

2

Seismic Wave Propagation and Earth models

Peter Bormann, Bob Engdahl and Rainer Kind

2.1 Introduction

The key data to be recorded by means of *seismic sensors* (Chapter 5) and *recorders* (Chapter 6) at seismological observatories (*stations* – Chapter 7, *networks* – Chapter 8, *arrays* – Chapter 9) are *seismic waves*, radiated by *seismic sources* (Chapter 3). Weak signals may be masked or significantly distorted by *seismic noise* (Chapter 4), which is usually considered disturbing and unwanted. Only in some special engineering-seismological applications is seismic noise also appreciated as a useful signal, from which some information on the structure, velocity and fundamental resonance frequency of the uppermost sedimentary layers can be derived (e.g. Bard, 1999). But most of what we know today of the structure and physical properties of our planet Earth, from its uppermost crust down to its center, results from the analysis of seismic waves generated by more or less localized natural or man-made sources such as earthquakes or explosions (Figs. 3.1 to 3.4). Either (repeatedly) solving the so-called forward (direct) or the inverse problem of data analysis (Fig. 1.1) achieves this.

It is not the task of the New Manual of Seismological Observatory Practice (NMSOP), to provide an in-depth understanding of the theoretical tools for this kind of analysis. There exist quite a number of good introductory (Lillie, 1999; Shearer, 1999) and more advanced textbooks (e.g., Aki and Richards, 1980 and 2002; Ben-Menahem and Singh, 1981; Bullen and Bolt, 1985; Dahlen and Tromp, 1998; Lay and Wallace, 1995; Kennett, 2001), and a variety of special papers and monographs related to specific methods (e.g. Fuchs and Müller, 1971; Červený et al., 1977; Kennett, 1983; Müller, 1985; Červený, 2001), types of seismic waves (e.g., Malischewsky, 1987; Lapwood and Usami, 1981) or applications (e.g., Gilbert and Dziewonski, 1975; Sherif and Geldart, 1995). Rather, we will take here a more phenomenological approach and refer to related fundamentals in physics and mathematical theory only as far as they are indispensable for understanding the most essential features of seismic waves and their appearance in seismic records and as far as they are required for:

- identifying and discriminating the various types of seismic waves;
- understanding how the onset-times of these phases, as observed at different distances from the source, form so-called travel-time curves;
- understanding how these curves and some of their characteristic features are related to the velocity-structure of the Earth and to the observed (relative) amplitudes of these phases in seismic records;
- using travel-time and amplitude-distance curves for seismic source location and magnitude estimation;
- understanding how much these source-parameter estimates depend on the precision and accuracy of the commonly used 1-D Earth models (see IS 11.1);

2. Seismic Wave Propagation and Earth models

- appreciating how these source parameter estimates can be improved by using more realistic (2-D, 3-D) Earth models as well as later (secondary) phase onsets in the processing routines; and
- being aware of the common assumptions and simplifications used in synthetic seismogram calculations that are increasingly used nowadays in seismological routine practice (see 2.5.4.4, 2.8, 3.5.3).

2.2 Elastic moduli and body waves

2.2.1 Elastic moduli

Seismic waves are elastic waves. Earth material must behave elastically to transmit them. The degree of elasticity determines how well they are transmitted. By the pressure front expanding from an underground explosion, or by an earthquake shear rupture, the surrounding Earth material is subjected to *stress* (compression, tension and/or shearing). As a consequence, it undergoes *strain*, i.e., it changes volume and/or distorts shape. In an inelastic (plastic, ductile) material this deformation remains while elastic behavior means that the material returns to its original volume and shape when the stress load is over.

The degree of elasticity/plasticity of real Earth material depends mainly on the *strain rate*, i.e., on the length of time it takes to achieve a certain amount of distortion. At very low strain rates, such as movements in the order of mm or cm/year, it may behave ductilely. Examples are the formation of geologic folds or the slow plastic convective currents of the hot material in the Earth's mantle with velocity on the order of several cm per year. On the other hand, the Earth reacts elastically to the small but rapid deformations caused by a transient seismic source pulse. Only for very large amplitude seismic deformations in soft soil (e.g., from earthquake strong-motions in the order of 40% or more of the gravity acceleration of the Earth) or for extremely long-period free-oscillation modes (see 2.4) does the inelastic behavior of seismic waves have to be taken into account.

Within its elastic range the behavior of the Earth material can be described by *Hooke's Law* that states that the amount of strain is linearly proportional to the amount of stress. Beyond its elastic limit the material may either respond with brittle fracturing (e.g., earthquake faulting, see Chapter 3) or ductile behavior/plastic flow (Fig. 2.1).

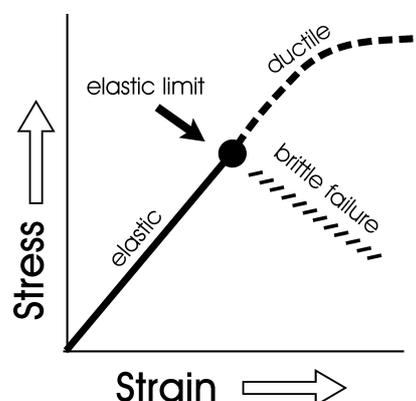


Fig. 2.1 Schematic presentation of the relationship between stress and strain.

Elastic material resists differently to stress depending on the type of deformation. It can be quantified by various elastic moduli:

- the *bulk modulus* κ is defined as the ratio of the hydrostatic (homogeneous all-sides) pressure change to the resulting relative volume change, i.e., $\kappa = \Delta P / (\Delta V/V)$, which is a measure of the *incompressibility* of the material (see Fig. 2.2 top);
- the *shear modulus* μ (or “rigidity”) is a measure of the resistance of the material to shearing, i.e., to changing the shape and not the volume of the material. Its value is given by half of the ratio between the applied shear stress τ_{xy} (or tangential force ΔF divided by the area A over which the force is applied) and the resulting shear strain e_{xy} (or the shear displacement ΔL divided by the length L of the area acted upon by ΔF), that is $\mu = \tau_{xy}/2 e_{xy}$ or $\mu = (\Delta F/A) / (\Delta L/L)$ (Fig. 2.2 middle). For fluids $\mu = 0$, and for material of very strong resistance (i.e. $\Delta L \rightarrow 0$) $\mu \rightarrow \infty$;
- the *Young’s modulus* E (or “stretch modulus”) describes the behavior of a cylinder of length L that is pulled on both ends. Its value is given by the ratio between the extensional stress to the resulting extensional strain of the cylinder, i.e., $E = (F/A) / (\Delta L/L)$ (Fig. 2.2 bottom);
- the *Poisson’s ratio* σ is the ratio between the lateral contraction (relative change of width W) of a cylinder being pulled on its ends to its relative longitudinal extension, i.e., $\sigma = (\Delta W/W) / (\Delta L/L)$ (Fig. 2.2 bottom).

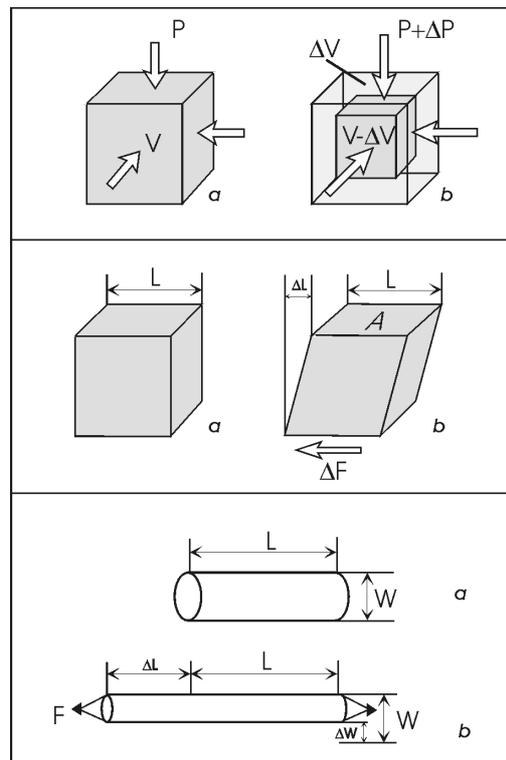


Fig. 2.2 Deformation of material samples for determining elastic moduli. Top: bulk modulus κ ; middle: shear modulus μ ; bottom: Young’s modulus E and Poisson’s ratio σ . a – original shape of the volume to be deformed; b – volume and/or shape after adding pressure ΔP to the volume V (top), shear force ΔF over the area A (middle) or stretching force F in the direction of the long axis of the bar (bottom).

2. Seismic Wave Propagation and Earth models

Young's modulus, the bulk modulus and the shear modulus all have the same physical units as pressure and stress, namely (in international standard (SI) units):

$$1 \text{ Pa} = 1 \text{ N m}^{-2} = 1 \text{ kg m}^{-1} \text{ s}^{-2} \quad (\text{with } 1 \text{ N} = 1 \text{ Newton} = 1 \text{ kg m s}^{-2}). \quad (2.1)$$

2.2.2 Stress-strain relationship

The most general linear relationship between stress and strain of an elastic medium is governed in the generalized *Hook's law* (see Equation (10) in the IS 3.1) by a fourth order parameter tensor. It contains 21 independent moduli. The properties of such a solid may vary with direction. Then the medium is called *anisotropic*. Otherwise, if the properties are the same in all directions, a medium is termed *isotropic*. Although in some parts of the Earth's interior anisotropy on the order of a few percent exists, isotropy has proven to be a reasonable first-order approximation for the Earth as a whole. The most common models, on which data processing in routine observatory practice is based, assume isotropy and changes of properties only with depth.

In the case of isotropy the number of independent parameters in the elastic tensor reduces to just two. They are called after the French physicist *Lamé* (1795-1870) the *Lamé parameters* λ and μ . The latter is identical with the *shear* modulus. λ does not have a straightforward physical explanation but it can be expressed in terms of the above mentioned elastic moduli and Poisson's ratio, namely

$$\lambda = \kappa - 2\mu/3 = \frac{\sigma E}{(1 + \sigma)(1 - 2\sigma)}. \quad (2.2)$$

The other elastic parameters can also be expressed as functions of μ , λ and/or κ :

$$E = \frac{(3\lambda + 2\mu)\mu}{(\lambda + \mu)} \quad (2.3)$$

and

$$\sigma = \frac{\lambda}{2(\lambda + \mu)} = \frac{3\kappa - 2\mu}{2(3\kappa + \mu)}. \quad (2.4)$$

For a *Poisson solid* $\lambda = \mu$ and thus, according to (2.4), $\sigma = 0.25$. Most crustal rocks have a Poisson's ratio between about 0.2 and 0.3. But σ may reach values of almost 0.5, e.g., for unconsolidated, water-saturated sediments, and even negative values of σ are possible (see Tab. 2.1).

The elastic parameters govern the velocity with which seismic waves propagate. The *equation of motion* for a continuum can be written as

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \partial_j \tau_{ij} + \mathbf{f}_i, \quad (2.5)$$

with ρ - density of the material, u_i – displacement, τ_{ij} – stress tensor and f_i – the body force term that generally consists of a *gravity term* and a *source term*. The gravity term is important at low frequencies in *normal mode seismology* (see 2.4), but it can be neglected for calculations of body- and surface-wave propagation at typically observed wavelengths. Solutions of Eq. (2.5) which predict the ground motion at locations some distance away from the source are called *synthetic seismograms* (see Figs. 2.54 and 2.55).

In the case of an inhomogeneous medium, which involves gradients in the Lamé parameters, Eq. (2.5) takes a rather complicated form that is difficult to solve efficiently. Also, in case of strong inhomogeneities, transverse and longitudinal waves (see below) are not decoupled. This results in complicated particle motions. Therefore, most methods for synthetic seismogram computations ignore gradient terms of λ and μ in the equation of motion by modeling the material either as a series of homogeneous layers (which also allows to approximate gradient zones; see *reflectivity method* by Fuchs and Müller, 1971; Kennett, 1983; Müller, 1985) or by assuming that variations in the Lamé parameters are negligible over a wavelength Λ and thus these terms tend to zero at high frequencies (*ray theoretical approach*; e.g., Červený et al., 1977; Červený, 2001). In homogeneous media and for small deformations the *equation of motion* for seismic waves outside the source region (i.e., without the source term f_s and neglecting the gravity term f_g) takes the following simple form:

$$\rho \ddot{\mathbf{u}} = (\lambda + 2\mu)\nabla\nabla\cdot\mathbf{u} - \mu\nabla\times\nabla\times\mathbf{u} \quad (2.6)$$

where \mathbf{u} is the displacement vector and $\ddot{\mathbf{u}}$ its second time derivative. Eq. (2.6) provides the basis for most body-wave, synthetic seismogram calculations. Although it describes rather well most basic features in a seismic record we have to be aware that it is an approximation only for an isotropic homogeneous linearly elastic medium.

2.2.3 P- and S-wave velocities, waveforms and polarization

The first term on the right side of Eq. (2.6) contains the scalar product $\nabla\cdot\mathbf{u} = \text{div } \mathbf{u}$. It describes a volume change (compression and dilatation), which always contains some (rotation free!) shearing too, unless the medium is compressed hydrostatically (as in Fig. 2.2 top). The second term is a vector product ($\text{rot } \mathbf{u} = \nabla\times\mathbf{u}$) corresponding to a curl (rotation) and describes a change of shape without volume change (pure shearing). Generally, every vector field, such as the displacement field \mathbf{u} , can be decomposed into a rotation-free (\mathbf{u}^r) and a divergence-free (\mathbf{u}^d) part, i.e., we can write $\mathbf{u} = \mathbf{u}^r + \mathbf{u}^d$. Since the divergence of a curl and the rotation of a divergence are zero, we get accordingly two independent solutions for Eq. (2.6) when forming its scalar product $\nabla\cdot\mathbf{u}$ and vector product $\nabla\times\mathbf{u}$, respectively:

$$\frac{\partial^2(\nabla\cdot\mathbf{u})}{\partial^2 t} = \frac{\lambda + 2\mu}{\rho} \nabla^2(\nabla\cdot\mathbf{u}^r) \quad (2.7)$$

and

$$\frac{\partial^2(\nabla\times\mathbf{u})}{\partial^2 t} = \frac{\mu}{\rho} \nabla^2(\nabla\times\mathbf{u}^d). \quad (2.8)$$

2. Seismic Wave Propagation and Earth models

Eqs. (2.7) and (2.8) are solutions of the wave equation for the propagation of two independent types of seismic *body waves*, namely *longitudinal (compressional - dilatational) P waves* and *transverse (shear) S waves*. Their velocities are

$$v_p = \sqrt{\frac{\lambda + 2\mu}{\rho}} = \sqrt{\frac{\kappa + 4\mu/3}{\rho}} \quad (2.9)$$

and

$$v_s = \sqrt{\frac{\mu}{\rho}}. \quad (2.10)$$

Accordingly, for a Poisson solid with $\lambda = \mu$ $v_p/v_s = \sqrt{3}$. This comes close to the v_p/v_s ratio of consolidated sedimentary and igneous rocks in the Earth's crust (see Tab. 2.1). Eqs. (2.9) and 2.10) also mean that P (*primary*) waves travel significantly faster than S (*secondary*) waves and thus arrive ahead of S in a seismic record (see Fig. 2.3). The Poisson's ratio is often used as a measure of the v_p/v_s ratio, namely

$$\sigma = (v_p^2/v_s^2 - 2)/2(v_p^2/v_s^2 - 1) \quad (2.11)$$

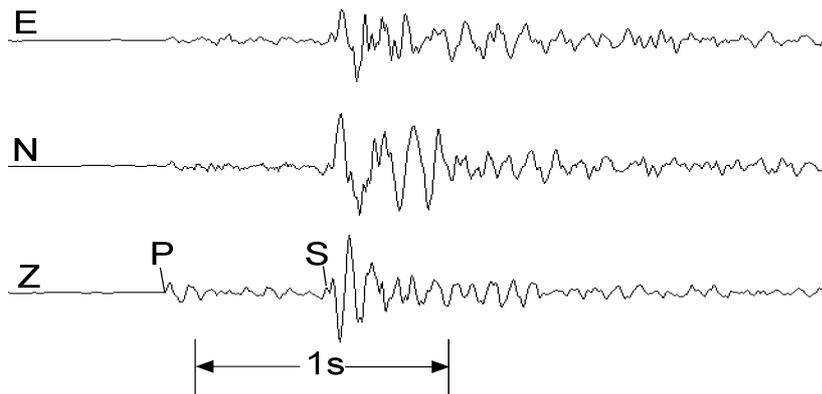


Fig. 2.3 The three components of ground-velocity proportional digital records of the P and S waves from a local event, an aftershock of the Killari-Latur earthquake, India (18.10.1993), at a hypocentral distance of about 5.3 km.

Note the simple *transient waveform (wavelet)* of P in the Z-component of Fig. 2.3. The waveform and duration of the primary body wave is related to the shape and duration of the source-time function. It is for an earthquake shear rupture usually a more or less complex displacement step (see Figs. 2.4 and 3.4) which can be described by the moment-release function $M(t)$ (see 3.5). In the *far-field*, i.e., at distances larger than the source dimension and several wavelengths of the considered signal, the related displacement $u(t)$ looks, in the idealized case, bell-shaped and identical with the moment-rate $\dot{M}(t)$ (or velocity source-time) function (see Fig. 2.4 middle). The base-width of this *far-field displacement source pulse* $u(t)$ corresponds to the duration of displacement at the source (for examples see Fig. 3.7). However, usually broadband seismometers record ground velocity $\dot{u}(t)$ instead of ground displacement. The recorded waveform then looks similar to the ones seen in Fig. 2.3 and Fig.

2.4 bottom. The period of the wavelet $\dot{u}(t)$ corresponds to the duration of the displacement of the source, τ_s . This waveform of primary body waves will be slightly changed due to frequency-dependent attenuation and other wave-propagation effects, e.g., those that cause phase shifts. But the duration of the body-wave ground-motion wavelet (or wave-group) will remain essentially that of the source process, independent of the observational distance, unless it is significantly prolonged and distorted by narrowband seismic recordings (see 4.2). We have made this point in order to better appreciate one of the principal differences in the appearance in seismic records of transient body waves on the one hand and of dispersed surface waves (see 2.3 and, e.g., Figs. 2.14 and 2.23) on the other hand.

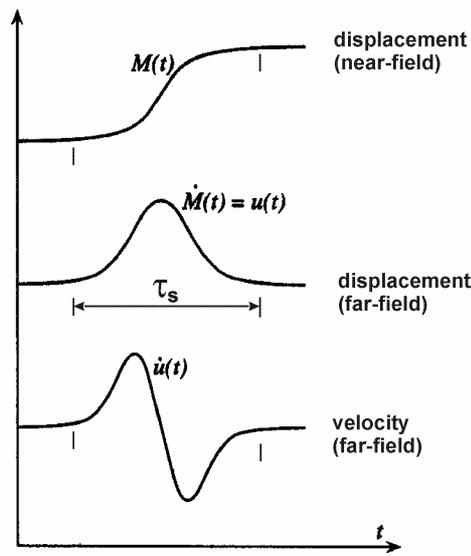


Fig. 2.4 Relationship between near-field displacement, far-field displacement and velocity from isotropic or double-couple source earthquake shear sources (modified from Shearer, Introduction to Seismology, 1999; with permission from Cambridge University Press).

Tab. 2.1 gives some approximate average values for the elastic moduli κ and μ , the density ρ and the seismic velocities v_p and v_s for air, water, ice and some selected Earth materials. The following general conclusions can be drawn from it:

- For the same material, shear waves travel always slower than compressional waves;
- The higher the rigidity of the material, the higher the P- and S-wave velocities;
- The rigidity usually increases with density ρ , but more rapidly than ρ . This explains why denser rocks have normally faster wave propagation velocities although $v^2 \sim 1/\rho$;
- Fluids (liquids or gasses) have no shear strength ($\mu = 0$) and thus do not propagate shear waves;
- For the same material, compressional waves travel slower through its liquid state than through its solid state (e.g., water and ice, or, in the Earth's core, through the liquid outer and solid inner iron core, respectively).

Seismic energy is usually radiated from localized sources with linear dimensions much smaller than the distance of observation. Therefore, seismic “wavefronts” from such “point sources,” i.e., the surfaces along which the propagating waves are oscillating *in phase*, are generally curved and the “seismic rays,” perpendicular to the wavefronts, are oriented in the

2. Seismic Wave Propagation and Earth models

radial directions of wave propagation. However, when the distance is large enough, the curvature of the wavefronts becomes so small that we can approximate them locally (e.g., within the aperture of a local seismic network or an array; see Chapters 8 and 9) by *plane waves* with parallel seismic rays.

Tab. 2.1 Typical values (averages and/or approximate ranges) of elastic constants, density, Poisson's ratio and seismic wave velocities for some selected materials, unconsolidated sediments, sedimentary rocks of different geologic age and igneous/plutonic rocks. Values for granite relate to 200 MPa confining pressure, corresponding to about 8 km depth, for basalt to 600 MPa (about 20 km depth), and for Peridotite, Dunite and Pyroxenite to 1000 MPa (about 30 km depth) (compiled from Hellwege, 1982; Lillie, 1999; and other sources).

Material or Geologic Formation	Bulk Modulus in 10^9 Pa	Shear Modulus in 10^9 Pa	Density in kg m^{-3}	Poisson Ratio	v_p in km s^{-1}	v_s in km s^{-1}	v_p/v_s
Air	0.0001	0	1.0	0.5	0.32	0	∞
Water	2.2	0	1000	0.5	1.5	0	∞
Ice	3.0	4.9	920	-0.034	3.2	2.3	1.39
Clastic sedimentary rocks					(1.4-5.3)		
Sandstone	24	17	2500	0.21	4.3	2.6	1.65
Salt	24	18	2200	0.17	4.6 (3.8-5.2)	2.9	1.59
Limestone	38	22	2700	0.19	4.7 (2.9-5.6)	2.9	1.62
Granite	56 (47-69)	34 (30-37)	2610 (2340-2670)	0.25 (0.20-0.31)	6.2 (5.8-6.4)	3.6 (3.4-3.7)	1.73 (1.65-1.91)
Basalt	71 (64-80)	38 (33-41)	2940 (2850-3050)	0.28 (0.26-0.29)	6.4 (6.1-6.7)	3.6 (3.4-3.7)	1.80 (1.76-1.82)
Peridotite, Dunit, Pyroxenite	128 (113-141)	63 (52-72)	3300 (3190-3365)	0.29 (0.26-0.29)	8.0 (7.5-8.4)	4.4 (4.0-4.7)	1.8 (1.76-1.91)
Metamorphic & igneous rocks					(3.8-6.4)		
Ultramafic rocks					(7.2-8.7)		
Cenozoic			1500-2100	0.38-<0.5	(0.2-1.9)		2.3 - 8
Cenozoic water saturated			1950	0.48	1.7	0.34	5
Cretaceous & Jurassic			2400-2500	0.28-0.43			1.8 - 2.8
Triassic			2500-2700	0.28-0.40			1.8 - 2.5
Upper Permian			2000-2900	0.23-0.31			1.7 - 1.9
Carboniferous				0.31-0.35			1.9 - 2.1

Fig. 2.5 depicts (exaggerated) the kind of displacements occurring from harmonic plane P and S waves. One clearly recognizes that P waves involve both a volume change and shearing (change in shape) while S-wave propagation is pure shear with no volume change. The P-wave particle motion is back and forth in the radial (R) direction of wave propagation

(*longitudinal polarization*) but that of the S wave is perpendicular (*transverse*) to it, in the given case oscillating up and down in the vertical plane (SV-wave). However, S waves may also oscillate purely in the horizontal plane (SH waves) or at any angle between vertical and horizontal, depending on the source mechanism (Chapter 3), the wave propagation history, and the incidence angle i_0 at the seismic station (see Fig. 2.27).

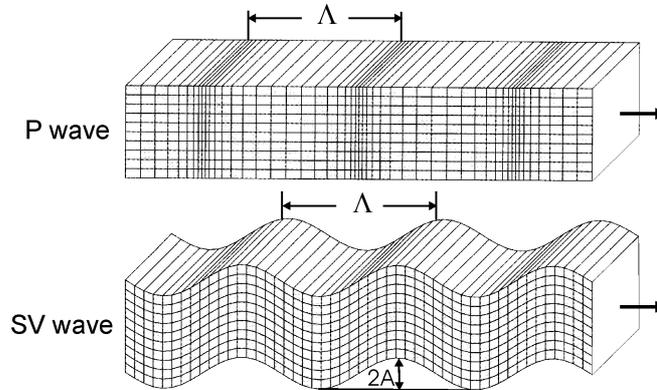


Fig. 2.5 Displacements from a harmonic plane P wave (top) and SV wave (bottom) propagating in a homogeneous isotropic medium. Λ is the wavelength. $2A$ means double amplitude. The white surface on the right is a segment of the propagating plane wavefront where all particles undergo the same motion at a given instant in time, i.e., they oscillate *in phase*. The arrows indicate the seismic rays, defined as the *normal* to the wavefront, which points in the direction of propagation (modified according to Shearer, Introduction to Seismology, 1999; with permission from Cambridge University Press).

The wavelength Λ is defined by the distance (in km) between neighboring wave peaks or troughs or volumes of maximum compression or dilatation (see Fig. 2.5). The wave period T is the duration of one oscillation (in s) and the frequency f is the number of oscillations per second (unit [Hz] = [s⁻¹]). The wavelength is the product of wave velocity v and period T while the wavenumber is the ratio $2\pi/\Lambda$. Tab. 2.2 summarizes all these various *harmonic wave parameters* and their mutual relationship.

Tab. 2.2 Harmonic wave parameters and their mutual relationship.

Name	Symbol	Relationships
Period	T	$T = 1/f = 2\pi/\omega = \Lambda/v$
Frequency	f	$f = 1/T = \omega/2\pi = v/\Lambda$
Angular frequency	ω	$\omega = 2\pi f = 2\pi/T = v \cdot k$
Velocity	v	$v = \Lambda/T = f \cdot \Lambda = \omega/k$
Wavelength	Λ	$\Lambda = v/f = v \cdot T = 2\pi/k$
Wavenumber	k	$k = \omega/v = 2\pi/\Lambda = 2\pi f/v$

In any case, the polarization of both P and S waves, when propagating in a homogenous and isotropic medium, is linear. This is confirmed rather well by particle motion analysis of real seismic recordings, if they are broadband (or long period). But higher frequencies, which are more strongly affected by local inhomogeneities in the Earth, show a more elliptical or irregular particle motion. Fig. 2.6 shows an example. While the rectilinearity of P is almost 1

2. Seismic Wave Propagation and Earth models

(0.95) in the BB record it is significantly less (0.82 as an average over 5 oscillations and down to 0.68 for some single oscillations) for the short-period filtered record.

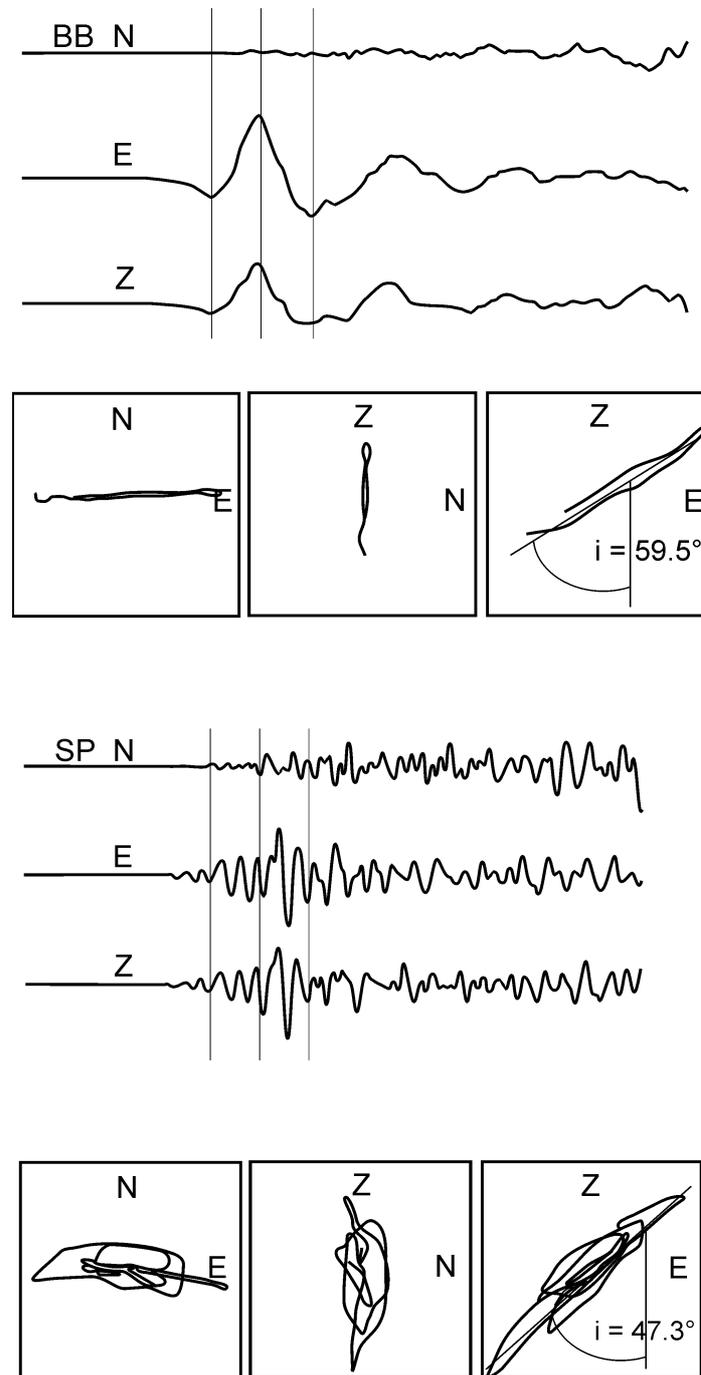


Fig. 2.6 3-component records at station MOX (top traces) and related plots of particle motion in the horizontal (N-E) plane and two vertical planes (Z-N and Z-E, respectively) of the P-wave onset from a local seismic event (mining collapse) in Germany (13.03.1989; $M_I = 5.5$; epicentral distance $D = 112$ km, backazimuth $BAZ = 273^\circ$). Upper part: broadband recording (0.1-5 Hz); lower part: filtered short-period recording (1- 5 Hz). **Note:** The incidence angle is 59.5° for the long-period P-wave oscillation and 47.3° for the high-frequency P-wave group.

S waves are also linearly polarized when propagating in homogeneous isotropic medium. However, in the presence of anisotropy, they split into a fast and slow component. These split waves propagate with different velocity that causes some time delay and related phase shift. Accordingly, the two split S-wave components superimpose to an elliptical polarization (Fig. 2.7). The orientation of the main axis and the degree of ellipticity are controlled by the fast and slow velocity directions of the medium with respect to the direction of wave propagation and the degree of anisotropy. Therefore, shear-wave splitting is often used to study S-wave velocity anisotropy in the Earth.

