1. Introduction

In what follows, we attempt to review progress made in the last few decades in the analysis of teleseismic and regional surface wave data for the retrieval of earthquake source parameters and global and regional Earth structure. This review is by no means exhaustive. We will rapidly skip over the early developments of the 1950s and 1960s that led the foundations of normal mode and surface wave theory as it is used today. We will not attempt to provide an exhaustive review of the vast literature on surface wave measurements and the resulting models, but rather focus on describing key theoretical developments that are relevant and have been applied to inversion. Since surface wave theory is closely related to that of the Earth’s normal modes, we will discuss the latter when appropriate. However, we make no attempt to extensively review normal mode theory, as this subject is addressed in a separate contribution (see Chapter 10 by Lognonné and Clévédé).

2. Background

Most of the long-period energy (periods greater than 20 s) generated by earthquakes and recorded at teleseismic distances propagates as surface waves. Most clearly visible on long-period seismograms are the successive, Earth-circling, dispersed wave trains of the fundamental mode. For moderate size earthquakes recorded at teleseismic distances ($M \sim 5.5$), only the surface waves propagating along the direct great circle path between the epicenter and the station have significant signal-to-noise ratio, mostly between 20 and 100 s period, and the dispersive and attenuative properties of these wave trains have been used extensively, since the 1950s to infer crust and upper mantle structure in different regions of the Earth. For earthquakes of magnitude 7 or larger, successive Earth-circling surface wave trains can be followed for many hours (Fig. 1), and are then either analyzed individually or combined to produce a spectrum of the Earth’s free oscillations by Fourier analysis of long time-series.

Most studied are fundamental mode Rayleigh waves, which correspond to $P-SV$ energy and have elliptical particle motions in the vertical plane containing the direction of propagation. These waves are well recorded on the quieter vertical component seismographs (Fig. 2). On the other hand, Love waves, which carry $SH$ energy, and are polarized horizontally in a direction perpendicular to the direction of propagation, require rotating the two horizontal records to extract the transverse component of motion. Love wave studies have suffered,
especially in the early days of analog recordings, from the more complex data processing required, and from the higher levels of background noise on horizontal components at long periods, due primarily to the influence of atmospheric pressure variations, inducing ground tilts. Fundamental mode Love and Rayleigh waves are generally well separated from other phases on the seismograms, and well excited by shallow, crustal earthquakes, while overtones travel at higher group velocities, appear as packets of mixed overtones (e.g., X phases, Jobert et al., 1977), and are better excited by deeper earthquakes (Fig. 3).

Surface waves recorded at teleseismic distances contain information about both the characteristics of the earthquake source and the structure of the Earth along the source station path. Separating these two effects has been one of the long-standing challenges faced by seismologists. Studies of the structure of the crust and upper mantle progressed rapidly in the 1950s and early 1960s, as the tools to measure group and phase velocities, and interpret them in terms of layered mantle and crust models, became readily available (e.g., Ewing et al., 1957; Brune et al., 1961a,b; Alsop, 1963). In these studies, source effects were generally eliminated by considering propagation between two or more stations aligned along the same great circle path, or, at longer periods, observation of consecutive Earth-circling wave trains at the same station. On the other hand, in early studies of earthquake sources, propagation effects were assumed to be known, and amplitudes were “equalized” to obtain the source radiation pattern and infer information about the fault orientation (e.g., Aki, 1960) and its directivity (Alterman et al., 1959; Ben-Menahem, 1961; Kanamori, 1970). At that time, the theoretical formulation for the excitation of surface waves and normal modes of the Earth was developed (Sato et al., 1962; Harkrider, 1964; Haskell, 1964), much stimulated by the occurrence of the great Chilean earthquake of 22 May 1960, and more quantitative studies of the effects of the earthquake source on spectra of surface waves followed. The association of a normal mode formalism (e.g., Gilbert, 1971) to compute dispersion and excitation of surface waves (and complete seismograms) with a moment tensor formalism to describe the earthquake source (e.g., Backus and Mulcahy, 1976; Mendiguren, 1977) has led to the present-day commonly used expressions and to a rapid development of source studies based on surface waves in the 1980s. A computational method (Takeuchi and Saito, 1972), following the theoretical approach of Saito (1967) based on Runge–Kutta matrix integration, has long been the main reference for the practical calculation of excitation for surface waves and normal modes in laterally homogeneous, elastic, flat or spherical Earth models. Later, different schemes, using different mathematical approaches (variational method) were developed (Wiggins, 1976; Buland and Gilbert, 1984). Today, another widely used code for spherical geometry and efficient to relatively short periods (10s) is based on a propagator matrix approach, in which minors of sets of solutions are used (Gilbert and Backus, 1966; Woodhouse, 1980a, 1988).
These theoretical advances were first applied to the analog data of the World Wide Standard Seismic Network (WWSSN) accumulated in the 1960s and 1970s. The IDA (International Deployment of Accelerometers) network, established in the mid-1970s (Agnew et al., 1976), provided the first long-period digital data, along with several stations installed and operated by the French (Jobert and Roult, 1976). The digital recording greatly facilitated the simultaneous analysis of many records, paving the way for large-scale tomographic studies of global structure and systematic teleseismic source studies. A major drawback, however, was the limited dynamic range of the IDA instruments, so that first-arriving low-frequency $R_1$ and $G_1$ wave trains would saturate for large earthquakes. This problem disappeared in the 1980s with the deployment by France of the high dynamic range, digital broadband GEO-SCOPE network (Romanowicz et al., 1984, 1991) and by the United States of the IRIS Global Seismic Network (e.g., Smith, 1986), gradually complemented by many broadband stations contributed by other countries through the Federation of Digital Seismic Networks (FDSN), which, starting in 1986, established high-level standards for broadband seismic sensors, recording systems, and data formats (e.g., Romanowicz and Dziewonski, 1986). The accumulation of high-quality data from numerous broadband stations has greatly contributed to the successes of global tomography and source moment tensor studies of the last 15 years.

3. Source Studies Using Surface Waves

To obtain the frequency spectrum of a single mode surface wave train from the expression of a seismogram obtained by summation of normal modes on a spherically symmetric Earth, one uses Poisson’s formula (e.g., Gilbert, 1976; Aki and Richards, 1980), which decomposes the modes into infinite trains of propagating surface waves traveling in opposite directions around the Earth. In this process, a high-frequency approximation is used, in which the phase velocity of a surface wave is related to the corresponding normal mode frequency by Jean’s formula:

$$C(\omega) = \frac{a\omega_i}{T + \sqrt{1/2}}$$

where $i$ is the angular order of the mode and $\omega_i$ its eigenfrequency, and $a$ is the radius of the Earth. In this high-frequency approximation, the surface waves propagate along the great circle path between the epicenter and the station and are sensitive to structure only along this great circle.

The most widely used theoretical framework for the interpretation of surface wave data was thus established in the 1970s. It is derived from a normal mode formalism using a high-frequency zeroth-order approximation, and leads to a simple expression for the spectrum of a single mode propagating surface wave at distance $\Delta$, azimuth $\theta$, and angular frequency $\omega$. Following Kanamori and Stewart (1976), and Nakanishi and Kanamori (1982):

$$U(\Delta, \theta, \omega) = U_s(\theta, \omega)S(\Delta)U_p(\Delta, \theta, \omega)F(\omega, \theta_0)D(\omega)I(\omega)$$  \hspace{1cm} (2)

where $U_s$ is the source spectrum, $U_p$ contains propagation effects, $I$ is the instrument response, $S(\Delta)$ the geometrical spreading term, and $F$ and $D$ express the source process as clarified below.

The propagation term $U_p$ can be expressed as (e.g., Romanowicz and Monfret, 1986)

$$U_p(\Delta, \theta, \omega) = \frac{1}{(\sin \Delta)^{1/2}} \exp(i\pi/4) \exp(im\pi/2) \times \exp[-i\omega\Delta/C(\omega, \theta)] \exp[-\eta(\omega, \theta)\Delta]$$  \hspace{1cm} (3)

where $m$ denotes the number of polar passages and $C(\omega, \theta)$, $\eta(\omega, \theta)$ are, respectively, the average phase velocity and attenuation coefficient along the source–station path.

The source spectrum $U_s$ is a linear combination of the moment tensor elements $M_{ij}$ of the source. In the notation of Kanamori and Stewart (1976), we have for Rayleigh waves:

$$U_s^R(\theta, \omega) = \frac{1}{2}(S_R + N_R)M_{zz} + \frac{1}{2}(2N_R - S_R)(M_{xx} + M_{yy}) + \frac{1}{2}P_R(M_{xy} - M_{yx}) \cos 2\theta - P_RM_{zz} \sin 2\theta + iQ_R(M_{zz} \cos \theta + M_{zz} \sin \theta)$$  \hspace{1cm} (4)

where $N_R$, $S_R$, $P_R$, $Q_R$ are the excitation functions, which are nonlinear functions of $\omega$ and of the depth $h$ of the source.

Likewise, for Love waves:

$$U_s^L(\theta, \omega) = P_L \left[ \frac{1}{2}(M_{xx} - M_{yy}) \sin 2\theta - M_{yy} \cos 2\theta \right] + iQ_L(-M_{zz} \sin \theta + M_{zz} \cos \theta)$$  \hspace{1cm} (5)

with depth-dependent excitation functions $P_L, Q_L$. $D(\omega)$ expresses the delay $\tau_D$ of the main faulting from the initial break (e.g., Nakanishi and Kanamori, 1982):

$$D(\omega) = \exp(-i\omega\tau_D)$$  \hspace{1cm} (6)

and $F(\omega, \theta_0)$ expresses the directivity term arising for the propagation of rupture. Ben Menahem (1961) derived an approximate expression for $F$ in the case of unilateral faulting, which has been shown to be exact if an equatorial coordinate system is used (Dziewonski and Romanowicz, 1977):

$$F(\omega, \theta_0) = \frac{\sin X}{X} \exp(-iX)$$  \hspace{1cm} (7)

where

$$X = \frac{\omega L}{2V} - \frac{\omega L}{2C} \cos(\theta - \theta_0)$$  \hspace{1cm} (8)

and $L$ is the length of fault, $C$ is the average phase velocity at frequency $\omega$, $V$ is the rupture velocity, and $\theta_0$ is the azimuth of the rupture direction with respect to the fault strike.
As seen from Eqs. (2) to (8), fundamental mode surface wave spectra contain information about the source moment tensor, source depth (centroid), source process time, and, under favorable circumstances for very large earthquakes, source directivity. However, propagation effects $U_{ij}$ have to be known or effectively eliminated.

