Chapter 2

Theoretical Framework

Theoretical foundations of the analytical methods used in this dissertation rest on the theory of elasticity and wave physics in material media with appropriate limiting and boundary conditions. Infinitesimal strain theory and generalized Hooke's law relating strains to internal stresses in isotropic homogeneous materials are valid approximations for the earth because of small displacements generally, and the random orientation of rock forming crystals which have the effect of statistically homogenizing heterogeneous materials over the dimensions (tens of meters) of a wavelength. When plugged in the equation for conservation of momentum, these stress-strain relations lead to the basic dynamical equation of elasticity. The latter when applied to unbounded media has two transient solutions: a scalar and a vector wave, implying that energy unleashed by any mechanical instability in such a medium travels radially outward from the source as longitudinal and shear waves. Solutions of the dynamical equation for inhomogeneous media constrained by the appropriate boundary conditions of continuity of stresses and strains across interfaces of homogeneous volumes of materials, lead to reflected and refracted waves whose respective amplitudes and therefore energies, can be thereby determined. In particular, when applied to the free boundary of a homogeneous medium such as the earth's surface where both the reflected and refracted waves must simultaneously exist and surface tractions must vanish, the SH wave amplitude is doubled whilst the P and SV waves interact to generate the retrograde elliptically polarized Rayleigh waves. When the half space below the free surface is inhomogeneous, the free
surface reflected SH wave combines with those reflected from internal boundaries to produce a trapped SH reverberation that appear at the surface as Love waves which are dispersed. Even the Rayleigh waves suffer dispersion in a vertically inhomogeneous or layered media owing to the appearance of finite scales that cause wavelength dependent interactions. The phase and group velocities of these surface waves (both Love and Rayleigh) are determined by the scale of the inhomogeneities and their elastic parameters can, in turn, be determined from these velocities abstracted from a set of seismograms generated at sites along the wave passage.

When dealing with inhomogeneous media however, $\lambda$ and $\mu$ being functions of position, the dynamical equation (with zero body force) cannot be reduced to a wave equation and a different approach has to be adopted using an approximation of the wave equation. One thus arrives at the 'Ray Equation' (the Eikonal) following wave trajectories which greatly facilitates the tracking of waves through inhomogeneous media although its range of validity is limited to situations where velocity gradients over the dimensions of a seismic wavelength are small. A concise but complete treatment of these basic theoretical formulations is therefore included in the following sections, together with those relating to converted phases at acoustic impedance boundaries, in particular, the shear waves generated at these boundaries in the earth's crust and upper mantle by near vertical incident P waves. The former when abstracted from the more convoluted seismograms, can be inverted to determine the seismic characteristics of the region traversed by these converted waves. Since for short amplitude seismic waves the earth behaves like a linear system, these converted shear waves and their multiples $Rs(t)$ appearing at the earth's surface can be regarded as the *convolved* product of the generator P-wave $P(t)$ incident at the base of an earth column and the overlying structure, and therefore as the earth column's response to the P-wave excitation. Theory shows that the generator teleseismic ($30^\circ < \Delta < 90^\circ$) P waves incident at a steep angle are transmitted without any significant
deviation from their original path or diminution in their amplitudes, which allows one to treat them as representing the vertical ground motion recorded at the surface. The converted shear wave time series \( Rs(t) \), on the other hand, are quite feeble compared with the much larger amplitude P-wave component appearing on broadband horizontal component seismograms. Fortunately, however, they can be neatly isolated by deconvolving the vertical component seismogram from its radial and transverse components, in turn, vectorially reconstituted from the north and the east component seismograms. The resulting time series \( Rs(t) \) contain the horizontal component of the generator P-wave followed by the converted shear wave phases and their multiples, and largely represent the properties of materials immediately beneath the receiver, aptly called a Receiver Function. A grid search method enables testing of the coherence of these multiply generated shear wave phases to constrain the depth of the seismic discontinuity and the average \( Vp/Vs \) value, whilst its inversion in terms of the shear wave speeds in a stack of equal thickness layers, yields the shear wave structure of the underlying earth. Accordingly, the topics covered in this chapter, include: the elastodynamic equation and the Eikonal (wave to rays), partitioning of energy at seismic discontinuities, converted phases, receiver functions, surface waves, inverse theory and seismic anisotropy.

2.1 Deformation of Elastic Media

Physical bodies when subjected to forces may suffer displacement or deformation. The latter involve changes in volume, shape or both due to rearrangement of contiguous matter. Response of macroscopic bodies to forces can be adequately analyzed using Newton's laws of classical mechanics for a particle by treating a material body as a continuous distribution of particles - a continuum - in which the point position of the particle is abstracted from an infinitesimal surrounding volume \( (dv) \) as this volume
approaches zero and its mass as concentration of the mass \((\rho \, dv)\) contained in this volume element as it shrinks to a point. Mathematically, this is realized with the help of the peaked Dirac Delta function \(\delta(R - R_0)\) defined as follows:

\[
\delta(R - R_0) = 1 \text{ for } R = R_0, \text{ and is zero for all other values of } R
\]

\[
\int \delta(R - R_0) \, dR = 1; \quad \int f(R) \delta(R - R_0) \, dR = f(R_0)
\]

Thus, applying Newton’s laws to a continuous distribution of infinitesimal volumes \(dv\), we can analyze the behaviour of material bodies when subjected to external forces.

In an undeformed body, the arrangements of constituent molecules are determined by the state of its thermal equilibrium. Any portion of such a body at constant temperature, is therefore in mechanical equilibrium and the sum of all forces therein is zero. When deformed, the equilibrium molecular arrangement within the body is altered requiring the creation of internal forces to return the body to equilibrium. These internal forces are sustained by interaction between molecules, as distinct from ‘molecular forces’ whose range of action extends only to molecular distances and is therefore of no consequence to macroscopic behaviour. For analyzing elastic behaviour of bodies, therefore, we may regard ‘internal forces’ as ‘neighbourhood acting forces’ impressed by any point (infinitesimal of volume) on to its neighbourhood points only or vice versa. Therefore, forces acting on any part of the body due to a surrounding part act only at their interface, and, in general, would vary with the orientation of the neighbourhood.

For proceeding further to analyze the response of material bodies to forces using Newton’s laws of mechanics, we need a quantitatively complete description of both forces acting over an infinitesimal volume representing a
point and the resulting Deformation in terms of relative displacements of
neighbourhood points.

2.1.1 Quantitative Measure of Deformation

We first seek a quantitative description of Deformation of an infinitesimal
volume of material centred at a point \( P(X) = i \times x_i \) which, in general, may
have suffered a displacement \( U(X) = i \times du_i \) accompanied by an altered
structure of its neighbourhood. Accordingly, we enquire as to what happens
to the neighbourhoods of the point \( X \) in this process, and more specifically, to
the variously directed infinitesimal \( dX = i \times dx_i \) vectors originally centred at
\( X \). We answer this by drawing upon our abstraction of the continuum, to
express the displacement \( U(X + dX) \) at a neighbouring point \( (X + dX) \) as a
Taylor’s expansion about \( U(X) \) retaining only the first term in view of the
infinitesimal value of \( |dX| \). Thus, writing the \( k \)th component of \( U(X + dX) \),
that is the displacement of the tip of \( dX \) as

\[
u_k(X + dX) = u_k(X) + \sum_i \left( \frac{\partial u_k}{\partial x_i} \right) dx_i,
\]

the deformation \( dX \) suffered by \( dX \) can be written in terms of the strain
components: \( S_k \)

\[dx_i' = S_k = u_k(X + dX) - u_k(X) = \sum_j \left( \frac{\partial u_k}{\partial x_j} \right) dx_j \]

