Chapter 4

Seismology

4.1 Historical perspective

1678 – **Hooke** Hooke's Law $\mathbf{F} = -c \cdot \mathbf{u}$ (or $\boldsymbol{\sigma} = \mathbf{E}\boldsymbol{\epsilon}$)

1760 – **Mitchell** Recognition that ground motion due to earthquakes is related to wave propagation

1821 – **Navier** Equation of motion

1828 – **Poisson** Wave equation
→ *P* & *S*-waves

1885 – **Rayleigh** Theoretical account surface waves
→ Rayleigh & Love waves

1892 – **Milne** First high-quality seismograph → begin of observational period

1897 – **Wiechert** Prediction of existence of dense core (based on meteorites → Fe-alloy)

1900 – **Oldham** Correct identification of *P*, *S* and surface waves

1906 – **Oldham** Demonstration of existence of core from seismic data

1906 – **Galitzin** First feed-back broadband seismograph

1909 – **Mohorovičić** Crust-mantle boundary

1911 – **Love** Love waves (surface waves)

1912 – **Gutenberg** Depth to core-mantle boundary : 2900 km

1922 – **Turner** location of deep earthquakes down to 600 km (but located some at 2000 km, and some in the air...)

1928 – **Wadati** Accurate location of deep earthquakes
→ Wadatai-Benioff zones

1936 – **Lehman** Discovery of inner core

1939 – **Jeffreys & Bullen** First travel-time tables
→ 1D Earth model

1948 – **Bullen** Density profile

1977 – **Dziewonski & Toksöz** First 3D global models

1996 – **Song & Richards** Spinning inner core?

Observations :

1964 ISC (International Seismological Centre) — travel times and earthquake locations

1960 WWSSN (Worldwide Standardized Seismograph Network) — (analog records)

1978 GDSN (Global Digital Seismograph Network) — (digital records)

1980 IRIS (Incorporated Research Institutes for Seismology)

4.2 Introduction

With seismology[1] we face the same problem as with gravity and geomagnetism; we can simply not offer a comprehensive treatment of the entire subject within the time frame of this course. The material is therefore by no means complete. We will discuss some basic theory to show how expressions for the propagation of elastic waves, such as P and S waves, can be obtained from the balance between stress and strain. This requires some discussion of continuum mechanics. Before we do that, let's look at a very brief – and incomplete – overview of the historical development of seismology. Modern seismology is characterized by alternations of periods in which more progress is made in theory development and periods in which the emphasis seems to be more on data collection and the application of existing theory on new and – often – better quality data. It's good to realize that observational seismology did not kick off until late last century (see section 4.1). Prior to that "seismology" was effectively restricted to the development of the theory of elastic wave propagation, which was a popular subject for mathematicians and physicists. For some important dates, see attachment above table (this historical overview is by no means complete but it does give an idea of the developments of thoughts). Lay & Wallace (1995) give their view on the current swing of the research pendulum in the following tables (with source related issues listed on the left and Earth structure topics on the right) :

Classical Research Objectives	
A. Source location (latitude, longitude, depth)	A. Basic layering (crust, mantle, core)
B. Energy release (magnitude, seismic moment)	B. Continent-ocean differences
C. Source type (earthquake, explosion, other)	C. Subduction zone geometry
D. Faulting geometry (area, displacement)	D. Crustal layering, structure
E. Earthquake distribution	E. Physical state of layers (fluid, solid)

Table 4.1: Classical Research objectives in seismology.

We will discuss some "classical" concepts and also discuss some of the more 'current ' topics. Before we can do this we have to deal with some basic theory. In principle, what we need is a formulation of the seismic source, equations to describe elastic wave propagation once motion has started somewhere, and a theory for coupling the source description to the solution for the equations of motion. We will concentrate on the former two problems. The seismic waves

[1] From the Greek words $\sigma\epsilon\iota\sigma\mu o\varsigma$ (seismos), earthquake and $\lambda o\gamma o\varsigma$ (logos), knowledge. In that sense, "earthquake seismology" is superfluous.

Current Research Objectives	
A. Slip distribution on faults	A. Lateral variations (crust, mantle, core?)
B. Stresses on faults and in the Earth	B. Topography on internal boundaries
C. Initiation/termination of faulting	C. Anelastic properties of the interior
D. Earthquake prediction	D. Compositional/thermal interpretations
E. Analysis of landslides, volcanic eruptions, etc	E. Anisotropy

Table 4.2: Current research objectives in seismology (after Lay & Wallace (1995))

basically result from the balance between stress and strain, and we will therefore have to introduce some concepts of continuum mechanics and work out general stress-strain relationships.

Intermezzo 4.1 SOME TERMINOLOGY

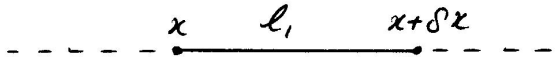
For most of the derivations we will use the Cartesian coordinate system and denote the position vector with either $\mathbf{x} = (x_1, x_2, x_3)$ or $\mathbf{r} = (x, y, z)$. The displacement of a particle at position $\mathbf{x}$ and time t is given by $\mathbf{u} = (u_1, u_2, u_2) = \mathbf{u}(\mathbf{x}, t)$, this is the vector distance from its position at some previous time t_0 (Lagrangian description of motion). The velocity and acceleration of the particle are given by $\dot{\mathbf{u}} = \partial\mathbf{u}/\partial t$ and $\ddot{\mathbf{u}} = \partial^2\mathbf{u}/\partial^2 t$, respectively. Volume elements are denoted by ΔV and surface elements by δS. Body (or *non-contact*) forces, such as gravity, are written as $\mathbf{f}$ and tractions by $\mathbf{t}$. A *traction* is the stress vector representing the force per unit area across an internal oriented surface $\delta\mathbf{S}$ within a continuum, and this is, in fact, the contact force $\mathbf{F}$ per unit area with which particles on one side of the surface act upon particles on the other side of the surface.
A general form of a wave equation is $\partial^2\mathbf{u}/\partial^2 t = c^2\partial^2\mathbf{u}/\partial^2 x$ or $\ddot{\mathbf{u}} = c^2\nabla^2\mathbf{u}$, which is a differential equation describing the propagation of a displacement disturbance $\mathbf{u}$ with speed c.

We will see that the fundamental theory of wave propagation is primarily based on two equations : Newton's second law ($\sum \mathbf{F} = m\mathbf{a} = m\partial^2\mathbf{u}/\partial^2 t$) and Hooke's *constitutive* law $\mathbf{F} = -c\mathbf{u}$ (stating that the extension of an elastic material results in a restoring force $\mathbf{F}$, with c the elastic (spring) constant (not wave speed as in the box above!). In one dimension, Hooke's law can also be formulated as the proportionality between stress σ and strain ϵ, with proportionality

factor E is Young's modulus : $\sigma = E\epsilon$. We will see that this linear relationship between stress and strain does not hold in 2D or 3D, in which case we need the so-called *generalized Hooke's Law*. For $\sum \mathbf{F} = m\mathbf{a}$ we have to consider both the non-contact body forces, such as gravity that works on a certain volume, as well as the contact forces applied by the material particles on either side of arbitrary and imaginary internal surfaces. The latter are represented by *tractions* ("stress vectors"). We therefore have to look in some detail at the definitions of stress and strain.

4.3 Strain

The strain involves both length and angular distortions. To get the idea, let's consider the deformation of a line element l_1 between x and $x + \delta x$.



Due to the deformation position x is displaced to $x + u(x)$ and $x + \delta x$ to $x + \delta x + u(x + \delta x)$ and l_1 becomes l_2.

The strain in the x direction, ϵ_{xx}, can then be defined as

$$\epsilon_{xx} = \frac{l_2 - l_1}{l_1} = \frac{u(x + \delta x) - u(x)}{\delta x} \tag{4.1}$$

If we assume that δx is small we can linearize the problem around the 'reference state' $u(x)$ by using a Taylor expansion on $u(x + \delta x)$:

$$u(x + \delta x) = u(x) + \left(\frac{\partial u}{\partial x}\right)\delta x + \mathcal{O}(\delta x^2) \approx u(x) + \left(\frac{\partial u}{\partial x}\right)\delta x \tag{4.2}$$

so that

$$\epsilon_{xx} = \left(\frac{\partial u(x)}{\partial x}\right) = \frac{1}{2}\left(\frac{\partial u(x)}{\partial x} + \frac{\partial u(x)}{\partial x}\right) \tag{4.3}$$

which represents the **normal strain** in the x direction. Similar relationships can be derived for the normal strain in the other principal directions and also for the **shear strain** ϵ_{xy} and ϵ_{xz} (etc), which involve the rotation of line elements within the medium.

The general form of the strain tensor ϵ_{ij} is

$$\begin{aligned} \epsilon_{ij} &= \frac{1}{2}\left(\frac{\partial u(x_i)}{\partial x_j} + \frac{\partial u(x_j)}{\partial x_i}\right) = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \\ &= \frac{1}{2}\left(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j}\right) = \epsilon_{ji} \end{aligned} \tag{4.4}$$

with normal strains for $i = j$ and shear strains for $i \neq j$. (In this discussion of deformation we do not consider translation and/or rotation of the material itself). Equation (4.4) shows that the strain tensor is symmetric, so that there the maximum number of different coefficients is 6.

4.4 Stress

Stress is force per unit area, and the principle unit is Nm^{-2} (or *Pascal* : $1\mathrm{Nm}^{-2} = 1\mathrm{Pa}$).

Similar to strain, we can also distinguish between **normal stress**, the force $\mathbf{F}_\perp$ per unit area that is perpendicular to the surface element $\delta\mathbf{S}$, and the **shear stress**, which is the force $\mathbf{F}_\parallel$ per unit area that is parallel to $\delta\mathbf{S}$ (see Fig. 4.1). The force $\mathbf{F}$ acting on the surface element $\delta\mathbf{S}$ can be decomposed into three components in the direction of the coordinate axes : $\mathbf{F} = (F_1, F_2, F_3)$. We further define a unit vector $\hat{\mathbf{n}}$ normal to the surface element $\delta\mathbf{S}$. The length of $\hat{\mathbf{n}}$ is, of course, $|\hat{\mathbf{n}}| = 1$.

For stress we define the **traction** as a vector that represents the total force per unit area on $\delta\mathbf{S}$. Similar to the force $\mathbf{F}$, also the traction $\mathbf{tt}$ can be decomposed into $\mathbf{t} = (t_1, t_2, t_3) = t_1x_1 + t_2x_2 + t_3x_3$. The traction $\mathbf{t}$ represents the total stress acting on $\delta\mathbf{S}$.

In order to obtain a more useful definition of the traction $\mathbf{t}$ in terms of elements of the stress tensor consider a tetrahedron. Three sides of the tetrahedron are chosen to be orthogonal to the principal axes in the sense that Δs_i is orthogonal to x_i; the fourth surface, $\delta\mathbf{S}$, has an arbitrary orientation. The stress working on each of the surfaces of the tetrahedron can be decomposed into components along the principal axes of the coordinate system. We use the following notation convention : the component of the stress that works on the plane $\perp x_1$ in the direction of x_i is σ_{1i}, etc.

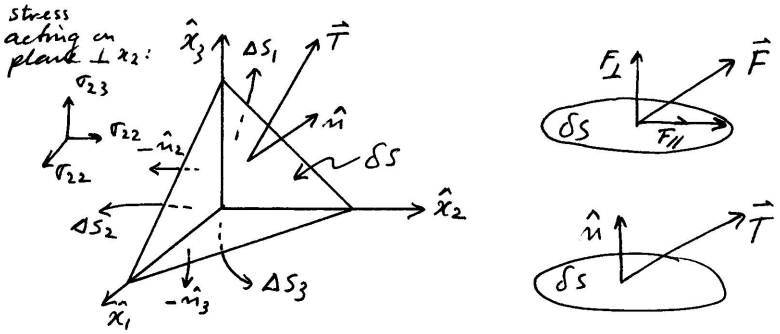


Figure 4.1: Stress balancing in the stress tetrahedron.

If the system is in equilibrium then a force $\mathbf{F}$ that works on $\delta\mathbf{S}$ must be cancelled by forces acting on the other three surfaces : $\sum F_i = t_i\delta S - \sigma_{1i}\Delta s_1 - \sigma_{2i}\Delta s_2 - \sigma_{3i}\Delta s_3 = 0$ so that $t_i\delta S = \sigma_{1i}\Delta s_1 + \sigma_{2i}\Delta s_2 + \sigma_{3i}\Delta s_3$. We know that the expression we are after should not depend on our choice of Δs nor on δS (since

the former were just chosen and the latter is arbitrary). This is easily achieved by realizing that δS and ΔS are related to each other : Δs_i is nothing more than the orthogonal projection of δS onto the plane perpendicular to the principal axis x_i : $\Delta s_i = \cos\varphi_i \delta S$, with φ_i the angle between $\hat{\mathbf{n}}$, the normal to δS, and x_i. But $\cos\varphi_i$ is in fact simply n_i so that $\Delta s_i = n_i \delta S$. Using this we get :

$$t_i \delta S = \sigma_{1i} n_1 \delta S + \sigma_{2i} n_2 \delta S + \sigma_{3i} n_3 \delta S \tag{4.5}$$

or

$$t_i = \sigma_{1i} n_1 + \sigma_{2i} n_2 + \sigma_{3i} n_3 \tag{4.6}$$

Thus : the i^{th} component of the traction vector $\mathbf{t}$ is given by a linear combination of stresses acting in the i^{th} direction on the surface perpendicular to x_j (or parallel to n_j), where $j = 1, 2, 3$;

$$t_i = \sigma_{ji} n_j \tag{4.7}$$

Conversely, an element σ_{ji} of the stress tensor is defined as the i^{th} component of the traction acting on the surface perpendicular to the j^{th} axis (x_j) :

$$\sigma_{ij} = t_i(x_j) \tag{4.8}$$

The 9 components σ_{ji} of all tractions form the elements of the stress tensor. It can be shown that in absence of body forces the stress tensor is symmetric $\sigma_{ij} = \sigma_{ji}$ so that there are only 6 independent elements :

$$\sigma_{ij} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{12} \\ \sigma_{31} & \sigma_{32} & \sigma_{13} \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ & \sigma_{22} & \sigma_{12} \\ & & \sigma_{13} \end{pmatrix} \tag{4.9}$$

The normal stresses are represented by the diagonal elements (i=j) and the shear stresses are the off diagonal elements ($i \neq j$). We can diagonalize the stress tensor by changing our coordinate system in such a way that there are no shear stresses on the surfaces perpendicular to any of the principal axes (see Intermezzo 4.2). The stress tensor then gets the form of

$$\sigma_{ij} = \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} \tag{4.10}$$

Some cases are of special interest :

- **uni-axial stress** : only one of the principal stresses is non-zero, e.g. $\sigma_1 \neq 0$, $\sigma_2 = \sigma_3 = 0$
- **plane stress** : only one of the principal stresses *is* zero, e.g. $\sigma_1 = 0$, $\sigma_2, \sigma_3 \neq 0$

- **pure shear** : $\sigma_3 = 0$, $\sigma_1 = -\sigma_2$
- **isotropic (or, hydrostatic) stress** : $\sigma_1 = \sigma_2 = \sigma_3 = p$ ($p = \frac{1}{3}(\sigma_1 + \sigma_2 + \sigma_3)$) so that the *deviatoric stress*, i.e. the deviation from hydrostatic stress is written as :

$$\sigma'_{ij} = \begin{pmatrix} \sigma_1 - p & 0 & 0 \\ 0 & \sigma_2 - p & 0 \\ 0 & 0 & \sigma_3 - p \end{pmatrix} \tag{4.11}$$

4.5 Equations of motion, wave equation, P and S-waves

With the above expression for the (symmetric) strain tensor (Eq. 4.4) and the definitions of the stress tensor σ_{ij} and the traction t_i, we can formulate the basic expression for the equation of motion :

$$\begin{aligned} \sum F_i &= \int_V f_i \, dV + \int_S t_i \, dS \\ &= \int_V f_i \, dV + \int_S \sigma_{ij} n_j \, dS = \int_V \rho \frac{\partial^2 u_i}{\partial t^2} \, dV = m a_i \end{aligned} \tag{4.12}$$

If we apply Gauss' divergence theorem[2], this can be rewritten as

$$\begin{aligned} \int_V \rho \frac{\partial^2 u_i}{\partial t^2} \, dV &= \int_V \left(f_i + \frac{\partial \sigma_{ij}}{\partial x_j} \right) dV \\ \rho \frac{\partial^2 u_i}{\partial t^2} &= f_i + \frac{\partial \sigma_{ij}}{\partial x_j} \end{aligned} \tag{4.13}$$

which is **Navier's equation** (also known as Cauchy's "law of motion" from 1827). For many practical purposes in seismology it is appropriate to ignore body forces so that the equation of motion is simplified to :

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial \sigma_{ij}}{\partial x_j} \quad \text{or} \quad \rho \ddot{u}_i = \sigma_{ij,j} \tag{4.14}$$

[2]Gauss' divergence theorem states that in the absence of creation or destruction of matter, the density within a region of space V can change only by having it flow into or away from the region through its boundary S :

$$\int_S \boldsymbol{t} . \, d\boldsymbol{S} = \int_V \bigtriangledown . \boldsymbol{t} \, dV$$

Intermezzo 4.2 DIAGONALIZATION OF A MATRIX

Many problems in (geo)physics can be simplified if we can diagonalize a matrix. Under certain conditions (almost always satisfied in geophysics), for any square matrix $\mathbf{A}$ of dimension n, there exists a $n \times n$ matrix $\mathbf{X}$ that diagonalize $\mathbf{A}$:

$$\begin{aligned} \mathbf{X^{-1}AX} &= \lambda = \mathbf{diag}(\lambda_\mathbf{1}, ..., \lambda_\mathbf{n}) \\ &= \begin{pmatrix} \lambda_1 & 0 & ... & ... & 0 \\ 0 & \lambda_2 & 0 & ... & 0 \\ 0 & ... & ... & ... & 0 \\ 0 & ... & 0 & \lambda_{n-1} & 0 \\ 0 & ... & ... & 0 & \lambda_n \end{pmatrix} \end{aligned} \tag{4.15}$$

This means that there exists a coordinate system in which $\mathbf{A}$ is diagonal. Diagonalizing $\mathbf{A}$ corresponds to finding this coordinate system and the values of the diagonal elements of $\mathbf{A}$ in this coordinate system. We can rewrite the last equation as follows :

$$\mathbf{AX} = \lambda\mathbf{X} \tag{4.16}$$

or

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{X} = \mathbf{0} \tag{4.17}$$

$\mathbf{I}$ is the Identity matrix. The λ_i $(i = 1, ..., n)$ are called the *eigenvalues* of $\mathbf{A}$ and the columns of $\mathbf{X}$ are formed by n *eigenvectors*. Diagonalizing a matrix is equivalent to finding its eigenvalues and eigenvectors. This is called an eigenvalue problem. Finding the eigenvalues can easily be done by solving the system of n linear equations and n unknowns (the λ_i) formed by Eq. 4.17. This has a non-trivial solution if the determinant is zero (this is called Cramer's rule) :

$$|\mathbf{A} - \lambda\mathbf{I}| = 0 \tag{4.18}$$

The eigenvectors can then be found by replacing the eigenvalues in the system of linear equations formed by Eq. 4.17. If all eigenvalues are different, the n eigenvectors are linearly independent and orthogonal. Otherwise, the eigenvalues are said to be *degenerate* and the number of independent eigenvectors is given by the number of independent eigenvalues. In the case of n independent eigenvalues, the eigenvectors can form a new orthogonal basis and they are called *principal axes*. If we change the coordinate system and use the system defined by the principal axes, matrix $\mathbf{A}$ becomes diagonal and its elements are given by the eigenvalues.
In the case of the stress tensor, equation 4.17 takes the form :

$$(\boldsymbol{\sigma} - \sigma\mathbf{I})\mathbf{n} = \mathbf{0} \tag{4.19}$$

The three eigenvalues (also called *principal stresses* and represented by the scalar σ) are thus found by solving :

$$|\boldsymbol{\sigma} - \sigma\mathbf{I}| = \begin{vmatrix} \sigma_{11} - \sigma & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} - \sigma & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} - \sigma \end{vmatrix} = 0 \tag{4.20}$$

This will give three values for σ (σ_1, σ_2 and σ_3). In the coordinate system formed by the three principal axes $\mathbf{n_i}$, the stress tensor is diagonal, as expressed in Eq. 4.10.

