UNIVERSIDADE FEDERAL DA BAHIA
CENTRO DE PESQUISA EM GEOFÍSICA E GEOLOGIA

AN INTRODUCTION TO SEISMOLOGY

Oldrich Novotny

Lecture notes for post-graduate studies
(preliminary text)

Instituto de Física
Instituto de Geociências

Salvador, Bahia, 1999

Contents

Preface

During the long historical development of the Earth sciences, several independent scientific disciplines have been constituted, such as geodesy, geology, geophysics, geochemistry and geography. Geophysics (physics of the Earth) is a branch of physics which investigates, by means of physical methods, the phenomena and processes occurring in the Earth and its immediate vicinity. According to analogies with physics, geophysics is further subdivided into smaller disciplines, such as gravimetry, seismology, geomagnetism and geoelectricity, geotermics and radioactivity of the Earth.

Seismology deals with earthquakes and with propagation of mechanical waves generated by earthquakes and by artificial sources. Similarly to other physical disciplines, seismology can also be divided into theoretical and observational (experimental) seismology, or into pure and applied seismology. In the present lecture notes we pay the main attention to theoretical aspects of pure seismology, although seismological observations and some problems of applied geophysics are also briefly discussed.

Seismological observations can be divided into two groups: *macroseismic* observations and *instrumental* (microseismic) observations. By the macroseismic observations we understand the observations of the effects of earthquakes on people, animals, construction objects and nature objects. No special instruments are used in macroseismic observations.

The construction of automatically registering seismographs in the second half of the 19^{th} century extended the possibilities of earthquake research tremendously. The seismographs made it possible to obtain the time dependence of earthquake vibrations, and also to register seismic waves of distant earthquakes. This opened possibilities for global studies of earthquakes, including earthquakes in uninhabited regions, such as oceans, high mountains and others. This progress in observational techniques has gradually turned seismology into the main discipline of studying the Earth's interior. Seismic methods also play a dominant role in many problems of geophysical prospecting.

The present text was prepared for the lecture "Seismology" which I read in 1998 and 1999 for post-graduate students at the Universidade Federal da Bahia (UFBA), Salvador, Brazil.

In preparing this text we have drawn mainly on the manuscript by Cerveny (1989), the lecture notes by Psencik (1994), and also on the textbooks by Båth (1979), Bullen and Bolt (1993), Fowler (1990), Richter (1958) and the lecture notes by Cerveny (1976, 1978). The text in the present lecture notes is more elementary than in the most textbooks mentioned above but, on the other hand, we have paid considerable attention to the mathematical and physical foundations of the corresponding seismological methods. We have attempted to emphasise relations to analogous problems in physics, rather than to describe seismological applications in detail. We hope that this text might be useful for many students of physics which start to specialise in geophysics. Nevertheless, some complicated formulae and their derivation (e.g., the Zöppritz equations) have been included into the text, although they are usually omitted in the standard textbooks.

Chapter 1 contains basic information on earthquakes, their effects, intensity scales, physical causes and problems of their predicting. Chapter 2 deals with the instrumental observations of seismic waves and with the elementary methods of their interpretations. The next three chapters are devoted to the computation of seismic rays and travel times in inhomogeneous media; the corresponding theory is based on Fermat's principle and on some analogies with analytical mechanics. Since the theory of seismic waves is based on continuum mechanics, in Chapter 6 we summarise the basic relations of continuum mechanics and their derivation. Chapters 7 to 11 deal with various types of elastic waves in homogeneous media, and Chapter 12 with the reflection and transmission of plane elastic waves at a plane interface. A more detailed description of the ray theory, including the computation of amplitudes, is given in Chapter 13.

The selection of the material in this lecture notes was partially influenced by the scientific orientation of the students, and so it does not cover all directions of the contemporary seismological research. Since the lecture "Attenuation and Dispersion of Elastic Waves" is also read to the students of the UFBA, we have omitted the theory of surface waves and matrix methods in seismology; we refer the reader to the lecture notes by Novotny (1999). Furthermore, we do not deal with free oscillations of the Earth, seismic instruments, interpretation of seismic data, and the constitution of the Earth. Only fragments of these problems are mentioned in the text. The reader may find a detailed description of these problems in the above-mentioned textbooks or other specialised literature.

The text of the lecture notes did not pass any language revision, which it necessarily needs.

I am very obliged to Prof. Vlastislav Cerveny for substantial contributions to this text. Prof. Cerveny offered me his manuscript, which forms essential parts of many chapters. I would like to express my gratitude also to the students and professors of the UFBA who contributed by their questions and comments to improvements of the text. To my wife, Šárka Novotná, I am obliged for the technical preparation of the text and figures. I shall be grateful to every reader for any critical comments and remarks to this text.

I would like to thank the Centro de Pesquisa em Geofísica e Geologia (CPGG/UFBA), Departamento de Geofísica Nuclear do Instituto de Física, and the Instituto de Geociências for providing me with favourable conditions for preparing this text. I wish to express my thanks especially to the CNPq (Conselho National de Desenvolvimento Científico e Tecnológico), to the Ministério de Educação e Cultura, and to the CPGG for providing me with the fellowships which made my stay at the Universidade Federal da Bahia possible.

Salvador, 1999

Oldrich Novotny

Chapter 1

Earthquakes

Earthquakes arise through a sudden release of energy within some confined region of the Earth. This energy is usually elastic strain energy, but may also be the gravitational energy, kinetic energy, chemical energy, etc.

Several types of earthquakes can be distinguished. The main of these types are as follows:

- Tectonic earthquakes, which are caused by the release of elastic strain energy. These earthquakes occur when the stresses in some region inside the Earth have accumulated to a value which exceeds the strength of the material. This leads rapidly to a fracture along some fault. The strongest world earthquakes belong to this category.
- Volcanic earthquakes. Many earthquakes of this type are often observed in volcanic regions, but they are usually small.
- Deep-focus earthquakes, indicating subduction zones (slabs) and having a similar character as tectonic earthquakes.
- Induced earthquakes, which are caused by a human activity.

Two basic terms related to the earthquake are the earthquake hypocentre and epicentre. The point where the rupturing process starts is called the hypocentre and the relevant time is called the hypocentral time. The hypocentre coordinates and the hypocentral time can by determined from seismometric records (from the first arrivals of seismic waves) by the procedure called the localisation of earthquakes. The vertical projection of the hypocentre onto the Earth's surface is called the epicentre. Both the hypocentre and epicentre are points.

In investigating the earthquakes, the following types of observations are used:

- macroseismic observations;
- instrumental observations using a world system of seismic stations;
- instrumental observations in the epicentral region.

1.1. Examples of large earthquakes

In order to form an idea on possible effects of earthquakes, let us give examples of some large earthquakes. The list given below contains earthquakes which were very strong from the point of view of their effects or released energy. The loss of lives is denoted by the sign † followed by the corresponding number. The quantity M, called the earthquake magnitude, characterises the strength of the earthquake from an energetic point of view. This quantity will be introduced later.

1556, 27.9. - **China, Province Shensi,** † 830 000. The largest catastrophic earthquake known from the historical times.

1737, 11.10. - **India, Calcutta**, † 300 000.

1755, 1.11. - **Portugal, Lisbon,** † 60 000. The sea wave, called the tsunami, reached the height of 20 m. Many people running to the coast after the earthquake were drowned in the great sea waves. The earthquake was felt up to central Europe; the mineral springs at the Teplice spa (Czech Republic) changed for several days.

1891, 28.10. - **Japan, Mino Owari,** † 7 200. A block was displaced at a length of about 100 km, horizontal displacements of up to 4 m, vertical ones of up to 7 m. It indicated a shear mechanism of strong earthquakes, which contradicted the hypotheses explaining earthquakes as underground explosions. This fact played an important role in the development of seismology.

1897, 12.6. - **India, Assam.** Visible amplitudes of the ground motion of up to 30 cm. Felt on an area of 4 million square kilometres; on an area of 75 000 km^2 all houses were damaged. Acceleration of up to 0.5 g, macroseismic intensity $I_0 = 12$ grades (see below), $M = 8.7$. The number of victims was relatively small († 1425).

1906, 18.4. - **USA, San Francisco,** † 750, $M = 8.2$. Fissures in the ground, horizontal displacements of up to 7 m. The earthquake occurred in the region of the San Andreas fault, which is geologically very active. Observed on an area of 10^6 km^2; damage of 524 million dollars. A great fire on the second and third days damaged practically the whole city. Survey retriangulations led to elastic-rebound theory of earthquakes, proposed by Reid. (If a new similar earthquake occurred at the same place now, damage would be much higher. Consequently, detailed monitoring and other investigations are carried out in the region of the San Andreas fault and other regions of California).

1908, 28.12. - **Italy, Messina,** † 83 000, $M = 7.5$. A well investigated earthquake, 0.2 g, tsunami of 12 m..

1920, 16.12. - **China, Province Kan-su,** † 100 000, $M = 8.5$. About 700 000 houses damaged.

1923, 1.9. - **Japan, Kwanto,** great Japanese earthquake, † 142 800, $M = 8.2$. Destruction in Tokyo and Yokohama. About a million houses destroyed or heavily damaged; damage of 2.8 billion dollars. Tsunami of 13 m. Earthquake Research Institute (Tokyo) founded as a consequence.

1948, 5.10. - **USSR, Turkmenia, Ashkhabad,** $M = 7.6$. Reports on anomalous behaviour of animals before the earthquake. An earthquake prediction programme adopted by Soviet seismologists in consequence.

1952, 4.11 - **USSR, Kamchatka.** The first observation of free oscillations of the Earth generated by this earthquake.

1960, 22.5. - **Chile**, † 5 700, $M = 8.5$. Damage of half a billion dollars. One of the strongest earthquakes of the 20th century.

1963, 26.7. - **Yugoslavia, Skopje,** † 1 074, $M = 6.0$, $ 300 million. A seismological institute for the Balkan region founded.

1964, 28.3. - **Alaska, Anchorage,** † 131, M = 8.6, $ 538 million. A huge earthquake from the point of view of released energy. (A Wiechert seismograph at the Prague seismic station, at a distance of many thousands of kilometres, fell out of its bearings).

1966, 26.4. - **USSR, Uzbekistan, Tashkent,** $ 300 million. A relatively weak, but shallow earthquake immediately below the city; construction of a new Tashkent. Changes in the contents of radon in underground waters were observed before the earthquake.

1970, 31.5. - **Peru,** † 52 000, M = 7.8, $ 507 million. A stony avalanche, released by the earthquake, buried the town of Yungai (a Czechoslovak alpinist expedition on Huascaran died under this avalanche).

1972, 23.12. - **Nicaragua, Managua,** † 5 000, $ 800 million.

1975, 4.2. - **China, Haicheng,** M = 7.3. The first predicted earthquake (on the basis of hydrological and other precursory phenomena). Evacuation of people preceded this destructive earthquake.

1976, 6.5. - **Italy, Friuli,** † 1 000, M = 6.5, $ 2 billion.

1976, 27.7. - **China, Tangshan,** † 665 000, M = 7.6. The town of Tangshan with 1.6 million inhabitants was practically destroyed, and a great damage occurred also in the vicinity, which is densely inhabited. The greatest earthquake of this century according to its effects.

1977, 4.3. - **Romania, Bucharest,** † 1 581, M = 7.2, $ 800 million. A little damage even to a distant nuclear power plant Kozloduy in Bulgaria; so that more strict building codes for nuclear power plants were adopted in some countries as a consequence.

1988, 7.12. - **USSR, Armenia, Spitak,** † 25 000.

1.2. Macroseismic observations

By macroseismic observations we mean field observations and observations of the effects of earthquakes by people (not by seismometers). The principal macroseismic effects of earthquakes are summarised in Tab. 1.1.

Table 1.1. Principal macroseismic effects of tectonic earthquakes; from Richter (1958).

Effect on	Primary	Secondary permanent	Secondary transient
Terrain	Regional warping, etc. Scarps Offsets Fissures, mole tracks, other trace phenomena Elevation or depression of coasts: changes in coast line	Landslides (slumps, flows, avalanches, lurches) Secondary fissures Sand craters Raising of posts and piles	Visible waves Perceptible shaking
Water	Damming; waterfalls;		Changes in well levels

	diversion Sag ponds Changes in wells, springs		Earthquake fountains Water over stream banks Seiches Tsunamis Seaquakes
Works of construction and	Offsets, and destruction or damage by rending or crushing; buildings, bridges, pipe lines, railways, fences, roads, ditches	Most ordinary damage to buildings, chimneys, windows, plaster	Creaking of frame Swaying of bridges tall structures
Loose objects		Displacement (including apparent rotation) Overturning, fall, projection (horizontal or vertical)	Rocking Swinging Shaking Rattling
Miscellaneous		Clocks stop, change rate, etc. Glacier affected Fishes killed Cable breaks	Nausea Fright, panic Sleepers wakened Animals disturbed Birds disturbed Trees shaken Bells rung Automobiles, standing or in motion, disturbed Audible sound Flashes of light

The macroseismic intensity, usually denoted by I, characterises the strength of the earthquake at a given place on the basis of its macroseismic effects. A number of different macroseismic scales have been set up for determining the macroseismic intensity.

The first widely adopted scale was the Rossi-Forel (R.F.) scale, which was set up in 1883, and had ten grades. However, an enormous range of intensity was put together at the highest level, X, and the description of some effects was too specifically European. These defects were largely removed in the Mercalli-Cancani-Sieberg scale (commonly abbreviated MCS). The original version of this scale was put forward by Mercalli in 1902 at first with ten grades, later with twelve grades following a suggestion by Cancani who also attempted to express these grades in terms of acceleration. An elaboration of the scale was published by Sieberg in 1923. Further modifications of the MCS scale were proposed by seismologists in the USA and Europe.

At present, the most commonly used macroseismic scales are as follows:

- MM....Modified Mercalli Scale, (modified by Wood and Neumann in 1931), 12 grades, used in the USA;
- MSK...Medvedev-Sponheuer-Karnik scale (1964), 12 grades, used in Europe;
- JMA...Japan Meteorological Agency scale, 7 grades, used in Japan.

The scales MM and MSK are very similar. The general characteristics of the MM scale are given in Tab. 1.2.

An up-dated version of the MSK scale was proposed by the European Seismological Commission in 1992 (Grünthal, 1993). The main changes in the up-dated scale are as follows: the effects on new types of constructions have been added (constructions from reinforced concrete and others), some irregularities in the scale have been removed (a large step between the previous grades VI and VII), and a systematic quantification of earthquake effects has been introduces, which facilitates a statistical processing of macroseismic data.

Table 1.2. Modified Mercalli scale of intensity (abridged); from Bullen and Bolt (1993).

I. Not felt except by a few under especially favourable circumstances.

II. Felt only by a few persons at rest, especially on upper floors of buildings. Delicately suspended objects may swing.

III. Felt quite noticeably indoors, especially on upper floors of buildings, but many people do not recognise it as an earthquake. Standing motor cars may rock slightly. Vibration like passing of truck. Duration estimated.

IV. During the day felt indoors by many, outdoors by few. At night some awakened. Dishes, windows, doors disturbed, walls make creaking sound. Sensation like heavy truck striking building. Standing motor cars rocked noticeably.

V. Felt by nearly everyone, many awakened. Some dishes, windows, etc., broken; a few instances of cracked plaster; unstable objects overturned. Disturbance of trees, poles, and other tall objects sometimes noticed. Pendulum clocks may stop.

VI. Felt by all; many frightened and run outdoors. Some heavy furniture moved; a few instances of fallen plaster or damaged chimneys. Damage slight.

VII. Everybody runs outdoors. Damage negligible in buildings of good design and construction; slight to moderate in well-built ordinary structures; considerable in poorly built or badly designed structures; some chimneys broken. Noticed by persons driving motor cars.

VIII. Damage slight in specially designed structures; considerable in ordinary substantial buildings, with partial collapse; great in poorly built structures. Panel walls thrown out of frame structures. Fall of chimneys, factory stacks, columns, monuments, walls. Heavy furniture overturned. Sand and mud ejected in small amounts. Changes well water. Disturbs persons driving motor cars.

IX. Damage considerable in specially designed structures; well-designed frame structures thrown out of plumb, great in substantial buildings, with partial collapse. Buildings shifted off foundations. Ground cracked conspicuously. Underground pipes broken.
X. Some well-built wooden structures destroyed; most masonry and frame structures destroyed with foundations; ground badly cracked. Rails bent. Landslides considerable from river banks and steep slopes. Shifted sand and mud. Water splashed (slopped) over banks.
XI. Few, if any, (masonry) structures remain standing. Bridges destroyed. Broad fissures in ground. Underground pipe-lines completely out of service. Earth slumps and land slips in soft ground. Rails bent greatly.
XII. Damage total. Waves seen on ground surfaces. Lines of sight and level distorted. Objects thrown upward into the air.

In processing the macroseismic observations, the following maps and parameters are usually determined:

a) Maps. The macroseismic intensities *I* observed at individual geographic locations are shown in the map, either by numbers or by various symbols.
b) Isoseismal curves. In the maps, the isoseismal curves are the boundaries between the observations of different intensity. For example, the isoseismal curve 7 is a boundary of all observations of 7° and higher. In other words, it separates observations of intensity 6° and 7°. An example of a map with isoseismal curves is given in Fig. 1.1.

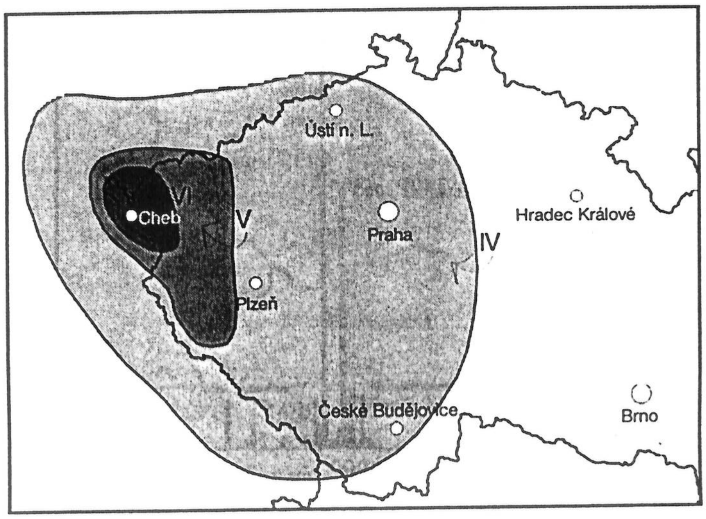


Fig. 1.1. The map of macroseismic intensities for the earthquake of December 21, 1985, in Western Bohemia, Czech Republic; magnitude 4.6.

c) Macroseismic epicentre. It is the region of highest observed macroseismic intensities. The macroseismic epicentre may be situated at a different place than the instrumental epicentre.

d) The macroseismic epicentral intensity I_0. It is the macroseismic intensity at the macroseismic epicentre. It characterises the strength of the earthquake.

e) The averaged radii of isoseismal curves. The averaged radius of the isoseismal curve I_n is usually denoted by r_n.

f) Shape of isoseismal curves. Usually these curves are not accurately spherical, which is a consequence of the geological structure and fault mechanism.

1.3. Problems of a physical interpretation of macroseismic intensity

The classification of earthquake effects by means of the macroseismic grades is important from social and economic points of view. It is also used for purposes of earthquake insurance, since insurance companies in many countries cover the damage caused by earthquakes if the macroseismic intensity exceeds a certain grade, e.g. grade VI. On the other hand, this "intensity" is not capable of simple quantitative definition, since it describes the earthquake effects in rather qualitative terms. Consequently, an interpretation of the macroseismic grades in terms of some physical quantities, such as the maximum acceleration or others, represents a complicated and problematic task. The problem consist in the fact that the macroseismic intensity depends in a complicated way not only on ground accelerations but also on the periods, duration and other features of the seismic waves. For example, the high-frequency shaking during the earthquakes in Western Bohemia (Fig.1.1) produces strong sound effects, which frighten people, but causes relatively a little damage to buildings, whose resonant frequencies are much lower.

Nevertheless, efforts have been made to associate the macroseismic intensity with accelerations or other parameters of the local ground shaking. These relations have usually been based on the Weber-Fechner psychological law, which-states that the feelings perceived by the human sense organs increase with an arithmetic sequence if the corresponding physical quantity (amplitude, concentration) increases with a geometrical sequence. Cancani assumed that a similar law holds also between the macroseismic intensity and acceleration of the ground motion, i.e. the macroseismic intensity increases with an arithmetic sequence if the acceleration increases with a geometrical sequence. As an illustration, we give the corresponding relation in Tab. 1.3; since buildings are more sensitive to horizontal accelerations, this component of the acceleration is considered there.

Table 1.3. The association of the macroseismic intensity I with the horizontal acceleration a_h in the Mercalli-Cancani-Sieberg scale.

I	$a_h\left(\mathrm{mm/s^2}\right)$
I	$0-2.5$
II	$2.5-5.0$
III	$5-10$
IV	$10-25$
V	$25-50$
VI	$50-100$
VII	$100-250$
VIII	$250-500$
IX	$500-1000$
X	$1000-2500$
XI	$2500-5000$
XII	>5000 (= 0.5 g)

The above-mentioned law can be expressed as

$$\log_{10} a = bI + c \ , \tag{1.1}$$

where a is the acceleration, I the macroseismic intensity, and b, c are constants. The relation in Tab. 1.3 is described approximately by the parameters $b = 1/3$ (three grades of intensity correspond to one order of acceleration) and c between –0.3 and –0.4 (if the boundary between grades I and II, i.e. $a_h = 2.5\,\mathrm{mm/s^2}$, is characterised by the intensity I = 1.5, etc.) Richter (1958) has presented similar values, $b = 1/3$ and $c = -1/2$, but very different values have been obtained by other authors. We can conclude that, despite of many attempts, this effort has not been successful and reliable formulae of type (1.1) have not been found. At present, such formulae are not usually used. For example, in the European Macroseismic Scale 1992, any relations of type (1.1) have been abandoned.