Expressed in a matrix form, the strain \( S \), suffered by the vector \( dX \) appears as
the result of an operation on \( dX \) by the matrix \( E \)

\[ S = E dX \]

with the matrix \( E \) itself decomposed as a sum of its 3 irreducible (Arfken,
2005) components \( E = h + c + e \), \( h \) behaving as a scalar, \( c \) as a vector \( (= \text{Curl } U) \), and \( e \) as a symmetrical tensor. Thus, it is easily seen that
\[
\varepsilon = \frac{1}{3} \left[ \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} \right]
\]

\[
h = \begin{bmatrix}
\varepsilon & 0 & 0 \\
0 & \varepsilon & 0 \\
0 & 0 & \varepsilon 
\end{bmatrix}, \text{ where } \varepsilon = \frac{1}{3} \left[ \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} \right]
\]

\[
c = \begin{bmatrix}
0 & \frac{1}{2} \left( \frac{\partial u_1}{\partial x_2} - \frac{\partial u_2}{\partial x_1} \right) & \frac{1}{2} \left( \frac{\partial u_1}{\partial x_3} - \frac{\partial u_3}{\partial x_1} \right) \\
\frac{1}{2} \left( \frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2} \right) & 0 & \frac{1}{2} \left( \frac{\partial u_2}{\partial x_3} - \frac{\partial u_3}{\partial x_2} \right) \\
\frac{1}{2} \left( \frac{\partial u_3}{\partial x_1} - \frac{\partial u_1}{\partial x_3} \right) & \frac{1}{2} \left( \frac{\partial u_3}{\partial x_2} - \frac{\partial u_2}{\partial x_3} \right) & 0
\end{bmatrix}
\]

\[
\varepsilon = \begin{bmatrix}
\frac{\partial u_1}{\partial x_1} & \frac{1}{2} \left( \frac{\partial u_1}{\partial x_2} + \frac{\partial u_2}{\partial x_1} \right) & \frac{1}{2} \left( \frac{\partial u_1}{\partial x_3} + \frac{\partial u_3}{\partial x_1} \right) \\
\frac{1}{2} \left( \frac{\partial u_2}{\partial x_1} + \frac{\partial u_1}{\partial x_2} \right) & \frac{\partial u_2}{\partial x_2} & \frac{1}{2} \left( \frac{\partial u_2}{\partial x_3} + \frac{\partial u_3}{\partial x_2} \right) \\
\frac{1}{2} \left( \frac{\partial u_3}{\partial x_1} + \frac{\partial u_1}{\partial x_3} \right) & \frac{1}{2} \left( \frac{\partial u_3}{\partial x_2} + \frac{\partial u_2}{\partial x_3} \right) & \frac{\partial u_3}{\partial x_3}
\end{bmatrix}
\]

The strain $S$ suffered by $dX$, thus turns out to be the sum of the operations on $dX$ by the matrices $h$, $c$ and $e$. The first of these leading to a uniform extension/contraction of all three components of $dX$ does not therefore alter the shape of the neighbourhood around $P(X)$, but only hydrostatic changes in volume, the cubical dilatation being given by the trace of $h$,

\[h_{\text{trace}} = \frac{\delta V}{V} = \frac{1}{3} \left[ \left( \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} \right) \right] \tag{2.4}
\]

The second operating on $dX$ transforms it to $\frac{1}{2} [\text{Curl} \times dX]$ which signifies rigid body rotation of the vector $dX$ by an amount equal to $\frac{1}{2} [\text{Curl}]$ about the vector $\text{Curl}$. The symmetric matrix $e$ with zero trace, alone is responsible for the non-hydrostatic shear deformation of the infinitesimal neighbourhood of $P(X)$ and is referred to as the deviatoric components of strain $e_{kl}$ which
happens to be symmetrical: $e_{kl} = e_{lk}$. The deviatoric strain suffered by $dX$ can thus be written as:

$$S_k = \sum_l e_{kl} dx_l$$  \hspace{1cm} (2.5)

where,

$$e_{kl} = \left[ \frac{1}{2} \left( \frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k} \right) - \delta_{kl} \right]$$  \hspace{1cm} (2.6)

and $\delta$ represents the Kronecker delta. It may be noted that the deviatoric strain tensor matrix $e$ like all symmetric matrices can be diagonalized by transformation to an appropriate set of orthogonal axes. These new axes then represent directions along which the strain at $P(X)$ is entirely longitudinal and are therefore called the Principal strains.

### 2.2. Quantitative Description of Force

It will be recalled that the equilibrium molecular arrangement of a body is altered when it suffers deformation, raising in consequence internal forces to create a new equilibrium state, and that these forces are 'near action' forces impressed by a point (infinitesimal volume element) only on the neighbouring points at their interfaces. To sharpen our focus on a mathematical description of the total force on a body, we isolate some part of it defined by an imaginary surface that separates it from the surrounding matter. The total force on this isolated volume of material is the sum of all the forces on all the infinitesimal volume elements contained within this isolated volume expressible as the volume integral $\int F dv$ where $F$ is the force density. In equilibrium, the forces which these various infinitesimal volume elements impress on their neighbouring elements as contact forces of action and reaction have a zero resultant, leaving only those acting at the imaginary surface of the isolated volume, exerted upon it by the parts that surround it.
The total force on this isolated portion of the body is thus the sum of forces exerted on all of its surface elements by its neighbours.

We further recall the well known vector relation that the volume integral of a scalar quantity which can be expressed as the divergence of a vector, can be transformed into an integral of the surface density of that vector over the surface that encloses this volume:

\[
\int (\nabla A) dv = \int A dS
\]

Making use of the above equivalence, we can express the total force \( F dv \) on any volume element in terms of the surface density of forces (Force per unit area) or 'tractions' impressed on the various elements of the surface that bounds it. However, since the integrand of the volume integral in this case is a vector with 3 scalar components \( f_k (k = 1, 2, 3) \) each one of which can be separately transformed into 3 surface integrals, we would require that each scalar component \( f_k \) is the divergence of a corresponding vector \( T_k \). Therefore, for a complete definition of the total force in the elementary volume, we require 3 vectors with components \( T_{\mu k} \), where the first subscript refers to the three tractions and the second to their respective components. The nine \( T_{\mu k} \) quantities thus distilled for the complete description of the force at a point in the body, are called the Stress Tensor or just stress at the point to which the infinitesimal volume shrinks in the limit. Tensors, originating from the word Tension whose analysis first led to their formulation, form a much more general class of entities. The total force density \( F(= i_k f_k) \) at a point within a body can thus be expressed as the divergence of the stress tensor \( T \):

\[
F(X) = - \nabla T(X)
\]

Or, in terms of their components:

\[
f_k = - \nabla T_{\mu k}
\]
where the negative sign signifies the convention that positive traction on \( \hat{n} \) is the force exerted by the material on the positive side of the surface vector \( \hat{n} \).