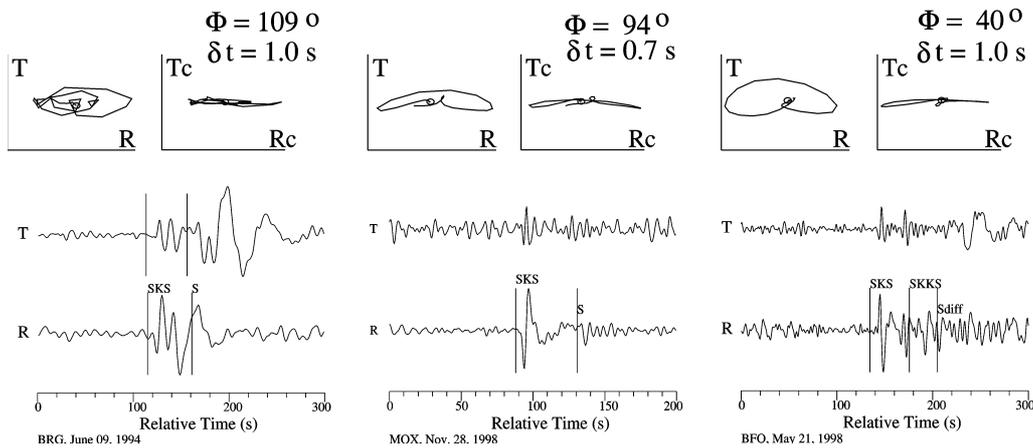


Fig. 2.7 Examples of SKS and SKKS recordings and plots of particle motion at three stations of the German Regional Seismograph Network. The horizontal radial (R) and transverse (T) components are shown. They were derived by rotation of the N-S and E-W horizontal components with the backazimuth angle. The T component at BFO has the same scale as the R component, while T is magnified two-fold relative to R at BRG and MOX. The top panels show the polarization in the R-T plane. Anisotropy is manifested in all three cases by the elliptical polarization. Linear polarization is obtained by correcting the R-T seismograms for the anisotropy effect using an anisotropy model where the direction of the fast shear wave is sub-horizontal and given by the angle Φ measured clockwise from north, and the delay time (in seconds) between the slow and the fast shear wave is given by δt (courtesy of G. Bock).

2.3 Surface waves

2.3.1 Origin

So far we have considered only body-wave solutions of the seismic wave equation. They exist in the elastic full space. However, in the presence of a free surface, as in the case of the Earth, other solutions are possible. They are called *surface waves*. There exist two types of surface waves, *Love waves* and *Rayleigh waves*. While Rayleigh (LR or R) waves exist at any free surface, Love (LQ or G) waves require some kind of a *wave guide* formed by a velocity increase with depth (gradient- or layer-wise). Both conditions are fulfilled in the real Earth.

SH waves are totally reflected at the free surface. Love waves are formed through constructive interference of repeated reflections of teleseismic SH at the free-surface (i.e., S3, S4, S5, etc.; see Fig. 2.42 and overlay to Figs. 2.48 and 2.49). They can also result from constructive interference between SH waves, which are *postcritically reflected* (see 2.5.3.5) within a homogeneous layer (or a set of i layers with increasing v_{si}) overlaying a half-space

2. Seismic Wave Propagation and Earth models

with higher velocity. The latter is the case of crustal layers, overlaying the upper mantle with a significant velocity increase at the base of the crust, called the “Mohorovičić-discontinuity” or *Moho* for short. The Moho marks the transition between the usually more mafic (often termed “basaltic”) lower crust and the peridotitic uppermost mantle (for related velocities see Tab. 2.1) and may, together with other pronounced intra-crustal velocity discontinuities give rise to the formation of complex guided crustal waves (see 2.3.3).

Generally, destructive interference of the upgoing and downgoing reflected SH waves will occur, except at certain discrete frequencies ω and n multiples of it (with n as an integer). The values of ω given for $n = 0$ are termed the *fundamental modes* while larger values of n define the *higher modes* or *overtones*. Fig. 2.8 (top) shows the horizontal (SH type) of displacement and linear polarization of the fundamental Love-wave mode as well as the exponential decay of its amplitudes with depth.

When a P (or SV) wave arrives at the surface the reflected wave energy contains (because of *mode conversion*, see 2.5.3.4) both P and SV waves. Lord Rayleigh showed more than 100 years ago that in this case a solution of the wave equation exists for two coupled inhomogeneous P and SV waves that propagate along the surface of a half-space. While Rayleigh waves show no dispersion in a homogeneous half-space, they are always dispersed in media with layering and/or velocity gradients such as in the real Earth. Rayleigh waves travel - for a Poisson solid - with a phase velocity $c = \sqrt{2 - 2/\sqrt{3}} v_s \approx 0.92 v_s$, i.e., they are slightly slower than Love waves. Therefore, they appear somewhat later in seismic records. The exact value of c depends on v_p and v_s . Since Rayleigh waves originate from coupled P and SV waves they are polarized in the vertical (SV) plane of propagation and due to the phase shift between P and SV the sense of their *particle motion* at the surface is *elliptical* and *retrograde* (counter clockwise). Fig. 2.8 (bottom) shows schematically the displacements for the fundamental mode of Rayleigh waves. They also decay exponentially with depth. The short-period fundamental mode of Rayleigh type in continental areas is termed Rg.

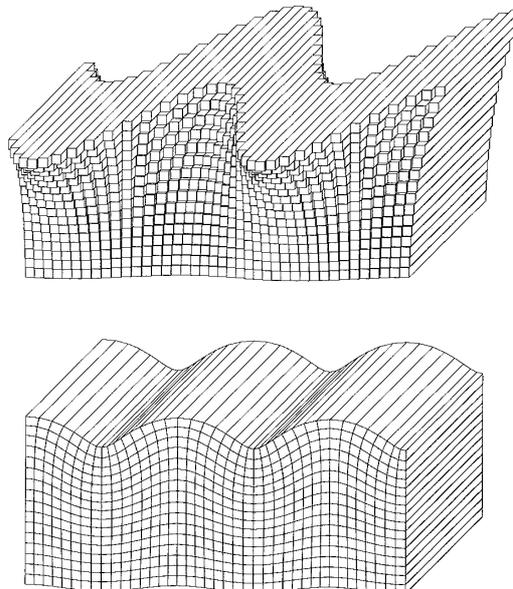


Fig. 2.8 Displacements caused by horizontally propagating fundamental Love (top) and Rayleigh waves (bottom). In both cases the wave amplitudes decay strongly with depth (from Shearer, Introduction to Seismology, 1999; with permission from Cambridge University Press).

2.3.2 Dispersion and polarization

The *penetration depth* below the surface increases with Λ . This is comparable with the frequency-dependent *skin effect* of electromagnetic waves propagating in a conducting medium with a free surface. Since the types of rocks, their rigidity and bulk modulus change (usually increase) with depth, the velocities of surface waves change accordingly since the longer waves “sense” deeper parts of the Earth. This results in a frequency dependence of their horizontal propagation velocity, called *dispersion*. Accordingly, while body-wave arrivals with no or negligibly small dispersion only (due to intrinsic attenuation) appear in seismic records as rather impulsive onsets or short transient wavelets (with the shape and duration depending on the bandwidth of the seismograph; see Chapter 4.2), the dispersion of surface waves forms long oscillating wave trains. Their duration increases with distance.

Usually, the more long-period surface waves arrive first (normal dispersion). But in some regions of the Earth low-velocity layers exist (e.g., the *asthenosphere* in the upper mantle; see the PREM model in 2.7, Fig. 2.53, in the depth range between about 80 and 220 km). This general trend may then be reversed for parts of the surface wave spectrum. Presentations of the propagation velocity of surface waves as a function of the period T or the frequency f are called *dispersion curves*. They differ for Love and Rayleigh waves and also depend on the velocity-depth structure of the Earth along the considered segment of the travel path (Fig. 2.9). Thus, from the inversion of surface wave dispersion data, information on the shear-wave velocity structure of the crust, and, when using periods up to about 500 s (mantle surface waves), even of the upper mantle and transition zone can be derived.

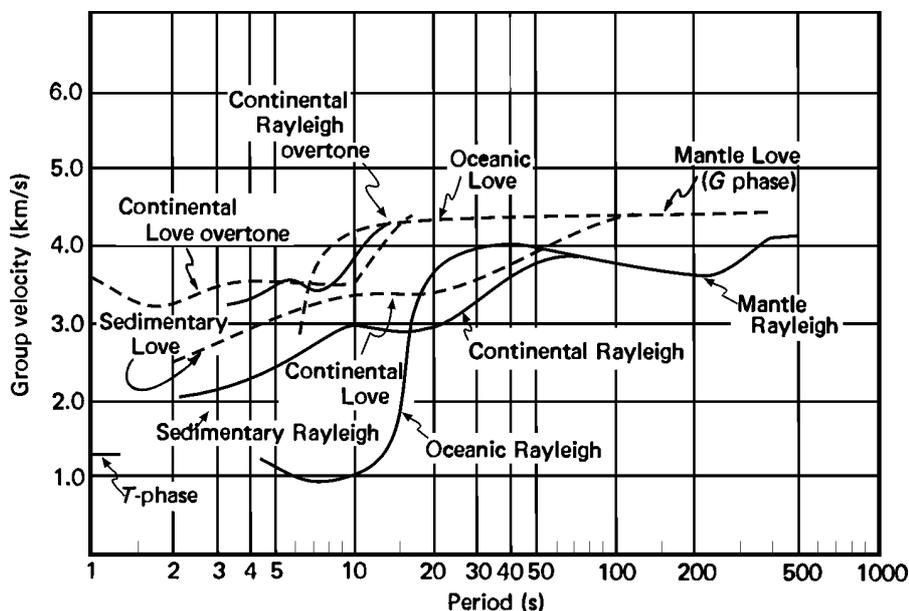


Fig. 2.9 Group-velocity dispersion curves as a function of period for Love and Rayleigh waves (fundamental modes and overtones) (from Bullen and Bolt, *An Introduction to the Theory of Seismology*, 1985; with permission from Cambridge University Press).

The large differences in crustal thickness, composition and velocities between oceanic and continental areas (Fig. 2.10) result in significant differences between their related average group-velocity dispersion curves (Fig. 2.9). They are particularly pronounced for Rayleigh

2. Seismic Wave Propagation and Earth models

waves. While the velocities for continental Rayleigh waves vary in the period range from about 15 and 30 s only from 2.9 to 3.3 km/s, they vary much more in oceanic areas (from about 1.5 to 4.0 km/s within the same period range). Consequently, LR wave trains from travel paths over continental areas are shorter and look more clearly dispersed because the various periods follow each other at shorter time differences (e.g., Figures 1d and 5a in DS 11.2). In contrast, LR wave trains with dominantly oceanic travel paths are much longer with almost monochromatic oscillations over many minutes (Fig. 2.11). Actually, the discovery of different surface-wave velocities along continental and oceanic paths were in the 1920s the first indication of the principle structural difference between oceanic and continental crust.

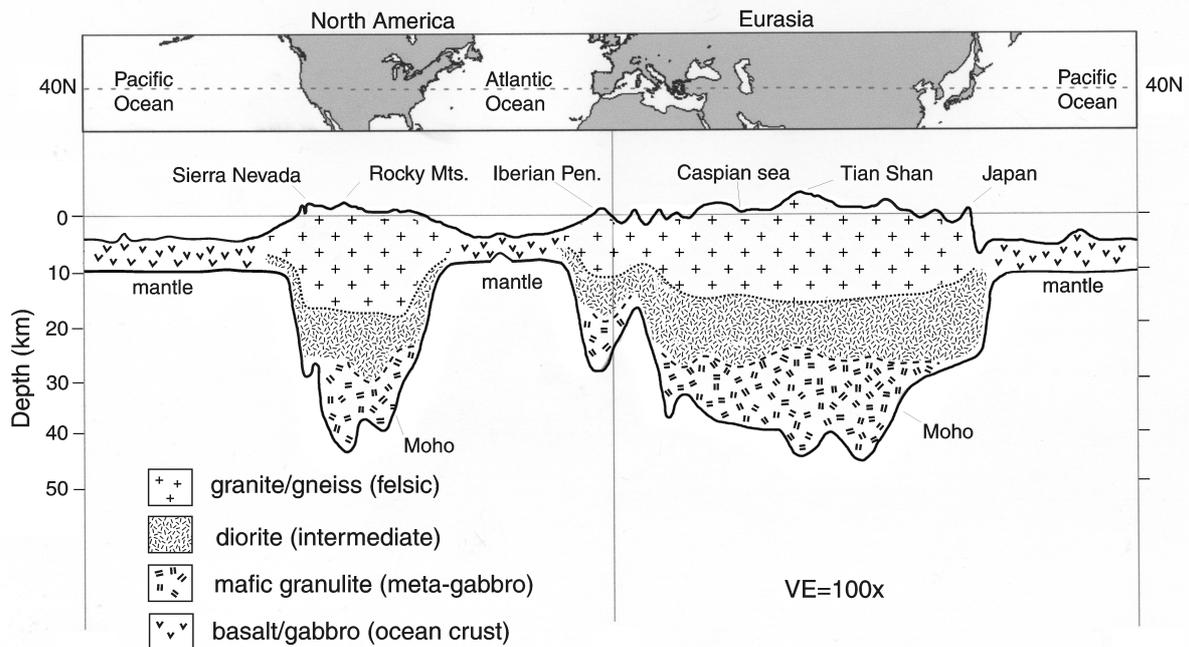


Fig. 2.10 Cross-section through the crust along 40° northern latitude. Note the different signatures for the upper (granitic), intermediate (dioritic) and lower (mafic) continental crust and the basaltic oceanic crust. The crustal base is termed “Moho” (according to its discoverer, the Croatian seismologist Andrija Mohorovičić). The P-wave velocity increases at the Moho from about 6.5-6.8 km/s to 7.8-8.2 km/s. The continental crust is about 25 to 45 km thick (average about 35 km) and has “roots” under young high mountain ranges which may reach down to nearly 70 km. The oceanic crust is rather thin (about 8 to 12 km) with a negligible upper crust (courtesy of Mooney and Detweiler, 2002).

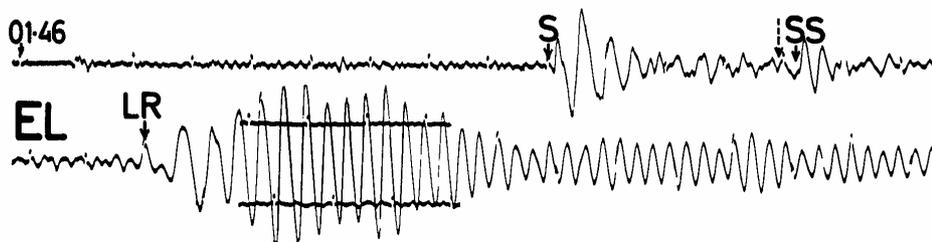


Fig. 2.11 Record of Rayleigh waves in the long-period vertical component at the station Toledo, Spain, from an earthquake in the Dominican Republic ($D = 6,622$ km; travel-path through the Atlantic Ocean) (courtesy of G. Payo, 1986).

Strictly speaking, when dealing with dispersive waves, one has to discriminate between the *group velocity* $U(T)$, with which the energy of the wave group moves and the *phase velocity* $c(T)$, with which the wave peaks and troughs at a given frequency travel along the surface. As seen from Fig. 2.12, $c(T)$ is always larger than $U(T)$. When comparing Figs. 2.9 and 2.12 the significant differences between dispersion curves calculated for a global 1-D Earth model like PREM (see 2.7, Fig. 2.53 and DS 2.1) and averaged measured curves for different types of crust or mantle models become obvious.

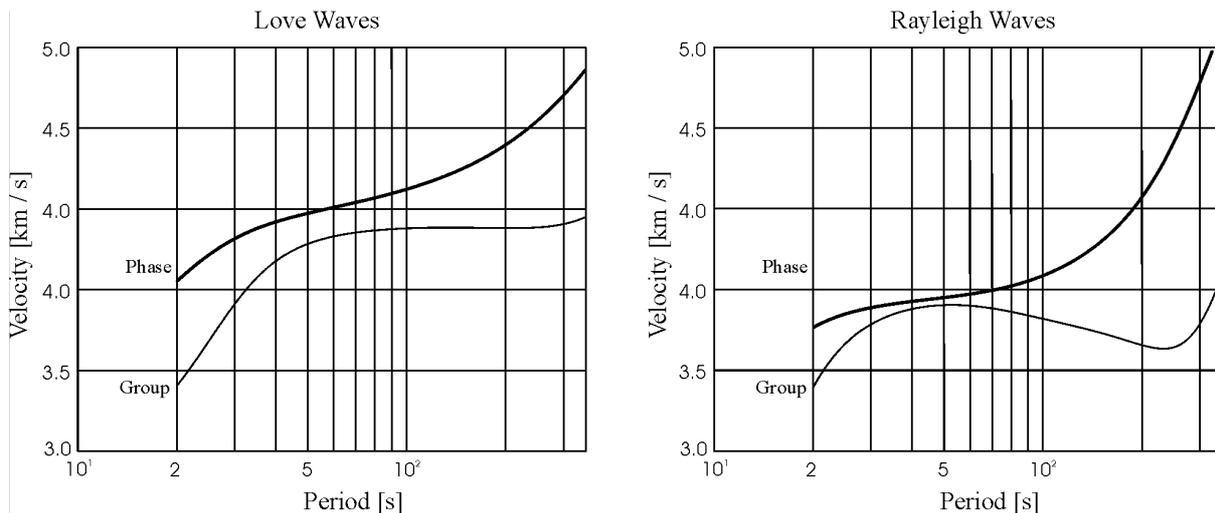


Fig. 2.12 Fundamental mode Love- and Rayleigh-wave dispersion curves computed for the PREM model with anisotropy (courtesy of Gabi Laske).

As shown in Fig. 2.13, the horizontal component of the fundamental Rayleigh wave mode will vanish in a uniform half-space at a depth of about $\Lambda/5$ and thus the particle motion becomes vertical linear. At larger depth the particle trajectories will be elliptical again, but with a *prograde* (clockwise) sense of motion. The amplitudes decay rapidly with depth. At a depth of $\Lambda/2$, the horizontal particle motion is about 10% of that at the surface while the vertical particle motion at the surface is reduced to about 30%.

Because of the strong decay of surface wave amplitudes with depth, earthquakes deeper than the recorded wavelengths will produce significantly reduced or no surface waves. The amplitude ratio between body and surface waves in broadband records is thus a reliable quick discriminator between shallow and deep earthquakes. For shallow teleseismic earthquakes the surface wave group has generally by far the largest amplitudes in broadband and long-period records (see Fig. 2.23). This is because of their 2-D propagation along the surface of the Earth and energy decay $\sim 1/r$ as compared to the 3-D propagation of body-waves and energy decay $\sim 1/r^2$. Also, the local maxima and minima in the group-velocity curves (Figs. 2.9 and 2.12) imply that surface wave energy in a wider period range around these extremes will travel with nearly the same velocity and arrive at a seismic station at about the same time, thus superimposing to large amplitudes. These amplitude maxima in the dispersive surface wave train are called *Airy phases*. For continental travel paths a pronounced Rayleigh wave Airy phase with periods around 20 s occurs which is rather stable and used for estimating the standard surface wave magnitude M_s (see 3.2.5.1). An example is given in Fig. 2.23. Long-period mantle Rayleigh waves have an Airy phase around $T \approx 220$ s (see Fig. 2.9).

2. Seismic Wave Propagation and Earth models

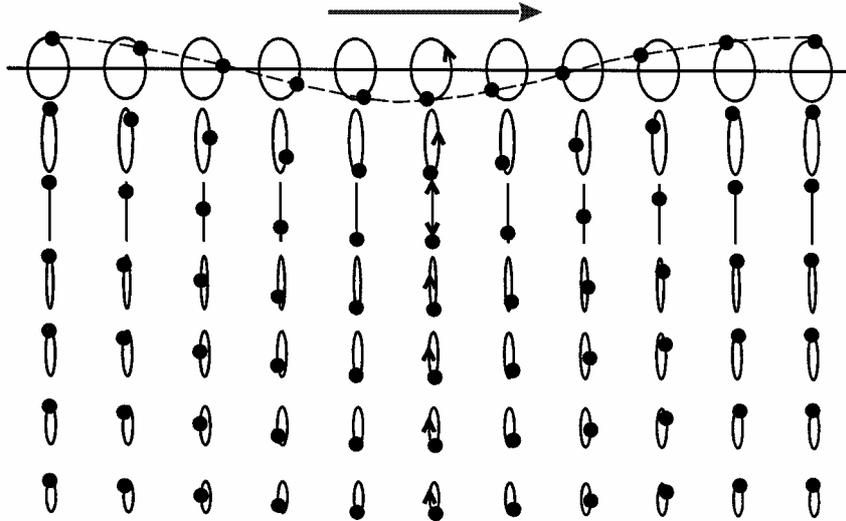


Fig. 2.13 Particle motion for the fundamental Rayleigh mode in a uniform half-space. Shown is one horizontal wavelength. At the surface the vertical motion is about 1.5 times larger than the horizontal motion. Note the change from retrograde to prograde sense of elliptical particle motion at a depth larger about $\Lambda/5$. The wave propagates from left to right. The dots show the position of the same particle at a fixed distance with time increasing from the right to the left (modified from Shearer, Introduction into Seismology, 1999; with permission from Cambridge University Press).

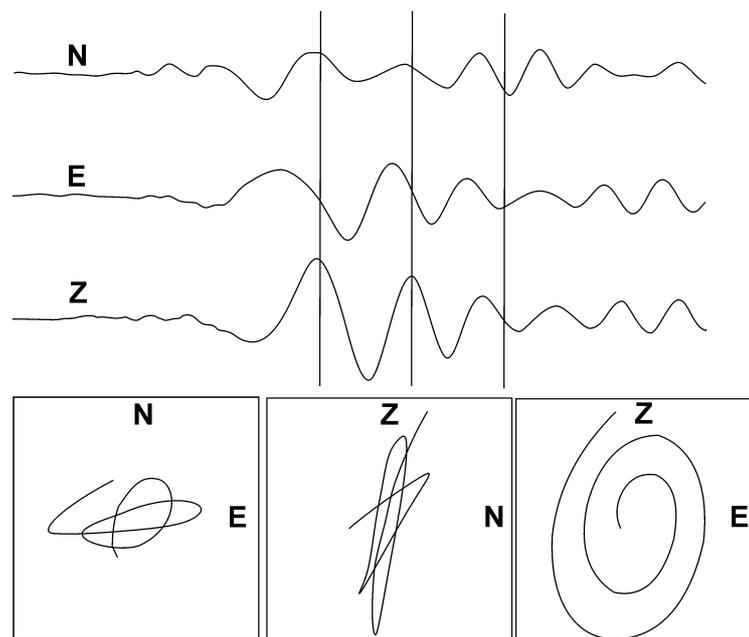


Fig. 2.14 3-component broadband records (top traces) and related plots of particle motion in the horizontal (N-E) plane and two vertical planes (Z-N and Z-E, respectively) of the surface-wave group of the same event as in Fig. 2.6 ($D = 112$ km; backazimuth BAZ = 273°).

Fig. 2.14 above shows, for the event in Fig. 2.6, the 3-component broadband record of the Rayleigh-wave group and the related particle motion trajectories in three perpendicular planes. There exists indeed a strikingly clear retrograde elliptical motion in the vertical-east

(Z-E) plane, which is in this case almost identical with the vertical plane of wave propagation (backazimuth 273°). Also the amplitude ratio vertical/horizontal component is 1.5, as theoretically expected. In the horizontal N-E plane, however, there is also some transverse energy present in this wave group. It is due to some SH energy present in the N-S component. Generally, one should be aware that the theoretically expected complete separation of LQ and LR waves in a homogeneous isotropic (horizontally layered) half-space is not absolutely true in the real Earth because of heterogeneity and anisotropy. This may cause the coupling of some Rayleigh-wave energy into Love waves and vice versa (see e.g., Malischewsky (1987) and Meier et al. (1997)), similar to S-wave splitting in the presence of anisotropy (see Fig. 2.7).

Higher mode surface waves have a different depth dependence than fundamental modes and sample deeper structure than that probed by fundamental modes of the same period.

2.3.3 Crustal surface waves and guided waves

Because of the broad maximum of the group velocity of Love-wave continental overtones with values around 3.5 km/s between about 3 and 10 s (see Fig. 2.9) they may appear in seismic records as an onset-like Lg-wave group with almost no dispersion, sometimes even in pairs (Lg1, Lg2) because of the nearby local minimum in the dispersion curve. Since the group velocity of Lg-waves is higher than that of continental fundamental modes with $T < 30$ s (see Fig. 2.9), they may appear in broadband records with an upper corner period around 20 s as clear forerunners to the surface wave group. The Lg-wave group, however, is not a pure continental Love wave but rather a complex *guided crustal* wave. It is caused by superposition of multiple S-wave reverberations between the surface and the Moho and SV to P and/or P to SV conversions as well as by scattering of these waves at lateral heterogeneities in the crust. Accordingly, Lg waves are also recorded on vertical components (see Fig. 2.15). Beyond epicentral distances of about 3° their amplitude maximum is usually well-separated from the earlier onset of Sg. Lg usually dominates seismic records of local and regional events and may propagate rather effectively along continental paths, in shield regions in particular, over a few thousand kilometers (see Fig. 2.15). Because of the stability of Lg amplitude-distance relationships in continental areas this phase is well suited for reliable magnitude estimates of regional events (see 3.2.6.6). The propagation of Lg may be barred by lateral changes in the velocity structure such as sedimentary basins, Moho topography, the transition between oceanic and continental crust or the boundaries between different tectonic units.

Near-surface seismic events such as industrial or underground nuclear explosions, rock-bursts etc. also generate short-period fundamental-mode Rayleigh waves, termed Rg. Rg waves show normal dispersion and have relatively large amplitudes on vertical components (see Fig. 2.16.). They are not excited by seismic events deeper than about one wavelength and thus a good discriminator between often man-made seismic sources near the surface and most natural earthquakes with depths most likely between 5 and 25 km (crustal earthquakes) or even deeper (intermediate or deep earthquakes in subduction zones). Rg is more strongly attenuated than the guided wave Lg. Its range of propagation is limited to less than about 600 km. However, up to about 200 km distance Rg may dominate the recorded wave train from local near-surface seismic events.

2. Seismic Wave Propagation and Earth models

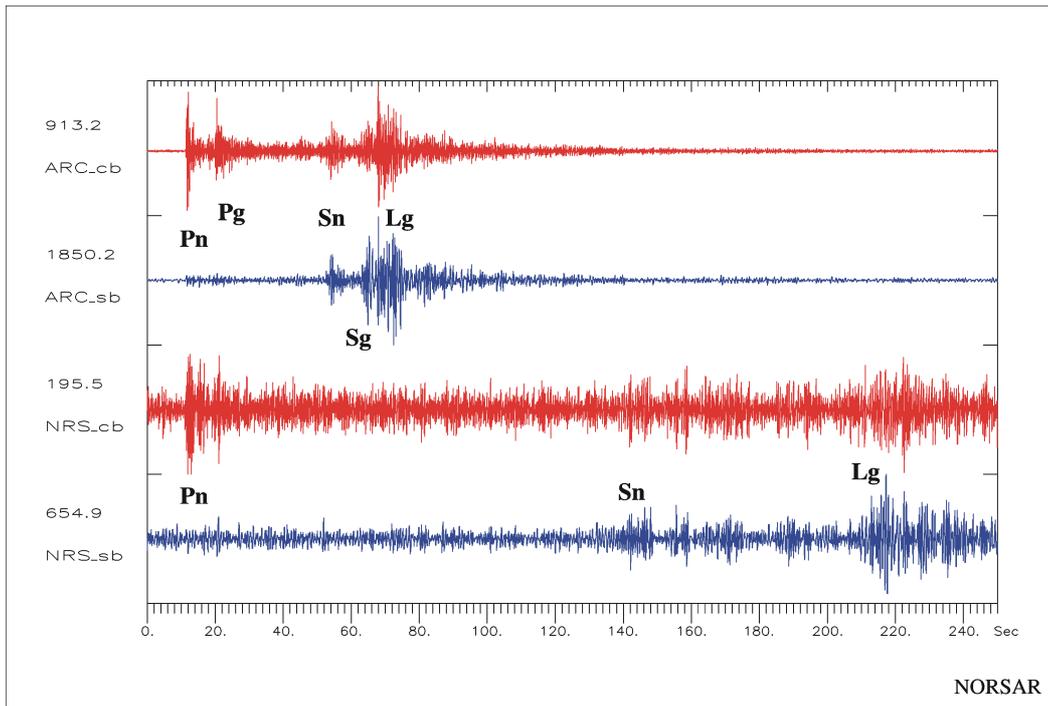


Fig. 2.15 Records of Lg, together with other crustal phases, in records of a Kola peninsula mining blast ($M_l = 2.4$) at the Norwegian array stations ARCES (distance $D = 391$ km; upper two traces) and NORES ($D = 1309$ km, bottom traces). cb and sb – P- and S-wave beams (see Chapter 9) of the vertical elements of the array, filtered with 2-8 Hz and 1-4 Hz, respectively (courtesy of J. Schweitzer).

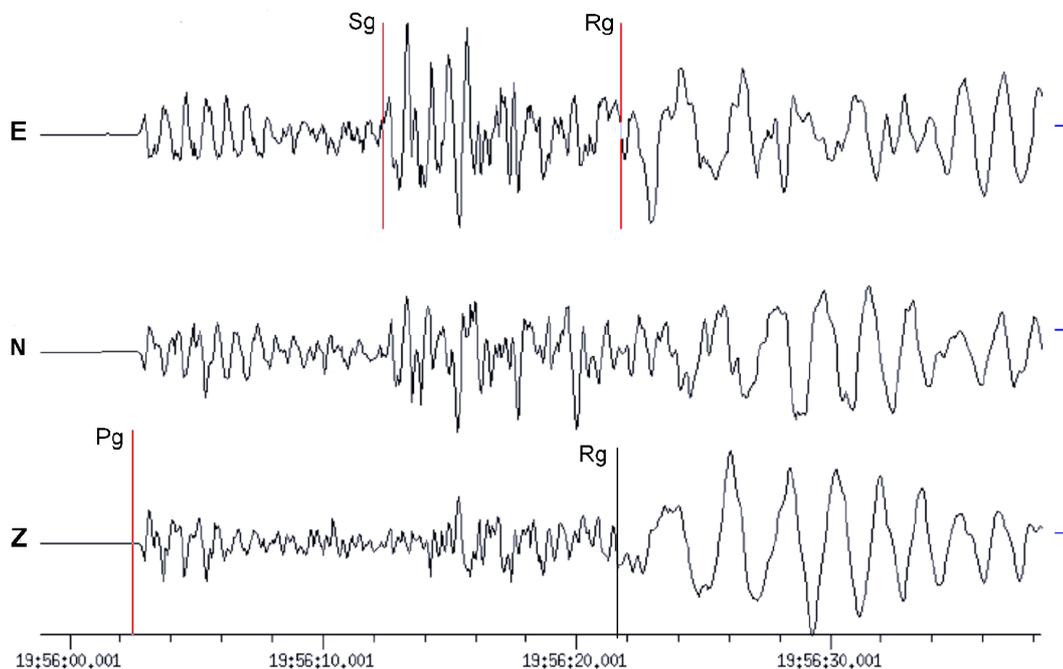


Fig. 2.16 Mining induced rock burst south of Saarbrücken, Germany, recorded at station WLF in Luxemburg ($D = 80$ km, $h = 1$ km, $M_l = 3.7$). Note the strong dispersive Rg phase.

