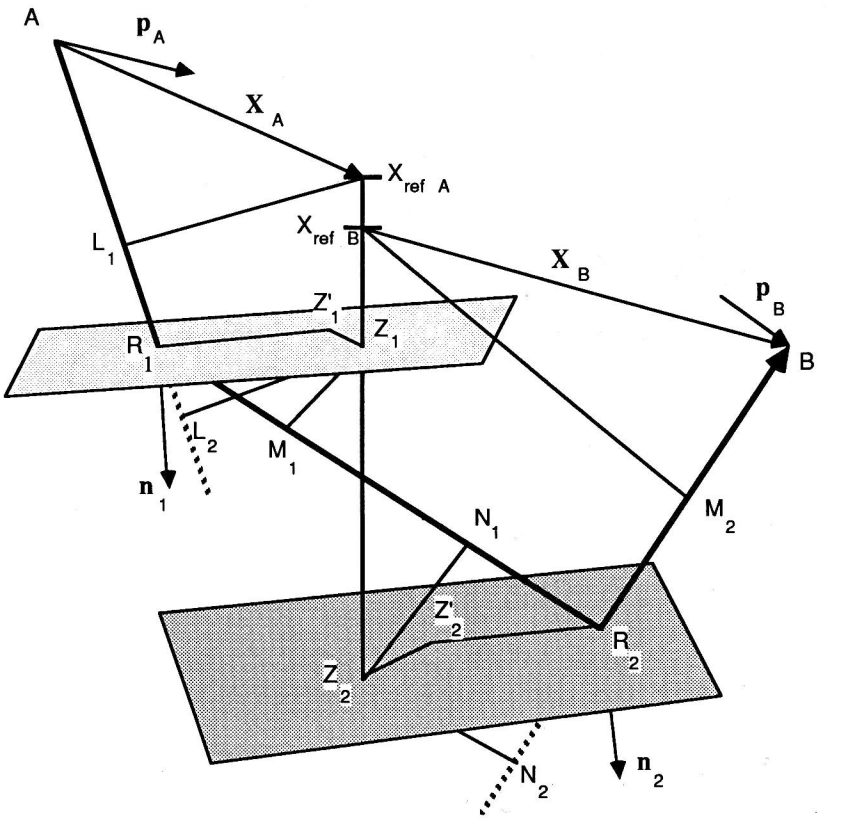


Fig. 43. The ray shown here from A to B is refracted at R_1 and reflected at R_2. A vertical reference line meets the two dipping interfaces at Z_1 and Z_2. X_{refA}, X_{refB} lie on this vertical line, at the horizontal levels of A and B respectively. For a plane wavefront propagating perpendicular to the ray, locations at L_1, M_1, N_1, and M_2 would put the wavefront also, in turn, at X_{refA}, Z_1, Z_2 and X_{refB}. Z_1L_2 and Z_2N_2 are perpendiculars to parts of the ray that are extrapolated out of the layer in which the real ray path is defined.

ray. Travel time along these segments is given schematically in fig. 44, and we find

$$T = \boldsymbol{p}_A \cdot \boldsymbol{X}_A + \boldsymbol{p}_B \cdot \boldsymbol{X}_B + \sum\nolimits_{ray} \xi_j z_j. \tag{6.3}$$

Equation (6.3) is almost the same as Diebold's (1987) result (eq. 6.1 above), but brings out the obvious generality that downgoing and upgoing ray legs need not cross the same interfaces. Pinchouts may occur. It is

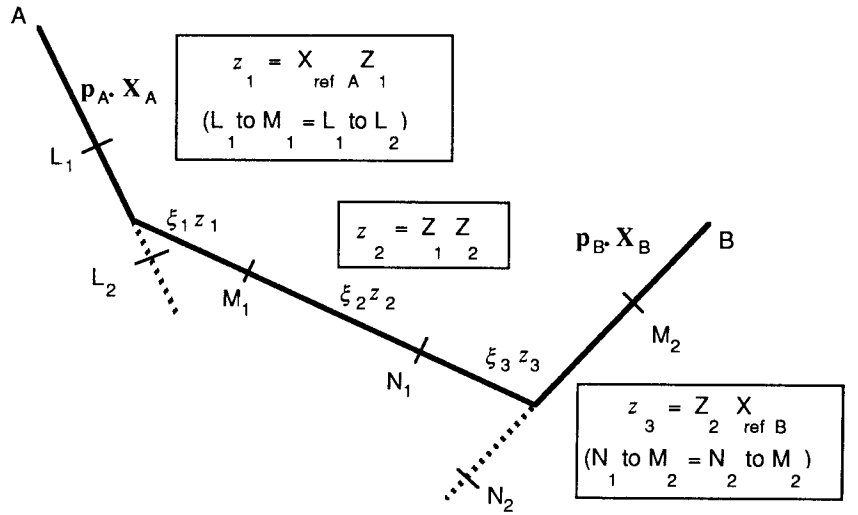


Fig. 44. Schematic of the ray path for fig. 43, showing propagation times for different segments. The z_j are effective layer thicknesses. Note: the three ray legs of the overall path would not in general lie in the same plane.

clear too that the method of proof can handle conversion between P and S, at any reflection or transmission. It is necessary only to choose the vertical slowness appropriately, whether for P or S, in each straight line ray segment. In forming the sum, $\sum \xi_j z_j$, it may seem natural to take positive signs always for the vertical slowness and for the layer thickness. For example, ξ_j and z_j would both appear to be positive for downgoing ray segments. Exceptions such as one shown in fig. 45 may occur. To the extent that a sign change appears appropriate for upgoing segments, each of ξ_j and z_j appears to require a sign change and so these changes might be ignored since the product is not affected. (Note: the effects of the factor $(\xi_{aj}+\xi_{bj})$ in eq. (6.1), and the factor 2 in eq. (6.2), are included in eq. (6.3) because the latter equation is taken along the ray path, and the former two equations are taken over depth.) However, in complicated situations where it is more difficult to keep track of signs, a more systematic procedure is required, as we next discuss.