Another important problem is the decrease of macroseismic intensity with distance from the source. First, let us derive several formulae for the energy of harmonic waves. Consider a unite volume of density ρ performing harmonic oscillations:

$$u = A \sin \omega t \ , \tag{1.2}$$

where u is the instantaneous displacement, A the amplitude, ω the angular frequency, and t the time. The instantaneous kinetic energy of the unit element is

$$\frac{1}{2}\rho \dot{u}^2 = \frac{1}{2}\rho\omega^2 A^2 \cos^2 \omega t \ ,$$

and the mean kinetic energy, averaged over a cycle, is

$$\frac{1}{4}\rho\omega^2 A^2 ,$$

since

$$\frac{1}{2\pi}\int_0^{2\pi}\cos^2 x\,\mathrm{d}x = \frac{1}{2\pi}\int_0^{2\pi}\sin^2 x\,\mathrm{d}x = \frac{1}{4\pi}\int_0^{2\pi}\left(\cos^2 x + \sin^2 x\right)\mathrm{d}x = \frac{1}{2} .$$

Assuming the mean kinetic and potential energies to be the same (virial theorem), the mechanical energy of the unit element (the energy density) is

$$\varepsilon = \frac{1}{2}\rho\omega^2 A^2 = 2\pi^2\left(A/T\right)^2 , \tag{1.3}$$

where T is the period. Thus the energy density is proportional to the square of the amplitude, $\varepsilon \sim A^2$.

Consider a point source of spherical harmonic waves in a homogeneous, perfectly elastic medium. Denote by r the distance from the source. Since the flux of energy through any sphere with its centre at the source is the same, we have

$$4\pi r^2 \varepsilon(r) = \text{const.}$$

Consequently, $\varepsilon(r) \sim 1/r^2$, and the amplitude of a spherical harmonic wave decreases with distance as

$$A(r) = \frac{K}{r} , \tag{1.4}$$

K being a constant.

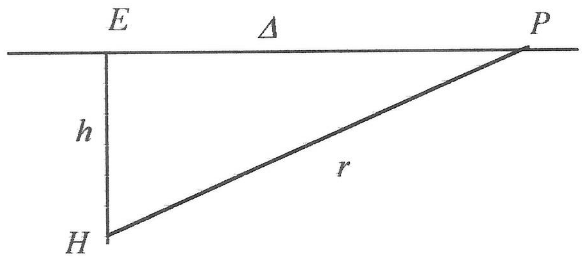


Fig. 1.2. The situation considered in the derivation of Gassmann's formula.

Using the notations given above, the amplitude of the acceleration in a harmonic wave is $\omega^2 A$. For a given frequency, formulae (1.1) and (1.4) yield

$$K^* - \log_{10} r = bI , \tag{1.5}$$

where K^* is a constant; $K^* = \log_{10}(\omega^2 K) - c$.

Let us compare the macroseismic intensity I_0 at the epicentre E with the intensity I_Δ at an epicentral distance Δ, see Fig. 1.2. Denoting the focal depth by h and using Eq. (1.5), we have

$$K^* - \log_{10} h = bI_0 ,$$

$$K^* - \log_{10}(h^2 + \Delta^2)^{1/2} = bI_\Delta .$$

By subtracting these equations and putting $b = 1/3$, we arrive at Gassmann's formula:

$$\frac{2}{3}(I_0 - I_\Delta) = \log_{10}\left(1 + \frac{\Delta^2}{h^2}\right) . \tag{1.6}$$

This formula can be used for estimating the local depth h from macroseismic data. It follows from this formula that, for a given difference of macroseismic intensities, a larger distance Δ between the isoseismals indicates a greater focal depth h. In other words, dense (resp. sparse) isoseismals indicate a shallow (resp. deep) earthquake focus. This qualitative conclusion follows, of course, also from simple geometrical considerations.

As a certain generalisation, a number of formulae of the type

$$a(I_0 - I_\Delta) = \log_{10}\left(1 + \frac{\Delta^2}{h^2}\right) \tag{1.7}$$

have been constructed for various regions, where a and h are constants to be fit (h may be interpreted as the focal depth). Note that all these formulae assume predominantly circular isoseismals. Since relations (1.7) are based on the problematic relation (1.1), all these formulae should be used with caution.

It is evident that the determination of the focal depth (and also the other parameters of the source) should be preferably based on instrumental observations. However, it happens very often that instrumental observations are not available from a close vicinity of the epicentre, and then formulae (1.7) should be used to obtain a first, rough estimate of the focal depth.

Some of the formulae in this section are now rather of a historical value. Nevertheless, we have paid them certain attention since similar logarithmic relations became very popular in studies of earthquakes, and many of them are still in use. We shall meet several of them below in this chapter.

1.4. Seismometers and seismic stations

The simplest instruments for recording earthquakes are called seismoscopes. They register earthquakes, but not the time dependence of the ground movement. The first seismoscopes had already been constructed in ancient China; Zhang-Heng, AD 132.

Seismographs record the ground movement at a particular point of the Earth's surface during an earthquake, as a continuous function of time.

First actually working seismographs were constructed by J. Milne in Japan around 1890. The seismographs are usually pendulums with some damping. The first seismographs were usually constructed as horizontal reverse pendulums with mechanical or optical registration. Later on, Golitsyn introduced electromagnetic seismographs, i.e. seismographs with electromagnetic registration. An important part of all seismographs is a good timing.

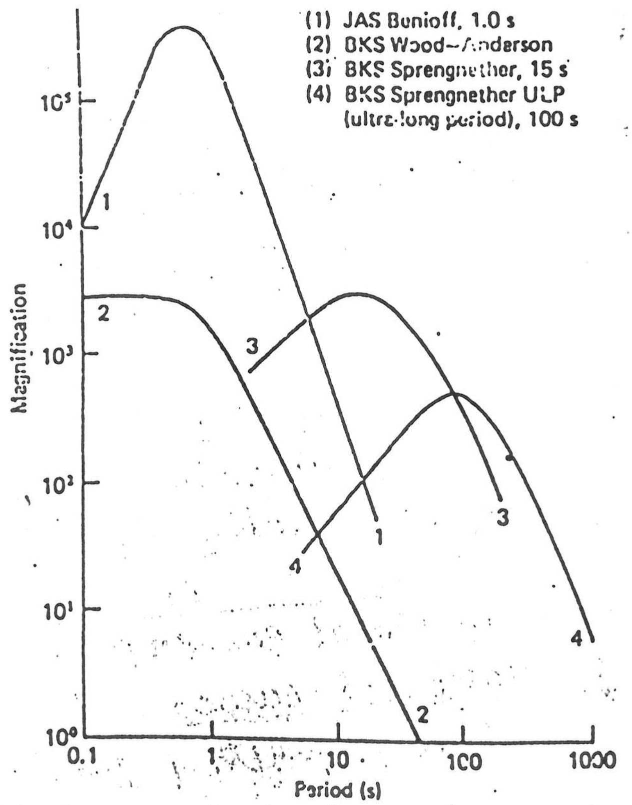


Fig.1.3. Magnification curves for the seismographs operated at Jamestown and Berkeley, California: 1) Benioff, 1.0 s; 2) Wood-Anderson; 3) Sprengnether, 15 s; 4) Sprengnether (ultra-long period), 100 s. Times refer to the resonant periods of the pendulums. Taken from Bullen and Bolt (1993).

The properties of the seismograph are usually characterised by the magnification curve of the seismograph: This curve shows the magnification of the harmonic ground motion versus period. The maximums of the magnification curve are situated at resonant periods. According to the character of these curves, the following types of seismographs are usually distinguished:

- short-period seismographs (resonant period close to 1 s);
- long-period seismographs (resonant period greater than 20 s);
- ultra-long-period seismographs (resonant period greater than 100 s); see Fig. 1.3.

Actually, seismologists are interested even in a broader range of periods, from T = 0.01 s (frequency f = 100 Hz) for local earthquakes to T = 3300 s (f = 0.0003 Hz) for free oscillations of the Earth. They also wish to record amplitudes in a range of 120 decibel ($10^6:1$).

The new developments in seismometry, from about 1970, are broad-band seismometers (broad ranges of frequencies and amplitudes). They are characterised by modern freed-back servosystems, usually with digital recording, and by a flat magnification curve.

Also quite different types of seismometers exist, such as Benioff strain meters or tiltmeters.

For a reliable determination of the position and further parameters of an earthquake, observations of one seismic stations are usually insufficient. Therefore, systems of seismic stations are commonly used. Let us mention some of these systems:

a) World-wide systems. The world system of seismic stations (~2000 stations) is not homogeneous (different instruments, different quality, non-homogeneous distribution of stations). Examples of world-wide systems are as follows:
   - WWSSW (World Wide Standard Seismic Network): Homogeneous instrumentation, short-period and long-period instruments, but non-digital registration (photographic registration). The system contains about 200 stations.
   - World wide digital systems, which have been built recently, e.g, GDSN (Global Digital Seismic Network), and several others.

b) Local. systems. They serve for the following purposes:
   - To record weak distant earthquakes (e.g., nuclear explosions).
   - To study local seismicity. Some local array are actually very small, e.g. in mines to follow rockbursts, etc.

Standard seismic stations determine, for each recorded wave, the following three parameters: arrival time, maximum ground amplitude, prevailing period. They send the data to international seismic centres.

Some preliminary processing of these data may be done at the station (preliminary determination of the epicentral distance from the difference between the arrival times of transverse and longitudinal waves, preliminary magnitude). The definitive processing is done at international centres.

International seismic centres perform:

a) localisation of earthquakes, including the determination of the hypocentral times;

b) determination of the magnitudes of earthquakes, or of some other instrumental quantities which characterise the strength of earthquakes.

The recent broad-band world-wide digital systems yield considerably more useful data, namely complete digital seismograms. Such complete seismograms are sent on magnetic tapes or telemetred to seismic centres, where they are processed in a considerably more sophisticated way.

1.5. Earthquake magnitude

The magnitude of an earthquake is an instrumental measure of the strength of the earthquake. The magnitude was first used by Richter, so that the earthquake magnitude scale is also popularly called the Richter scale.

The magnitude is mostly determined from the maximum ground motions (amplitudes) of individual waves using the relation

$$\boxed{M = \log\frac{A}{T} + \sigma(\Delta,h) + \sum_{(i)} c_i}, \tag{1.8}$$

where A is the maximum ground amplitude in micrometers, T the corresponding period, Δ the epicentral distance, h the focal depth, $\sigma(\Delta,h)$ an empirical calibration function (calibration curve), c_i are various corrections, mainly a station correction. The logarithm in the formula is the ordinary logarithm to the base 10, i.e. $\log(A/T) = \log_{10}(A/T)$. The same logarithms are used in the rest of this chapter.

The calibration curve $\sigma(\Delta,h)$ reduces the amplitudes observed at the epicentral distance Δ to some reference epicentral distance (according to the original definition it was the epicentral distance of 100 km). The intention is to obtain the same magnitude from stations at different epicentral distances. The magnitude is thus a number characteristic of the earthquake and independent of the location of the recording station. From this point of view, the magnitude differs substantially from the macroseismic intensity, which depends on the place of observation (the macroseismic intensity characterises the earthquake effects at the given place).

To determine $\sigma(\Delta,h)$, the amplitude-distance curves $A = A(\Delta)$ must be first found. The calibrations curves have an international character, they are approved and recommended by international organisations (e.g. by the International Association of Seismology and Physics of the Earth Interior, IASPEI).

Several types of magnitude have been set up. The most important of them are as follows:

M_S surface wave magnitude (determined from seismic waves of periods ~ 20 sec);

m_b body wave magnitude (periods ~ 5 - 10 sec);

M_L local magnitude.

The local magnitude was originally set up by Richter. It is based on the records of the torsion Anderson-Wood seismometers with the resonant period $T_0 = 0.8$ s and static magnification $V = 2800$. Thus, the local magnitude M_L is based on considerably higher frequencies ($T_0 = 0.8$ s) than the other two magnitudes. Its numerical value is rather close to the value of the macroseismic intensity I_0.

If we speak on magnitude M, without specifying the type, we usually mean $M = M_S$. Typical values of seismic magnitudes are as follows:

the highest observed magnitudes	:	$M \sim 9$;
earthquake catastrophes	:	$M > 8$;
weak earthquakes	:	$M < 4$;
microearthquakes	:	$M < 3$.

In case of very small earthquakes (rockbursts in coal mines), we may even have $M < 0$.

An approximate empirical formula for the relation between the magnitude M and the epicentral macroseismic intensity I_0 has the form

$$M \sim \frac{2}{3} I_0 + 1.2 \log h - 1.1 \ , \tag{1.9}$$

where h is the focal depth in kilometres. Thus, for $h \sim 10$ km, we have roughly $M = \frac{2}{3} I_0$.

At the end of this section, let us add several historical remarks. Richter set up the local magnitude in 1935 for earthquakes in southern California. He originally defined the magnitude by the relation

$$M = \log \frac{A(\Delta)}{A_0(\Delta)} = \log A(\Delta) - \log A_0(\Delta) \ , \tag{1.10}$$

where A is the maximum trace amplitude for a given earthquake at a given distance Δ as recorded by the standard Anderson-Wood instrument (the parameters of which are given above), and A_0 is that for a particular earthquake selected as standard. The reference level A_0 was taken to be one micrometer at a distance of 100 km. This specification means that, for example, an earthquake recording with trace amplitude of 1 mm measured on a standard seismogram at 100 km, is assigned the magnitude $M = 3$.

The passage from the original formula (1.10) to the contemporary formula (1.8) went through several steps. First, a calibration function $\sigma = -\log A_0$ as a function of epicentral distance Δ was established on the basis of observed amplitude-distance curves. Then, in order to unify measurements carried out by different instruments, the amplitudes on standard seismograms were replaced by the amplitudes of the ground motion. Finally, the amplitude A was replaced by the ratio A/T. The quantity A/T has the advantage of being simply related, in theory at least, to the mechanical energy of its wave group; see formula (1.3).

Let us also mention some terminological and mathematical motivations, leading to the introduction of the earthquake magnitude, as described by Richter (1958): "The term magnitude was selected by analogy with the corresponding usage in astronomy. The scale of star magnitudes is also logarithmic, though on a less simple basis; in a sense it is reversed, since the greater the magnitude the fainter the star. The earthquake magnitude scale follows the more obvious course of assigning the larger number to the larger earthquake. Logarithmic scales are in use in other fields; examples are the decibel scale in acoustics and the *p*H scale for hydrogen-ion concentration". As we have seen in Sec. 1.3, even in seismology there were previous attempts to introduce a logarithmic relation, namely a relation between macroseismic intensity and acceleration.

1.6. Other instrumental measures of the strength of earthquakes

The magnitude, being rather a simple and rough measure of the strength of an earthquake, is not sufficient for some purposes. Therefore, also other instrumental measures of the strength of earthquakes have been introduced. We shall describe here two of them.

1.6.1. Seismic moment

The seismic moment, M_0, is an analogue of the moment of a force (torque) in rigid body mechanics. For a simple fault source it can be expressed as

$$M_0 = \mu \Delta u \sum \ , \tag{1.11}$$

where μ is the rigidity of the medium, Δu is the average rupture displacement along the rupture plane, and Σ is the area of the rupture plane. The dimension of the seismic moment is newton.meter, $[M_0]$ = N.m. The maximum observed values of M_0 are around 10^{23} N.m.

In some cases, the seismic moment can be estimated simply by substituting field measurements into (1.11). For example, for the 1906 San Francisco earthquake, putting $\mu = 4 \times 10^{10}\ \mathrm{N/m^2}$, rupture length L = 400 km, rupture depth h = 10 km, fault offset Δu = 5 m, we get the seismic moment of about 8×10^{20} N.m.

However, the seismic moment M_0 is more frequently determined from the Fourier amplitude spectrum S of the complete seismogram. A typical shape of such a spectrum is shown schematically in Fig. 1.4. The frequency corresponding to the intersection of the horizontal low-frequency asymptote and the sloping high-frequency asymptote is called the corner frequency and is generally denoted by f_0. For $f > f_0$, the spectrum usually decays as f^{-2}, for $f < f_0$ it is flat. Thus, the displacement spectrum near the source may be described approximately by three basic parameters: the spectral amplitude at zero frequency, the corner frequency and the spectral slope at frequencies above the corner frequency (on a log–log plot). It can be shown that each of these parameters is a measure of basic properties of the source, namely the source power, fault length and source time function, respectively.

The seismic moment M_0 can be determined from the spectrum for $f < f_0$. For example, in the case of *SH* waves, if the source radiation pattern is not considered, we can use the simple formula

$$M_0 = 4\pi\rho\beta^3 r S_0 \ ,$$

where ρ is the density of the medium, β the shear wave velocity, r the distance of the receiver from the source, and S_0 the low-frequency asymptote of the amplitude spectrum.

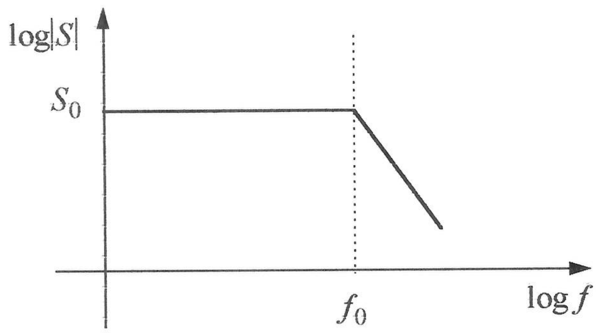


Fig. 1.4. A schematic form of the amplitude spectrum of the ground motion displacement: f is the frequency, $|S|$ the Fourier amplitude spectrum.

The advantage of the seismic moment over the magnitude consists in the fact that the magnitude is saturated at $M \sim 8-8.6$, i.e. the magnitude does not distinguish well between big earthquakes. As an example, let us mention the following two earthquakes:

- California, 18.4.1906, $M_S = 8.3$, $M_0 = 10^{21}$ Nm;
- Chile, 22.5.1960, $M_S = 8.3$, $M_0 = 2.4 \times 10^{23}$ Nm.

Thus, in terms of the seismic moment, the earthquake in Chile was about 240 times greater than the earthquake in California, but in terms of the magnitude they were the same.

Recently, also the so-called moment magnitude M_w has been introduced. It is evaluated from M_0 by the relation

$$M_w = 0.67 \log M_0 - 6.03 , \qquad (1.12)$$

where M_0 is given in N.m. For the above example, we would obtain $M_w = 8.0$ for California and $M_w = 9.6$ for Chile.

The advantage of M over M_0 is that its determination is considerably simpler; it does not require computation of spectra. Many station are equipped only with non-digital seismographs, and the computations of spectra requires digitalisation of records.

1.6.2. Seismic energy

The seismic energy, E_S, is the energy released during the earthquake in the form of seismic waves. It does not represent the total energy released by the earthquake, as a part of the energy is also released in the form of heat, etc. The seismic energy is probably less than 50 per cent of the total energy released by the earthquake (perhaps $\sim 1/3$).

If we wish to be more strict, we must define some reference sphere with its centre at the hypocenter and calculate the energy flux through this sphere. If we take a sphere with a different radius, the complete energy flux may be different due to absorption. Riznichenko introduced the radius of the reference sphere of 100 km.

Determination of the seismic energy from seismic records is very complicated, and many simplifying assumptions must be made. In spite of that, E_S has been determined for many earthquakes. An approximate empirical relation between E_S and M has the form

$$\log E_S = 5.24 + 1.44 M \ , \qquad (1.13)$$

where E_S is in joules. A unit increase in M thus corresponds to a 28-fold increase in energy. For large earthquakes, $M \sim 8.5$, we obtain $E_S \sim 10^{17} - 10^{18}$ J.

As an example, let us derive here a very simple formula for determining the seismic energy from seismic records. Consider a receiver inside a homogeneous medium at a distance r from a point source which radiates seismic waves symmetrically in all directions (this assumption of spherical symmetry about the focus is of course not realised with actual earthquakes). The energy flux through a sphere with its centre at the source and passing through the receiver is $Sc\varepsilon$, where $S = 4\pi r^2$ is the area of the sphere, c is the velocity of seismic waves, and ε is the energy in the unit volume, given by (1.3). The total seismic energy E_S is then obtained by the integration over time (time duration of the oscillations on the seismogram):

$$E_S = 8\pi^3 r^2 \rho \int c A^2 T^{-2} \, dt \ .$$

In particular, for observations at the epicentre we must put $r = h$, where h is the focal depth. Formulae of this type formed the basis for assessing energy in earthquakes. It should be noted that A denotes the amplitude in the incident waves, but it is assumed that this amplitude is of the same order as that of the observed ground motion. (Since the motion on the Earth's surface is the superposition of incident and reflected waves, in a more exact approach the observed amplitudes should be recalculated to obtain the amplitudes of the incident waves, see Chap. 5).

Assuming the maximum value of A/T to be a characteristic of the whole seismogram, and all seismograms to have a comparable time duration, it follows from the above-mentioned relation that $\log E_S \sim 2\log(A/T)_{\max} \sim 2M$. Consequently, we get approximately

$$\log E_S = c_1 + c_2 M \ ,$$

where $c_2 \sim 2$. The more accurate value of this coefficient in (1.13) is lower.