2.3.4 Mantle surface waves

Love and Rayleigh waves travel along great circle paths around the globe. Surface waves from strong earthquakes may travel several times around the Earth. They are termed *global surface waves*. The first surface wave group arriving at a seismic station at the epicentral distance Δ° will have taken the shorter great circle while the later arrival has traveled the major arc path over $360^\circ - \Delta^\circ$ (Fig. 2.17).

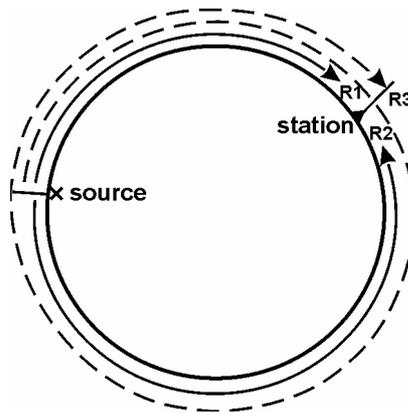


Fig. 2.17 Great circle paths for the first three arrivals of global Rayleigh waves.

These arrival groups are called R1, R2, R3, R4 etc. for Rayleigh waves and G1, G2, G3, G4 etc. for Love waves, respectively. R3 (or G3) have traveled over $360^\circ + \Delta^\circ$ and R4 over $720^\circ - \Delta^\circ$ etc. Fig. 2.18 gives an example for long-period records of P, SV, SH, R1, R2, G1 and G2 in the vertical (Z) and the two rotated horizontal components (radial R and transverse T). As expected, P appears only on Z and R while S has both SV and SH energy. The Love wave groups G1 and G2 are strongest in T and arrive ahead of R1 in R2, which are visible only on the R and Z components. But Fig. 2.18 is also a good example for inverse (negative) dispersion in the Rayleigh-wave groups. Their periods range from about 60 s to almost 200 s, with the longest periods arriving at the end of the R1 and R2 groups. This is just the period range of inverse dispersion according to Fig. 2.9 for both continental and oceanic mantle Rayleigh waves. This inverse dispersion is not seen in records of classical WWSSN long-period seismographs with a peak magnification around 15 s because the corresponding periods are filtered out by the system response of these seismographs.

Further, one should note in Fig. 2.18 that these surface waves originate from an earthquake in the Tonga trench subduction zone at a depth of $h = 230$ km. This seems to contradict the above statement, that no or only weak surface waves can be observed from deep earthquakes. However, there is no contradiction. As discussed above, the depth of penetration (and thus constructive interference) of surface waves increases with their wavelength. For the periods considered in Fig. 2.18 Λ ranges between about 230 and 880 km, i.e., it is comparable or larger than the source depth. Therefore, we still can expect significant surface wave energy in that period range for the largest amplitudes in Fig. 2.18. However, no periods below 50 s, as recorded in classical narrow-band long-period records, are recognizable in these surface-wave trains.

2. Seismic Wave Propagation and Earth models

With modern very broadband (VBB) recording systems of high dynamic range (see Chapter 5) it is possible to record such long-period global mantle surface waves up to about R7, riding on oscillations of solid Earth's tides of even longer period (more than 12 hours). Fig. 2.19 shows a striking example. The successive groups of R reveal an exponential decay of amplitudes. This allows the determination of the intrinsic frequency-dependent attenuation in the crust and mantle (see 2.5.4.2).

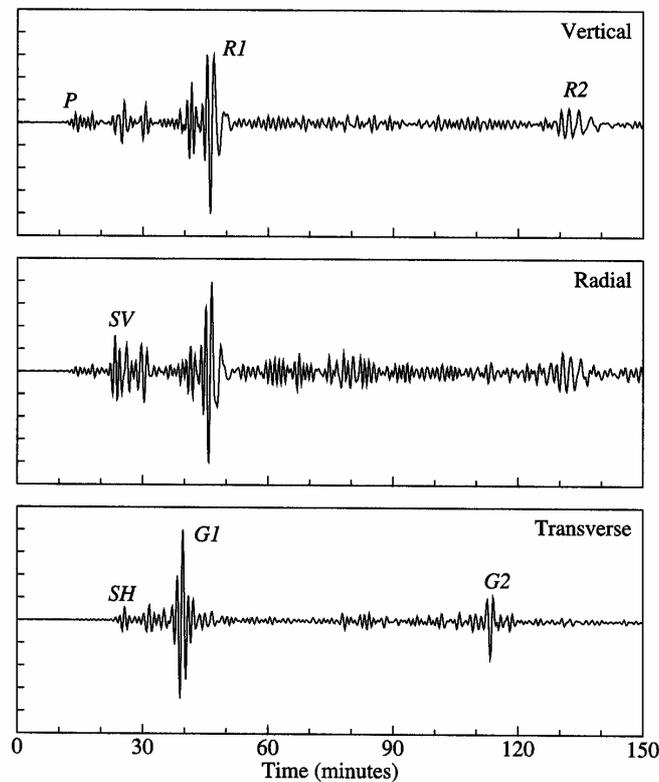


Fig. 2.18 Records of the March 11, 1989 Tonga trench earthquake ($h = 235$ km) in the Z, R and T components of the IRIS/IDA station NNA in Peru ($D = 93.7^\circ$) (from Shearer, Introduction to Seismology, 1999; with permission from Cambridge University Press).

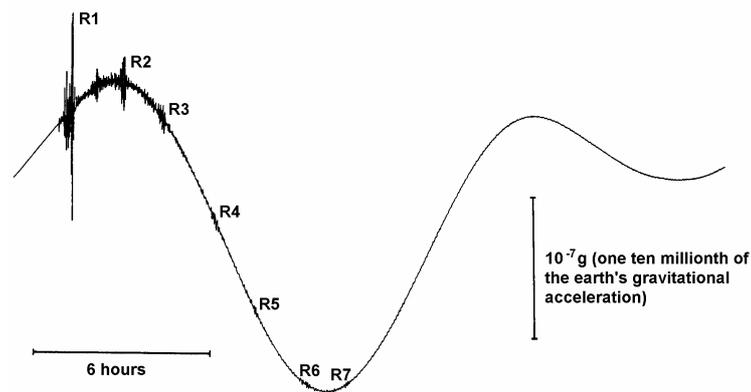


Fig. 2.19 Example of a very broadband (VBB) record with high dynamic range by the STS1 seismograph operated by the Nagoya University, Japan. The seismic wave groups from a magnitude 8.2 earthquake in the Kermadec Islands (October 20, 1986) are superimposed to solid Earth's tides (modified from a pamphlet of the Japanese Global Seismology Subcommittee for the POSEIDON project).

2.4 Normal modes

Since the Earth is not an infinite half-space but a finite body, all wave motions must be confined too. Body waves are reflected back from the surface into the Earth, surface waves orbit along great circle paths. Thus, there will be a multitude of different seismic phases arriving at a given point on the surface. Depending on their timing and periods they will interfere with each other, either in a more destructive or more constructive manner. The latter will be the case at certain resonant frequencies only. They are termed the Earth's *normal modes* and provide another alternative way of representing wave propagation. An analogy is the standing wave modes of a vibrating string fixed at both ends (Fig. 2.20). The lowest frequency is called the fundamental mode; the higher modes are the overtones. This can be treated as an eigenvalue problem: the resonant frequencies are called *eigenfrequencies*; the related displacements are termed the *eigenfunctions*.

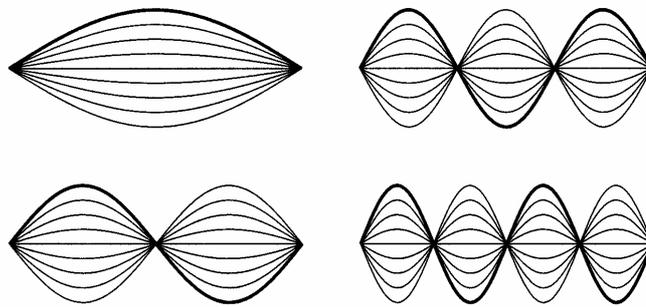


Fig. 2.20 The first four modes of vibration of a string between fixed endpoints (from Shearer, *Introduction to Seismology*, 1999; with permission from Cambridge University Press).

One should be aware of the following points about normal modes in observatory practice:

- any wave motion within the Earth may be expressed as a sum of normal modes with different excitation factors;
- there exist, in analogy to P/SV (Rayleigh) and SH (Love) -wave motion, *spheroidal modes* and *toroidal modes*, respectively;
- toroidal modes involve no radial motion and are only sensitive to the shear velocity;
- spheroidal modes have both radial and horizontal motion. They are sensitive to both compressional and shear velocities;
- long-period spheroidal modes are sensitive to gravity and thus provide information about the density structure of the Earth that may not be obtained in any other way;
- the ellipticity of the Earth, its rotation as well as its 3-D velocity variations will cause a splitting of the eigenfrequency spectral lines. Thus the investigation of normal mode splitting allows to constrain 3-D structures of the Earth;
- normal modes do (besides PKPdf amplitudes) provide information about the shear-wave velocity of the *inner core*;
- the decay of mode amplitudes with time has provided important information about the attenuation properties of the Earth at very long periods;
- normal modes provide a complete set of basis functions for the computation of synthetic seismograms for surface-wave and long-period body-wave seismology.

2. Seismic Wave Propagation and Earth models

Therefore, the collection of high-quality broadband data that also allow retrieval of normal modes is an important function of high-tech seismological broadband observatories. This requires very stable installation conditions, for horizontal seismometers in particular, e.g., in boreholes (see 7.4.5) or deep non-producing underground mines in order to reduce near surface tilt noise caused by barometric pressure variations. The latter may also be filtered out by correlating parallel recordings of seismometers and micro-barometers (e.g., Warburton and Goodkind, 1977; Beauduin et al., 1996; see Fig. 2.21).

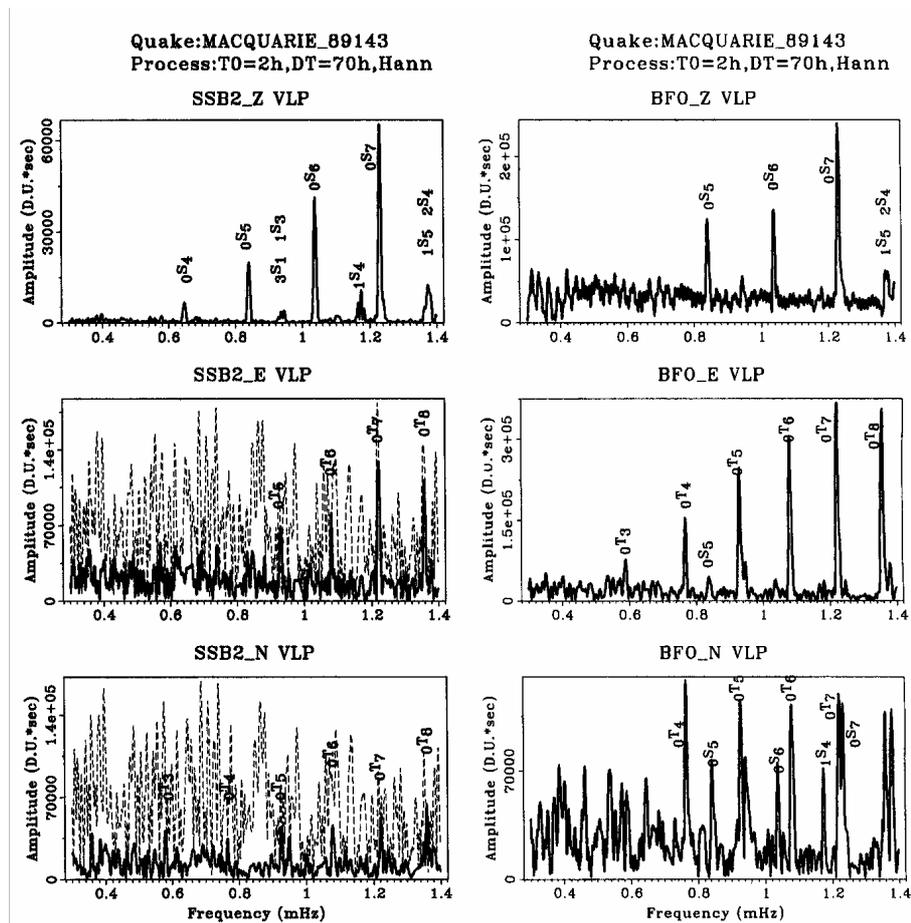


Fig. 2.21 Normal mode spectra excited by an $M_s = 8.2$ earthquake in the Macquarie Island region and recorded with STS1 at the stations SSB2 in France and BFO in Germany. BFO is located in an old silver mine and has very low tilt noise. The latter is high at SSB2 (broken lines) but could be filtered out (solid lines) by correlation with micro-barometric recordings (reproduced from Beauduin et al., *The Effects of the Atmospheric Pressure Changes on Seismic Signals ...*, Bull. Seism. Soc. Am., Vol. 86, No. 6, Fig. 8, page 1768, 1996; © Seismological Society of America).

In fact, normal mode analysis in the *spectral domain* (see Chapter 4.1) is the only practical way to examine seismic records at very long periods (> 500 s) and thus with wavelengths of 2000 and more kilometers. But normal mode studies themselves are beyond the scope of routine data analysis at seismological observatories and will not be considered in this Manual. (For further readings see Gilbert and Dziewonski, 1975; Aki and Richards, 1980 and 2002; Lapwood and Usami, 1981; Lay and Wallace, 1995; Dahlen and Tromp, 1998; Kennett, 2001).

First observations of some normal modes were made in conjunction with the strongest earthquake of the 20th century (Chile, 1960). Since then, further progress in seismometry and data analysis have permitted the identification of over a thousand modes and on that basis, to significantly refine velocity, density and attenuation models of the Earth (see 2.7; PREM model). Fig. 2.22 shows the patterns of surface and radial motions related to some of the spheroidal and toroidal modes. Their general nomenclature is ${}_nS_l$ and ${}_nT_l$. n is the number of zero crossings of amplitudes with depth while l is the number of zero (nodal) lines on the surface of the sphere.

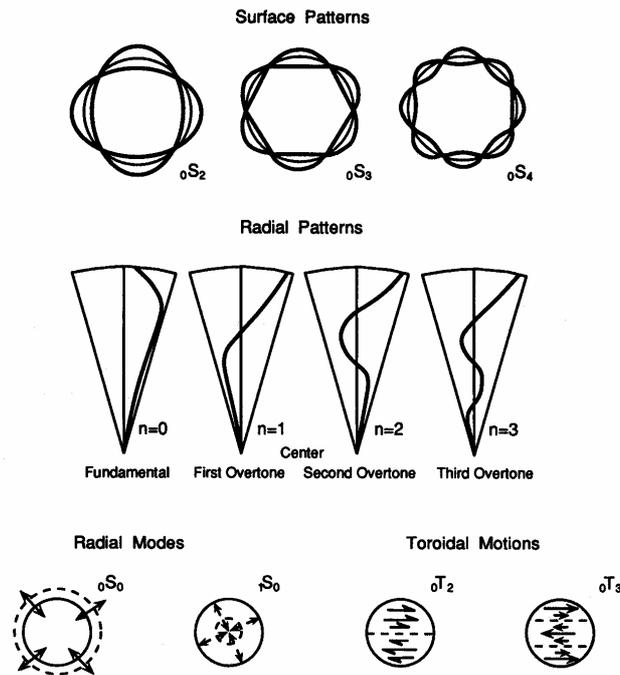


Fig. 2.22 Top: Surface and radial patterns of motions of spheroidal modes. Bottom: Purely radial modes involve no nodal pattern on the surface but have nodal surfaces at depth. Toroidal modes involve purely horizontal twisting of the Earth. For explanation of mode nomenclature see text (after Bolt, 1982; from Lay and Wallace, 1995, Fig. 4.24, p. 160; with permission of Elsevier Science (USA)).

Accordingly, the fundamental spheroidal “breathing” mode of the Earth is ${}_0S_0$ because it represents a simple expansion and contraction of the Earth. It has a period of about 20 min. ${}_0S_2$ has the longest period (≈ 54 min) and describes an oscillation between an ellipsoid of horizontal and vertical orientation, sometimes termed “rugby” mode. The toroidal mode ${}_0T_2$ corresponds to a purely horizontal twisting motion between the northern and southern hemisphere and has a period of about 44 min. Overtones ${}_iS$ and ${}_iT$ with $i = 1, 2, 3, \dots$ have one, two, three or more nodal surfaces at constant radii from the center of the Earth across which the sense of radial or twisting motions reverses.

In summary, strong earthquakes can make the planet Earth ring like a bell. Seismologists may be compared with experienced bell-makers who are able to infer from the complex sound spectra of a bell not only its size and general shape but also the composition of the alloy of which it is made.

2.5 Seismic rays, travel times, amplitudes and phases

2.5.1 Introduction

So far we have introduced seismic body and surface waves. We have learned why these different wave types travel with different velocities through and consequently appear in the seismogram at different times. We have seen that body waves form short transient wavelets (see Figs. 2.3 and 3.7), in contrast to the prolonged and dispersed wave trains of surface waves (e.g., Figs. 2.11 and 2.23). Fig. 2.23 shows a seismic record of an earthquake 73° away. Besides the discussed primary body and surface waves (P, S, LQ, and LR), several additional arrivals are marked in the seismogram and their symbols are given. These energy pulses are mainly caused by reflection or mode conversion of primary P or S waves either at the free surface of the Earth or at velocity-density discontinuities inside the Earth.

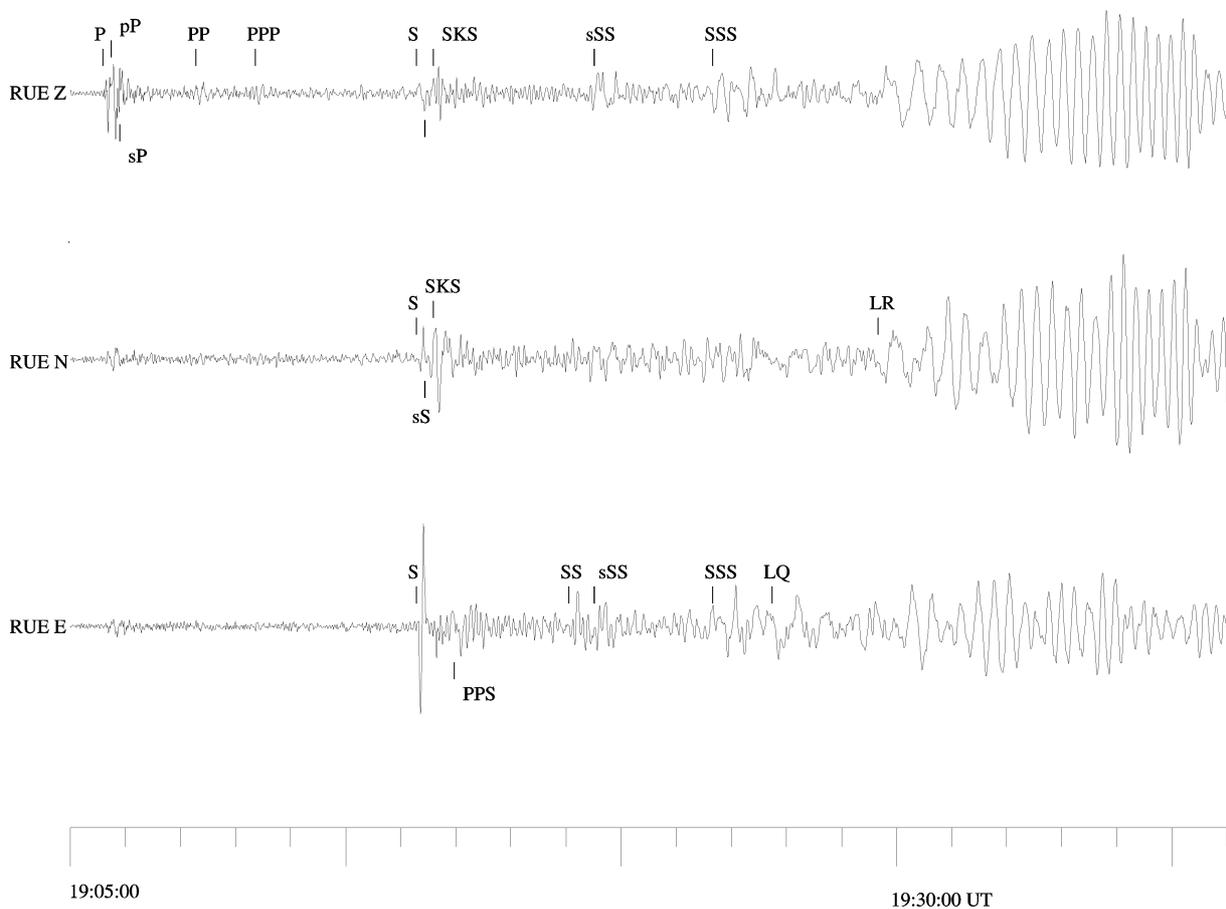


Fig. 2.23 Digital broadband record of the Seattle $M_w = 6.8$ earthquake on 28 February 2001 at the station Ruedersdorf (RUE), Germany (epicentral distance $D = 73^\circ$). Note the detailed interpretation of secondary phase onsets.

A proper understanding of these arrivals is essential for a correct phase identification that in turn is of great importance for event location (see IS 11.1) and magnitude determination (see 3.2 and EX 3.1) but also for later determination of seismic velocities inside the Earth. We will

introduce and use the concept of seismic rays to understand and illustrate the formation and propagation of these different wave arrivals.

Seismic ray theory is a very convenient and intuitive way to model the propagation of seismic energy and in particular of body waves. It is generally used to locate earthquakes and to determine focal mechanisms and velocity structure from body wave arrivals. Seismic ray theory is essentially analogous to optical ray theory, including phenomena like ray-bending, focusing and defocusing.

Using ray theory, it is important to keep in mind that it is an approximation that does not include all aspects of wave propagation. Ray theory is based on the so-called *high-frequency approximation* which states that fractional changes in velocity gradient over one seismic wavelength are small compared to the velocity. In other words, we may use ray theory only if the dimensions of structures to be considered are larger than the seismic wavelengths used.

These conditions are valid for most parts of the Earth (see global model in Fig. 2.53) and for the wavelengths that are usually recorded and analyzed in seismological observatory practice. The problem of relatively sharp boundaries, as for example the crust-mantle interface (Moho - discontinuity), discontinuities in the upper mantle, and the core-mantle boundary (CMB) or the inner-core boundary (ICB) can be tackled by matching the boundary conditions between neighboring regions in which the ray solutions are valid.

2.5.2 Huygen's and Fermat's Principle and Snell's Law

In classical optics, *Huygen's Principle* governs the geometry of a wave surface. It states that every point on a propagating wavefront can be considered the source of a small secondary wavefront that travels outward at the wave velocity in the medium at that point. The position of the wavefront at a later time is given by the tangent surface of the expanding secondary wavefronts. Since portions of the primary wave front, which are located in relatively high-velocity material, produce secondary wavefronts that travel faster than those produced by points in relatively low-velocity material, this results in temporal changes of the shape of the wavefront when propagating in an inhomogeneous medium. Since rays are defined as the normals to the wavefront, they will change accordingly. Rays are a convenient means for tracking an expanding wavefront. Fig. 2.24 depicts the change of direction of a plane wavefront and associated ray when traveling through a discontinuity which separates two homogeneous media with different but constant wave propagation velocity.

Fermat's Principle governs the geometry of the raypath. It states that the energy (or ray) will follow a *minimum time path*, i.e., it takes that path d between two points, which takes an extreme travel-time t (i.e., the shortest or the longest possible travel time, with $\partial t / \partial d = 0$). Such a path is called *stationary*. In case of a stationary time path there exist three possibilities, depending on the value (sign) of the higher derivatives of $\partial t / \partial d$:

for	$\partial^2 t / \partial d^2 > 0$	the ray follows a true <i>minimum time path</i> ,
for	$\partial^2 t / \partial d^2 < 0$	the ray follows a <i>maximum time path</i> and
for	$\partial^2 t / \partial d^2 = 0$	i.e., in case of an inflection point of the travel-time curve, the ray follows a <i>minimax time path</i> .

2. Seismic Wave Propagation and Earth models

Different kinds of seismic waves follow different time paths, e.g., the reflected waves pP (see Fig. 2.43) a true minimum path, the PP or the SKKS reflection (Fig. 2.42) a minimax path and the reflected wave P'P' (PKPPKP) (Fig. 2.44) a true maximum path. Note that the character of the stationary path influences the character (phase shift) of the reflected waveform. Whenever a seismic ray travels in some parts of its raypath as a maximum time ray, it touches a *caustic*. This caustic can be a focusing point (see 2.5.3.3 or 2.5.3.4) or a surface along which seismic rays superimpose each other (see 2.5.4.3). In any case prominent phase distortion can be observed and has to be taken into account during the analysis of seismograms.

2.5.2.1 Snell's Law for a flat Earth

From *Fermat's Principle* follows, with some simple geometry and mathematics, *Snell's Law* of wave refraction (e.g., Aki and Richards 1980 and 2002; Lay and Wallace, 1995; Shearer, 1999; Červený, 2001; Kennett, 2001):

$$\sin i/v = s \sin i = s_x = 1/v_{\text{app}} \equiv p = \text{constant} \quad (2.12)$$

where i is the angle of incidence, measured between the ray and the vertical (see Fig. 2.24), v is the velocity of wave propagation in the medium, $s = 1/v$ is called slowness, and p is the so-called ray parameter, $v/\sin i = v_{\text{app}}$ is the apparent horizontal wave propagation velocity in x -direction with $v_{\text{app}} = \infty$ for $i = 0$ (vertical incidence of the ray) and $s_x = 1/v_{\text{app}}$ is the horizontal component of the slowness vector s . Note, however, that p is constant for laterally homogeneous media only. In Fig. 2.24 the refraction of a seismic wavefront and of a related seismic ray across the interface of two half spaces with different but constant seismic velocities v_1 and v_2 is sketched. Such an instantaneous velocity jump is called first-order discontinuity. Because the ray parameter must remain constant across the interface, the ray angle has to change:

$$\sin i_1/v_1 = \sin i_2/v_2 = s_1 \sin i_1 = s_2 \sin i_2. \quad (2.13)$$

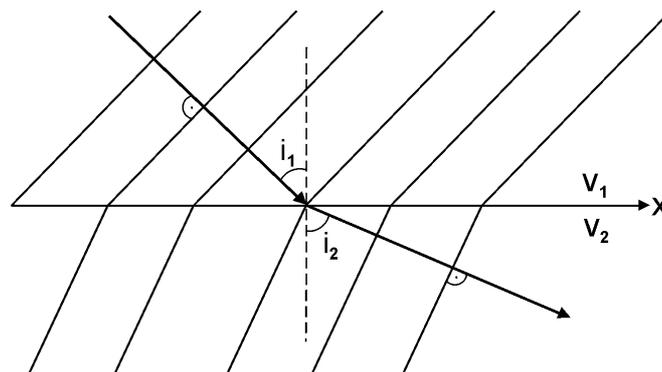


Fig. 2.24 A plane wavefront with the associated ray crossing a medium boundary with $v_2 > v_1$. The ray in medium two is refracted away from the vertical, i.e., $i_2 > i_1$.

2.5.2.2 Snell's Law for the spherical Earth

Above, a flat-layered case was considered. Yet the Earth is a sphere and curvature has to be taken into account at distances greater than about 12° . In this case the ray parameter has to be

modified. In Fig. 2.25 a ray is sketched in a sphere composed of two concentric shells 1 and 2 of different but constant velocity v_1 and v_2 or slowness $s_1 = 1/v_1$ and $s_2 = 1/v_2$, respectively. At the first interface between medium 1 and 2, Snell's Law must be satisfied locally, i.e.,:

$$s_1 \sin i_1(r_1) = s_2 \sin i_2(r_2) \quad (2.14)$$

for $r_1 = r_2$. Inside shell 2, however, despite $v_2 = \text{const.}$, the incidence angle changes as the ray progresses, namely, $i_1(r_1) \neq i'_2(r'_2)$. If we project the ray in medium 2 further to its turning point where $r = r_{\min}$ we see from the set of right triangles that the following relationship holds:

$$s_1 r_1 \sin i_1 = s_2 r'_2 \sin i'_2.$$

This is true along the entire ray path and we can generalize

$$s r \sin i = r \sin i/v \equiv p, \quad (2.15)$$

which is the modified Snell's Law for a spherical Earth.

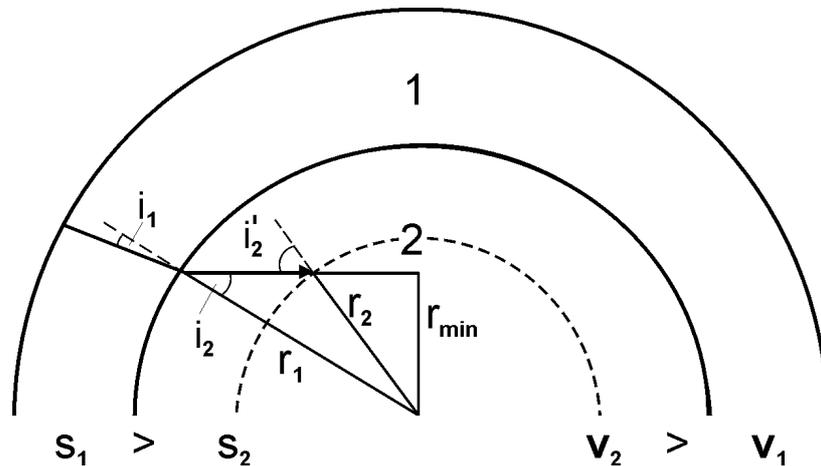


Fig. 2.25 Ray geometry for an Earth model consisting of two spherical shells of constant but different velocity v_1 and v_2 .

2.5.3 Rays and travel times in laterally homogeneous (1-D) media

2.5.3.1 Velocity gradient

It is true for most parts of the Earth that the seismic velocity increases with depth due to compaction of the material. Consider a ray travelling downwards through a stack of layers with constant velocities $v_i = 1/s_i$ each, however, increasing layer velocities with depth (Fig. 2.26). Applying Snell's law

$$p = s_1 \sin i_1 = s_2 \sin i_2 = s_3 \sin i_3 \dots \quad (2.16)$$

we can derive the incidence angle i , that is continuously increasing with depth, and finally approaching 90° . At $i = 90^\circ$ the ray is at its *turning point* tp .