The components of \( T_{\mu} \), can be written in a matrix form as follows:

\[
T_{\mu} = \begin{bmatrix}
T_{11} & T_{12} & T_{13} \\
T_{21} & T_{22} & T_{23} \\
T_{31} & T_{32} & T_{33}
\end{bmatrix}
\]

(2.10)

An intuitively appealing approach to visualizing the description of stress at a point within a given body is to consider dividing the body by an imaginary surface passing through the point under consideration. Each part of this divided body acts on the other at that point exerting equal and opposite forces. But, we describe the force vector at a point on one of the two divided portions under consideration in terms of its surface density or traction:

\[
\lim_{\delta S \to 0} [F/S]
\]

where \( F \) is the force vector acting over an infinitesimal contact area \( S \) around the point of the imaginary surface. However, the traction at the point, for any given configuration of applied forces, would depend on the orientation of the imaginary dividing surface through the point and would therefore furnish an incomplete description of the force acting thereon. This arbitrariness can, however, be removed by incorporating three traction vectors \( T_{\mu} \) on three imaginary orthogonal planes through the point, thereby involving nine scalar quantities. These nine components of the stress tensor, \( T_{\mu} \) at a point completely describe the force acting at that point because they can be used to calculate the traction across any arbitrarily oriented surface \( \hat{n} \) through the point. We shall follow the convention that the first subscript of \( T_{\mu} \) denotes the planes: \( x_1, x_2, x_3 \), and the second denotes the components along the respective axes.
When a body is made of constituent volume elements that are susceptible to additional forces because of the presence of some specific fields e.g., masses in a gravitational field [e.g. \( \rho g \text{dv} \)] or charged and magnetically polarized particles in an electromagnetic field, there would appear these 'body' forces \( f \text{dv} \) that must be added to the stresses described above.

For equilibrium, \( (F/\text{dv}=\bar{F}) \) must balance all other forces. In general, therefore

\[
F + f = \rho \left( \frac{\partial^2 U}{\partial t^2} \right)
\]  
(2.11)

Further, the conditions of i) rotational equilibrium of the infinitesimal volume about the various axes can be shown to require that the stress tensor be symmetrical, that is,

\[
T_{\mu} = T_{\nu}
\]  
(2.12)

and, ii) the equilibrium of an elementary tetrahedron bounded by the three axial planes and an arbitrarily oriented surface \( \hat{n} \) require that

\[
T_{\nu} = \sum_k T_{n_k \hat{n}_k}
\]  
(2.13)

which enables one to calculate the jth component of the traction at a point across an arbitrary surface \( \hat{n} \), given the stress tensor \( T_{\mu} \) at that point.

Like the strain tensor, the symmetric stress tensor \( (T_{ij} = T_{ji}) \) can be decomposed into its irreducible parts,

\[
\begin{bmatrix}
T_{11} & T_{12} & T_{13} \\
T_{21} & T_{22} & T_{23} \\
T_{31} & T_{32} & T_{33}
\end{bmatrix} = [p] + [T_{\text{shear}}]
\]  
(2.14)

where \( p = \frac{1}{3} \sum_k T_{kk} \)  
(2.15)
The deviatoric stress tensor \([T^{\text{dev}}]\), being symmetric can be diagonalized like the deviatoric strain components to yield the Principal (normal) stresses at a point.

### 2.3 Elastic Energy and Relationship between Stress and Strain

The relationship between stress and strain, called the constitutive relation, is an important characteristic of the medium and can be elucidated from thermodynamic arguments. Consider an elastic body in equilibrium that suffers an infinitesimal strain characterized by the displacement of its constituent points by \(\delta U(R)\). The work \(\delta W\) done by the applied stresses against the elastic restoring forces in the course of this infinitesimal strain is:

\[
\delta W = \int (f \delta U) \, dv + \int (T \delta U) \, ds \tag{2.17}
\]

Since \((T, \delta u) ds = \delta u_i \{T_{i1} n_1 + T_{i2} n_2 + T_{i3} n_3\} ds\)

\[
= \delta u_i \{T_{i1} n_1 + T_{i2} n_2 + T_{i3} n_3\} ds \tag{2.18}
\]

\[
\therefore \int (T, \delta U) \, ds = \int (T, \delta U) \, ds = \int \nabla (T \delta U) \, dv \tag{2.19}
\]

But \(\nabla (T_{k1} \delta u_i) = \delta u_i \nabla T_{k1} + T_{k1} \nabla u_i\)

\[
\therefore \frac{\delta W}{dt} = d / dt \left[ (T_1 + \nabla T_{k1}) \delta u_1 \right] dv + d / dt \left[ (T_{k1} \nabla \delta u_1 + T_{k2} \nabla \delta u_2 + T_{k3} \nabla \delta u_3) \right] dv \tag{2.20}
\]
According to the first law of thermodynamics, the internal energy may change with deformation and the energy balance for the work done on the body is:

\[
\text{rate of doing mechanical work } (\delta w) + \text{rate of heating } \delta Q = \text{rate of increase of both kinetic } (\delta E_k) \text{ and internal energies } (\delta E_i)
\]

(2.21)

Where \(\delta E_i\) is the resulting change in internal energy. Substituting for \(\delta W\) from (2.20) in the above equation and noting that, in view of (2.11), the first integral on its RHS equals \((1/2)(d/dt)[2\rho(\partial U/\partial t)(\partial U/\partial t)/dt] = d/dt(\delta E_k)\), and that the second is the product \(T_{jk}\delta e_{jk}\) of the strain and the corresponding strain tensor, we obtain using (2.21),

\[
\delta E_i = \delta Q + T_{j}l\delta e_{j}l
\]

(2.22)

For adiabatic deformation \(\delta Q = 0 = T\delta S\) (for reversible processes), and (2.22) reduces to

\[
\delta E_i = T_{\mu}\delta e_{\mu} = dW \text{ if } \delta(\text{kinetic energy}) \text{ is extremely small}
\]

(2.23)

The internal energy density, \(E_i\), at any point of the body must depend on the local state of strain.

Or,

\[
E_i = E_i(e_{\mu}), \text{ and } \delta E_i = (\partial E_i/\partial e_{\mu})\delta e_{\mu}
\]

(2.24)

Comparing (2.23) and (2.24), we identify \(E_i\) as the strain energy function of the strain components that evokes stresses according to \(T_{\mu} = (\partial E_i/\partial e_{\mu})\). When deformations are small, the stress components can be expressed as linear functions of strain (Generalized Hooke's law).

Or,

\[
T_{\mu} = C_{\mu\nu\lambda}e_{\lambda}\mu
\]

(2.25)
where $C_{\mu
u}$ are called the elastic constants of the medium. The symmetry of $e$ and $T$, ensures that $C_{\mu\nu} = C_{\nu\mu} = C_{\mu\nu\mu} = C_{\nu\mu\mu}$. 

Further, from (2.25) we deduce that 

$$
(\partial T_{\mu\nu}/ \partial e_{\rho\sigma}) = (\partial^2 E_{\lambda}/ \partial e_{\mu\lambda} \partial e_{\rho\sigma}) = (\partial T_{\rho\sigma}/ \partial e_{\mu})
$$

Or, 

$$
C_{\rho\sigma\nu \mu} = C_{\rho\sigma\nu \mu} = C_{\rho\sigma\nu \mu} = C_{\rho\sigma\nu \mu} = C_{\nu\rho\mu \sigma} = C_{\nu\rho\mu \sigma} = C_{\nu\rho\mu \sigma}
$$

Therefore, the number of elastic constants necessary to specify the relation between stress and strain of any arbitrarily chosen material is reduced to 21. Rock materials of the earth, generally made up of randomly oriented crystals except where exceptional thermal and stress regimes may orient them preferentially, can be treated as being isotropic on the scale of a seismic wave length of a few kilometers. This condition drastically reduces the number of elastic constants to just two. For, the most general isotropic 4th order tensor such as $C_{\mu\rho\nu\sigma}$ having the (2.26) symmetries, can be expressed (Jeffereys) as:

$$
C_{\mu\rho\nu\sigma} = \lambda \delta_{\rho\mu} \delta_{\nu\sigma} + \mu (\delta_{\rho\sigma} \delta_{\mu\nu} + \delta_{\rho\mu} \delta_{\sigma\nu})
$$