In order to correct for propagation, different approaches need to be taken depending on the size of the earthquake and period range considered. At very long periods ($T > 180$ sec) and for large earthquakes ($M > 6.5$), propagation effects can be accounted for approximately in the framework of a spherically symmetric reference Earth model. At shorter periods and for smaller earthquakes, corrections on individual source–station paths need to be known much more accurately. We thus discuss these two domains separately.

### 3.1 Teleseismic Studies of Moderate Size Events ($M \sim 5–6$)

Today significant efforts are still expended toward surface wave “path calibration,” in particular in the framework of CTBT research (e.g., Stevens and McLaughlin, 2000). However, several methods have been developed to deal with poorly known path effects, with particular applications to the study of moderate earthquakes, for which good signal-to-noise records of 20–100 sec surface waves are obtained at teleseismic distances. In this case, source duration and directivity can generally be neglected, since their effect is significant only at shorter periods. A standard procedure is to use a “reference event,” for which a reliable mechanism has been obtained independently, to obtain path corrections to specific stations. These path corrections are then used to infer source parameters of other neighboring events of interest. Following the approach of Weidner and Aki (1973), Patton (1977) developed an iterative method to infer simultaneously depth, source mechanism, and propagation effects for a group of closely located events observed teleseismically. On the path to a given station, the propagation effects are assumed to be the same for all events. In adjacent steps of the procedure, path effects and source characteristics are alternately improved. In the source improvement step, the linear system in $M_j$ [Eqs. (4) and (5)] is solved successively for different depth values, and a solution is declared corresponding to the depth that provides the best fit to the data in a least-squares sense. The imaginary parts of Eqs. (4) and (5) generally do not contribute to the depth determination, as, for shallow earthquakes, the eigenfunctions corresponding to the moment tensor elements $M_{zz}$ and $M_{xz}$ are very small (and go to zero at the free surface). This also results in poor constraints on these two moment tensor elements for shallow ($h < 20$ km) earthquakes if only fundamental mode surface waves are used.

The method of Patton (1977) is limited by the fact that it requires several earthquakes located in the same area to be well recorded teleseismically at the same set of stations. Moreover, Aki and Patton (1979) have shown that, in order to obtain reliable moment tensor solutions from data in the period range 20–100 sec, it is necessary to know average phase velocities on source station paths with an accuracy of 0.5%. When path corrections in the phase are inaccurate, the method breaks down because a clear minimum in the residuals/depth curves cannot be defined. On the other hand, because path effects on amplitudes are less coherent, an average-$Q$ model is sufficient if only amplitude data are used (Mendiguren, 1977). Tsai and Aki (1971) first showed how the amplitude spectrum of Rayleigh waves (and to a lesser extent Love waves) contained the signature of depth of source in the period range 10–100 sec, in the form of a “hole” in the spectrum, which appears at a given period for a specific source mechanism. However, when amplitudes only are used, the inverse problem becomes nonlinear in the moment-tensor, requiring the knowledge of a reasonable starting solution.

A procedure that greatly relaxes the constraint of accurate path corrections in the phase was proposed by Romanowicz (1982a), based on the following observation: For a given source and at a given frequency $\omega_k$, the real ($\alpha$) and imaginary ($\beta$) parts in Eqs. (4) and (5) are functions of the azimuth $\theta$ only, and can be written in the form

$$\alpha(\omega_k, \theta) = A_k + B_k \cos 2\theta + C_k \sin 2\theta$$

$$\beta(\omega_k, \theta) = D_k \cos \theta + E_k \sin \theta$$

where $A_k$, $B_k$, $C_k$, $D_k$, and $E_k$ depend on frequency. By virtue of the uniqueness of the Fourier decomposition of continuous functions, at each frequency $\omega_k$, given a set of values of $\alpha$ and $\beta$ for different azimuths $\theta$, these coefficients are uniquely determined. The Fourier expansions of $\alpha$ and $\beta$ in azimuth contain other coefficients, of degree $n > 2$, that arise from imperfect knowledge of path corrections. In solving Eqs. (9), these other terms are eliminated. Thus only the very long-wavelength ($n \leq 2$) terms of the phase velocity “map” in the region containing sources and stations need to be known accurately, an increasingly reachable goal today, thanks to improvements in global surface wave tomography.

The inversion procedure thus proceeds in “two steps”: First, Eqs. (9) are solved at a set of frequencies $\omega_1, \ldots, \omega_k$, and second, the following system of equations is solved for depth $h$, searching, as previously, for a minimum in the residuals/depth curve. Thus, for example, for Rayleigh waves:

$$A_k = \frac{1}{\rho} S_\nu(\omega, h) M_{zz}$$

$$B_k = \frac{1}{\rho} P_\nu(\omega, h) (M_{yy} - M_{zz})$$

$$C_k = -P_\nu(\omega, h) M_{xy}$$

$$D_k = Q_\nu(\omega, h) M_{xz}$$

$$E_k = Q_\nu(\omega, h) M_{yz}$$

Here we have assumed that there is no volume change at the source, so that $\sum M_{ij} = 0$ and $M_{xz} + M_{yz}$ is replaced by
−Mzz in Eq. (2). While justified for most earthquakes, this assumption needs to be relaxed when a volumetric component of the source is sought. To resolve the latter requires multimode observations, and Eqs. (2) to (8) show that resolving Mxx, Myy, and Mzz separately involves many trade-offs with structure effects (e.g., Dufumier and Rivera, 1997). Romanowicz (1982a) showed that this procedure allows the determination of accurate source parameters for small events (M < 5.5), for which only surface wave observations are available seismically, provided one event of larger magnitude is available, with a well-constrained mechanism, in a source region ~1500 km in aperture. This method was later extended to the nonlinear inversion of surface wave amplitude data (Romanowicz and Suarez, 1983).

3.2 Global Studies of Large Earthquakes

At very long periods (150 < T < 320 sec) and for large earthquakes (M > 6.5), spherically symmetric Earth models can be used to correct for propagation effects in the phase of “mantle” waves (as surface waves are called in this period range), and the biases introduced by neglecting lateral heterogeneity are relatively small, except for very shallow events, for which the source phase needs to be known with greater accuracy. In the last 15 years, the availability of global 3D tomographic models of the upper mantle of increasing accuracy has not only facilitated very long-period source studies but has also made it possible to extend the period range to shorter periods (down to ~120 sec) and to efficiently use more alternative methodologies, based on time-domain waveform inversion, that are no longer restricted to the fundamental mode, thus providing more accurate estimation of the source depth. Such a waveform approach was introduced by Dziwonski et al. (1981), who combined waveforms of mantle waves at periods greater than 120 sec with waveforms of body waves at periods greater than 80 sec. This forms the basis of a new routine procedure that serves to construct the widely used Harvard centroid moment tensor (CMT) catalog.

The development of methodologies to invert fundamental mode mantle wave data has nevertheless continued, following Kanamori and Given (1981), who showed how the spectra of mantle waves, sampled at only a few frequencies, could be used to rapidly determine moment tensors of large earthquakes, when the depth of the event is known. Romanowicz and Guillemant (1984) extended the approach of Romanowicz (1982a) to show how centroid depth could be accurately determined using mantle waves even in a spherically symmetric reference Earth model. For the size of earthquakes considered here, the source process time (Nakanishi and Kanamori, 1982) cannot be neglected and there can be significant trade-offs with source depth. However, the source process time can be estimated, as a function of frequency, prior to the inversion by computing phase differences between three wave trains at each individual station (e.g., Furumoto, 1979; Furumoto and Nakanishi, 1982). For example, if Rayleigh wave trains R2n, R2m+1 and R2n+2 are used, with R2m+1 traveling in the opposite direction to the other two trains, we have, for the corresponding phases Φn:

\[ \Phi_{2n} + \Phi_{2m+1} + (n + m)(\Phi_{2n} - \Phi_{2n+2}) + 2\Phi_{\text{intra}} = -\omega^2 \tau + k2\pi \]

(11)

where k is an arbitrary integer and the source time τ is defined as [Eqs. (6), (7), and (8)]

\[ \tau = \frac{L}{V} + 2\pi \]

(12)

Expression (11) is independent of structure within the framework of the high-frequency approximation in which it is derived. This procedure requires that at least three mantle wave trains traveling in opposite directions have adequate signal-to-noise ratio, which can be rather restrictive. Romanowicz and Monfret (1986) proposed an approach that requires the availability of only one mantle wave train.

Noting from Eqs. (3) to (9) that, if \( U_s \) is the observed spectrum corrected for instrument response and propagation, then,

\[ U_s(\theta, \omega)e^{i\phi} = U_s(\theta, \omega) = \alpha + i\beta \]

(13)

where

\[ \Phi_{\tau} = \frac{\omega\tau}{2 + \delta\phi} \]

(14)

\( \delta\phi \) is the residual phase shift due to inaccurate propagation corrections in the phase, and \( \alpha \) and \( \beta \) are the real and imaginary parts of Eqs. (4) and (5), respectively. For specified, incremental values of \( \tau_j, j = 1, \ldots, p \) of \( \tau \), the system of Eqs. (4) and (5) is solved for different frequencies \( \omega_j \). The residual of this inversion for each frequency is summing to obtain a squared residual as a function of \( \tau \), and is found for the value of \( \tau \) corresponding to the minimum value of this residual. The second step of the inversion proceeds as described above [Eqs. (9) and (10)]. Unlike the three-train method of Furumoto and Nakanishi (1982), because of the trade-off with Earth structure as seen in Eq. (14), the determination of the source time by this method depends on the accuracy of the Earth model used, and specifically, on errors in the Earth model that contribute a constant phase shift as a function of azimuth, that is, long-wavelength features. A regionalized global model, such as that of Okal (1977), was shown to be sufficient to obtain stable source time estimates. With the increased precision of currently available 3D tomographic models, this is no longer an issue. This method can be extended to the case of large earthquakes with significant directivity, by including a parameter search over the azimuth \( \theta_0 \) of the fault and the rupture length L. For very large earthquakes, Kuge et al. (1996) showed that source complexity (spatiotemporal changes in the source mechanism) sometimes...
needs to be invoked to reconcile source mechanisms obtained separately using Love and Rayleigh waves. The spectral domain inversion of fundamental mode mantle waves was further tested against Earth models and applied to large earthquakes by Zhang and Kanamori (1988a,b), showing generally consistent results with the Harvard CMT solutions. An alternative approach to study the spatiotemporal characteristics of large earthquakes from surface wave data has been proposed by Bukchin (1995), who used a representation of the source in terms of higher-order moments.