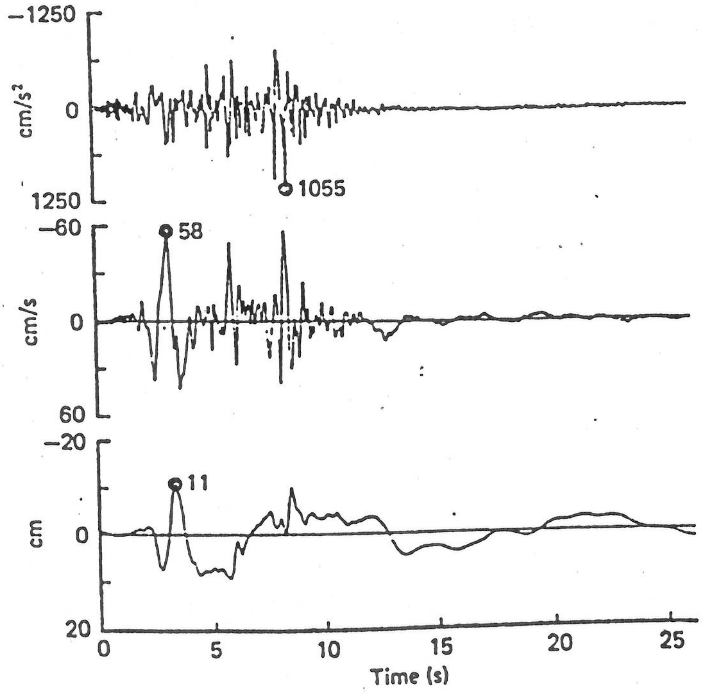


Fig. 1.3. Ground motions recorded on the abutment of Pacoima Dam in the 1971 San Fernando, California earthquake From top: acceleration, velocity, displacement. (Taken from Bullen and Bolt (1993)).

1.7. Seismic instrumental observations in epicentral regions

There are several types of seismic system for observations in epicentral regions:

a) Systems with standard short-period seismometers (see Sec. 1.4). They are used to fast localisation of earthquakes, determination of magnitudes, etc.
b) Systems of accelerometers (strong motion seismology). Their characteristics are as follows: waiting regime, high frequencies, small magnification. Such systems make it possible to investigate details of rupture processes (asperities, barriers, etc.). They are broadly used in seismic engineering. Note that the maxima of the acceleration may reach $10\,\mathrm{m/s^2}$, the velocity about 1 m/s, the displacement several meters. These values may be even higher for catastrophic earthquakes. A typical record of an accelerometer is shown in Fig. 1.5; the velocity record and the displacement record are obtained by integration.

1.8. Seismicity

The word "seismicity" is used in different meanings. Usually, seismicity is understood to mean seismic activity. It can be related to the whole Earth, or to a selected region.

Seismic activity of the world (or of a selected region) is described by the group of the following five parameters, corresponding to individual earthquakes (five parameters for one earthquake):

φ, λ, h	position of the hypocentre,
H	hypocentral time,
M or I_0 or M_0 or E_S	some parameter measuring the intensity of the earthquake.

Thus, the seismic activity is a 5-dimensional function. Several of its projections are described in the following subsections.

1.8.1. Maps of epicentres

The map of epicentres shows the projection of earthquake hypocentres onto the (φ, λ)-plane. These maps may be related to a selected range of depth, selected range of magnitudes, and selected time interval.

A world map of epicentres is shown in Fig. 1.6 for the time interval 1961-1967, depths of 0-100 km, and magnitudes $M > 4$. The figure shows that the great majority of earthquakes are concentrated in several belts:

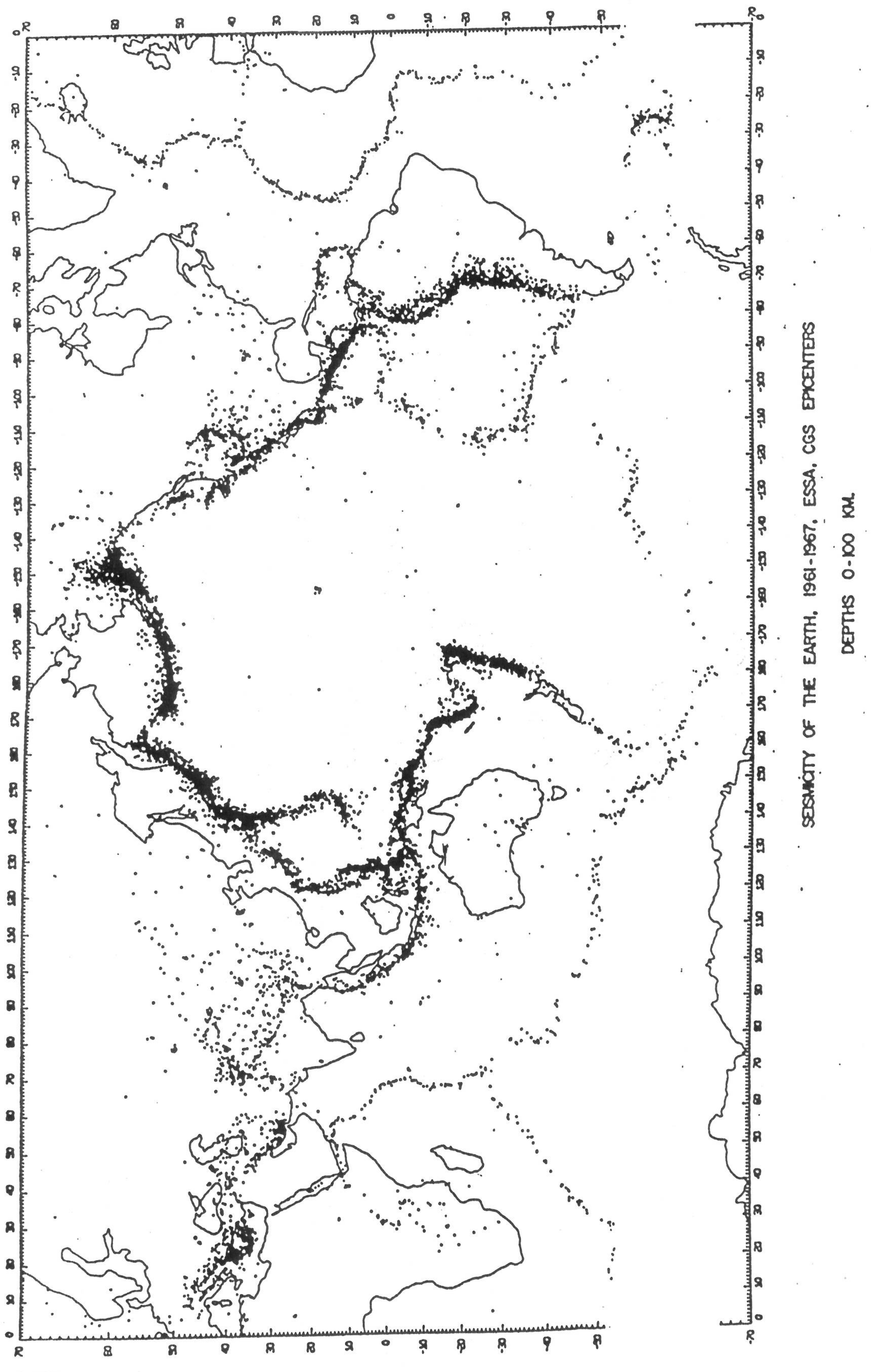


Fig. 1.6. Geographical distribution of earthquakes.

- The circum-Pacific belt, including the regions of New Zealand, New Guinea, Philippines, Japan, the Kurile Islands, Kamchatka, the Alentian Islands, Alaska, and the western regions of North and South America. About 80 per cent of the total energy released in earthquakes comes from earthquakes in this belt.
- The Mediterranean-Himalayan belt, passing from the region of Portugal and north Africa through Italy, Greece, Turkey, Central Asia, the Himalayan region to Indonesia, where it joins the first belt. The energy released in earthquakes from this belt is about 15 per cent of the total.
- Other belts, mainly along mid-oceanic ridges, which clearly indicate boundaries of lithospheric plates.

The maps of epicentres represent a basic material for plate tectonics. Such maps are also of primary importance for earthquake engineering.

1.8.2. Depth distribution of earthquakes

The data on the depth distribution of hypocentres represent a basic material for place tectonics. The position of hypocentres of deep earthquakes indicates the subduction of lithospheric plates.

According to the focal depth, h, earthquakes are traditionally divided as follows:

1) Shallow earthquakes $(h < 60\,\text{km})$. The great majority of earthquakes originate within this depth interval. About 85 per cent of the total energy released in earthquakes comes from these earthquakes. All the strongest earthquakes belong to this category.
2) Intermediate earthquakes $(h \sim 60 - 300\,\text{km})$. About 12 per cent of earthquake energy is released in these earthquakes.
3) Deep-focus earthquakes $(h \sim 300 - 700\,\text{km})$. These earthquakes commonly occur in the so-called Benioff zones, that dip into the Earth. Such zones are found in the regions of Japan, Vanuatu (the New Hebrides), the Tonga Islands, Alaska, along the South American Andes, and in some other regions. The greatest focal depths are of about 700 km (a more accurate estimation gives $h \sim 680\,\text{km}$). At greater depths, no earthquakes have been observed. Only about 3 per cent of the total earthquake energy comes from the deep-focus earthquakes.

1.8.3. Time sequences of earthquakes

The earthquakes are not usually isolated in time and space. If we consider a specified locality, we can follow the time sequences of earthquakes. The standard time sequence is as follows: foreshocks, main shock, aftershocks.

There are also different time sequences. For example, the time sequence of volcanic earthquakes usually does not contain a dominant main shock. We then speak on earthquake swarms. Similar earthquake swarm may be sometimes observed even in non-volcanic regions.

1.8.4. Magnitude - frequency relations

These relations characterise the time frequency of earthquakes as functions of magnitude. The corresponding empirical relations are usually assumed to have the form

$$\log N = a - bM\,, \tag{1.14}$$

where N is the number of earthquakes for some interval around the magnitude M in a given region (or for the whole Earth), and for a selected time interval (e.g. one year). Particularly the quantity b is of a great importance in seismology (it is also used for purposes of earthquake prediction).

The relation (1.14) for the whole Earth, the time interval equal to one year, and the magnitudes in the range $\langle M - 0.5, M + 0.5 \rangle$, has the form

$$\log N = 8.73 - 1.15M\,. \tag{1.15}$$

It can be seen from this relation that the number of earthquakes increases more than 10 times when the magnitude is decreased by one.

In general, the number of earthquakes to be felt $(m_b > 4)$ is about 20 000 on the whole Earth within one year. The number of earthquakes with $M > 8$ is one or two within one year.

In addition to N, it is also interesting to study the quantity

$$\varepsilon = NE_S\,, \tag{1.16}$$

which represents the seismic energy released in the region under consideration (or on the whole Earth) in a selected time interval and magnitude range. Using (1.13) and (1.15), we obtain

$$\log \varepsilon = 13.97 + 0.29M \tag{1.17}$$

for the whole Earth, time interval equal of one year, and $\langle M - 0.5, M + 0.5 \rangle$.

It follows from formula (1.17), since the coefficient with M is positive, that the most energy is released mainly in several big earthquakes. With decreasing M, energy ε decreases.

The total seismic energy released within one year on the whole Earth is about 5×10^{17} J.

1.8.5. Migration of earthquakes

The seismically active regions often slowly move from a place to a place. We speak on a migration of earthquakes. (This term does not have anything common with the migration in seismic prospecting!).

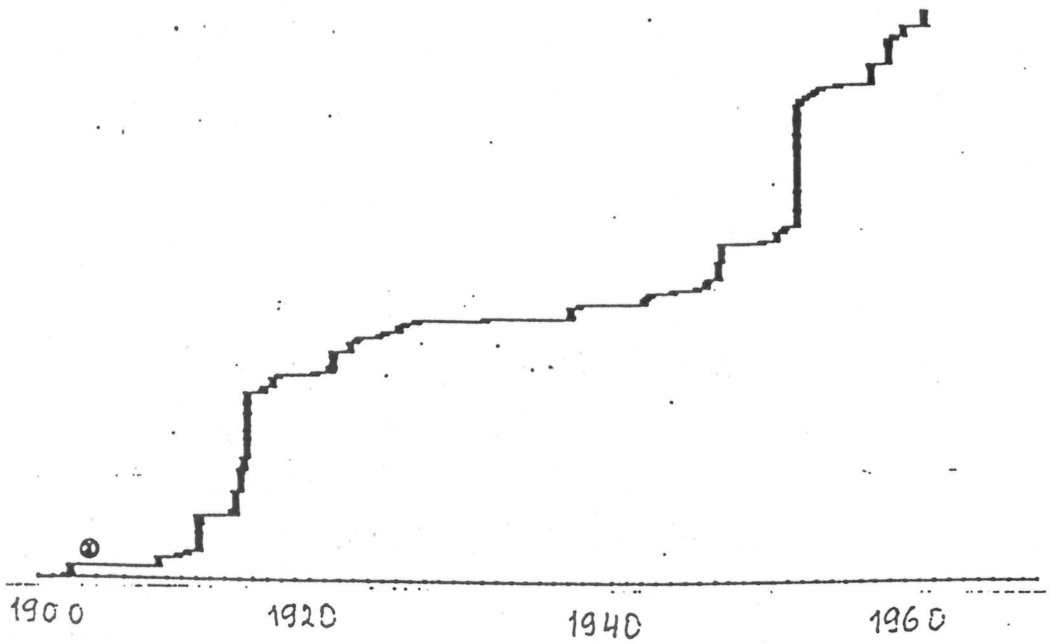


Fig. 1.7. Cumulative graph (Benioff curve) for one region of the Mediterranean Sea (The Ionic Islands). The regions of higher seismic activity are clearly visible.

1.8.6. Cumulative graphs of seismic energy

We can sum up a power of the seismic energy released in individual earthquakes, $\sum_{(i)} (E_S)^n$, as a function of time in some region or on the whole Earth. Most commonly, we consider $\sum_{(i)} \sqrt{E_S}$, i.e. $n = 1/2$. Such cumulative graphs are known also as the Benioff curves. They show the periods of higher seismic activity and seismically quiet periods. They are influenced mainly by largest earthquakes. An example of the cumulative graphs is shown in Fig. 1.7.

The square of energy, usually used in the cumulative graphs $(n = 1/2)$, can be substantiated by the following considerations. It is known from continuum mechanics that the density of elastic energy, accumulated in a deformed elastic medium, can be expressed as

$$\varepsilon = \frac{1}{2}\tau_{ij}e_{ij} = \frac{1}{2}c_{ijkl}e_{kl}e_{ij} \;;$$

here e_{ij} is the strain tensor, τ_{ij} is the stress tensor, c_{ijkl} the elastic coefficients, and $\tau_{ij} = c_{ijkl}e_{kl}$ according to the generalised Hooke's law (Chap. 2). Thus, the elastic energy is a quadratic function of deformation, so that $\sqrt{E_S}$ is a measure of deformation. Further, since it is assumed that tectonic processes lead to a gradual increase of deformation, it is natural to study the cumulative value of $\sqrt{E_S}$ as a function of time.

1.8.7. Seismic gaps

By a seismic gap we call a region of temporarily decreased seismic activity, with a strong seismic activity in its vicinity. The seismic gap is often a place where the movement of the lithosperic plate remains fixed locally for some time, whereas the plate moves in the vicinity. Assuming a uniformity of crustal displacements, it indicates an increasing stress (cumulating of stress). It is a very dangerous region; the seismic gap may be a place of a future great earthquake.

1.8.8. Local effects. Seismic microzoning

The same earthquakes has often strong effects at one place, and small effects at some other close place. The differences are mainly due to topographic, geological and soil conditions close to the Earth's surface. Very dangerous are, for example, sedimentary basins. A great role is played by resonance effects in the uppermost layers. It depends strongly on the prevailing frequencies of the movements.

The mapping of normalised expected seismic effects (assuming the same incident seismic wave) is called seismic microzoning. Microzoning maps have a great importance in seismic engineering.

1.8.9. Seismic risk

In a simplified way, we can say that the seismic risk is the probability that the earthquake effects at a given locality will exceed some critical value within a given time interval (e.g. within 10 years, 50 years, etc.).. It is investigated using probability methods. It has as great importance in seismic engineering. As an example, we give a map of seismic risk in Fig. 1.8.

A similar quantity is the seismic hazard, which describes the probable economic loss (expressed in money) caused by earthquakes at a given locality within a given time interval. The seismic hazard depends, besides the seismic risk, also on population density and other factors. Methods of seismic hazard assessment belong also to the main problems of seismic engineering.

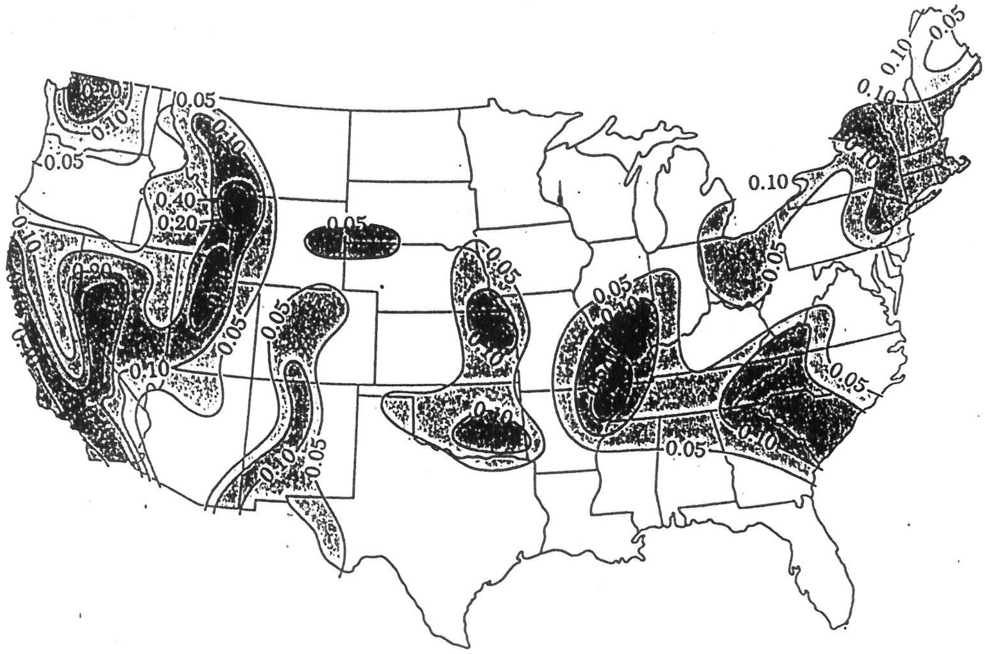


Fig. 1.8. A seismic-risk map for the Union States, prepared in 1976 to 1977. The contours indicate effective peak, or maximum, acceleration levels (values are in decimal fractions of gravity) that might be expected to be exceeded during a 50-year period. (From Bullen and Bolt (1993)).

1.8.10. Induced seismicity

Earthquakes can also be caused by human activity. Such earthquakes may be connected with:

- Excavation of mines, which causes rockbursts. They are, in fact, small tectonic earthquakes.
- Filling of large water reservoirs, dams, etc. (such as Aswan in Egypt).
- Injection of fluids in deep wells.
- Large underground (mainly nuclear) explosions.

The induction mechanism is as follows. The regional strain release is triggered by small changes in the local stress field (caused by human activity). This leads to a fracture or to a fault slip.

1.9. Tsunamis

After certain earthquakes under the sea, very long water waves are sometimes generated. They are called tsunamis (from Japanese). On the open sea, their amplitudes do not exceed several meters (usually they do not exceed 1 m), their wavelength is of the order of hundreds of kilometres and the velocity of several hundreds of kilometres per hour. They may be, however, very dangerous close to the shore, particularly in bays of *U* or *V* shape. They can reach even a height of about 20-30 m and destroy all structures. In the Pacific Ocean, a warning system is organised (SSWWS - Seismic Sea Wave Warning System, with its centre at Honolulu).

1.10. Extraterrestrial seismology

Seismometers have already been installed also on two celestial bodies, namely on the Moon and Mars.

Moon (6 sites, 1969). The seismometers detected 600 - 3 000 moonquakes per year. Three types of moonquakes have been recognised:

- Impacts of meteorites and other bodies.
- Shallow moonquakes $(h < 100\,\mathrm{km})$.
- Deep moonquakes $(h \sim 800 - 1000\,\mathrm{km})$. The cause of these moonquakes is not known.

The main differences of the lunar seismograms, in comparison with the terrestrial seismograms, are higher frequencies and longer duration of the records (often over one hour); see Fig.1.9. It indicates a greater scattering and lower attenuation of seismic waves in the Moon (it may be attributed to the absence of water on the Moon).

Mars (2 seismometers, 1976). The seismograms were complicated due to strong winds and oscillations, but at least one event is considered to be a marsquake.