(2.27)

So that only the $C_{\mu\rho\nu\sigma}$, $C_{\nu\rho\mu\sigma}$ and $C_{\mu\rho\mu\nu}$ terms survive and can be expressed in terms of just two elastic constants $\lambda$ and $\mu$, called the Lame's constants.

$$
C_{\mu\rho\nu\sigma} = (\lambda + 2\mu) \delta_{\rho\mu} \delta_{\nu\sigma}; C_{\nu\rho\mu\sigma} = \lambda \text{ and } C_{\mu\rho\mu\nu} = \mu
$$

(2.28)

The stress - strain relations, in turn, are thus distilled into

$$
T_{\mu\rho} = \lambda \theta \delta_{\mu\rho} + 2\mu e_{\mu\rho}
$$

(2.29)

Where

$$
\theta = e_{11} + e_{22} + e_{33}
$$

(2.30)
In turn, the various physically evocative elastic modulii, namely the Young’s modulus ($Y$), rigidity ($\mu$), incompressibility ($k$), and Poisson’s ratio ($\sigma$) can be expressed in terms of $\lambda$ and $\mu$ or in terms of each other. For example, the Poisson’s ratio can be written as:

$$\sigma = \frac{\lambda}{2(\lambda + \mu)} = \frac{\lambda}{(3k - \lambda)} = \frac{(3k - 2\mu)/2(3k + \mu)}{(3k - Y)/6k} \tag{2.31}$$

(2.23) and (2.25) also enable one to show that the work done by stresses $T_\eta$ in developing a state of strain equal to $e_\eta$ is equal to $\frac{1}{2}(T_\eta \cdot e_\eta)$.

### 2.4 The Dynamic equation

Using the constitutive equations (2.29), we can now write Newton’s equation for the conservation of momentum

$$\rho U = F + (\lambda + 2\mu)\nabla(\nabla \cdot U) - \mu\nabla \times \nabla \times U \tag{2.32}$$

(2.32) is the full equation of motion for an elastic medium from which we can deduce solutions to various specific problems of motion and deformation. For example, by replacing $F$ in the above equation by a concentrated point force at $R_0$ directed along $\hat{a}$

$$F = |F|\hat{a} \delta(R - R_0)$$

and ignoring inertial forces, we obtain the fundamental Kelvin solution which is the key to constructing solutions of a wide variety of problems involving a complex set of body forces. Another most important solution to the basic problem of elastic wave propagation is obtained by treating the body forces to be zero and seeking transient solutions of the resulting equation by considering the most general displacement vector composed of a conservative (gradient of a scalar) and a
solenoidal (curl of a vector) vector \( \mathbf{U}(X) = (\nabla \phi + \nabla \times \psi) \). Thus both \( \phi \) and \( \psi \) are found to satisfy the wave equation

\[
\nabla^2 \chi = \left( \frac{1}{V_x} \right) \left[ \frac{\partial^2 \chi}{\partial t^2} \right]
\]

where

\[
V^2_\phi = \left( \frac{1}{\lambda + 2\mu} \right) / \rho \quad \text{and} \quad V^2_\psi = \mu / \rho
\]

(2.34)

Solutions to the above wave equation can be determined by separation of the three spaces and the time coordinates. Thus, writing \( \chi = [X(x)Y(y)Z(z)T(t)] \) in the rectangular coordinate system and its substitution in (2.33), one obtains the following set of 4 coupled equations:

\[
\ddot{X} = k_1^2 \Delta X; \quad \ddot{Y} = k_2^2 \Delta Y; \quad \ddot{Z} = k_3^2 \Delta Z; \quad \ddot{T} = \omega^2 \Delta T
\]

Where, \( \left[ k_1^2 + k_2^2 + k_3^2 \right] = \left( \frac{\omega}{V_x} \right)^2 = K^2 \), and \( k_i \) are the cartesian components of the wave number vector \( \mathbf{K} \) which defines the propagation direction of the wave. The solution of (2.33) can now be written in a straightforward manner as:

\[
\chi = A \exp \left[ \pm j(\omega t - K \cdot X) \right]
\]

(2.35)

where \( \chi \) will be recognized as a forward travelling wave whose form is reproduced at a subsequent position \( (X + dX) \hat{k} \) and time \( (t + dt) \) provided it moves with a velocity \( (dX/dt) = (\omega/|K|) \). \( A \) signifies the amplitude of \( \chi \) and the quantity \( (\omega t \pm K \cdot X) \) describes its phase with respect to a zero phase reference at \( t = 0 \). In 3 dimensions as the wave radiates outwards from the source, there would be many points \( X_i \) in space whereat the phase \( (\omega t \pm K \cdot X_i) \) for any given time \( t \) would be the same. The locus of all these equiphase \( X_i \) points would be a surface called the wavefront \( S(X_i) \). We may, therefore write (2.35) in 3 dimensions as:
\[ \chi(X_1,t) = A \exp\left[ j(\vec{k} \cdot \nabla S(X_1) - \omega t) \right] \tag{2.36} \]

where \( \hat{k} \) is the unit vector normal to the wavefront at each point, defining the direction of propagation of that element of the wavefront analogous to the 'ray' in optics. The orientation of the ray at any time is therefore normal to the surface \( S(t) = \text{constant} \) or along \( \nabla S \).

Since the earth is largely a spherically layered medium the Fermat path between an earthquake source and a receiver lies in the plane connecting the two through the center of the earth. Assuming the \( x_1 \) axis to lie opposite the back azimuth of an arriving wave and the \( x_i \) axis to be vertical (positive downwards), one could further simplify the above solution to 2 dimensions:

\[ \psi(X_1,t) = A \exp\left[ j(\omega t \pm k_1 x_1 \pm k_2 x_2) \right] \tag{2.37} \]

This corresponds to a plane wave advancing along the vector \( \vec{K} \), the surface of constant phase \( \phi_0 \) at a given time \( t = 0 \), that is the wavefront, being given by the plane containing the \( x_1 \) axis and the line \( [x_1 = [\phi_0 - k_1 x_1]/k_2] \). If we define the angle \( \theta \) between the \( \vec{k} \) vector and the \( x_1 \) axis as the angle of incidence of the wave, we obtain

\[ k_1 = \omega \sin \theta / V_x = \omega p \quad \text{and} \quad k_2 = \omega \sin \theta / V_z = \omega \eta \]

\[ \tag{2.38} \]

where \( p \) will be recognized as the reciprocal of the apparent horizontal velocity called the horizontal slowness or the 'ray parameter' and \( \eta \) the vertical slowness. It is often more convenient to write the solution of the wave function in terms of \( p \) and \( \eta \).

Or, for a plane wave incident from below the surface propagating along the positive \( x_i \),

\[ \psi(X_1,t) = A \exp\left[ j(\omega x_1 - \eta x_1 - t) \right] \tag{2.39} \]

The \( \phi \) solution of the above equation represents a scalar wave polarized along the direction of propagation \( \hat{k} \), and is identified with the longitudinal or the P-wave. The \( \psi \) solution, on the other hand, represents a vector shear wave.
also traveling along $k$, but having its displacement components (from $U = \nabla \times \psi$) perpendicular to $k$ in the $x_i - x_j$ plane, constituting the SV component. The $x_j$ component involving purely horizontal displacements which in this case is zero, is called the SH component.

2.5 Short Wave Approximation: The Eikonal Equation

Solutions of the wave equation for any given instant of time yield the surface of constant phase (wavefront) normal at any of its point to the direction of propagation, so that all particle displacements thereon are exactly in step. Wavefronts evoke an intuitive understanding of radiating energy particularly in inhomogeneous media. In analogy with light waves, seismic waves can be shown to approximate as rays (the Eikonal equation) with respect to large-scale heterogeneities in the earth.