### 3.3 Regional Distance Source Studies

While most of the modern formalism and methodology for intermediate-period surface wave inversion for source parameters was in place by the mid-1980s, they have only been recently adapted to the case of earthquakes observed at regional distances. This has been made possible by the rapid expansion, in the last ten years, of regional broadband networks in seismically active regions, such as, for example, TERRA-scpe (Thio and Kanamori, 1995) or the Berkeley Digital Seismic Network (BDSN; Romanowicz et al., 1993) in southern and northern California, respectively, MEDNET in the Mediterranean region (Giardini et al., 1993), and also in Japan. In the regional case, target frequent earthquakes have smaller magnitude (generally $M < 4.5$) and the period range of interest is $\sim$10–60 sec (Fig. 4). Here, an important aspect is to obtain information about the earthquake source rapidly to guide local emergency response, and this has led to the development of automated procedures (e.g., Pasyanos et al., 1993). The theoretical framework is essentially the same as was developed for teleseismic observations at intermediate (Romanowicz, 1982a) or long periods (Kanamori and Given, 1981; Romanowicz and Monfret, 1986). Some automatic procedures involve the comparison of a spectral domain inversion of fundamental mode surface waves with a time domain complete waveform inversion, at periods longer than 10 sec (e.g., Pasyanos et al., 1993).

With the improvement of global upper mantle models, it is now also possible to extend to smaller magnitudes ($M_W > 4.5$) the time-domain CMT inversion methodology developed by Dziewonski et al. (1981), making use of intermediate-period surface waves (Ritsema and Lay, 1993). Arvidsson and Ekström (1998) use a low-pass cutoff at $\sim$45 sec and consider the fundamental mode surface waves as traveling waves. The phase is corrected for propagation effects using recent global phase velocity maps.

### 4. Structure Studies Using Surface Waves

#### 4.1 Fundamental Mode Studies

Fundamental mode surface waves are well suited to study the elastic structure of the crust and upper mantle, which can be deduced from their group and/or phase dispersion properties. They allow the sampling of vast areas of the globe that are otherwise devoid of seismic stations and sufficiently strong earthquake sources, such as the oceans.

It is beyond the scope of this chapter to review the numerous studies that have used surface waves to infer lateral variations of seismic velocities in the crust and upper mantle, since the 1950s and up to this day. Many early studies documented the correlation of seismic velocity variations with surface tectonic features, using regional measurements of phase and group velocities of fundamental mode Love and Rayleigh waves in the period range 20–100 sec (e.g., Knopoff, 1972; Kovach, 1978) or at longer periods (e.g., Toksöz and Anderson, 1966; Dziewonski, 1971; Wu, 1972). Originally, group velocity dispersion was obtained by measuring the times of arrival $t(\omega)$ of peaks and troughs of waves in a dispersed wave train. The time between two successive peaks would give the half period $(T/2 = \pi/\omega)$, and the group velocity $U(\omega)$ would be computed as

$$U(\omega) = \frac{X}{t(\omega)}$$

where $X$ is epicentral distance in kilometers and $t(\omega)$ is measured with respect to the earthquake’s origin time. Since the early 1970s, the computation of “energy diagrams,” as

![FIGURE 4](image-url)
described below, has greatly facilitated group velocity dispersion measurements (e.g., Fig. 5).

On the other hand, phase velocity $C(\omega)$ is obtained from the phase $\Phi$ of the Fourier spectrum of a dispersed wave train that has been corrected for the contribution of the source and the instrument [Eq. (2)]:

$$C(\omega) = \frac{\lambda}{t_0} = \frac{\Phi - N - (m/4) - 1/8}{\omega}$$  \hspace{1cm} (16)

where $t_0$ is the start time of the Fourier window with respect to the event’s origin time, $m$ is the number of polar passages, and $N$ is an integer arising from the $2\pi$ indeterminacy of the phase. This integer is determined first at long periods to obtain reasonable values of phase velocity compatible with well-constrained global models (e.g., PREM; Dziewonski and Anderson, 1981). It is then successively obtained at decreasing periods in such a way as to obtain a smooth phase velocity curve. This can become a problem at periods shorter than 30 sec, where small variations in phase velocity correspond to rapid cycling of the phase. The source phase also needs to be accurately corrected for.

Most early studies circumvented the issue of separating source and propagating effects by making measurements using the “two-station method,” in which dispersion was measured between two stations approximately aligned with the epicenter on a great circle path, thus eliminating the common source phase. If, in addition, the two stations were located within a relatively homogeneous geological province, such measurements were called “pure path” and led directly to the determination of elastic velocity structure beneath that province. An extension of this method to “many stations” to infer structure beneath an array of stations spanning a geologically homogeneous region was also devised and used extensively in the 1960s (e.g., Knopoff et al., 1967). To access remote regions devoid of stations but with relatively frequent earthquakes of magnitude 5.5 or larger, such as the Tibet Plateau, a “two-event” method was also devised. The success of this approach, which relied on an accurate independent knowledge of the source phase, was only possible under very restrictive conditions. In particular, the alignment of the three points (two epicenters, one receiver) had to be within 3° of the great circle path (e.g., Romanowicz, 1982b).

**FIGURE 5** Example of group velocity dispersion diagrams obtained for the fundamental Rayleigh wave mode on the vertical record of the Chile M$_s$ 6.8 earthquake of 15 Oct. 1997 at Geoscope station INU, using a multiple filtering approach as perfected in the FTAN method (e.g., Lander, 1989; Levshin et al., 1994). (a) Dispersion diagram before filtering; (b) same after filtering. The group velocity dispersion curve obtained follows the maxima of energy as delineated by the gray scale contours. (c) Corresponding time-domain seismogram before and after variable filtering. The “raw” seismogram is bandpass filtered in the period 50–400 sec. (Courtesy of Boris Bukchin.)
At longer periods, especially beyond the group velocity minimum at 200–250 sec, fundamental mode dispersion measurements were hampered by interference with higher modes (a problem also in the vicinity of the group velocity maximum around 60 sec), and the fact that the longest-period waves are not sufficiently well dispersed on first-arriving trains. In the early 1970s sophisticated filtering techniques were developed (Dziewonski et al., 1969; Cara, 1973a,b) to isolate a surface wave mode along its group velocity curve (Fig. 5). This approach, later perfected by Levshin and collaborators in the FTAN method (e.g., Lander, 1989; Levshin et al., 1994), involves two steps. In the first step, an “energy diagram” is formed, in which the energy contained in the seismogram is plotted as a function of period and time (i.e., group velocity). The maxima of the 2D diagram thus obtained delineate the group velocity curve of the dispersed surface wave modes present in the seismogram in the time window and frequency window considered. In a second step, the time-domain seismogram is then filtered using multiple filters centered, at each point in time, on the frequency corresponding to the maximum of energy at that time. The possibility of simultaneously inverting group and phase velocity data has been discussed by Yanovsky and Ditmar (1990).

By allowing the extraction of a mode branch over many consecutive wave trains, these techniques also resulted in better measurements of the spheroidal and toroidal eigenfrequencies for the fundamental and several higher modes (e.g., Jobert and Roulit, 1976). With the advent of digital recording in the mid-1970s and the expansion of global digital long-period and later broadband networks, the processing of the relatively long time-series needed to measure surface waves became much easier and opened the way in the 1980s to large-scale and global studies of upper mantle structure.

Large-scale studies first proceeded according to a regionalization scheme, in which it was assumed that the depth variation of seismic velocities is the same throughout each tectonic province. These studies confirmed and extended to longer periods (and hence larger depths) early results on the age dependence of structure in the oceans (e.g., Forsyth, 1975; Montagner and Jobert, 1983; Nishimura and Forsyth, 1989).

The constraint of regionalization was soon relaxed, after Dziewonski et al. (1977) introduced the expansion of lateral heterogeneity into a basis of spherical harmonics (applied first, however, to the lower mantle using body wave travel time data).

On the global scale, as in source studies, two different approaches emerged, one based on dispersion measurements in the frequency domain, the other based on time-domain waveform inversions. Dispersion measurements in the frequency domain have been focused primarily on the fundamental mode, and thus limited in resolution beyond depths of 350–400 km. This approach has led to several generations of collections of global phase velocity maps at discrete frequencies. These maps can then be inverted jointly, in a second step, to infer 3D S-velocity structure of the upper mantle. This approach was pioneered by Nakanishi and Anderson (1983) and Nataf et al. (1986), and is now well established, with different research groups introducing various improvements in the measurement technique leading to increasingly high-resolution maps (e.g., Trampert and Woodhouse, 1995; Laske and Masters, 1996; Zhang and Lay, 1996; Ekström et al., 1997). Trampert and Woodhouse (1995) designed an automatic method to measure phase and amplitude of surface waves in the period range 40–150 sec and invert the resulting dataset to obtain global phase velocity maps expanded in spherical harmonics up to degree 40. On the other hand, Ekström et al. (1997) used phase-matched filters to isolate the fundamental mode and make global dispersion measurements in the period range 35–150 sec and obtain models of phase velocity also up to degree 40 in spherical harmonics. In these studies, lateral variations of structure seem to be well resolved down to wavelengths of ~1000–2000 km. At the long-period end the period range is limited by the difficulty of separating consecutive wave trains, and at the short period end by the increased complexity of surface wave propagation in the strongly heterogeneous crust and uppermost mantle, resulting in lateral refractions and multipathing. The latter are not taken into account by the simple high-frequency, great circle propagation assumptions underlying the interpretation procedures.