Note that body forces such as gravity cannot always be ignored in – what is known as – **low-frequency seismology**. For instance, gravity is an important restoring force for some of Earth's free oscillations. We can also introduce a body force term to describe the seismic source.

We've derived Eq. 4.14 using index notation. Let's state it in vector form. The acceleration is proportional to the divergence of the stress tensor (see Intermezzo 4.3) :

$$\rho\ddot{\mathbf{u}} = \boldsymbol{\nabla} \cdot \boldsymbol{\sigma} \tag{4.21}$$

Equation (4.14) represents, in fact, three equations (for i=1,2,3) but there are more than three unknowns (the 6 independent elements of the stress tensor σ_{ij} plus density ρ. In this general form the equation of motion does not have a unique solution. Also, we have introduced forces and tractions but we not yet specified *how* the material reacts to the applied (non-)contact forces. We need some physics to help us out. Specifically, we need to know the relationship between stress and strain, i.e. a *constitutive* relationship.

Intermezzo 4.3 DIVERGENCE OF A TENSOR

We know how to define the divergence of a vector. The divergence of a tensor is simply the generalization to higher dimensions of the divergence of a vector (remember that a vector is nothing more than a tensor of dimension 1).
The divergence of a vector $\mathbf{v}$ is a scalar denoted by $\bigtriangledown.\mathbf{v}$ and given by :

$$\bigtriangledown.\mathbf{v} = \sum_{\mathbf{i}} \frac{\partial \mathbf{v_i}}{\partial \mathbf{x_i}} = \frac{\partial \mathbf{v_1}}{\partial \mathbf{x_1}} + \frac{\partial \mathbf{v_2}}{\partial \mathbf{x_2}} + \frac{\partial \mathbf{v_3}}{\partial \mathbf{x_3}} \tag{4.22}$$

Similarly, the divergence of a dimension 2 tensor is a vector whose components are given by :

$$(\bigtriangledown.\boldsymbol{\sigma})_j = \sum_{i} \frac{\partial \sigma_{ij}}{\partial x_i} = \frac{\partial \sigma_{1j}}{\partial x_1} + \frac{\partial \sigma_{2j}}{\partial x_2} + \frac{\partial \sigma_{3j}}{\partial x_3} \tag{4.23}$$

And we can further generalize : the divergence of a *n*-dimension tensor is a tensor of dimension *n-1* obtained in a way similar to Eq. 4.23.

In one-dimension this relationship is given, as mentioned before, by $\sigma = E\epsilon$ (or $\sigma_i = E\epsilon_i$, where E is the Young's modulus, which is the ratio of uniaxial stress to strain in the same direction, i.e. a measure of the resistance against extension. A simple example demonstrates that in more dimensions this scalar proportionality breaks down. Imagine an elastic band : if one stretches this band in one direction, say the x_1 direction, than the band will extend in that direction. In other words there will be strain e_{11} due to stress σ_{11}. However, the strap will also thin in the x_2 and x_3 directions; so $e_{22} = e_{33} \neq 0$ even though $\sigma_{22} = \sigma_{33} = 0$.

Clearly, a simple scalar relationship between the stress and strain tensors is invalid : $\sigma_{ij} \neq E\epsilon_{ij}$. Somehow we must express the elements of the stress tensor as a linear combination of the elements of the strain tensor. This linear combination is given by a 4^{th} order tensor c_{ijkl} of elastic constants :

$$\sigma_{ij} = C_{ijkl}\epsilon_{kl} \tag{4.24}$$

This form of the constitutive law for linear elasticity is known as the **generalized Hooke's law** and C is also known as the stiffness tensor. Substitution of eq (4.24) in (4.14) gives the wave equation for the transmission of a displacement disturbance with wave speed dependent on density ρ and the elastic constants in C_{ijkl} in a general elastic, homogeneous medium (in absence of body forces) :

$$\rho\ddot{u}_i = \rho\frac{\partial^2 u_i}{\partial t^2} = \frac{\partial}{\partial x_j}\left[C_{ijkl}\frac{\partial u_k}{\partial x_l}\right] = C_{ijkl}\frac{\partial}{\partial x_j}\frac{\partial u_k}{\partial x_l} = C_{ijkl}u_{k,lj} \tag{4.25}$$

In three dimensions, a fourth order tensor contains $3^4 = 81$ elements. What did we gain by doing all this? After all, we mentioned above that we needed to introduce a constitutive relationship in order to solve the wave equation (Eq. 4.14) since the number of equations was less than the number of unknowns. Now we have arrived at a situation (Eq. 4.25) in which we have 3 equations to solve for 82 unknowns (density + 81 elastic moduli), so the introduction of physics does not seem to have helped us at all! The situation improves once we consider the intrinsic symmetry of the tensors involved. The symmetry of the stress and strain tensors leads to symmetry of the elasticity tensor : $C_{ijkl} = C_{ijlk} = C_{jilk}$. This reduces the number of independent elements in C_{ijkl} to 6×6=36. It can also be demonstrated (with less trivial arguments) that $C_{ijkl} = C_{klij}$, which further reduces the number of independent elements in C_{ijkl} to 21. This represents the most general (homogeneous) anisotropic medium (anisotropy in this context means that the relationship between stress and strain is dependent on the direction i).

By restricting the complexity of the medium we can further reduce the number of independent elements of the elasticity tensor. For instance, one can investigate special cases of anisotropy by allowing directional dependence in a plane perpendicular to certain symmetry axes only. We will come back to this later.

The simplest case is a homogeneous, isotropic medium (i.e. no directional dependence of elastic properties), and it can be shown (see, e.g., Malvern (1969)) that in this situation the general form of the 4^{th} order (linear) elasticity tensor is

$$C_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \tag{4.26}$$

where λ and μ are the only two independent elements; λ and μ are known as Lamé's (elastic) constants (or moduli), after the French mathematician G. Lamé. (The Kronecker (delta) function $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \neq j$). Substitution of Eq. (4.26) in (4.24) gives for the stress tensor

$$\sigma_{ij} = C_{ijkl}\epsilon_{kl} = \lambda\delta_{ij}\epsilon_{kk} + 2\mu\epsilon_{ij} = \lambda\delta_{ij}\Delta + 2\mu\epsilon_{ij} \tag{4.27}$$

with Δ the **cubic dilation**, or volume change. This form of Hooke's law was first derived by Navier (1820-ies). The Lamé constant μ is known as the **shear modulus or rigidity** : it is a measure of the resistance against shear or torsion of the medium. The shear modulus is large for very stiff material, but is small for media with low viscosity ($\mu = 0$ for water or for liquid metallic iron in the outer core). The other Lamé constant, λ, does not have much (general) physical meaning by itself, but defines important elastic parameters in combination with the shear modulus μ. Of most interest for us right now is the definition of κ, the **bulk modulus** or **incompressibility** : $\kappa = \lambda + 2/3\mu$. The bulk modulus is a measure of the resistance against volume change : $\kappa = -\partial P/\partial\Delta$, with P the pressure and Δ the cubic dilatation, and is large when the change in volume is small even for large (hydrostatic) pressure. The minus sign is necessary to keep $\kappa > 0$ since $\Delta < 0$ when $P > 0$. For isotropic media other important elastic parameters, such as the Poisson's ratio, i.e. the ratio of lateral contraction to longitudinal extension, and Young's modulus can also be expressed as linear combinations of λ and μ (or κ and μ). We can readily see that the stress tensor consists of terms representing (resistance to) either changes in volume or shear (or torsion).

stress : effects of volume change + torsion (or shear) of material

This is a fundamental result and it underlies, what we will see below, the formulation of wave propagation in terms of compressional (dilatational) P and transversal (shear) S-waves.

With the above constitutive relationships we can now derive the equation that describes wave propagation in a homogeneous, isotropic medium

$$\rho\frac{\partial^2 u_i}{\partial t^2} = (\lambda + \mu)\frac{\partial}{\partial x_i}\frac{\partial u_k}{\partial x_k} + \mu\nabla^2 u_i \tag{4.28}$$

which represents a system of three equations (for i=1,2,3) with three unknowns (ρ, λ, μ). Note that for practical purposes in seismology these parameters are not really constant; in Earth they are functions of position $\mathbf{r}$ and vary significantly, in particular with depth.

4.6 P and S-waves

There are several ways to demonstrate that solutions of the equation of motion essentially consist of a dilatational and a rotational term, the P and S-waves, respectively. Using vector notation the equation of motion is written as

$$\rho\ddot{\mathbf{u}} = (\lambda + \mu)\nabla(\nabla \cdot \mathbf{u}) + \mu\nabla^2\mathbf{u} \tag{4.29}$$

or, by making use of the vector identity

$$\nabla^2\mathbf{u} = \nabla(\nabla \cdot \mathbf{u}) - (\nabla \times \nabla \times \mathbf{u}), \tag{4.30}$$

we can write the equation of motion as :

$$\rho\ddot{\mathbf{u}} = \underset{\substack{\uparrow \\ \text{dilatational}}}{(\lambda+2\mu)\nabla(\nabla\cdot\mathbf{u})} \; \underset{\substack{\uparrow \\ \text{rotational}}}{-\mu(\nabla\times\nabla\times\mathbf{u})} \tag{4.31}$$

which is a system of three partial differential equations for a general displacement field $\mathbf{u}$ through an unbounded, homogeneous, and isotropic medium.

In general, it is difficult to solve this system directly for the displacement $\mathbf{u}$. Typically, one tries to decompose the general wave equation into separate equations that relate to *P*- and S-wave propagation. One approach is to eliminate directly any rotational contributions to the displacement by taking the divergence of Eq. (4.31) and using the property that for a vector field $\mathbf{a}$, $\nabla\cdot(\nabla\times\mathbf{a}) = 0$. Similarly we can eliminate the dilatational contributions by taking the rotation of (4.31) and using the identity that, for a scalar field μ, $\nabla\times\nabla\mu = \mathbf{0}$. Assuming no body force $\mathbf{f}$, we get :

- Taking the **divergence** leads to

$$\rho\frac{\partial^2(\nabla\cdot\mathbf{u})}{\partial t^2} = (\lambda+2\mu)\nabla^2(\nabla\cdot\mathbf{u}) \tag{4.32}$$

  or, with $\nabla\cdot\mathbf{u} = \Theta$,

$$\boxed{\frac{\partial^2\Theta}{\partial t^2} = \alpha^2\nabla^2\Theta} \tag{4.33}$$

  which is a *scalar* wave equation that describes the propagation of a volume change Θ through the medium with wave speed

$$\alpha = \sqrt{\frac{\lambda+2\mu}{\rho}} = \sqrt{\frac{\kappa+4/3\mu}{\rho}} \tag{4.34}$$

  In general $\kappa = \kappa(\mathbf{r})$, $\mu = \mu(\mathbf{r})$, $\rho = \rho(\mathbf{r}) \Rightarrow \alpha = \alpha(\mathbf{r})$

- Taking the **rotation** leads to

$$\rho\frac{\partial^2(\nabla\times\mathbf{u})}{\partial t^2} = (\lambda+2\mu)\nabla\times\nabla(\nabla\cdot\mathbf{u}) - \mu\nabla\times(\nabla\times\nabla\times\mathbf{u}) \tag{4.35}$$

  which, with $\nabla\times\nabla(\nabla\cdot\mathbf{u}) = \mathbf{0}$ and the vector identity as used above (and again using $\nabla\cdot(\nabla\times\mathbf{a}) = \mathbf{0}$), leads to :

$$\boxed{\frac{\partial^2(\nabla\times\mathbf{u})}{\partial^2 t} = \beta^2\nabla^2(\nabla\times\mathbf{u})} \tag{4.36}$$

This is a *vector* wave equation that describes the transmission through a medium of a rotational disturbance $\nabla \times \mathbf{u}$ with wave speed

$$\beta = \sqrt{\frac{\mu}{\rho}} \tag{4.37}$$

In general $\mu = \mu(\mathbf{r})$, $\rho = \rho(\mathbf{r}) \Rightarrow \beta = \beta(\mathbf{r})$

The dilatational and rotational components of the displacement field are known as the P and S-waves, and α and β are the P and S-wave speed, respectively.

Another (more elegant) way to see that solutions of the wave equation are in fact P and S-waves is by realizing that any vector field can be represented by a combination of the gradient of some scalar potential and the curl of a vector potential. This decomposition is known as **Helmholtz's Theorem** and the potentials are often referred to as **Helmholtz Potentials**. Using Helmholtz's Theorem we can write for the displacement $\mathbf{u}$

$$\mathbf{u} = \nabla\Phi + \nabla \times \mathbf{\Psi} \tag{4.38}$$

and

$$\nabla.\mathbf{\Psi} = 0 \tag{4.39}$$

with Φ a rotation-free scalar potential (i.e. $\nabla \times \Phi = 0$) and $\mathbf{\Psi}$ the divergence-free *vector potential*. Substitution of (4.38) into the general wave equation (4.31) (and applying the vector identity (4.30)) we get :

$$\nabla[(\lambda + 2\mu)\nabla^2\Phi - \rho\ddot{\Phi}] + \nabla \times [\mu\nabla^2\mathbf{\Psi} - \rho\ddot{\mathbf{\Psi}}] = 0 \tag{4.40}$$

which is a third-order differential equation[3]. Equation (4.40) can be satisfied by requiring that both

$$(\lambda + 2\mu)\nabla^2\Phi - \rho\ddot{\Phi} = 0 \tag{4.41}$$

which is a scalar wave equation for the propagation of the rotation-free displacement field Φ with wave speed

$$\alpha = \sqrt{\frac{\lambda + 2\mu}{\rho}} = \sqrt{\frac{\kappa + 4/3\mu}{\rho}} \tag{4.42}$$

and

$$\mu\nabla^2\mathbf{\Psi} - \rho\ddot{\mathbf{\Psi}} = 0 \tag{4.43}$$

[3]Strictly speaking this is not the way to formulate the problem. The need to solve third-order differential equations could have been avoided if the problem was set up in a different way by making use of what is known as **Lamé's theorem**. This also involves Helmholtz potentials. See, for instance, Aki & Richards, **Quantitative Seismology** (1982) p. 67-69. This mathematical correctness is, however, not required for a basic understanding of the decomposition in P and S terms

which is a vector wave equation for the propagation of the divergence-free displacement field $\mathbf{\Psi}$ with wave speed

$$\beta = \sqrt{\frac{\mu}{\rho}} \tag{4.44}$$

Comparing Eq. 4.33 and 4.41, we can identify Φ with the volume change ($\boldsymbol{\nabla} \cdot \mathbf{u}$ is called the cubic dilatation). Similarly, $\mathbf{\Psi}$ can be identified with the rotational component of the displacement field by comparing Eq. 4.36 and 4.42.

It is often much easier to solve the wave equations (4.41) and (4.43) than to solve the equation of motion directly for $\mathbf{u}$, and from the solution for the potentials the displacement $\mathbf{u}$ can then determined directly by Eq. (4.38). Note that even though P and S-waves are often treated separately, the *total* displacement field comprises both wave types.

Let's now consider a Cartesian coordinate system with z oriented downward, x parallel with the plane of the paper, and y out of the paper. We'll make the x-z plane the special plane of the problem. Because $\partial/\partial y \Phi \hat{\mathbf{y}} = 0$, we can write :

$$\boldsymbol{\nabla}\Phi = \frac{\partial}{\partial x}\Phi\hat{\mathbf{x}} + \frac{\partial}{\partial z}\Phi\hat{\mathbf{z}} \tag{4.45}$$

and

$$\boldsymbol{\nabla} \times \mathbf{\Psi} = \begin{vmatrix} \partial/\partial x & \partial/\partial y & \partial/\partial z \\ \Psi_x & \Psi_y & \Psi_z \\ \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \end{vmatrix} \tag{4.46}$$

Therefore,

$$\begin{aligned} u_x &= \frac{\partial \Phi}{\partial x} - \frac{\partial \Psi_y}{\partial z} \\ u_y &= \frac{\partial \Psi_z}{\partial x} - \frac{\partial \Psi_x}{\partial z} \\ u_z &= \frac{\partial \Phi}{\partial z} + \frac{\partial \Psi_y}{\partial x} \end{aligned} \tag{4.47}$$

The displacement direction from Φ is in the x-z plane and it is compressional — Φ is the P-wave potential. P wave propagation is thus rotation-free and has no components perpendicular to the direction of wave propagation, $\mathbf{k}$: it is a longitudinal wave with particle motion in the direction of $\mathbf{k}$. In contrast, the particle motion associated with the purely rotational S-wave is in a plane perpendicular to $\mathbf{k}$: transverse particle motion can be decomposed into vertical polarization, the so-called SV wave, and horizontal polarization, the so-called SH-wave (see Fig. 4.2) The displacement u_y from the SV-wave potential is in the same plane. In this formulation, u_y could just as well have been called the S_H-wave potential with displacement direction perpendicular to the x-z-plane.