For simplicity, we consider the propagation of compressional waves in heterogeneous media. The displacement $u(= \nabla \phi)$ can be expressed in terms of the P-wave scalar potential $\phi$ as

$$\nabla^2 \phi - \frac{1}{\alpha^2} \frac{\partial^2 \phi}{\partial t^2} = 0$$

where the P-wave velocity, $\alpha$, is a function of position, $\alpha = \alpha(X) = [\omega / k(X)]$.

We assume a harmonic solution of the form

$$\phi = A(x) e^{-i\omega T(x)}$$

where $\omega T(X)$ represents the wave phase $= (\omega t \pm k.X)$, and $A$ the local amplitude. Substituting this value of $\phi$ in the wave equation and dividing out the constant $e^{-i\omega T(x)}$ factor, we obtain
\[ \nabla^2 A - \omega^2 A |\nabla T|^2 = -i \left[ 2 \omega A \cdot \nabla T^* + \omega A \nabla^2 T \right] - \frac{Am^2}{\alpha^2} \] (2.42)

Equating the real and imaginary parts, we obtain

\[ \nabla^2 A - \omega^2 A |\nabla T|^2 = -\frac{Am^2}{\alpha^2} \] (2.43)

\[ 2\omega A \cdot \nabla T + \omega A \nabla^2 T = 0 \] (2.44)

Dividing (2.43) by \( Am^2 \) and rearranging, we obtain

\[ |\nabla T|^2 - \frac{1}{\alpha^2} = -\frac{V^2 A}{Am^2} \] (2.45)

We now adopt the high frequency approximation that \( \omega \) is large and \( 1/\omega^2 \) can therefore be ignored,

\[ \therefore |\nabla T|^2 \approx \frac{1}{\alpha^2} \] (2.46)

This can be written in a more general form as \( |\nabla T|^2 = \frac{1}{c^2} \), called the Eikonal equation, where \( c \) is the local wave speed, \( \alpha(X) \) in the case of P-wave and \( \beta(X) \) in the case of S-waves.

The equation can also be expressed as \( |\nabla T|^2 = S^2 \) where \( S=1/c \) is called the slowness. Accordingly, we can write

\[ |\nabla T(X)|^2 = (\partial/\partial_x)^2 + (\partial/\partial_y)^2 + (\partial/\partial_z)^2 = S(X)^2 \] (2.47)

The phase factor \( T \) has a gradient whose amplitude is equal to the local slowness. The surface \( T(X) = \text{constant} \), defines surfaces on which the phase is constant i.e. wavefronts. Radiating energy along \( \nabla T(X) \) is perpendicular to the surface \( T(X) = \text{constant} \), and are therefore defined as rays.

\[ \nabla T = \hat{S} = S \] (2.48)

where \( \hat{S} \) marks the unit vector in the local ray direction and \( S \) is the slowness vector. The function \( T(X) \) has the units of time because the wavefronts
advance with local slowness along a direction parallel to the rays, and is simply the time required for a wavefront to reach the point X.

2.6 Partitioning of Energy across seismic discontinuities

When a body wave, P or S, traveling in a homogeneous medium encounters a seismic discontinuity across which the velocity changes, it is refracted according to Fermat's principle that requires the ray parameter \( p \) to be conserved by changing its direction to compensate for the change in velocity. In turn, this causes the stresses and displacements associated with the incident and refracted rays on the two sides of the boundary to change and become discontinuous unless supplemented by additional waves generated at the boundary. In general, therefore, the energy of an incident P or S-wave at a boundary is required to be partitioned into 4 new generated waves at the boundary in order to ensure the continuity of displacements and stresses across it. These are: the reflected P, reflected S, refracted P and refracted S, each obeying the Snell's law with respect to the parent wave. The S waves being of transverse nature, their particle motion can be split into two orthogonal components: one in the plane of propagation and the other normal to it. These S wave components are respectively denoted by SV and SH. The latter being confined to the plane perpendicular to the plane of propagation of P and S waves, are completely decoupled from the P wave particle displacement. Therefore the partitioning of an SH wave energy at a boundary will involve only two daughter waves: reflected and refracted SH waves.

Consider an incident P wave traveling in the \( x_1, x_2 \) direction, impinging on a horizontal boundary at an angle \( \theta \) (figure below). We can write the displacement potentials corresponding to all 5 waves: the parent and the 2 reflected ones in the first medium and the two refracted ones in the second, in the form (2.37) or (2.39), and require that the associated normal stresses and
displacements at each pair of points on either side of the boundary must be equal.

\[
\begin{align*}
\text{P wave} & \quad \gamma \quad \text{Sv} \\
\alpha_1, \beta_1, \rho_1 & \quad \alpha_2, \beta_2, \rho
\end{align*}
\]

It may be noted that since the particle motion associated with the incident P-wave is confined to the \( x_1 - x_1 \) plane, it will not generate any SH wave (particle motion in the \( x_1 - x_2 \) plane) at the interface. However, since the P wave displacements alone cannot combine to yield continuous displacements or tractions across the boundary, additional particle motion in the \( x_1 x_3 \) plane would be required at each point of the boundary to make up for the difference and this is exactly what the reflected and a refracted SV wave provides.

The P and SV wave potentials, \( \phi \) and \( \psi \), for the 5 wave components,

\[
\begin{align*}
\phi_{(\text{layer1})} &= \phi_{\text{incident-ray}} + \phi_{\text{reflected-ray}} \\
\phi_{(\text{layer2})} &= \phi_{\text{refracted}} \\
\psi^- &= \psi_{\text{reflected}} \\
\psi^+ &= \psi_{\text{refracted}}
\end{align*}
\]

can, in turn, be expressed by:

\[
\varphi_{\text{incident}} = A_1 \exp[i\omega(px_1 + \eta \alpha_1 x_3 - t)]
\]
CHAPTER II THEORETICAL FRAMEWORK

\( \phi_{\text{reflected}} = A_2 \exp[i\omega(px_1 - \eta_\alpha x_3 - t)] \)
\( \phi_{\text{refracted}} = A_3 \exp[i\omega(px_1 + \eta_\alpha x_3 - t)] \)
\( \psi_{\text{reflected}} = B_2 \exp[i\omega(nx_1 - \eta_\beta x_3 - t)] \)
\( \psi_{\text{refracted}} = B_3 \exp[i\omega(nx_1 + \eta_\beta x_3 - t)] \)  \hspace{1cm} (2.54)

Particle displacements associated with each of these waves can be derived for any general point on either side of the boundary from (2.72 & 2.73), using the assumed Helmholtz relation: \( U(X) = (\nabla \phi + \nabla \times \psi) \) and the various stress components from the relation \( [T_i] = \lambda \delta_{ij} + \mu \left\{ \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right\} \). Equating the stresses and displacements at points on either side of the boundary, yield the ratios of post-conversion amplitudes (\( A_2, A_3, B_1, B_2 \)) with respect to the incident amplitude (\( A_1 \)), respectively called the reflection and transmission coefficients. These coefficients determine the proportion of the incident energy partitioned among the derived waves. Calculations of these coefficients for a general boundary between two solid media involves complex algebraic manipulations but can be found in various texts (e.g. Lay & Wallace).