Still in the frequency domain, but aiming at exploiting both the real and imaginary parts of the spectrum, Dziewonski and Stein (1982) developed a method to retrieve both phase velocity and attenuation averaged over complete great circle paths. These authors considered the frequency-dependent transfer function $T(\omega)$ between two consecutive fundamental mode Rayleigh wave trains traveling in the same direction along the great circle linking the epicenter and the station (say, $R_n$ and $R_{n+2}$) and related it to the average dispersion and attenuation over the great circle path:

$$T(\omega) = \frac{R_{n+2}}{R_n}$$

$$T(\omega) = \exp\left\{-\pi a \left[ i \left( 2k(\omega) - \frac{1}{a} \right) + \frac{\omega q(\omega)}{2U(\omega)} \right] \right\}$$

(17)

where $k(\omega) = \omega / C$ is the real wavenumber, $U(\omega)$ is group velocity, $q(\omega)$ is the inverse of the quality factor, and the extra term $i\pi$ accounts for polar passages. The transfer function $T(\omega)$ is linearized and compared to that predicted by a reference model, to retrieve perturbations in phase velocity $C$ and attenuation $q$ along the great circle.

In order to proceed to a time-domain waveform inversion, synthetic seismograms needed to be computed in a reference model and perturbed. First-order perturbation theory had been developed in the 1970s, culminating in the work of Woodhouse and Dahlen (1978) and Woodhouse (1980b). Without going into detail beyond the scope of this chapter, we will briefly point out the key steps that led to the formalism routinely used today. An asymptotic expression relating the
observed frequency shift $\delta \omega_k$ of a normal mode $k$ observed on a given source station path to the underlying average elastic structure over the corresponding great circle path was derived by Jordan (1978) and Dahlen (1979):

$$\delta \omega = \frac{1}{2\pi} \int_0^{2\pi} \delta \omega_k(s) \, ds$$  \hspace{1cm} (18)

where the local frequency shift $\delta \omega_k(s)$ represents an integral over depth of the difference between the local elastic structure and the reference Earth model at point $s$ along the source-station great circle, weighted by the depth sensitivity kernels of mode $k$ (e.g., Woodhouse and Girmius, 1982). The local frequency shift can be related to the phase velocity perturbation using Jeans’s formula. A perturbed synthetic seismogram could then be formed as a sum over normal modes:

$$u(\Delta, t) = \text{Re} \left[ \sum_k A_k^I(\Delta) \exp(i(\omega_k + \delta \omega_k) \phi \omega_k t) \right]$$  \hspace{1cm} (19)

where $\Delta$ is epicentral distance in radians, and the amplitude term $A_k^I$, frequency $\omega_k$, and attenuation $\omega_k$ are computed in the reference spherically symmetric Earth model. $\omega_k$ and $\delta \omega_k$ are the real and imaginary parts, respectively, of the average frequency shift along the complete great circle path. However, this expression only reflects the effect of heterogeneity that is symmetric with respect to the center of the Earth, in agreement with the fact that, to zeroth order in the asymptotic approximation, normal modes are only sensitive to that part of lateral heterogeneity (e.g., Woodhouse, 1983).

Relating observed waveforms of a single (say, the first arriving) surface wave to a synthetic normal mode seismogram presented a challenge. From the practical point of view, this stumbling block was removed by Woodhouse and Dziewonski (1984), who introduced a “distance shift” that depended on the structure of the minor arc in the computation of the seismogram in a slightly heterogeneous Earth, as follows:

$$u(\Delta, t) = \text{Re} \left[ \sum_k A_k^I(\Delta + \delta \Delta) \exp(i(\omega_k + \delta \omega_k) \phi \omega_k t) \right]$$  \hspace{1cm} (20)

where the summation is over all modes $k$, and

$$\delta \Delta = \frac{\omega_k a}{(l + 1/2) U} (\delta \omega - \delta \omega_k)$$  \hspace{1cm} (21)

Here $U$ is the group velocity of the mode $k$, $a$ is the radius of the Earth, and $\delta \omega_k$ is the minor arc average of the local frequency shift $\delta \omega$, defined as

$$\delta \omega = \frac{1}{\Delta} \int_0^{\Delta} \delta \omega(s) \, ds$$  \hspace{1cm} (22)

Theoretical proofs of Eqs. (20) and (21) were later given independently by Romanowicz (1987), Park (1987), and Mochizuki (1986a,b), in the framework of zeroth-order asymptotic coupling theory (Woodhouse, 1983; Tanimoto, 1984). The minor arc average $\delta \omega$ arises as a consequence of coupling of neighboring modes along a single dispersion branch, due to lateral heterogeneity, in contrast to the great circle average, which is obtained through coupling within isolated mode multiplets (Jordan, 1978; Woodhouse, 1983; Romanowicz and Roult, 1986).

The frequency-domain approach involves a two-step procedure: First, determine phase velocity maps of a single surface wave mode branch at individual frequencies; second, invert the dispersion curves obtained at every point on the globe to obtain the laterally varying depth distribution of shear velocities (since surface waves are primarily sensitive to shear velocity). On the other hand, the waveform approach is a single-step approach that obtains the 3D structure directly from the seismograms. Moreover, it can also be used for multimode waveforms, the approach favored by Woodhouse and Dziewonski (1984) and later extended, in combination with body wave travel time measurements, to whole-mantle tomography (e.g., Su et al., 1994).

Whatever the approach, a significant issue regarding the inversion of fundamental mode surface waves is that of crustal corrections. Indeed, surface waves are sensitive to shallow crustal structure even at long periods (e.g., Dziewonski, 1971). Most studies perform crustal corrections in the framework of linear perturbation theory. However, Montagner and Jobert (1988) showed that the effect of strongly varying crustal structure is nonlinear and proposed a more accurate correction procedure based on a tectonic regionalization. The nonlinear part comes primarily from very large lateral variations in depth to Moho and other crustal discontinuities (for example, depth to Moho can vary by a factor of 4 between oceans and continents). The crustal contribution is computed in two steps. First, $i = 1 - n$ regional reference models are considered, and phase velocities as well as partial derivatives with respect to the parameters and discontinuity depths are computed for each of these models. In a second step, linear corrections are applied for each point along the source-station path, taking into account perturbations of the actual crustal elastic parameters and discontinuity depths with respect to the local tectonic model. The contribution $\Delta \Phi$ to the observed phase, due to crustal structure, is thus of the form:

$$\Delta \Phi = \sum \left[ \int_S ds(M) \frac{\delta v_i}{C_i} - \int_S ds(M) \frac{\delta v_i}{C_i^2} \right]$$  \hspace{1cm} (23)

where $\delta v_i(M)$ is 1 if point $M$ belongs to a tectonic region $i$, and 0 otherwise; $C_i$ is the phase velocity in crustal model $i$; $h_k$ is the depth of the $k$th discontinuity of the tectonic model $i$; and $\delta h_k$
is the perturbation in that depth at point $M$. Also, $p_{ri}$ is
the $r$th elastic parameter in tectonic model $i$, and $\delta p_{ri}$
is the perturbation in that parameter at point $M$. Montagner
and Jobert (1988) give a slightly different expression for
the crustal contribution, arguing that only the lateral varia-
tions in discontinuity depths need to be treated in a nonlinear
fashion.

Developing accurate crustal models (e.g., Mooney et al.,
1998) remains a challenge for large-scale surface wave
inversions for structure, as does the accurate treatment of
crustal effects.

### 4.2 Higher-Mode Surface Wave Inversion

While well separated in the time domain from other mode
branches and therefore well suited for single mode analysis
techniques, fundamental mode surface waves have several
shortcomings: at intermediate periods (say, 20–150 sec) their
sensitivity to structure below about 200 km is poor, whereas
longer-period mantle waves, which reach down to the top of
the transition zone, have poor spatial resolution. In any case,
resolving structure in the upper mantle transition zone, which
is also poorly sampled by body waves, requires the analysis of
higher-mode surface waves, whose sensitivity is larger at
these depths (Fig. 6). They are also a powerful tool for inves-
tigating structures where low velocity zones may be present
(e.g., Kouch and Anderson, 1964).

In some specific frequency windows, and for some specific
source excitations, it has been possible to isolate and mea-
sure the dispersion of the first higher Rayleigh wave modes,
either at very short period, where they are well separated
on the seismogram (e.g., Crampin, 1964), or with the help
of time-variable filtering at periods between 100 and
200 sec (e.g., Roult and Romanowicz, 1984). In general, how-
ever, higher-mode surface waves overlap in the time–fre-
quency domain, and single mode dispersion methods therefore
cannot be applied. For example, in the period range 80–
150 sec, Rayleigh modes 3, 4, and 5 are well excited by
intermediate-depth earthquakes, and are observed on seismo-
grams as single energetic wave packets, labeled “X-phase” by
Jobert et al. (1977).

In the 1970s, in order to isolate higher modes of surface
waves, similar array methods were developed independently
by Nolet (1975) and Cara (1973b, 1978) and applied in the
period range 20 to 100 sec to paths across Eurasia and the
Pacific Ocean, respectively.

These methods require a linear regional array of stations
approximately aligned with the epicenter (and not in a nodal
direction) to separate modes in the $(\omega, k)$ domain, where $k$
is wavenumber. After correction for the instrument response, an
array stack is formed (e.g., Nolet, 1975):

$$S(\tilde{k}, \omega) = \left| \sum_n F_n(\omega) \exp(i\Phi_n)H_n(\tilde{k}, \omega) \right|^2$$  \hspace{1cm} (24)

where the “array response” is

$$|H_n(\tilde{k}, \omega)|^2 = \left[ \frac{1}{N} \sum_j \exp\left[ i(k_n(\omega) - \tilde{k})\Delta_j \right] \right]^2$$  \hspace{1cm} (25)

and $F_n(\omega)$, $\Phi_n(\omega)$, and $k_n(\omega)$ are, respectively, the amplitude
spectrum, initial phase, and wavenumber of mode $n$. $\Delta_j$ is
the epicentral distance to station $j$, and the spectrum of the multi-
mode signal at station $j$ is

$$W_j(\omega) = \sum_n F_n(\omega) \exp\{i(k_n(\omega)\Delta_j + i\Phi_n(\omega))\}$$  \hspace{1cm} (26)

The array-response has a peak at $\tilde{k} = k_n$. These peaks are
measured from plots of contours of the function $S(\tilde{k}, \omega)$ in the
$(\omega, \tilde{k})$ plane, from which dispersion curves $k_n(\omega)$ are then
derived (Fig. 7).