2. Seismic Wave Propagation and Earth models

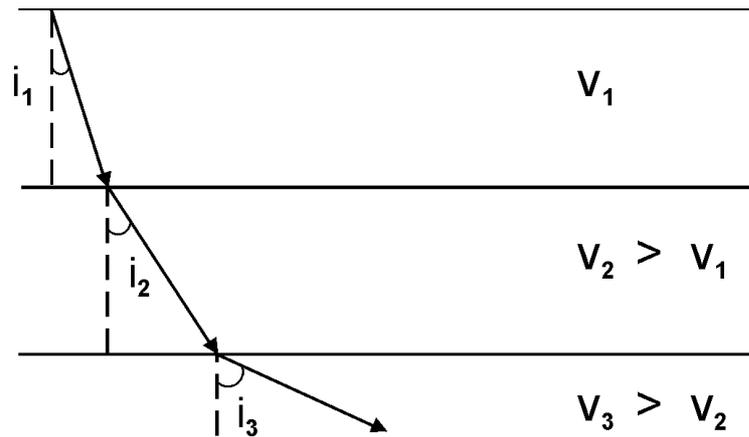


Fig. 2.26 Ray through a multi-layered model with constant velocity within the layers but increasing velocity with depth of the layers. The ray angle i increases accordingly with depth.

This can be generalized by modeling a velocity gradient with depth as a stack of many thin layers with constant velocity. Rays and travel times for this case are sketched in Fig. 2.27. The plot of arrival times t versus distance x is generally called the *travel-time curve*. The tangent dt_i/dx_i on the travel-time curve at any distance x_i corresponds to the inverse of the horizontal wave propagation velocity $1/v_{appi}$ and thus to the ray parameter p_i of that ray which comes back to the surface at x_i . Because of $\sin i = \sin 90^\circ = 1$ at the turning point of the ray, we can determine the velocity v_{tp} at the turning point of the ray either from the gradient of the travel-time curve at x_i via $p_i = dt_i/dx_i = 1/v_{tp}$ or by knowing the sub-surface velocity v_{oi} at station x_i and measuring the incidence angle i_{oi} at that station ($v_{tp} = v_{oi} / \sin i_{oi}$).

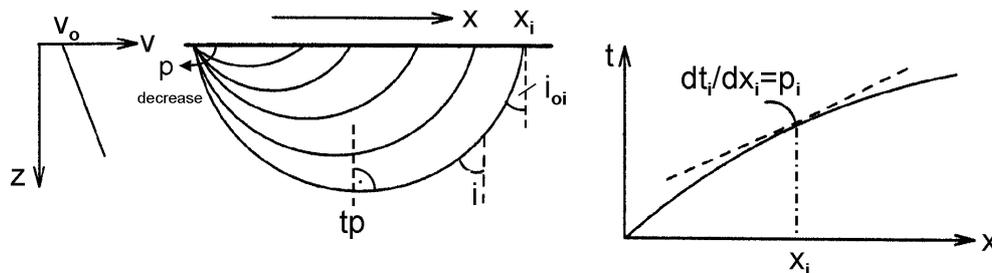


Fig. 2.27 Raypaths (middle) and travel-time curve (right) for a model with velocity v gradually increasing with depth z (left). The incidence angle i increases continuously until it reaches 90° at the turning point tp , then the rays turn up again to reach the surface at x_i . On the travel-time curve each point comes from a different ray with a different slowness and ray parameter p . The gradient of the tangent on the travel time curve at x_i is the ray parameter $p_i = dt_i/dx_i$. In the considered case of modest velocity increase with depth the distance x increases with decreasing p . The related travel-time curve is termed *prograde*.

2.5.3.2 Effect of a sharp velocity increase

Next we consider the effect of a sharp velocity increase, which may be an increase in gradient (second-order discontinuity) or an instantaneous velocity jump (first-order discontinuity). Fig.

2.28a shows on the left side a hypothetical velocity-depth model in the upper crust of the Earth together with the related seismic rays and on the right the corresponding travel-time curves in the reduced-time presentation $t_{red} = t - (x/v_{red})$. Usually travel-time increases with distance. Consequently, presenting absolute travel-time curves for large epicentral distances would require very long time-scales. Also, small changes in dt/dx are then not so well recognizable. Therefore, in order to reduce the time scale and to increase the resolution of changes in slowness, travel-time curves are often represented as *reduced travel-time curves*, in which $t_{red} = t - x/v_{red}$ is plotted (for some constant v_{red}) as a function of x . The reduction velocity v_{red} is usually chosen so as to be close to the mean velocity in the considered depth range or of the considered seismic phase. Its reduced travel-time is then constant and positive or negative slowness deviations are clearly recognizable.

In the ray diagram of Fig. 2.28a one recognizes that at certain distances x , rays with different incidence angles may emerge. Modest velocity gradients in the upper and lower part of the velocity profile result in rays which return to the surface with increasing distance x for decreasing ray parameter p . This produces prograde travel-time branches (yellow and green branches in the t_{red} - x plot). In contrast, a strong velocity gradient leads to decreasing x with decreasing p and thus to a receding (retrograde) travel-time branch (red). Thus, a strong gradient zone between two weak gradient zones results in a *triplication* of the travel-time curve. The endpoints of the triplication are called *caustics*. At the caustics (positions x_1 and x_2) rays, which have left the source under different take-off angles, arrive at the surface at the same time. This causes a focusing of energy, large amplitudes and a waveform distortion (see 2.5.4.3). Fig. 2.28b shows qualitatively, with the same color coding as in Fig. 2.28a, the changes in the ray parameter p with distance x for the prograde and retrograde travel-time branch(es) of a triplication.

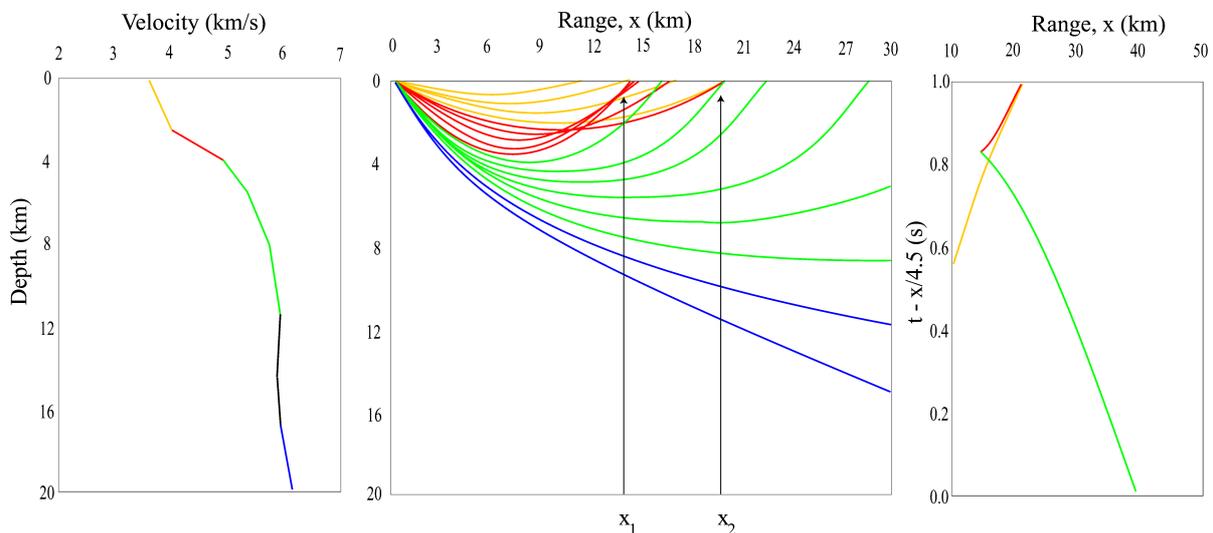


Fig. 2.28a **Left:** Velocity-depth profile in a model of the upper crust with a strong velocity gradient between about 2.5 and 6 km depth and related seismic rays from a surface source. **Right:** ray diagram and t_{red} - x relation for the given model; $v_{red} = 4.5$ km/s. Note the differently colored segments of the velocity-depth distribution and of the travel-time branches that relate to the seismic rays given in the same color. Yellow and green: prograde travel-time curves, red: retrograde travel-time curve. Note the two lowermost blue rays that have already been affected by a low-velocity zone below 10 km depth (courtesy of P. Richards.)

2. Seismic Wave Propagation and Earth models

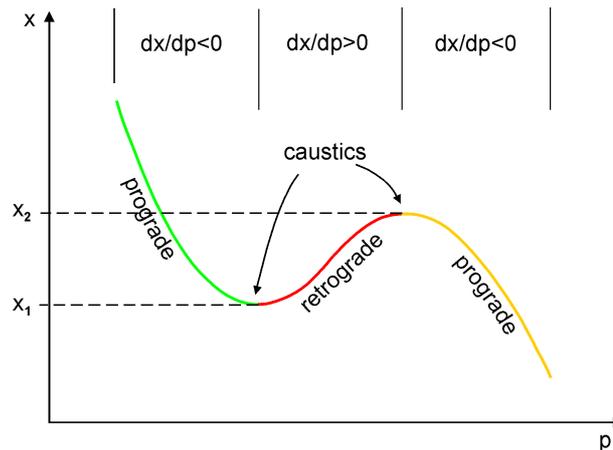


Fig. 2.28b Distance x as a function of ray parameter p for triplications. Note that the colors in this diagram correspond to the colors of the related rays and velocity segments in Fig. 2.28a.

The gradient of the retrograde travel-time branch and the position x_1 and x_2 of the caustics are controlled by the thickness and the velocity-gradient in this strong-gradient zone. Similar triplications develop in the presence of first-order discontinuities with positive velocity jump. In this case the retrograde branch relates to the postcritical reflections from such a discontinuity (see 2.5.3.6 and Fig. 2.32). The identification and quantification of such first- and second-order discontinuities is of greatest importance for the understanding of related changes in physical and/or compositional properties in the Earth. This necessitates, however, that not only first arrivals of seismic waves but also later, secondary arrivals are identified and their amplitudes measured. Since the latter may follow rather closely to the former, their proper identification and onset-time measurement may be difficult in very narrow-band filtered recordings because of their strong signal distortion (see figures in 4.2).

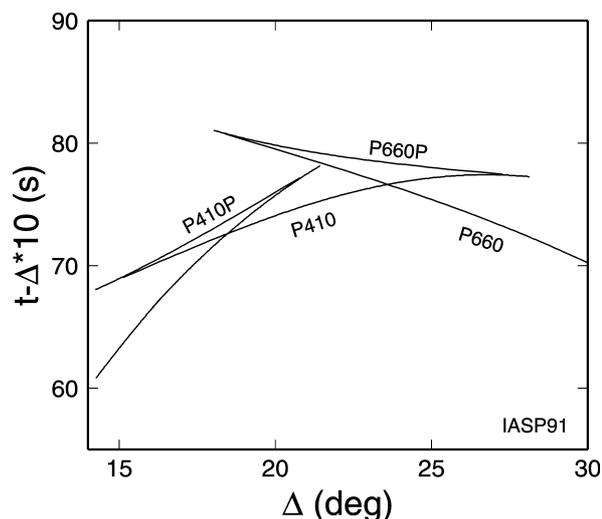


Fig. 2.29 Triplications of the P-wave travel-time curve (here in reduced presentation) due to the 410 km and 660 km upper mantle/transition zone discontinuities, calculated according to the IASP91 velocity model (Kennett and Engdahl, 1991) (see 2.7, Fig. 2.51). The P waves diving directly below the 410 km (660 km) are called P410 (P660); the phases P410P and P660P are the overcritical reflections from the outer side of these discontinuities, respectively.

Two of the most pronounced velocity and density increases occur at about 410 and 660 km below the surface (see 2.7, Figs. 2.51 and 2.53). They mark the lower boundary of the upper mantle and of the transition zone from the upper mantle to the lower mantle, respectively. Both are caused by phase transitions of the mantle material into states of higher density at critical pressure-temperature (P-T) conditions. These two pronounced discontinuities result in triplications of the P-wave travel-time curves in the distance range between about 14° and 28° (see Fig. 2.29) associated with a strong increase of P-wave amplitudes around 20° (so-called 20° discontinuity; see also Fig. 3.13).

2.5.3.3 Effect of a low-velocity zone

Velocity generally increases with depth due to compaction, however, lithologic changes or the presence of water or melts may result in low-velocity zones (LVZ). Fig. 2.30 shows the effects of an LVZ on seismic rays and the travel-time curve. The latter becomes discontinuous, forming a shadow zone within which no rays emerge back to the surface. Beyond the shadow zone the travel-time curve continues with a time off-set (delay) from a caustic with two branches: one retrograde branch (blue) beginning with the same apparent horizontal velocity as just before the beginning of the shadow zone and another prograde branch with higher apparent velocity (smaller dt/dx). This is shown in Fig. 2.30 which is in fact a continuation of the model shown in Fig. 2.28a towards greater depth. One recognizes a low-velocity zone between 12 and 18 km depth. The related ray diagram clearly shows how the rays that are affected by the LVZ jump from an arrival at distance 79 km to 170 km, and then go back to a caustic at 110 km before moving forward again. The related prograde travel-time branches and rays have been color-coded with green, blue and violet. The corresponding t_{red} - x plot on the right side of Fig. 2.30 shows nicely the travel-time offset and caustic beyond the shadow zone with two branches: a) retrograde (blue) and b) prograde (violet).

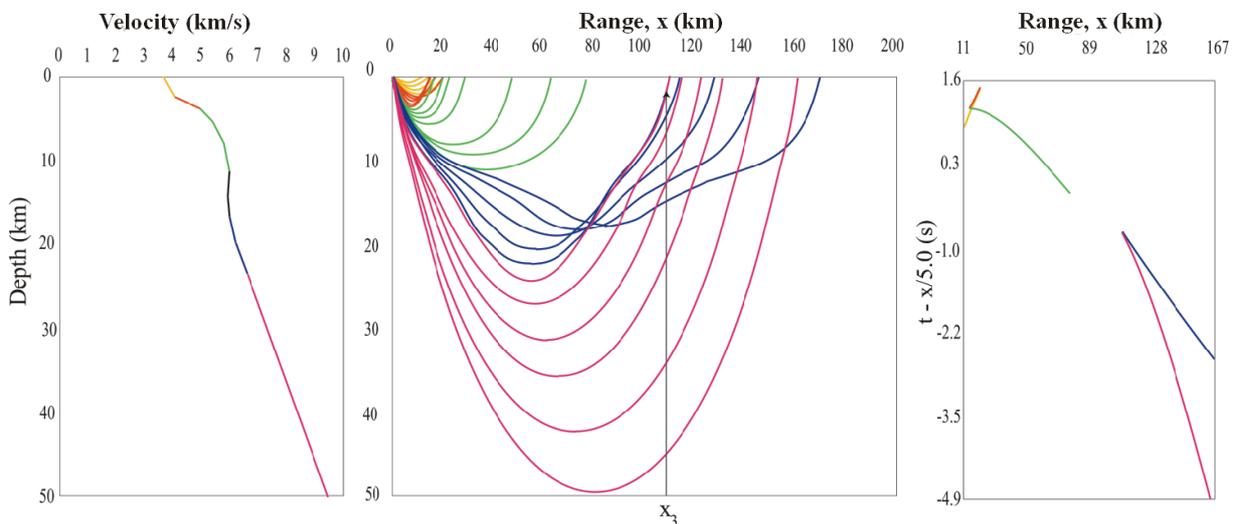


Fig. 2.30 **Left:** Velocity-depth profile and seismic rays in the crust with a low-velocity zone between $12 \text{ km} < h < 18 \text{ km}$ depth. The black segment in the velocity-depth curve produces the shadow zone. **Right:** ray diagram and t_{red} - x relation for the considered model. The reduction velocity is $v_{red} = 5.0 \text{ km/s}$. Note the additional colored travel-time branches which relate to the seismic rays given in the same color. Green and violet: prograde travel-time curves, blue and red: retrograde travel-time curves. There is a caustic at distance x_3 . Therefore, the end of the shadow has strong amplitudes (courtesy of P. Richards).

2. Seismic Wave Propagation and Earth models

An outstanding example for an LVZ, which shows these feature very clearly, is the outer core. At the core-mantle boundary the P-wave velocity drops from about 13.7 km/s in the lowermost mantle to about 8 km/s in the liquid outer core. This causes a shadow zone for short-period direct P waves between around 100° and 144° , however slightly “illuminated” by reflected arrivals from the inner-core boundary (PKiKP) and by rays that have been refracted backward to shorter distances (retrograde travel-time branch) due to the strong velocity increase in the inner core (phase PKPdf = PKIKP) (see Fig. 11.59). The travel-time branch PKPab corresponds qualitatively to the blue branch and the branch PKPdf beyond the caustic to the violet branch in Fig. 2.30 (compare with overlay to Fig. 2.47). There may exist, however, also LVZ's in the crust and in the upper mantle (asthenosphere; see PREM model in Fig. 2.53). Low-velocity zones are often more pronounced in S-wave velocity than in P-wave velocity because material weakening due to (partial) melting reduces more strongly the shear modulus μ than the bulk modulus κ (see Eqs. (2.9) and (2.10)).

2.5.3.4 Refraction, reflection, and conversion of waves at a boundary

So far we have only considered transmission of seismic waves at a boundary. However, generally not all energy is transmitted; parts are reflected or converted. If a P wave hits a boundary between different seismic velocities, four different waves may be generated: a transmitted P wave; a converted transmitted S wave purely polarized in the vertical plane of propagation (SV-wave); a reflected P wave; and a reflected converted SV wave (Fig. 2.31). The geometry of these waves is also governed by Snell's Law:

$$\sin i/v_{p1} = \sin j/v_{s1} = \sin i'/v_{p2} = \sin j'/v_{s2}. \quad (2.17)$$

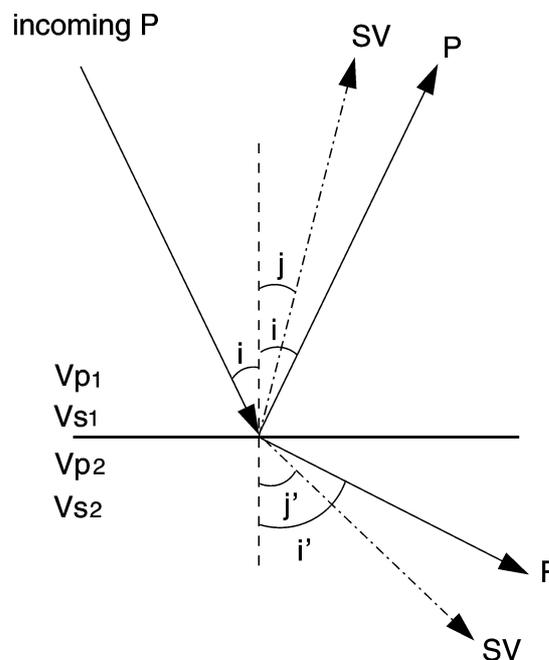


Fig. 2.31 An incident P wave at a solid-solid boundary (shown is the case $v_1 < v_2$) generates a reflected and a transmitted P wave and a reflected and transmitted SV wave. Snell's Law governs the angular relationship between the rays of the resultant waves.

In the case of an SH wave hitting the boundary, which is purely polarized in the horizontal plane, there is only a transmitted and a reflected SH wave, but no conversion into P or SV possible. If a single incident wave is split into multiple scattered waves, energy must be partitioned between these waves. Coefficients governing the partitioning between transmitted, reflected, and converted energy will generally depend on the incidence angle of the incoming wave and the *impedance contrast* at the boundary. Impedance is the product of wave velocity and density of the medium. Derivation of the expressions for reflection, transmission, and conversion coefficients is beyond the scope of this book. We refer, e.g., to the classic textbook of Aki and Richards (1980 and 2002) for a complete treatment and to Müller (1985) or Shearer (1999) for a condensed overview. The same applies to the following considerations below on seismic energy, amplitudes and phases.

2.5.3.5 Seismic rays and travel times in homogeneous models with horizontal and tilted layers

Below we consider a horizontal two-layer model above a half-space. Within the layers and in the half space the wave velocities are constant with $v_1 < v_2 < v_3$. The discontinuities between them are of first order, i.e., with instantaneous velocity jumps (see Fig. 2.32). For an incidence angle $i_1 = i_{cr}^1$, with $\sin i_{cr}^1 = v_1/v_2$ and $v_2 > v_1$, no wave energy can penetrate into the layer 2, because $\sin i_2 = 1$ and thus $i_2 = 90^\circ$. The angle i_{cr} is called the *critical angle* because for $i > i_{cr}$ all energy incident at a first-order discontinuity is totally reflected back into the overlying layer. However, part of it may be converted. The point in the travel-time curve at which a critically reflected ray (reflection coefficient 1) comes back to the surface is termed the *critical point* x_{cr} . The travel-time curve has a caustic there. Reflected rays arriving with $i < i_{cr}$ are termed *precritical* (or steep angle) reflections (with reflection coefficients < 1), those with $i > i_{cr}$ as *postcritical*, *supercritical* or *wide-angle reflections* (with a reflection coefficient = 1) (see Fig. 2.32). However, in this case the reflection coefficient becomes a complex number which results in the above discussed phase distortion of overcritical reflections. Note that the travel-time hyperbola of the reflected waves from the bottom of the first layer (red curve) merges asymptotically at larger distances with the travel-time curve of the direct wave in this layer (yellow curve).

Seismic rays incident with $i_n = i_{cr}^n$ on the lower boundary of layer n are refracted with $i_{n+1} = 90^\circ$ into the boundary between the two layers n and $n+1$. They form so-called seismic head waves (green and blue rays and travel-time curves, respectively, in Fig. 2.32). Head waves are inhomogeneous boundary waves that travel along the discontinuity with the velocity of layer $n+1$ and radiate upward wave energy under the angle i_{cr}^n . The full description of this kind of wave is not possible in terms of ray theory but requires a wave-theoretical treatment. In the real Earth, with non-ideal first-order layer boundaries, true *head waves* will hardly exist but rather so-called *diving waves* which slightly penetrate - through the high-gradient zone between the two media - into the underlying high-velocity medium. There they travel sub-parallel to the discontinuity and are refracted back towards the surface under an angle $\approx i_{cr}$. In terms of travel time there is practically no difference between a diving wave and a pure head-wave along a first-order velocity discontinuity; diving waves, however, have usually larger amplitudes.

2. Seismic Wave Propagation and Earth models

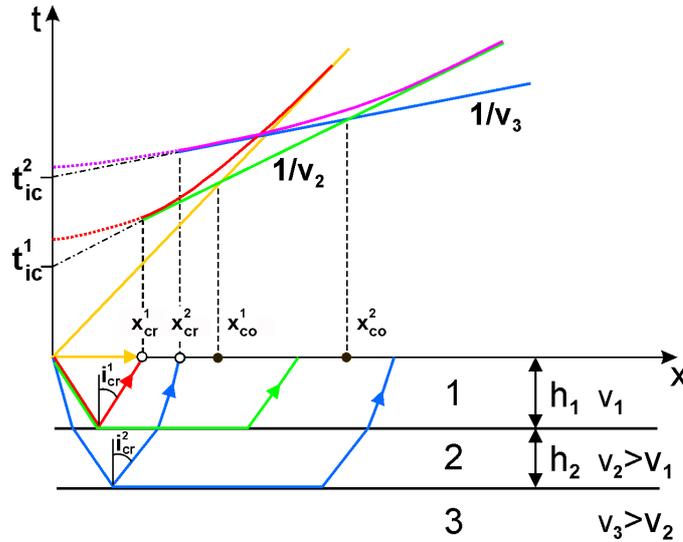


Fig. 2.32 Schematic local travel-time curves (time t over distance x from the source) for a horizontal two-layer model with constant layer velocities v_1 and v_2 , layer thickness h_1 and h_2 over a half-space with velocity v_3 . Other abbreviations stand for: t_{ic}^1 and t_{ic}^2 – intercept times at $x = 0$ of the extrapolated travel-time curves for the “head-waves”, which travel with v_2 along the intermediate discontinuity between the layers 1 and 2 and with v_3 along the discontinuity between layer 2 and the half-space, respectively. x_{cr}^1 and x_{cr}^2 mark the distances from the source at which the critically reflected rays from the bottom of the first and the second layer, respectively, return to the surface. Beyond x_{co}^1 and x_{co}^2 the head-waves from the bottom of the first and the second layer, respectively, become the first arriving waves (x_{co} - crossover distance). Rays and their corresponding travel-time curves are shown in the same color. The full red (violet) travel-time curve relates to the supercritical reflections ($i > i_{cr}$) from the intermediate (lower) discontinuity while the dotted red (violet) travel-time curve refers to the respective pre-critical ($i < i_{cr}$) steep angle reflections.

In the case of horizontal layering as in Fig. 2.32 the layer and half-space velocities can be determined from the gradients dt/dx of the yellow, green and blue travel-time curves which correspond to the inverse of the respective layer velocities. When determining additionally the related intercept times t_{ic}^1 and t_{ic}^2 by extrapolating the green and blue curves, or with help of the crossover distances x_{co}^1 and x_{co}^2 , then one can also determine the layer thickness h_1 and h_2 from the following relationships:

$$h_1 = 0.5 x_{co}^1 \sqrt{\frac{v_1 + v_2}{v_1 + v_2}} = 0.5 t_{ic}^1 \frac{v_1 \cdot v_2}{\sqrt{v_2^2 - v_1^2}} \quad \text{and} \quad h_2 = \frac{t_{ic}^2 - 2h_1 \sqrt{v_3^2 - v_1^2} (v_1 \cdot v_2)^{-1}}{2\sqrt{v_3^2 - v_2^2} \cdot (v_2 \cdot v_3)^{-1}}. \quad (2.18)$$

For the calculation of crossover distances for a simple one-layer model as a function of layer thickness and velocities see Equation (6) in IS 11.1.

In the case where the layer discontinuities are tilted, the observation of travel-times in only one direction away from the seismic source will allow neither the determination of the proper sub-layer velocity nor the differences in layer thickness. As can be seen from Fig. 2.33, the intercept times, the cross-over distances and the apparent horizontal velocities for the critically refracted head-waves differ when observed down-dip or up-dip from the source although their total travel times to a given distance from the source remain constant.

Therefore, especially in controlled-source seismology, *countershot* profiles are deliberately designed so as to identify changes in layer dip and thickness.

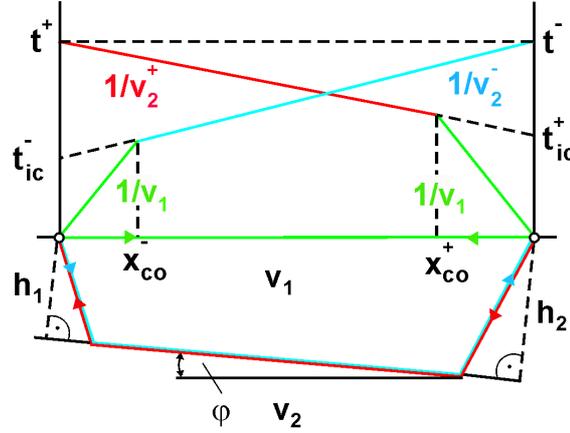


Fig. 2.33 Schematic travel-time curves for direct waves and head waves in a single-layer model with inclined lower boundary towards the half-space. Note the difference between up-dip and down-dip observations (“countershot profile”). t_{ic}^- and v_2^- are the intercept time and related apparent velocity of the down-dip head wave, t_{ic}^+ and v_2^+ the respective values for the up-dip travel-time curve.

For the considered simple one-layer case in Fig. 2.33 the dip angle ϕ and the orthogonal distance h_1 to the layer boundary underneath the seismic source on the left can be determined from the following relations:

$$\phi = \frac{1}{2} [\sin^{-1} (v_1/v_2^-) - \sin^{-1} (v_1/v_2^+)] \quad (2.19)$$

and

$$h_1 = \frac{1}{2} t_{ic}^- [v_1 v_2 / \sqrt{(v_2^2 - v_1^2)}]. \quad (2.20)$$

2.5.3.6 Wiechert-Herglotz inversion

In the case of velocity $v = f(z)$ increasing monotonously with depth z , as in Fig. 2.27, a continuous travel-time curve is observed because all rays return back to the surface. The epicentral distance $x = D$ of their return increases with decreasing p , i.e. $dx/dp < 0$. The related travel-time curve, with $dt/dx > 0$ is termed *prograde*. In this case an exact analytical solution of the inverse problem exists, i.e., when knowing the apparent horizontal velocity $c_x(D) = v_0/\sin i_0 = dD/dt$ at any point D , we know the velocity v_{tp} at the turning point of the ray that returns to the surface at D . Thus we can calculate the depth $z(p) = z_{tp}$ of its turning point. The following relations were given by Wiechert and Herglotz in 1910 for the return distance $D(p)$ and the depth of the turning point $z(p)$ of a given ray:

$$D(p) = 2 \int_0^{z(p)} \frac{p v(z)}{\sqrt{1 - p^2 v(z)^2}} dz \quad (2.21)$$

and

$$z(p) = \frac{1}{\pi} \int_0^D \cosh^{-1} \frac{c_x(D)}{c_x(x)} dx \quad (2.22)$$

2. Seismic Wave Propagation and Earth models

Note, however, that the velocity $v_{tp}(p)$ determined from dx/dt at distance $x = D$ does always relate to the respective depth half way between source and station! Nevertheless, practically all one-dimensional Earth models have been derived this way assuming that lateral variations of velocity are negligible as compared to the vertical velocity variations.

2.5.4 Amplitudes and phases

2.5.4.1 Energy of seismic waves

The energy density E contained in a seismic wave may be expressed as the sum of kinetic (E_{kin}) and potential (E_{pot}) energy densities :

$$E = E_{kin} + E_{pot} . \quad (2.23)$$

The potential energy results from the distortion of the material (strain; see. Figs. 2.2 and 2.5) working against the elastic restoring force (stress) while the kinetic energy density is

$$E_{kin} = \frac{1}{2} \rho a_v^2 , \quad (2.24)$$

where ρ is the density of the material, $a_v = A \omega \cos(\omega t - kx)$ is the ground-motion particle velocity, with A - wave amplitude, ω - angular frequency $2\pi f$ and k - wavenumber. Since the mean value of \cos^2 is $\frac{1}{2}$ it follows for the average kinetic energy density $\bar{E}_{kin} = \frac{1}{4} \rho A^2 \omega^2$, and with $E_{kin} = E_{pot}$ in case of an isotropic stress-strain relationship in a non-dispersive (closed) system for the average energy density

$$\bar{E} = \frac{1}{2} \rho A^2 \omega^2 . \quad (2.25)$$

The energy-flux density per unit of time in the direction of wave propagation with velocity v is then

$$E_{flux} = \frac{1}{2} v \rho A^2 \omega^2 \quad (2.26)$$

and the total energy-flux density \mathbf{E}_{flux} through a small surface area dS of the wavefront bounded by neighboring rays which form a *ray tube*

$$\mathbf{E}_{flux} = \frac{1}{2} v \rho A^2 \omega^2 dS . \quad (2.27)$$

When considering only waves with wavelengths being small as compared to the inhomogeneities of the medium of wave propagation (high-frequency approximation), then we can assume that the seismic energy only travels along the rays. According to the energy conservation law, the energy flux within a considered ray tube must remain constant although the surface area dS of the wavefront related to this ray tube may vary along the propagation path due to focusing or defocusing of the seismic rays (compare Fig. 2.28). Considering at different times two surface patches of the propagating wavefront $dS_1 \neq dS_2$, which are bounded by the same ray tube, and assuming that v and ρ are the same at these two locations then

$$A_1/A_2 = (dS_2/dS_1)^{1/2} , \quad (2.28)$$

i.e., the amplitudes vary inversely as the square root of the surface area of the wavefront patch bounded by the ray tube. Thus amplitudes increase due to ray focusing, which is particularly strong at caustics (see 2.5.3.2) and decrease when the wavefront spreads out.