Sequential evaluation of vertical slownesses . For horizontal plane layering, vertical slownesses are simply $\sqrt{(v^{-2} - p^2)}$ in a layer with body-wave speed v (whether for P or S): the scalar horizontal slowness p, also called the ray parameter, is (by Snell's Law) unchanged throughout the medium. We shall refer to this as the 1D problem. It has been studied for decades, and is described in sections 1.3 and 1.4. (See also Spencer, 1960, and

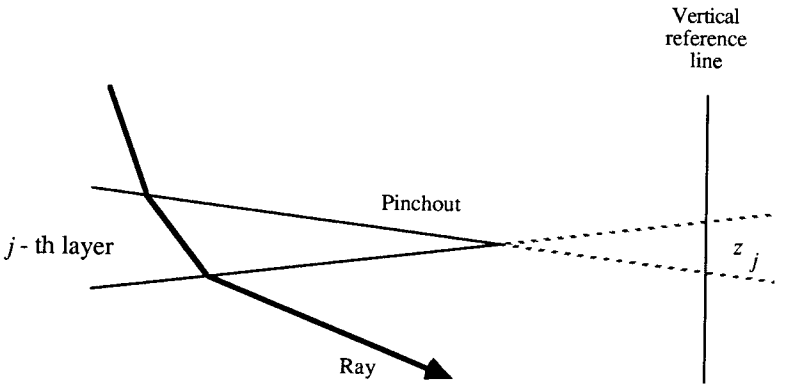


Fig. 45. Schematic in 2D, for a 3D situation in which pinchout occurs. The vertical reference line meets extrapolations of the two interfaces, defining a "layer thickness" that would enter eq. (6.3) with a negative value if the ray passed down through the jth layer (i.e. $\xi_j > 0, z_j < 0$, in this case).

Helmberger, 1968, for discussion of generalized rays; and Aki and Richards, 1980, chapters 5, 6, 7 and 9.)

For the 3D problem, horizontal slowness should be thought of as a 2D vector, and it is changed in general from one ray segment to the next, since interfaces need not be horizontal. For applications of eq. (6.3) it is useful to have each vertical slowness ξ_j defined in terms of the initial horizontal slowness $\boldsymbol{p}_A$, at the source. (In some cases of data analysis, for example when $\boldsymbol{p}_B$ is a measured quantity, the horizontal slowness at the receiver is the more natural independent variable.)

For the 3D problem let the sequence of layers traversed by the ray have, in turn, body wave velocities $v_A = v_0, v_1, v_2, v_3, \ldots, v_j, \ldots, v_J = v_B$. See fig. 46. At some interfaces there is reflection; at some, transmission. Some velocities in the list may be equal (e.g., adjacent velocities are equal for a P-wave reflection). The list can be a mix of P- and S-wave velocities, signifying conversion on reflection or transmission. A generalized ray is specified by this list, together with designation of the interaction occurring at each interface (reflection or transmission, conversion or no conversion).

Let the normals for interfaces "met" by the ray path be $\boldsymbol{n}_1$, $\boldsymbol{n}_2$, $\boldsymbol{n}_3, \ldots, \boldsymbol{n}_j, \ldots, \boldsymbol{n}_J$. The remaining details of Earth structure concern layer thicknesses, specified by the sequence $z_1, z_2, z_3, \ldots, z_j, \ldots, z_J$ of "layer thicknesses" measured on the vertical reference line between interfaces (extrapolated as necessary, for pinchouts: see fig. 45).

The travel-time formulas, equs. (6.1) or (6.3), do not require explicit knowledge of the length of ray segments in each layer. It is relatively simple to work out the sequence of slowness vectors $\boldsymbol{s}_0, \boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3, \ldots, \boldsymbol{s}_j, \ldots, \boldsymbol{s}_J$