This method is nevertheless limited in its application to a
few regions of the world with relatively dense, linear arrays.
The condition of alignment of the array in a narrow azimuthal
range was later relaxed by Okal and Jo (1985, 1987), by cor-
recting for the azimuthal variations of the initial phase
$\Phi_n(\omega)$. This approach still suffered from lack of accuracy in
the reading of the maxima in the $S(\tilde{k}, \omega)$ plots, in particular due
to the presence of large side-lobes.

FIGURE 6 Depth profiles of partial derivatives of Rayleigh wave
fundamental and first three higher modes with respect to $S$ velocity at
a period of $\sim$100 sec, computed in the PREM model (Dziewonski
and Anderson, 1981). At this period, the fundamental mode sensitivity
peaks around 150 km depth and is negligible below 300 km depth. On
the other hand, higher modes have significant sensitivity throughout
the transition zone, and even, for modes 2 and 3, at the top of the
lower mantle. (Courtesy of Yuancheng Gung.)
In the meantime, following a cross-correlation technique first proposed by Dziwnowski et al. (1972), a waveform-based method involving the comparison of single record observed and synthetic seismograms was developed by Lerner-Lam and Jordan (1983) and later improved by Cara and Lévêque (1987) and Lévêque et al. (1991). In this approach, branch cross-correlation functions (bccf’s) are formed between a particular single-mode synthetic and the observed seismogram, as follows. The observed seismogram \( s(t) \) is written as the sum over overtone branches \( u_n(t) \),

\[
s(t) = \sum_n u_n(t) \tag{27}
\]

and a synthetic seismogram \( \tilde{s}(t) \) is computed for a reference spherically symmetric model suitable as a starting model for the average structure along the source–station path considered:

\[
\tilde{s}(t) = \sum_n \tilde{u}_n(t) \tag{28}
\]

The matched filter is the synthetic seismogram \( \tilde{u}_m(t) \) for a particular mode branch \( m \), so that the objective function to be minimized is formed as

\[
\tilde{O}_m(\tau) = (s(t) - \tilde{s}(t)) \cdot \tilde{u}_m(t) \tag{29}
\]

The objective function should be peaked near \( \tau = 0 \), and the displacement of the peak from \( \tau = 0 \) is a function of the difference between the observed and the computed dispersion. Partial derivatives with respect to model parameters can be formed to then invert for elastic model perturbations and obtain an average structure along the specific source–station path. Cara and Lévêque (1987) added a bandpass filter to further improve the resolution, and Lévêque et al. (1991) added a procedure to invert the envelope and phase of the filtered cross-correlograms. Indeed, this formalism lends itself also to the derivation of secondary observables such as group or phase velocities, which have a more linear dependence on the elastic structure, an approach discussed by Cara and Lévêque (1987) and also further developed by Gee and Jordan (1992). The main drawback of this cross-correlation methodology is the contamination of the single mode objective function by interference from other modes.
Other waveform inversion approaches have been developed that do not try to separate individual higher modes. Nolet (1990) introduced the “partitioned waveform inversion” approach in which inversion for elastic structure proceeds in two steps. Path integral parameters are defined and retrieved by nonlinear waveform fitting and, in a second step, inverted linearly for elastic structure. The main goal of this approach is to reduce the number of parameters to be fit in the nonlinear part of the inversion, and thereby to simplify the computations and increase their robustness. This method has subsequently been applied by Zielhuis and Nolet (1994) to retrieve upper mantle structure under central Europe, by van der Lee and Nolet (1997) in North America, and by Zielhuis and van der Hilst (1996) in Australia.

Alternatively, following the waveform modeling approach of Woodhouse and Dziewonski (1984), Li and Romanowicz (1995, 1996) introduce a global waveform modeling approach, based on a theoretical normal mode formalism that includes coupling across different mode branches (Li and Taninoto, 1993). This formalism is particularly appropriate for body waves, in that it involves the computation of broadband sensitivity functions centered around the ray path, a more correct approach than the standard surface wave path average approximation, which averages kernels laterally between the source and the receiver. However, for higher-mode surface wave trains with sensitivity to the top of the lower mantle, across-branch mode coupling also starts to matter if increased resolution of structure is to be attained (e.g., Mégínn and Romanowicz, 1998). A similar mode-coupling formalism, using a propagating wave approach in the frequency domain, has been derived by Marquering and Snieder (1995) and Zhao and Jordan (1998).

Efforts to extract single mode observables using a waveform approach continued in the 1990s. Stutzmann and Montagner (1993, 1994) directly compare observed and synthetic waveforms of multimode wavepackets for paths between several earthquakes approximately in the same location, but at different depths, and for a given station. They set up an iterative, nonlinear inversion, in which they assume that the only unknowns are perturbations of the phase velocities of the fundamental mode and first few higher modes (in general three) along each path. They thus retrieve path-dependent dispersion properties, which, in a second step, they combine in a global inversion for structure. The drawback of this method is the constraint on path selection imposed by the requirement of finding several well-recorded neighboring earthquakes of different depths, which limits the number of paths that can be processed around the Earth, and therefore the resolution of the model.

Recently, van Heijst and Woodhouse (1997, 1999) proposed a “mode-branch stripping” method based on a bcfe approach combined with a frequency-stepping procedure. These authors proceed iteratively to extract phase velocity and amplitude information, as a function of frequency, for the well-excited modes contained in a higher-mode wavepacket. They start at low frequency and proceed to higher frequencies, taking advantage of the smoothness of dispersion as a function of frequency, and thus are able to avoid the indeterminate $2\pi$ phase shifts that can be a problem at higher frequencies. By computing synthetics, they also determine the respective strengths of the different mode branches, and start the procedure by fitting the strongest mode branch. The improved synthetics for this mode branch are then subtracted from the data and the procedure is repeated for the next strongest mode with data and synthetics “stripped” of the already fitted preceding mode. At each step, the objective function to be minimized for mode $q$ is then

$$Q_q(\tau) = (s^{-p_1}p_1\cdots(t) - \tilde{s}^{-p_1}p_1\cdots) * \tilde{u}_q(m)$$

where $s^{-p_1}p_1\cdots(t)$ and $\tilde{s}^{-p_1}p_1\cdots$ are the observed and synthetic seismograms, respectively, stripped of the modes $p_1, p_2, \ldots$ that have already been fit.

van Heijst and Woodhouse (1997) argue that this approach not only provides better resolution for a given path than the original bcfe approach, but, by allowing the determination of single mode dispersion and amplitude with many fewer restrictions on the paths than the method of Stutzmann and Montagner (1993), it also leads to better sampling of the globe and, not being limited to theoretical assumptions on great circle path propagation, leaves open the possibility of extracting information on off-path propagation.

### 4.3 Upper Mantle Anisotropy from Fundamental Mode Surface Waves

The incompatibility of dispersion curves measured for Love and Rayleigh waves has provided some of the earliest evidence for anisotropy in the crust and upper mantle (e.g., Anderson, 1961; Aki and Kaminuma, 1963; McEvilly, 1964). Its widespread character was confirmed in the 1970s and 1980s, mainly for fundamental modes, and in the oceans (e.g., Forsyth, 1975; Schlue and Knoopp, 1977; Mitchell and Yu, 1980; Montagner, 1985), but also for higher modes (Levêque and Cara, 1983). This discrepancy can be explained by introducing a transversely isotropic medium with vertical symmetry axis, and it is in this framework that the widely used Preliminary Reference Earth Model (PREM) has been constructed (Dziewonski and Anderson, 1981). Studies of the Love/Rayleigh discrepancy have been extended to the global scale and at long periods (100–250 sec) by Nataf et al. (1984), who first mapped the global lateral variations of transverse isotropy as a function of depth, expanded in spherical harmonics up to degree 6.

Recently, Ekström and Dziewonski (1998) used global Love and Rayleigh wave dispersion measurements in the period range 35–300 sec, complemented by long period waveform and travel time data, to invert separately for $V_{SH}$ and $V_{SV}$ in
the upper mantle, and thus obtain a measure of the growth distribution of transverse isotropy. Their model ("S20"), which is expanded laterally in spherical harmonics up to degree 20, confirms the widespread presence of this type of anisotropy in the upper mantle, and particularly singles out a strong anomaly in the central Pacific Ocean.

In addition, surface wave dispersion has also been shown to vary with azimuth, another indication of anisotropy (e.g., Forsyth, 1975; Suetogu and Nakanishi, 1987). The global azimuthal variations of Rayleigh and Love wave dispersion at long periods (100–250 sec) were first mapped by Tanimoto and Anderson (1984, 1985), who showed that the fast direction appears to correlate with flow directions in the mantle.

Surface waves thus show two manifestations of anisotropy: (1) variations of phase and group velocities with azimuth (azimuthal anisotropy), and (2) inconsistent dispersion curves for azimuthally averaged Love and Rayleigh waves (transverse isotropy).

The theoretical expressions for the propagation of surface waves in an anisotropic plane layered medium was studied by Crampin (1970) and Smith and Dahlen (1973). The latter provided expressions for the azimuthal dependence of Love and Rayleigh waves in a slightly anisotropic medium. This formalism was extended to the case of a spherical Earth in a normal mode and spherical harmonics framework by Tanimoto (1986) and Mochizuki (1986b). Romanowicz and Sniider (1988) developed an equivalent formalism that does not require a spherical harmonics expansion and thus is applicable to regional studies, and Park (1997) generalized this to include source terms and compute complete synthetic seismograms in the Born approximation.

Here, we only present the basic asymptotic expressions that relate dispersion to anisotropic elastic structure. To first order in anisotropy, and at frequency \( \omega \), the azimuthal variation of phase velocity (Love or Rayleigh wave) is of the form:

\[
C(\omega, \theta) = A_1(\omega) + A_2(\omega) \cos 2\theta + A_3(\omega) \sin 2\theta + A_4(\omega) \cos 4\theta + A_5(\omega) \sin 4\theta
\]

where \( \theta \) is the azimuth of the wavenumber vector defined clockwise from north.