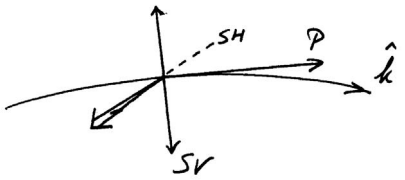


Figure 4.2: P and S waves : partical motion and propagation direction.

4.7 From vector to scalar potentials – Polarization

Using the Fourier transform, we show that the vector decomposition with Φ and $\mathbf{\Psi}$ can be reduced into three equations with the scalar potentials Φ, Ψ_{SV} and Ψ_{SH} (waves are typically described by oscillatory functions, i.e. complex exponentials. It is therefore natural to move the analysis to the frequency domain, i.e. to use Fourier transforms). We will write $\mathbf{u}(\mathbf{r}, t)$ for the time and space domain displacement, and $\mathbf{u}(\mathbf{r}, \omega)$ for the displacement in space and frequency. The transformation to the frequency domain is done by means of the (temporal) Fourier transform, which is defined as :

$$\mathbf{u}(\mathbf{r}, \omega) = \int_{-\infty}^{\infty} \mathbf{u}(\mathbf{r}, t) e^{i\omega t}\, dt \tag{4.48}$$

$$\mathbf{u}(\mathbf{r}, t) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \mathbf{u}(\mathbf{r}, \omega) e^{-i\omega t}\, d\omega \tag{4.49}$$

It is easy to see how the time derivative in a partial differential equation (PDE) brings out a factor of $i\omega$. This can be verified using the PDE obtained in section 4.6 :

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = (\lambda + 2\mu)\nabla(\nabla \cdot \mathbf{u}) - \mu \nabla \times (\nabla \times \mathbf{u}) \tag{4.50}$$

The separation of the equation of motion into two parts was done in section 4.6. It can also be done in the frequency domain : using Eq. 4.49 and Eq. 4.38, Eq. 4.50 (the equation of motion) becomes :

$$\omega^2 \Phi = -\alpha^2 \boldsymbol{\nabla} \cdot \mathbf{u} \tag{4.51}$$

$$\omega^2 \mathbf{\Psi} = -\beta^2 \boldsymbol{\nabla} \times \mathbf{u} \tag{4.52}$$

We thus easily get :

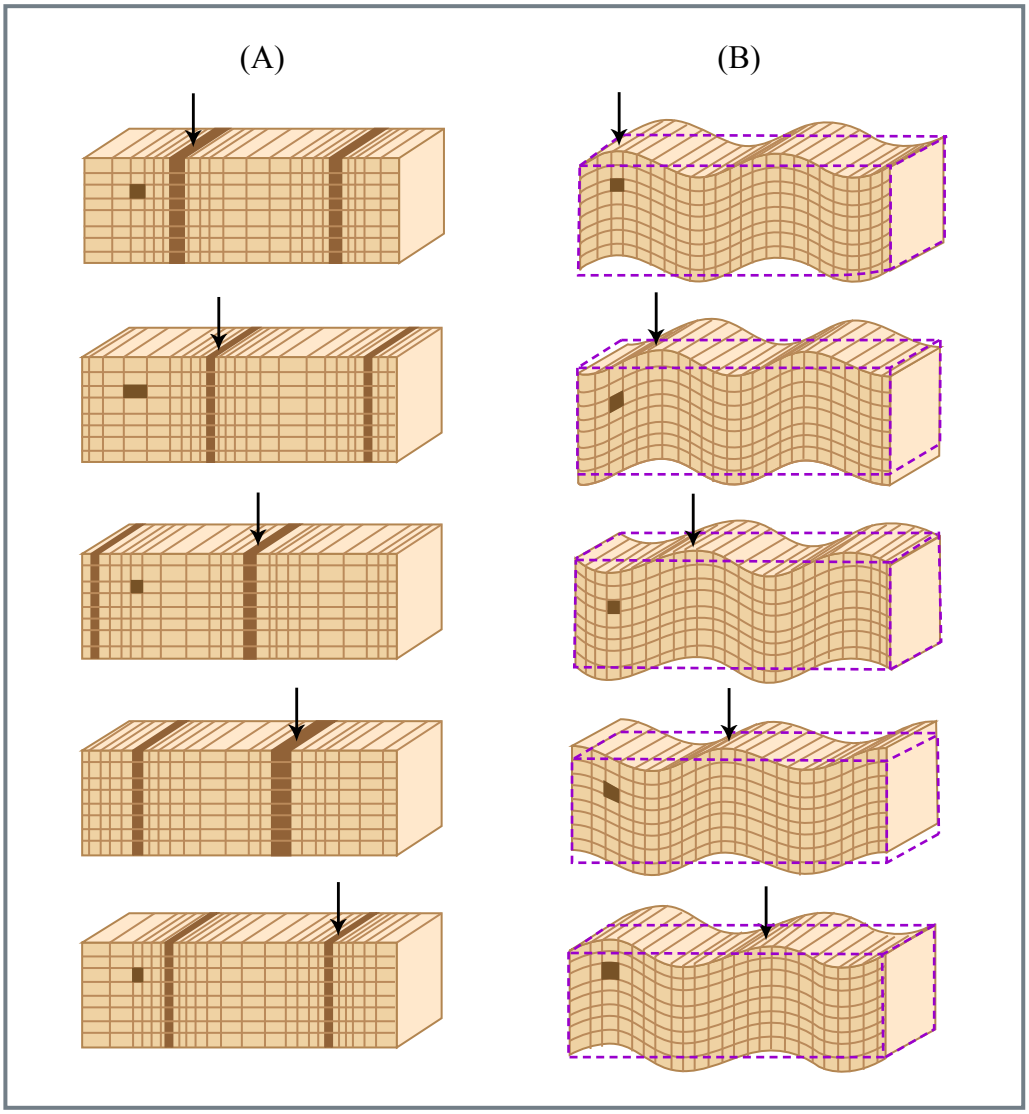


Figure by MIT OCW.

Figure 4.3: Successive stages in the deformation of a block of material by P-waves and SV-waves. The sequences progress in time from top to bottom and the seismic wave is travelling through the block from left to right. Arrow marks the crest of the wave at each satge. (a) For P-waves, both the volume and the shape of the marked region change as the wave passes. (b) For S-waves, the volume remains unchanged and the region undergoes rotation only.

$$\begin{aligned} \alpha^2\nabla^2\Phi &= -\omega^2\Phi \\ \beta^2\nabla^2\Psi_{SV} &= -\omega^2\Psi_{SV} \\ \beta^2\nabla^2\Psi_{SH} &= -\omega^2\Psi_{SH} \end{aligned} \tag{4.53}$$

We now have ordinary differential equations (ODEs), also known as Helmholtz equations, which are much easier to solve than PDEs. (NB one can readily see that $\nabla\Phi$ would lead to $-ik\Phi$ and $\nabla^2\Phi$ to $k^2\Phi$; therefore $k_\alpha^2 = frac\omega^2\alpha^2$, $k_\beta^2 = frac\omega^2\beta^2$.)

4.8 Solution by separation of variables

In a way, we've solved the wave equation by realizing that we could reduce it to an ordinary differential equation using the Fourier transform. So we knew the solution would be a complex exponential in the time variable (it is a "natural" way of describing a wave). We will now justify the validity of this approach by attempting to solve the following partial differential equation :

$$c^2\nabla^2\Phi = \frac{\partial}{\partial t^2}\Phi \tag{4.54}$$

without resorting to the Fourier transform.

If we propose a solution by separation of variables :

$$\Phi = X(x)Y(y)Z(z)T(t) \tag{4.55}$$

and plug Eq. 4.55 into Eq. 4.54, we obtain :

$$\frac{1}{X}\frac{d^2X}{dx^2} + \frac{1}{Y}\frac{d^2Y}{dy^2} + \frac{1}{Z}\frac{d^2Z}{dz^2} - \frac{1}{c^2T}\frac{d^2T}{dt^2} = 0 \tag{4.56}$$

The partial derivatives are regular derivatives now : we went from a PDE to ordinary differential equations (ODE). Each of these terms needs to be constant. We can pick these constants (ω^2 for T, and k_x^2, k_y^2 and k_z^2 for the spatial functions) but they will not be independent (they are linked to one another through Eq. 4.56). If we pick ω, k_x and k_y, then k_z is not independent anymore and satisfies :

$$k_z^2 = \frac{\omega^2}{c^2} - k_x^2 - k_y^2 \tag{4.57}$$

With those constants, it is easy to show that X, Y, Z and T are oscillatory functions :

$$\begin{aligned} X &\sim \exp(\pm ik_x x) \\ Y &\sim \exp(\pm ik_y y) \\ Z &\sim \exp(\pm ik_z z) \\ T &\sim \exp(\pm i\omega t) \end{aligned} \tag{4.58}$$

We have obtained solutions to the wave equation. Of course any linear combination of particular solutions leads to the general solution, and also we need to pick the sign in Eq. 4.58 (from the boundary conditions). Relation 4.57 is called a *dispersion relation.* k_x, k_y and k_z can be seen as the cartesian component of a vector $\mathbf{k}$ and Φ can be written as an oscillatory function of the type

$$\Phi \propto \exp(i(\mathbf{k}\cdot\mathbf{r} - \omega t)) \tag{4.59}$$

These waves are called *plane waves* and $\mathbf{k}$ is the direction of wave propagation.

4.9 Plane waves

We've called functions of the type $\exp(i(\mathbf{k}\cdot\mathbf{r}-\omega t))$ plane waves. Let's look at a few characteristics.

TRAVELING WAVES Let's first notice that plane waves are of the general form describing traveling waves :

$$\Phi(x,t) = f(x-ct) + g(x+ct) \tag{4.60}$$

with f and g *arbitrary* functions, provided that they are twice differentiable with regard to space and time (and that the second derivatives are continuous). After all, they need to solve $\ddot{\Phi} = c^2\nabla^2\Phi$. This is referred to as **d'Alembert's solution**. The function $f(x-ct)$ represents a disturbance propagating in the positive x direction with speed c. The function $g(x+ct)$ represents a disturbance propagating in the negative x axis : this part of the solution will be ignored in the following, but it must be taken into account when dealing with wave interference.

WAVELENGTH
With $k = 2\pi/\lambda$, the spatial part can be manipulated as follows:

$$e^{i\frac{2\pi}{\lambda}x} = e^{i\frac{2\pi}{\lambda}x}e^{i2\pi N} = e^{i\frac{2\pi}{\lambda}(x+N\lambda)} \tag{4.61}$$

to show that λ is indeed the wave length — after this distance, the displacement pattern repeats itself.

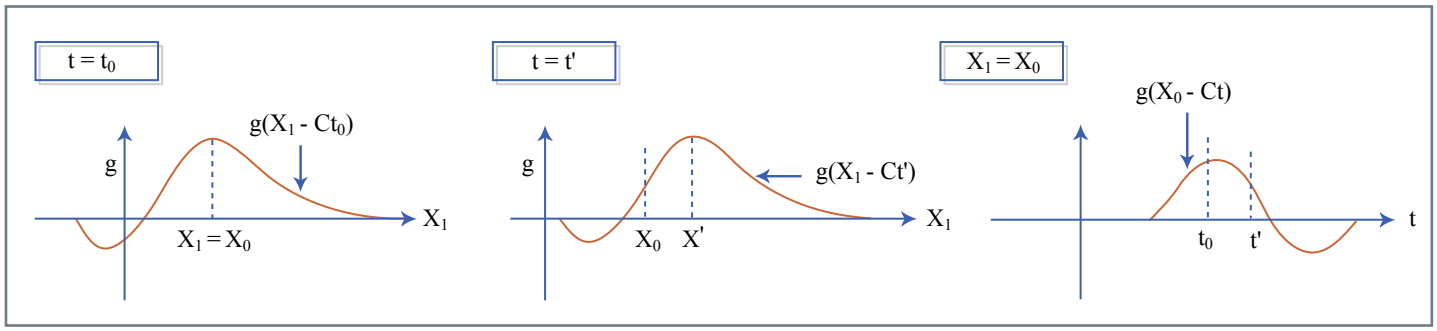


Figure by MIT OCW.

Figure 4.4: Plane waves : propagating disturbances.

PHASE
With increasing time t the argument of function f does not change provided that x also increases (hence the propagation in the positive x axis). In other words, if the argument remains constant it means that the shape defined by function f translates through space. The argument of f, $x-ct$, is referred to as the phase; one can define the wave front as the propagating function for a given

value of the phase. That c is the phase velocity is easily obtained by considering a constant phase at times t and t' ($x-ct = x'-ct' \Rightarrow c = (x'-x)/(t'-t) \approx dx/dt \equiv$ speed).

Wavefront

A wavefront is a surface through all points of equal phase, i.e. a surface connecting all points at the same travel time T from the source (see Fig. 4.5). In other words, at the wave front, all particles move in phase. Rays are the normals to the wave fronts and they point in the direction of wave propagation. The use of rays, ray paths, and wave fronts in seismology has many similarities with optics, and is called geometric ray theory.

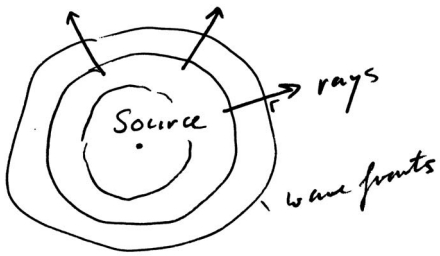


Figure 4.5: Seismic wavefronts.

Plane waves have plane wave fronts. The function Φ remains unchanged for all points on the plane perpendicular to the wave vector : indeed, on such a plane, the dot product $\mathbf{k} \cdot \mathbf{r}$ is constant.

At distances sufficiently far from the source body waves can be model-led as plane waves. As a rule of thumb : observer must be more than 5 wave lengths away from source to apply far field — or plane wave — approximation. Closer to the source one would need to consider spherical waves. Note that a seismogram corresponds to the recording of $\mathbf{u} = \mathbf{u}(\mathbf{r}_0, t)$ at a fixed position $\mathbf{r}_0$; i.e. the displacement as a function of time that records the passage of a wave group past $\mathbf{r}_0$.

Polarization direction

The polarization direction is different from the propagation direction. As already mentionned in sections 4.8 and 4.7, all waves propagate in the direction of their wave vector $\mathbf{k}$. The P-wave displacement ($\boldsymbol{\nabla}\Phi$) is parallel with the $\mathbf{k}$. The SV-displacement ($\boldsymbol{\nabla} \times (\hat{\mathbf{y}}\Psi)$) is perpendicular to this, in the $x - z$ plane, and the SH-displacement is out of the plane.

To indicate explicitly the propagation in the direction of or perpendicular to wave vector $\mathbf{k}$, one sometimes also writes

$$\begin{array}{cc} \text{for } P\text{-waves:} & \text{for } S\text{-waves:} \\ \Phi(\mathbf{r},t) = A_n \mathbf{k}\, e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} & \Phi(\mathbf{r},t) = \mathbf{B}_n \times \mathbf{k}\, e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \end{array} \tag{4.62}$$

Low- and high-frequency seismology

The variables used to describe the harmonic components are related as follows;

Angular frequency	$\omega = kc$
Wavelength	$\lambda = cT = 2\pi/k$
Wavenumber	$k = \omega/c$
Frequency	$f = \omega/2\pi = c/\lambda$
Period	$T = 1/f = \lambda/c = 2\pi/\omega$

Seismic waves have frequencies f ranging roughly from about 0.3 mHz to 100 Hz. The longest period considered in seismology is that associated with fundamental free oscillations of the earth : $T \approx 59$ min. For a typical wave speed of 5 km/s this involves signal wavelengths between 15,000 km and 50m. A loose subdivision in seismological problems is based on frequency, although the boundaries between these fields are vague (and have no physical meaning) :

low frequency seismology	f <20 mHz	$\lambda > 250$ km
high frequency seismology	50 mHz $< f <$10 Hz	0.5 km $< \lambda <$ 100 km
exploration seismics:	$f > 10$ Hz	$\lambda <$ 500 m

4.10 Some remarks

1. The existence of P and S-waves was first demonstrated by Poisson (in 1828). He also showed that P and S-type waves are, in fact, the only solutions of the wave equations for an *unbounded* medium (a 'whole' space), so that $\mathbf{u} = \nabla\Phi + \nabla\times\mathbf{\Psi}$ provides the complete solution for the displacement in an elastic, isotropic and homogeneous medium. Later we will see that if the medium is not unbounded, for instance a half-space with perhaps some stratification, there are more solutions to the general equation of motions. Those solutions are the surface (Rayleigh and Love) waves.

2. Since $\kappa > 0$ and $\mu \geq 0 \Rightarrow \alpha > \beta$: P-waves propagate faster than shear waves! See Fig. 4.6.

3. It can be shown that independent propagation of the P and S-waves is only guaranteed for sufficiently high frequencies (the so-called **high-frequency approximation**, "high frequency" in the sense that spatial variations in

elastic properties occur over much larger distances than the wavelength of the waves involved) underlies most (but not all) of the theory for body wave propagation).

4. The three components of the wave field (P, SV, and SH-waves, see section 4.7 for more details) can be recorded completely with three orthogonal sensors. In seismometry one uses a vertical component [Z] sensor along with two horizontal component sensors. In the field the latter two are oriented along the North-South [N] and East-West [E] directions, respectively. Fig. 4.7 is an example of such a three-component recording; we will come back to this in more detail later in the course.