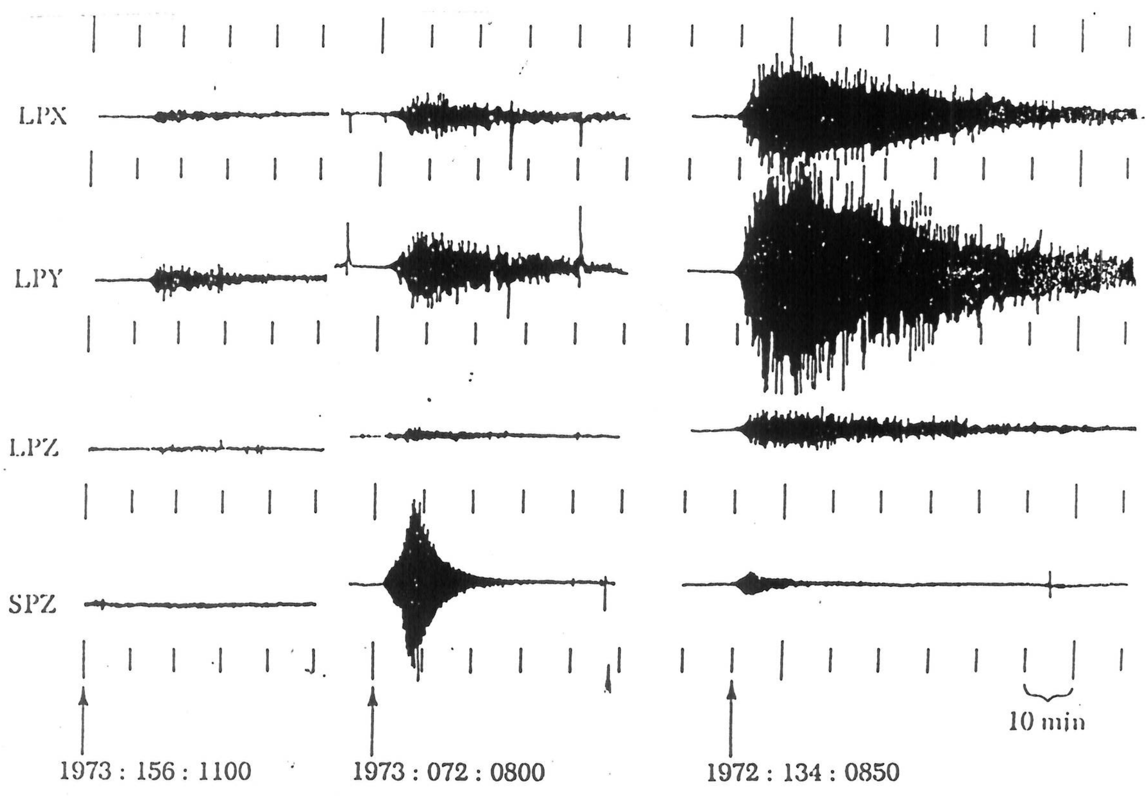


Fig. 1.9. Seismogram from three types of moonquakes recorded at the Apollo 16 station. LPX, LPY, and LPZ are the three long-period components, and SPZ is the short-period vertical component. The first column shows a deep-focus moonquake; the center column, a shallow moonquake; the third column shows records of the impact of a meteoroid on the lunar surface. (From Bolt (1988)).

1.11. Prediction of earthquakes

Earthquakes are a very complex phenomenon and their prediction is rather complicated. Under the prediction, we normally understand the prediction of: a) the position, b) the time, c) the magnitude; of future earthquakes. The most complicated is the prediction of the time of an earthquake. A very weak external trigger may initiate a rupture, as soon as the stress conditions are favourable for a generation of an earthquake.

Many possible forecasting symptoms have been investigated in detail. Still, however, no clear prediction possibilities have been found. Seismology studies primarily the time changes of the following parameters:

- seismic regime of the locality under consideration (number and intensity of weak earthquakes as a function of time);
- seismic velocities, particularly the changes of the v_P/v_S ratio;
- strain and stress fields;
- ground water level (or intensity of springs);
- chemical composition of the underground water (particularly the concentration of radon);
- electric conductivity and other electromagnetic quantities;
- moreover, anomalous behaviour of animals is also studied.

Several earthquakes were actually predicted, e.g., in Haicheng, China, M = 7.3, on February 4, 1975. The city of Haicheng was avacuated several days before the earthquake, and then it was destroyed by the earthquake. This is, however, more or less an exception, since many other predictions were not successful.

The problem of the prediction of earthquakes remains open, probably for a long time.

1.12. Mechanism of an earthquake

1.12.1. Elastic rebound theory

Reid studied the San Andreas fault before and after the 1906 San Francisco earthquake. He concluded that the earthquake motion is due to waves radiated from a spontaneous slippage on active geological faults. His explanation of the earthquake is now called the elastic rebound theory of earthquakes. According to this theory, the tectonic earthquake occurs in such a region inside the Earth where the stresses have accumulated to the point exceeding the strength of the material.

1.12.2. Double couple model

It is now common to infer the character of faulting in an earthquake from observed distributions of the polarities of the first onsets of longitudinal waves

(P waves), arriving at the Earth surface. The first onsets of P waves are called anaseismic (compressional) and denoted by "+" if the direction is away from the epicentre. In the opposite case, when the first arrival is towards the epicentre, they are called kataseismic (dilatational) and denoted by "–". In the definition, we always consider the component of the displacement vector into the line connecting the epicentre and the point of observation (Fig. 1.10).

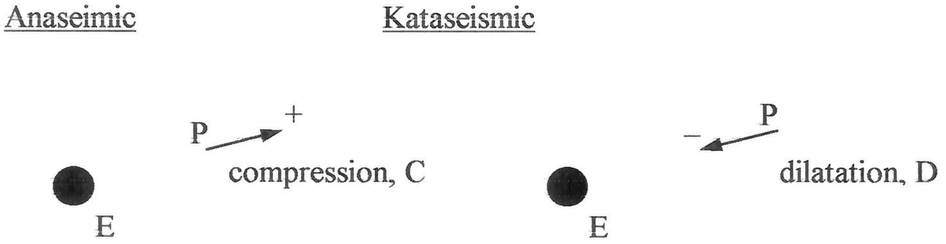


Fig. 1.10. Polarities of the first onsets of longitudinal waves: E is the epicentre, P is the point of observation.

The anaseisms and kataseisms can be plotted into a map. For a vertical fault and a homogeneous Earth, the distribution of the signs may be as shown in Fig. 1.11. This is a typical quadrant distribution of anaseisms and kataseisms. We may separate the regions of "+" and "–" by lines, which are called nodal lines.

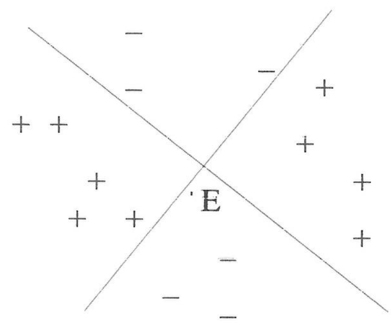


Fig. 1.11. Quadrant distribution of anaseisms and kataseisms for an earthquake on a vertical fault.

Such a polarity pattern can be explained by a point source represented by a system of forces at the source. In our case, the best representation is obtained by two couples of equal and opposite moment, and the forces of one couple being at right angles to the other couple. The forces of the couples are oriented along the nodal lines (Fig. 1.12). We speak on a double couple model.

If we assume that the fault is vertical and the medium is homogeneous, one nodal plane then contains the fault. However, it is not possible to determine, from P wave observations only, which nodal curve of the two contains the fault.

An alternative system of forces which would explain the observations of compressions (+) and dilatations (–) is shown in Fig. 1.13. Here P denotes pressure, T tension. The hatched quadrants correspond to tension, the remaining, not hatched quadrants denote pressure. Such a simple system of forces has a very appealing geological interpretation and is often used to study the distribution of tensions and pressures within seismically active regions using the fault plane solutions for earthquakes.

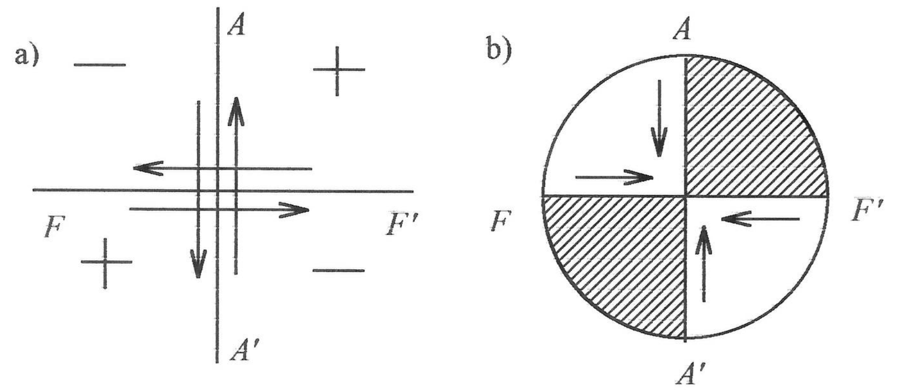


Fig. 1.12. a) Plan view of horizontal displacement on vertical fault $A-A'$ or $F-F'$ and resulting distribution of compressions (+) and dilatations (–); b) corresponding fault-plane diagram. (According to Bullen and Bolt (1993)).

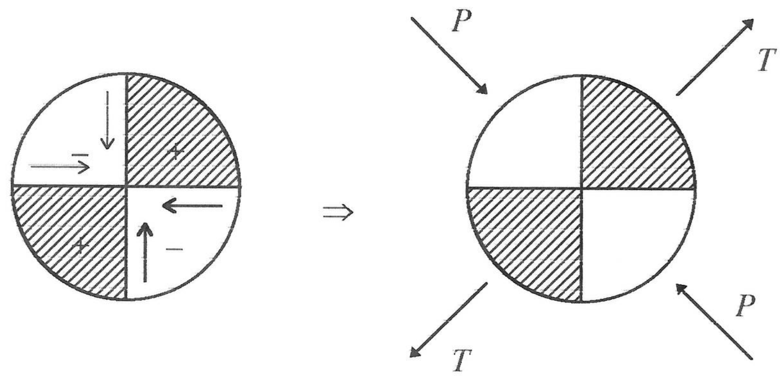


Fig. 1.13. An alternative explanation of the quadrant distribution of P wave onsets.

1.12.3. Source mechanism estimation. Fault plane solutions

The faults within the Earth, however, are not in general vertical. Even more, the Earth is not homogeneous, and the rays, along which the energy of P waves propagates, are not straight lines. For this reason, the situation is usually more complicated than shown above.

The "+" and "–" observations and the nodal lines are usually projected on the so-called focal sphere (Fig. 1.14). This sphere has its centre at the hypocentre H, and the radius equal to unity. Usually only one half of the focal sphere (the lower hemisphere) is considered. The stations observing "+" or "–" are projected upon the focal sphere by ray tracing. The focal sphere, with the observed "+" and "–", is then processed and the nodal lines are constructed. After this, the focal sphere is again projected onto a horizontal plane by some sort of stereographic or other projection. The nodal lines are no more straight lines in this case, and the angle between them is not the right angle. The results may be, for example, as shown in Fig. 1.15. The hatched areas again show the compression (tension *T*) and the non-hatched the dilatation (pressure *P*).

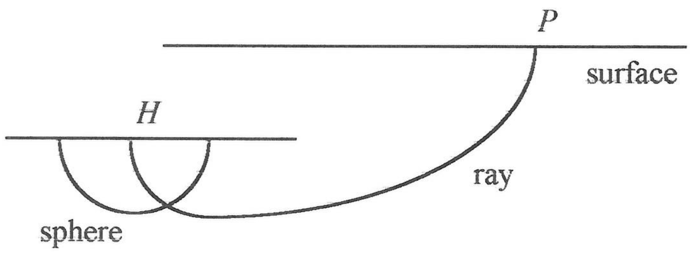


Fig. 1.14. Projections of the signs of the onsets of seismic waves onto the unit sphere.

Such fault plane solutions are then plotted into maps at relevant epicentres (or, at least, they are in some way connected with the epicentres); see an example in Fig. 1.16. In general, such maps show focal mechanisms in some region. The focal mechanisms are not distributed randomly, but usually very systematically. They indicate the stress conditions in the given region.

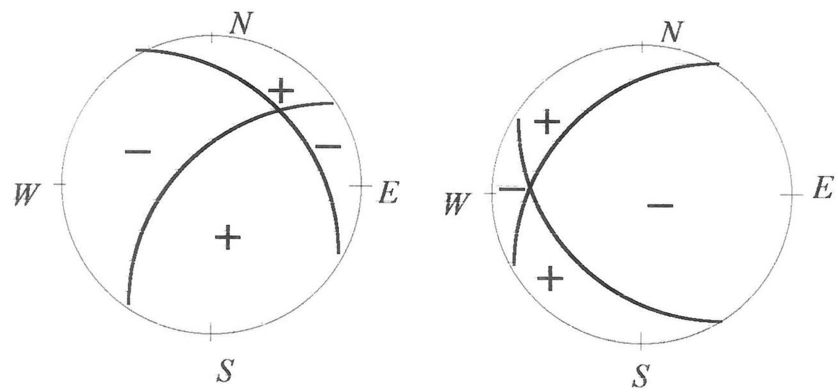


Fig. 1.15. Fault plane solutions for tectonic earthquakes.

1.12.4. Other source parameters and their estimation

From seismic measurements, it is also possible to get other information about the source, partly theoretically and partly empirically.

In the preceding sections, we have learned how to determine the magnitude M, seismic moment M_0, seismic energy E_S and the focal mechanism (fault plane solution). Here we shall discuss the determination of some other, additional quantities.

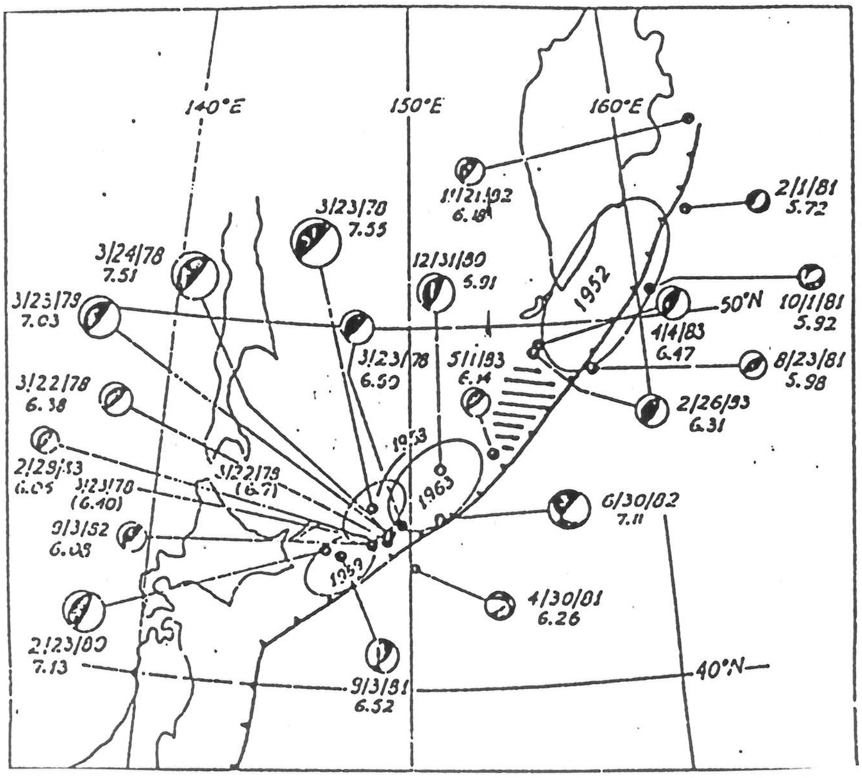


Fig.1.16. Earthquake mechanisms in the region of the Kurile Islands, 1977-83. For each earthquake under consideration, the date and magnitude are shown. (According to Kanamori and Dziewonski).

An integral and more detailed characteristic of the physical process along the fault plane, related to a point source approximation, is the moment tensor M_{ij} (i, j = 1, 2, 3). The moment tensor is symmetric, $M_{ij} = M_{ji}$, so that we have six independent components. The moment tensor is a function of time (or of the frequency, if we work in the frequency domain). The moment tensor can be obtained from observations of complete seismograms, not only from travel times and amplitudes.

Another useful information on the parameters of the source may be obtained from the Fourier amplitude spectrum of the complete earthquake seismogram (Sec. 1.6). It is the corner frequency f_0, which can be used to estimate the dimension of the rupture plane Σ. In general, by a semiempirical method, the following relation was derived for the average length L of the rupture plane,

$$L = 0.7\frac{\beta}{f_0}, \qquad (1.18)$$

where β is the shear wave velocity (if f_0 is in Hz and β in km/s, then L is in km). For big earthquakes, observed f_0 may be close to 0.01 Hz, and we obtain $L \sim 10^2$ km.

Remember formula (1.11) for the seismic moment M_0, i.e. $M_0 = \mu \, \Delta u \, \Sigma$. Writing approximately $\Sigma = L^2$, we obtain roughly

$$\Delta u = \frac{M_0}{\rho \beta^2 L^2} \quad . \tag{1.19}$$

This gives an <u>average rupture</u> Δu if M_0 is determined from the low-frequency asymptote of the amplitude spectrum (Sec. 1.16) and L from (1.18). As an example, for big earthquakes $(M_0 \sim 10^{23}\,\text{Nm})$, we obtain $\Delta u \sim 10^0 - 10^1$ m. For smaller earthquakes $(M_0 \sim 10^{16}\,\text{Nm})$, we obtain $L \sim 1$ km and $\Delta u \sim 0.1$ m.

Let us now consider a circular fracture of a radius p. By theoretical methods, it is possible to find the relation between the seismic moment M_0 and radius p:

$$M_0 = \frac{16}{7} \Delta\sigma \, p^3 \ , \tag{1.20}$$

where $\Delta\sigma$ is the <u>stress drop</u> (the stress prior to the earthquake minus the stress after the earthquake). As p can be approximately determined in the same way as L, the above relation can be used to determine $\Delta\sigma$. The stress drop can also be calculated by other methods, e.g. from the seismic energy:

$$E_S = \frac{\Delta\sigma \, M_0}{2\mu} \quad . \tag{1.21}$$

This relation can be used to calculate $\Delta\sigma$ if E_S and M_0 are known, or to calculate E_S if $\Delta\sigma$ and M_0 are known.

Stress drops for many earthquakes have been estimated. The values range from 10^5 to 10^7 Pa.

1.12.5. Recent developments in the investigations of the mechanism of the seismic source

1) <u>Experimental data</u>. The material used in recent studies of the seismic source is usually as follows:

- Complete seismograms (in a digital form), containing a broad range of amplitudes and frequencies.

- Accelerograms (also in a digital form) recorded close to the source. Such investigation in epicentral regions are called the strong motion seismology. System of such seismograms, obtained by array measurements, are also used.
- Measurements of slow movements of the plates.
- Measurements of other physical fields in epicentral regions (electric, magnetic, geothermic, etc.).
- Laboratory measurements and investigations of the fracturing process, etc.

2) Problems. The main problems which are studied are as follows:
- The details of the rupture process along the rupture plane and in time. Studies of barriers and asperities. Studies of the duration of the rupture process, including specific phenomena, such as slow earthquakes (duration of the rupture process for many tens of seconds) or quiet earthquakes (they do not generate body and surface waves, but excite free oscillations of the Earth).
- The kinematics and dynamics of the moving dislocation source. Moment tensor, radiation pattern, directivity, etc.
- The physical mechanism of fracturing.
- The relation of tectonic processes to physical processes along the rupture plane.
- Relation to other physical fields. The possibilities to use such studies for the prediction.

From a theoretical point of view, the theory and numerical modelling of the physical processes in the seismic source, particularly the generation of seismic waves by a tectonic source, are based mostly on general representation theorems (Aki and Richards, 1980).

1.13. Hypotheses and theories on the origin of earthquakes

1.13.1. Review of main hypotheses

A great number of models have been proposed for explaining the earthquake phenomena. Let us mention models of the following authors (more details can be found in Gokhberg et al., 1982):

• 1910	Reid	elastic rebound theory;
• 1967	Ulomov	model adopted after the 1966 Tashkent earthquake;
• 1968	Riznichenko	energy model;
• 1971	Myachkin et al.	unstable avalanche crack generation;
• 1972	Nur	dilatancy-diffusion model;
• 1974	Stuart	diffusionless dilatancy model;
• 1974	Brady	

- 1979, 1980 Dobrovolskiy model with an inclusion;
- 1979, 1980 Gold methane model.

The most of the present models of earthquake phenomena have been influenced by the pioneering work by H. F. Reid, who investigated in detail the 1906 San Francisco earthquake. His research resulted in the elastic rebound theory of earthquakes; see the description given above in this chapter. This theory formed a framework of all theories which followed. A further support for these hypotheses came later from the concepts of the tectonics of lithospheric plates. At the same time, i.e. in the 1960th and 1970th, some countries adopted extensive programmes of predicting earthquakes (Japan, the former Soviet Union, the USA). It further increased the interests in theories of earthquake phenomena. Let us describe briefly the main models of that time.

The model of the unstable avalanche crack generation, proposed in the USSR, is based on extensive laboratory experiments on fracturing of rocks. When a rock is deformed under an increasing stress, cracks start to develop inside the material at a certain moment. Some of these cracks are then joining together, forming a main fault, whereas the others are closing. Along the main fault, the rupture of material then occurs.

The dilatancy-diffusion model, used in the USA and other countries, is similar to the previous model, but a great importance is ascribed to the role of water. According to this model, the preparation of an earthquake goes through the following stages:

1) Building up of elastic strain.
2) Dilatancy. The opening of cracks leads to an increase of volume. This phenomenon is called the dilatancy.
3) Diffusion. An influx (diffusion) of water into the cracks leads to a decrease of the strength of the material (the material can slip on wet surfaces).
4) Earthquake.

In some later models, the role of fluids is even greater. For example, Gold's hypothesis emphasises the importance of the release of gases in geodynamic processes. The hypothesis explains the earthquakes by the vertical transport of methane from the mantle. However, observations have not proved this hypothesis. Nevertheless, it opened space for similar speculations on a possible role of other gases, such as carbon dioxide, water vapour, and even hydrogen in certain cases. These questions are still open, despite of the fact that many observations seem to support these hypotheses.

1.13.2. Earthquake phenomena and properties of complex systems

The complexity of earthquake phenomena led seismologists also to detailed investigations of analogous complicated models and processes. Let us mention two of them.