Montagner and Nataf (1986) provided expressions for the coefficients \( A_i(\omega) \), which are depth integral functions, in terms of the following combinations of standard cartesian elastic coefficients \( C_{ij} \), for both Love and Rayleigh waves:

**Constant term** \((A_1)\):

\[
A = \rho V_{ph}^2 = \frac{1}{9} (C_{11} + C_{22}) + \frac{1}{4} C_{12} + \frac{1}{2} C_{66}
\]

\[
C = \rho V_{pH}^2 = C_{33}
\]

\[
F = \frac{1}{4} (C_{13} + C_{23})
\]

\[
L = \rho V_{22}^2 = \frac{1}{4} (C_{44} + C_{55})
\]

\[
N = \rho V_{44}^2 = \frac{1}{8} (C_{11} + C_{22} - \frac{1}{2} C_{12} + \frac{1}{2} C_{66})
\]

**2\(\theta\) azimuthal term**:

\[
(A_2)
\]

\[
\cos 2\theta
\]

\[
B_x = \frac{1}{2} (C_{11} - C_{22})
\]

\[
B_y = C_{16} + C_{26}
\]

\[
G_x = \frac{1}{4} (C_{55} - C_{44})
\]

\[
G_y = C_{54}
\]

\[
H_x = \frac{1}{2} (C_{13} - C_{23})
\]

\[
H_y = C_{36}
\]

**4\(\theta\) azimuthal term**:

\[
(A_4)
\]

\[
\cos 4\theta
\]

\[
E_x = \frac{1}{4} (C_{11} + C_{22} - \frac{1}{2} C_{12} - \frac{1}{2} C_{66})
\]

\[
E_y = \frac{1}{2} (C_{16} - C_{26})
\]

The term \( A_1 \), independent of azimuth, involves only the five independent combinations of elastic coefficients needed to describe a transversely isotropic medium with vertical symmetry axis, labeled \( A, C, F, L, N \) (Love, 1927; Takeuchi and Saito, 1972). Montagner and Nataf (1986) showed that the partial derivatives of all the other terms only require the computation of partial derivatives with respect to the five parameters of a reference transversely isotropic Earth, thus providing the means to invert the azimuthal variations of surface wave phase velocities. They also estimated the different terms for realistic upper mantle models and showed that for Love waves, the \( N \) kernel is much larger than \( L \) and that the \( 4\theta \) term is dominant, making it difficult to use Love waves to constrain azimuthal anisotropy on the global scale, whereas for Rayleigh waves the \( 2\theta \) term prevails. Finally, because Rayleigh waves are sensitive both to shallow crustal and deeper mantle anisotropy, it is important to use a wide frequency range to resolve the depth dependence of anisotropy using surface waves.

Montagner and Nataf (1988) showed how this formalism could be used for the inversion of surface wave dispersion data including the azimuthal terms, under the assumption that the material possesses a symmetry axis (orthotropic medium). In this case the 3D model can be described using seven parameters (plus density): the five parameters \( A, C, F, L, N \) describing transverse isotropy, and two angles describing the orientation in space of the axis of symmetry. These authors used the inversion algorithm of Tarantola and Valette (1982) in its continuous form, as adapted to surface waves by Montagner (1986) with the introduction of an appropriate spatial correlation function, providing the description of the model and its error distribution on a grid of points (rather than a global basis function expansion), suitable for regional studies. This methodology, called “vectorial tomography,” was first applied by Montagner and Jobert (1988) to retrieve lateral heterogeneity and anisotropy, described at each point by an amplitude scalar, and a direction vector (under the hypothesis of orthotropy), to the study of the Indian Ocean, and the results were interpreted in terms of flow in the mantle.
This procedure involves two steps: (1) determining maps of azimuthal variations of anisotropy at individual frequencies, and (2) inverting the retrieved coefficients locally for heterogeneity and anisotropy variations with depth. A similar approach was subsequently applied to many other regions (e.g., Mocquet et al., 1989; Roulle et al., 1994; Grout et al., 1998; Pillet et al., 1999), and extended to the global study of heterogeneity and anisotropy in the upper mantle by Montagner and Tanimoto (1990, 1991). In this latter case, an approximation was introduced for the calculation of the data covariance matrix in the Tarantola and Valette (1982) framework, to make the computations manageable for a global dataset. While some questions linger about trade-offs between lateral heterogeneity and anisotropy in this type of inversion, Montagner and collaborators have shown that they can explain their datasets with fewer parameters when azimuthal anisotropy is considered than when it is ignored.

In order to circumvent the trade-off between lateral heterogeneity and azimuthal anisotropy, Park and Yu (1992, 1993) have looked for other diagnostic effects of anisotropy in long-period surface waves, such as waveform anomalies caused by Rayleigh–Love coupling, which generates “quasi-Love” waves on vertical components and “quasi-Rayleigh” waves on transverse components. Yu and Park (1994) have documented such observations, best seen in nodal directions of strike-slip sources, in the Pacific Ocean and inferred small scale variations in anisotropy related to tectonic features.

### 4.4 Effects of Scattering and Non-Great Circle Path Propagation

Until now, most regional and global models using surface waves have been derived using the standard “path-average” approximation, which is a high-frequency asymptotic approximation in which the propagation is assumed to be confined to the great circle path between the source and the receiver. This is valid only if the wavelength of lateral variations of structure is long with respect to that of the surface waves considered.

In fact, over the years, there have been many observations indicating that lateral heterogeneity is strong enough to cause departures from this simple hypothesis, as evidenced, for example, at short periods, where 20 sec surface waves sensitive to shallow crustal structure consistently show multipathing (e.g., Capon, 1970; Bungum and Capon, 1974; Fig. 8), or at long periods (T > 100 sec), where amplitude anomalies have been widely documented, with, for example, later-arriving trains showing larger amplitudes than the ones preceding them (Fig. 9), which cannot be explained by lateral variations of anelasticity (e.g., Lay and Kanamori, 1985; Roulle et al., 1986; Park, 1987).

Several approaches have been developed to try to explain these effects and exploit them to obtain better constraints on lateral variations of structure. The principal ones are cast in the framework of three different formalisms, depending on the application: ray theory, scattering theory, and a coupled-mode formalism. Each of these methods lends itself to various degrees and, under specific circumstances, to inversion.

Working in a 2D cartesian framework for the description of surface wave propagation in a smoothly laterally varying medium, Woodhouse (1974) introduced the concept of local modes. These are the surface wave modes corresponding to a laterally homogeneous model, which locally has the depth distribution of the laterally varying model. If the medium is smooth, the local mode branches propagate as independent wave trains, the dispersion and displacement of which are modified according to the evolution of the local modes. However, when the lateral variations are sharp (for example, in the presence of a structural discontinuity such as an ocean–continent boundary), the coupling of the local modes cannot be neglected, and its strength depends on the width of the structural transition zone (Kennett, 1972). Kennett (1984)
derived a formalism for mode-coupling in 2D laterally varying structures using a representation of the displacement field on the modes of a reference, laterally homogeneous structure, which still requires continuity in the lateral variations. The coupled, reference mode approach has been extended to the 3D scattering case and for certain types of structures, in cartesian coordinates, by Bostock (1991, 1992) and Bostock and Kennet (1992).

To overcome the restriction of smooth lateral variations, coupled local modes need to be considered. In this case, the wavefield is expanded in terms of a laterally varying local mode basis. A specific case, that of Love waves in a one-layered laterally varied structure, was studied by Odom (1986), whereas Maupin (1988) generalized the coupled local mode approach to any 2D structure. In this case, no reference laterally homogeneous structure is needed.

While 2D mode coupling can provide exact computations of coupling and conversion, and is appropriate for certain applications such as the study of continental margins, a scattering formalism has been, so far, more easily implemented in the context of inversion for the study of 3D heterogeneity in relatively smooth media.

Several applications of ray theory have been proposed to study the departure of surface wave paths from the conventional great circle approximation. Theoretical calculations were proposed by Woodhouse (1974), Babich et al. (1976), and Hudson (1981). Using a simple transformation from cartesian to spherical coordinates, Jobert and Jobert (1983) traced low-frequency Earth-circling mantle waves in a smoothly varying heterogeneous model using a combination of spherical Earth normal mode theory and a Gaussian beam computation scheme (Červený et al., 1982), and showed significant amplitude and travel time anomalies for a model of heterogeneity with strength of ~5%. At the regional, shorter period scale, Tanimoto (1990) showed how to compute the full wavefield of surface waves in a smoothly varying heterogeneous medium using an approximate scalar wave equation, and applied this to illustrate the distortion of surface waves propagating across California. Numerical computations showing the strong effects of lateral heterogeneity on the surface wavefield, in the framework of ray theory, were also performed by Yanovskaya and Roslov (1989). Tromp and Dahlen (1992a,b) developed a JWKB theory for the propagation of monochromatic surface waves in an Earth model with smooth lateral variations. These studies, however, did not result in any inversion procedures.

On the other hand, Yomogida and Aki (1985) used the Gaussian beam approach to compute intermediate period (20–80 sec) fundamental mode surface waveforms. They showed that, in this Born approximation framework, the amplitudes of surface waves depend on the second spatial derivatives of the phase velocity distribution, whereas the ray path depends on the first spatial derivatives and the phase on the phase velocity distribution itself. Based on this approach, they developed an inversion method for amplitude and phase data and applied it to retrieve phase velocity anomalies in the Pacific Ocean (Yomogida and Aki, 1987). The same dependence of phase and amplitude anomaly on the underlying phase velocity structure was derived independently by Woodhouse and Wong (1986) using ray theory based on the formalism of Woodhouse (1974). These authors showed that amplitude anomalies, as observed for long-period mantle waves on multiple Earth-circling paths, can be caused by focusing or defocusing due to lateral heterogeneity. Wong (1989) applied this theory to the retrieval of very long-period global mantle wave phase velocity maps from measured phase and amplitude anomalies. Since they depend on transverse gradients of structure along the propagation path, the amplitude anomalies can help constrain smaller-scale lateral variations of structure. The expression for the phase \(\delta \Phi\), amplitude \(\delta A/A_0\), and polarization \(\nu\) anomalies obtained in this framework, at epicentral distance \(\Delta\), are of the forms, respectively:

\[
\delta \Phi = \frac{1}{\Delta} \int_0^\Delta \frac{\partial \xi (\pi/2, \phi)}{c_0} d\phi
\]

\[
\frac{\delta A}{A_0} = \left( \frac{1}{2 \sin \Delta} \right) \int_0^\Delta \left[ \frac{\sin (\Delta - \phi)}{1 - \sin (\Delta - \phi)} \right] d\phi
\]

\[
- \frac{\partial}{c_0} \left( \frac{\partial \xi}{c_0} \right) \sin \phi d\phi
\]
\[ \nu(\Delta) = -\csc(\Delta) \int_{0}^{\phi} \sin \phi \bar{\phi} \left[ \delta c(\pi/2, \phi) \right] \, d\phi \]  

(37)

In these equations, the great circle path under consideration has been rotated to lie along the Equator, along which the angle \( \phi \) is measured (\( \phi = 0 \) at the epicenter and \( \phi = \Delta \) at the receiver). \( \nu \) is related to the azimuth \( \Theta \) along the great circle path by \( \nu = \tan(\Theta) \). The local perturbation in phase velocity at a point along the source receiver path is \( \delta c/c_0 \). We note that the phase \( \delta \Phi \) depends only on the phase velocity along the source station path [Eq. (35)]. On the other hand, the amplitude and polarization depend on first and second derivatives of the phase velocity, that is, on gradients of the phase velocity distribution, which in turn depend on short-wavelength features of that distribution. The expression under the integration sign in Eq. (36) actually corresponds to the transverse gradient in the direction perpendicular to the source station great circle (e.g., Romanowicz, 1987). Thus, amplitude and polarization anomalies potentially provide constraints on shorter-wavelength structure. Romanowicz (1987) showed that these expressions are equivalent, for relatively short propagation times, to those obtained for the perturbation to a low-frequency seismogram using a normal mode formalism, under an asymptotic approximation up to order \( 1/l \) in the development of spherical harmonics \( Y_m^n \).