Also, for a spherical wavefront (e.g., body-wave propagation in a homogeneous isotropic medium) the surface area grows with r^2 and for a cylindrical wavefront (e.g., for surface waves) with distance r only. Accordingly, the wave-amplitude decay is in the former case $\sim r^{-1}$ and in the latter case $\sim r^{-1/2}$. This difference in *geometrical spreading* is the main reason for the domination of surface wave amplitudes in seismic records of shallow events (see Fig. 2.23 above and Fig. 3.13).

However, wave amplitudes will also change, even in the absence of geometrical spreading, when density ρ and velocity v vary at different locations along the ray path. We then get

$$A_1/A_2 = [(\rho_2 v_2)/(\rho_1 v_1)]^{1/2}. \quad (2.29)$$

The product ρv is termed the *impedance* of the material and $(\rho_2 v_2)/(\rho_1 v_1)$ is the *impedance contrast* between the two adjacent media m_1 and m_2 . The latter largely controls the reflection and transmission coefficients at the media discontinuity. From Eq. (2.29) it follows that seismic amplitudes will increase as waves propagate into media of lower density and wave propagation velocity. This has two important implications. On the one hand, seismic stations on hard bedrock tend to record smaller amplitudes and thus to slightly underestimate event magnitudes as compared to stations on average or soft-soil conditions. On the other hand, ground shaking from strong earthquakes is usually more intense on top of unconsolidated sediments as compared with nearby rock sites. Additionally, reverberations and resonance within the unconsolidated near-surface layers above the basement rocks may significantly amplify the amplitudes at soft-soil sites. This may increase significantly local seismic hazard.

2.5.4.2 Wave attenuation

Amplitudes of seismic waves are not only controlled by geometrical spreading or focusing and by the reflection and transmission coefficients that occur at discontinuities. Besides this, wave amplitudes may be reduced because of energy loss due to inelastic material behavior or internal friction during wave propagation. These effects are called *intrinsic attenuation*. Also, scattering of energy at small-scale heterogeneities along the travel paths may reduce amplitudes of seismic waves. In the case of such *scattering attenuation*, however, the integrated energy in the total wavefield remains constant, while *intrinsic attenuation* results in loss of mechanical wave energy, e.g., by transformation into heat. The wave attenuation is usually expressed in terms of the dimensionless *quality factor* Q

$$Q = 2\pi E/\Delta E \quad (2.30)$$

with ΔE the dissipated energy per cycle. Large energy loss means low Q and vice versa, i.e., Q is inversely proportional to the attenuation. In a simplified way we can write for the decay of source amplitude A_0 with distance x

2. Seismic Wave Propagation and Earth models

$$A = \frac{A_0}{x^n} e^{-\frac{\omega t}{2Q}} = \frac{A_0}{x^n} e^{-\frac{\pi x}{QTv}} = \frac{A_0}{x^n} e^{-\frac{\omega x}{2Qv}}, \quad (2.31)$$

with A_0/x^n – the geometrical spreading term, $\exp(-x t/2Q) = \exp(-\pi/Q T v)$ the attenuation term, ω - angular frequency $2\pi/T$, T – period of wave, t – travel time, v – propagation velocity of wave, and n – exponential factor controlled by the kind of geometric spreading. According to experimental data, n varies between about 0.3 and 3, depending also on the type of seismic wave and distance range considered.

In ray theoretical methods, attenuation may be modeled through the use of the parameter t^* that is defined as the integrated value of the travel time divided by $1/Q$

$$t^* = \int_{\text{path}} \frac{dt}{Q(\vec{r})}, \quad (2.32)$$

where \vec{r} is the position vector. We can then write Eq. (2.31) as

$$A(\omega) = A_0(\omega) e^{-\omega t^*/2}. \quad (2.33)$$

Note, that P-wave attenuation Q_α and S-wave attenuation Q_β differ. They are related to the shear attenuation Q_μ and the bulk attenuation Q_κ by the relationships

$$Q_\beta = Q_\mu \quad \text{and} \quad 1/Q_\alpha = 4(\beta/\alpha)^2/3Q_\mu + [1 - 4(\beta/\alpha)^2/3]/Q_\kappa. \quad (2.34)$$

with P-wave velocity $\alpha = v_p$ and S-wave velocity $\beta = v_s$. In the Earth shear attenuation is much stronger than bulk attenuation. While Q_μ is smallest (and thus shear attenuation strongest) in the upper mantle and the inner core, Q_κ is generally assumed to be infinite, except in the inner core. While the P- and S-wave velocities are rather well known and do not differ much between different Earth models, the various model assumptions with respect to Q_α and Q_β as a function of depth still differ significantly (see Fig. 2.53). According to the PREM model, Q_μ is 600 for less than 80 km depth. It then drops between 80 and 220 km to 80, increases to 143 from 220 to 670 km, and is 312 for the lower mantle below 670 km depth.

In practice, it is difficult to separate intrinsic attenuation and scattering Q . Particularly in local earthquake records, which are strongly affected by scattering on crustal inhomogeneities, scattering Q dominates. Scattering Q is usually determined from the decay of coda waves following Sg (SmS) onsets (e.g., Fig. 2.40) and is called accordingly Q_c . A full discussion on these topics can be found, e.g., in Aki and Richards (1980; pp. 170-182).

In this context it should be mentioned that amplitudes of S waves are generally about five times larger than those of P waves (see Fig. 2.3). This follows directly from Eq. (3.2) in Chapter 3 or from the far-field term of the Green's function when modeling earthquake shear sources (see Equation (24) in the IS 3.1) taking into account that $v_P \approx v_S \sqrt{3}$ (see this Chapter, Eq. (2.9)). Also, the periods of S waves are longer than those of P waves, again by at least a factor of $\sqrt{3}$, due to the differences in wave propagation velocity and the related differences in the corner frequencies of the P- and S-wave source spectrum. Additionally, S waves are much stronger attenuated than P waves (see following section), thus filtering out higher frequencies more strongly. It should also be noted that S waves do not propagate in the fluid outer core

because of vanishing shear modulus (see Fig. 2.53). Therefore, no direct S waves are observed beyond 100° epicentral distance.

The discussed differences in amplitude-distance relationships have to be compensated by wave-type dependent calibration functions in order to be able to derive comparable magnitude values for seismic events based on amplitude readings from different types of seismic waves (see 3.2).

2.5.4.3 Phase distortions and Hilbert transform

As shown in Fig. 2.27 seismic rays will curve in the case of a vertical velocity gradient and thus seismic wavefronts will no longer be planar. Nevertheless, locally, between adjacent rays defining a *ray tube*, the wavefront still can be considered as a plane wavefront. In the case of strong gradients, retrograde travel-time branches will develop because rays bend stronger, cross each other and the wavefront folds over itself at the turning point (Fig. 2.34). Accordingly, a local plane wavefront traveling through a strong vertical velocity gradient will experience a constant, frequency-independent $\pi/2$ phase advance at the turning point. The envelope of turning points of these crossing bended rays is termed an *internal caustic surface*. Because of the $-\pi/2$ phase shift the up-going plane wave is the *Hilbert transform* of the down-going wave. More generally, whenever a ray has a non-pure minimum raypath (see 2.5.2) it touches such a caustic. Consequently, its pulse shape is altered (see Fig. 2.35). Example: The Hilbert transform of a pure sine wave is a cosine wave. In the case of seismic waves this phase shift by $-\pi/2$ has to be applied to each single frequency represented in the seismic pulse. This results in the known pulse shape alterations.

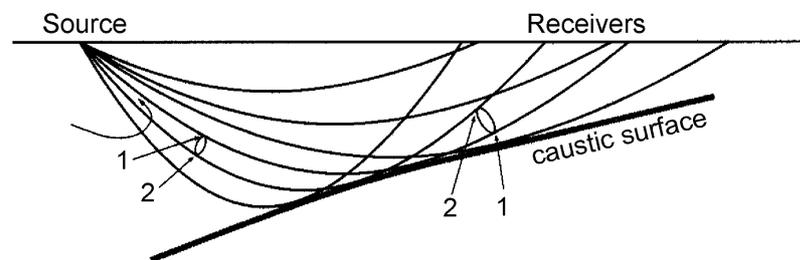


Fig. 2.34 In a medium with steep vertical velocity gradient, seismic rays with larger take-off angles from the source turn back towards the source thus forming a retrograde branch of a travel-time curve. The crossing of ray paths forms an internal caustic surface that produces a $-\pi/2$ phase shift in the waveforms (according to Choy and Richards, 1975; modified from Shearer, Introduction to Seismology, 1999; with permission from Cambridge University Press).

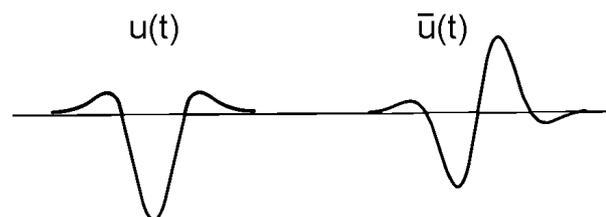


Fig. 2.35 Left: a typical seismic pulse; right: its Hilbert transform.

2. Seismic Wave Propagation and Earth models

Generally, in the case of a steep velocity gradient producing a retrograde travel-time branch, ray theory predicts that seismic wave arrivals along this branch are Hilbert transformed compared to the prograde branches. One should note, however, that also in case of relatively weak vertical velocity gradients, which do not produce a retrograde travel-time branch for the direct wave, the related reflected phase might nevertheless be Hilbert transformed (see Fig. 2.36 for PP waves). On the other hand, when the gradient becomes too steep, or in case of a first-order velocity step discontinuity, the postcritical reflection coefficients for such an interface involve a continuous change in phase with ray angle. The phase shift may then acquire any value other than a constant $-\pi/2$ phase shift.

Without exception, all the distorted waveforms bear little or no resemblance to the original waveforms. Accordingly, neither their onset times (first arrival of energy) nor the relative position of peaks and troughs of the distorted waveforms appear at the times that are theoretically predicted by ray theory. This biases onset-time picking, related travel-time determinations as well as waveform correlations between primary and Hilbert transformed phases. Therefore, modern digital data-analysis software can routinely apply the inverse Hilbert transform to phases distorted by internal caustics. The following major teleseismic body-wave phases are Hilbert transformed: PP, PS, SP, SS, PKPab, pPKPab, sPKPab, SKKSac, SKKSdf, PP', S'S'ac. For the nomenclature of these phases and their travel paths see Fig. 2.42 and IS 2.1). However, many phases pass caustics several times in the Earth and then the final pulse shapes are the sum of all internal caustic effects.

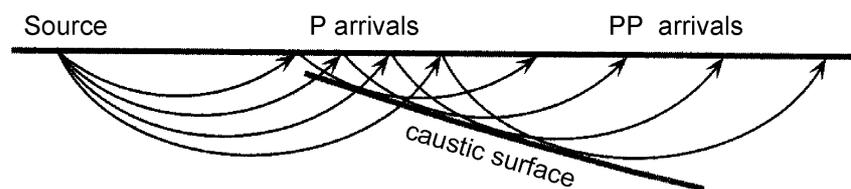


Fig. 2.36 Ray paths for the surface reflected phase PP. Note that the rays after the reflection points cross again and form an internal caustic. Accordingly, PP is Hilbert transformed relative to P and additionally has an opposite polarity (phase shift of π) due to the surface reflection (from Shearer, Introduction to Seismology, 1999; with permission from Cambridge University Press).

2.5.4.4 Effects not explained by ray theory

As mentioned above, ray theory is a high-frequency approximation that does not cover all aspects of wave propagation. Although detailed wave-theoretical considerations are beyond the scope of this Manual we will shortly mention three major phenomena that are of practical importance and not covered by ray theory.

Head waves

As mentioned in 2.5.4.1, seismic waves impinging at a discontinuity between the layers n and $n+1$ with $v_n < v_{n+1}$ under the critical incidence angle i_c with $\sin i_c = v_n/v_{n+1}$ are refracted into this discontinuity with the angle $i_{n+1} = 90^\circ$. There they travel along (or just below) this discontinuity with the velocity v_{n+1} of the lower faster medium. Such inhomogeneous waves

are usually referred to as *head waves*. They have the unique property to transmit energy back into the upper medium at exactly the critical angle i_c . However, the amplitudes of true head-wave are rather small as compared to direct, reflected or converted waves. The travel-time curve of a head wave is a straight line with the slope of $1/v_{n+1}$ (see Fig. 2.32). This provides a convenient and direct measure of the sub-discontinuity velocity. Head waves are of particular importance for crustal studies and in the analysis of seismic records from local and regional seismic events (see 2.6.1).

Seismic Diffraction

Diffraction, analog to optics, is the phenomenon of transmission of energy by non-geometric ray paths. In optics, the classic example is the diffraction of light “leaking” around the edge of an opaque screen. In seismology, diffraction occurs whenever the radius of curvature of a reflecting interface is less than a few wavelengths of the propagating wave. Seismic diffraction is important for example in steep-angle reflection data in the presence of sharp boundaries. But there are also long-period diffracted waves such as Pdif and Sdif which are “bended” around the core-mantle boundary into the core shadow zone beyond about 100° epicentral distance. Only little short-period P- and S-wave energy is observed in this shadow zone. In fact, the edge of a discontinuity/impedance contrast acts like a secondary source according to Huygen’s principle and radiates energy forward in all directions.

Diffractions can also be understood from the standpoint of *Fresnel zones*. This concept states that waves are not only reflected at a considered point of the discontinuity (like a seismic ray) but also from a larger surrounding area. The radius of the so-called first Fresnel zone is about $\frac{1}{2}$ wavelength around a considered ray arriving at a station, i.e., the range within which reflected energy might interfere constructively. The wavelength-dependent width of this Fresnel zone also determines the geometrical resolution of objects/impedance contrasts that can be at best achieved by seismic (or optical) methods.

Since the real Earth may significantly deviate from simplified global one-dimensional models, scattering and diffraction effects render not only amplitudes but also travel times of more low-frequency waves sensitive to the 3-D structure off the seismic rays. This has to be taken into account when making use of recent developments of automated travel-time measurement techniques which use cross-correlation of observed body wave phases in digital broadband records with the corresponding synthetic phases possible in spherical Earth. Marquering et al. (1999) showed that for an SS wave observed at an epicentral distance of 80° , near-surface heterogeneities situated more than 15° from the bounce point at 40° can exert a significant influence upon the travel time of an SS wave. They conclude that geometrical ray theory, which has been a cornerstone of seismology for about a century and proven useful in most practical applications, including earthquake location and tomography, is, however, valid only if the scale length of the 3-D heterogeneities is much greater than the seismic wavelengths. In other words, the validity of ray theory is based on a high-frequency (short-period) approximation. However, intermediate-period and long-period seismic waves, with wavelength of the order of 100 – 1000 km, already have comparable scale lengths with 3-D anomalies in current global tomographic models. When investigating smaller 3-D structures and applying new methods of waveform correlation, these wave-theoretical considerations gain growing importance, probably even in future observatory routines.

Scattering of seismic waves

Often the primary arrivals are followed by a multitude of later arrivals that can not be explained by simple 1-D models (Fig. 2.37). The complex wave train following the primary

2. Seismic Wave Propagation and Earth models

arrival is called coda. Coda arrivals are produced by scattering, that is, the wavefield's interaction with small-scale heterogeneities. Heterogeneity at different length scales is present almost universally inside the Earth. Seismic coda waves can be used to infer stochastic properties of the medium, i.e., scale amplitude of the average heterogeneities and to estimate coda Q_c which is particularly needed for correcting source spectra prior to deriving spectral source parameters from records of local events (see exercise EX 3.4).

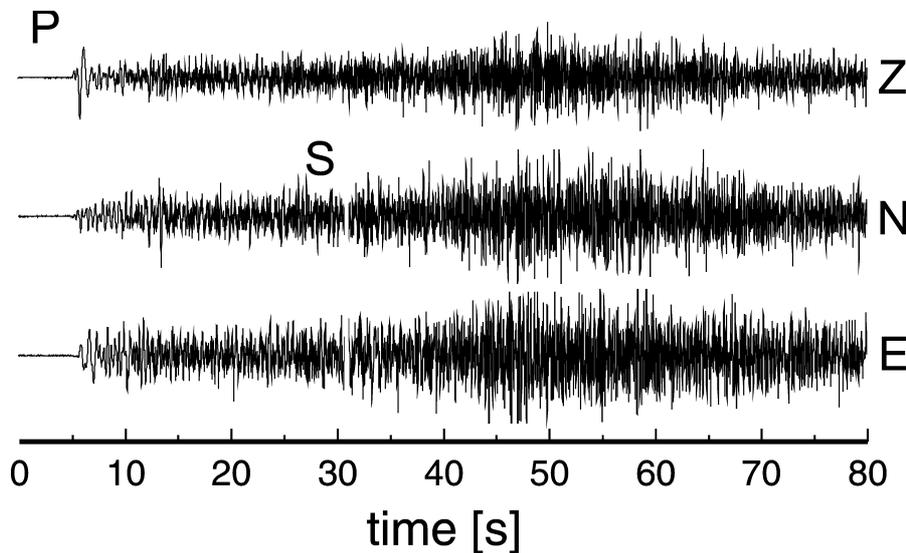


Fig. 2.37 Three-component seismogram of a local, 100 km deep earthquake recorded at a portable station on the active volcano Lascar in northern Chile. The P-wave arrival is followed by coda waves produced by heterogeneous structure in the vicinity of the volcano (courtesy of B. Schurr, 2001).

2.6 Seismic phases and travel times in the real Earth

The basic types of horizontally propagating seismic surface waves (Rayleigh waves, Love waves, and their higher modes; see 2.3) remain more or less unchanged with growing distance. Surface waves, however, do not form seismic phases (wavelets) with well-defined onsets and duration but rather dispersed wave trains. Due to the dispersion their duration increases with distance. Occasionally, surface wave trains of relatively high frequencies, as generated by shallow local events, may additionally be prolonged significantly due to lateral reverberations when propagating through strong lateral velocity contrasts in the crust (see Fig. 2.38). This phenomenon was used by Meier et al. (1997) to establish a tomography with reflected surface waves.

In contrast, seismic body waves, which propagate three-dimensionally, are more strongly affected than surface waves by refraction, reflection and mode conversion at the main impedance contrasts in the radial direction of the Earth. This gives rise, with growing distance, to the appearance of more and more secondary seismic body-wave phases following the direct P- and S-wave arrivals in seismic records. And since body waves show no dispersion in the considered frequency range below a few Hz these phases can usually be well observed and discriminated from each other as long as their travel-time curves do not overlap.

All of these secondary phases have a special story to tell about the geometrical and physical properties of the discontinuities which they encountered during their travel through the Earth's interior and which have shaped their waveforms and influenced their amplitude and frequency content. Therefore, the proper identification and parameter or waveform reporting about later phases in seismic records to relevant data centers is an important duty of seismological observatories. In addition, the complementary use of secondary phases significantly improves the precision and accuracy of seismic event locations, their source depth in particular (see Figure 7 in IS 11.1). In the following, we will introduce the main types of seismic body-wave phases that can generally be observed at local, regional and teleseismic distance ranges. They should be recognized and reported by the personnel at seismic observatories or data analysis centers. Basic features of their travel-time curves, polarization and frequency range of observation, which can guide their identification, will be presented.

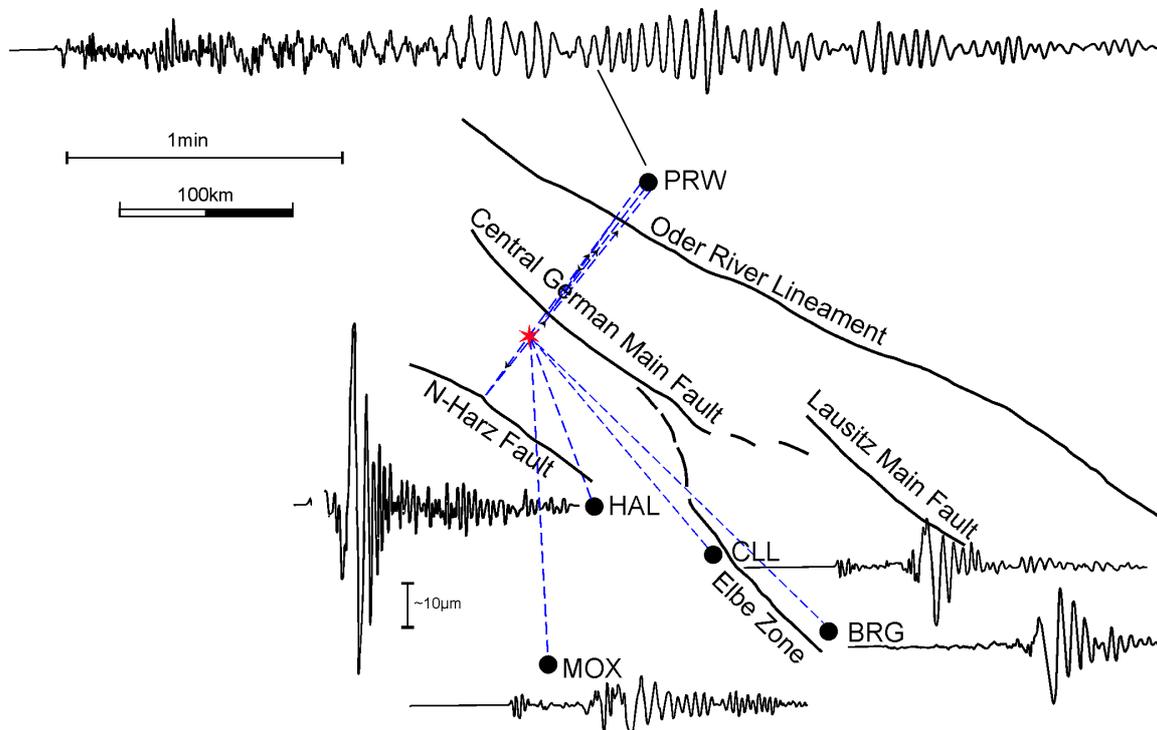


Fig. 2.38 Ray paths of surface waves (broken lines) from a mining collapse (star) to several seismic stations in the eastern part of Germany. Note: Records at stations along travel paths that have not or only once crossed some of the main tectonic faults in the area, are rather short. They have only one surface-wave maximum. In contrast, at station PRW, which is at the same epicentral distance as HAL, the seismic record is about four times longer and shows four surface-wave groups due to multiple reflections at several pronounced fault systems (compiled from data provided by H. Neunhöfer (1985; and personal communication)).

2.6.1 Seismic phases and travel times from local and regional seismic events

Seismic waves arriving at stations at local distances of up to about 150 km or regional distances of up to about 15° ($1^\circ = 111.2$ km) from the seismic source have traveled exclusively or dominantly through the crust or the sub-crustal uppermost mantle. The crust varies strongly in its thickness (see Fig. 2.10), petrologic composition and internal structure

2. Seismic Wave Propagation and Earth models

due to folding and faulting processes in the past. The resulting strong heterogeneities in its physical properties at scale length of several decameters to several km cause intensive scattering of P and S waves in the typical frequency range for the recording of near seismic events (about 0.5 to 50 Hz). Therefore, primary wave onsets are usually followed by signal-generated noise or coda waves that make it difficult to identify later seismic phases reflected or refracted from weaker intra-crustal discontinuities. It is usually only the significant velocity increase of about 20% at the base of the crust towards the upper mantle (*Mohorovičić discontinuity*, or *Moho* for short), which produces first or later wave onsets besides the direct P and S waves that are strong enough to be recognizable above the ambient or signal-generated noise level. Only in some continental regions may an intermediate discontinuity, named the *Conrad discontinuity* after its discoverer, produce recognizable critically refracted ($P_b = P^*$; $S_g = S^*$) or reflected waves (see Fig. 2.39). Accordingly, for purposes of routine seismological observatory practice, it is usually sufficient to represent the crust as a horizontal one-layer model above the half-space (upper mantle).

The currently most common global 1-D Earth model IASP91 (Kennett and Engdahl, 1991; see 2.7) assumes a homogeneous 35 km thick two-layer crust with the intermediate crustal discontinuity at 20 km depth. The respective average velocities for the upper and lower crust and the upper mantle are for P waves 5.8 km/s, 6.5 km/s and 8.04 km/s, and for S waves 3.36 km/s, 3.75 km/s and 4.47 km/s, respectively. The impedance contrast at the Conrad discontinuity and the Moho is about 1.3. Fig. 2.39 is a simplified depiction of such a two-layer crust and of the seismic rays of the main crustal/upper mantle phases to be expected. These are: P_g , S_g , P_b , S_b , P_n , S_n , P_mP and S_mS . For a detailed definition of the named phases see IS 2.1.

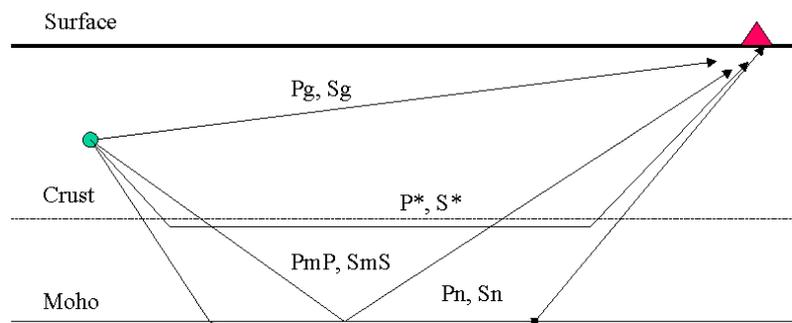


Fig. 2. 39 A simplified model of the crust showing the ray traces of the main “crustal phases” observed for near (local and regional) earthquakes. Note: $P^* = P_b$ and $S^* = S_g$.

The apparent horizontal velocity of the reflected P_mP and S_mS waves varies with distance according to their changing incidence angle on the surface. Their travel-time branches form hyperbolae that approach asymptotically the travel-time curves for P_g and S_g (or P_b and S_b) with increasing distance (see Fig. 2.40). Note that P_n and head waves have usually smaller amplitudes than P_g and S_g , at least for distances up to about 300 km. P_n can be usually identified above the noise level only when it becomes the P-wave first arrival. At larger distances, because of the stronger attenuation of upper crustal P_g and S_g and with P_n and S_n being less attenuated upper mantle diving phases, P_n and S_n may become clear P and S first arrivals (see Fig. 2.15). Beyond the critical point (at about 70-80 km distance for an average

crust) the supercritically reflected waves PmP and SmS have generally the largest amplitudes, however, arriving always closely after Pg and Sg, their onset times can usually not be picked reliably enough as to be of value for earthquake location. Therefore, these phases are usually not explicitly reported in routine observatory practice. However, reporting of Pg, Sg, Pn and Sn, if recognizable, is a must. This also applies to the reporting of the maximum amplitudes in records of near seismic events for the determination of local magnitudes M_l (see 3.2.4). Depending on source depth too, this amplitude maximum may be related to Sg/SmS, Lg, or Rg (see Figs. 2.15, 2.16 and 2.40).

Travel-time curves for the phases Pn, Pg, Sn, Sg and Lg for distances up to 400 km are given in Figure 4 of Exercise EX 11.1. These curves relate to an average single-layer crust for Central Europe. From the global Earth model IASP91, given in Datasheet DS 2.1, one may calculate respective travel-time curves for a two-layer crustal model. However, such global crustal travel-time curves may not be representative at all for certain regions and may serve as a starting model only to work with. It is one of the main tasks of operators of local and regional seismic networks to derive from their own carefully analyzed data of near events not only local/regional magnitude calibration functions (see 3.2.4) but also average local/regional travel-time curves. The latter will not only allow significantly improved seismic event locations but may later serve also as starting models for 3-D tomographic studies of crustal heterogeneities.

Fig. 2.40 shows real short-period seismic network records of two local earthquakes in Switzerland in the distance range between about 10 km and 180 km along different profiles together with the modeling of their reduced travel-time curves and inferred structural profiles. While one event was at a depth of only 5 km, the other event was about 30 km deep. The first one was observed by stations situated up-dip while the latter event was observed down-dip. One sees striking differences in the shape and gradient of the travel-time curves and in the crossover distance between Pg and Pn, in particular. In the case of the deeper event near to the Moho, Pn becomes the first arrival beyond 70 km distance, whereas for the shallower event Pn outruns Pg only at more than 130 km epicentral distance. In both cases Pg (Sg) and/or PmP (SmS) are the prominent P and S arrivals. The Pn first arrivals are relatively small. No Pb, Sb or reflected waves from a mid-crustal discontinuity are recognizable in Fig. 2.40. Note, however, that depending on the orientation of the earthquake rupture and thus of the related radiation characteristic of the source, it may happen that a maximum of energy is radiated in the direction of the Pn ray and a minimum in the direction of the Pg ray. Then the usual relationship $A_{Pn} < A_{Pg}$ may be reversed (examples are given in 11.5.1).

Misinterpretation of Pn as Pg or vice versa may result in large errors of event location. Therefore, one should have at least a rough idea at which distance in the region under study, depending on the average crustal thickness and velocity, one may expect Pn to become the first arrival. A “rule-of-thumb” for calculating the crossover distance x_{co} is given in Equation (6) of IS 11.1. For an average single-layer crust and a surface source, $x_{co} \approx 5 z_m$ with z_m – Moho depth. However, as demonstrated with Fig. 2.40, x_{co} is only about half as large for near Moho earthquakes and also the dip of the Moho and the direction of observation (up- or down-dip) does play a role. Yet, lower crustal earthquakes are rare in a continental (intraplate) environment. Mostly they occur in the upper crust.

Rules-of-thumb for calculating the source distance from the travel-time differences Sg-Pg and Sn-Pn are given in Eqs. (11.1) and (11.2).