Accordingly,
\[
T_{PP} = \left[ 2 \rho_1 \eta_\alpha F(\alpha_1 / \alpha_2) \right] / D \]  \hspace{1cm} (2.55)
\[
T_{PS} = \left[ 2 \rho_1 \eta_\alpha Hp(\alpha_1 / \beta_2) \right] / D \]  \hspace{1cm} (2.56)
\[
R_{PP} = \left[ (b \eta_\alpha - c \eta_\alpha_2)F - (a + d \eta_\alpha_1 \eta_\beta_2)Hp^2 \right] / D \]  \hspace{1cm} (2.57)
\[
R_{PS} = -\left[ 2 \eta_\alpha_1(ab + cd \eta_\alpha_2 \eta_\beta_2)Hp(\alpha_1 / \beta_1) \right] / D \]  \hspace{1cm} (2.58)

where \( D = EF + GHp^2 \)  \hspace{1cm} (2.59)
\( E = b \eta_\alpha_1 + c \eta_\alpha_2 \)  \hspace{1cm} (2.60)
\( F = b \eta_\beta_1 + c \eta_\beta_2 \)  \hspace{1cm} (2.61)
\( G = a - d \eta_\alpha_1 \eta_\beta_2 \)  \hspace{1cm} (2.62)
\( H = a - d \eta_\alpha_2 \eta_\beta_1 \)  \hspace{1cm} (2.63)
\[ a = \rho_2 \left( 1 - 2 \beta_2^2 p^2 \right) - \rho_1 \left( 1 - 2 \beta_1^2 p^2 \right) \]  
(2.64)

\[ b = \rho_2 \left( 1 - 2 \beta_2^2 p^2 \right) - 2 \rho_1 \beta_1^2 p^2 \]  
(2.65)

\[ c = \rho_1 \left( 1 - 2 \beta_1^2 p^2 \right) + 2 \rho_2 \beta_2^2 p^2 \]  
(2.66)

\[ d = 2 \left( \rho_2 \beta_2^2 - \rho_1 \beta_1^2 \right) \]  
(2.67)

Products of the material density and the wave velocity such as \((\rho \alpha)\) or \((\rho \beta)\) represent the associated acoustic impedances of the medium. In the case of an incident P-wave at a horizontal boundary, the 4 coefficients for the converted phases can be simply determined for the case of vertical incidence \((I = 0 = \rho)\).

From (2.55-2.60), there are no reflected or transmitted shear waves because both \(R_{n,\alpha}\) and \(T_{n,\alpha}\) are zero. The coefficients of reflected and refracted P waves are respectively reduced to:

\[ T_{pp} = \left[ 2 \rho_1 \eta_{\alpha_1} F(\alpha_1 / \alpha_2) \right] / D = \left[ \left( (2 \rho_1 \alpha_1) / (\rho_1 \alpha_1 + \rho_2 \alpha_2) \right) \right] \]  
(2.68)

\[ R_{pp} = \left[ \left( b \eta_{\alpha_1} - c \eta_{\alpha_2} \right) F - \left( a + \rho \eta_{\alpha_1} \eta_{\beta_2} \right) H p^2 \right] / D \]

\[ = \left[ \left( (\rho_2 \alpha_2 - \rho_1 \alpha_1) / (\rho_1 \alpha_1 + \rho_2 \alpha_2) \right) \right] \]  
(2.69)

It may be noted that if the wave is traveling from a higher acoustic impedance medium to a lower one, the reflected wave has a negative coefficient signifying a phase reversal at the boundary. This is reflected in the sum of the quantities in (2.68) and (2.69) being equal to unity. Computation of \(T_{pp}\) and \(R_{pp}\) for different values of the incident angle \(i\), shows that these coefficients have very small gradients near \(i=0\). In fact the coefficients change very little even for incident angles as large as \(\sim 20^\circ\).
2.7 Converted Phases as Response to Time Invariant Linear System

As stated in the previous section, converted waves derived from a parent incident wave at a boundary, can also be expressed as the output of a linear system. The three main characteristics of a time invariant linear system are:

(i) **Superposition:** If \( g_1(t) \) is the output of an input \( f_1(t) \), and \( g_2(t) \) is the output of \( f_2(t) \), then the output of \( [f_1(t) + f_2(t)] \), is \( [g_1(t) + g_2(t)] \).

(ii) **Scaling:** If \( g_1(t) \) is the output of an input \( f_1(t) \), then \( a \) \( g_1(t) \) is the output of \( a \) \( f_1(t) \).

(iii) **Shifting (time invariance):** If \( g_1(t) \) is the output of an input \( f_1(t) \), then \( g_1(t-t_1) \) is the output of \( f_1(t-t_1) \).

Also, the response of a linear system to a unit impulse function called the Dirac-Delta function is known as the transfer function. A Delta function \( \delta(t-t_0) \) is defined as a function that is zero for all values of \( t \) except at \( t = t_0 \) where it tends to infinity in a manner that

\[
\int_{-\infty}^{\infty} \delta(t-t_0) dt = 1, \text{ and } f(t) = \int_{-\infty}^{\infty} f(\tau) \delta(t-\tau) d\tau \tag{2.70}
\]

Consider two time series \( f(t) \) and \( g(t) \). Then the convolution of these functions is defined as

\[
f(t) * g(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau) d\tau \tag{2.71}
\]

If \( g(t) = G[f(t)] \) is output of a linear system to an input \( f(t) \), we first express \( f(t) \) as an integral involving the delta function as
\[ f(t) = \int_{-\infty}^{\infty} f(\tau) \delta(t-\tau) d\tau \quad (2.72) \]

and the response of the system to the input \( |f(\tau)\delta(t-\tau)| \) where \( |f(\tau)| \) is the scalar value of \( f(t) \) at \( t = \tau \).

Thus the output \( g(t) = G[f(t)] \) of the system to the input

\[ f(t) = \int_{-\infty}^{\infty} f(\tau) \delta(t-\tau) d\tau \quad (2.73) \]

is

\[ G[f(t)] = \int_{-\infty}^{\infty} f(\tau) h(t-\tau) d\tau \quad (2.74) \]

which is known as the convolution integral and the process is called convolution. An inverse process called deconvolution can then yield a knowledge of either the input or the system function from a knowledge of the output or data.

### 2.8 Receiver Function

A Receiver Function is the record of converted Ps phases from underlying acoustic boundaries or the corresponding Impulse function response. As a P wave travels through the crust and upper mantle, it encounters velocity contrasts where the P wave can be reflected, transmitted and/or converted to as SV wave. Assuming that the distance between the earthquake events and seismometer station is large enough, the angle of incidence of the P wave will be close to vertical within the lithosphere beneath the station and so will be most strongly recorded on the vertical component seismogram, whereas the converted shear waves would be most prominent in the horizontal components.
Calculation of receiver function therefore would involve a rotation of the horizontal components to lie parallel and orthogonal to the great circle path along the Earth's surface, between the source and the receiver. The signal recorded by a seismometer can then be viewed as a convolution in the time domain:

\[ X_i(t) = S(t) * I(t) * E_i(t) \]

where \( Z, R \) and \( T \) represent the vertical, radial and transverse components of motion respectively, \( S(t) \) represents the near source and mantle propagation effects, \( I(t) \) is the instrument response and \( E_i(t) \) is the Earth's response below the receiver.

As the SH phases are a minor feature of the vertical seismogram (Langston, 1979) for teleseismic events, we can write

\[ X_Z(t) = S(t) * I(t) \]

Therefore a receiver function can be isolated from the radial or tangential seismograms by deconvolving the source and instrument responses, which corresponds to the vertical seismogram. Figure 2.1 shows the phases that arrive at a seismometer station as a result of interaction with a single crustal layer over an upper mantle half-space, and the corresponding radial receiver function.
Figure 2.1: (top) A series of seismic phases being generated at a single sub-receiver interface. Solid lines represent the P-wave and dotted lines represent the S-wave. (bottom) The radial receiver function corresponding to the simple earth model shown in top.