In particular, the normal mode derived expression for the amplitude perturbation is

\[ \frac{\delta A}{A} = -\frac{a \Delta}{U} \left( \frac{\bar{D}}{2(l + 1/2)} \right) \]  

(38)

where \( U \) is group velocity and \( \bar{D} \) is the minor-arc average of the second transverse derivative of the local frequency perturbation along the source–station great circle path. These expressions for amplitude anomalies were considered by Romanowicz (1990) and Durek et al. (1993) in attempts to separate the effects of focusing and intrinsic attenuation on the global scale.

Laske (1995) and Laske and Masters (1996) used expression (37) for the azimuth anomaly to interpret observed polarization anomalies, measured using a multistep technique (Park et al., 1987), in terms of lateral variations of phase velocities on the global scale and thus retrieve shorter-scale variations than can be obtained with the same global distribution of observations using phase data alone. Polarization analysis using the same multistep technique had been proposed previously for the study of lateral heterogeneity by Lerner-Lam and Park (1989) and Paulissen et al. (1990).

Expression (35) for phase anomaly is correct to first order in lateral heterogeneity. Pollitz (1994) calculated that the second-order contribution to the phase and concluded that it is unimportant for global phase velocity models expanded up to degree 12 but is potentially important for rougher models (degree of expansion \( \geq 16 \)). That study suggested systematic bias in phase velocity maps that do not account for the second-order effect, though the potential magnitude of that bias requires further exploration. This effect arises from structure gradients perpendicular to the great circle path, suggesting that it should be considered jointly with polarization measurements in global phase velocity inversions.

In the framework of scattering theory, the single-scattering Born approximation, developed for surface waves by Snieder (1986) in a flat Earth geometry and Snieder and Nolet (1987) in a spherical Earth geometry, has been the subject of many studies. Born scattering is well suited for inversion since the scattered wavefield depends linearly on structural perturbations. Indeed, following Snieder (1988a), the surface wave displacement field can be decomposed into a “path average” part \( u_0^{\text{pava}} \), computed in a classical fashion, and a Born perturbation \( u^{\text{born}} \).

\[ u = u_0^{\text{pava}} + u^{\text{born}} \]  

(39)

where

\[ u^{\text{born}} = \int \int \int \sum_{\nu, \sigma} \mathbf{p}^{\nu} e^{i(k_0 \Delta \Delta + \pi/4)} S^{\nu, \sigma} e^{i(k_0 \Delta \Delta + \pi/4)} (\sin \Delta_2)^{1/2} (\sin \Delta_2)^{1/2} \times (\mathbf{E}^n; \mathbf{M}) \sin \theta \, d\theta \, d\phi \]  

(40)

where \( \sigma, \nu \) are the excited and scattered modes, respectively, \( \mathbf{p}^{\nu} \) is the polarization vector at the receiver, \( \Delta_1 \) and \( \Delta_2 \) are the angular distances to the scatterer from the source and the receiver, respectively, and the scattering matrix \( \Sigma_{\nu, \sigma} \) depends on elastic structural parameters via coupling kernels \( K \).

\[ S_{\nu, \sigma} = \int_{0}^{a} (K_{\nu, \sigma}^{\mu, \mu} + K_{\nu, \sigma}^{\mu, \rho} + K_{\nu, \sigma}^{\rho, \rho} + K_{\nu, \sigma}^{\rho, \lambda} + K_{\nu, \sigma}^{\lambda, \lambda}) \, dr \]  

(41)

where \( \delta \mu, \delta \rho, \) and \( \delta \lambda \) are 3D perturbations to the reference laterally homogeneous model described by \( \mu(r), \rho(r), \lambda(r) \). For intermediate- and long-period surface waves, perturbations in S velocity \( \beta \) need only be considered, and \( S_{\nu, \sigma} \) reduces to

\[ S_{\nu, \sigma} = \int_{0}^{a} K_{\nu, \sigma}^{\mu, \rho} \, d\beta \]  

(42)

If \( d \beta \) is expanded using 3D spatial basis functions, this leads to the linear expression

\[ \delta u^{\text{born}} = \sum_{i} \gamma_i a_i \]  

(43)

where \( a_i \) can be evaluated in the reference medium, and \( \gamma_i \) are the expansion coefficients of \( \delta \beta \) in the 3D spatial basis.

Snieder (1988a) showed how to set up a regional tomographic inversion using many fundamental surface waveforms. This method has been applied to the area of Europe and the Mediterranean by Snieder (1988b) and to North America by Alsina et al. (1996), and has been extended to the case of
multimode wave trains by Meier et al. (1997). In the latter two cases, the partitioned waveform method of Nolet (1990) has been used in conjunction with Born scattering, to minimize the computational effort involved.

Snieder and Romanowicz (1988) generalized this scattering formalism in the case of spherical earth geometry, and in a normal mode context, by applying an operator formalism as introduced by Romanowicz and Roult (1986) and Romanowicz (1987), avoiding expansion in spherical harmonics of the 3D Earth model, as previously employed (Woodhouse and Gurnis, 1982), and thus making it applicable to the case of any single scatterer (and not only to smooth heterogeneity). In this formalism, the addition theorem of spherical harmonics is used:

\[ \sum_{m} Y_{l}^{m}(\theta, \phi) J_{l}^{m}(\theta', \phi') = \gamma Y_{l}^{0}(\lambda) \sum_{m} Y_{l}^{m}(\theta, \phi) Y_{l}^{m}(\theta', \phi') \]

where \( l \) is the angular order of a normal mode \( k \), \((\theta_{e}, \phi_{e})\), \((\theta_{s}, \phi_{s})\) and \((\theta, \phi)\) are the coordinates of the epicenter, the receiver, and the scattering point, respectively (Fig. 10), and \( \gamma = \frac{2l+1}{4\pi} \). Angular distances \( \lambda \) and \( \beta \) are defined in Figure 10. The interaction coefficient \( Z_{KK'}^{mm'} \) between modes \( K, K' \) and their singlets \( m, m' \) may then be written simply as

\[ Z_{KK'}^{mm'} = \sum_{i=0}^{\infty} \int \int \frac{\partial p_{i}(X^{0}(\lambda)) \partial \omega_{KK'}^{0}(\theta, \phi) \partial p_{i}(X^{0}(\beta))}{d\Omega} \]

where the operators \( \partial p_{i} \) are linear combinations of differential operators acting on the coordinates \((\theta, \phi)\) of the running point, and incorporate source and receiver effects, computed in the reference spherically symmetric model. The integral is over the unit sphere, and \( \partial \omega_{KK'} \) are local frequency perturbations, describing depth-integrated effects of perturbations in the elastic model. To zeroth order in an asymptotic expansion in orders of \( 1/l \), the summation over \( i \) reduces to only one term \( (i = 0) \). Expression (45) has the same form as Eq. (40) and they are equivalent in the surface wave, high-frequency limit (Snieder and Romanowicz, 1988). Romanowicz and Snieder (1988) extended the same formalism to the case of a general anisotropic perturbation.

The Born approach is flexible as it is applicable to general 3D structures. However, it is a poor approximation in the case of strong heterogeneity, or when the region of scattering is large, in which case multiple scattering can play an important role.

On the other hand, Friederich et al. (1993) proposed an enhanced scattering theory for surface waves in a cartesian geometry, in which multiple forward scattering and single backward scattering are included. They use a potential formalism that can be summarized through the consideration of a scattering kernel \( K_{n}^{m} \) that includes all the details of the regional scattering problem relevant to two modes \( n \) and \( m \), depending only on the structural parameters. The surface wave displacement \( u \) for a mode \( m \) is decomposed as

\[ u_{m}(x, y, z) = U_{m}(z) \Phi_{m}(x, y) \]

where \( \Phi_{m} \) is the displacement potential and \( U \) is a vertical eigenfunction

\[ \Phi_{m}(x, y, z) = \int \int dx' dy' K_{m}^{m}(x, y|x', y') \Phi_{m}(x', y') \]

where \((x, y)\), \((x', y')\) are the coordinates of the observation point and the scattering point respectively, \( \Phi_{m} \) is the displacement potential of mode \( m \) and \( \Phi_{m}^{m} \) is the scattered potential of mode \( n \) generated by mode \( m \). The computation is simplified by discretizing the model into cells and assuming a piecewise plane wave approximation in each cell. When \( \Phi_{m} \) is the potential corresponding to a reference 1D model of mode \( m \), this is equivalent to single scattering theory. Friederich et al. (1993) use the reference potential for the row of cells closest to the source, compute the scattered potential \( \sum \Phi_{m} \) at all other rows, add the scattered wavefield and use this as starting potential to compute the scattered wavefield in the next row, and so on. They show that the computation thus includes all multiply forward scattered waves, but is incomplete for the backscattered field. This is not a problem in most cases, however, except when sharp, reflecting discontinuities are present. Subsequently, Friederich (1999) extended this approach to the case of spherical geometry and normal mode summation in a global Earth framework, using the operator formalism for the mode coupling terms of Romanowicz (1987). To limit the

\[ \text{FIGURE 10} \quad \text{Geometry for single scattering corresponding to Eq. (45).} \quad \text{S and R are the locations of the epicenter and the receiver, respectively.} \quad \text{Q is a generic point on the surface of the sphere, and P is the pole of the source receiver great circle.} \quad \text{(Reproduced with permission from Romanowicz, 1987.)} \]
amount of computation, they compute only the seismic displacement wavefield between the $S$ wave and the end of the first-arriving surface wave, at epicentral distances shorter than 50°.