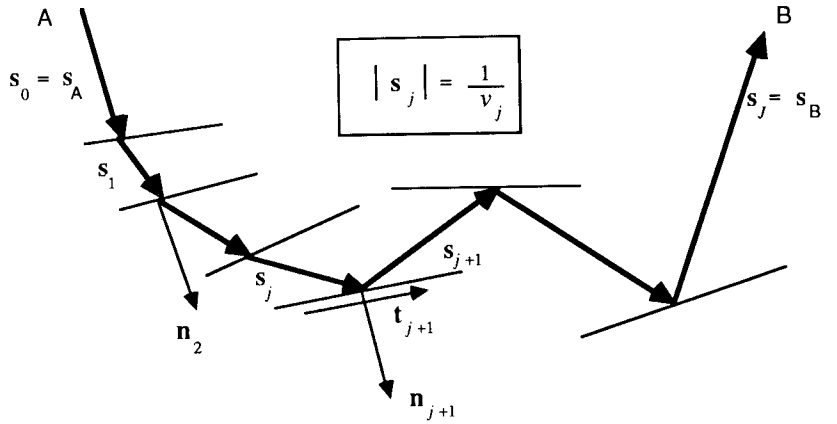


Fig. 46. 2D schematic of a 3D ray path, showing the labelling of layer velocities, slowness vectors, and interface normals. Interface $j+1$ lies between layers j and $j+1$. We here presume that source and receiver lie on horizontal surfaces (the zero-th and $J+1$th "interfaces", with normals $\boldsymbol{n}_0, \boldsymbol{n}_{J+1}$ that are vertical). Let Z_j be the depth of the point on the vertical reference line cut by the jth interface. Then, the "thickness" z_j of the jth layer is $Z_{j+1} - Z_j$. Vertical slowness ξ_j is the z-component of $\boldsymbol{s}_j$. In forming $\xi_j z_j$ with these definitions, note that vertical slownesses and "layer thicknesses" can be negative.

on each of the $J+1$ ray segments, since this does not depend on layer thicknesses. It does of course depend on the sequence of interface orientations. When specifying vector components in detail, for slowness $\boldsymbol{s}$ and interface normals $\boldsymbol{n}$, a useful right-handed coordinate system is (North, East, Down) for (x, y, z). Thus, for conventional definitions of dip δ, measured down from horizontal, and strike ϕ_s, measured clockwise round from North, the components of $\boldsymbol{n}$ are $(\sin\delta \sin\phi_s, -\sin\delta\cos\phi_s, \cos\delta)$. (See Aki and Richards, 1980, Fig. 4.13 for a convention on ranges of δ and ϕ_s: in what follows, $\boldsymbol{n}$ always has a non-negative $\boldsymbol{z}$ component.)

We are now in a position to give the sequence of vertical slownesses $\xi_0, \xi_1, \xi_2, \xi_3, \ldots, \xi_j, \ldots, \xi_J$ given the initial horizontal slowness vector $\boldsymbol{p}_A$ with (North, East) components (p_{Ax}, p_{Ay}).

Thus, since

$$s_0 = (p_{Ax}, p_{Ay}, \xi_0) \quad \text{and} \quad |\boldsymbol{s}_0| = \frac{1}{v_0}, \tag{6.4}$$

it follows that

$$\xi_0 = \sqrt{\frac{1}{v_0^2} - p_{Ax}^2 - p_{Ay}^2}, \tag{6.5}$$

so $\boldsymbol{s}_0$ and ξ_0 are known as functions of $(p_{Ax}, p_{Ay}) = \boldsymbol{p}_A$.

In general (see figs. 46 and 47), $\boldsymbol{s}_j$ and $\boldsymbol{s}_{j+1}$ are related by the rule

$$\boldsymbol{s}_j - (\boldsymbol{s}_j \cdot \boldsymbol{n}_{j+1})\boldsymbol{n}_{j+1} = \boldsymbol{s}_{j+1} - (\boldsymbol{s}_{j+1} \cdot \boldsymbol{n}_{j+1})\boldsymbol{n}_{j+1} = \boldsymbol{t}_{j+1}. \tag{6.6}$$

This is an expression of Snell's Law, equating the projection (along the interface) of incident and scattered wave slownesses. Equation (6.6) defines $\boldsymbol{t}_{j+1}$, the along-interface slowness vector (fig. 47). The equation can be turned into a explicit expression for $\boldsymbol{s}_{j+1}$ in terms of $\boldsymbol{s}_j$, noting that

$$\boldsymbol{s}_{j+1}.\boldsymbol{n}_{j+1} = \pm\sqrt{\frac{1}{v_{j+1}^2} - \boldsymbol{t}_{j+1} \cdot \boldsymbol{t}_{j+1}} \tag{6.7}$$

where the positive root is taken if the ray departs from the interface on the same side as $\boldsymbol{n}$, and the negative root if departure is on the other side from $\boldsymbol{n}$. Combining the last two equations,

$$\boldsymbol{s}_{j+1} = \boldsymbol{s}_j - (\boldsymbol{s}_j.\boldsymbol{n}_{j+1})\boldsymbol{n}_{j+1} \pm \sqrt{\frac{1}{v_{j+1}^2} - \frac{1}{v_j^2} + (\boldsymbol{s}_j.\boldsymbol{n}_{j+1})^2}\,\boldsymbol{n}_{j+1} \tag{6.8}$$

$$\text{and } \xi_{j+1} = z - \text{component of } \boldsymbol{s}_{j+1} = \boldsymbol{s}_{j+1} \cdot (0, 0, 1). \tag{6.9}$$

Equations (6.4)–(6.9) give a method for computing all relevant vertical slownesses along the generalized ray path, as a function of initial horizontal slowness components $(p_{Ax}, p_{Ay}) = \boldsymbol{p}_A$. The magnitude of each of the $\boldsymbol{t}_{j+1}$ is all that is needed locally to evaluate specific reflection/transmissions coefficients (see Aki and Richards, 1980, chapter 5, for detailed formulas for these coefficients). If $\boldsymbol{p}_B$ is needed, it can be found from the horizontal components of $\boldsymbol{s}_J$.