In general, of course there would be some energy scattered away from the Fermat plane, and would appear on the transverse component of the seismogram. Its amplitude can thus be deemed to provide a measure of the departure of the underlying earth structure from the ideal radially symmetrical, or over a limited region, a layered flat earth model.

As our objective is to investigate the shear wave speed structure of the earth to a depth of the sharpest discontinuity, and this information is contained in $E_R(t)$, we first need to abstract this function from the seismogram before proceeding to invert this function to obtain the velocity structure.
There exist several methods of extracting receiver function. As for example, the water level stabilized frequency domain deconvolution process (Clayton and Wiggins, 1976), the time domain approach (e.g. Gurrola et al., 1995 and Sheehan et al., 1995) and the iterative time domain deconvolution procedure (Pablo Ligorria and Ammon, 1999).

Frequency domain deconvolution technique exploits the fact that convolution of two functions in the time domain is equal to the product of their Fourier Transforms of each. So, we can take the Fourier transforms of the quantities to get the corresponding equation in the frequency domain. Finally, by taking the Inverse Fourier Transform of $E_R(t)$ one can recover $E_R(t)$.

Alternatively, one can abstract $E_R(t)$ directly without forward and backward Fourier transformation, by recognizing the fact that this function when convolved with the vertical yields the radial, and therefore building a function iteratively, which when convolved with the vertical seismogram would yield the radial. Specifically, the process is based on the source-time function estimation algorithms of Kikuchi and Kanamori (1991). The process begins by cross-correlating the vertical and radial component seismograms to establish the lag of the first and largest spike in the receiver function. Convolution of the current estimate of the receiver function with the vertical component seismogram is then subtracted from the radial component seismogram to form an updated radial seismogram. This means that after one iteration there exists a synthetic receiver function containing only one spike which models the direct P arrival, and a new radial seismogram lacking the direct P arrival. The same process is then repeated several times using the iteratively updated radial seismograms to look for the subsequent spike lags and amplitudes. At each iteration a misfit function between the vertical seismogram and receiver function convolution, with the real radial component seismogram is computed. By repeating the above procedure, we
finally get a train of delta functions which when convolved with \( E_Z(t) \) gives a function that is very close to \( E_R(t) \), i.e., the desired receiver function.

2.9 Surface Wave Dispersion, Phase and Group Velocity

Surface wave dispersion measurements have been employed for a long time to understand the average structure of the crust and upper mantle. Surface wave group and phase velocities are mainly sensitive to the shear structure of the underlying crust and upper mantle. A pulse shaped waveform become spread out (dispersed) as the different frequency components become spatially separated when the different frequency components of the wave packet travel at different velocities. The associated phase velocity of each frequency component defines the wave speed of that particular frequency. Also the various frequency components of the signal will undergo interference effects leading to the overall canceling out of energy except at particular times defined by the group velocity of the wave as the waveform, originally a pulse of energy, become dispersed.

Consider a seismic signal with only two frequencies \( \omega_1 \) and \( \omega_2 \). Then we have

\[
U(x, t) = \cos(\omega_1 t - k_1 x) + \cos(\omega_2 t - k_2 x) \tag{2.77}
\]

where \( U(x, t) \) is the observed displacement due to the seismic signal, \( \omega_1 = \omega - \delta \omega \), \( \omega_2 = \omega + \delta \omega \), \( k_1 = k - \delta k \), \( k_2 = k + \delta k \) and \( \omega, k \) are the average frequency and wave number respectively.

Substituting the values of \( \omega_1, \omega_2, k_1, k_2 \) in equation (2.94), we obtain

\[
U(x, t) = 2 \cos(\omega t - k x) + \cos(\delta k x - \delta \omega t) \tag{2.78}
\]

In the above equation \( \omega \) is the average frequency and \( \delta \omega \) is the modulated frequency of the resultant waveform defined by the equation (2.106). The
wave speed for the shorter period signal is \( c = \frac{\sigma}{k} \) (phase velocity) and for the longer period is \( u = \frac{\partial \sigma}{\partial k} \) (group velocity) (Lay and Wallace, 1995).

2.10 Inverse Theory

Most of the physical problems require estimation of system parameters, which cannot be obtained directly from the observed data. This inverse estimation depends on the availability of a physical theory, that defines the physical influences by relating the values of the parameters \([m]\) to observed data \([d]\), in a form \( Gm = d \), where \( m \) may represent a column vector of \( m \) quantities to be estimated, \( d \) a column vector of \( n \) data sets \( d_{i,n} \), and \( G \) the Green's Function matrix \( G_{jk} \). It is very rare to have a unique inverse \((G^{-1})\) of \( G \) such that \( G \, G^{-1} = G^{-1} \, G = I \), which enables one to calculate \( m \) from the data \( d \) by the inverse relation \( m = G^{-1} \, d \).

In spite of the fact that the data (observable) are completely free from error, \( G \) must be square and non-singular to avail a unique inverse \( G^{-1} \) of \( G \). Since all physical measurements involve some errors, experiments are usually designed in a manner that the presence of errors in the data is indicated, and the measurements are made in large numbers than the model parameter set. Such a problem is known as an over-determined problem and the solution of such problems are obtained by using minimization techniques.

2.10.1 Least Square Solution

Consider that \( G \, m = d \). Since this is an over determined system, \( G \) will not be a square matrix and consequently will have no unique inverse.
Let $G'$ be an acceptable inverse of $G$ i.e. $[G'G = I]$.

Then we can write

$$G'(Gm) = G' \hat{d} = \hat{m}$$  \hspace{1cm} (2.79)

$$G\hat{m} = \hat{d}$$  \hspace{1cm} (2.80)

Now we minimize

$$|d - G\hat{m}| \hspace{0.5cm} \text{i.e.} \hspace{0.5cm} (d - G\hat{m})^2 = E$$  \hspace{1cm} (2.81)

in order to make sure that $\hat{m}$ is very close to $m$.

Replacing $\hat{m}$ by $m$, we obtain

$$E = (d - Gm)^T (d - Gm)$$

$$= \sum_{i} \left[ d_i - \sum_{j} G_{ij} m_j \right] \left[ d_i - \sum_{k} G_{ik} m_k \right]$$

$$= \sum_{i} \left[ d_i d_i - d_i \sum_{j} G_{ij} m_j - d_i \sum_{k} G_{ik} m_k + \sum_{i} \sum_{j} G_{ij} G_{ik} m_j m_k \right]$$

$$= \sum_{j} \sum_{k} \left[ m_j m_k \sum_{i} G_{ij} G_{ik} \right] - 2 \sum_{i} \left[ m_i \sum_{j} G_{ij} d_j \right] + \sum_{i} d_i d_i$$  \hspace{1cm} (2.82)

Now we demand a model $m (m_q)$ for which

$$\frac{\delta E}{\delta m_q} = 0.$$  \hspace{1cm} (2.83)

As such

$$\frac{\delta E}{\delta m_q} = \sum_{j} \sum_{k} \left[ \delta_{jq} m_k + m_j \delta_{kq} \right] \sum_{i} G_{ij} G_{ik} - 2 \sum_{i} \delta_{jq} \sum_{j} G_{ij} d_j$$

$$= \sum_{j} \sum_{k} \left[ m_k \sum_{i} G_{iq} G_{ik} + m_j \sum_{i} G_{ij} G_{iq} \right] - 2 \sum_{i} G_{iq} d_i$$

$$= 2 \sum_{k} m_k \sum_{i} G_{iq} G_{ik} - 2 \sum_{i} G_{iq} d_i$$  \hspace{1cm} (2.84)
We can express the equation in matrix form as

$$G^T G m - G^T d = 0$$ \hspace{1cm} (2.85)