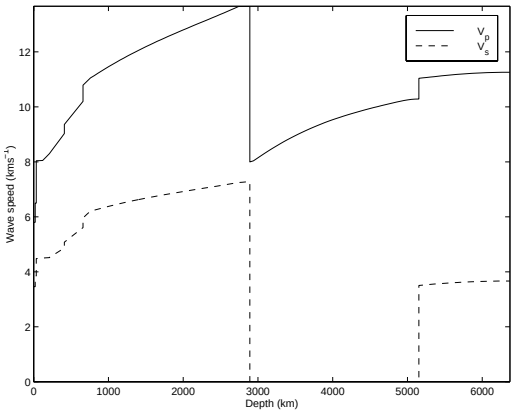


Figure 4.6: P and S wave speed in the ak135 Earth model.

4.11 Nomenclature of body waves in Earth's interior

At this stage it is useful to introduce the jargon used to describe the different types of body wave propagation in Earth's interior. We will get back to several wave propagation issues in more detail after we have discussed the basics of ray theory and the construction and use of travel time curves. There are a few simple basic "rules", but there are also some inconsistencies :

- Capital letters are used to denote body wave propagation (transmission) *through* a medium. For example, P and S for the compressional and shear waves, respectively, K and I for outer and inner core propagation of compressional waves (K for German 'Kerne'; I for Inner core), and J for shear wave propagation in the Inner Core (no definitive observations of this seismic phase, although recent research has produced compelling evidence for its existence).

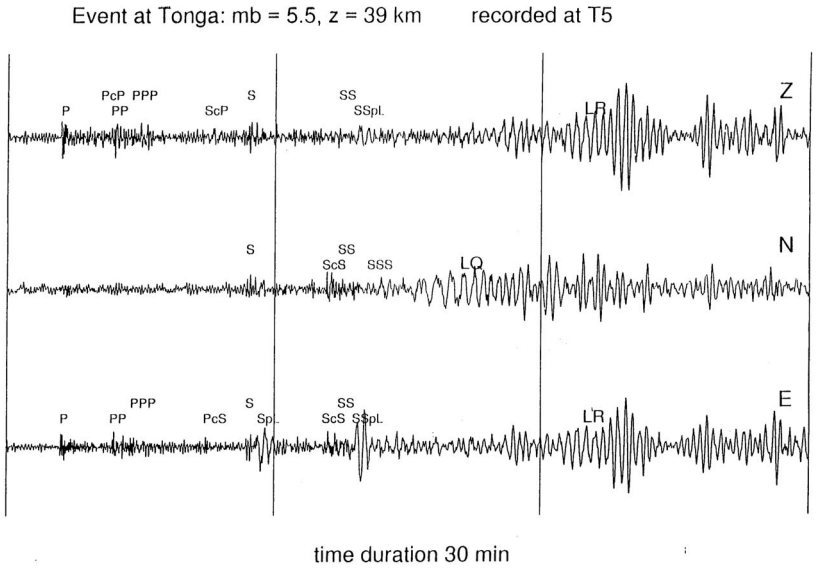


Figure 4.7: Example of a three-component seismic record

- Lower case letters are either used to indicate either **reflections** (e.g., c for the reflection at the CMB, i for the reflection at the ICB, and d for reflections at discontinuities in the mantle, with d standing for a particular depth (e.g., '410' or '660' km), or **upward propagation** of body waves before they are reflected at Earth's surface (e.g., s for an upward traveling shear wave, p for an upward traveling P wave). Note that this is *always* used in combination of a transmitted wave : for example, the phase pP indicates a wave that travels upward from a deep earthquake, reflects at the Earth's surface, and then travels to a distant station.

P	a P-wave in the mantle
S	an S-wave in the mantle
K	a P-wave through the outer core
I	a P-wave through the inner core
J	an S-wave through the inner core
c	a reflection from the mantle–outer core boundary
i	a reflection from the outer core–inner core boundary
p	a P-wave reflected from the surface of the earth close to the earthquake focus
s	an S-wave reflected from the surface of the earth close to the earthquake focus
LR	a Rayleigh wave
LQ	a Love wave

Figure 4.8: Nomenclature of body waves

4.12 More on the dispersion relation

We have already introduced the concept of dispersion (Eq. 4.57). Searching for a solution by separation of varibles, we have seen that the solution to the wave equation is an exponential both in the time and space domain. We had, however,already shown the oscillatory behavior of the solution in the time domain by using the time Fourier transform. In this section, we go one step further. Predicting that the solution will be a complex exponential in the spatial domain as well, we will investigate what insight the spatial Fourier-transform will bring us. Time and space are linked through the wave equation (it is a PDE) – the linkage between them is by the dispersion relation which we are deriving here.

As definition for the spatial Fourier transform and its inverse, we take

$$\Phi(\mathbf{k},\omega) = \int_V \Phi(\mathbf{r},\omega) e^{-i\mathbf{k}\cdot\mathbf{r}}\, d^3\mathbf{r} \tag{4.63}$$

and

$$\Phi(\mathbf{r},\omega) = \frac{1}{(2\pi)^3} \int_K \Phi(\mathbf{k},\omega) e^{i\mathbf{k}\cdot\mathbf{r}}\, d^3\mathbf{k} \tag{4.64}$$

The integrations are over all of physical space V ($dxdydz$) and all of wave vector space K ($dk_x dk_y dk_z$), respectively. The dot product $\mathbf{k}\cdot\mathbf{r} = k_q x_q$ with the Einstein summation convention. Remember also that $k_p^2 = k_p k_p = |\mathbf{k}| = k^2$.We need the Laplacian of Φ, this is given by :

$$\nabla^2\Phi = \frac{\partial^2}{\partial x_p \partial x_p} = \frac{1}{(2\pi)^3} \int_K \Phi(\mathbf{k},\omega) e^{ik_q x_q} i^2 k_p^2\, d^3\mathbf{k} \tag{4.65}$$

Comparison with Eq. 4.54 leads to (call α or β now c) :

$$-k^2 + \frac{\omega^2}{c^2} = 0 \quad \text{or} \quad |\mathbf{k}| = \left|\frac{\omega}{\alpha}\right| \tag{4.66}$$

We can quickly convert this dispersion relation into something you're all familiar with : with $k = 2\pi/\lambda$ and $f = \omega/(2\pi)$, we get $\lambda f = c$: the frequency of a wave times its wave lengths gives the propagation speed. We will discuss this in more detail below.

The complete solution to the wave equation is thus given by inverse transformation of $\Phi(\mathbf{r},\omega)$ as follows :

$$\Phi(\mathbf{r},t) = \frac{1}{(2\pi)^4} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \Phi(k_x,k_y,\omega,z) e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\, dk_x\, dk_y\, d\omega \tag{4.67}$$

There are three independent quantities involved here (not four) : k_x, k_y and ω, and their relationship is given by the dispersion equation. In other words,

$$\mathbf{k}\cdot\mathbf{r} = k_x x + k_y y + z\left(\frac{\omega^2}{c^2} - k_x^2 - k_y^2\right)^{1/2} \tag{4.68}$$

It's important to see Eq. 4.67 as what it is : a superposition (integral) of plane waves with a certain wave vector and frequency, each with its own amplitude. The amplitude is a coefficient which will have to be determined from the initial or boundary conditions.

We thus have seen that the dispersion equation can be obtained either by solving the wave equation by separation of variables or by introducing the time and spatial Fourier transforms.

4.13 The wave field — Snell's law

In this section, we'll use plane wave displacement potentials to solve a simple problem of wave propagation. Not only will we understand why and how reflections, refractions and phase conversions happen, but we'll also derive an important relation for plane waves in planar media known as Snell's law.

Let's start with a plane P-wave incident on the free surface, making an angle with the normal i. We can identify the P-wave with its wave vector. In our case, we know that

$$k_x = \left|\frac{\omega}{\alpha}\right| \sin i \quad \text{and} \quad k_z = -\left|\frac{\omega}{\alpha}\right| \cos i \tag{4.69}$$

Two kinds of boundary conditions are used in seismology — there are the kinematic ones, which put constraints on the displacement, and the dynamic ones, which constrain the stresses or tractions. The free surface needs to be traction-free. We remember that the traction vector was given by dotting the stress tensor into the normal vector representing the plane on which we are computing the tractions : $t_i = \sigma_{ij} n_j$. For a normal vector in the positive z-direction, the traction becomes :

$$\mathbf{t}(\mathbf{u}, \hat{\mathbf{z}}) = (\sigma_{xz}, \sigma_{yz}, \sigma_{zz}) \tag{4.70}$$

For isotropic materials, we have seen the following definition for the stress tensor :

$$\sigma_{ij} = \lambda(\boldsymbol{\nabla} \cdot \mathbf{u})\delta_{ij} + \mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \tag{4.71}$$

Tractions due to the P wave

We know that the displacement is given by the gradient of the P-wave displacement potential Φ (see Eq. 4.47) :

$$\mathbf{u} = \boldsymbol{\nabla}\Phi = \left(\frac{\partial \Phi}{\partial x}, 0, \frac{\partial \Phi}{\partial z} \right) \tag{4.72}$$

Therefore the required components of the stress tensor are :

$$\sigma_{xz} = 2\mu\frac{\partial^2\Phi}{\partial x\partial z} \tag{4.73}$$
$$\sigma_{yz} = 0 \tag{4.74}$$
$$\sigma_{zz} = \lambda\nabla^2\Phi + 2\mu\frac{\partial^2\Phi}{\partial^2 z} \tag{4.75}$$

Tractions due to the SV wave

The displacement is given as the rotation of the Ψ potential (see Eq. 4.47) :

$$\mathbf{u} = \left(-\frac{\partial\Psi}{\partial z}, 0, \frac{\partial\Psi}{\partial x}\right) \tag{4.76}$$

For the stress tensor, we find :

$$\sigma_{xz} = \mu\left(\frac{\partial^2\Psi}{\partial x^2} - \frac{\partial^2\Psi}{\partial z^2}\right) \tag{4.77}$$
$$\sigma_{xz} = 0 \tag{4.78}$$
$$\sigma_{zz} = 2\mu\frac{\partial^2\Psi}{\partial x\partial z} \tag{4.79}$$

Tractions due to the S_H wave

The S_H wave, as we've seen, has only one component in this coordinate system :

$$\mathbf{u} = (0, u_y, 0) \tag{4.80}$$

and the stress tensor components are given by

$$\sigma_{xz} = 0 \tag{4.81}$$
$$\sigma_{yz} = \mu\frac{\partial u_y}{\partial z} \tag{4.82}$$
$$\sigma_{zz} = 0 \tag{4.83}$$

Comparing Eqs. 4.75 and 4.79, we see how P and SV waves are naturally coupled. In this plane-wave plane-layered case, the P-wave had energy only in the x- and z-component, and so did SV. Upon reflection and refraction, energy can be transferred from the incoming P-wave to a reflected P-wave and a reflected SV-wave. No S_H waves can enter the system — they have all their energy on the y-component.

Analogously to Eq. 4.69, we can represent the incoming P, the reflected P and the reflected SV wave by the following slownesses :

$$P^{\text{inc}} = \left(\frac{\sin i}{\alpha}, 0, \frac{-\cos i}{\alpha}\right) \tag{4.84}$$

$$P^{\text{refl}} = \left(\frac{\sin i^*}{\alpha}, 0, \frac{\cos i^*}{\alpha}\right) \tag{4.85}$$

$$SV^{\text{refl}} = \left(\frac{\sin j}{\beta}, 0, \frac{\cos j}{\beta}\right) \tag{4.86}$$

Thus the total P-potential Φ is made up from the incoming and reflecting P-wave, and the shear-wave potential Ψ is given by the reflected SV-wave. All of them, of course, have the plane wave form, so that we can write :

$$\Phi^{\text{inc}} = A \exp\left[i\omega\left(\frac{\sin i}{\alpha}x - \frac{\cos i}{\alpha}z - t\right)\right] \tag{4.87}$$

$$\Phi^{\text{refl}} = B \exp\left[i\omega\left(\frac{\sin i^*}{\alpha}x + \frac{\cos i^*}{\alpha}z - t\right)\right] \tag{4.88}$$

$$\Psi^{\text{refl}} = C \exp\left[i\omega\left(\frac{\sin j}{\alpha}x + \frac{\cos j}{\alpha}z - t\right)\right] \tag{4.89}$$

As pointed out before, there are no kinematic boundary conditions on the free surface. The displacement of the free surface is unconstrained, and above it there is no displacement at all. The dynamic boundary conditions, however, are non-trivial. The tractions must vanish on the free surface : so $\sigma_{xz} = \sigma_{yz} = \sigma_{zz} = 0$ at $z = 0$. It is easy to see that, with $z = 0$, the sum of the three plane wave displacement potentials will be of the type

$$\begin{aligned} A \exp\left[i\omega\left(\frac{\sin i}{\alpha}x - t\right)\right] &+ B \exp\left[i\omega\left(\frac{\sin i^*}{\alpha}x - t\right)\right] \\ &+ C \exp\left[i\omega\left(\frac{\sin j}{\alpha}x - t\right)\right] \end{aligned}$$

Hence, for this sum to be zero for all x and t, we need :

$$\boxed{\frac{\sin i}{\alpha} = \frac{\sin i*}{\alpha} = \frac{\sin j}{\beta} \equiv p} \tag{4.90}$$

Thus, for plane waves in plane-layered media, the whole system of rays is characterized by a common horizontal slowness. This is true for the whole wave field of reflected and transmitted (refracted) waves. Eq. 4.90 is known as Snell's law and p is called the **ray parameter**. In the following paragraph, a more general principle called Fermat's principle is used to prove Snell's Law.

4.14 Fermat's Principle and Snell's law

An important principle in optics is Fermat's principle, which governs the geometry of ray paths. This principle states that a wave propagating from position A to position B follows a path of **stationary time**. The principle of stationary time plays a fundamental role in high frequency seismology. Note that stationary time does not necessarily mean *minimum* time; it can also be a maximum time.

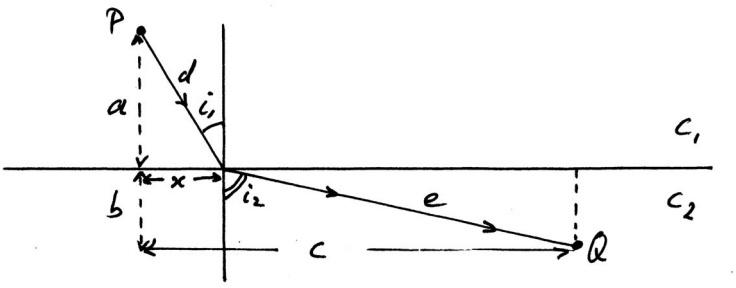


Figure 4.9: The principle of stationary time.

Consider Fig. 4.9. A ray leaves point P that is in a medium with wave speed c_1 and travels to point Q in a medium with wave speed c_2. What path will the ray take to Q? Since the wave speeds in the media are constant the ray path in each medium is a straight line, so that in this simple case the geometry is completely defined by the positions of P, Q, and the point x where the ray crosses the interface.

The travel time on an arbitrary path between P and Q is given by

$$t_{P-Q} = \frac{d}{c_1} + \frac{e}{c_2} = \frac{\sqrt{a^2 + x^2}}{c_1} + \frac{\sqrt{b^2 + (c-x)^2}}{c_2} \tag{4.91}$$

For the path to be a stationary time path (i.e. time is maximum or minimum) we simply set the spatial derivative of the travel time to zero :

$$\frac{dT}{dx} = 0 = \frac{x}{c_1\sqrt{a^2 + x^2}} - \frac{c-x}{c_2\sqrt{b^2 + (c-x)^2}} \tag{4.92}$$

and note that

$$\frac{x}{\sqrt{a^2 + x^2}} = \sin i_1 \quad \text{and} \quad \frac{c-x}{\sqrt{b^2 + (c-x)^2}} = \sin i_2 \tag{4.93}$$

This gives Snell's law :

$$\boxed{\frac{\sin i_2}{c_2} = \frac{\sin i_1}{c_1} \equiv p} \tag{4.94}$$

p is called the **ray parameter**.

One can expand on this simple geometry and consider many more layers, but the result is the same : *the ray parameter p is constant along the entire ray*! As a ray enters material of increasing velocity, the ray is deflected toward the horizontal; if the ray enters material with lower velocity, the ray is deflected to the vertical. In seismology the angle between the ray and the vertical is referred to as the **angle of incidence** (also, *take-off angle*).

4.15 Ray geometries of the wave field

For most applications we have to deal with a complex wave field : in each layer of a stratified medium there can be 6 different body wave groups : the up- and down-going P, SV, and SH-waves. The propagation of such a wave field through a stratified medium (a stack of horizontal layers or spherical shells in which the wave speed is constant) is controlled by Snell's law (Fermat's Principle) and boundary conditions.

The wave field is determined by **reflections**, **refractions**, and **phase conversions**; for instance, a down-going P wave can reflect at an interface and part of its energy can be transmitted to the other side, and part of its energy can (or often *has to be*) converted to SV-wave energy (see Fig. 4.10).

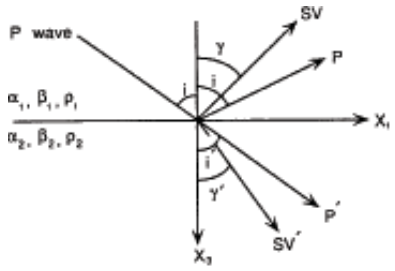


Figure 4.10: Ray conversions at interfaces.

The incidence angles of the reflected and refracted waves that compose this complex wave field are controlled by an extended form of Snell's law. For this example, Snell's law is :

$$\frac{\sin i_1}{\alpha_1} = \frac{\sin j_1}{\beta_1} = \frac{\sin i_2}{\alpha_2} = \frac{\sin j_2}{\beta_2} \equiv p \tag{4.95}$$

This generalization of Snell's law shows an important concept that the whole system of seismic waves produced by reflection and transmission of plane waves in a stratified medium is characterized by the value of their common horizontal slowness, or the ray parameter p. It can also be used directly to determine the angles for critical reflection and refraction. The ray parameter is constant not only for a single ray, but for the entire wave field generated by reflection and refraction of an incoming P or S-wave.