Wielandt (1993) showed that the measured (or dynamic) phase velocity of teleseismic surface waves in a regional context is different from the structural phase velocity in the region considered. The former depends strongly on the geometry of the incoming wavefield, which can be distorted by structure along the propagation path outside of the area of study. Based on the work of Friederich et al. (1993), Friederich and Wielandt (1995) proposed a method to invert jointly for incoming wavefields and the heterogeneous phase velocity structure within a study region. In this approach, they include the multiple forward and backward scattering terms. They use a plane wave decomposition of the incoming wavefield and they reduce the nonuniqueness of the joint inversion by applying an “energy criterion,” constraining the energy of the total wavefield averaged over all the events considered to be equal to the mean squared amplitude at the stations, also averaged over all events. This method has been applied to fundamental mode surface waves in southern Germany by Friederich (1998) and in northern California by Pollitz (1999). Notably, Pollitz (1999) showed how to relate the inverted 2D phase velocity maps to 3D structure, taking into account both isotropic [$i = 1$ in Eq. (45)] and nonisotropic scattering interactions. The application of these promising methodologies is still in an early development stage, as they present many computational challenges.

### 4.5 Surface Wave Attenuation Measurements and Inversion for Upper Mantle Anelastic Structure

Surface wave amplitude measurements can be used to try to retrieve information about the anelastic structure of the crust and upper mantle. In the absence of perturbing effects due to scattering and focusing, the amplitude spectrum of a mono-mode wave train $i$, usually the fundamental mode, can be written as

$$ A_i(\omega) = A_0(\omega) \exp(-\eta(\omega)X_i) \quad (48) $$

where $X_i$ is the epicentral distance in kilometers and $A_0(\omega)$ represents the amplitude at the source. The attenuation coefficient can be related to $Q$ through (e.g., Aki and Richards, 1980)

$$ \eta(\omega) = \frac{\omega}{2C(\omega)Q(\omega)} \quad (49) $$

where $C$ is phase velocity. Surface wave measurements provide the primary constraints on attenuation structure (both variations with depth and lateral variations) in the crust and uppermost mantle. To investigate the radial structure in $Q$ over a wider depth range, long-period surface wave or normal mode attenuation measurements can be used. Both the radial and the lateral structures in $Q$ are, however, to this day, less well constrained than elastic structure, due to a large extent to the contamination of amplitude data by scattering and focusing effects due to wave propagation in a heterogeneous elastic Earth, as described in the previous section. There still exists an apparent discrepancy (on the order of 15%, Fig. 11) between the measurements of fundamental mode $Q$ obtained using a standing wave approach (e.g., Masters and Gilbert, 1983; Smith and Masters, 1989; Widmer et al., 1991) and those using a propagating wave approach (Dziewonski and Stein, 1982; Romanowicz, 1990, 1995; Durek et al., 1993; Durek and Ekström, 1996). This was investigated by Durek and Ekström (1997), who proposed that the discrepancy could be due to measurement techniques, with the presence of noise leading to an overestimation of $Q$ from normal modes, which require the computation of spectra over long time-series (typically 24 hours). On the other hand, Masters and Laske (1997) have questioned the accuracy of surface wave measurements of $Q$ at very long periods, in particular due to the difficulty of selecting appropriate time windows to isolate the fundamental structure.

**FIGURE 11** Fits of various mantle $Q_n$ model predictions to spheroidal mode data (Sailor and Dziewonski, 1978; Widmer et al., 1991; Resovsky and Ritzwoller, 1998) and Rayleigh wave data (Dziewonski and Stein, 1982; Durek et al., 1993; Romanowicz, 1995; Masters and Laske, 1997), illustrating the discrepancy between measurements obtained using standing wave and propagating wave approaches, respectively. Models QL6 and QM1 were constructed by Durek and Ekstrom (1996) and Widmer et al. (1991) respectively. (Reproduced with permission from Romanowicz and Durek, 2000.)
mode, in the presence of overlapping wave trains. This issue remains to be resolved.

Regional studies of amplitudes of fundamental mode surface waves in the period range 5–100 sec (Anderson et al., 1965; Mitchell, 1975; Canas and Mitchell, 1978, 1981) and of $L_g$ waves (Xie and Mitchell, 1990a,b) have long established the existence of large variations of $Q$ in the crust and uppermost mantle, correlated with tectonic provinces and in particular with the age of the oceans, and with time elapsed since the latest tectonic activity on continents. A recent review can be found in Mitchell (1995). At longer periods, large lateral variations also exist (e.g., Nakanishi, 1978; Dziewonski and Stein, 1982). These lateral variations can be an order of magnitude larger than those in elastic velocity.

Progress in using surface wave data to constrain global 3D anelastic structure of the upper mantle has been slow because, as mentioned previously, of the inherent difficulty of measuring attenuation in the presence of focusing and scattering effects that can be as large as or larger than anelastic effects and depend strongly on the short-wavelength details of the elastic structure. Indeed, as discussed previously [Eq. (36)] to first order these effects depend on the transverse gradients of structure along the great circle path linking epicenter and station, which in turn depend on terms of the form $s^t C_{st}$, where $C_{st}$ are the coefficients of an expansion in spherical harmonics of the elastic model ($s$ and $t$ are the degree and order of the corresponding spherical harmonic $Y_{st}$).

If the elastic structure of the Earth were accurately known to short wavelengths, one could first correct for its effects on the amplitude.

In the meantime, indirect methods have been used to minimize contamination of amplitudes by unwanted elastic effects. Romanowicz (1990) and Durek et al. (1993) took advantage of the fact that, in the linear approximation, focusing and anelastic effects could be separated by combining measurements over several consecutive wave trains, since attenuation effects are always additive whereas focusing depends on the direction of propagation. However, the longer the wave path the more the waves are affected by 3D elastic structure, and the harder it is to account for that in a simple approximate fashion. Another source of bias comes from uncertainties in the amplitude at the source. Romanowicz (1994) developed a method that involves computing attenuation coefficients in two ways: (1) using first-arriving trains only; (2) using the first three wave trains. The first measurement involves a source bias, but is less contaminated by elastic effects. On the other hand, in the second measurement, the source effect has been canceled out; however, the attenuation measurement is less accurate, and generally shows large variations with frequency, due to increased elastic effects over the longer paths. Comparison of the two $\gamma(\omega)$ curves allows determination of a source correction factor and thus allows one to obtain a relatively accurate attenuation measurement using first- and second-arriving trains only.

Progress in the retrieval of accurate $Q$ information from long-period surface wave data for global modeling is coupled to the development of efficient techniques for the modeling of elastic scattering effects. Recently, Billien et al. (2000) have proposed an approach in which phase and amplitude measurements of Rayleigh waves at intermediate periods (40–150 sec) are jointly inverted for elastic and anelastic structure, taking into account the first-order asymptotic focusing term [Eq. (36)]. It is still not clear, however, whether such a first-order, “smooth” approximation is sufficient to rule out biases due to scattering by small-scale heterogeneity.

Another issue is to account for dispersion in velocities due to attenuation, using an absorption band model for the Earth. This has been found to explain the discrepancy between average velocity models obtained using high-frequency body waves and surface waves (Kanamori and Anderson, 1977). However, most of the frequency dependence in surface waves is related to the effects of depth-dependent elastic structure. Surface waves alone cannot resolve the frequency dependence of $Q$. More detailed recent reviews on global $Q$ structure can be found in Romanowicz (1998) and Romanowicz and Durek (2000).

5. Conclusions

We have reviewed the evolution of the main methodologies based on the inversion of surface wave data over the last several decades. We find that much progress has been made in the last 20 y in the development of surface wave methods to retrieve both source parameters and Earth structure using surface wave data. This progress has built upon fundamental theoretical developments of the 1960s and 1970s, during which the basis for the computation of earthquake source excitation as well as modeling of surface wave dispersion properties in terms of the Earth’s structure was established. In the 1980s, surface wave studies have benefited from progress in computer speed and capacity, which, together with the deployment of a new generation of digital, broadband, high dynamic range global and regional seismic networks, has opened the way to large-scale source and structure inversions, with recent efforts to include anisotropy and anelastic contributions, as well as to systematically exploit information contained in overtones.

Most inversions, so far, have been performed under the assumption that the effects of propagation in the laterally heterogeneous Earth can be accounted for using the simple great circle path average approximation, which assumes that surface waves are sensitive only to the structure in the vertical plane containing the source and the receiver, and to the laterally averaged structure along this path. Beyond this approximation, there are two particularly important issues. One is the computation of more accurate kernels for propagation in the vertical plane containing the source and the receiver. This is important for waveform inversion that
includes body waveforms, and is relevant, in the context of this review, to the inversion of overtone waveforms sensitive to the mantle transition zone and greater depths. The first step in this direction has been the incorporation of crossbranch coupling effects computed asymptotically to zeroth order, both in a normal mode and in a propagating wave framework. This has started to be put into practice in the context of global waveform inversion (Li and Romanowicz, 1996; Mégünin and Romanowicz, 2000), and should help improve the resolution of shorter-wavelength features of tomographic models. The other issue is that of how to account for scattering and focusing effects, which can be viewed as departures of the wave path from the source–receiver great circle. Theoretical progress on how to describe these effects accurately has been steady, evolving from the consideration of single scattering to the more complex question of multiple scattering. Transition into practice (i.e., into large-scale inversions of real data) still presents many computational challenges. It is, however, a necessary step if the ever-increasing broadband data collection assembled in the last quarter of the 20th century is to be fully exploited to gain improved resolution at shorter wavelengths.

Acknowledgments

This paper was written while the author was on sabbatical leave from UC Berkeley, visiting at the Département de Seismologie, Institut de Physique du Globe in Paris, France. The author thanks three reviewers who contributed to improving the quality of this review.

References