For S -waves, the question of polarization also arises (see fig. 47). Between interfaces j and $j+1$ for an S-wave, let particle motion be in the direction ν_j (a unit vector, perpendicular to $\boldsymbol{s}_j$ because the motion

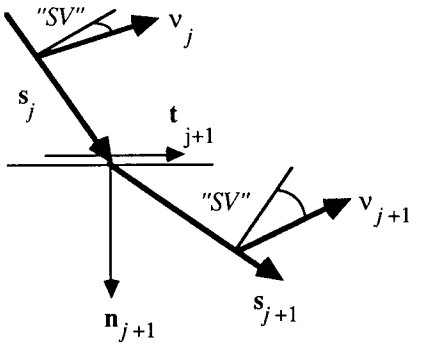


Fig. 47. The direction of vector S-motion is quantified by unit vectors ν_j , ν_{j+1} , transverse to the ray. Vectors $\boldsymbol{s}_j$, $\boldsymbol{s}_{j+1}$, $\boldsymbol{n}_{j+1}$, $\boldsymbol{t}_{j+1}$ all lie in a common plane, labelled Π (say): the projection of ν_j in this plane gives the fraction of incident S-displacement that locally is "SV". Vector subtraction of this in-plane component, or alternatively the projection of ν_j on to the normal to Π, gives the incident local "SH" component.

is transverse to the ray). The projection of ν_j on to the plane containing $(\boldsymbol{s}_j, \boldsymbol{s}_{j+1}, \boldsymbol{n}_{j+1}, \boldsymbol{t}_{j+1})$ then gives the fraction of S-displacement to be regarded locally as incident SV. Incident SH is given by a vector subtraction. On emergence from the interface, separately using SV and SH coefficients of reflection or transmission for the local SV and SH components, vector S-displacement is reconstituted from local SV and SH to determine the new direction ν_{j+1} of transverse particle motion propagated with slowness $\boldsymbol{s}_{j+1}$ to the next interface.

Note that computations for the sequence of slowness vectors (and for the directions of particle motion) depend only on wave speeds in each layer, and on the orientations of interfaces. Specifically, these computations do not depend on layer thicknesses. Because of this invariance, one can say that the forward computation of travel time, for a given source and a given take-off direction, can be done without ray tracing. It appears that in general one does not know what receiver positions will be reached by this ray. However, in the next section it is shown how to find the take-off direction for a ray that reaches a particular receiver.

6.4. Ray calculations without ray tracing

6.4.1. One dimensional

There are many conceivable uses of the travel-time formula, eq. (6.3). An elementary use concerns fitting a ray between a fixed source and receiver,

and finding the geometrical spreading, without actually doing any ray tracing. To illustrate this use, the method will first be reviewed for the problem of rays in 1D structures.

Thus, in the frequency domain, the wavefield for a generalized ray in a structure that varies solely with depth can be written

$$W(X_0,\omega) = \int_\Gamma f(p,\omega)\exp[\mathrm{i}\omega\{T(p,\omega) - pX(p,\omega) + pX_0\}]\,\mathrm{d}p \quad (6.10)$$

where Γ is a path of integration in the p-plane (see discussion of eq. (4.12)); $T(p,\omega) - pX(p,\omega)$ is the vertical integral of vertical slowness along the ray path,

$$T - pX = \tau = \int_{\mathrm{ray}} \sqrt{\frac{1}{v^2} - p^2}\,\mathrm{d}z, \quad (6.11)$$

in which $v(z)$ is the velocity (complex, if there is attenuation, and frequency-dependent if there is dispersion); and $f(p,\omega)$ includes the radiation pattern as well as relevant reflection/transmission coefficients for the generalized ray of interest.

Both plane-stratified and spherically-stratified media can be studied via (6.10). Richards (1973) obtained examples of eq. (6.10) even for rays that have turning points, rather than reflection from a specific interface. From eq. (4.17), but with the notation of (6.10), we expect stationary points to occur at values p_0 of the ray parameter, which are solutions of the equation $X(p) = X_0$.

In order to obtain the ray-theory approximation for an array of receivers, consider the following six steps:

(i) One fixes ω and computes f and $\tau(p)$ for a set of equally spaced real p-values, in a range of the real ray-parameter axis that is close to where one expects complex saddle points to lie, for the range of distances (i.e. X_0-values) at which one wishes to know W. Let the increment in p here be Δp.

(ii) At fixed X_0, one searches through the sampled values of $\Re[T - pX + pX_0]$ to see at which of the discrete points in p it has its least value (or greatest value, for certain kinds of ray). Label this discrete point as p_{js}. The complex saddle must presumably lie near this point of the real p-axis. If there is no attenuation, the saddle will lie on the real p-axis if there is a real ray between source and receiver.

(iii) Find the values (complex, if there is attenuation) of the three constants T_0, p_0, and $DXDP$ that best fit the sampled phase factor according to

$$T - pX + pX_0 = T_0 - \frac{1}{2}(p - p_0)^2 \times DXDP \tag{6.12}$$

in the vicinity of p_{js}. In this way, one obtains in p_0 an estimate of the position of the complex saddle; in T_0 an estimate of the complex travel time; and in $DXDP$ (signifying dX/dp at the saddle) an estimate of the relevant constant, complex in general, needed to evaluate geometrical spreading. Note that if one just uses p_{js} itself, and the points just before and just after it, closed form expressions that in practice are often very accurate can easily be given for p_0, T_0, and $DXDP$. Aki and Richards (1980) define the function $J(p) = T - pX + pX_0$ (see their page 423). In terms of this sampled function,

$$J_{js-1} = J(p_{js-1}) = T_0 - (p_{js} - p_0)^2 \times DXDP/2 + \Delta p.(p_{js} - p_0) \times DXDP - (\Delta p)^2 DXDP/2,$$

$$J_{js} = J(p_{js}) = T_0 - (p_{js} - p_0)^2 \times DXDP/2,$$

$$J_{js+1} = J(p_{js+1}) = T_0 - (p_{js} - p_0)^2 \times DXDP/2 - \Delta p.(p_{js} - p_0) \times DXDP - (\Delta p)^2.DXDP/2.$$