4.16 Travel time curves and radial Earth structure

We have been developing some basic theory and concepts of body wave seismology. One of the major objectives of seismology is to extract structural information about Earth's structure from the observed data, the seismograms. We will discuss some rather classical techniques to do this.

Snell's Law

We derived Snell's law for a "flat" Earth :

$$\boxed{\frac{\sin i_1}{c_1} = \frac{\sin i_2}{c_2} = \ldots = \frac{\sin i_n}{c_n} = \text{constant} \equiv p, \text{ the ray parameter}} \quad (4.96)$$

The ray parameter is constant along the entire ray path, and is the same for all rays (reflections, refractions, conversions) associated with the same incoming ray. The ray parameter plays a very important role in seismology.

Snell's Law shows that the ray parameter is inversely proportional to velocity, or proportional to 1/velocity, which is the **slowness**. In seismology it is often more convenient to use *slowness* instead of *wave speed*. One significant advantage of the slowness vector is that it can be added vectorially, whereas this is not always justified (in our context) for the velocity.

$$\mathbf{s} = (s_1, s_2, s_3) = s_1\mathbf{x}_1 + s_2\mathbf{x}_2 + s_3\mathbf{x}_3 \quad (4.97)$$

The vector summation for velocity can give practical problems : consider, for instance, the plane wave that propagates in the direction $\mathbf{k}$. The apparent velocity c_1 measured at the surface (from observations at several stations) is *larger* than the true velocity c : with i the angle of incidence, $c_1 = c/\sin i > c$, so that $c \neq c_1 + c_3$.

From Fig. 4.11 we can easily derive two other important relationships :

$$\sin i = \frac{ds}{dx_1} = c\frac{dt}{dx_1} = \frac{c}{c_1} \Rightarrow \boxed{p = \frac{\sin i}{c} = \frac{dt}{dx_1} = \frac{1}{c_1}} \quad (4.98)$$

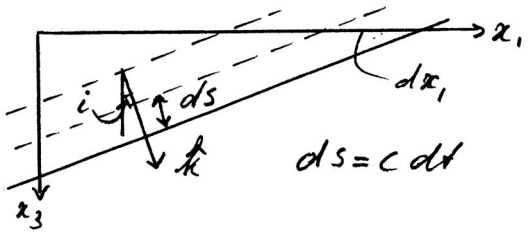


Figure 4.11: Derivation of Snell's law.

1. The ray parameter p is $1/c_x$, which is referred to as the **horizontal slowness**!
2. the ray parameter is simply the derivative of the travel time T with horizontal distance. This will prove to be of major importance (and convenience!).

For a spherical earth we can derive a relationship for the ray parameter that is similar to Eq. (4.98), the "only" difference being the 'scale' factor r :

$$\boxed{p = r\frac{\sin i}{v(r)}} \tag{4.99}$$

where r is the radius to any point along the ray path, and $v(r)$ the wave speed at that radius. It can also be shown that (with Δ the angular distance)

$$p = \frac{\partial T}{\partial \Delta} \tag{4.100}$$

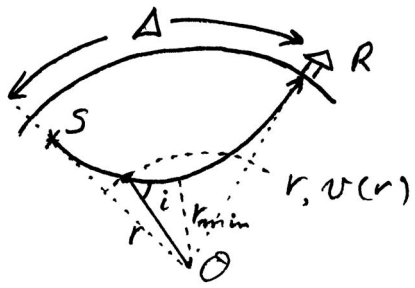


Figure 4.12: Ray parameter in spherical geometry

Notice the similarity between the definition of the ray parameter as the spatial derivative of travel time for the "flat" (Eq. 4.98) and spherical earth (Eq. 4.100)! Beware : For a flat earth the unit of ray parameter is s/km (or s/m), for the spherical earth it is either s/rad or just s or s/deg, so even though the definitions are completely equivalent there are differences in units!

With the definition for the ray parameter in a spherical Earth (Eq. 4.99) we can also get a simple expression that relates p to the minimum radius (or maximum depth) along the ray path : this point is known as the turning or bottoming point of the ray. A "turning ray" is the spherical Earth equivalent of the "head wave" (see Fig. 4.12).

$$\frac{r_{\min}\sin 90}{v(r_{\min})} = \boxed{\frac{r_{\min}}{v(r_{\min})} = p} \tag{4.101}$$

Under the assumption of a reference earth model for seismic wave speeds we can determine the horizontal distance traveled by the ray (e.g., from 4.98) and the depth to the turning point (from Eq. 4.101) once we know the ray parameter. Before showing how the ray parameter can be determined from observed data, let me mention another important concept based on the ray parameter :

Travel time curves

Eq. (4.98) indicates that the ray parameter, i.e. the *horizontal slowness*, can be determined from seismic data by determining the difference in travel time of a phase arrival at two adjacent stations. Ideally one uses an *array* of instruments to do this accurately.

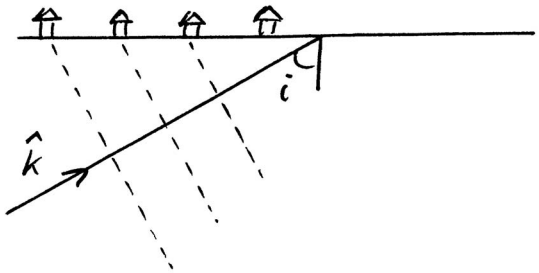


Figure 4.13: Determination of the ray with a seismometer array

In other words, one can determine the value of the ray parameter directly from the travel time curve, which represents the variation of travel time as a function of distance : $T(X)$ or $T(\Delta)$. A travel time curve can be constructed by arranging observed records of ground motion due to the same explosion or earthquake as a function of distance. In such a **record section** the travel time curve of a particular phase is just the curve that connect onset times of that phase in all records. One could also construct a travel time curve by using many measurements, *phase picks*, of the travel time of particular phases, say the P-phase, at different distances from the source. Seismologists try to find simple models of radial variations of wave speed that produce travel time curves consistent with the observed data. "Theoretical" travel time curves in this sense are thus best fitting curves determined from some reference model of seismic wave speeds.

Well known models for the Earth's depth dependent structure are the **Preliminary Reference Earth Model** (PREM) by Dziewonski & Anderson (1981), and the more recent **iasp91** model (Kennett & Engdahl, 1991). Typically, this fitting is not done by trial and error but by means of inversion of either the travel times or the travel time curves. A classical approach that is discussed in most text books is the one first applied by Herglotz and Wiechert in the beginning of this century. They were the first to invert travel time data for simple radially stratified models of seismic wave speed, and their technique has been used for decades. The first comprehensive model and the corresponding travel time tables was published by Jeffreys & Bullen (1939/1940). In fact their model, known as the **JB model**, is still being used for routine earthquake location by the International Seismological Centre in the U.K.

The ray parameter of a seismic wave (group) arriving at a certain distance can be thus be determined from the *slope* of the travel time curve. The straight line tangent to the travel time curve at Δ can be written as a function of the

intercept time σ and the slope p :

$$p = \frac{\partial T}{\partial \Delta} \Rightarrow T(\Delta) = \sigma + \frac{\partial T}{\partial \Delta}\Delta = \sigma + p\Delta \tag{4.102}$$

and this equation forms the basis of what is known as the $\sigma - p$ method.

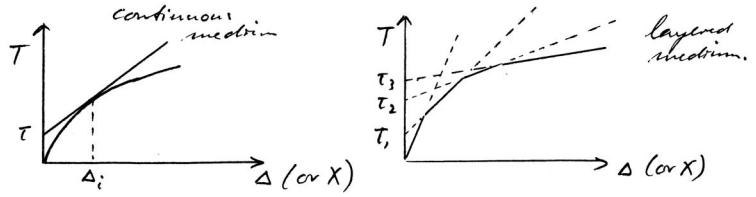


Figure 4.14: Determination of the ray parameter from the travel-time curve

The (local) slope of the travel time curve contains important information about the horizontal slowness, and thus about the wave speed, and the intercept time σ, the **zero offset time**, contains information about the layer thickness. This property is exploited in exploration seismics, where we typically deal with travel time "curves" that consist of segments of straight lines (see Fig. 4.14).

Another piece of information that can be obtained from travel time curves is contained in the *second* derivative of the travel time curve with distance, or the variation of ray parameter with distance $\partial p/\partial \Delta$. This quantity controls the *amplitude* of the arrivals. To see this, consider a situation (that we will discuss in more detail below) in which rays with different incident angles at the source (and receiver) are somehow focused to travel to the same seismographic station so that the amplitude increases. In that case, $\delta p \neq 0$ but $\delta \Delta = 0$ so

$$\frac{\partial p}{\partial \Delta} = \frac{\partial^2 T}{\partial \Delta^2} \rightarrow \infty \tag{4.103}$$

In other words, the larger $\partial p/\partial \Delta$, the more energy arrives at a small distance range $\delta \Delta$, and the higher the amplitude. In real life the amplitude of seismic waves is always finite, and this reveals, in fact, one of the shortcomings of ray theory. If rays are assumed to be infinitesimally narrow the theoretical amplitude can go to infinity, but in practice the amplitude remains finite as a result of the interference of the waves that arrive at the same time.

4.17 Radial Earth structure

In a spherical earth we typically encounter three important situations that are characterized by the geometry of the ray paths, the travel time curves $T(\Delta)$, the variation of the ray parameter with distance $p(\Delta)$, and the $\sigma(p)$ curves. In the following, just imagine what happens if you "shoot" rays from an earthquake source at the surface to increasing distances. In other words, you start of with

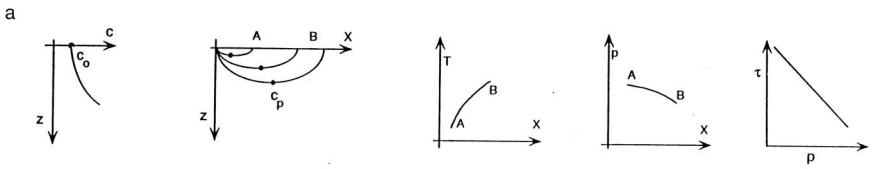


Figure 4.15: Case 1 : Wave speed monotonously increases with depth.

a large take-off angle and you analyze what happens when you decrease this angle (i.e. let the ray dive steeper into the Earth).

1. The situation that applies to most depth ranges in the Earth's interior is that of a steady increase in seismic wave speed (see Fig. 4.15) so that:

   - Ray paths : the rays sample progressively deeper regions in the Earth,
   - $T(\Delta)$: and arrive at progressively larger distances.
   - $p(\Delta)$: the slope of the travel time curves decrease monotonically with increasing distance (i.e. the ray parameter decreases for waves traveling to larger distances), so there are no significant changes in amplitude (other than those due to geometrical spreading!) ($\partial p/\partial \Delta < 0$).
   - the intercept time σ decreases with increasing ray parameter (decreasing distance!)

   A look at the travel time curves suggests that this situation is indeed very common and describes the overall character of the curves pretty well.

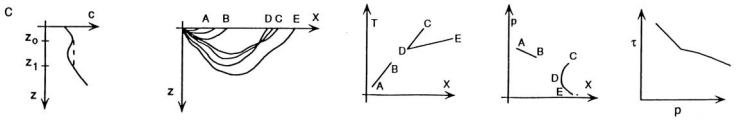


Figure 4.16: Case : The presence of a low-velocity zone.

2. The first important deviation from this situation is when there is a decrease in wave speed with increasing depth or decreasing radius (see Fig. 4.16). This gives rise to some interesting effects.

   - Ray paths : The rays will still sample progressively deeper regions when the ray parameter decreases, but the pattern is more complex. Initially (i.e. above the depth where the wave speed drops) the behavior is the same as in the general situation above. However, when the ray parameter decreases further the rays interact with the low velocity zone. (A sufficient condition for the 'low velocity zone' is that $\partial v/\partial r < v/r$.) The decrease in wave speed results in the deflection of the ray toward the vertical and the rays do not turn within the low velocity zone; they only reflect back to the Earth's surface to

be recorded by seismometers when the wave speed increases again. The corresponding waves arrive significantly farther away from the source than the ones with only a slightly less ray parameter. (You can also say that here we have a situation where $\delta p \approx 0$ but $\delta\Delta \neq 0$ so that the amplitude is zero.) Initially, some rays may reflect at the top of the "base" of the low velocity zone so that energy is projected to shorter distances with a further decrease in ray parameter (incidence angle), but eventually, the effect of the low wave speed zone is no longer felt and the rays sample deeper regions and behave in a manner similar to the general situation.

In terms of ray geometry : there will be a region in the Earth's interior that is not sampled.

- $T(\Delta)$: The travel time curve will reveal a **shadow zone**, a region where (according to our simplified – ray – theory based on the high frequency approximation) no phases arrive. There will be a small distance where two phases can arrive : the wave reflected from the base of the low velocity zone and the direct arrival which is the wave that turns beneath the LVZ.
- $p(\Delta)$: Initially, p will decrease with increasing distance ($\partial p/\partial\Delta < 0$), and $p(\Delta)$ is continuous. When p decreases so that the ray is refracted through the LVZ two things happen :
  (a) the $p(\Delta)$ curve is no longer continuous since the ray defined by the incrementally smaller p arrives at a different distance, and
  (b) with decreasing p the distance initially decreases because of the reflection ($\partial p/\partial\Delta > 0$). If p decreases even further the "normal" behavior is established again ($\partial p/\partial\Delta < 0$).
- Amplitude : The amplitude is zero in the shadow zone (the $p - \Delta$ curve is horizontal), but becomes large for arrivals at a distance just outside the shadow zone corresponding to rays that bottom just beneath the LVZ (the $p - \Delta$ curve is vertical).

The two most important regions in the Earth where this happens are the low velocity layer beneath oceanic lithosphere and at the transition from the mantle to the outer core (for P-waves).

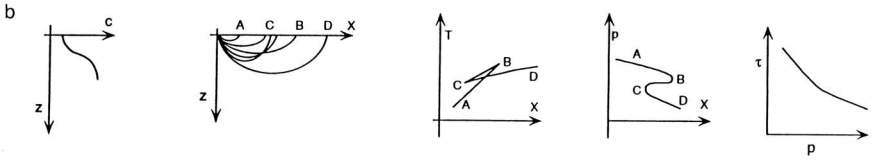


Figure 4.17: Case 3 : A sharp increase in wave speed with depth.

3. The second important deviation from the "normal" situation is when there is a region where the wave speed increases rapidly with increasing depth : $\partial v/\partial r >> 1$ (see Fig. 4.17). Let's for the discussion assume that the increase in wave speed occurs instantly, i.e. that there is a **seismic discontinuity** in $\partial v/\partial r$ (the function $v(r)$ itself is – of course – continuous; this situation is also known as a first order discontinuity), but you must realize that similar effects occur when the gradient in wave speed is steep.

   - Ray paths : For large incidence angles the rays turn above the discontinuity. These form the direct rays. When the incidence angle (or, equivalently, the ray parameter) decreases the rays will reflect at the interface. The ray with the smallest ray parameter that does not reflect is called the **grazing ray**. The rays that are reflected from the interface form arrivals at shorter distances those corresponding to the grazing ray. This leads to a situation where there is a distance range where we have arrivals of **both** the direct and the reflected waves. The situation is slightly more complicated because when the ray parameter continues to decrease, there is a critical angle where the rays no longer reflect but refract into the deeper earth. From that point onward, the behavior of the rays is as one would expect from the "normal" situation, and the rays go to larger distances. The reflection will cause the ray paths to cross which causes a **caustic** and results in large amplitudes of the phase arrivals.

   - $T(\Delta)$: The corresponding travel time curve is complicated. In the distance range between the arrival of the waves associated with the grazing ray and the critical ray there are, in fact, three arrivals : the direct phase propagating through the medium above the interface, the reflected phase, and the refracted wave that propagates in part in the medium beneath the interface. This distance range it, therefore, called the **triplication range** because there are, in fact, three arrivals.

   - $p(\Delta)$: For large ray parameters the behavior is as in the standard situation; a gradual increase in distance with decreasing p ($\partial p/\partial\Delta < 0$). When p becomes smaller than that of the grazing ray the reflection causes the distance to decrease with decreasing p ($\partial p/\partial\Delta > 0$), but when p decreases further and becomes smaller than the for the critical ray the distance increases again ($\partial p/\partial\Delta < 0$).

   - Amplitude : there are two points in the $p(\Delta)$ curve where $\partial p/\partial\Delta$ becomes very large (in ray theory the slope can go to infinity!). These two points correspond to the ray parameter for the grazing and critically refracted rays, respectively. Consequently, the amplitude of the phase arrival will be large on either end of the triplication range.

Final remarks

It is clear that the $\tau - p$ curves are the only curves associated with travel time curves that are continuous in all circumstances, and this is a very attractive property in, for instance, inversion of travel time information for Earth's structure. In fact, this curve also plays a central role in the computation of synthetic seismograms with the so-called WKBJ approximation.

A significant body of research is based on the arrival times of first arriving, direct phases such as P. In triplication zones there are typically more than two arrivals; there can be as many as 5 when triplication zones due to discontinuities at different depths overlap. The identification problem is aggravated due to the effect of the caustics on the amplitude : near the cusps in the travel time curve the later arriving triplication phases have significantly higher amplitude than the first arrival and for small signal to noise ratio in the data (for instance when there's a small earthquake) the first arrival that can be identified in the record can, in fact, be a later arriving phase. This causes substantial scatter in the arrival time data in these distance ranges.

The difficulty of phase identification in the triplications due to upper mantle discontinuities and the related uncertainty in the geometry of the ray paths involved has important implications for the imaging of upper mantle structure, which is more difficult than the imaging of lower mantle structure, and for the accurate location of earthquake hypocenters using these data.

In seismological literature one encounters the terms **regional** and **teleseismic distances**. The precise boundary between these distances is not well defined. It basically refers to the distance ranges where effects of an upper mantle low velocity layer and the discontinuities are (regional) or are not (teleseismic) significant. Regional distance is the distance where the associated rays bottom in the upper mantle and transition zone (i.e. above 660 km depth) and this is about 25° to 30°, with exact values dependent on the reference Earth model used. Teleseismic arrivals refer to arrivals beyond the triplication range and refer to turning rays in the lower mantle.