2. Seismic Wave Propagation and Earth models

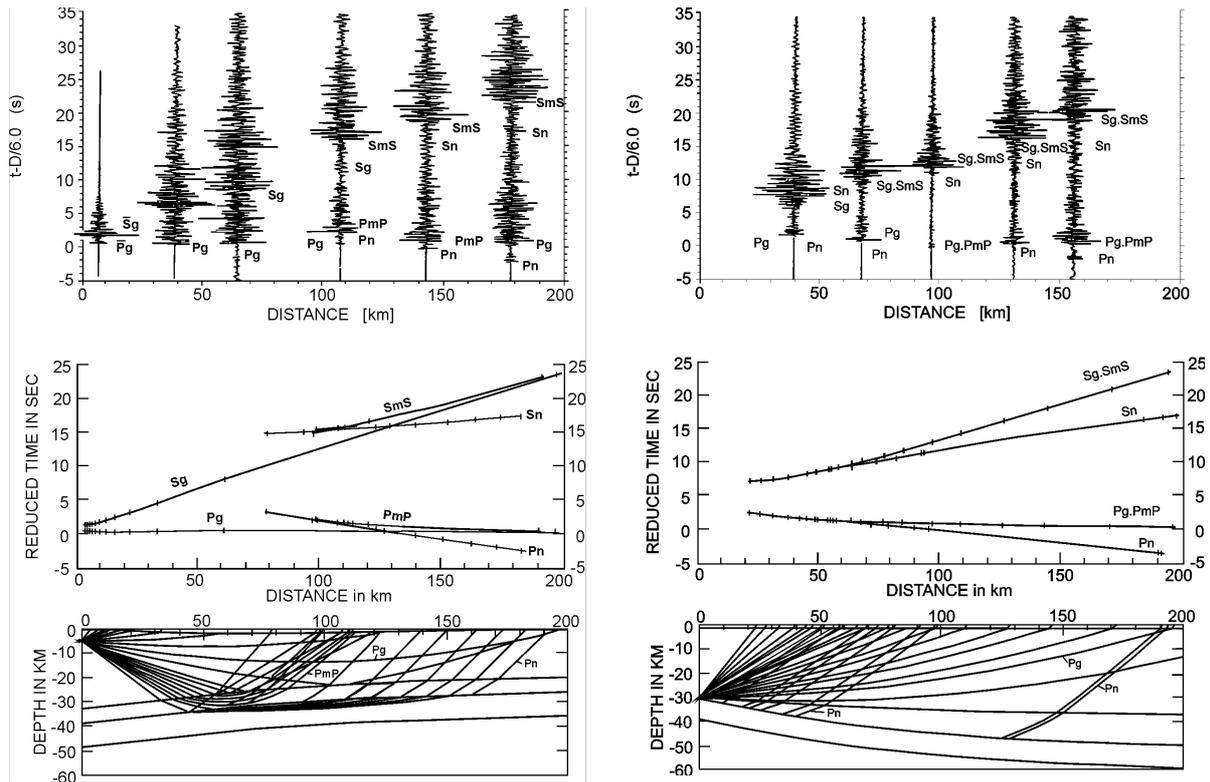


Fig. 2.40 Records (above) of two regional earthquakes of Oct. 9, 1986 at Sierre (left) and of July 7, 1985 at Langenthal, Switzerland together with the calculated reduced travel-time curves (middle) and ray-tracing crustal models which best fit the observations (below), redrawn and complemented from *Anatomy of Seismograms*, Kulhánek, plate 4, pp. 83-84, © 1990 (with permission from Elsevier Science).

Sometimes, very strong onsets after Pg, well before Sn or Sg can be expected, may be related to depth phases (e.g., sPmP; Bock et al., 1996). This may complicate proper interpretation of the local phases as well and can usually not be solved in routine analysis. Also be aware that in the case of sub-crustal earthquakes, which are common in subduction zones, none of the crustal phases discussed above exist. In this case, the first arriving longitudinal and shear wave onsets are termed P and S, respectively, as for teleseismic events (see Fig. 2.41).

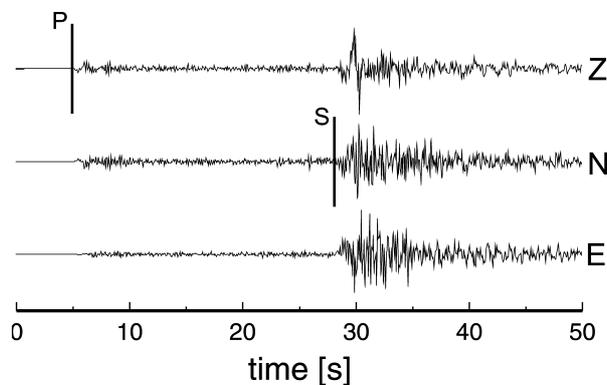


Fig. 2.41 P- and S-wave onsets from a local earthquake in northern Chile at a depth of 110 km and a hypocentral distance of about 240 km (courtesy of B. Schnurr, 2001).

2.6.2 Seismic phases and travel times at teleseismic distances

Seismic waves arriving at distances beyond 10° up to about 30° have mainly traveled through the upper mantle (from Moho to about 410 km depth) and the transition zone to the lower mantle (between about 410 km and 660 km depth). The strong discontinuities which mark the upper and lower boundary of the transition zone are associated with strong increases in seismic impedance (i.e., of both velocity and density; see Fig. 2.53). This results in two remarkable triplications of the travel-time curve for P waves (see Fig. 2.30) and S waves, which give rise to complicated short-period waveforms of P and S with rather long duration (up to about 10 and more seconds) and consisting of a sequence of successive onsets with different amplitudes.

For epicentral distances $D > 30^\circ$ P and S waves are followed by an increasing number of secondary waves, mainly phases, which have been reflected or converted at the surface of the Earth or at the core-mantle boundary. Fig. 2.42 depicts a typical collection of possible primary and secondary ray paths together with a three-component seismic record at a distance of $D = 112.5^\circ$ that relates to the suit of seismic rays shown in red in the upper part of the cross section through the Earth. The phase names are standardized and in detail explained in IS 2.1

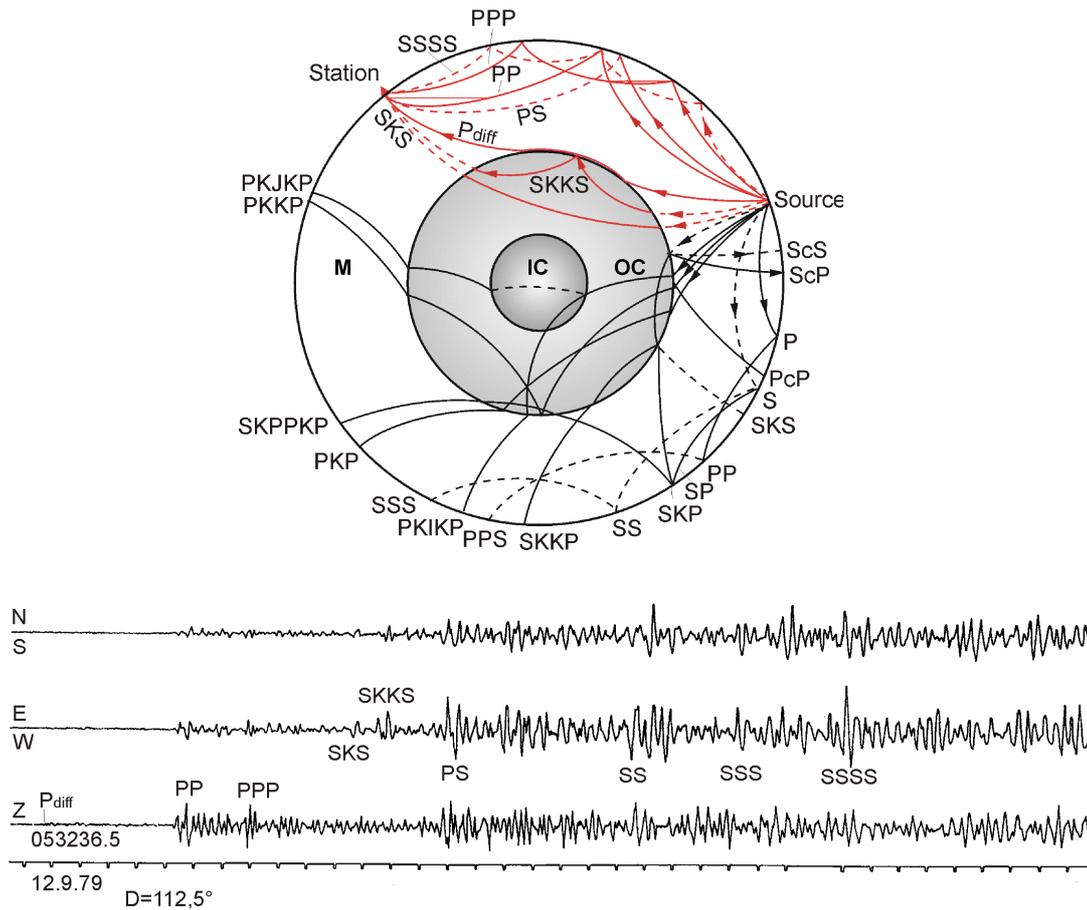


Fig. 2.42 Top: Seismic ray paths through the mantle (M), outer core (OC) and inner core (IC) of the Earth (above) with the respective phase symbols according to the international nomenclature (see IS 2.1, also for detailed ray tracing). Full lines: P rays; broken lines: S rays. Related travel-time curves are given in Fig. 2.46 and the transparency to Fig. 2.48. Red rays relate to the 3-component Kirnos SKD broadband seismograms recorded at station MOX, Germany (bottom) of body-waves from an earthquake at an epicentral distance of 112.5° .

2. Seismic Wave Propagation and Earth models

In the case of deep earthquakes the direct P wave that leaves the source downward will arrive at a teleseismic station first. It will be followed, depending on the source depth, up to about 4.5 min later by other phases that has left the source upward. These phases, reflected and eventually converted at the free surface of the Earth or an ocean bottom (e.g., pP, sP, pPP, sPP, pPKP, etc.), at the free surface of the ocean (e.g., pwP) or from the inner side of the Moho (e.g., pmP) are the so-called depth phases. Their proper identification, onset-time picking and reporting is of crucial importance for reliable determination of source depth (see 6.1 and Figure 7 in Information Sheet 11.1). Differential travel-time tables pP-P and sP-P are given in the Exercise EX 11.2. For the definition of these phases see also IS 2.1.

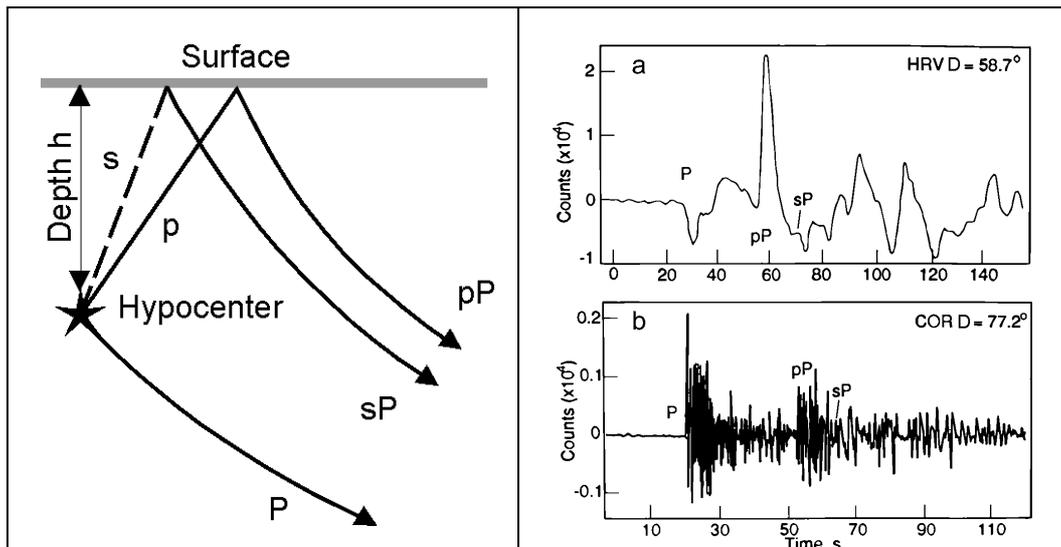


Fig. 2.43 Left: Different ray paths of a direct teleseismic P wave and of its depth phases. Right: Records of depth phases of the May 24, 1991 Peru earthquake (hypocentral depth $h = 127$ km); a) broadband record and b) simulated short-period recording (the right figure is a corrected cutout of Fig. 6.4 of Lay and Wallace, *Modern Global Seismology*, p. 205, © 1995; with permission of Elsevier Science (USA)).

However, the identification of depth phases is rather difficult for shallow crustal earthquakes because their onsets follow rather close to the direct phase, thus superposing with their wavelets. They may, however, be discriminated by waveform modeling with variable source depth (see subchapter 2.8, Fig. 2. 56).

Between about 30° and 100° epicentral distance P and S have traveled through the lower mantle, which is characterized by a rather smooth positive velocity and density gradient (see Fig. 2.53). In this distance range, seismograms are relatively clearly structured with P and S (or beyond 80° with SKS) being the first, prominent longitudinal and transverse wave arrivals, respectively, followed by multiple surface and core-mantle boundary (CMB) reflections or conversions of P and S such as PP, PS, SS and PcP, ScP etc. (see Fig. 2.42 and 2.48 with overlay). Within about 15 to 35 min after the first P arrival multiple reflections of PKP from the inner side of the CMB (PKKP; P3KP) or from the surface (PKPPK = P'P') may be recognizable in short-period seismic records. Their ray traces are shown in Fig. 2.44 and many more, with record examples, in 11.5.3.

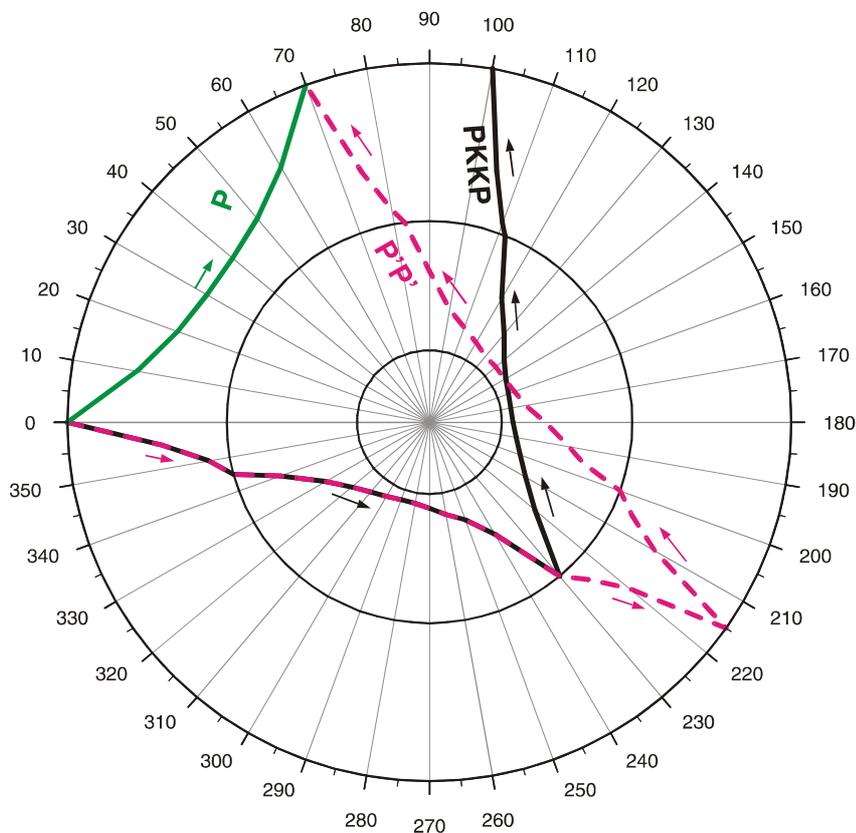


Fig. 2.44 Ray paths of PKKP and P'P' (= PKPPKP) with respect to the direct P phase (courtesy of S. Wendt, 2001).

Beyond 100° , only P-wave rays, which entered the outer core after strong downward refraction, will reach the surface. This is due to the dramatic reduction of the P-wave velocity at the CMB from about 13.7 km/s in the lowermost mantle to 8.0 km/s in the upper outer core. Thus, P waves form a *core shadow*. However, long-period P-wave energy is diffracted around the CMB into this shadow zone. According to the new IASPEI nomenclature of phase names (see IS 2.1) the diffracted P wave is termed Pdif, however the old phase symbol Pdiff is still widely used.

The amplitudes of Pdif are comparably small thus making PP the strongest longitudinal arrival up to nearly 144° (see Figs. 2.42, 2.55, 11.60 and 11.61). PKP has a caustic at 145° causing strong amplitudes comparable with those of P at much shorter distances around 50° (see Fig. 3.13) and separates into different branches beyond the caustic (see Figs. 2.45, 11.62 and 11.63).

In more detail, the types of seismic phases appearing in the various distance ranges and their peculiarities are discussed in Chapter 11 where many record examples are given both in the main text and in complementary Datasheets (DS 11.1 to 11.3).

2. Seismic Wave Propagation and Earth models

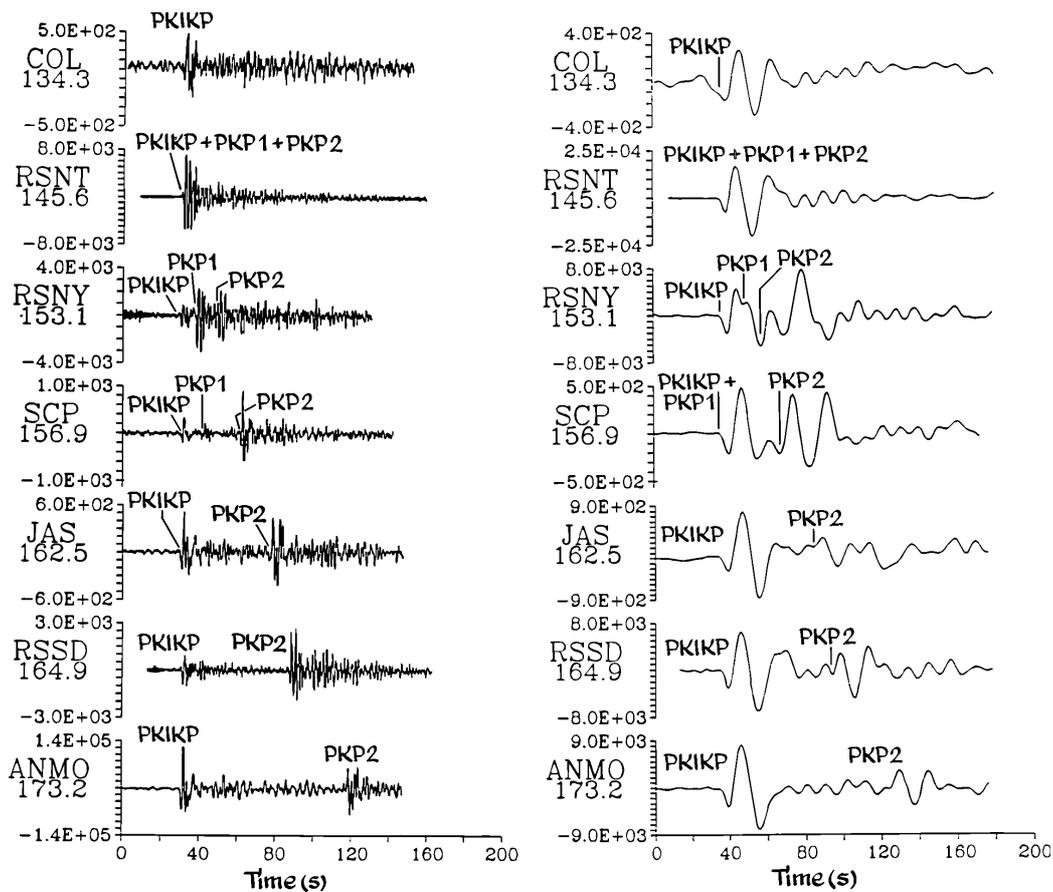


Fig. 2.45 Short-period (left) and long-period (right) seismograms for the Mid-Indian Rise earthquake on May 16, 1985 ($M = 6.0$, $h = 10$ km) in the range $D = 145.6^\circ$ to 173.2° . (From Kulhánek, *Anatomy of Seismograms*, plate 55, pp. 165-166, ©1990; with permission from Elsevier Science). Note: The figure above gives still the old names of the core phases. According to the new IASPEI phase names PKP2 should be replaced by PKPab, PKP1 by PKPbc and PKIKP by PKPdf (see IS 2.1, also for the detailed ray tracing of these phases).

The first discernable motion of a seismic phase in the record is called the *arrival time* and the measurement of it is termed *picking* of the arrival (see 11.2.2). Up to now, arrival time picking and reporting to international data centers is one of the major operations of data analysts at seismic stations or network centers. Plotting the time differences between reported arrival times and calculated origin times over the epicentral distance, seismologists were able to construct travel-time curves for the major phases and to use them to infer the average radial velocity structure of the Earth (see 2.7). In Fig. 2.46 (left) more than five million travel-time picks, archived by the International Seismological Centre (ISC) for the time 1964 to 1987, have been plotted. Most time picks align nicely to travel-time curves, which match well with the travel-time curves theoretically calculated for major seismic phases on the basis of the IASP91 model (Fig. 2.46 right).

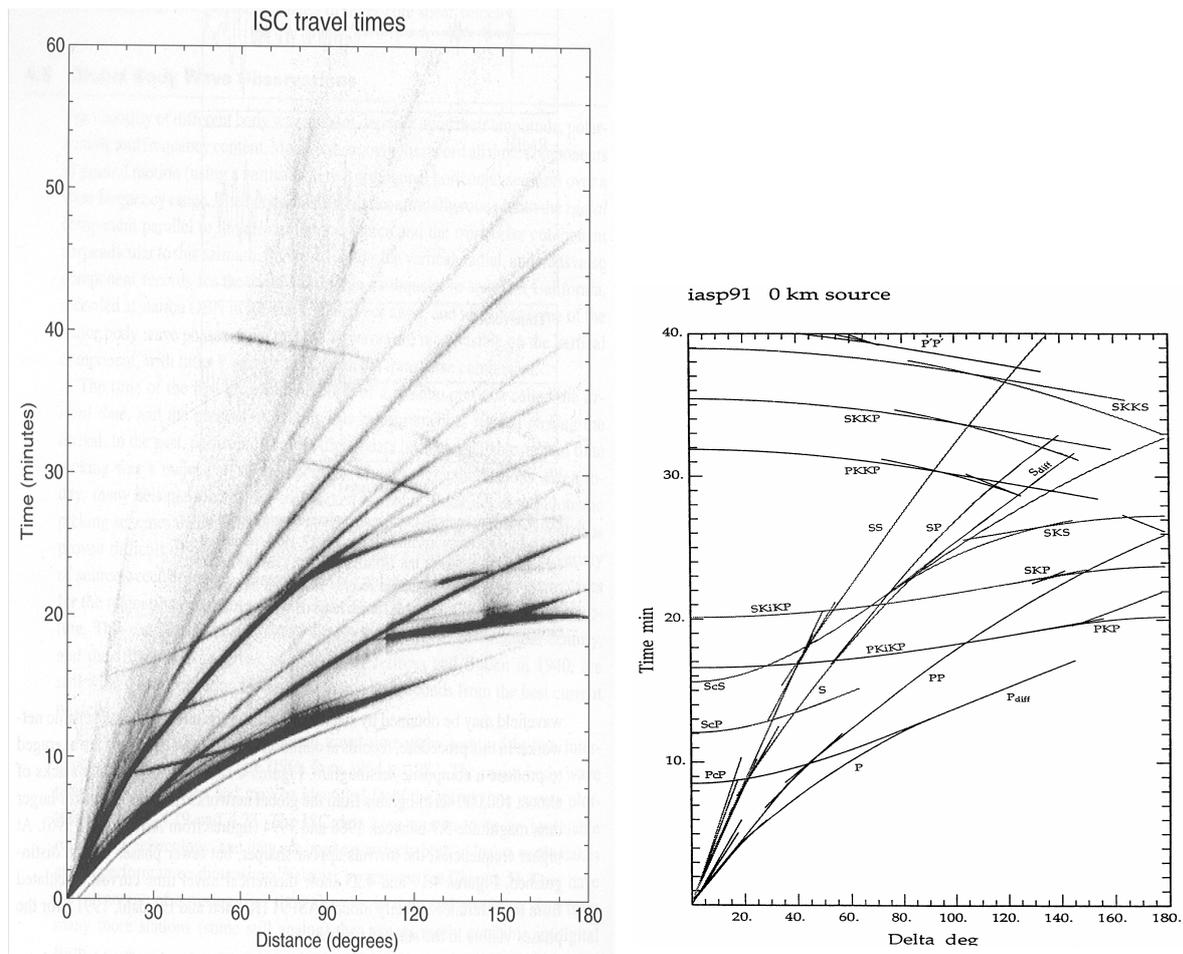


Fig. 2.46 Left: Travel-time picks collected by the ISC between 1964 and 1987 for events shallower than 50 km. (From Shearer, *Introduction to Seismology*, 1999; with permission from Cambridge University Press). Right: IASP91 travel-time curves for surface focus (from Kennett, 1991).

An even more complete picture of the entire seismic wavefield may nowadays be obtained by stacking data from modern digital seismic networks. For this, records at common source-receiver ranges are averaged to produce a composite seismogram. Stacks of almost 100,000 seismograms from the global digital networks are plotted in Fig. 2.47 (for short-period records with periods $T < 2\text{s}$) and Fig. 2.48 (for long-period records with $T > 10\text{s}$). Although the arrivals appear sharper at higher frequencies, much fewer later phases can be distinguished in short-period records. The late arriving reflected core phases P'P' (PKPPKP), PKKP, however, and higher multiples of them, are discernable in short-period records only. Note that the relative darkness with which the “curves” appear against the gray background is a measure of the relative frequency with which these phases can be observed above the noise level. The transparent overlays to the figures give the nomenclature for the visible phases in these stacks together with the more complete calculated travel-time curves according to the IASP91 velocity model (Kennett and Engdahl, 1991). They match very well.

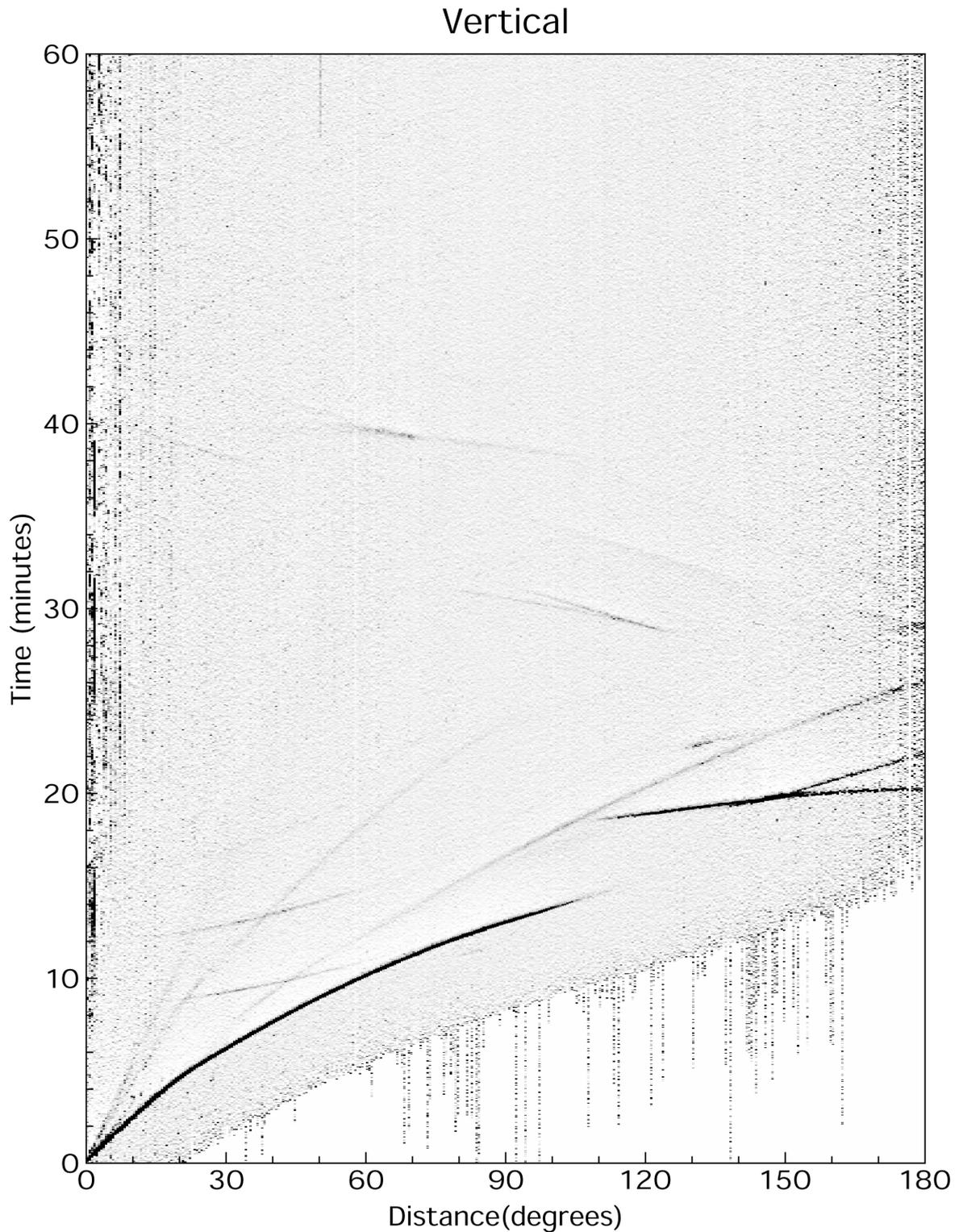
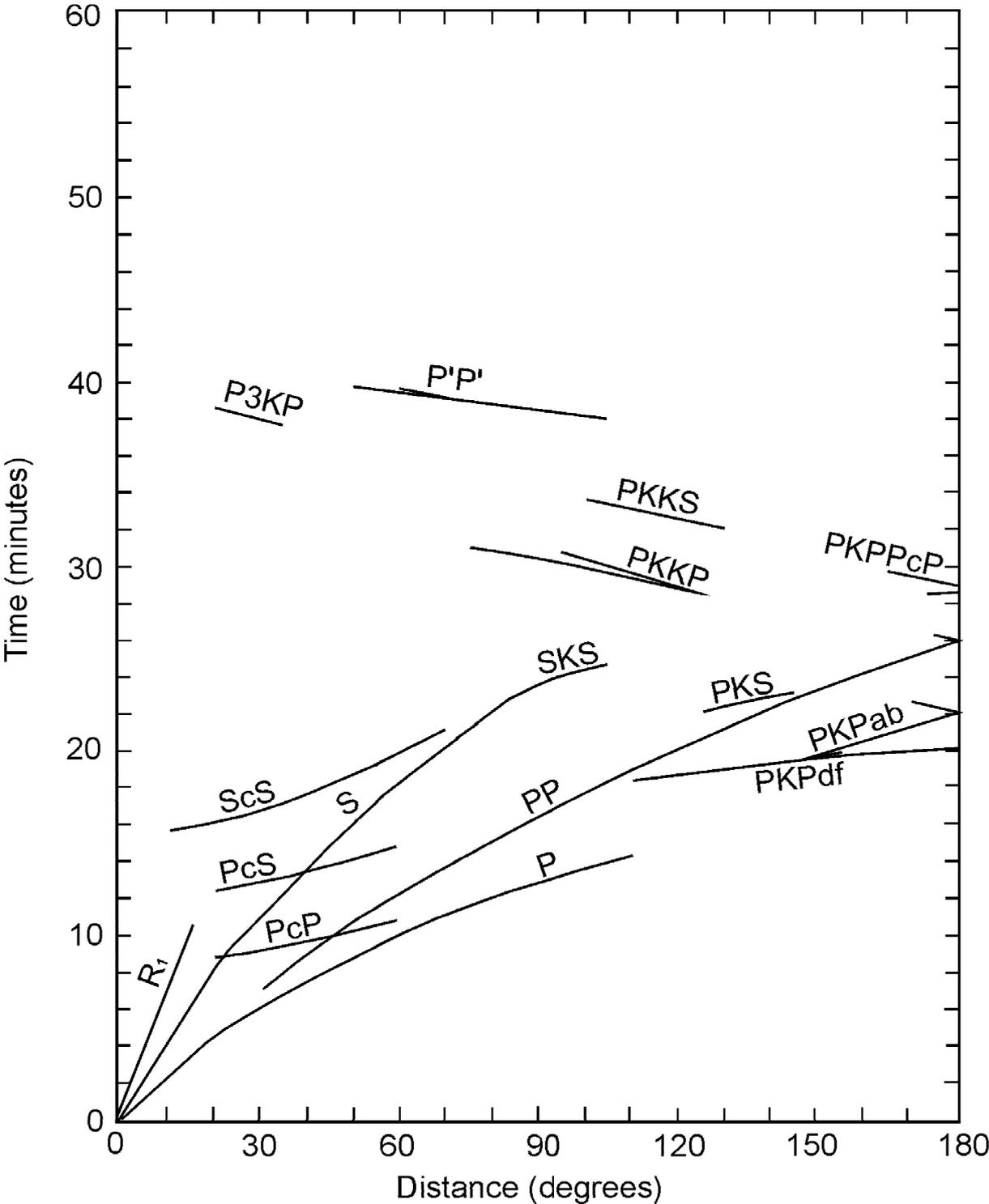


Fig. 2.47 A stack of short-period filtered (<2 s), vertical component data from the global networks between 1988 and 1994. See the overlay for the phase names and for the travel-time curves calculated using the IASP91 model (Kennett and Engdahl, 1991) (from Astiz et al., *Global Stacking of Broadband Seismograms*, *Seismological Research Letters*, Vol. 67, No. 4, p. 12, © 1996; with permission of Seismological Society of America).