Therefore, the closed form expressions for the desired constants are

$$DXDP = -[J_{js+1} - 2J_{js} + J_{js-1}]/(\Delta p)^2,$$

$$p_0 = p_{js} + [J_{js+1} - J_{js-1}]/(2.\Delta p.DXDP)$$

and finally

$$T_0 = J_{js} + (p_{js} - p_0)^2 \times DXDP/2. \tag{6.13}$$

(iv) Interpolate from $f(p_{js}, \omega)$ and $f(p_{js+1}, \omega)$ to evaluate $f(p_0, \omega)$.

(v) Claim that the saddle point approximation to W is

$$W(X_0, \omega) = f(p_0, \omega)\sqrt{\frac{2\pi}{\mathrm{i}\omega DXDP}}\,\mathrm{e}^{\mathrm{i}\omega T_0}. \tag{6.14}$$

Choose the next value of X_0, and loop back to the second step to do various distances. Choose the next value of frequency ω, and loop back to the first step. This loop is necessary only if Q is frequency-dependent, and if there is allowance for body-wave dispersion. For some types of anelasticity, it may be possible to abbreviate this step and estimate more directly the global frequency-dependence of the saddle-point approximation.

(vi) Finally, one can go to the time domain:

$$W(X_0, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} W(X_0, \omega)\,\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}\omega. \tag{6.15}$$

Note that the whole effort is accomplished using real values of ray-parameter, at the time-consuming stage of tabulating f and τ at discrete p values. Often, the dependence on frequency in eq. (6.14) is sufficiently simple that eq. (6.15) can be evaluated explicitly. Since the final time series is real, (6.15) can be addressed using only positive frequencies.

The method allows easily for investigation of different attenuation-dispersion pairs. This is handled just by insertion of some appropriate rule for evaluating $v(z, \omega)$, such as (4.24), and making allowance for the fact that $\tau = \tau(p, \omega)$. The method is simple and rapid to execute, and has been found quite accurate for 1D problems (Richards, 1985, pages 197-199). A series of papers by Borcherdt (including Borcherdt et al., 1986: see also Richards, 1984) has drawn attention to the need for more careful handling of attenuation. Specifically he has advocated working with plane waves that may propagate in a direction that differs from the direction of most rapid attenuation. (The "direction of propagation" is the direction of most rapid phase increase.) He has shown that such waves have properties that much conventional analysis (i.e., that based on plane waves propagating in the direction of maximum attenuation) cannot reproduce. Fortunately, the method based on equs. (6.10) through (6.14), which finds stationary values of the integrand at complex values of horizontal slowness, gets around Borcherdt's valid objections to conventional analysis, yet does so for attenuating media in a way that requires little change from computational experience with elastic media.

6.4.2. Three dimensional

Each of steps (i)–(vi) above can be expected to carry over to 3D problems in layered structures with dipping planar interfaces. The major differences will concern the need to work with two components of horizontal slowness, $(p_{Ax}, p_{Ay}) = \boldsymbol{p}_A$. Let us outline part of the theory in a particularly simple case, choosing the vertical reference line to lie through the receiver so that $\boldsymbol{p}_B = \mathbf{o}$. Without ambiguity we can then drop the subscript A from $\boldsymbol{p}_A$, writing $\boldsymbol{p} = (p_x, p_y)$ for the horizontal slowness at the source. We now have two scalar independent variables, and it is instructive to use them to reproduce the known field of a point source in a homogeneous medium. Thus, consider

$$W(\boldsymbol{X}_0, \omega) = \frac{\exp(\mathrm{i}\omega R/\alpha)}{R},$$

where $R = |\boldsymbol{X}_0|$, which describes a spherical wave at $\boldsymbol{X}_0 = (X_0, Y_0, Z_0)$, expanding from the origin.

Then it is known that W can be expressed as

$$W(\boldsymbol{X}_0, \omega) = \frac{\mathrm{i}\omega}{2\pi} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\xi} \exp[\mathrm{i}\omega(p_x X_0 + p_y Y_0 + \xi|z|)]\, \mathrm{d}p_x\, \mathrm{d}p_y$$

$$\text{where} \quad \xi = \sqrt{\frac{1}{v^2} - p_x^2 - p_y^2}. \tag{6.16}$$

(For derivation of a related result, see (4.7), or Aki and Richards, 1980, pp197-199.) Defining J by

$$J(p_x, p_y) = \boldsymbol{p}.\boldsymbol{X}_0 + \tau(p_x, p_y),$$

which equals $(p_x X_0 + p_y Y_0 + \xi|z|)$ here in eq. (6.16), we can expect to find an expansion about the stationary point as

$$J(p_x, p_y) = T_0 + A(p_x - p_{x0})^2 + 2B(p_x - p_{x0})(p_y - p_{y0}) + C(p_y - p_{y0})^2. \tag{6.17}$$