When waves pass through caustic (i.e. the arrivals on the receding branches of the travel time curves, for instance the PKP_{AB} phase and the reflections off a seismic discontinuity) the wave form will be distorted due to a 90° phase shift in the phase term $i\xi(\mathbf{r}, t)$. This will cause additional complications in picking the arrival time by hand. A better way is to generate synthetic waveforms that have the same phase shifts and apply cross correlation techniques.

4.18 Surface waves

Introduction

We have seen before that the solutions of the equations of motion in an unbounded, homogeneous, isotropic medium are remarkably simple and that the total displacement field due to a stress imbalance is completely accounted for by propagating P and S-waves. We also discussed how this body wave field becomes increasingly complex in the presence of interfaces, for instance the Earth's (free) surface, and the first order seismic discontinuities such as the Moho, the 410 and 660 km discontinuities, the CMB, and the ICB. The total P- and S-displacement field is then composed of up and downgoing SV and SH waves and their interaction is controlled by the reflection and transmission coefficients and by the boundary conditions at the interfaces.

In a bounded medium there is another important class of seismic waves, the *surface waves*; these are caused by the interaction of body waves with the free surface. Specifically, the interaction of the P-SV field with the free surface results in **Rayleigh** waves (after Lord Rayleigh, 1842-1919) whereas the interaction of the SH wave field with the free surface combines with internal layering to produce **Love waves** (after mathematician A. E. H. Love, 1843-1940, who predicted the existence of these waves in 1911). Later we will see that both the body waves and the surface waves can be represented by — and are equivalent with — a superposition of the **normal modes of free oscillation** of the Earth and it is important to be aware of the intimate relationship between these seemingly separate descriptions of wave propagation in the Earth's interior (see Table 4.3). All body waves propagating in the Earth's interior have counterparts in both propagating surface waves or standing free oscillations. However, each representation has distinct advantages for studying specific problems related to Earth's structure and the seismic source.

Body waves	Surface waves	Free oscillations
P-SV waves	Rayleigh waves	Speroidal modes
SH waves	Love waves	Toroidal modes

Table 4.3: Body waves, surface waves and free oscillation equivalencies.

General properties of surface waves

Surface waves propagate along the Earth's surface. This seems like a rather trivial statement but it has important implications for the amplitude of surface waves.

The cylindrical expansion of the wave front of the waves along the Earth's surface implies that the energy of surface waves decreases as 1 over r, with r the distance between the source and the position of the wave front. The amplitude of surface waves, related to the square root of the energy, therefore falls of as 1 over $\sqrt{r}$. In contrast, the geometrical spreading of body waves in the Earth's

interior implies that the energy decays as 1 over r^2 so that the amplitude of body waves decays as 1 over r. As a result of the difference in geometrical spreading, the amplitude of surface waves is typically much larger than that of body waves, in particular at larger distances from the source. (The distance from source to receiver is typically referred to as the *epicentral distance*).

Another implication of horizontal wave propagation and energy conservation is that surface waves are **evanescent**, i.e., the amplitude decays with increasing depth and goes to zero for very large depths. As a rule of thumb: the (fundamental mode of) surface waves are most sensitive at a depth $z = \lambda/3$, with λ the wave length, and their sensitivity becomes very small for $z > \lambda$. For example, at a period of $T = 100$s, the wavelength is about 450 km. Those waves are most sensitive in the upper 180 km of the mantle (where the shear wave speed is about 4.5 km/s).

The fact that the amplitude of surface waves decays with depth as 1 over λ means that long wave length (or low frequency) waves are more sensitive to deeper structure than high frequency waves. In combination with the fact that, in general, the wave speed changes with depth, this explains why surface waves are **dispersive**: surface waves of different frequency propagate with different wave speeds.

Due to the dispersion, the wave form will change with increasing distance from the source so that it becomes less clear what is meant if one talks about the velocity of surface waves; to understand dispersion it will be necessary to consider two definitions of propagation velocity: **group and phase velocity**.

The surface waves are typically of substantially lower frequency than the body waves. Owing to the low frequency (sometimes in the same range as the eigenfrequencies of man made constructions) and their large amplitude, surface waves typically cause most of the earthquake damage to buildings.

Rayleigh waves

Interference between P and SV waves near the free surface[4] causes a type of displacement known as **Rayleigh waves**. Since the SV wave speed β is smaller than the P wave speed α there is an angle of incidence for an incoming SV wave that produces a critically refracted P wave, which propagates horizontally along the interface (see Fig. 4.18)

In other words, P-wave energy is trapped along the surface in a natural way, i.e., it does not require any particular wave speed variations at depth (Rayleigh waves can, in principle, exist in a half space). To conserve energy the amplitude of the horizontally propagating P wave must decrease with depth and vanish at some point, i.e., a critically refracted P wave is an *evanescent* wave.

[4]The boundary condition at the free surface is that the traction on that surface vanishes. It is convenient to take n_3 as the direction normal to the Earth's surface, so that $\sigma_{13} = \sigma_{23} = \sigma_{33} = 0$ and $T_3 = \sigma_{i3} n_i = 0$

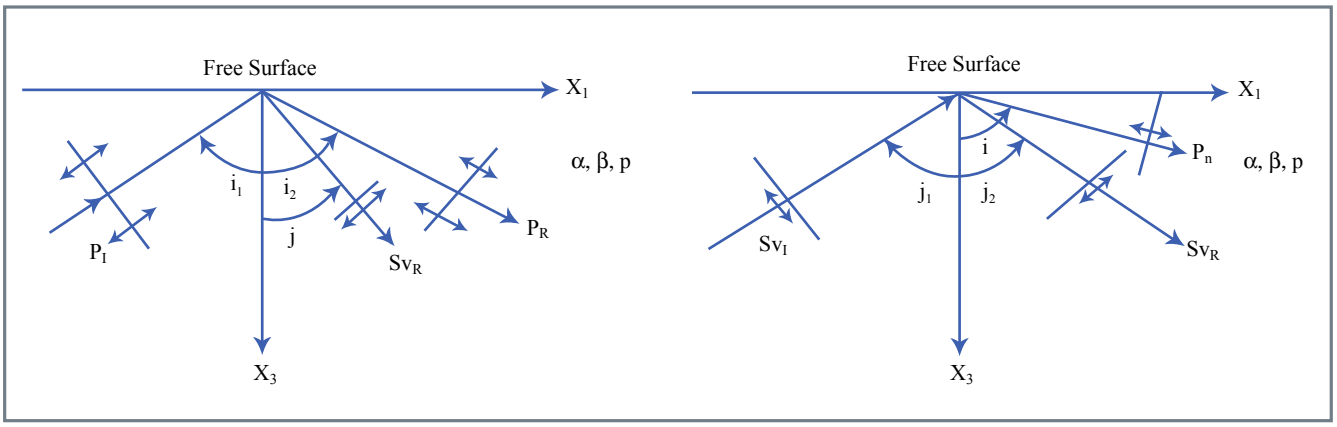


Figure by MIT OCW.

Figure 4.18: Free-surface interactions of an incident P and S wave.

Intermezzo 4.4 EVANESCENT WAVES

From analysis of a displacement potential ϕ it can be shown that the amplitude $A(z)$ of a horizontally propagating, critically refracted P-wave decays with increasing depth.
Consider the potential

$$\phi = A(z)e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} = A(z)e^{i\omega(px+\eta_\alpha z-t)} \tag{4.104}$$

with $\mathbf{k}$ the wave number vector and p and η_α the horizontal and vertical components of the P-wave slowness. From the vector properties of the slowness it follows that $p^2 + \eta_\alpha^2 = 1/\alpha^2$. The horizontal slowness p (the ray parameter!), is constant for the entire wave field generated by the incoming SV wave, which has a wave speed $\beta < \alpha$. In the case that $p = 1/c > 1/\alpha$ then

$$\eta_\alpha = \sqrt{\frac{1}{\alpha^2} - p^2} = i\sqrt{p^2 - \frac{1}{\alpha^2}} = i\hat{\eta_\alpha} \tag{4.105}$$

so that

$$\phi = B(z)e^{i\omega(px-t)}e^{-\hat{\eta_\alpha}\omega z} \tag{4.106}$$

A similar expression can be given for the SV-wave, with η_β instead of η_α. The fact that the argument of the exponential component of the amplitude factor is real has important implications for the admissible wave speeds. Since the wave number $\hat{\eta}\omega = k_z$ is related to $|\mathbf{k}| = 2\pi/\lambda$, with λ the wavelength, it also follows that the amplitude decay with depth is larger for small wave lengths than for long wave lengths, and this is of fundamental importance for the understanding of the *dispersion* of surface waves. (NB the horizontally propagating, evanescent P-wave must interfere everywhere with SV-waves; this can be achieved if there is an incoming SV-wavefield but for Rayleigh waves the evanescent P-wave interferes with a horizontally propagating, and thus also evanescent, SV-wave.)

Along the interface the critically refracted P-wave exists simultaneously with the incident SV-wave; in fact, the evanescent P-waves alone do not satisfy the stress-free boundary conditions and they *cannot* propagate along the interface without coupling to SV. The interference of P and SV-wave produces a particle

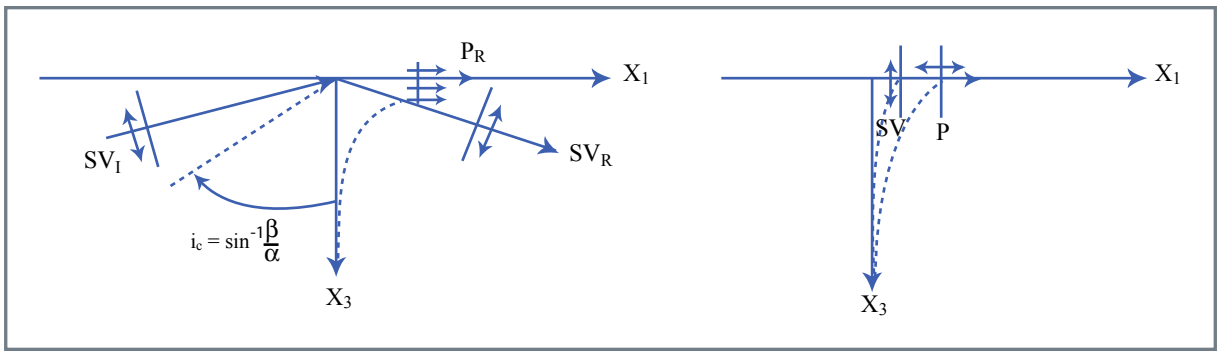


Figure by MIT OCW.

Figure 4.19: Evanescent waves; left evanescent P wave; right evanescent S wave. Amplitude decays exponentially with increasing distance from the interface.

motion in the $x - z$ plane that is *retrograde* at shallow depth, but changes to *prograde* at larger depth (see Fig. 4.20). This is similar to the particle motion in ocean waves.

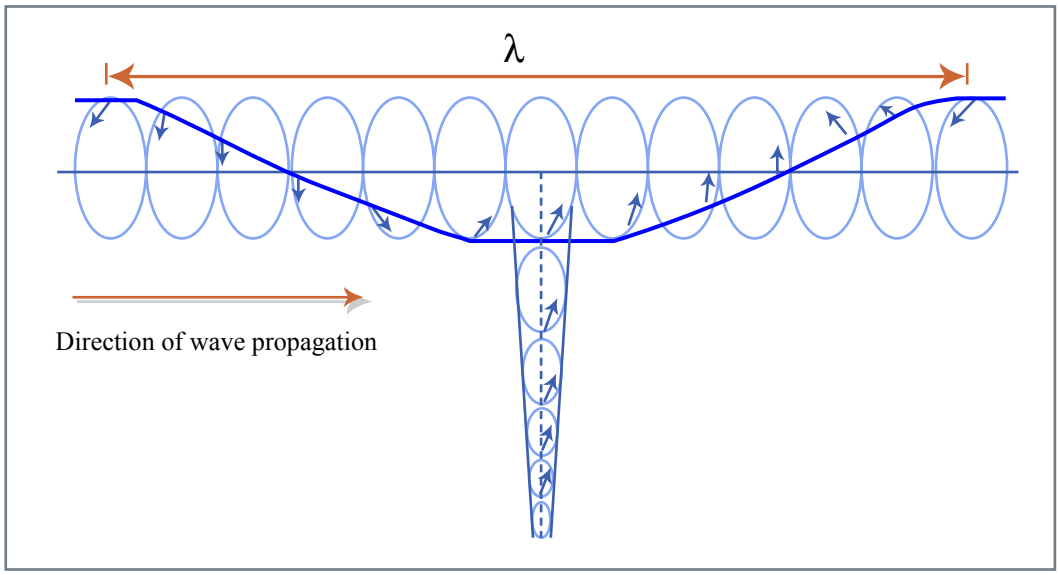


Figure by MIT OCW.

Figure 4.20: Elliptical particle motion for Rayleigh wave propagation.

The Rayleigh wave can thus be observed at both the vertical (in the direction of z) and horizontal (radial, i.e., in the direction of x) components of the displacement field (see also Fig. 4.21).

Love waves

Another type of surface wave, the **Love wave**, is formed by interaction of the *SH*-wavefield and the free surface. In contrast to the critically refracted waves that interfere to produce Rayleigh waves, there is no critical refraction of *SH*-waves (angle of incidence = angle of reflection) and in order to satisfy the boundary conditions there must be total reflection of the *SH*-waves at the free surface. *SH* energy can thus not be trapped near the surface in a half space. In order for Love waves to exist *SH* energy has to be reflected back to the surface

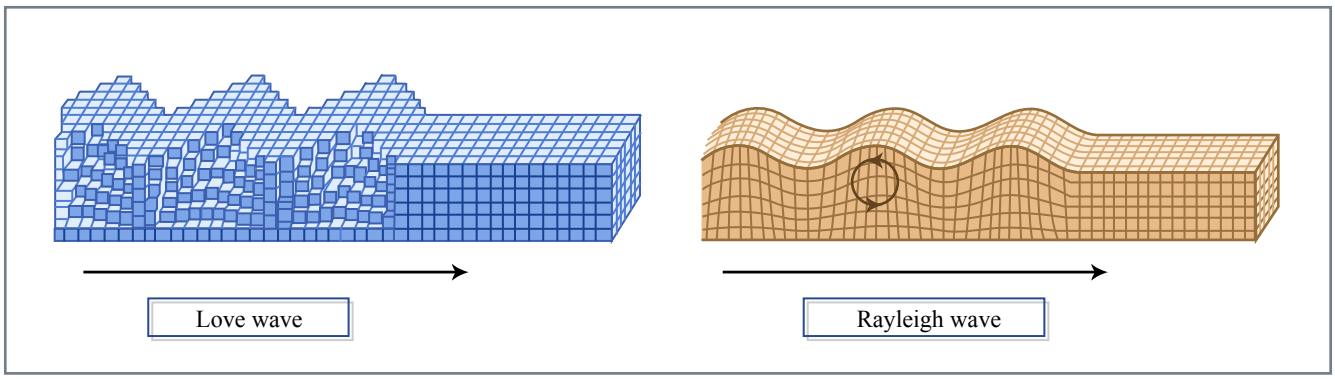


Figure by MIT OCW.

Figure 4.21: Love and Rayleigh wave displacement.

by a wave speed gradient at some depth; there must be a layer over a half space with the shear wave speed in the layer lower than in the half space. If the shear wave speed increases with depth a **wave guide** is formed in which rays are multiply reflected between the free surface and the turning points of the rays. In general, some energy may leak into the half space (if the form of *SH* body waves), unless the incoming *SH*-ray strikes the reflecting interface at (post) critical angles so that — effectively —- a head wave is formed and all energy is trapped within the wave guide (see Fig. 4.22). The headwave is also evanescent, and its amplitude decreases in with increasing depth beneath the layer (see box).

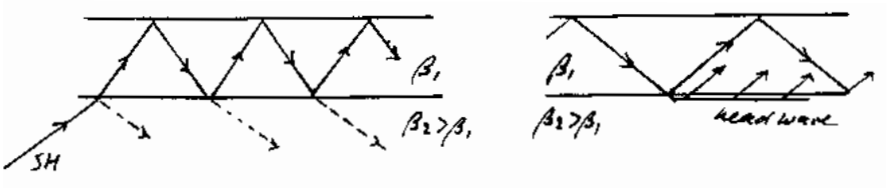


Figure 4.22: Trapped waves in the crust.

Since Love waves are interfering *SH*-waves, the particle motion is purely horizontal, in the x_2, or y, direction. Wave guides formed by a low-wave speed layer over a faster half space occur naturally in the Earth; the wave speed in the crust is larger than that in the mantle beneath the Moho, and at larger depths there can be a low velocity zone — in particular beneath oceanic lithosphere — that can cause efficient Love-wave propagation. Love waves are observed only on the transverse component (parallel to x_2) of the displacement field.

Propagation speed

From looking at data we can make an important observation: Love waves arrive before Rayleigh waves. Love waves propagate intrinsically faster than Rayleigh waves, see below, but the difference is not large enough to explain the observed advance of the Love wave arrival. Since Love waves involve only horizontal displacement whereas Rayleigh waves are composed of *P*-waves and vertically

polarized SV-waves, the observed advance of the Love waves suggests a form of **seismic anisotropy** with faster wave propagation in the horizontal plane than in the vertical direction (a situation known as *transverse isotropy*).

It can be shown, using the information given in the box below, that for horizontally propagating waves to be evanescent they must travel with a propagation velocity c that is always smaller than the compressional wave speed α, $c = 1/p < \alpha$, and also smaller than the shear wave speed $\beta, c = 1/p < \beta$. If $1/p \rightarrow \beta$ the amplitude of the surface waves no longer decays with depth and conservation of energy is then achieved by the leaking of energy into the half space in the form of body waves (SV in the case of Rayleigh waves and SH in the case of Love waves). If this happens one speaks of **leaky modes**.