1) Brittle fracturing of rocks. Contrary to many slow geological processes, the earthquake are characterised by very fast motions. Therefore, it seems that the earthquake could be considered as a brittle fracturing of rocks. Such fracturing experiments can easily be performed at many laboratories. These laboratory experiments have revealed many important factors of the fracturing process, e.g. the role of inhomogeneities of the material, existence of tensile cracks at the ends of the fault, space and time development of the fracturing process, acoustic emission, changes of physical parameters in deformed materials, etc. Nevertheless, some authors hesitate to adopt the brittle fracturing of rocks as an appropriate model of tectonic earthquakes, since principal differences also exist between these phenomena. In particular, the following differences should be mentioned:
   a) Earthquakes often repeated practically at the same place. Such a phenomenon is not known from the laboratory experiments where the fracturing process proceeds to new places. To explain the repetition of earthquakes at the same place, we must admit the existence of sufficiently fast healing processes which join the material at the fault and restore the previous situation (mineralised water, heat, pressure).
   b) Redistribution of stresses in solid materials proceeds at the velocities of seismic waves, i.e. at the velocities of the order of kilometres per second. Consequently, the aftershocks should follow the main shock within several seconds. However, the observed aftershock sequences often last from many days to many months. Hence, some slow processes controlling the redistribution of stresses are needed. The diffusion of fluids seems to be one of possible candidates.

Thus, we may conclude that the brittle fracturing of rocks, as a model of tectonic earthquakes, requires substantial modifications and supplements.

2) Deterministic chaos. Many phenomena in nature contain deterministic and pronounced random components. In this case we often speak about the deterministic chaos. Such a situation occurs in many non-linear systems, where small changes of boundary or initial conditions may lead to substantial changes in the solution. A small change of some parameter may trigger a series of further processes. As a typical example, we could mention the meteorological phenomena. It is supposed that earthquakes belong also to this category of processes. For example, the space and time distribution of earthquakes exhibits certain deterministic features (e.g. a concentration of earthquakes along main geological faults), but also quite random components. The information on these distributions is contained in seismic catalogues. Hence, many seismologists are convinced that these catalogues and some maps should be sufficient for estimating a probable position and time of future earthquakes if the deterministic chaos is better understood. On the other hand, other investigators are more sceptic and rely rather on monitoring of various parameters, such as deformations, tilts, stresses, changes of physical parameters, water level, composition

and radioactivity of underground waters, etc. However, such measurements are more complicated and more expensive.

1.13.3. Meditations on the research of earthquakes

David Hilbert, one of the greatest mathematicians of the 20th century, organised famous seminars on atomic physics in Heildelberg in the 1920th and 1930th. On opening the seminars, he often asked the question: "Can anybody tell me what the atom is?" Such a provocative, apparently naive and simple question then stimulated very deep discussions on various aspects of the atomic research and on the essence of atomic phenomena. It contributed significantly to the progress in atomic physics.

Paraphrasing Hilbert, we should also ask the question: "Can anybody tell us what the earthquake is?" However, such a question is not heard at seminars on earthquake phenomena, since seismologists are probably convinced that the principles of the earthquake phenomena are well known, and only some details need to be added. Nevertheless, our failures in predicting earthquakes indicate something else, namely that our research is still more or less in a blind alley and enters into particulars, so that we are unable "to see the wood for the trees". It may indicate that we have not recognised some of the fundamental phenomena yet.

Kitaygorodskiy, a Russian physicists, has divided the physical theories into three categories according to their quality. He has distinguished the theories which:

1) describe;
2) explain;
3) predict.

Let us demonstrate this classification on examples from celestial mechanics:

1) Ptolemaios' theory. It is a typical example of a theory of the first category. This theory described the planetary motions quite formally by means of the so-called epicycles. An independent system of epicycles was needed for each planet.
2) Kepler's theory. Kepler reduced the number of necessary rules to the three well-known laws, which explain the planetary motions.
3) Newton's theory of gravitation. This theory is capable of predicting new phenomena. Let us mention the famous discovery of the planet Neptune by Leverrier and Adams in 1846. At present, this theory is used also for predicting the trajectories of space-crafts and artificial satellites.

In comparison with celestial mechanics, our present earthquake theories are rather primitive. They can be included into the first or, the second category of Kitaygorodskiy's classification. And for an earthquake theory of the third category, we are still waiting.

The main practical goal the research of earthquakes is earthquake prediction. Certain predictions can be made on any level of the scientific knowledge, but their precision and reliability depends fundamentally on the quality of the scientific theory used. Also three categories of the prediction theories, analogous to the above-mentioned classification, can be distinguished (Bullen and Bolt, 1993):

1) descriptive prediction, using processes of interpolation and extrapolation (e.g. predictions based on the hypothesis of seismic gaps);
2) inductive prediction, requiring a general theory (elastic rebound theory, dilatancy-diffusion theory);
3) deductive prediction, using a more detailed general theory, from which consequences are determined as logical or mathematical steps (in earthquake prediction we have not arrived at this stage yet).

Appendix A:

MACROSEIMIC INTENSITY SCALE

(very simplified description with a schematic summary)
Abbreviations: P - people, O - objects, B - buildings, N - nature

	Notation	Effects	Effects on P O B N
I.	Not felt	Registered by instruments only.	– – – –
II.	Scarcely felt	Felt by people at rest on upper floors.	+ – – –
III.	Weak	Felt by people at rest. Slight swinging of hanging objects.	+ + – –
IV.	Largely observed	Felt by many people. Swinging of hanging objects. Rattling of dishes, glasses, windows, doors.	+ + – –
V.	Strong	Some dishes, windows, etc., broken. Cracks in plaster.	+ + + –
VI.	Slightly damaging	Furniture may be shifted. Fall of pieces of plaster.	+ + + –
VII.	Damaging	Fall of chimneys. Cracks in many walls.	+ + + –
VIII.	Heavily damaging	Failure of individual walls (non-structural ones). Changes in well water.	+ + + +
IX.	Destructive	Partial structural failure. Cracks in the ground.	+ + + +
X.	Very destructive	Many masonry buildings destroyed. Ground badly cracked, rails bent.	+ + + +
XI.	Devastating	Many buildings of reinforced concrete destroyed. Broad fissures in ground.	+ + + +
XII.	Completely devastating	Practically all structures above and below ground are destroyed. Changes in the face of the landscape.	+ + + +

Appendix B:

Translation of the macroseismic questionnaire which is used by the Geophysical Institute of the Czech Academy of Sciences, Prague

MACROSEISMIC QUESTIONNAIRE

Glosses: □ □ □
YES NO I DON'T KNOW, I CAN'T ANSWER

Name of the observer:..
employment:...
address:...
phone:.......................

Earthquake: year.........month..........day..........hour..........minute..........
place of observation: municipality..................town □
street..........................village □
district........................

Observation: outdoors □ indoors □ floor:........
type of the building: wood □ stone □ brick □ prefab □
reinforced concrete □ another □
foundation of the building: normal □ on a plate □ on piles □
ground: clay □ rock □ filling □ another □

<u>Description of the macroseismic effects on the **man:**</u>
observer's position: he was standing □ sitting □ lying □
observer's feeling: loss of equilibrium □ swinging □ surprise □
unpleasant feeling □ suspense □ anxiety □ fear □
panic □ wakened □
reaction of the other people:..
how many people observed the shock: all □ most □ several □ only you □
I don't know □
further data:...
accompanying sounds: boom □ vibration □ hum □ whistle □
sounds similar to an explosion □
sounds similar to a movement of a heavy truck or tank □
other data on accompanying sounds:..
light effects □ description of the light effect:..................................
other observed phenomena:..
behaviour of animals:...

Description of the macroseismic effects on **furnishing**:
hanging objects: swinging □ turning □ fall □
dishes: clinging □ rattling □ fall □
water in containers: oscillated □ spilled □
windows: rattling □ swung open or shut □
doors: rattling □ swung open or shut □
pendulum clock: stopped □ started going □ lost or gained □
light pieces of furniture: swinging □ displacement □ overturning □
heavy pieces of furniture: swinging □ displacement □ overturning □
bells rang: small □ large □
further data:...

Description of the macroseismic effects on **buildings**:
damage to the roof □ cracks in plaster □ cracks in a wall □
fall of pieces of plaster □ fall of a wall □ collapse of a building □
damage to chimneys: cracks □ turning □ fall □
further data:...

Description of the macroseismic effects on the earth's **surface**:
landslides □ cracks in sand □ cracks in clay □ cracks in rock □
change of the water level in wells □ change of water streams □
other changes:...

Other supplementary data:
direction of motion:...
observation of other people:...
....
remarks:...
....

We ask you for kindly filling up and sending the questionnaire. Your data will be used for investigating earthquakes and earthquake risk on our territory. We also ask for the names of other persons who can give further information as observers, especially if they live in another place. Give your address and telephone number always, even if you have not felt the earthquake.

We emphasise that every, even the least report has its value for further processing. We are interested also in negative observations (see above), in information on previous earthquakes, reports in chronicles, etc. We thank you for your willingness and co-operation, which enables us to collect valuable data for studying the seismicity of the Czech Republic.

Geophysical Institute, Acad. Sci. of the Czech Republic

Seismological Department, Geophysical Inst., Acad. Sci. of the Czech Republic,
Bocni II, 141 31 Prague 4
e-mail: seis@ig.cas.cz http;//seis.ig.cas.cz

Appendix C:

Hang this up. ~ Follow these tips.

27 things to help you survive an earthquake

Californians are constantly aware of the potential of an earthquake creating damage and creating dangerous conditions. So if we don't properly prepare, the next quake may cause greater personal damage than necessary. Each item listed below won't stop the next earthquake but it may help you survive in a better way.

4 basics to do during an earthquake

1) STAY CALM.
2) Inside: Stand in a doorway, or crouch under a desk or table, away from windows or glass dividers.
3) Outside: Stand away from buildings, trees, telephone and electric lines.
4) On the road: Drive away from underpasses/overpasses; stop in safe area; stay in vehicle.

6 basics to do after an earthquake

1) Check for injuries - provide first aid.
2) Check for safety - check for gas, water, sewage breaks; check for downed electric lines and shorts; turn off appropriate utilities; check for building damage and potential safety problems during after shocks such as cracks around chimney and foundation.
3) Clean up dangerous spills.
4) Wear shoes.
5) Turn on radio and listen for instructions from public safety agencies.
6) Don't use the telephone except for emergency use.

14 survival items to keep on hand

1) Portable radio with extra batteries.
2) Flashlight with extra batteries.
3) First Aid Kit - including specific medicines needed for members of your household.
4) First Aid book.
5) Fire extinguisher.
6) Adjustable wrench for turning off gas and water.
7) Smoke detector properly installed.
8) Portable fire escape ladder for homes/apartments with multiple floors.
9) Bottled water - sufficient for the number of member in your household.
10) Canned and dried foods sufficient for a week for each member of your household. Note: Both water and food should be rotated into normal meals of household so as to keep freshness. Canned goods have a normal shelf-life of one year for maximum freshness.
11) Non-electric can opener.

12) Portable stove such as butane or charcoal. Note: Use of such stoves should not take place until it is determined that there is no gas leak in the area. Charcoal should be burned only out of doors. Use of charcoal indoors will lead to carbon monoxide poisoning.
13) Matches.
14) Telephone numbers of police, fire, and doctor.

3 things you need to know

1) How to turn off gas, water and electricity.
2) First Aid.
3) Plan for reuniting your family.

The best survival is a prepared survival

American Red Cross

Chapter 2

Observations of Seismic Waves

The mechanical waves which are generated by earthquakes or explosions can propagate through the interior of the Earth's and along its surface. These waves, called seismic waves, are recorded by seismograph station the world over, provided that the released energy at the source has been big enough.

The record of a seismograph is called the seismogram. Seismograms usually show complicated wave motions, and may be of long duration (minutes, hours), especially in the case of distant earthquakes.

The term of the "teleseism" is frequently used in this connection. A teleseism is an earthquake recorded by a seismograph at a great distance. By international convention, this distance is required to be over 1 000 km from the epicentre. The distance over 1 000 km is thus also referred to as the teleseismic distance. Earthquakes recorded nearer the recording station are "near earthquakes" or "local earthquakes" (Richter, 1958).

2.1 Structure of a Seismogram

A schematic form of a seismogram of a distant earthquake is shown in Fig. 2.1, and an example of a real seismogram is shown in Fig. 2.2.

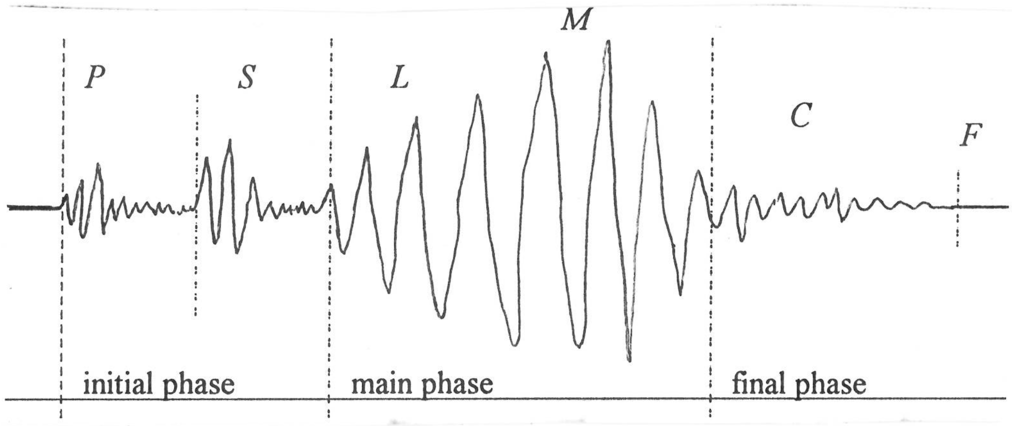


Fig. 2.1. A schematic form of a seismogram.

The first record of a distant earthquake was obtained in 1889. This earthquake occurred in Japan and the record was written in Potsdam, Germany. At first, the complicated form of seismograms was ascribed to the complexity of earthquake sources. However, as soon as further experimental data were accumulated, an alternative explanation was adopted, namely that the initial disturbances at earthquake sources were relatively simple and of short duration, but the recorded complexity arose between hypocentre and station. Consequently, the waves of different types, travelling along different paths and at different velocities, had to be considered in order to explain the observations.

It was found very soon that the records of distant earthquakes showed two very distinct wave groups, the first being characterised by a very feeble motion, the second by a much larger motion. These parts of the seismogram were termed the "preliminary tremor" and "main shock", respectively. Note that the "main shock" was also termed "large waves", the "principal portion" or "principal earthquake". These two part of the seismogram were at first interpreted as longitudinal and transverse waves, known from the theory of elasticity. However, this interpretation had to be modified when Oldham (1900) recognised two distinct parts in the preliminary tremor.

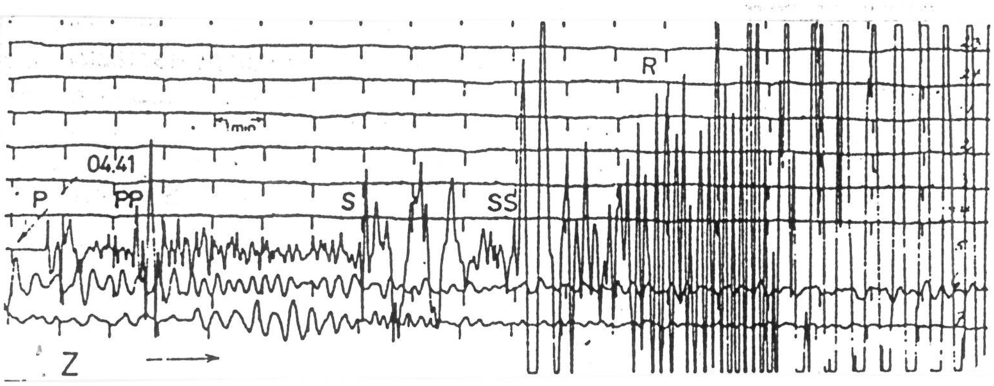


Fig. 2.2. An earthquake in China on November 13, 1965, recorded at Kiruna, Sweden. The distance between the neighbouring time marks represents 1 minute. The higher-mode Rayleigh waves are exceptionally pronounced. (After Båth (1979)).

Note that the individual parts of the seismogram, or even the individual waves, are frequently referred to as *seismic phases*. Thus, we may speak of the initial, principal (main) and final phases of a seismogram, etc. (Fig 2.1).

An international terminology, based on Latin designation, was adopted for reporting the normal type of a seismogram. The following abbreviations were introduced for the individual seismic phases:

- *P* (*undae primae*) for the primary waves (first waves, the first preliminary tremor);
- *S* (*undae secundae*) for the secondary waves (the second preliminary tremor);
- *L* (*undae longue*) for large, long-period waves of the principal phase;
- *M* (*undae maximae*) for the maximum of the seismogram;
- *C* (*cauda*, or *coda* from Italian) for decreasing later waves;
- *F* (*finis*) for the approximate end of the recorded disturbance.

Special notations are also used to describe the type of the onset of a seismic phase:

- *i* (*impetus*) denoting a sharp onset when the beginning of the seismic phase is clearly seen;
- *e* (*emersio*) denoting a gradual onset.

The letters *i* and *e* are written before the letters *P*, *S* and others, denoting the corresponding seismic phases. For example, *iP* denotes a sharp onset of the *P* phase, *eP* denotes a gradual onset of the *P* phase, etc. These abbreviations are frequently used in seismic bulletins.

In the following sections of this chapter we shall give the physical interpretation of the individual seismic phases in terms of seismic waves propagating within the Earth. The corresponding solutions of the elastodynamic equations will be derived in the following chapters.

2.2 Body Waves

Seismic body waves propagate through the body of the Earth. It follows from the theory of elasticity that there are two principal types of body waves:

1) *Longitudinal waves*, also called compressional, dilatational or irrotational waves. These waves involve compression and rarefaction of the material as the wave passes through it, but not rotation. The particles of the medium, through which the longitudinal wave is passing, vibrate about the equilibrium position in the same direction as the direction of wave propagation (Fig. 2.3). These waves are the analogue of sound waves in the air. Longitudinal waves are identified with the *P* phase on seismograms. Consequently, longitudinal waves are also commonly called *P* waves.
2) *Transverse waves*, also called shear, rotational or equivoluminal waves. These waves involve shearing and rotation of the material as the wave passes through it, but no volume change. The particle motion is perpendicular to the direction in which the wave is travelling. These waves are identified with the *S* phase on seismograms, so that they are usually termed *S* waves. The *S* wave motion can be split into a horizontally polarised motion termed *SH* and a vertically polarised motion termed *SV*.

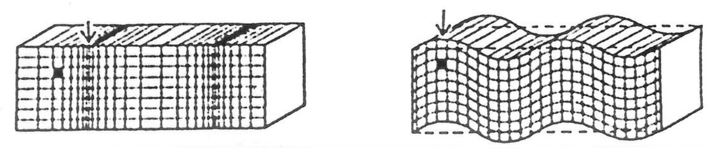


Fig. 2.3. Deformations of the medium when body waves are passing from left to right: *P* wave (on the left), and *SV* wave (on the right). (After Fowler (1994)).

Body waves are reflected and transmitted at interfaces where the elastic coefficients and/or density change. This increases the number of waves recorded on seismograms (see below).

2.3 Surface Waves

In a homogenous, isotropic and unlimited medium, only longitudinal and transverse waves can propagate. If the medium is bounded, another type of waves, so-called *surface waves*, can be guided along the surface of the medium. Surface waves do not penetrate deeply into the medium, the depth of their penetration being usually comparable with the wavelength. These waves usually form the principal phase of the seismogram. There are two types of surface elastic waves:

1) *Rayleigh waves*, named after Rayleigh, who predicted their existence in 1887. The particle motion in these waves is confined to a vertical plane containing the direction of propagation (Fig. 2.4a). Near the surface of a homogeneous half-space this is a retrograde vertical ellipse (anticlockwise for a wave travelling to the right). Thus, Rayleigh waves are elliptically polarised waves. These waves can therefore be recorded by both the vertical and horizontal components of the seismometer. Rayleigh waves are denoted by *LR* or *R* (*L* for long; *R* for Rayleigh).
2) *Love waves,* named after Love, who predicted their existence in 1911. The particle motion of these waves is transverse and horizontal, so that they can only be recorded by horizontal seismometers (Fig. 2.4b). Love waves are denoted by *LQ* or *Q* (*Q* for Querwellen, German, meaning "transverse waves"). As opposed to Rayleigh waves, Love waves cannot propagate in a homogeneous half-space. Love waves can exist only if, in general, the *S*-wave velocity increases with depth. These waves propagate by multiple internal reflections of horizontally polarised *S* waves (*SH* waves) in a near-surface medium. Hence, Love wave represents the interference phenomenon of *SH* waves.

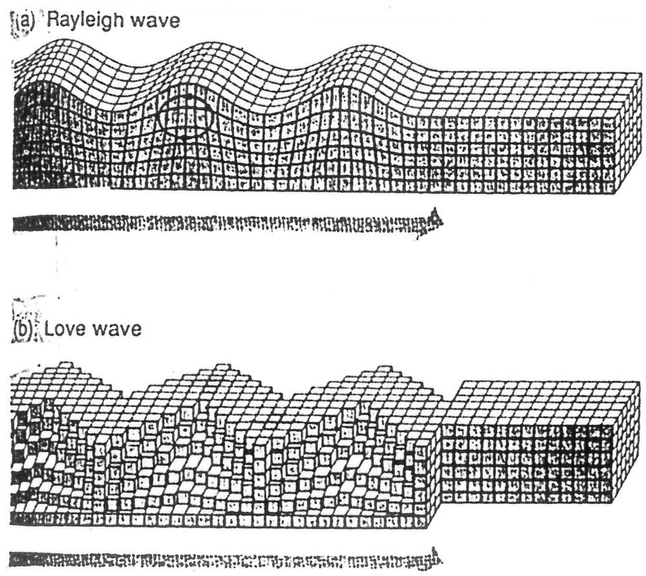


Fig. 2.4. The particle motion for surface waves: (a) Rayleigh waves and (b) Love waves. (After Fowler (1994)).