In the present case, because it is so simple, we know $T_0 = R/\alpha$, and we can even find analytic expressions for the Taylor series coefficients. They are

$$A = -\frac{|z|}{2\xi_0^3}\left(\frac{1}{v^2} - p_{y0}^2\right), B = -\frac{|z|}{2\xi_0^3} p_{x0} p_{y0}, C = -\frac{|z|}{2\xi_0^3}\left(\frac{1}{v^2} - p_{x0}^2\right).$$

But in the general case, using a table of values for $\tau(p_x, p_y)$, the six constants $T_0, A, B, C, p_{x0}, p_{y0}$ may be recovered numerically, using essentially the same method as given in eqs. (6.12) and (6.13) for a scalar horizontal slowness. The stationary phase approximation is

$$W(\boldsymbol{X}_0, \omega) = \frac{\mathrm{i}\omega\, \mathrm{e}^{\mathrm{i}\omega T_0}}{2\pi\xi_0} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp[\mathrm{i}\omega\{A(p_x - p_{x0})^2 + 2B(p_x - p_{x0})(p_y - p_{y0}) + C(p_y - p_{y0})^2\}]\, \mathrm{d}p_x\, \mathrm{d}p_y, \tag{6.18}$$

and after some straightforward work it can be shown that this double integral is merely

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp[\{\ldots\}]\, \mathrm{d}p_x\, \mathrm{d}p_y = -\frac{\mathrm{i}\pi}{\omega}\sqrt{\frac{1}{AC - B^2}}. \tag{6.19}$$

Substitution of eq. (6.19) into (6.18) gives an analytic approximate solution for $W(\boldsymbol{X}_0, \omega)$ that is easily taken into the time domain. In the present case, where explicit formulae for A, B, C are known, we can go on to show that $\sqrt{[1/(AC - B^2)]} = 2\xi_0^2 . v/|z|$, and since $\xi_0 . v/|z| = 1/R$ it turns out for the point source in a homogeneous medium that the stationary phase approximation returns the exact solution,

$$W(\boldsymbol{X}_0, \omega) = \frac{\exp\left(\mathrm{i}\omega\frac{R}{\alpha}\right)}{R}.$$

Having given elementary results in this section for a 1D medium, and for a point source regarded as an integral over two horizontal slownesses, eq. (6.16), we next give some more general results for the 3D problem.

6.5. A representation for generalized rays

For a generalized ray in a stack of layers with randomly dipping planar interfaces, we can write the response in the frequency domain as

$$W(\boldsymbol{X}_0,\omega) = \int\int \text{plane waves}$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(\omega).F(\boldsymbol{p}_\mathrm{A}).\prod(\boldsymbol{p}_\mathrm{A}).\exp\Big[\mathrm{i}\omega(\boldsymbol{p}_\mathrm{A}.\boldsymbol{X}_\mathrm{A} + \boldsymbol{p}_\mathrm{B}.\boldsymbol{X}_\mathrm{B} + \sum_{\text{ray}}\xi_j z_j)\Big].C(\boldsymbol{p}_\mathrm{B}).\,\mathrm{d}p_x\,\mathrm{d}p_y. \tag{6.20}$$

Here, $S(\omega)$ is the source spectrum; $F(\boldsymbol{p}_\mathrm{A})$ is the radiation pattern; $(\Pi\boldsymbol{p}_\mathrm{A})$ is the product of transmission coefficients appropriate to the generalized ray path; and $C(\boldsymbol{p}_\mathrm{B})$ is a receiver dependent factor that is appropriate for the particular response under analysis (e.g., the vertical displacement, or a particular strain component). The double integral is thought of as a decomposition of the source into its plane wave components, recombining these plane waves after their passage through the medium. It is necessary to find the "along-interface" slownesses to evaluate factors within $\Pi(\boldsymbol{p}_\mathrm{A})$, and also to find $\boldsymbol{p}_\mathrm{B}$ as a function of $\boldsymbol{p}_\mathrm{A}$. However, as shown above in eq. (6.4) – (6.9), this can be done systematically in a way that does not require ray tracing. The distances $\boldsymbol{X}_0, \boldsymbol{X}_\mathrm{A}, \boldsymbol{X}_\mathrm{B}$ in eq. (6.20) are independent of $\boldsymbol{p}_\mathrm{A}$.

We can show that the double integral in eq. (6.20) has a stationary integrand for the special value of slowness pA that does shoot a ray from A (the source) to B (the receiver). The demonstration requires a notation that distinguishes between the fixed value of $\boldsymbol{X}_\mathrm{B}$, and a functional relationship such as $\boldsymbol{X}_\mathrm{B}(\boldsymbol{p}_\mathrm{A})$ or $\boldsymbol{X}_\mathrm{B}(\boldsymbol{p}_\mathrm{B})$, giving the dependence of range on slowness. Let us loosely use $\boldsymbol{X}_\mathrm{B}(\boldsymbol{p})$ to refer to either of the latter functions.