So Rayleigh waves always propagate with a speed that is *lower* than the shear wave speed. For a half space with shear wave speed β_1, the propagation speed of the Rayleigh wave is about $0.9\beta_1$. (In the Earth the situation is more complicated because of the radial variation of both P and S-wave speed: if the wave speed gradually increases with depth from $c = \beta_1$ at the surface to $c = \beta_2$ in the half space: $0.9\beta_1 < c_{\text{Rayleigh}} < 0.9\beta_2$). We will see below that the surface-wave propagation speed depends on the wave length, and thus on frequency, of the wave (dispersion). For Love waves it is slightly different. Here it's the head wave that is evanescent; for high-frequency waves (short wavelengths) the evanescent head wave hardly penetrates into the half space (suppose a shear wave speed of β_2) so that the propagation speed is dominated by SH-propagation in the layer over the half space (propagation speed $c = \beta_1$). For longer period Love waves, the head wave is sensitive to as much larger depth range and the propagation speed gets closer to the shear wave speed in the half space (β_2). Thus: $\beta_1 < c_{\text{Love}} < \beta_2$.

4.19 Sensitivity kernels

For evanescent waves such as Rayleigh and Love waves we have seen that long wavelength waves penetrate deeper into the half space than short-wavelength waves. As a rule of thumb, at a depth of 0.4 λ the amplitude is reduced to $1/e$ of its value at the surface, and wave propagation is influenced by structure anywhere in this depth interval. How exactly structure in a certain depth interval influences a wave of a particular frequency is described by a **sensitivity kernel**. They represent the maximum partical motion at a certain depth as a function of frequency, which can be computed from a reference Earth model. A few examples are given below.

These kernels are a sort of **Green's functions** and they are typically convolved with (a model of) Earth structure in order to synthesize observables such as waveforms. (Note: we have seen someting like this before: in travel time tomography I mentioned that one solves the system of equations given by — in matrix notation — $\mathbf{Am} = \mathbf{d}$, with $\mathbf{m}$ the model vector and $\mathbf{d}$ the data vector. The matrix $\mathbf{A}$ contains the kernels and is therefore sometimes referreed to as the *sensitivity matrix*. In the case of travel-time tomography the kernels, the

elements of **A** are simply the path length of a ray in a certain block.)

4.20 Excitation of surface waves

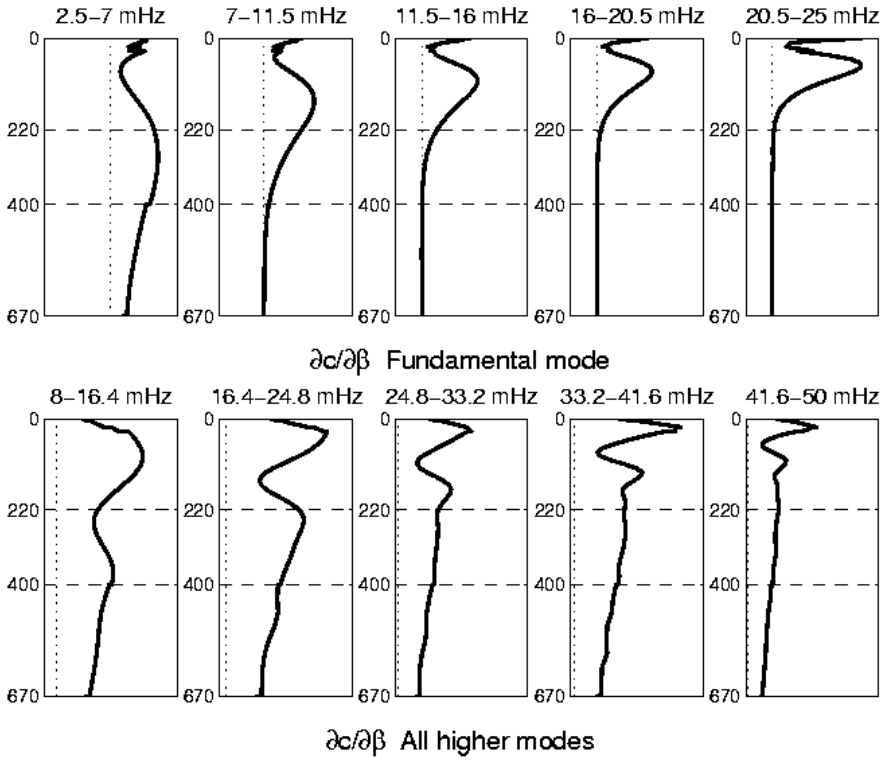


Figure 4.23: Phase speed sensitivity kernels.

Fig. 4.23 can be used to understand in qualitative sense the excitation of surface waves by earthquakes. In general, the position of the earthquake (i.e. the depth in our case of depth-dependent media) determines which modes can be excited. A **fundamental mode** has no displacement deeper than a certain depth; by reciprocity, a source (assume a white spectrum of the source so that it can — in principle — excite all frequencies) that is located at those large depth will not cause displacement of that fundamental mode at the surface.

4.21 Dispersion: phase and group velocity

The dependence of the depth of penetration on the period is described by the sensitivity kernels. If the wave speed is constant in the half space the waves associated with different kernels travel with the same wave speed and thus arrive

at the same time at a receiver at some distance from the source. But if, as is the case in Earth, the P and S-wave speed changes with depth, the longer period waves arrive at a different time than the shorter period waves. In Earth, the propagation speed of Rayleigh waves is thus frequency-dependent, and the waveform changes with increasing or decreasing distance from the source. This frequency dependence of propagation speed is called *dispersion*. Love waves are always dispersive since they cannot exist unless there is a layer over a half space, with the shear wave speed in the half space larger than in the overlying layer.

As a result of dispersion the surface waveform changes with varying distance from the source, and it is clear that one can no longer describe the wave propagation with a single wave speed. We describe the propagation velocity of the part of the waveform that remains constant, such as the onset of the phase arrrival, a peak, or a trough (see discussion of plane waves) with the **phase velocity** $c = \omega/k$. Wave packages with different frequencies travel at different velocities and their interference results in a phenomenon known as **beating** (see Interm): the propagation velocity of the envelope, which is related to the energy, of the resulting wave train is called the **group velocity U**.

Peaks or troughs in the wave form, or the onset of a particular phase arrival in the seismogram, all propagate with the **phase velocity**. In fact, we have seen this before when we discussed travel time curves of the body waves, which depend on the phase velocity. The phase velocity can thus be measured directly from travel time curves (recall that the horizontal slowness p can be determined from the slope of the travel time curve at a certain distance).

In Fig. 4.25 the dashed lines through A, B, etc. are travel time curves for those phases. But note that the frequency of those phases change with distance, so that the waveform changes. For instance, with increasing distance, the first arriving phase (A) is composed of waves with larger frequencies (because they sample deeper).

The **group velocity** is constant for a given frequency ($d\omega = 0$). Thus the group velocity of surface waves of a particular frequency defines a straight line through the origin and through the signal of that particular frequency on records of ground motion at different distances. The group velocity decreases as the frequency increases. As a result, high frequency phases become less and less pronounced with increasing distance from the source (or time in the seismogram).

The group velocity is very important: the energy in surface waves propagates mainly in the constructively interfering wave packets, which move with the group velocity.

Narrow-band filtering can isolate the wave packets with specific central frequencies (see Fig.4.26), and the group velocity for that frequency can then be determined by simply dividing the path length along the surface by the observed travel time. This technique can be used for the construction of dispersion curves (see Sec. 4.22).

Intermezzo 4.5 GROUP VELOCITY

Consider two harmonic waves with the same amplitude but slightly different frequencies (ω_1 and ω_2), wave numbers k_1 and k_2, and phase velocities $k_1 = \omega_1/c_1$ and $k_2 = \omega_2/c_2$ (see Fig. 4.24). These waves combine to give the total displacement

$$u(x,t) = \cos(k_1 x - \omega_1 t) + \cos(k_2 x - \omega_2 t). \tag{4.107}$$

If we define ω as the average between ω_1 and ω_2 so that $\omega_1 + \delta\omega = \omega = \omega_2 - - - \delta\omega$, and $k_1 + \delta k = k = k_2 - - - \delta k$, with $\delta\omega \ll \omega$ and $\delta k \ll k$, insert it into (4.107) and apply the cosine rule $2\cos x \cos y = \cos(x+y) + \cos(x-y)$, we obtain

$$u(x,t) = 2\cos(kx - \omega t)\cos(\delta k x - - - \delta\omega t) \tag{4.108}$$

This is the product of two cosines, the second of which varies much more slowly than the first. The second cosine 'modulates' the amplitude of the first. The propagation speed of this 'envelope' is given by $U(\omega) = \delta\omega/\delta k)$. In the limit as $\delta\omega \to 0$ and $\delta k \to 0$

$$U(\omega) = \frac{d\omega}{dk} = c + k\frac{dc}{dk} = c - - - \lambda\frac{dc}{d\lambda} \tag{4.109}$$

The group velocity is related to interference of waves with slightly different phase velocities; in other words U depends on c and on how c varies with frequency (or wavelength or wave number). In the earth $dc/d\lambda > 0$ so that the group velocity is typically smaller than the phase velocity.

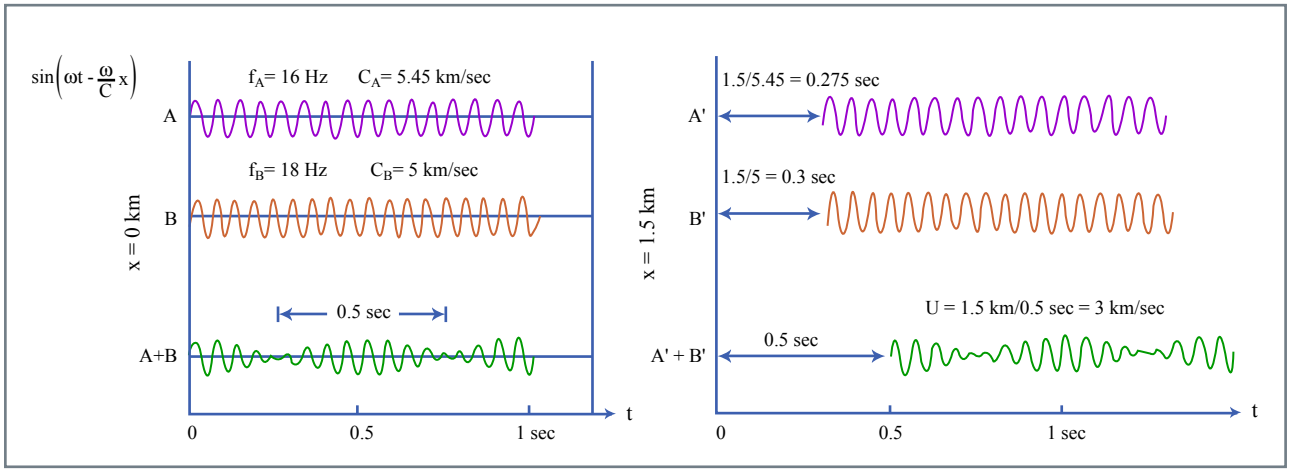


Figure by MIT OCW.

Figure 4.24: Two harmonic waves with the same amplitude but slightly different frequencies. The resulting "beating" is visible in the lowermost trace.

4.22 Dispersion curves

We have seen that the radial variation of shear wave speed causes dispersion of the surface waves. This means that the observed surface wave dispersion

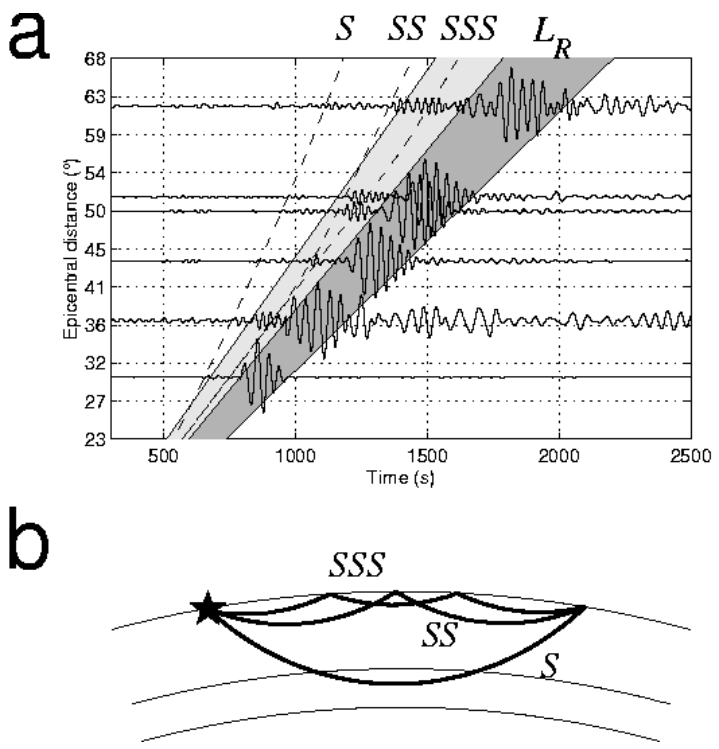


Figure 4.25: Group velocity windows and phase veclocity curves.

contains structural information about the radial variation of seismic properties. A plot of the group or phase velocity as a function of frequency is called a **dispersion curve**. Their diagnostic value of 1D structure has been explored in great detail. Typically, the curves produced from observed records are matched with standard curves computed from an assumed reference Earth model that can have a structure that is characteristic for a certain type of upper mantle (e.g., old/young continents, old/young oceans, etc.). Such analyses have produced the first maps of the thickness of oceanic lithosphere which revealed the increase in thickness with increasing age of the lithosphere (or distance from the ridge), and also underlie the discovery of the Low Velocity Zone (LVZ) at a depth of about 100 to 200 km beneath most oceans and beneath the younger parts of continents. Fig. 4.27 shows a variety of typical dispersion curves for different tectonic provinces.

4.23 Seismology: free oscillations

Like any bounded medium, the Earth can "ring like a bell" and after occurrence of a big earthquake it can oscillate in **normal modes** with discrete (eigen)frequencies. Normal modes of the Earth were predicted to exist in the

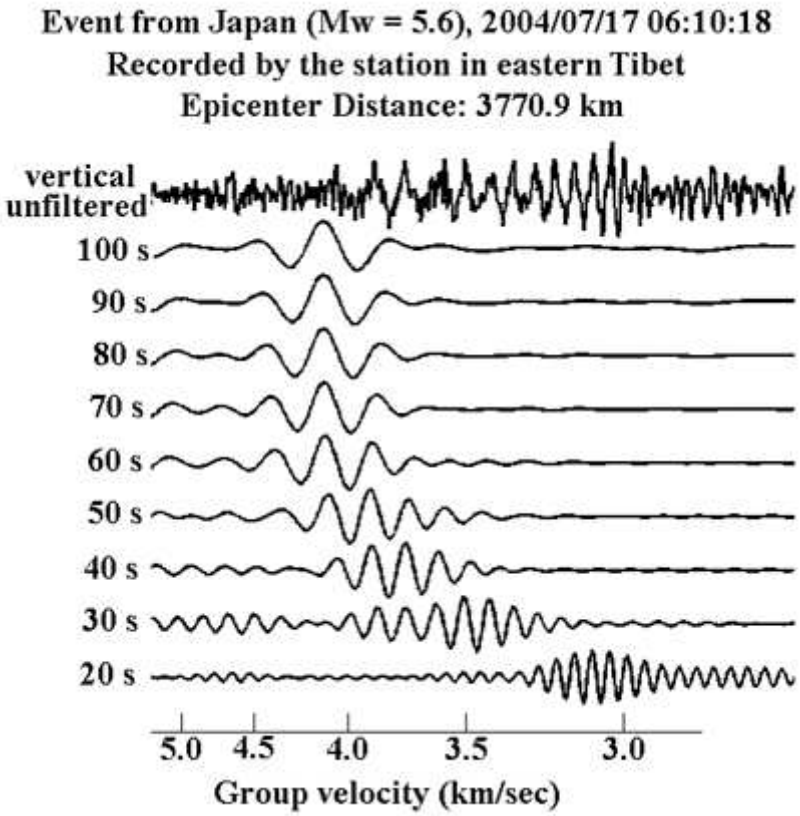


Figure 4.26: Frequency-band filtering of seismograms.

early part of the 19$^{\text{th}}$ century when mathematicians (Poisson, Rayleigh) studied elastic wave propagation extensively. However, in absence of sensitive long-period seismometers the normal models of free oscillation of the Earth remained undetected until the Benioff strain seismometer recorded the low-frequency signal due to a great earthquake in Kamchatka (1952). With the global network of highly sensitive broad-band seismometers many (many more than 1500) normal modes have now been observed and identified.

The "tone" of the ringing contains information about the structure of the Earth's interior. Since the entire Earth is involved in the free oscillations, the normal modes are more sensitive to average properties and whole-earth structure than to local anomalies. Of particular relevance is also that the low-frequency waves have to do work against gravity so that records of the modes contain information about the density distribution within the Earth. For these reasons the normal modes have played a central role in the development of global reference models for seismic properties.