Seismic surface waves are generated best by shallow earthquakes. Deep earthquakes and nuclear explosions do not generate comparable surface waves. The absence of strong surface waves on seismograms is thus used as one criterion for discriminating nuclear explosions from earthquakes.

2.3.1 Some differences between body waves and surface waves

Surface waves are usually larger in amplitude and longer in duration than body waves. Surface waves arrive after main *P* and *S* waves, because their velocities are lower than those of body waves. The larger amplitudes of surface waves can easily be explained in the following way.

Consider harmonic oscillations of a unit volume in an elastic medium:

$$u = A\cos(\omega t) \ ,$$

where u is the displacement, A its amplitude, ω the angular frequency and t the time. The kinetic energy of this volume element is

$$\frac{1}{2}\rho\left(\frac{\mathrm{d}u}{\mathrm{d}t}\right)^2 = \frac{1}{2}\rho\omega^2 A^2 \sin^2 \omega t \ ,$$

ρ being the density. Calculate the mean value of this energy, averaged over a cycle. As the mean value of $\sin^2 x$ is

$$\frac{1}{2\pi}\int_0^{2\pi} \sin^2 x \,\mathrm{d}x = \frac{1}{2} \ ,$$

the mean kinetic energy of the volume element is

$$E_k = \frac{1}{4}\rho\omega^2 A^2 \ .$$

According to Rayleigh's principle, the mean kinetic energy and mean elastic potential energy are the same. Consequently, the mechanical energy (the sum of the kinetic and potential energies) of the vibrating element is

$$E = \frac{1}{2}\rho\omega^2 A^2 \ . \tag{2.1}$$

Thus, the amplitude of a harmonic wave is proportional to the square of this energy density, $A \sim \sqrt{E}$.

As the energy in body waves diverges in three dimensions, whereas the energy in surface waves only in two dimensions (Fig. 2.5), surface waves acquire a continually increasing preponderance at a great distance from the source. Let us give a rough estimate of this effect.

At distance x from a source, *O*, the area of a spherical wavefront is $4\pi x^2$; see Fig. 2.5a. By conservation of energy, the energy density in a body wave at distance x is thus proportional to $1/x^2$. Consequently, its amplitude is proportional to $1/x$. As opposed to it, the area of a cylindrical wavefront is

$2\pi xz$, where z characterises the depth of penetration (Fig. 2.5b). By conservation of energy, the energy density of a surface wave at distance x is thus proportional to $1/x$, and the amplitude to $1/\sqrt{x}$. Hence, the amplitudes of body waves decrease with the distance from the source approximately as $1/x$, whereas the amplitudes of surface waves decrease approximately as $1/\sqrt{x}$. Note that these amplitude estimates are only approximate, because many effects have not been taken into account, such as the influence of inhomogeneities of the medium, absorption and dispersion.

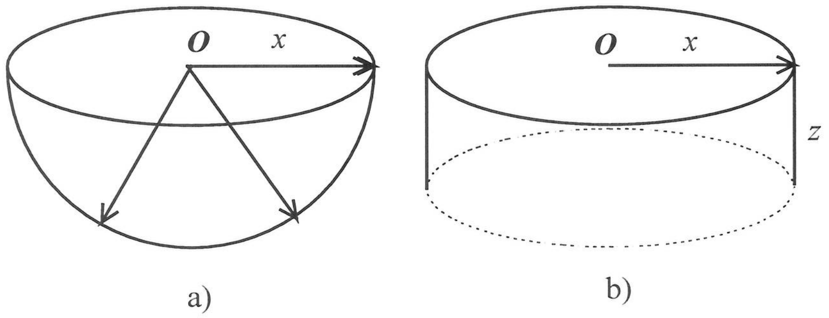


Fig. 2.5. Propagation of the energy of (a) body waves and (b) surface waves.

Surface waves are not, in principle, a new type of waves but only a result of interference of body waves. As we have already mentioned, Love waves originate from interference of *SH* waves. Rayleigh waves are formed by interfering *P* and *SV* waves. However, the interfering waves, forming both Love waves and Rayleigh waves, include also so-called inhomogeneous body waves (see the next chapters). Inhomogeneous waves are connected, e.g., with the total reflection; such waves are generated in the faster medium when the total reflection occurs in the slower medium. Since inhomogeneous waves diminish with the distance from the interface, the amplitudes of surface waves are also reduced at large depths. The decrease depends mainly on the wavelength. For the fundamental modes of surface waves, the amplitudes at a depths of one wavelength are of the order of 0.1 of the surface amplitude (usually 0.1 to 0.2 of the surface value).

The depth of penetration of fundamental modes may be estimated by the value of $\lambda/3$, λ being the wavelength. As the velocities of surface waves, generated by distant earthquakes, are about $c = 4\,\mathrm{km\,s^{-1}}$, another simple rule of thumb may be proposed, namely that the depth of penetration

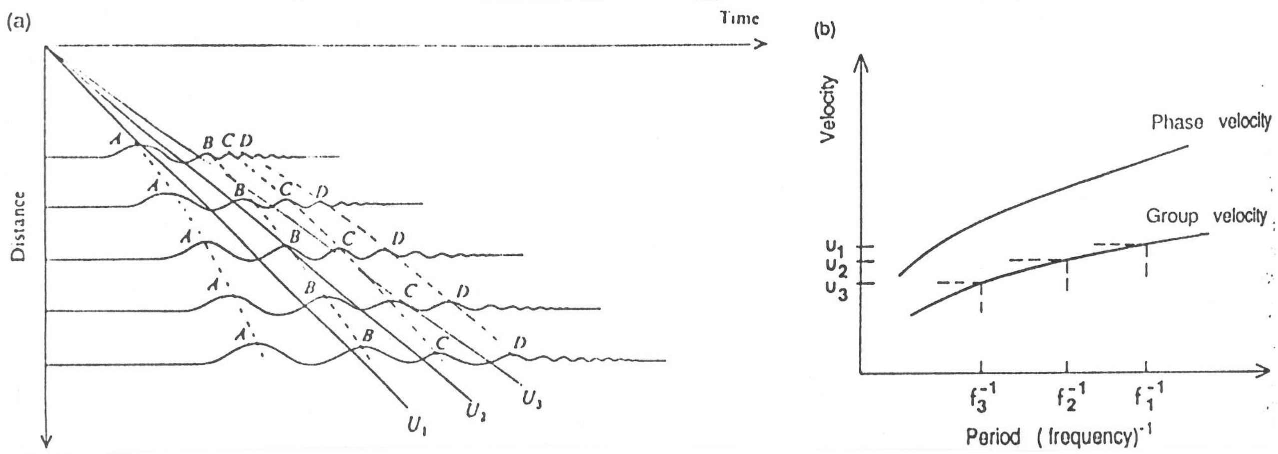
$$h \approx \frac{\lambda}{3} = \frac{cT}{3}$$

in kilometres is numerically equal to period T in seconds. Thus, we may say roughly that, e.g., a surface wave with a period of 50 s penetrates to a depth of about 50 km, etc. Therefore, surface waves with periods lower than about 50 s may be used to study the Earth's crust, and surface waves with longer periods to study the mantle structure.

We have already mentioned the differences between body and surface waves in their velocities, periods, amplitudes, polarisation and the paths of propagation. However, there is another remarkable feature of surface waves, namely their dispersion. Although body waves in real media are also dispersive, the dispersion of surface waves is much more pronounced; see below.

2.3.2 Dispersion of surface waves

Surface waves are usually *dispersive*, which means that their velocity depends on frequency. This is the consequence of their interference character. The only exception of non-dispersive surface waves are Rayleigh waves in a homogeneous half-space. Love waves do not exist in this simple medium.



a) b)

Fig. 2.6. Propagation of dispersive surface waves: (a) spreading of the wave train with the distance from the source; (b) phase- and group-velocity dispersion curves for the example shown in (a). (After Fowler (1994)).

The dispersion causes that a wave train changes its shape as it travels (Fig. 2.6). The first surface waves to arrive are of those frequencies that have the highest velocities. The waves of the other frequencies arrive later according to their velocities. Consequently, seismograms at increasingly larger distances from an earthquake are increasingly spread out.

Surface waves propagating in the Earth have generally higher velocities for long periods, and lower velocities for short periods. This is closely related to the general increase of the *P*- and *S*-wave velocities with depth. Namely, as surface waves with long wavelengths (long periods) penetrate to larger depths, their velocities are influenced by the medium with higher body-wave velocities. Consequently, long-period surface wave acquire higher velocities than short-period waves (Fig. 2.6). A real seismogram, beginning with long periods which gradually decrease, is reproduced in Fig. 2.7. Since the waves in this case propagated approximately from south to north, the surface waves recorded with the *EW*-component seismograph are Love waves.

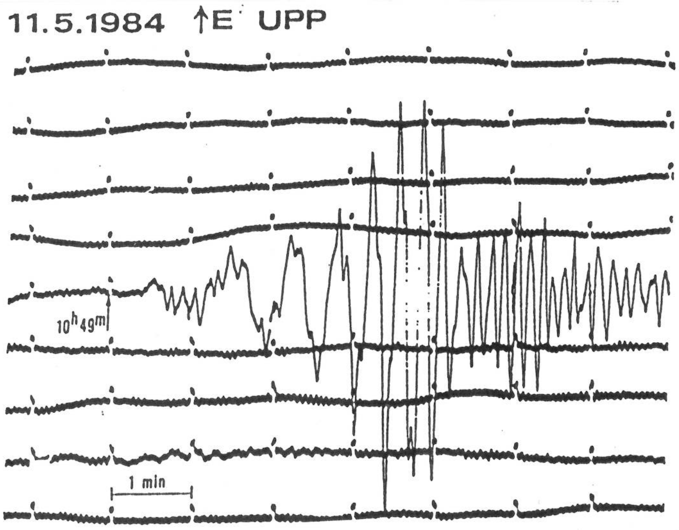


Fig. 2.7. An earthquake in Italy on May 11, 1984, recorded at Uppsala, Sweden. The epicentral distance was about 2 000 km. (After Novotny et al. (1997)).

The velocity with which any particular phase travels (e.g., a peak or trough) is called the *phase velocity*. However, the periods of peaks and troughs are not constant, but increase as the wave travels. This is shown in Fig. 2.6a by the dashed curves which link the subsequent peaks *A*, *B*, *C* and *D*. The phase velocity of a selected peak (say peak *A*) at any particular distance is the slope of the dashed curve at that distance. The slopes of all these dashed lines indicate that, in this example, the phase velocity increases with period (Fig. 2.6b).

The drawing of the curves which link the subsequent peaks or troughs represents a very simple method of determining the phase velocities from observed seismograms. Consider two seismograms recorded along a profile which passes through the epicentre. The following approximate formula for computing phase velocity c may then be used (Savarensky, 1975):

$$c\left(\frac{T_1 + T_2}{2}\right) = \frac{\Delta_2 - \Delta_1}{t_2 - t_1}, \qquad (2.2)$$

where Δ_1 and Δ_2 are the epicentral distances of the corresponding stations (measured along the great circle arc linking station and earthquake), t_1 is the arrival time of a particular peak at the first station, T_1 is its period, t_2 and T_2 are the arrival time and period of the same peak at the second station. Note that this phase velocity is ascribed to the mean period $(T_1 + T_2)/2$. Such an approximation is sufficient in most applications. Another, relatively simple method of determining phase velocities from observations uses the phase spectra of the corresponding seismograms; we shall not derive the corresponding formula here.

We have seen in the previous examples that a smoothed seismogram of surface waves has a character of a quasi-harmonic wave with a variable period.

Therefore, at any time we may speak of the instantaneous period of the surface wave. The points of the same instantaneous periods in Fig. 2.6a are linked with the solid lines (in this example, these lines are the straight lines with slopes U_1, U_2 and U_3). The slopes of these lines define another velocity, termed the *group velocity*. We shall denote it by U. Since it is difficult to find visually the same instantaneous periods on two seismograms, this method of determining the group velocity, as indicated by the solid lines in Fig. 2.6a, is not usually used in practice. Various time-frequency analyses are usually used for this purpose.

It can be shown that the energy of surface waves, associated with a particular period, propagates with the group velocity (Brillouin, 1960). It follows from the corresponding theory that the group velocity is given by the relation (Bullen, 1965; Savarensky, 1975; Novotny, 1999)

$$U = \mathrm{d}\omega / \mathrm{d}k \ , \tag{2.3}$$

where ω is the angular velocity, $k = \omega / c$ the wavenumber, and c the phase velocity. Considering the reciprocal value,

$$\frac{1}{U} = \frac{\mathrm{d}k}{\mathrm{d}\omega} = \frac{\mathrm{d}(\omega/c)}{\mathrm{d}\omega} = \frac{1}{c} - \frac{\omega}{c^2} \frac{\mathrm{d}c}{\mathrm{d}\omega} \ ,$$

we arrive at another formula for the group velocity,

$$U = \frac{c}{1 - \dfrac{\omega}{c} \dfrac{\mathrm{d}c}{\mathrm{d}\omega}} \ . \tag{2.4}$$

Other formulae are as follows:

$$U = \frac{c}{1 - \dfrac{f}{c} \dfrac{\mathrm{d}c}{\mathrm{d}f}} = \frac{c}{1 + \dfrac{T}{c} \dfrac{\mathrm{d}c}{\mathrm{d}T}} = c - \lambda \frac{\mathrm{d}c}{\mathrm{d}\lambda} \ , \tag{2.5}$$

where f is the frequency, T the period, and $\lambda = cT$ the wavelength.

2.4 Free Oscillations of the Earth

Large earthquakes excite free oscillations of the Earth, when the Earth vibrates like a giant bell. These oscillations were first recorded after the large earthquake in Kamchatka in 1952, but their periods were first determined after the large earthquake in Chile on May 22, 1960.

There are two independent types of free oscillations: toroidal, or torsional, oscillations (T) and spheroidal oscillations (S). Toroidal oscillations are perpendicular to the radius vector (with its origin at the Earth's centre), but spheroidal oscillations have generally both radial and tangential components.

The simplest spheroidal oscillation is a purely radial motion. Both types of oscillation have an infinite number of modes (or, as in music, overtones).

Free oscillations can be detected by strain meters, tiltmeters and gravimeters (gravimeters are able to record spheroidal vibrations only). Standard seismometers are not suitable for detecting these long-period motions.

Like all free vibrations, free oscillations of the Earth are standing waves. The spheroidal waves arise through mutual interference of propagating Rayleigh waves, and the toroidal oscillations arise from Love waves in the same way. Therefore, the spheroidal oscillations represent a discrete long-period continuation of Rayleigh waves, whereas the toroidal oscillations represent an analogous extension of Love waves. By measuring the periods of free oscillations, the dispersion curves of surface waves can be extended out to the maximum period of about 3 200 s for the *S* modes and to about 2 000 s for the *T* modes.

The periods of free oscillations depend on the distribution of elastic parameters and density within the Earth. The observations of free oscillations contributed to improving the velocity models and, especially, the density models of the Earth.

2.5 Microseisms

A certain type of seismic noise is permanently present on seismograms. This noise is known as microseisms. There are various sources of microseisms, such as sea waves, meteorological factors (variations in the atmospheric pressure, wind), traffic, vibrations of heavy machines, swinging of high buildings and others. Relatively intensive microseisms have periods of about 6 s. The amplitudes of microseisms usually decrease with depth below the surface. Hence, to reduce microseisms, seismometers are frequently located in boreholes.

The physical nature of microseisms is not quite clear. They seem to consist of Rayleigh and Love waves, including their higher modes. At some places, the body-wave component of microseisms is also significant.

Localities with a low level of microseisms are required for placing seismic stations, accelerators of elementary particles, electron microscopes and other precise instruments. Investigations of microseisms have also been used in seismic microzoning. It has been found, for example, that the places with a high level of microseisms usually coincide with the places where increased macroseismic intensities are observed during earthquakes.

2.6 Travel-Time Curves of Body Waves

The first processing of a seismogram usually consists in determining the arrival times of the individual seismic phases which can be recognised. The arrival times are written into a column in the preliminary seismic bulletin. If the type of some phases can be determined (e.g., *P* or *S*), this type is also written in the preliminary bulletin.

BUDAPEST

Date	Phase	Heure de Greenwich h	m	s	Période s	Amplitude A_N μ	A_E μ	A_Z μ	Δ km	Remarques
31. N—S	P	21	8	23					83,8°	Philippines
	e		9	37					9300	
	iS		18	42	5	5				
	e		19	12						
	e			55						
	eSSS		26	42						
	eL		42							
	M		49		16	5				
	F	22	10							
E—W	iP	22	8	24						
	i			39						
	PP		11	39						
	iS		18	39						
	ePS		19	15						
	PPS			53						
	SS		24							
	eL		47							
	F	22	10							
Juin 5. N—S	P	17	10	2					97,5°	Iles de
	ePP		13	27					10800	Kiou-Siou
	SKKS		20	14						
	iS			39						
	ePS		22	5						
	ePPS			38						
	eSSP		27	30						
	eSSS		31	23						
	eL		38							
	M		43		24	78				
	M		48		17	47				
	M		50,5		14	27				
	M		52		16	43				
	M		53		16	40				
	F	18	30							
E—W	P	17	10	2						
	i			15						
	ePP		12	52						
	SKKS		20	16						
	iS			40						
	e		21	20						
	ePPS		22	49						
	eSSP		27	33						
	eSSS		30	5						
	eL		38							
	M		48		16		35			
	M		50		14		51			
	M		52		15		58			
	M		53		14		17			
	F	17	40							
6. N—S	P	16	16	41					32,5°	
	i		17	8					3610	
	e		21	38						

Fig. 2.8. Page from seismological station bulletin, Budapest, June 5, 1951. (After Richter (1958)).

If the onsets of the *P*- and *S*-phases are clearly seen on the seismogram, the epicentral distance of the earthquake can be determined from the time difference between the arrival times of these phases. Then, if the epicentral distance is known, the earthquake magnitude can also be determined. The preliminary seismic bulletins are regularly sent to the international seismological centres.

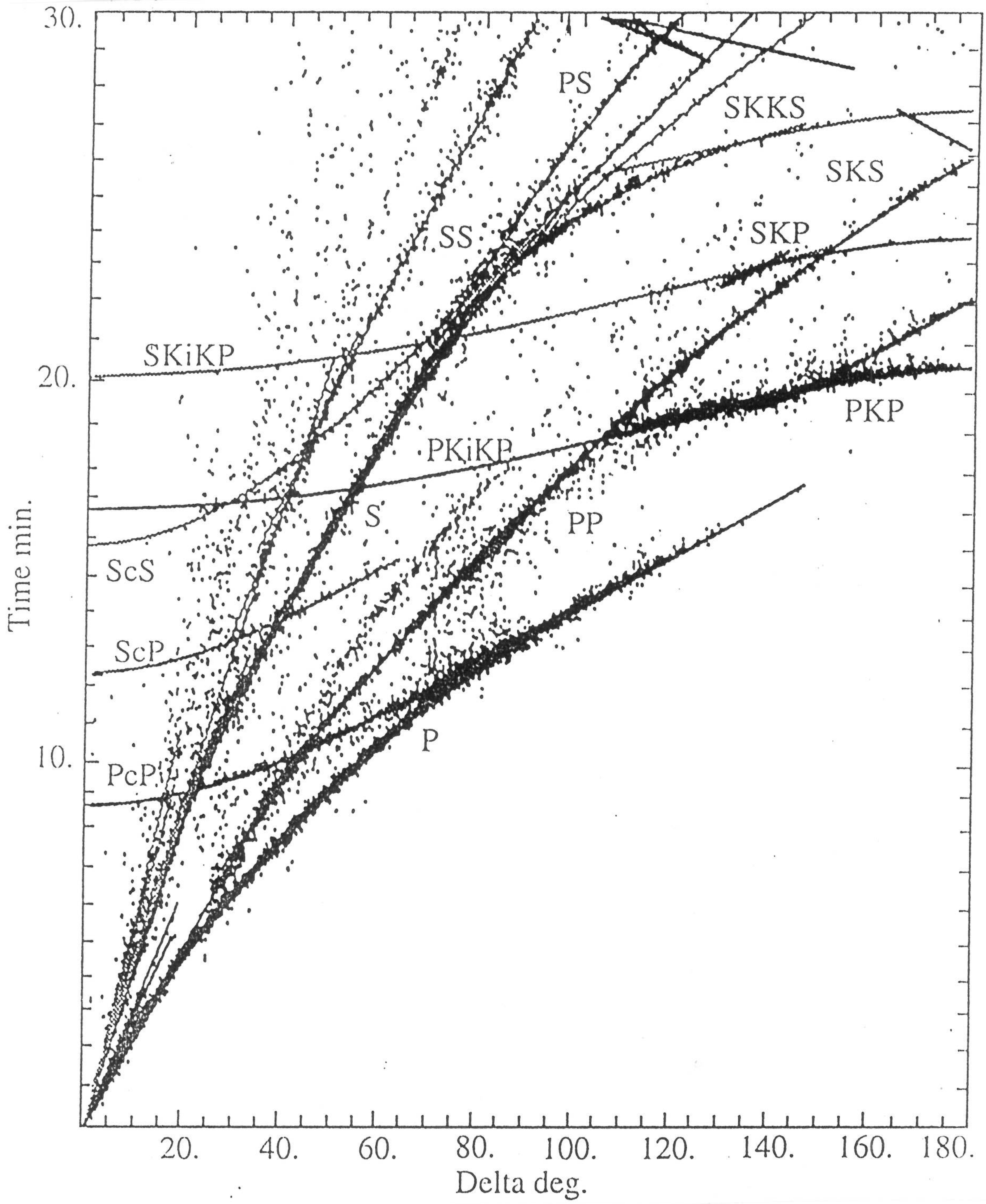


Fig. 2.9. Graph of travel times of seismic phases identified in the IASPEI 1991 Seismological Tables. (After Stacey (1992)).