Thus, defining

$$J = \boldsymbol{p}_A.\boldsymbol{X}_A + \boldsymbol{p}_B.\boldsymbol{X}_B + \sum\nolimits_{\text{ray}}\xi_j z_j(\boldsymbol{X}_\mathrm{A}\text{and } \boldsymbol{X}_\mathrm{B} \text{ fixed}), \tag{6.21}$$

and substituting for $\sum_{\text{ray}}\xi_j z_j$ from eq. (6.3), we have

$$J = T - \boldsymbol{p}_\mathrm{B}\cdot[\boldsymbol{X}_\mathrm{B}(\boldsymbol{p}) - \boldsymbol{X}_\mathrm{B}(\text{fixed})]. \tag{6.22}$$

Let us now make a small perturbation from $\boldsymbol{p}_\mathrm{A}$ to $\boldsymbol{p}_\mathrm{A}+\Delta\boldsymbol{p}_\mathrm{A}$, and consider the effect on J in eq. (6.22). Note that $\boldsymbol{X}_\mathrm{A}$ is still fixed, because this merely specifies the position (with respect to the source) at which "layer thicknesses" are defined. But the dependent variables $\boldsymbol{p}_\mathrm{B}$, T, $\boldsymbol{X}_\mathrm{B}$ and J become $\boldsymbol{p}_\mathrm{B}+\Delta\boldsymbol{p}_\mathrm{B}, T+\Delta T, \boldsymbol{X}_\mathrm{B}+\Delta\boldsymbol{X}_\mathrm{B}$, and $J+\Delta J$. From any of the figs. 40 through 44, it is clear that (to first order) $\Delta T = \boldsymbol{p}_\mathrm{B}.\Delta\boldsymbol{X}_\mathrm{B}$. It follows that

$$J + \Delta J = T + \Delta T - (\boldsymbol{p}_\mathrm{B} + \Delta\boldsymbol{p}_\mathrm{B}).[\boldsymbol{X}_\mathrm{B}(\boldsymbol{p}) + \Delta\mathrm{X}_\mathrm{B} - \boldsymbol{X}_\mathrm{B}(\text{fixed})]$$
$$= J - \Delta\mathrm{p}_\mathrm{B}.[\boldsymbol{X}_\mathrm{B}(\boldsymbol{p}) - \boldsymbol{X}_\mathrm{B}(\text{fixed})]. \tag{6.23}$$

Hence,

$$\Delta J = 0 \quad \text{if} \quad \boldsymbol{X}_\mathrm{B}(\boldsymbol{p}) = \boldsymbol{X}_\mathrm{B}(\text{fixed}), \tag{6.24}$$

(compare with eq. (4.17)) and so the integrand of (6.20) has stationary phase, for the slowness value that shoots a ray to the actual position of the receiver.

Parts of this analysis are similar to ideas contained in Frazer and Phinney (1980) and Burdick and Salvado (1986), notably our eq. (6.22). However, a key part is different in that the concept of a sum over vertical slownesses is retained even for a 3D varying medium. The phase within eq. (6.20) can easily and exactly be computed via (6.21), which in turn has an interpretation in terms of rays and stationary points, via (6.22). Much can therefore be retained, of the decades of experience in understanding waves in plane parallel layering.

6.6. A simple example

In order to verify at the simplest level one of the most elementary uses of the method given above, and to suggest further improvements, we can compare the method with exact ray theory for the reflections from a single dipping layer. We compare reflection travel times, horizontal slownesses, and geometrical spreading for the situation shown in fig. 48 (500 m source - receiver distances), and fig. 49 (2500 m). These calculations were done for the same discretized "look-up" table for $\sum_{\text{ray}}\xi_j z_j$, using the range ± 0.2 s/km with 40 points for each of the two cartesian components of $\boldsymbol{p}_\mathrm{A}$. We see from fig. 48 that travel time errors begin to show in the fifth significant digit; horizontal slowness errors in the third significant digit; and geometrical spreading errors in the second significant digit. Errors are seen to be somewhat larger for fig. 49, though still insignificant for most practical purposes.

Our numerical method is based on finding the six unknown constants in eq. (6.17). For figs. 48 and 49 we have taken a primitive approach, merely

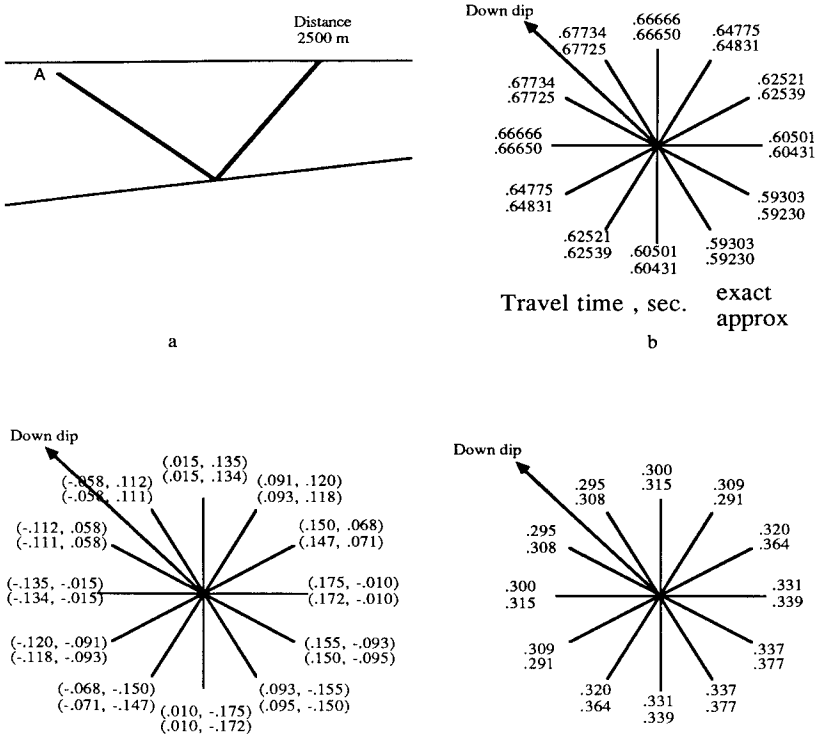


Fig. 49. Computation for the reflection from a dipping layer. As in fig. 48, except that the receivers are 2500 m horizontally from the source. For a horizontal reflector at the same depth below the source, travel time would be 0.63906 s; geometrical spreading would be 0.313 km^{-1}.

tions. Such problems are suitable for further study in light of the methods of this section 1.6. However, for broad classes of media involving planar bur non-parallel interfaces, our method is directly applicable.