Overlay to Fig. 2.47

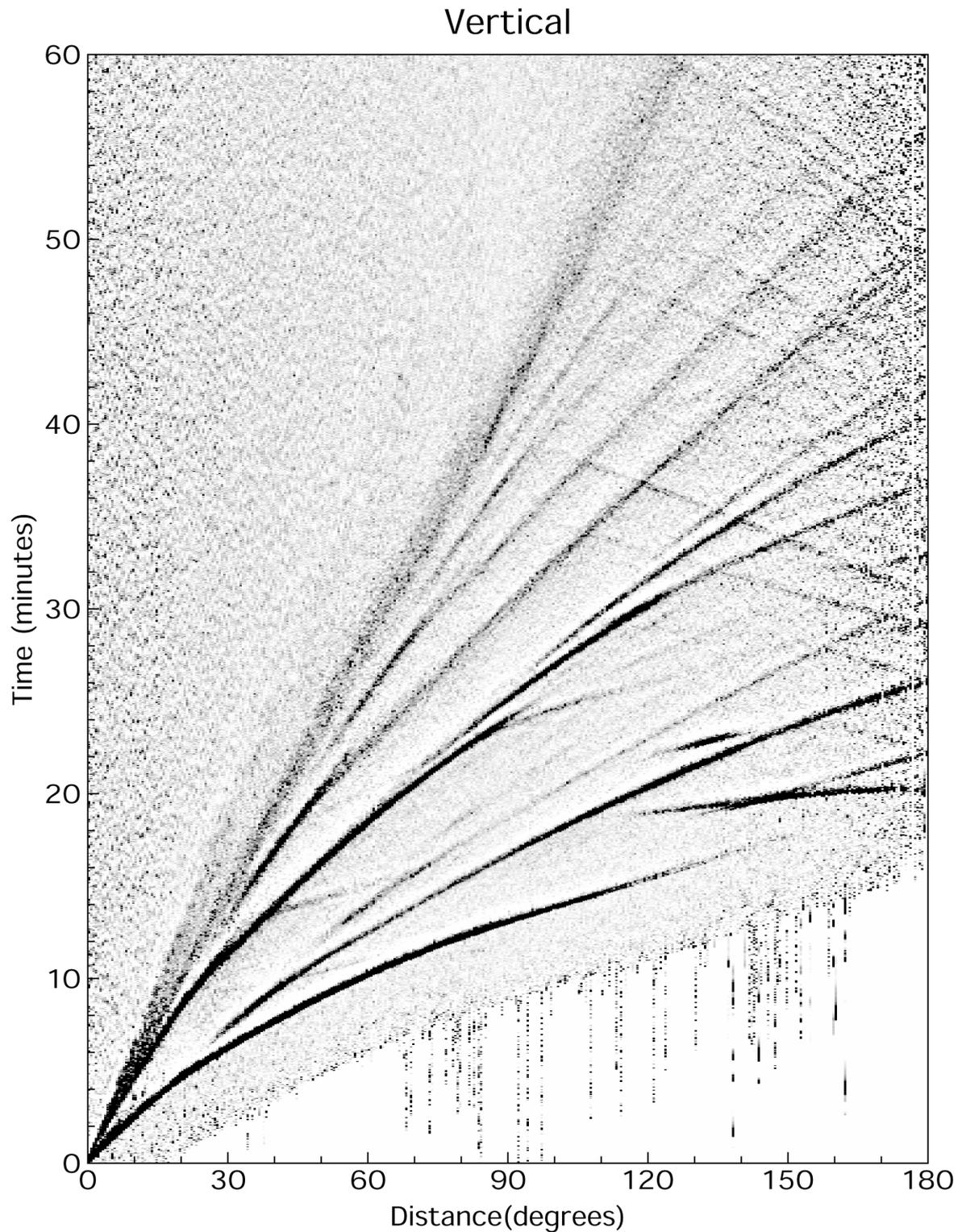


Fig. 2.48 A stack of long-period (> 10 s), vertical component data from the global networks between 1988 and 1994. See the overlay for the phase names and for the travel-time curves calculated for all types of phases (see also Fig. 2.49) using the IASP91 model (Kennett and Engdahl, 1991) (from Astiz et al., *Global Stacking of Broadband Seismograms*, *Seismological Research Letters*, Vol. 67, No. 4, p. 14, © 1996; with permission of Seismological Society of America).

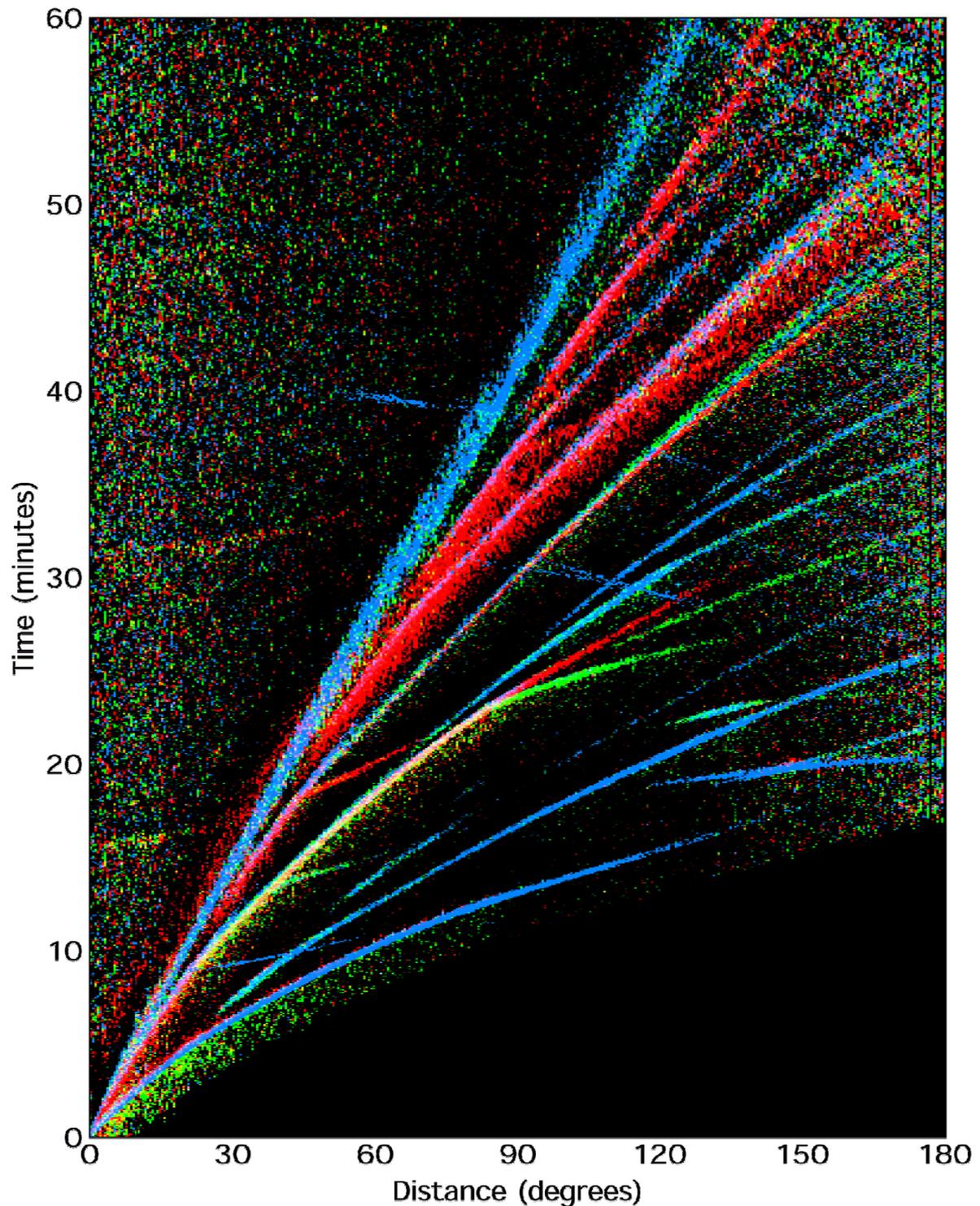


Fig. 2.49 Global travel-time curves for shallow earthquakes as produced by stacking broadband seismograms. Seismic phases are shown in different colors depending on their polarization (blue: vertical motion; green: radial-horizontal; and red: transverse-horizontal) (courtesy of L. Astiz).

Additionally, Fig. 2.49 reveals that the polarization of these various phases differs. While all primary longitudinal phases and all from P or K to S converted phases and vice versa appear

on vertical and radial-horizontal components only, multiple reflected S waves, which lose with each reflection more and more of their SV polarized energy due to conversion into P (or K at the CMB), become more and more transversely polarized. Primary S, however, has significant energy on both horizontal components that are oriented either parallel to the backazimuth to the source (radial) or perpendicular to it (transverse). Direct P waves, polarized in the direction of ray propagation, have in the teleseismic range dominating vertical components because of their steep incidence angle, which gets smaller and smaller with increasing distance (see e.g., PKP phases). PP, P3 and higher multiples may, however, have significant energy in the radial component too. These examples illustrate that the visibility and discrimination of body wave phases in seismic records depends on their relative amplitude, polarization and frequency content. All of these criteria have to be taken into account, besides the differences in travel-times, when analyzing seismic records.

2.7 Global Earth models

In the first part of the 20th century travel-time models for seismic phases, empirically derived from historical data, were rudimentary at best. One of the earliest travel-time model, the Zoeppritz tables (Zoeppritz, 1907) were applied by Herbert Hall Turner in a version as published by Galitzin (1914) to locate earthquakes for the ‘Bulletin of the British Association of the Advancement of Science, Seismology Committee’ for the years 1914 until 1917. During the 1920s, Turner gradually expanded these tables for newly discovered phases and better phase observations, often suggested and derived by Beno Gutenberg. These Zoeppritz-Turner tables were in use to locate earthquakes for the International Seismological Summary (ISS) from 1918 to 1929. This situation greatly improved with the introduction of the Jeffreys-Bullen (J-B) tables (Jeffreys and Bullen, 1940), which provided a complete, remarkably accurate representation of P, S and other later-arriving phases. Like the Gutenberg-Richter travel-time tables, the J-B tables were developed in the 1930s using reported arrival times of seismic phases from a sparse global network of stations, many of which often had poor time-keeping. Once the travel times of the main phases had been compiled, smoothed empirical representations of these travel times were inverted using the Herglotz-Wiechert method to generate a velocity model. The travel times for other phases were then determined directly from the velocity model. As a testament to the careful work that went into producing the J-B tables, they are still being used by the International Seismological Centre (ISC) and by the U. S. Geological Survey National Earthquake Information Center (NEIC) for routine earthquake location.

Although the limitations of the J-B tables were known for some time, it was not until the early 1980’s that a new generation of models was constructed in a completely different way. Instead of establishing smoothed, empirical representations of phase-travel times, inverse modeling was used to construct one-dimensional models for structure that fit phase travel times reported in the ISC Bulletin since 1964 and other parametric data. The Preliminary Reference Earth Model (PREM) of Dziewonski and Anderson (1981) was the most important member of this generation of new global 1-D models. However, PREM was constructed to fit both body-wave travel-time and normal-mode data, so it was not generally thought to be especially useful for earthquake location. In fact, soon afterwards Dziewonski and Anderson (1983) published a separate analysis of just P waves in an effort to produce an improved travel-time table.

2. Seismic Wave Propagation and Earth models

In 1987 the International Association of Seismology and Physics of the Earth's Interior (IASPEI) initiated a major international effort to construct new global travel-time tables for earthquake location and phase identification. As a result of this effort two models were developed: IASP91 (Kennett and Engdahl, 1991); and SP6 (Morelli and Dziewonski, 1992). Although differences in predicted travel times between these two models were small, some effort was still required to reconcile the travel times of some important, well-observed seismic phases before either of these models could be used by the ISC and NEIC for routine earthquake location. The upper mantle part of the IASP91 model was fitted to summary P and S wave travel times, binned in 1° intervals of epicentral distance, published by Dziewonski and Anderson (1981, 1983) (Fig. 2.50).

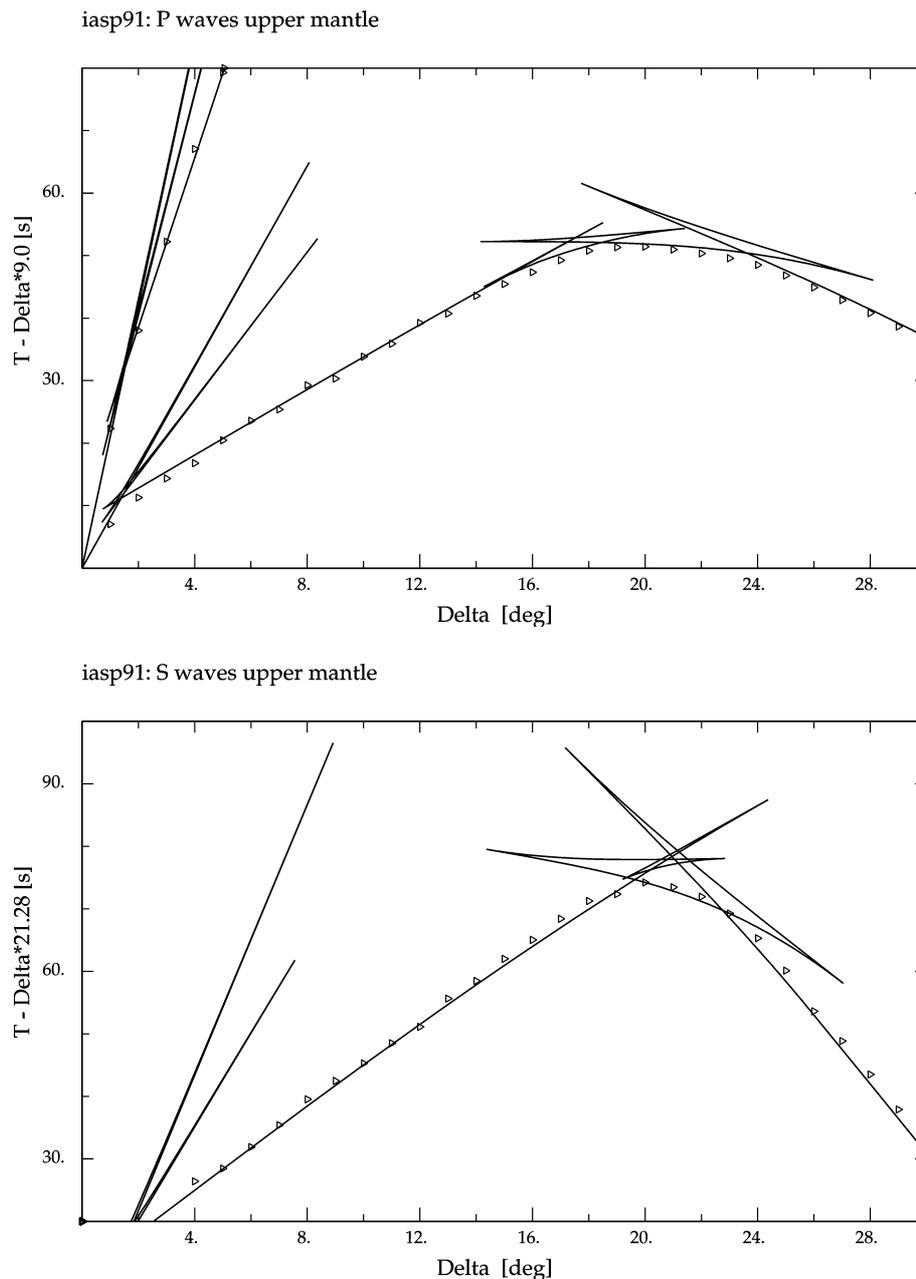


Fig. 2.50 Fitting of IASP91 upper mantle travel times as a function of epicentral distance to the summary first-arrival travel times of P (top) and S waves (bottom) according to Dziewonski and Anderson (1981, 1983) in time-reduced presentation (from Kennett and Engdahl, 1991).

As shown in Fig. 2.51, the IASP91 upper mantle differed substantially from PREM and, in particular, IASP91 had no mantle low-velocity zone for either P or S waves. Although this did run counter to the prevailing ideas about upper mantle structure, it did have a practical advantage for locating events because the upper mantle travel times in IASP91 were not discontinuous. Characteristics of the main upper mantle discontinuities were also different from previous models. In IASP91 the 210 km discontinuity was essentially absent. The 410 km and 660 km discontinuity velocity jumps in IASP91 were slightly greater in amplitude than in PREM. Path coverage was generally more uniform in the lower mantle, so these parts of the IASP91 P and S models were considered to be more representative of the average Earth. P structure was reasonably well constrained, except near the core-mantle boundary, but the complication of interfering phases put a practical limit on the amount and quality of data constraining S structure. Nevertheless, IASP91 seems to have done a reasonably good job of representing teleseismic travel times, as indicated by the analysis of arrival-time data from well-constrained explosions and earthquakes (Kennett and Engdahl, 1991).

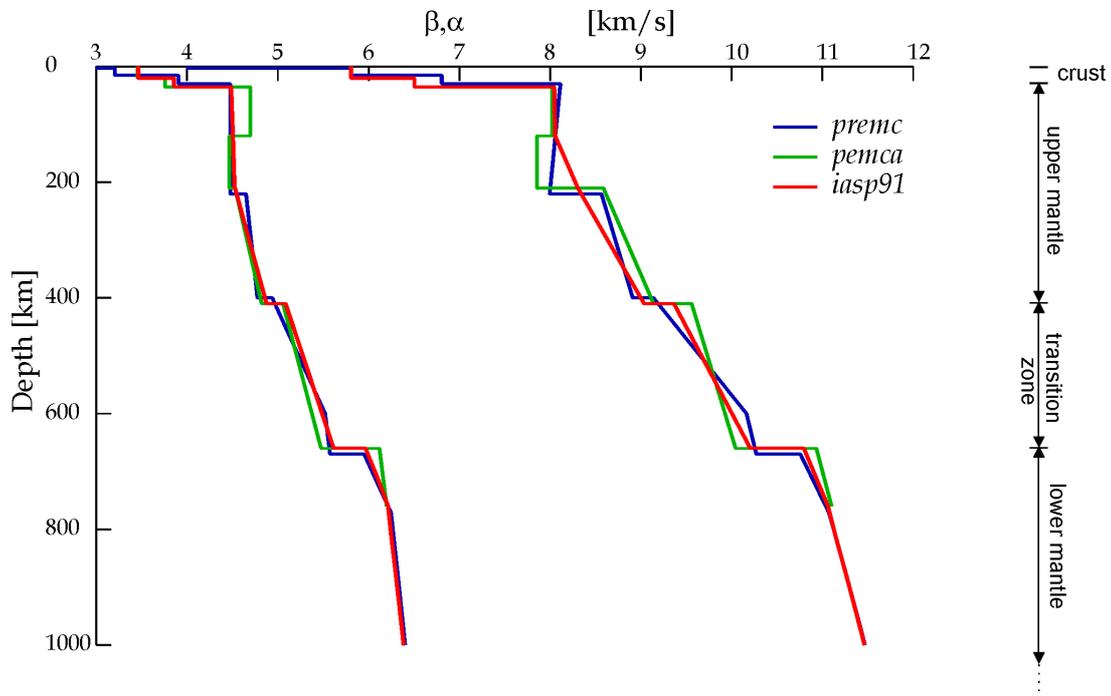


Fig. 2.51 Comparison of upper mantle velocity models for IASP91, PEMCA, and PREM (from Kennett and Engdahl, 1991;). Left: β - speed of S wave; right: α - speed of P wave.

Morelli and Dziewonski (1993) developed an alternative model (SP6) using the same model parameterization and upper mantle model as Kennett and Engdahl (1991). In their approach, they solved for multiple source-region station corrections averaged over 5° areas to account for lateral heterogeneity in an approximate manner. They then derived new sets of summary travel times for lower mantle and core P and S phases binned in 1° intervals of epicentral distance, and inverted those summary times for 1-D P and S velocity models. Although lower mantle P and S in the resulting model was generally comparable to IASP91, the models differed in that SP6 had slightly lower velocity gradients with depth and correspondingly higher velocity jumps at the 660-km discontinuity. Moreover, SP6 had incorporated a pronounced negative velocity gradient in the D'' region, a layer 100-150 km thick just above

2. Seismic Wave Propagation and Earth models

the core-mantle boundary. The SP6 model fitted the S data and all core-phase observations significantly better than IASP91. The differences in the seismic velocities between the models were significant for the core, owing to the addition of substantial core phase data in the construction of SP6.

The most significant differences between these new models and the older J-B travel-time model are in the upper mantle and core. The upper mantle is highly heterogeneous. Hence, velocities and major discontinuities in the upper mantle of recent models such as IASP91 and SP6 are set at values, which give an effective average representation of velocities for waves traveling out to 25° (see Kennett and Engdahl, 1991). The core models for IASP91 and SP6 predict more accurately than the J-B model the observed travel times of later-arriving core phases bottoming in the lowermost part of the outer core.

These models also resolve a long-standing problem in that the relocation of nuclear tests using the J-B travel-time model results in incorrect estimates of the origin times of nuclear explosions by about -1.8 sec. This error will propagate into all derived travel times and may affect the procedure of phase association. Kennett and Engdahl (1991) resolved this error in the absolute travel time (or "baseline" error) by fitting the IASP91 model to the mean teleseismic residual estimated from the origin times and hypocenters reported for explosions and well-constrained earthquakes by "test event" contributors. As a result, the times of teleseismic P and S waves for the IASP91 model now appear to be in better agreement with the travel time data than the times predicted by the J-B model. The IASP91 model has been adopted as the global reference model for the International Data Centre in Vienna established under the 1996 Comprehensive Nuclear-Test-Ban Treaty (CTBT).

Subsequently, Kennett et al. (1995) began with the best characteristics of the IASP91 and SP6 models and sought to enhance the data quality by improving the locations of a carefully selected set of geographically well-distributed events. The basic strategy was to use a location algorithm developed by Engdahl et al. (1998) with a IASP91 model modified to conform to the SP6 core to relocate events and improve phase identifications using only first arriving P phases and re-identified depth phases (pP, pwP and sP). The resulting set of smoothed empirical relations between travel time and epicentral distance for a wide range of re-identified seismic phases was then used to construct an improved reference model for the P and S radial velocity profile of the Earth (AK135). A composite residual plot (Fig. 2.52) shows that the model AK135 provides a very good fit to the empirical times of 18 seismic phases. The baseline and trend of S is well presented and most core phase times are quite well matched. Thus, for improved global earthquake location and phase association, there has been convergence on effective global, radially symmetric P- and S-velocity Earth models that provide a good average fit to smoothed empirical travel times of seismic phases.

The primary means of computing travel times from such models is based on a set of algorithms (Buland and Chapman, 1983) that provide rapid calculation of the travel times and derivatives of an arbitrary set of phases for a specified source depth and epicentral distance. In the mantle, AK135 differs from IASP91 only in the velocity gradient for the D" layer and in the baseline for S wave travel times (about -0.5 sec). Significant improvement in core velocities relative to earlier model fits was also realized. Inner core anisotropy, as discussed in the literature, is not yet accounted for in any of the newer 1-D Earth models. However there are so few reported arrivals of PKPdf at large distances along the spin axis of the Earth that the effects of this anisotropy in earthquake location are negligible.

The model AK135 has since been used for further reprocessing of the arrival time information (Engdahl et al., 1998). The reprocessed data set and the AK135 reference model have formed the basis of much recent work on high-resolution travel-time tomography to determine three-dimensional variations in seismic wave speed (e.g., Bijwaard et al., 1998). However, it is important to recognize that none of these models can properly account for the effect of lateral heterogeneities in the Earth on teleseismic earthquake location. Most deeper than normal earthquakes occur in or near subducted lithosphere where aspherical variations in seismic wave velocities are large (i.e., on the order of 5-10%). Such lateral variations in seismic velocity, the uneven spatial distribution of seismological stations, and the specific choice of seismic data used to determine the earthquake hypocenter can easily combine to produce bias in teleseismic earthquake locations of up to several tens of kilometers (Engdahl et al., 1998). For a review of recent advances in teleseismic event location, with the primary emphasis on applications using one-dimensional velocity models such as AK135, the reader is referred to Thurber and Engdahl (2000). The most accurate earthquake locations are best determined using a regional velocity model with phase arrival times from a dense local network, which may differ significantly (especially in focal depth) from the corresponding teleseismic locations.

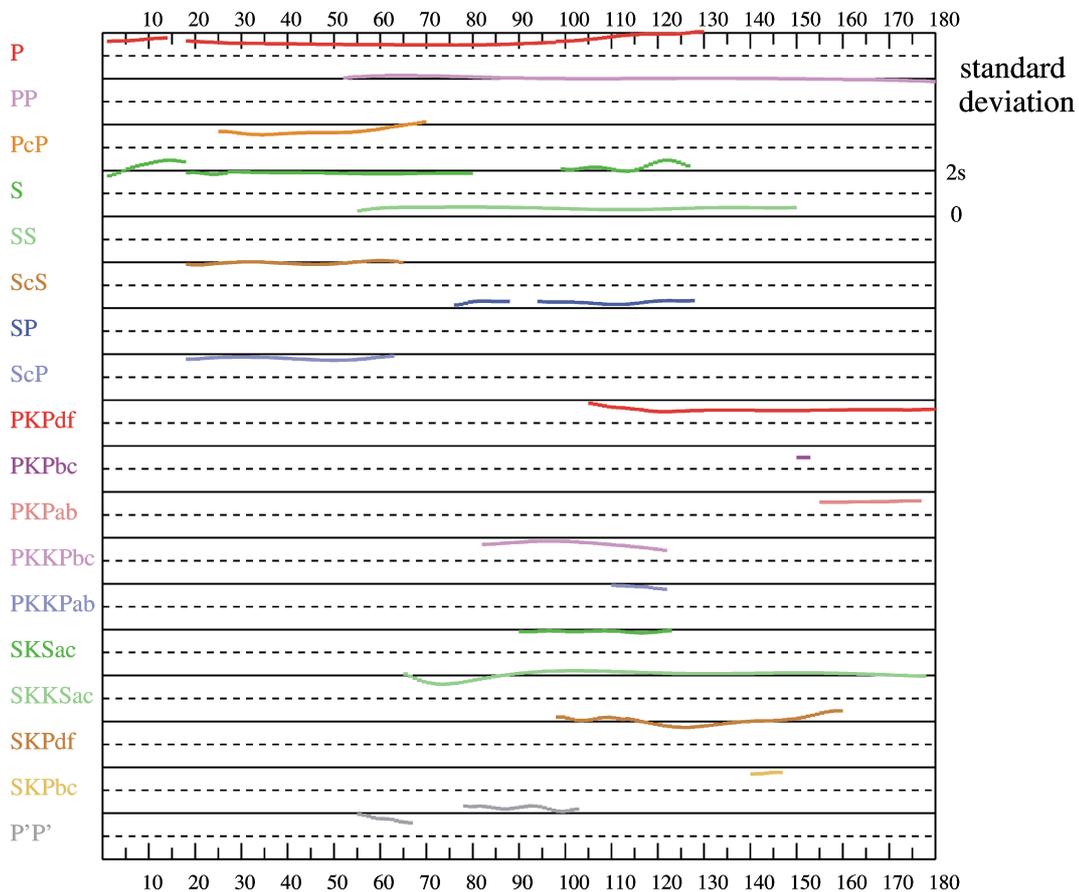


Fig. 2.52 Composite display of the estimates of standard deviations for the empirical travel times used in the construction of the AK135 velocity model (Kennett et al., 1995).

The AK135 wave speed reference model is shown in Fig. 2.53. However, though the P- and S-wave speeds are well constrained by high-frequency seismic phases, more information is needed to provide a full model for the structure of the Earth.