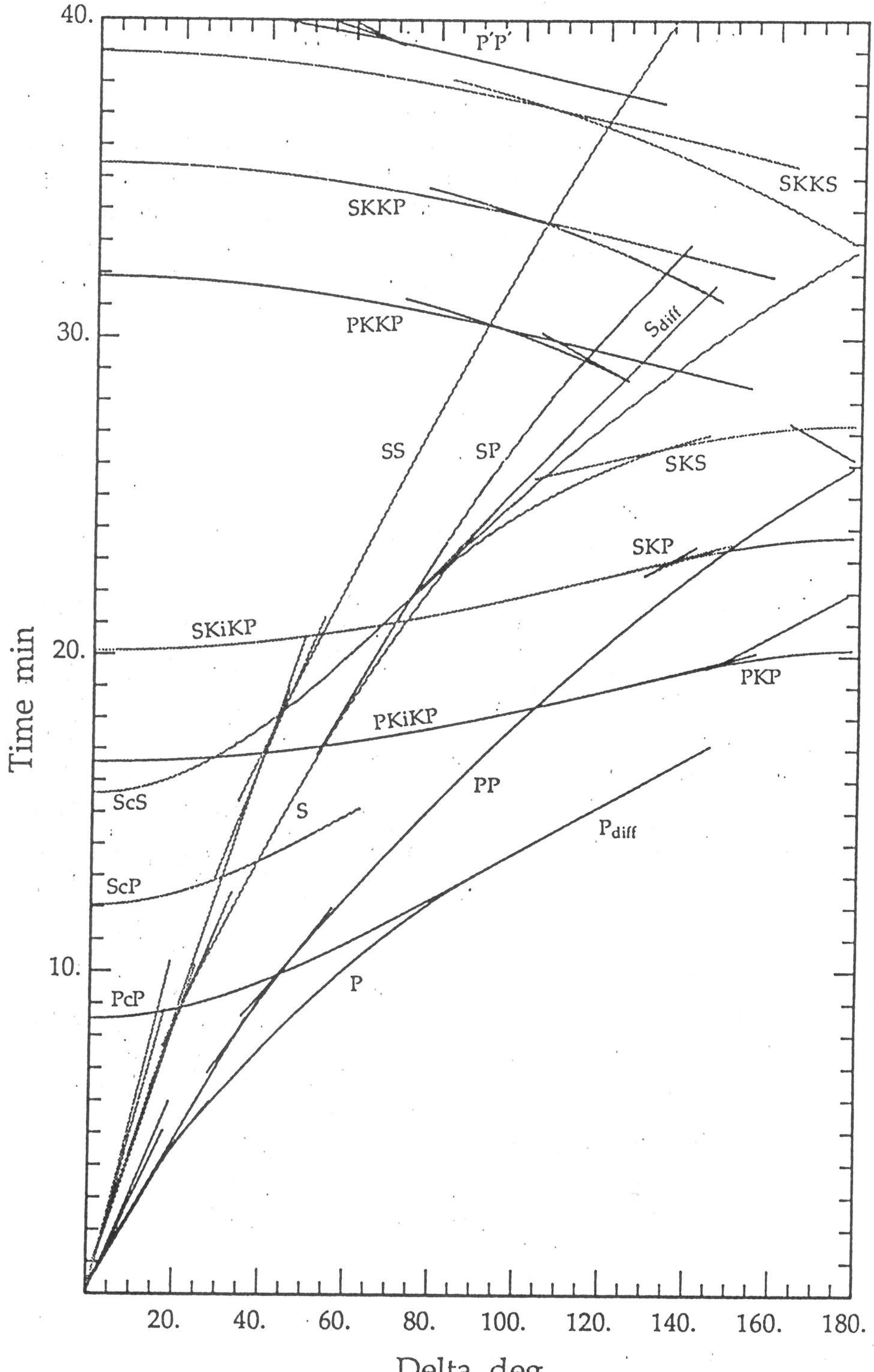


Fig. 2.10. Travel-time curves. (After Stacey (1992)).

The seismological centres carry out the localisation of earthquakes, i.e. computing the geographical coordinates, depth and origin time for each earthquake. As usual, these localisations are based only on the first arrivals of *P* waves. The results of the localisations are then sent back to the seismic stations to be used in preparing their final bulletins. Many new data are added into these bulletins, such as the identifications of further seismic phases, epicentral distances, azimuths, earthquake magnitudes, etc. An example of a final seismic bulletin is shown in Fig. 2.8.

The seismological centres also calculate the tables and graphs of the travel times for the individual seismic phases as functions of the epicentral distance (Figs. 2.9 and 2.10). An example of such tables for several epicentral distances is given in Tab. 2.1. These tables and graphs are used at seismological stations to determine epicentral distances and to identify further seismic phases.

The following simple rule, referred to as Laska's rule, may be used to estimate roughly the epicentral distances of distant earthquakes:

$$\Delta = \left(t_S - t_P\right) - 1 , \tag{2.6}$$

where epicentral distance Δ is given in megametres (thousands of kilometres), and the arrival times t_S for the *S* wave and t_P for the *P* wave are given in minutes. For example, if $t_S - t_P$ is 10 minutes, Laska's rule yields an epicentral distance of 9 000 km. In this case, Tab. 2.1 yields a more accurate value of $\Delta = 80° = 8890\,\text{km}$.

Table 2.1. Travel times (min:sec) of several seismic phases for a focal depth of 25 km. After the tables by Richter (1958); the travel times for $\Delta = 103°$ have been extrapolated.

Δ°	*P*	*S*	*S–P*	*PcP*	*ScS*
0	0:04	0:08	0:04		15:30
10	2:28	4:17	1:49		15:39
20	4:34	8:12	3:38		16:04
30	6:11	11:05	4:54	9:08	16:45
40	7:36	13:39	6:03	9:32	17:39
50	8:55	16:03	7:08	10:11	18:42
60	10:06	18:19	8:13	10:50	19:53
70	11:12	20:21	9:09	11:32	21:09
80	12:11	22:11	10:00	12:17	22:30
90	13:00	23:49	10:49	13:01	
100	13:46	25:14	11:28		
103	14:00	25:40	11:40		

2.7 Elementary Interpretation of Travel-Time Curves

One of the most important problems in seismology since the beginning of the 20th century has been the determination of travel times with the highest possible accuracy. Such travel-time curves have been required for the following principal reasons:

- These curves find practical applications in localising earthquakes; see the previous section.
- Travel-time curves can be used to determine the distribution of seismic velocities within the Earth; see below. This velocity distribution represents fundamental information on the internal structure of the Earth.

Here we shall deal briefly with elementary interpretations of travel-time curves. More accurate methods will be described in Chapter 3. These methods will enable us to construct the rays of individual seismic waves and to determine the distribution of seismic velocities within the Earth with a higher accuracy.

The most important branches of the travel-time curves are those for the direct *P* wave and the direct *S* wave; see Figs. 2.9 and 2.10. As usual, the epicentral distance in these figures is expressed as the angular distance (as the angle between the radius vectors of epicentre and receiver, the radius vectors being constructed from the Earth's centre).

2.7.1 Seismic division of the Earth

When the epicentral distance approaches about 103° (for ordinary shallow earthquakes), the rays of both *P* and *S* waves graze a discontinuity. This conclusion follows from the fact that beyond this distance the amplitudes of *P* and *S* waves decrease rapidly. This discontinuity is the boundary of the Earth's core. The part of the Earth above this discontinuity is called the Earth's mantle. Another part of the Earth, called the Earth's crust, can be distinguished on the basis of observations of near earthquakes. Hence, on the basis of the properties of the *P*- and *S*-wave travel times, we immediately arrive at the basic division of the Earth into the *Earth's crust*, *mantle* and *core*. This division represents the simplest seismic model of the Earth.

Both *P* and *S* waves can propagate in the Earth's crust and mantle. However, only *P* waves can propagate in the core (more accurately, in the outer core), whereas no *S* waves propagating in the core have been observed. Since transverse waves can propagate is solids and not in liquids, we may immediately conclude that the *Earth's crust and mantle are solid*, but the *outer core is liquid.* We shall see later that the inner core is again solid.

2.7.2 Travel times in a homogeneous sphere

Now, let us derive elementary quantitative estimates following from the travel times in Tab. 2.1. Assume the outer part of the Earth to be homogenous up to the core. Seismic rays are then straight lines, and elementary geometry can be used. The situation is shown in Fig. 2.11.

Consider source *O* and receiver *A* to be located at the Earth's surface, and denote their distance by *s* (Fig. 2.11a). It follows from triangle *OCB* that

$$\frac{s}{2} = R \sin\frac{\Delta}{2} \,, \tag{2.7}$$

where R is the Earth's radius and Δ the angular epicentral distance. Since $s = vt$, where v is the velocity and t the corresponding travel time of the wave propagation from O to A, we get the following equation for the travel-time curve:

$$t = \frac{2R}{v} \sin\frac{\Delta}{2} \,. \tag{2.8}$$

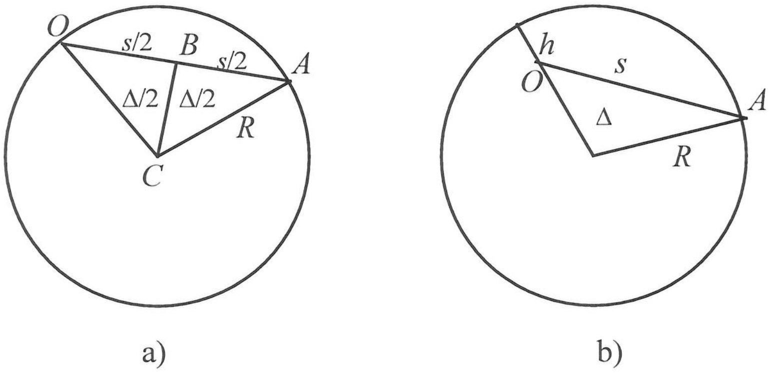


Fig. 2.11. Propagation of a direct wave inside a homogeneous Earth: O is the source, A the receiver, B the central point of the abscissa OA, C the Earth's centre, R the Earth's radius, Δ the angular epicentral distance, s the direct distance between the source and receiver, h the source depth ($h = 0$ in the left figure).

This is a very simple example of the equation of a travel-time curve, where the travel time is expressed by an analytical formula. Later on, we shall encounter more complicated situations when it will not be possible to express the travel time explicitly, but only in a parametric form.

In geophysics we frequently speak of *forward problems* and *inverse problems*. By a forward problem we understand the determination of a certain physical quantity (gravity or magnetic anomaly, travel time, synthetic seismogram, etc.) for a given model. The opposite problem, i.e. the determination of the parameters of a model from measured data, is called the inverse problem.

Equation (2.8) represents a very simple example of solving a forward problem. For a given model of the medium, i.e. for given values of R and v, Eq. (2.8) makes it possible to calculate travel time t as a function of epicentral distance Δ.

Express Eq. (2.8) as an equation for the unknown velocity v:

$$v = \frac{2R}{t} \sin\frac{\Delta}{2} \,. \tag{2.9}$$

Assuming the Earth's radius R to be known, this equation solves the corresponding inverse problem, i.e. makes it possible to determine the unknown velocity v. In other words, Eq. (2.9) solves the inverse problem for the travel time of a direct wave propagating in a homogeneous sphere from a surface source (Fig. 2.11a). This is an exceptionally simple case of solving inverse problems. Most seismic inverse problems are much more complicated, and their solution cannot be expressed in an explicit form. These problems usually lead to systems of algebraic or differential equations, which are frequently non-linear.

Put $R = 6370\,\text{km}$ for the mean radius of the Earth. Insert the travel times from Tab. 2.1 into Eq. (2.9), and ignore the fact that these travel times correspond to the source depth $h = 25\,\text{km}$, whereas in Eq. (2.9) we assume $h = 0$. The velocities obtained for the individual epicentral distances are given in Tab. 2.2. It can be seen that the velocities increase with the increasing epicentral distance. This means that the velocities of both P and S waves increase with depth, which is, however, in contradiction with our initial assumption on the homogeneity of the medium.

If the velocities in the mantle increase with depth, the seismic rays are not rectilinear, but curved and convex downward; see Fig. 2.13 below. Consequently, a more advanced theory of wave propagation is needed to interpret the observed travel-time curves. Nevertheless, the velocities in Tab. 2.2 indicate approximately the range of velocities in the crust and mantle, namely the P-wave velocities $v_P \approx 7 - 12\,\text{km}\,\text{s}^{-1}$, and the S-wave velocities $v_S \approx 4 - 7\,\text{km}\,\text{s}^{-1}$. A more detailed distribution of these velocities is shown in the graphs at the end of Chapter 3.

Table 2.2. Interpretation of the travel times from Tab. 2.1 under the assumption of rectilinear seismic rays: Δ is the epicentral distance, v_P and v_S are the mean velocities of P and S waves.

$\Delta°$	$v_P\left(\text{km}\,\text{s}^{-1}\right)$	$v_S\left(\text{km}\,\text{s}^{-1}\right)$
10	7.5	4.3
20	8.1	4.5
30	8.9	5.0
40	9.6	5.3
50	10.1	5.6
60	10.5	5.8
70	10.9	6.0
80	11.2	6.2
90	11.5	6.3
100	11.8	6.4
103	11.9	6.5

In Fig. 2.11a we have considered a surface source only. The corresponding formulae may easily be generalised to a source at depth h (Fig. 2.11b). From the cosine theorem,

$$s^2 = R^2 + (R-h)^2 - 2R(R-h)\cos\Delta \ , \tag{2.10}$$

and using

$$1-\cos\Delta = 2\sin^2\frac{\Delta}{2} \ , \tag{2.11}$$

we get

$$s^2 = 4R^2\left(1-\frac{h}{R}\right)\sin^2\frac{\Delta}{2} + h^2 \ , \tag{2.12}$$

which is a generalisation of Eq. (2.7) for a non-zero source depth, h. The equation of the travel-time curve now takes the form

$$t = \frac{2R}{v}\sqrt{\left(1-\frac{h}{R}\right)\sin^2\frac{\Delta}{2} + \left(\frac{h}{2R}\right)^2} \ . \tag{2.13}$$

The application of this formula to the travel times in Tab. 2.1 yields the results which are very similar to those in Tab. 2.2, so that we shall not present them here.

Simplified estimates can also be obtained for the depth of the core-mantle boundary, still assuming the medium above the boundary to be homogeneous. If the seismic ray at the epicentral distance of 103° grazes this boundary, it follows from Fig. 2.11a that the radius of the core is $r_c = R\cos(103°/2) \approx 3970\,\mathrm{km}$. Consequently, for the depth of this discontinuity below the Earth's surface we get $d_c = R - r_c \approx 2400\,\mathrm{km}$. However, if we consider seismic rays curved downwards, this depth will be larger than 2 400 km.

Another simple estimate of the core depth follows from the travel times of seismic waves reflected at the core-mantle boundary. For example, the travel time of the *ScS* wave (the *S* wave which was reflected at the core-mantle boundary and returned back again as the *S* wave) is $15^{\mathrm{m}}30^{\mathrm{s}} = 930\,\mathrm{s}$ for $\Delta = 0°$ (Tab. 2.1). Considering the mean velocity of *S* waves in the mantle to be about $6.5\,\mathrm{km\,s^{-1}}$ (see the last column in Tab. 2.2), we get the length of the corresponding ray $s = 930 \times 6.5\,\mathrm{km} = 6045\,\mathrm{km}$. The depth of the core-mantle boundary is one half of this length, i.e. $d_c \approx 3000\,\mathrm{km}$. This estimate is already very close to the correct value of this depth, which is 2 900 km.

2.8 Principal Seismic Discontinuities in the Earth

In addition to the main branches of the travel-time curves for *P* and *S* waves, many other branches can be seen in Figs. 2.9 and 2.10. These further branches are connected mainly with the existence of discontinuities, at which seismic waves are reflected and transmitted (refracted). The main of these discontinuities are:

- the Earth's surface;
- the Mohorovicic discontinuity, separating the Earth's crust and mantle;

- the core-mantle boundary (CMB);
- the inner-core boundary (ICB), separating the outer and inner cores.

The seismic model of the Earth, marked by these seismic discontinuities, consists of the Earth's crust, mantle, outer core and inner core (Fig. 2.12).

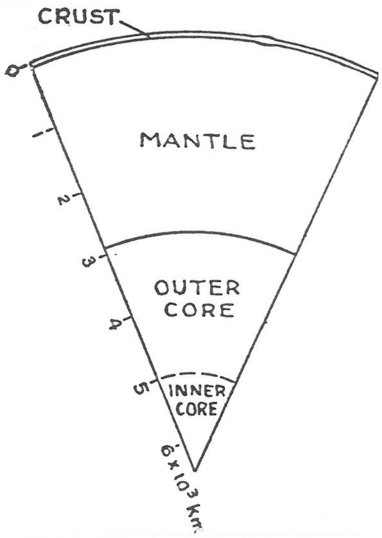


Fig. 2.12. Seismic model of the Earth. The depths are given in thousands of kilometres. (After Richter (1958)).

The *Mohorovicic discontinuity*, briefly called the *Moho*, marks the base of the Earth's crust. This subcrustal discontinuity was discovered by A. Mohorovicic in 1909 on the basis of seismograms of an earthquake with its epicentre in Croatia not far from his station at Agram (now Zagreb). The depth of the Moho on the continents is around 35 km, but large regional variations in this depth exist; from less than 25 km in some basins to more than 60 km in the regions of high mountains. This indicates the tendency of the Earth's crust to attain the isostatic equilibrium according to Airy's model (Heiskanen and Vening Meinesz, 1958; Novotny, 1998). The Moho is seismically less pronounced than the Earth's surface or the CMB, i.e. the waves reflected from the Moho are weaker than reflections from the Earth's surface or from the CMB.

The notion of a "thin" crust overlying a molten interior of the Earth was current and popular for a long time till the second half of the 19th century. This notion had to be modified when the measurements of tidal deformations of the Earth and measurements of the motions of the poles (variations of latitudes) allowed to calculate the mean rigidity for the whole Earth. It appeared that the Earth has a mean rigidity of the same order as that of steel. This means that a substantial part of the Earth must be solid. Nevertheless, the mean rigidity is small. If the whole Earth were composed of material similar to the rocks of the crust, its mean rigidity would be much larger than the observed mean. Wiechert concluded than there must be a large interior part of the Earth, the properties of which should approach those of a liquid, and it might be in the form of a central core. However, the radius of this core remained undetermined. The first, sufficiently accurate value of this radius was obtained by B. Gutenberg in 1913 at Göttingen on the basis of seismic observations. He found that the depth of the surface of the core is very close to 2 900 km. Seismic studies also brought direct evidence that the core is at least partly fluid, as expected by Wiechert, since it does not transmit transverse waves.

The existence of another important boundary below the 5 000 km depth was suggested by Miss I. Lehmann in 1936 in Copenhagen. Lehmann's paper; called P', has probably the shortest title in science. This discovery perhaps would have rated a Nobel Prize in physics for the analogous detection of a new atomic particle (Bolt, 1988). Additional seismological work in the next decades has indicated that the inner core is a solid body as opposed to the liquid outer core. This conclusion follows from investigations of seismic body waves and, independently, also from observations of the periods of free oscillations of the Earth.

2.9 Principal Types of Seismic Waves Propagating within the Earth

In this section we shall restrict ourselves to body waves of teleseisms. We shall not distinguish here the Earth's crust, but we shall consider it as a part of the mantle. The individual waves connected with the discontinuities within the Earth's crust will be discussed in the next Section 2.10.

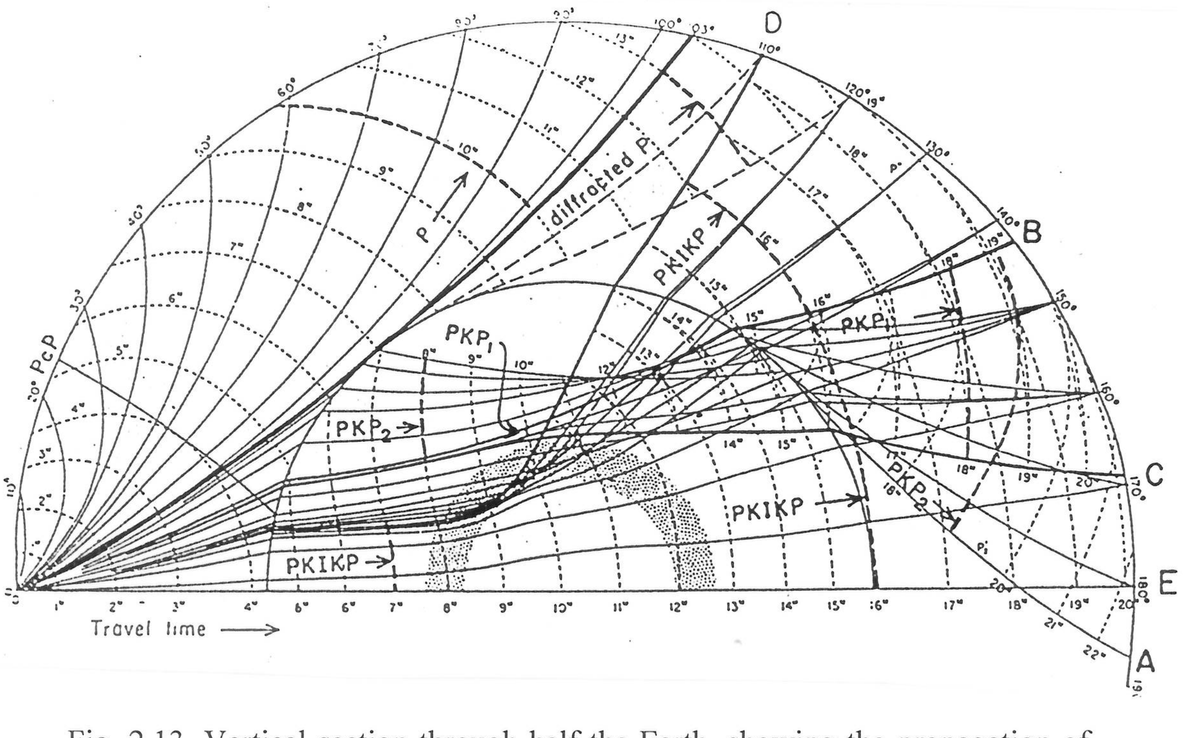


Fig. 2.13. Vertical section through half the Earth, showing the propagation of the longitudinal waves from a source in the left corner of the figure. The core-mantle boundary is marked by a circular arc at a depth of 2 900 km and the transition to the inner core by a shaded zone at about 5 000 km depth. (After Gutenberg; from Båth (1979)).