Acknowledgements

I thank Göran Ekström for supplying fig. 4, John Orcutt and Michael Hedlin for supplying fig. 37, and Dean Witte for help with computations behind figs. 48 and 49. I also thank Albert Tarantola and Claude Mercier for much help with preparation of the printed version of these lectures.

These lectures describe new work in the theory of wave propagation, that was supported by NSF Grant EAR 87-08564.

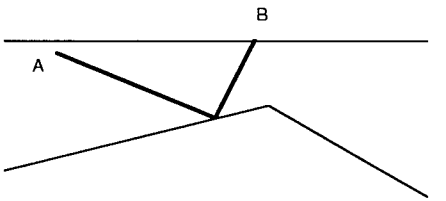


Fig. 50. A ray departing from A is reflected from a locally plane interface, and received at B. But this reflection would not be expected to be present at much greater distances, since the reflector would be missed by rays departing nearly horizontally from A.

References

Adachi, R., Kumamoto J. Sci., Ser. A., 2, 18-23, 1954.

Aki, K., and P.G. Richards, Quantitative Seismology – Theory and Methods, 2 volumes, W.H. Freeman and Co., San Francisco, 1980.

Azimi, Sh.A., A.V. Kalinin, V.V. Kalinin, and B.L. Pivovarov, Izvestiya, Physics of the Solid Earth, 88-93, February 1968.

Borcherdt, R.D., G. Glassmoyer, and L. Wennerberg, J. Geophys. Res., 91, 11503-11518, 1986.

Burdick, L.J., Geoph. J. Roy. astr. Soc., 80, 33-55, 1985.

Burdick, L.J., and C.A. Salvado, J. Geophys. Res., 91, 12482–12496, 1986.

Chapman, C.H., Geophys. Res. Let., 3, 153–156, 1976.

Chapman, C.H., and J.A. Orcutt, Rev. Geophys., 23, 105–163, 1985.

Choy, G.L., and J. Boatwright, Bull. Seism. Soc. Amer., 71, 691–711, 1981.

Choy, G., and V.F. Cormier, J. Geophys. Res., 91, 7326 7342, 1986.

Dahlen, F.A., and R.I. Odom, Geophys., 49, 478–480, 1984.

Diebold, J.D., Geophys., 52, 1492–1500,1987.

Dziewonski, A.M., and D.L. Anderson, Phys. Earth Plan. Int., 25, 297–356, 1981.

Eisner, E., Geophys., 48, 1132 1134, 1983.

Frazer, L.N., and R.A. Phinney, Geoph. J. Roy. astr. Soc., 63, 691–717, 1980.

Gladwin, M.T., and F.D. Stacey, Phys. Earth Plan. Int., 8, 332–336, 1974.

Helmberger, D.L., Bull. Seism. Soc. Amer., 58, 179–214, 1968.

Hong, T.-L., and D.L. Helmberger, Bull. Seism. Soc. Amer., 67, 995 1008, 1977.

Johnson, S.J., Geophys., 41, 418–424, 1976.

Langston, C.A., and D.E. Blum, Bull. Seism. Soc. Amer., 67, 693–711, 1977.
McLaughlin, K.L., and L.M. Anderson, The Vela Program; a 25 year Review of Basic Research, ed. A.U. Kerr, Defense Advanced Research Projects Agency, Washington, DC, 433–442, 1985.
Müller, G., J. Geophys., 85, 153–174, 1985.
O'Connell, R.J., and B. Budiansky, Geophys. Res. Let., 5, 5–8, 1978.
O'Neill, M.E., and D.P. Hill, Bull. Seism. Soc. Amer., 69, 17–25, 1979.
Rial, J.A., and V.F. Cormier, J. Geophys. Res., 85, 2661–2668, 1981.
Richards, P.G., Geophys., 36, 798–809, 1971.
Richards, P.G., Geoph. J. Roy. astr. Soc., 35, 243–264, 1973.
Richards, P.G., Bull. Seism. Soc. Amer., 74, 2157–2165, 1984.
Richards, P.G., Bull. Seism. Soc. Amer., 66, 701–717, 1976.
Richards, P.G., Seismic wave propagation effects – development of theory and numerical modelling, Chapter in The Vela Program; a 25 year Review of Basic Research, ed. A.U. Kerr, Defense Advanced Research Projects Agency, Washington, DC, 183–251, 1985.
Richards, P.G., and W. Menke, Bull. Seism. Soc. Amer., 73, 1005–1021, 1983.
Spencer, T.W., Geophys., 25, 625–641, 1960.
Spudich, P., and J.A. Orcutt, J. Geophys. Res., 85, 1409–1434, 1980.
Strick, E., Geophys., 35, 387–403, 1970.
Wiggins, R.A., Geoph. J. Roy. astr. Soc., 46, 1–10, 1976.
Woodhouse, J.H., and A.M. Dziewonski, J. Geophys. Res., 89, 5953–5986, 1984.