2. Seismic Wave Propagation and Earth models

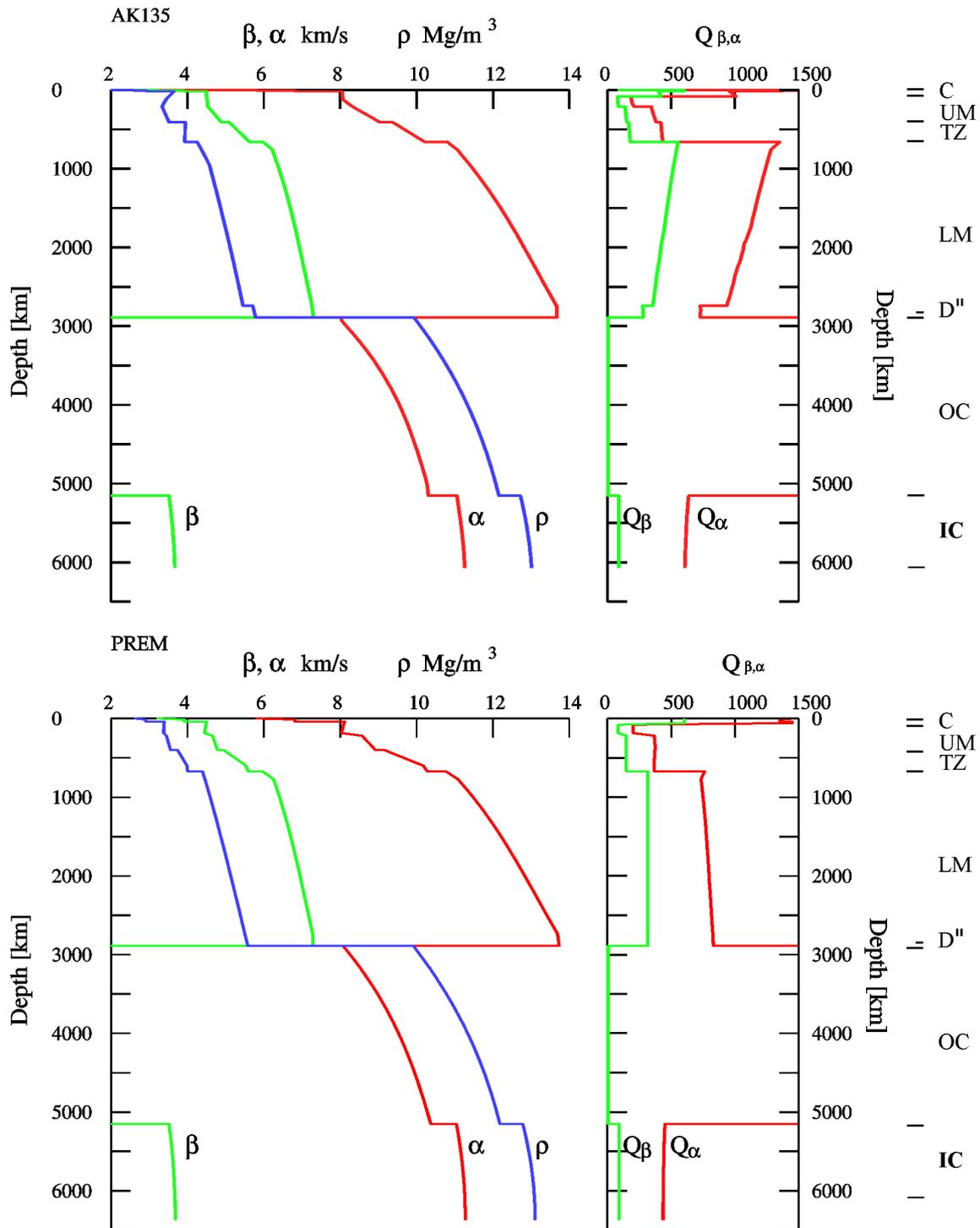


Fig. 2.53 Radial symmetric reference models of the Earth. Top: AK135 (seismic wave speeds according to Kennett et al. (1995), attenuation parameters and density according to Montagner and Kennett (1996); Bottom: PREM (Dziewonski and Anderson, 1981). α - and β : P- and S-wave velocity, respectively; ρ - density, Q_{α} and $Q_{\beta} = Q_{\mu}$ - “quality factor” Q for P and S waves. Note that wave attenuation is proportional to $1/Q$. The abbreviation on the outermost right stand, within the marked depth ranges, for: C – crust, UM – upper mantle, TZ – transition zone, LM – lower mantle, D''-layer, OC – outer core, IC – inner core.

In particular, any reference model should also include the density and inelastic attenuation distributions in the Earth. Work by Montagner and Kennett (1996) provided these parameters which, although known less precisely than the seismic velocities, are needed because it makes the model suitable for use as a reference to compute synthetic seismograms (see 2.8) without requiring additional assumptions. Nevertheless, the primary use of AK135 (and IASP91) remains earthquake location and phase identification. The PREM model of Dziewonski and Anderson (1981), also shown for comparison in Fig. 2.53, forms the basis for many current studies on global Earth's structure using quantitative exploitation of seismic waveforms at longer periods. It is the objective of the 'Working Group on Reference Earth models' in the 'IASPEI Commission on Earth Structure and Geodynamics' to retrieve a new 1-D reference Earth model for many depth-depending parameters which is also in agreement with observations of the Earth's normal modes.

The IASPEI 1991 Seismological Tables (Ed. Kennett, 1991) are now out of print. The more recent global P- and S-wave velocity and density model AK135, and the related body wave travel-time tables and plots are available via <http://rses.anu.edu.au/seismology/ak135/intro.html> and can be downloaded or printed in postscript. Additionally, software for travel-time routines and for corrections of the ellipticity of the Earth can be obtained via <http://rses.anu.edu.au/seismology/ttsoft.html>.

2.8 Synthetic seismograms and waveform modeling

A good measure of the advancement made by a scientific discipline is its ability to predict the phenomena with which it is dealing. One of the goals of seismology, as stated already over a hundred years ago by Emil Wiechert, is to understand every wiggle on the seismogram. This requires, as sketched in Fig. 1.1 of Chapter 1, an understanding and quantitative modeling of the contributions made to the seismic record (the output) by the various subsystems of the complex information chain: the source effects (input), the propagation effects (medium), the influence of the seismograph (sensor) and of the data processing. It is possible nowadays to model each of these effects quite well mathematically and thus to develop procedures for calculating *synthetic seismograms*. While the modeling of the seismometer response (see Chapter 5) and of the source effects (see 3.5 and IS 3.1) have been outlined in more detail in this Manual, it is beyond the scope of a handbook on observatory practice to go into the depth of wave propagation theory. Here we have to refer to pertinent textbooks such as Aki and Richards (1980 and 2002), Kennett (1983, 2001, 2002), Lay and Wallace (1995), Dahlen and Tromp (1998) or, for some condensed introduction, to Shearer (1999). Below we will only sketch some of the underlying principles, refer to some fundamental approaches, discuss their potential and shortcomings and give a few examples of synthetic seismogram calculation and waveform modeling for near and teleseismic events.

Based on advanced theoretical algorithms and the availability of powerful and fast computers the calculation of synthetic seismograms for realistic Earth models is becoming more and more a standard procedure both in research and in advanced observatory routines. Such calculations, based on certain model assumptions and parameter sets for the source, propagation path and sensor/recorder are sometimes referred to as the solution of the *direct* or *forward problem* whereas the other way around, namely, to draw inferences from the observed data itself on the effects and relevant parameters of propagation path and source is termed the *inverse problem* (see Fig. 1.1). With the exception of a few specialized cases of direct analytical solutions to the inverse problem (such as using the Wiechert-Herglotz

2. Seismic Wave Propagation and Earth models

inversion (Eqs. (2.21) and (2.22)) for calculating the velocity-depth distribution of the medium from the observed travel-time curves), most inverse problems are solved by comparing synthetic data with observed ones. The model parameters are then changed successively in an iterative process until the differences between the observed and the synthetic data reach a minimum. The procedure of comparing synthetic and observed seismograms is known as *waveform modeling*. It can be used in routine practice for better identification of seismic phases and more reliable onset-time picking in case of noisy data. Additionally, more and more advanced seismological data centers, such as NEIC, now make use of waveform fitting for fast seismic moment tensor and other source parameter solutions, such as source depth (see 3.5.6.1).

The underlying mathematical tool for constructing synthetic seismograms is the *linear filter theory*. The seismogram is thus treated as the output of a sequence of linear filters, each accounting for relevant aspects of the seismic source, propagation path and sensor/recorder. Accordingly, the seismogram $u(t)$ can be written as the result of convolution of three basic filters, namely:

$$u(t) = s(t) * g(t) * i(t), \quad (2.35)$$

where $s(t)$ is the signal from the seismic source, $g(t)$ is the propagation filter, and $i(t)$ is the overall instrument response. These basic filters can in fact be broken down into various sub-filters, each accounting for specific effects of the source (such as source radiation directivity, source-time function), the propagation medium (such as structure and attenuation) or the instrument (such as sensor and recorder). This makes it possible to study in detail the effects of a specific parameter or process on the character of the seismogram, e.g., the effects of the shape and bandwidth of the seismograph response on the recording (see 4.2) or of the source depth, rupture orientation or time-history of the rupture process on the signal shape (see pp. 400-412 in Lay and Wallace, 1995). With respect to the propagation term in Eq. (2.35) it may be modelled on the basis of a full wave-theoretical approach, solving Eq. (2.5) for 1-D media consisting of stacks of homogeneous horizontal layers. The complete response of such series of layers may be described by matrixes of their reflection and transmission coefficients and a so-called *propagator algorithm* (Thomson, 1950 and Haskell, 1953) or by generalized reflection and transmission coefficients for the entire stack as in the reflectivity method by Fuchs and Müller (1971), Kennett (1983), Müller (1985). Another, ray theoretical approach (e.g., Červený et al., 1977; Červený, 2001) is possible when assuming that variations in the elastic parameters of the media are negligible over a wavelength and thus these gradient terms tend to zero at high frequencies. While pure ray tracing allows one only to model travel-times, the assumption of so-called "*Gaussian beams*", i.e., "ray tubes" with a Gaussian bell-shaped energy distribution, permits the modeling of both travel-times and amplitudes and thus to calculate complete synthetic seismograms also for non-1-D structures. While a decade ago limited computer power allowed one to model realistically only relatively long-period teleseismic records, it is now possible to compute complete short-period seismograms of up to about 10 Hz or even higher frequencies. Several program packages (e.g. Fuchs and Müller, 1971; Kind, 1978; Kennett, 1983; Müller 1985; Sandmeier, 1990; Wang, 1999) permit one to compute routinely for given source parameters and, based on 1-D Earth models, synthetic seismograms for both near field and teleseismic events.

Two examples of synthetic seismogram sections in reduced travel-time presentation are shown below. Fig. 2.54 shows records for the local/regional distance range between 50 and 350 km with P, S and surface waves in the frequency range between about 0.5 and 2 Hz. Fig. 2.55 compiles synthetic records for longitudinal and some converted phases with frequencies between about 0.1 and 0.3 Hz in the teleseismic distance range between 32° and 172°. The

earth-flattening approximation of Müller (1977) is used to transform the flat layered model into a spherical model. This approximation does not permit calculation of phases travelling close to the center of the Earth. The theoretical record sections are noise-free and have simpler waveforms than most real seismograms, owing to the assumption of a simple source function. Fig. 2.54 does not show signal-generated codas of scattered waves that are so typical for short-period records of local events.

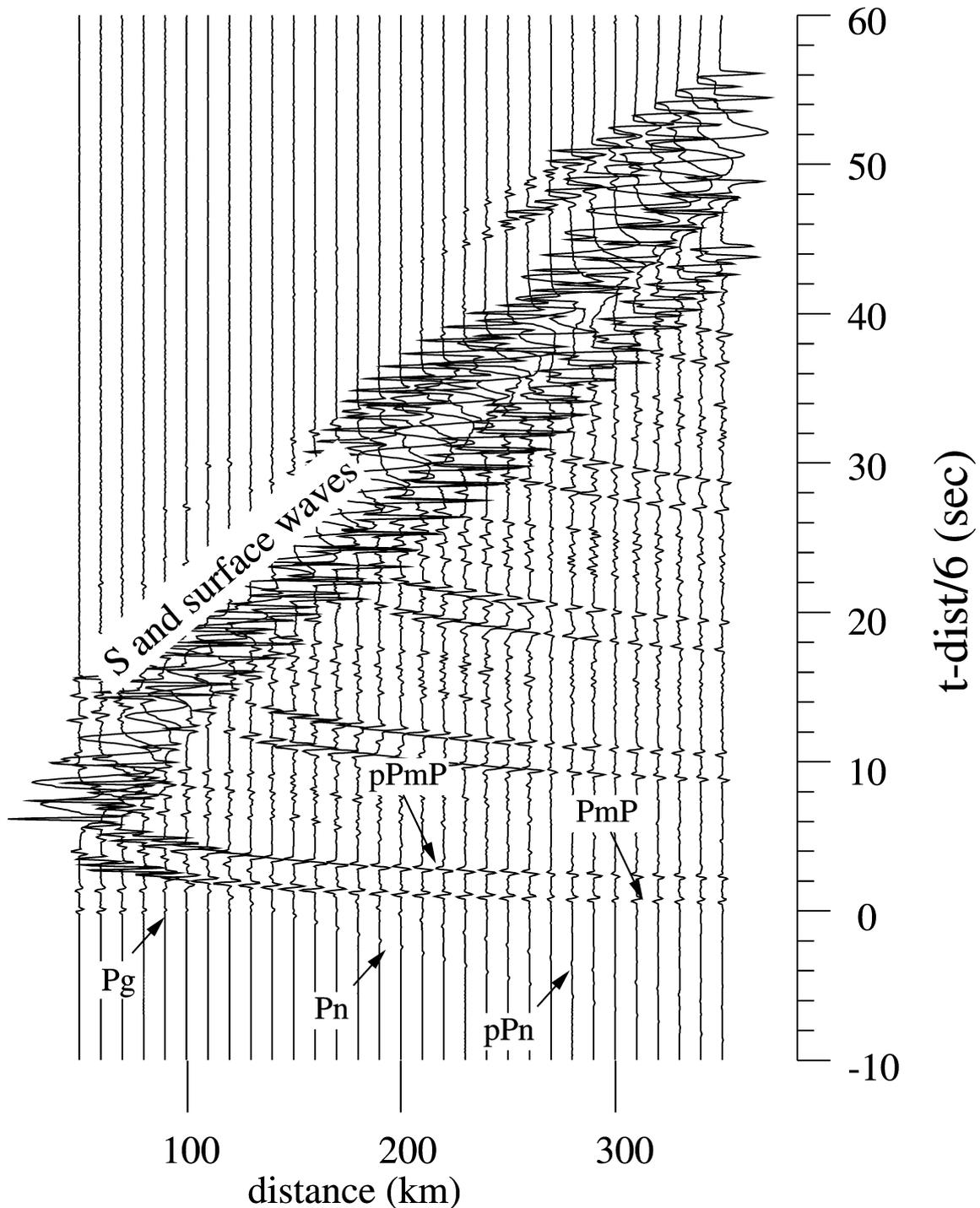


Fig. 2.54 Synthetic seismogram sections in the distance range 50-350 km, calculated for a hypothetical explosive source at 6 km depth in a homogenous single layer crustal model of 30 km thickness. For the calculation the program by Kind (1978) was used.

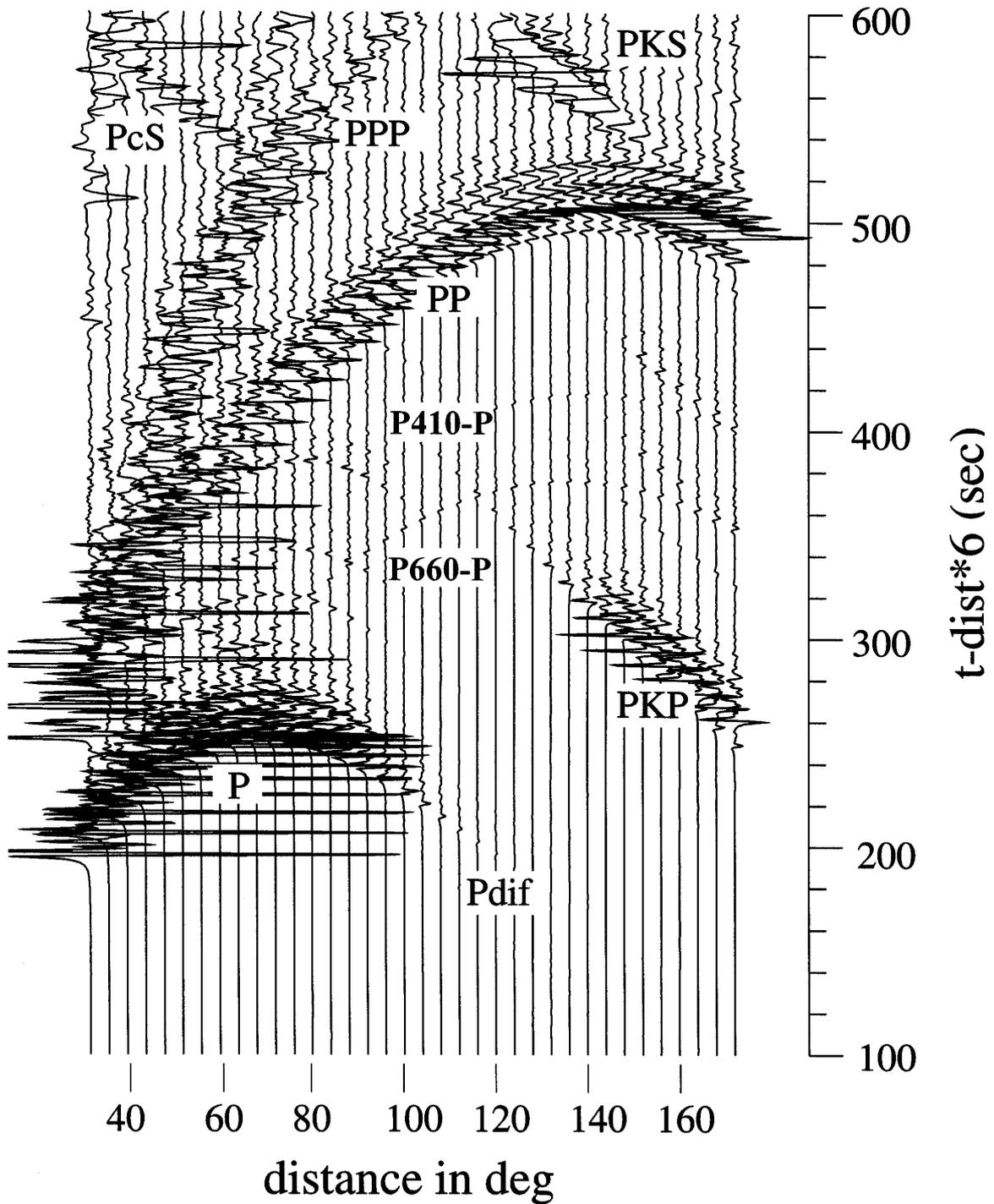


Fig. 2.55 Long-period synthetic seismic record section for the epicentral distance range 36°-166°, assuming a surface explosion and wave-propagation through the IASP91 model (Kennett and Engdahl, 1991). For the calculation the program by Kind (1978) was used.

2.8 Synthetic seismograms and waveform modeling

The synthetic record sections shown in Figs. 2.54 and 2.55 provide some general insights into basic features of seismograms in these two distance ranges such as:

- the overcritical Moho reflections PmP have the largest amplitudes in the P-wave part of near seismic recordings, with maximum amplitudes near the critical point around 70 km;
- Pg is the first arrival up to about 140 km (for a crustal thickness of 30 km) with amplitudes decaying rapidly with distance in this simple model example;
- since the travel-time curve of PmP approaches that of Pg asymptotically for larger distances, it may be difficult to separate Pg from Pm in real Earth for distances larger than about 100 km (see Fig. 2.40);
- Pn takes over as first arrival beyond about 140 km with generally rather weak amplitudes and higher apparent velocity;
- Sg (and in case of shallow events also surface waves, e.g., Rg) has (have) much larger amplitudes than the various types of direct, refracted or reflected P waves in records of local/regional events;
- the core shadow due to the strongly reduced P-wave velocities in the outer core is indeed clearly developed in epicentral distances between about 100° and 140°, however, long-period diffracted P waves may still be observable as relatively weak first arrivals up to 120° and more;
- PP is the first strong wave arrival in the core shadow range and, if Pdif or the weak inner-side reflections of P from the 660km or 410 km discontinuities (phase names P660-P and P440-P, respectively) are buried in the noise, PP can easily be misinterpreted as P-wave first arrival;
- the caustic of PKP around 145° produces very strong amplitudes comparable to those of P between about 50° to 70°;
- the branching of PKP into three travel-time branches beyond the caustic is well reproduced in the synthetic seismograms;
- converted core reflections (PcS) and converted core refractions (PKS) may be rather strong secondary later arrivals in the P-wave range between about 35°-55° and in the core-shadow range between about 120°-140°, respectively.

The following figures illustrate the potential of waveform modeling. Depth phases are not only very useful for determining the focal depth from teleseismic records, they are also frequently observed at regional distances and permit accurate depth determinations. Fig. 2.56 shows the ray paths for the phases Pn, pPn, sPn and sPP in a single layer crust from an event at depth h , as recorded in the distance range beyond 150 km, when Pn appears as the first arrival. Fig. 2.57 (left) shows the theoretical seismograms for all these phases at a distance of 210 km and as a function of source depth. It is easy to identify the depth phases. Fig. 2.57 (right) presents a compilation of the summation traces of all available vertical component records of the Gräfenberg array stations for the 1978 Swabian Jura (Germany) earthquake (September 3, 05:09 UT; $M_l = 6.0$) and for several of its aftershocks. All these events have been recorded at an epicentral distance of about 210 km. Depth phases sPn were observed in most records. From the correlation of sPn in neighboring traces it becomes obvious that the source depth migrated within 5 hours from the main shock at $h = 6.5$ km to a depth of only about 2-3 km for the aftershock at 10:03 UT.

2. Seismic Wave Propagation and Earth models

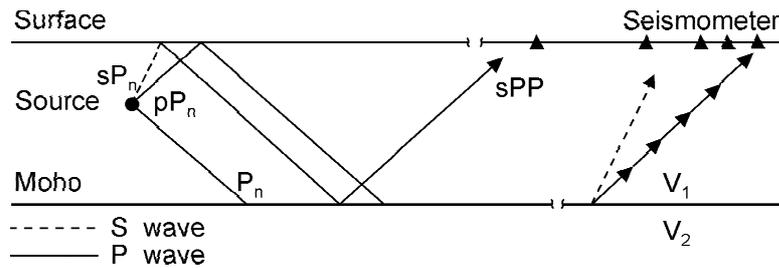


Fig. 2.56 Ray path of the sPn phases.

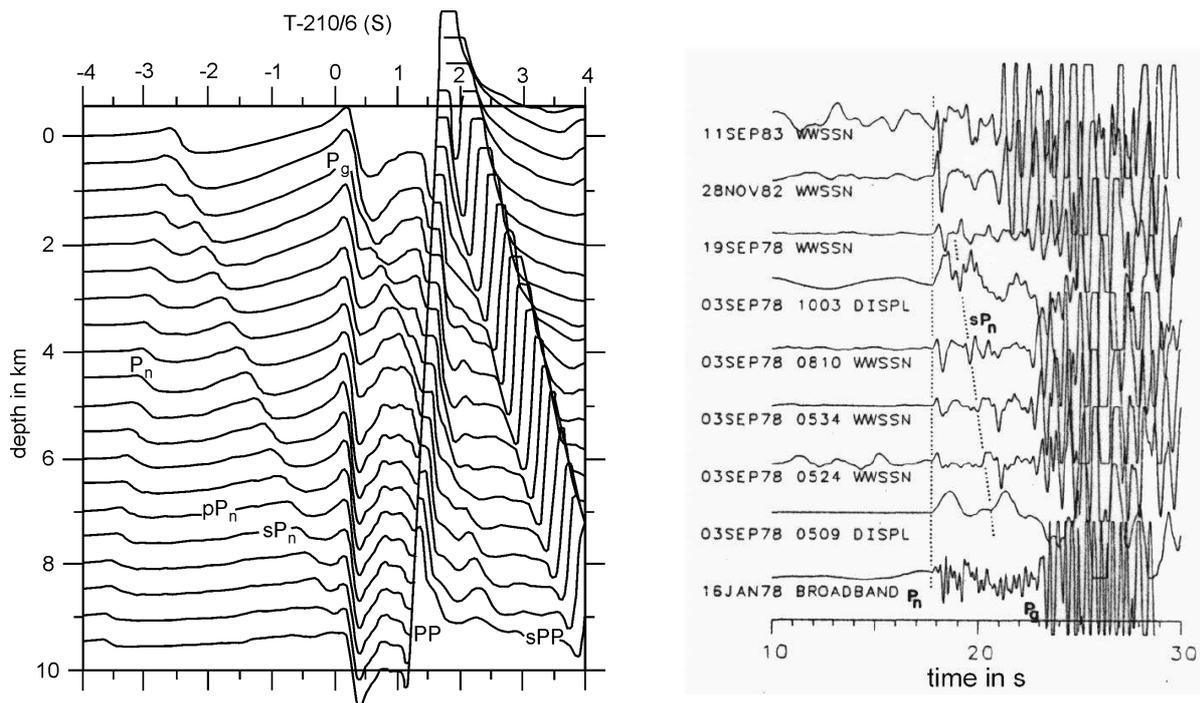


Fig. 2.57 **Left:** Theoretical seismograms in reduced travel-time presentation at 210 km epicentral distance as function of source depth for a single-layer crust (as in Fig. 2.56) of 30 km thickness. A clear depth phase sPn is recognizable between Pn and Pg; **right:** Gräfenberg records of Swabian Jura events in southern Germany. Epicentral distance is 210 km. Between the Pn and Pg arrival, a clear depth phase sPn can be observed. These observations indicate that after the main shock on September 3 at 05:09 the aftershocks migrated from 6.5 km depth to 2-3 km depth within 5 hours (from Kind, 1985).

Langston and HelMBERGER (1975) studied the influence of hypocenter depth h , type of source mechanism, source-time function and of stress drop on seismic waveforms. The superposition of P, pP and sP, which follow close one after another in the case of crustal earthquakes, make it difficult to separate these individual phases properly in more long-period teleseismic records and to pick the onset times of the depth phases reliably. However, because of the pronounced changes in the waveform of this P-wave group as a function of depth, one may be able to constrain also the source depth of distant earthquakes rather well by waveform modeling with an accuracy of about 5 km. On the other hand, one should be aware that there is a strong trade-off between source depth and the duration of the source-time function. A deeper source with source function of shorter duration may be similar to a shallower source

2.8 Synthetic seismograms and waveform modeling

with a longer source function. For simple sources, broadband data may help to overcome much of this trade-off. For complex source functions, however, these may trade-off with differences in source depth if only data from single stations are available. Using data from several stations instead could reduce this problem.

Generally, waveform modeling is much more powerful than first-motion focal mechanism determinations (see 3.4) in constraining fault orientation. Even with only a few stations and limited azimuthal coverage around the source superior results may be achieved. This is of particular importance for a fast determination of source parameters. Additionally, by comparing predicted and observed amplitudes of waveforms, the seismic moment can be determined rather reliably (see 3.5). Fig. 2.58 shows an example of waveform modeling in the teleseismic distance range for records of the 1989 Loma Prieta earthquake in different azimuth around the source. From the best fitting synthetics, the source-time function, fault strike ϕ , dip δ , rake λ and seismic moment M_0 were estimated. However, Kind and Basham (1987) could show that even with the broadband data from only one teleseismic station good estimates of fault depth, strike, dip and rake could be derived from waveform modeling.

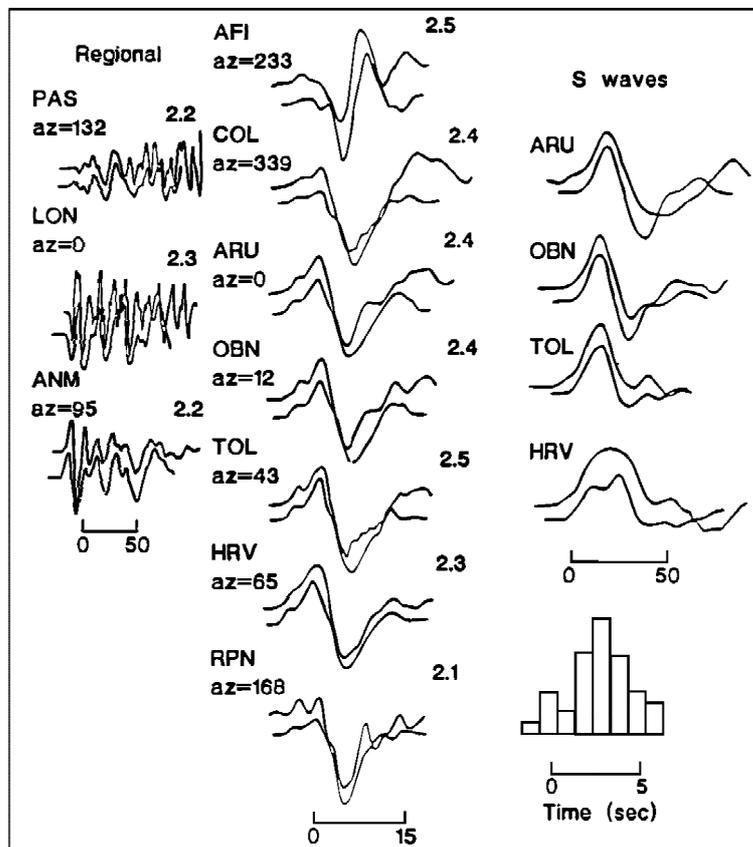


Fig. 2.58 Results of waveform modeling for the 1989 Loma Prieta earthquake. Depicted are the pairs of observed (top trace) and synthetic waveforms (bottom trace) for long-period Pn (left column), teleseismic P (middle column) and SH waves (right column). The time function used is shown at the lowermost right side. From the inversion of these data the following source parameters were determined: $\phi = 128^\circ \pm 3^\circ$, $\delta = 66^\circ \pm 4^\circ$, $\lambda = 133^\circ \pm 7^\circ$, and the moment $M_0 = 2.4 \times 10^{19}$ Nm (reproduced from Wallace et al., 1991, A broadband seismological investigation of the 1989 Loma Prieta, California, earthquake: Evidence for deep slow slip?, Bull. Seism. Soc. Am., Vol. 81, No. 5, Fig. 2, page 1627; 1991; © Seismological Society of America).

Acknowledgments

The authors thank P. Shearer, P. Malischewski and J. Schweitzer for careful proof-reading and many valuable suggestions. Thanks go also to M. Baumbach, B.L.N. Kennett, H. Neunhöfer, B. Schurr, J. Schweitzer and K. Wylegally for making some figures or related data for their production available.

Recommended overview readings (see References under Miscellaneous in Volume 2)

Aki and Richards (1980 and 2002)
Bullen and Bolt (1985)
Chapman (2002)
Kennett (2001 and 2002)
Lay and Wallace (1995)
Lognonne and Clevede (2002)
Sato et al. (2002)
Shearer (1999)