A second important implication of normal modes is that the displacement of

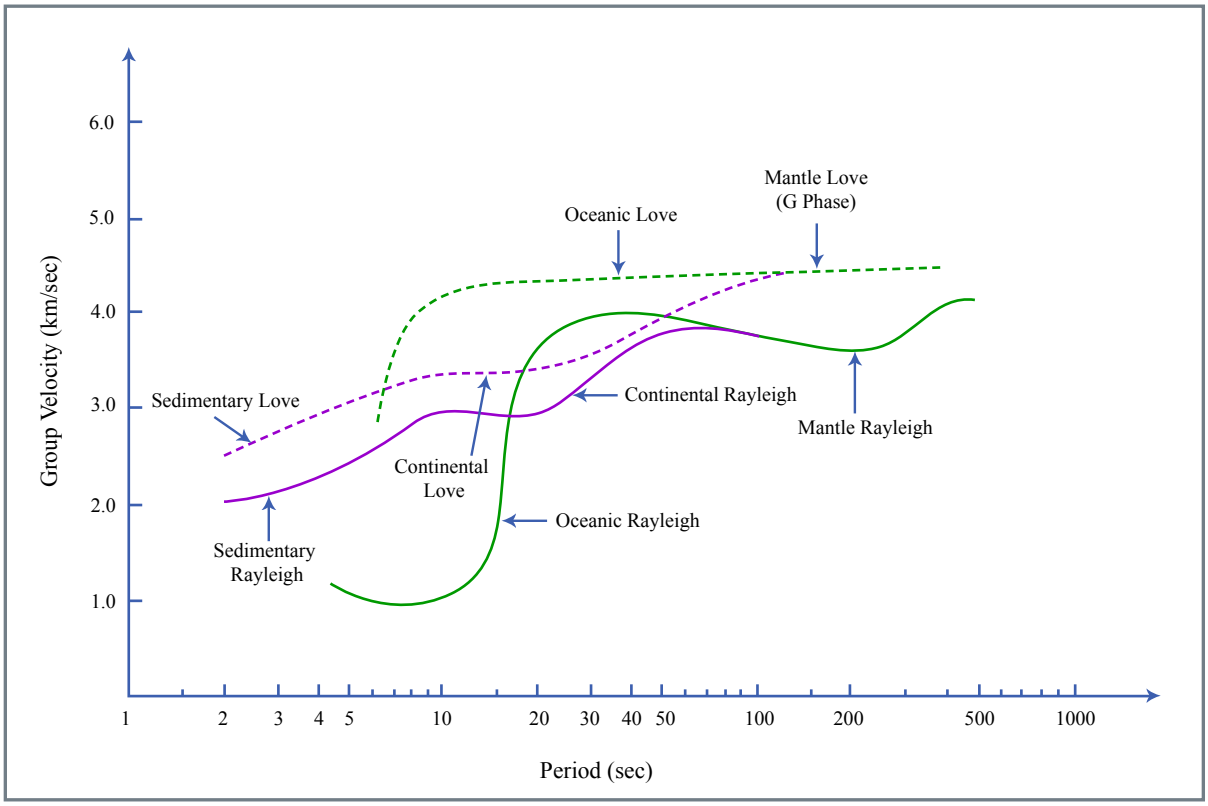


Figure by MIT OCW.

Figure 4.27: Dispersion curves for different tectonic provinces.

any number of normal modes can be summed as a Fourier series, with certain weights for the different frequencies, in order to construct synthetic seismograms (a technique known as **mode summation**). In fact, body and surface-wave propagation can be simulated by superposition of a sufficient number of fundamental and higher modes. In the discussion of surface waves we considered a "flat" Earth and an infinite half space (overlain, in case of Love waves, by a low wave speed wave guide). This is only useful to derive some fundamental properties, in particular at relatively short periods ($T < 200$s), but for long period surface waves , which penetrate deep into the Earth's interior and for the interference of waves that have propagated along the circumference of the Earth, one must take sphericity into account. The surface waves were characterized by their frequency ω and wave number k. We did not consider boundaries of the medium other than the free surface, and the frequency was taken as the independent variable: for each frequency there are only certain discrete wave numbers $k = k_n(\omega)$ for which the boundary conditions could be satisfied. Instead we could have formulated the problem in terms of discrete eigenfrequencies $\omega = \omega_n(k)$ with k the independent variable. This formalism makes more sense for the discussion of free oscillations of the Earth, since the medium is bounded. In the spherical geometry the "horizontal wave number" k is fixed at certain discrete values by the finite lateral extend of the medium. One often uses the angular wave number l instead of k, with l zero or a positive integer (see Fig. 4.28).

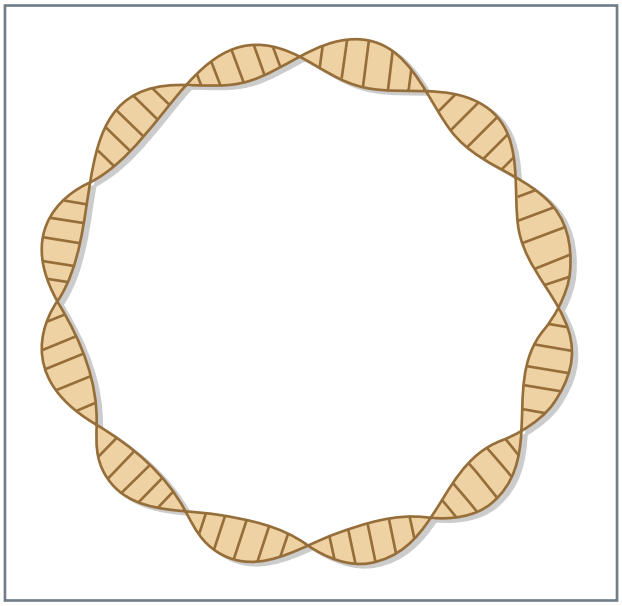


Figure by MIT OCW.

Figure 4.28: Standing waves in a spherical Earth.

Normal modes and overtones

To get some insight in the problem, let's consider the simple situation of vibrations of a string held fixed at either end. The motions in the string must obey the 1D wave equation, with c the phase velocity:

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} \tag{4.110}$$

The general solution of this equation is

$$u(x,t) = Ae^{i\omega(t+\frac{x}{c})} + Be^{i\omega(t-\frac{x}{c})} + Ce^{-i\omega(t+\frac{x}{c})} + De^{-i\omega(t-\frac{x}{c})} \tag{4.111}$$

The constants $A - D$ can be determined from the boundary conditions, i.e. the fixed end points: $u(0,t) = u(L,t) = 0$. The first gives $A = -B$ and $C = -D$. The condition at $x = L$ then gives

$$(Ae^{i\omega t} - Ce^{-i\omega t})2i\sin\left(\frac{\omega L}{c}\right) = 0 \tag{4.112}$$

which has nontrivial solutions for $\omega L/c = (n+1)\pi$, $n = 0, 1, 2, 3, \cdots \infty$. These discrete frequencies, labeled ω_n, are called the eigenfrequencies of this bounded system. The corresponding displacements, Eq. (4.110), are the eigenfunctions or normal modes of the system and are of the form $u = \exp(i\omega_n t)\sin(\omega_n x/c)$. The **fundamental mode** is given for $n = 0$, and has no internal nodes (where $u = 0$) within the system; $n > 0$ corresponds to **higher modes** or **overtones**, which have n internal nodes. It is important to realize that the motion of each of the modes occurs without horizontal motion of the nodes: they are standing waves and the modes themselves don't propagate horizontally. However, constructive interference of the coexisting vibrations corresponds to traveling waves. We have previously said that P and S-waves are the complete solutions to the wave equation, and it can be shown that the normals modes of free oscillations are, in fact, not fundamentally different from the body waves. Normal modes can be used to describe body wave propagation. Indeed, any propagating disturbance can be represented by an infinite weighted sum of the eigen frequencies (Fourier series!) so that normal mode summation can be used to simulate propagating waves such as body waves and surface waves:

$$u(x,t) = \sum_{n=0}^{\infty}\left(A_n e^{i\omega_n t} + B_n e^{-\omega_n t}\right) \tag{4.113}$$

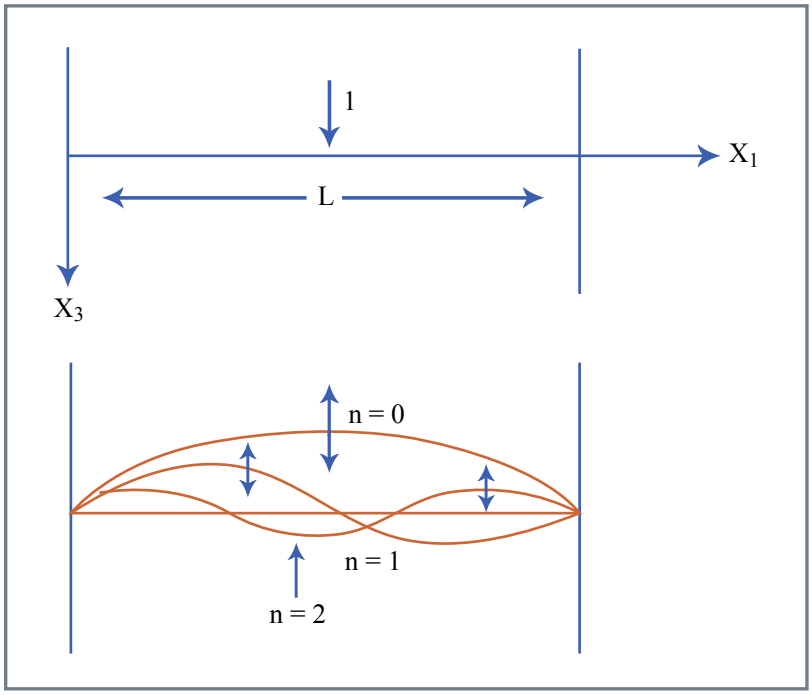


Figure by MIT OCW.

Figure 4.29: A string under tension. Fundamental mode is given by n = 0; n = 1, 2, ... are the over tones.

Power spectrum

The individual modes can, in general, not be observed directly from the seismograms. Free oscillations are studied with spectral techniques. If one was to take a Fourier transform of a sufficiently long record of ground motion, typically many hours or even days, one gets a power spectrum that reveals the distinct eigenfrequencies of the Earth's free oscillations (see Fig. 4.30).

Nomenclature of normal modes

Normal modes of free oscillation are just the solutions of the wave equation in a spherical coordinate system and the nomenclature of the modes is therefore based on spherical harmonics. Recall that the gravity and magnetic potentials were, in fact, summations of modes with different coefficients (Gaussian coefficients in the case of the magnetic potential). The expression of mode summation is similar to the spherical harmonic expressions used when we discussed, for instance, the geoid and the magnetic field with two differences: (1) the normalization of the harmonic coefficients are typically specific to each application (seismology, gravity, geomagnetism), but don't worry about that now, and (2) instead of doing the summation from $m = 0$ to l with two (Gaussian) coeffi-

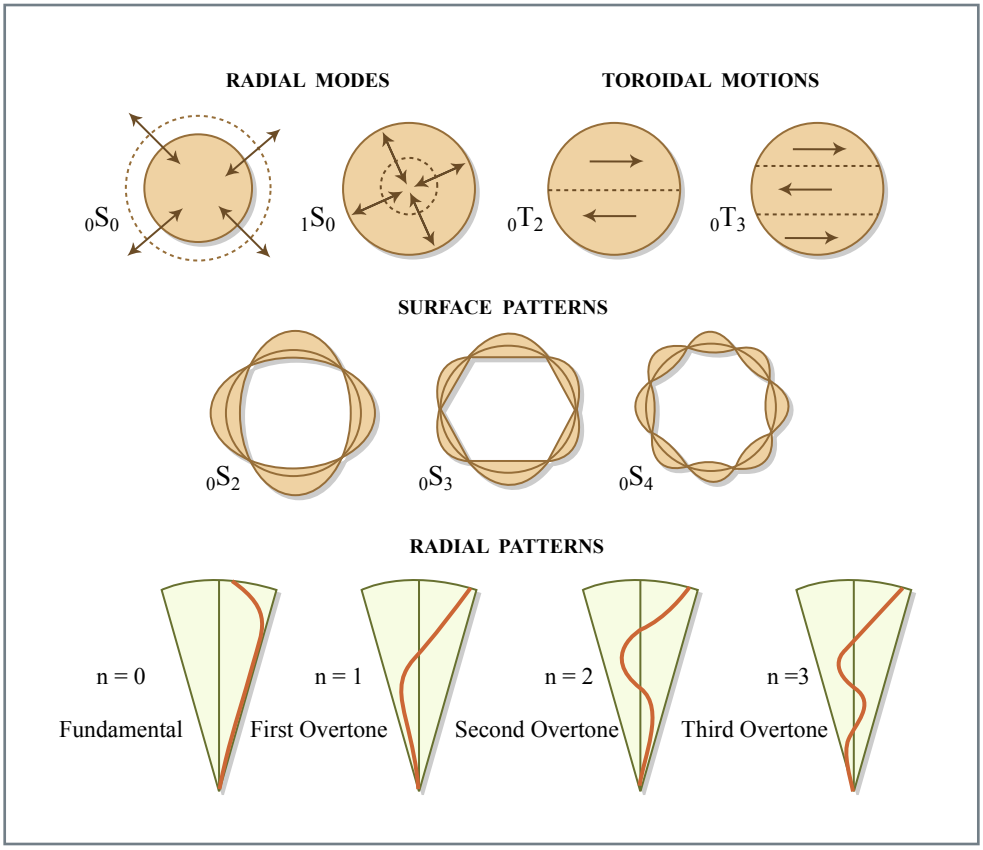


Figure by MIT OCW.

Figure 4.30: Surface and nodal patterns of free oscillations.

cients, in seismology one typically uses a notation that sums from $m = -l$ to $+l$: in both cases there are $2l + 1$ coefficients (this is called a $2l + 1$ **degeneracy**).

There are two basic types of free oscillation (1) **spheroidal modes**, which are analogous to the *P*-*SV*-system and the Rayleigh waves and have a component of motion parallel to the radius from the Earth's center; and (2) **toroidal** or torsional modes involving shear motions parallel to the Earth's surface, analogous to *SH* and Love waves. Spheroidal modes involve expansion and contraction of (parts of) the Earth, whereas toroidal modes involve differential rotation of parts of the globe. Gravity does not influence the toroidal motion but long-period spheroidal oscillations do involve significant work against gravity; observation of these modes can therefore yield information about the Earth's gross density structure.

The toroidal and spheroidal modes are labeled ${}_nT_l$ and ${}_nS_l$, respectively, where n indicates the number of nodes along the radius of the Earth[5] Torsional modes are only sensitive to shear wave speed; spheroidal modes are sensitive

[5]The latter would be true if the Earth was homogeneous and uniform; in reality it is more complicated. The behavior of normal modes in the Earth is complicated by stratification, the existence of a fluid outer core, by the rotation of the sphere, and, of course, by deviations from sphericity (3D structure + anisotropy).

to compressional and shear wave speed and density – n is the overtone number — and l (the angular order or degree or wave number) indicates the number of nodal planes on the surface (see Fig. 4.31).

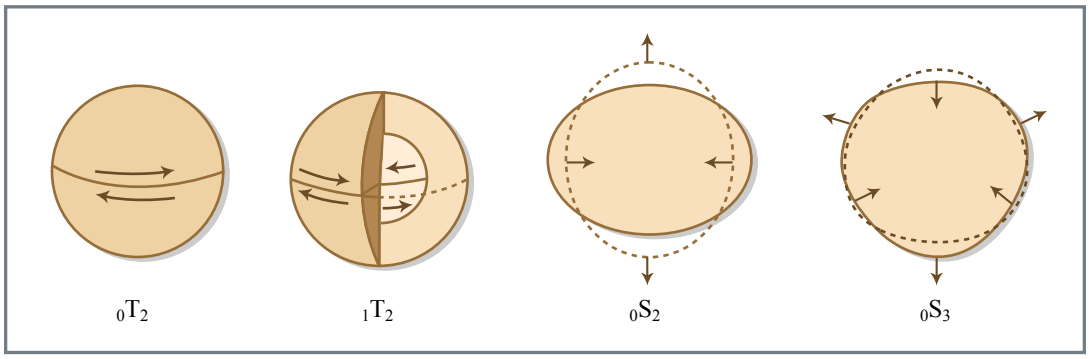


Figure by MIT OCW.

Figure 4.31: Different toroidal modes ($_0T_2$, $_1T_2$; top) and spheroidal modes ($_0S_2$, $_0S_3$; bottom).

For example, the mode $_0T_2$ corresponds to alternating twisting of the entire upper and lower hemisphere of the spherical body; the mode $_1T_2$ corresponds to similar twisting of the center of the sphere, but now with twisting in the reverse direction of the outer part of the sphere (see Fig. 4.32). The modes with $n = 0$ sense the gross mantle structure, and the modes with increasing n are, in general, sensitive to elastic properties at different depths in the sphere. For toroidal modes, the poles have no motion, counting as the $l = 1$ term. The mode $_0T_1$ cannot exist. Spheroidal modes with $l = 0$ have no nodal planes at the surface and are therefore sometimes called **radial modes**. The mode $_0S_0$ involves expansion and contraction of the sphere as a whole; mode $_0S_2$ has two equatorial bands of zero displacement, $_0S_3$ has three nodal lines etc. (see Fig. 4.32).

Mode	Period (s)	Mode	Period (s)
$_0S_0$	1277.52	$_0T_2$	2636.38
$_0S_2$	3223.25	$_0T_{10}$	618.97
$_0S_{15}$	426.15	$_0T_{20}$	360.03
$_0S_{30}$	262.09	$_0T_{30}$	257.76
$_0S_{45}$	193.91	$_0T_{40}$	200.95
$_0S_{60}$	153.24	$_0T_{50}$	164.70
$_0S_{150}$	66.90	$_0T_{60}$	139.46
$_1S_2$	1470.85	$_1T_2$	756.57
$_1S_{10}$	465.46	$_1T_{10}$	381.65
$_2S_{10}$	415.92	$_2T_{40}$	123.56

Table 4.4: Oscillation periods of some normal modes.

Table 4.4 gives the periods of some of the observed modes. The normal mode with the longest period is the spheroidal mode $_0S_2$, with a period of about 54

minutes. In the last 4 decades many modes have been identified. This also is a game of matching the observed spectra with model predictions, identifying the modes, using that to improve the reference Earth models, and the improved starting models may then allow the identification of previously unknown modes.

Normal mode splitting: aspherical Earth's structure

We have used the notation of modes in terms of S and T and the degree l and the overtone number n, for instance ${}_0S_2$. Just as in the use of spherical harmonics to describe the gravity and magnetic fields we also have the order m in seismology. (As a reminder: there are l nodal lines at the surface: there are m nodal lines along great circles (m=0 gives the **zonal** harmonics) and there are thus $l - m$ nodal lines along latitude. For $l \neq m$: tesseral harmonics). For each angular degree l there are $2l + 1$ values for m. In a spherically symmetric, non-rotating body the $2l + 1$ modes have the same eigenfrequency, the modes correspond to a single peak in the spectrum — the overlapping peaks are known as **multiplets** — and this redundancy is the reason why the superscript m is usually ignored in the notation. However, the 5 different modes that constitute ${}_0S_2$ have different angular moments and when the body is rotating the $2l + 1$ peaks, or **singlets** do not exactly overlap any more. This phenomenon is known as the **splitting** of the modes. The split modes have eigenfrequencies that are very close together so that interference occurs.

Splitting can be caused by rotation, but also by aspherical Earth's structure such as lateral variation in isotropic seismic properties (due to dynamic processes in the mantle) or by seismic anisotropy. Conversely, the analysis of splitting in the power spectra can give invaluable information about 3D structure and anisotropy.