Thus, the least square solution of the inverse problem $G m = d$ is given by

$$m = (G^T G)^{-1} G^T d.$$ \hspace{1cm} (2.86)

### 2.10.2 Finding of Velocity Structure of the Earth using Minimization Technique

Simultaneous inversion of receiver function and surface wave dispersion can be employed to determine a single velocity model beneath the earth surface. This inversion scheme seeks to find a minimum to the following objective function,

$$S = [(1-p)N_r + pN_s] \left\{ \frac{1}{N_r} \sum_{i=0}^{N_r} \left( \frac{O_{ri} - P_{ri}}{\sigma_{ri}} \right)^2 + \frac{p}{N_s} \sum_{j=0}^{N_s} \left( \frac{O_{sj} - P_{sj}}{\sigma_{sj}} \right)^2 \right\}$$ \hspace{1cm} (2.87)

where $S$ is the standard error between the observed and predicted data, $p$ is a weighting value, $O_{ri}$ is the $i^{th}$ observed receiver function measurement, $O_{sj}$ is the $j^{th}$ observed surface wave dispersion measurement and $P_{ri}, P_{sj}$ are the corresponding predicted values based on the current velocity model. $N_r$ and $N_s$ are the number of receiver function and surface wave measurements respectively and $\sigma_{ri}$ and $\sigma_{sj}$ are the standard error limits for the $i^{th}$ and $j^{th}$ pieces of receiver function and surface wave information respectively.

Consider the current model $M_0$ and the residual matrix $[d]$ corresponding to the model. Then the matrix $[d]$ is represented by a linear combination of changes to the current model as follows:

$$w[G] [\Delta m] = [d]$$ \hspace{1cm} (2.88)
where \( [G] \) is the partial derivative matrix with respect to changes in the current model, \([\delta m]\) is the model correction vector and \(w\) is the weighting system.

If we expand the above equation it will take the form as follows:

\[
\begin{bmatrix}
w_r & 0 \\
0 & w_s
\end{bmatrix}
\begin{bmatrix}
\frac{\delta r}{\delta m_1} & \frac{\delta r}{\delta m_2} & \cdots & \frac{\delta r}{\delta m_N} \\
\frac{\delta x}{\delta m_1} & \frac{\delta x}{\delta m_2} & \cdots & \frac{\delta x}{\delta m_N}
\end{bmatrix}
\begin{bmatrix}
\Delta m_1 \\
\cdots \\
\Delta m_N
\end{bmatrix} = [d]
\]

\( (2.89) \)

where

\[
w_r = \left[ (1-p)N_r + pN_s \right] \left( \frac{1-p}{N_r \sigma_r} \right)^{1/2}
\]

\( (2.90) \)

\[
w_s = \left[ (1-p)N_r + pN_s \right] \left( \frac{p}{N_s \sigma_s} \right)^{1/2}
\]

\( (2.91) \)

The equation (2.117) can be written in a short notation as

\[
[d] = \begin{bmatrix} w_{r, \text{res}} \\ w_{s, \text{res}} \end{bmatrix}
\]

\( (2.92) \)

where

\[
w_{r, \text{res}} = w_r \sum_{i=1}^{N} \left[ \frac{\delta r}{\delta m_i} \Delta m_i \right]
\]

\( (2.93) \)

and

\[
w_{s, \text{res}} = w_s \sum_{i=1}^{N} \left[ \frac{\delta x}{\delta m_i} \Delta m_i \right]
\]

\( (2.94) \)

Singular value decomposition technique is employed to minimize the objective function \( S \). Being the no-linearity involved in the forward problem, an iterative sequence of linearised inversion steps are used to approach a minimum of \( S \).
2.11 Seismic Anisotropy

Seismic rays, by analogy to optical rays also suffer birefringence when crossing material media with directional properties. Geological materials invariably exhibit this property because of prevailing stresses and fractures in the earth or due to preferential alignment of anisotropic crystals such as olivine which is one of the principal minerals in the earth's mantle. A seismic shear wave entering upon an anisotropic region thus splits into two phases with polarizations and velocities that mimic the properties of the anisotropic media. The phases, polarized into fast and slow directions, progressively separate in time as they propagate through the anisotropic media. This split is preserved in any isotropic segment along the ray path and can be detected as a time delay between two horizontal components of motion, its polarity and amplitude being strongly affected by the azimuth of an arriving wave. The existence of such split S-waves, whenever detected, therefore signifies the presence of anisotropy and an analysis of the polarity and amplitudes of the split phases, yields knowledge of the anisotropy characteristics of the subsurface.

In an anisotropic medium, there are three plane waves that propagate with three different speeds and with mutually perpendicular polarizations. Because of the near parallelism (not more than 10° away from parallel and perpendicular to the propagation direction), the three waves are often referred to as the quasi-P and quasi-S waves. Quasi-SV and quasi-SH are occasionally used instead for the polarizations that are closest to being in and out of the sagittal plane respectively.

The polarizations and speeds of the two quasi-S waves are determined by both the properties of the medium and the propagation direction through the medium. The time delay ($\Delta t$) between the two quasi-S waves depends on both
the path length in the anisotropic material and the difference in speed between the two quasi-S waves. This can be obtained by

\[
\delta t = L\left(\frac{1}{V_{S1}} - \frac{1}{V_{S2}}\right)
\]

(2.95)

where \(V_{S1}\) and \(V_{S2}\) represent the speeds of the two quasi-shear waves given the direction of propagation and the material properties and \(L\) is the length of the anisotropic path traversed. The differences between \(V_{S1}\) and \(V_{S2}\) are often reported as a percent anisotropy \(K_s\), calculated as

\[
K_s = \frac{200(V_{S1} - V_{S2})}{(V_{S1} + V_{S2})}
\]

(2.96)

However, this anisotropy should be distinguished from the intrinsic anisotropy of the material, which is defined as the percent difference between the maximum and minimum velocities (Birch, 1961).

A shear wave traversing in an anisotropic layer splits according to equation (2.112). The splitting is preserved if the shear wave then subsequently traverses an isotropic layer and the polarization of the first shear wave (\(\phi\)) and \(\delta t\) can be determined. \(\phi\) and \(\delta t\) are called the shear wave splitting parameters, or simply the splitting parameters. The amplitude of the fast and slow shear waves is directly proportional to the amplitude of the initial \(S\) wave in the fast and slow directions, respectively (the \(S\) wave polarization). After passing through a single anisotropic layer and entering an isotropic layer, the \(P\) wave has the same linear particle motion as before, but the \(S\) wave has a characteristic shape that depends on \(\delta t\) compared to the period. For large splitting or short periods the two split shear waves may be entirely separated from each other and the polarization diagrams take on a cruciform appearance with the initial polarization from the fast quasi-\(S\) wave and one from the slow quasi-\(S\) wave (e.g., Keith and Crampin, 1977). On the other hand, for smaller splitting or longer periods the two split waves are separated
by only a fraction of a period and the initial portion of the particle motion is linear and polarized in the fast direction, but the majority of the waveform has elliptical particle motion. With assumptions about the type and degree of anisotropy, \( \phi \) determines the orientation of the anisotropic symmetry system, and \( \delta \) gives the thickness of the anisotropic layer. Furthermore, by knowing the relationship between the orientation of the system and the strain, one can infer the orientation and magnitude of the strain.