The various individual wave groups on seismograms represent direct waves and various combinations of reflections and refractions. Each wave is denoted by a particular combination of symbols, chiefly letters, which stand for the successive segments of the ray in order from source to station. Some of the possible ray paths for seismic waves penetrating the Earth are shown in Fig. 2.13 and in simplified forms in Figs. 2.14 and 2.15. In the mantle and core, the velocities increase with depth, so that the rays bend away from the normal. However, the decrease of the velocity at the mantle-core boundary causes that the rays refracted into the core are bent towards the normal.

The notation of seismic waves went through certain historical development, which should briefly be mentioned. The type of a wave along any segment was denoted by P for longitudinal and S for transverse waves. Thus, the direct longitudinal and transverse waves in the mantle (their rays are convex downwards) were denoted simply as phases P and S, respectively. Reflection at the surface of the Earth was indicated simply by the succession of the chief symbols, *PP*, *SS*, *PS* and *SP*; see Fig. 2.14. Multiple reflections at the surface were denoted by *PPP*, *SSS*, etc. This notation of seismic waves in the mantle has been retained up to now.

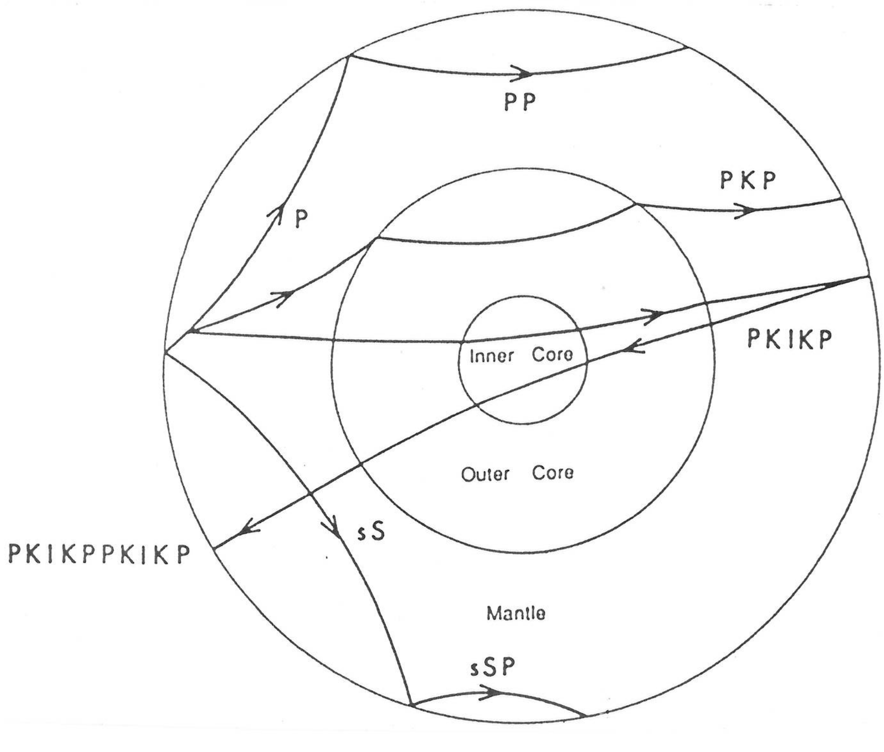


Fig. 2.14. Some of the possible ray paths for seismic waves penetrating the Earth. (After Fowler (1994)).

All incidences of a wave at the surface of the core were originally denoted by the small letter c, which was included into the code of the wave. Consequently, a longitudinal wave which passed down through the mantle and was reflected at the core surface back into the mantle as a longitudinal wave was designated *PcP*. A transverse wave reflected back as a transverse wave was designated *ScS*. For converted waves reflected at the core surface, we have analogous notations *PcS* and *ScP* (a wave is called a converted one if it changes at an interface from a longitudinal wave into a transverse wave or vice versa). These notations are also used at present.

Now, consider waves penetrating into the outer core. A longitudinal wave which passed down through the mantle and the outer core and then up through the mantle was originally denoted, according to the above-mentioned rule, by *PcPcP*. The other similar combinations are as follows (note that only *P* waves can propagate in the liquid outer core): *ScPcS*, *PcPcS* and *ScPcP*. These notations are logical and sufficiently elementary, but rather long and not easy to be quickly decoded. Since the letter *P* surrounded by letters *c* from left and right denotes a longitudinal wave in the outer core, the group *cPc* was later replaced by one letter, *K* (*Kern* from the German, meaning core). Therefore, the above notations are now abbreviated as follows: *PKP*, *SKS*, *PKS* and *SKP*. Analogous separate notations were later approved officially also for the longitudinal and transverse waves passing through the inner core.

2.9.1 Direct and transmitted waves

Consequently, the present situation in denoting the longitudinal and transverse waves within the Earth can be summarised as follows (Tab. 2.3):

- letters *P* and *S* are only used to denote the waves in the mantle;
- *K* denotes a longitudinal wave in the outer core (transverse waves do not propagate there);
- *I* and *J* denote longitudinal and transverse waves in the inner core.

Table 2.3. Notation of longitudinal and transverse waves in different parts of the Earth.

	mantle	outer core	inner core
longitudinal wave	*P*	*K*	*I*
transverse wave	*S*	–	*J*

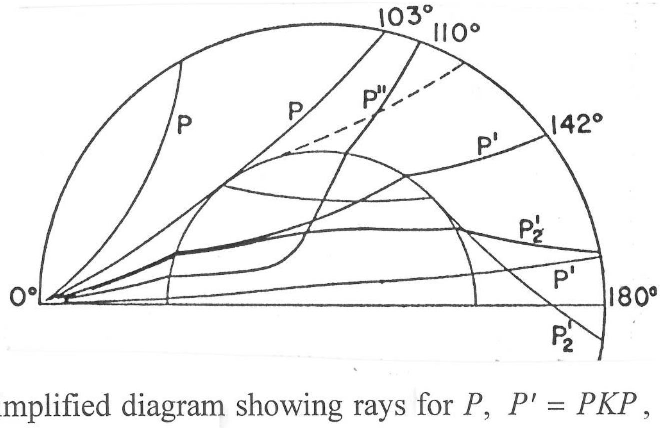


Fig. 2. 15. Simplified diagram showing rays for *P*, $P' = PKP$, $P'' = PKIKP$. (After Richter (1958)).

Thus, the waves passing through the inner core are designated *PKIKP*, *PKJKP*, etc. Still briefer notations have also been introduces, such as P' for *PKP* and P'' for *PKIKP*, see Fig. 2.15. However, now these very brief notations are not usually used.

2.9.2 Reflected waves

As mentioned above, the reflections from the outer side of the core are denoted by the letter *c*, which is inserted into the code of the corresponding wave. Analogously, letter *i* is used to denote reflections from the outer side of the inner core (Figs. 2.9, 2.10 and Tab. 2.4). Moreover, reflections from the upper side of the Moho are denoted by letter *M*. For example, the wave $P^{M}p$ is the longitudinal wave reflected from the Moho. Note that this wave has frequently been used to determine the thickness of the crust in deep seismic soundings.

Table 2.4. Waves reflected only once from the outer sides of the main discontinuities: CMB is the core-mantle boundary, ICB is the inner-core boundary, and the symbol is the letter included into the wave code.

Discontinuity	Symbol	Examples
CMB	*c*	*PcP*, *ScS*
ICB	*i*	*PKiKP*, *SKiKS*

Reflections from the inner sides of the discontinuities are not indicated by interposing specific letters. These reflections are evident from the succession of the chief symbols (Fig. 2.14 and Tab. 2.5).

Table 2.5. Waves reflected only once from the inner sides of the main discontinuities.

Discontinuity	Examples
Earth's surface	*PP*, *SS*, *PS*, *SP*, *SKSP*
CMB	*PKKP*, *SKKS*, *PKKS*, *SKKP*
ICB	*PKIIKP*, *SKIIKS*

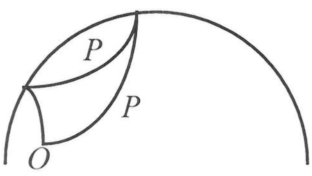


Fig. 2.16. Waves *P* and *pP* propagating from source *O*.

In addition to the reflections already mentioned, special types of reflections at the surface of the Earth are denoted by small letters *p* and *s*, such as *pP*, *pS*, *sP* or *sS*, see Figs. 2.14 and 2.16. The wave *pP* travelled up from the focus as a *P* wave, was reflected at the Earth's surface close to the focus and continued as a *P* wave to the observer (Fig. 2.16). As opposed to it, wave *PP* started to travel down from the focus and was reflected at approximately half a way between the focus and observer (Fig. 2.14). For shallow earthquakes, it is usually difficult to distinguish the phase *pP* from phase *P*. However, for deep earthquakes, these phases can be distinguished, and the difference of their travel times is frequently used to determine the focal depth.

2.10 Seismic Waves at Short Epicentral Distances

At short epicentral distances, the wavefield is influenced significantly by the structure of the crust, and further branches of the travel-time curves may be distinguished. There is a critical epicentral distance, generally in the range from 100 to 150 km, where seismograms change their character. Stations at shorter distances register the *P* and *S* waves as initial sharp phases followed by smaller motion. Beyond the critical distance, the order of theses phases is reversed. The *P* and *S* waves begin there with relatively small and long-period motion, designated P_n and S_n, which is followed by a larger and shaper impulse of a shorter period. Mohorovicic found that the travel-time curve of P_n is continuous with that for the *P* wave at teleseismic distances (continuous with the normal *P* wave, which substantiates the notation by P_n). The sharp phase was called him $\overline{P}$. The apparent velocity of $\overline{P}$ is about 6 km s^{-1}, of P_n about 8 km s^{-1} (the velocities found in 1909 were lower).

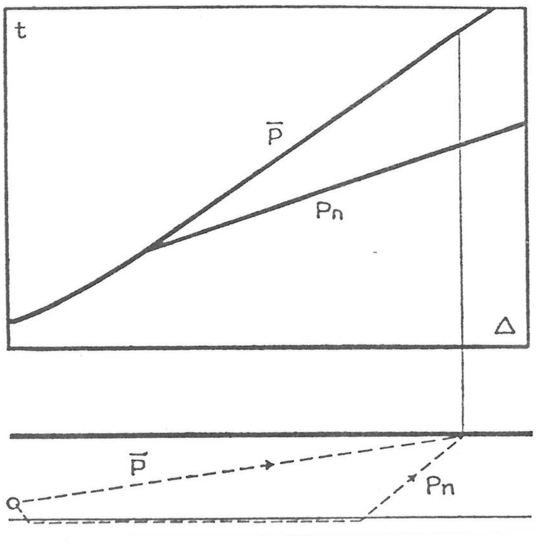


Fig. 2.17. Travel-time curves of $\overline{P}$ and P_n, and their interpretation according to Mohorovicic. (After Richter (1958)).

The simplest form of the explanation adopted by Mohorovicic is shown in Fig. 2.17. Neglecting the curvature of the Earth, consider a layer of constant velocity v_1, overlying a half-space of constant velocity v_2, where $v_1 < v_2$. In our case, $v_1 = 6 \text{ km s}^{-1}$ and $v_2 = 8 \text{ km s}^{-1}$. The wave $\overline{P}$ is interpreted as the direct wave propagating to the recording station (through the medium of the lower velocity). The wave P_n is refracted horizontally below the discontinuity, where the velocity is higher. Consequently, there is a critical distance beyond which the higher velocity compensates for the longer path, and P_n arrives before $\overline{P}$. Since the angle of refraction is 90°, the angle of incidence above the boundary, *i*, is given by $\sin i = v_1/v_2 = 0.75$, whence $i = 49°$ approximately. The theoretical travel-time curve for P_n is a straight line with slope $1/v_2$. The travel-time curve for $\overline{P}$ is a hyperbola asymptotic (at large epicentral distances Δ) to a straight line through the origin with slope $1/v_1$.

It should be, however, noted that the explanation of the refraction like that postulated for P_n has serious theoretical complications. Such waves, which propagate along the interface in the faster medium and radiate their energy back into the slower medium, really exist but their theory is rather complicated. They are called the *head waves*; we shall return to this problem in the next section.

Later studies revealed further complexity of seismograms at short epicentral distances, which led to subdividing the crust into more layers. In Europe, Conrad observed a small sharp impulse between P_n and $\overline{P}$ with the apparent velocity near 6.5 km s^{-1}, which he named P^*. The corresponding discontinuity is referred to as the Conrad discontinuity. Jeffreys and others accepted the Conrad discontinuity as separating predominantly granitic layer above it from a basaltic layer. According to this interpretation, $\overline{P}$ propagates in the granitic layer, and P^* is connected with the basaltic layer. Consequently, an alternative notation, P_g, was also proposed for $\overline{P}$, and P_b for P^*. A schematic representation of these notions is shown in Fig. 2.18. Note that analogues to $\overline{P}$, P^* and P_n have also been found in the *S*-wave group of seismograms.

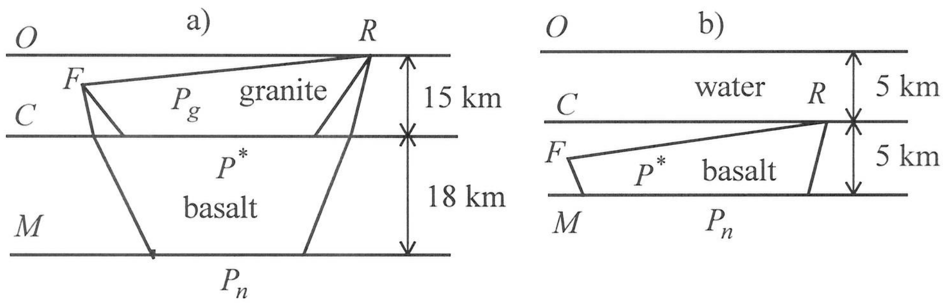


Fig. 2.18. Principles of wave propagation in the classical models of the continental crust (a) and the oceanic crust (b): *F* is the focus, *R* the receiver, *O* the Earth's surface, *C* the Conrad discontinuity, *M* the Mohorovicic discontinuity. (Modified after Båth (1979)).

Later studies of the crustal structure have not confirmed the existence of the Conrad discontinuity in many regions. Also the petrological division of the crust into the granitic and basaltic layers seems to be oversimplified and far from reality. Nevertheless, significant differences have been found in seismic properties of the upper and lower crust. Seismic reflection methods have revealed that the upper crust is rather transparent for seismic waves, whereas the lower crust is more reflective. This can be explained by different physical properties of the upper and lower crust. The *upper crust* is considered to be *brittle*, which is also in agreement with the appearance of many earthquakes in this part of the crust. As opposed to it, the *lower crust* is considered to be *ductile*, containing sliding surfaces which reflect seismic waves.

2.11 Head Waves

The P_n wave, as show in Fig. 2.17, belongs to a special category of waves which are referred to as head waves. These waves cannot be explained in the framework of the geometric theories of wave propagation, such as geometric optics or geometric seismic (as the zero-approximations of the ray theory). Head waves do not also exist in any theory of plane waves.

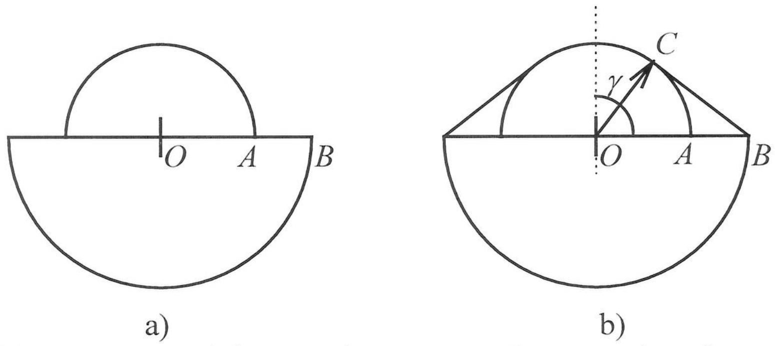


Fig. 2.19. Waves generated by a point source, O, at an interface of two homogeneous media: a) wave surfaces in the standard ray approximation; b) wave surfaces in exact wave theories, including the head wave.

In order to explain head waves, we must consider spherical waves. To simply the problem, consider a point source located at an interface of two homogeneous media (Fig. 2.19). Along the opposite sides of the interface, the waves propagate at different velocities. At a certain time, the wave propagating in the slower (upper) medium arrived at point A, whereas the wave in the faster (lower) medium arrived at more distant point B. The corresponding wave surfaces are shown in Fig 2.19a. A rather complicated situation occurs along the abscissa AB. According to the standard ray theory, the medium under AB moves but the medium over AB remains at rest, because (in this simple ray theory) the energy of seismic waves does not propagate perpendicularly to the rays "tubes".

The situation along abscissa AB in Fig 2.19a is, however, inconsistent with the usual physical notions, which require the continuity of the displacement and stress at the interface. Also according to Huygen's principle, each point of the wave surface becomes a source of elementary waves. Hence, according to this principle, the disturbance propagating along OB generates elementary waves which propagate also into the upper medium. The envelope of these waves forms a new wave, termed the *head wave*, having a conical wave surface as shown in Fig. 2.19b by abscissa *BC* (and by the symmetrical abscissa in the left part of the figure).

It can easily be shown that angle γ between the ray of the head wave and the normal to the interface is the critical angle. This follows from triangle *OBC*, becouse

$$\sin\gamma = \frac{OC}{OB} = \frac{v_1}{v_2}, \tag{2.14}$$

where v_1 and v_2 are the velocities in the upper and lower media, respectively. Formula (2.14) is the special form of Snell's law for the critical angle.

Head waves can be described in the framework of higher approximations of the ray theory or other exact wave theories. Although the travel-time curve of the P_n wave can be interpreted as that of a head wave, some inconsistencies are encountered in their amplitudes. Namely, the observed P_n waves are stronger than the theoretical head waves generated at an interface of two homogeneous media. In order to overcome this problem, more complicated models of the medium must be considered. For example, if the velocity under the discontinuity increases slightly with depth, an interference head wave of a higher intensity appears in such a medium, which can explain the observations. Also some other waves propagating in vertically inhomogeneous media have properties similar to those of P_n waves, in particular refracted waves; see the next chapter.

Chapter 3

Simple Ray Theory, Based on Fermat's Principle

The contemporary theory of elastic waves is based on the theory of elasticity, whose basic equations were derived by Navier, Cauchy and Poisson in the 1820^{th}; for details we refer the reader to the review in Novotny (1999). The theory of light propagation is based on the theory of the electromagnetic field, which was formulated by Maxwell even later, in 1873. The earlier theories of wave phenomena had, therefore, to start from other, simpler laws. The most important of them is Fermat's principle.

3.1 Fermat's Principle

According to the original formulation from the 17^{th} century, *Fermat's principle* states that a disturbance propagates between two points along the path for which the travel time is minimum:

$$\boxed{\int_A^B \frac{\mathrm{d}s}{v} = \min.}\,, \tag{3.1}$$

where ds is a path element and v the corresponding velocity. A generalised form of Fermat's principle will be given in the next section, but the simple form (3.1) will be sufficient in many cases.

Fermat's principle lost much of its previous importance when the more general theories, the theory of elasticity and the theory of the electromagnetic field, were worked out. Namely, Fermat's principle can be derived from these general theories as a special relation, valid under certain conditions. We shall deal with these problems later on. Nevertheless, in solving some special problems, Fermat's principle still represents a valuable basic relation. Its advantages are as follows:

- Relative simplicity of the basic equations. We shall see that Euler's equations for the extremal of Fermat's functional are much simpler than the equations of motion of a continuum.
- Fermat's principle can be applied to very general models of the medium, including complicated three-dimensional models.
- Universal applicability of Fermat's principle to wave phenomena in various branches of physics. The equations for rays and travel times, which will be derived in this and the next chapters, can be applied not only to seismic waves, but also to acoustic waves in the atmosphere, acoustic waves in the ocean, propagation of light, radio waves, and to some other problems.

On the other hand, Fermat's principle has also serious disadvantages, in particular:

- This principle does not describe the wave field in full. It describes only some so-called kinematic characteristics of the wave field (rays, travel times), but not dynamic characteristics (amplitudes, polarisation). In order to estimate amplitudes, additional "laws" or rules must be added to Fermat's principle. We shall discuss them at the end of this chapter.

- Fermat's principle describes the propagation of a separate disturbance, assuming its velocity is known. However, we encounter also more complicated problems when the velocity must be determined at the beginning (waves in anisotropic media, and others), or the wave field is of a complicated chapter. For example, one velocity is not sufficient to describe the propagation of dispersive waves, but two velocities, the phase and group velocity, are needed. However, only one velocity is considered in Fermat's principle.

3.2 Fermat's and Hamilton's Principles

Generally, many various waves may arrive at a receiver, such as a direct wave, refracted waves, reflected waves and others. To describe this situation, Fermat's principle (3.1) must be generalised. We shall formulate Fermat's principle in the following generalised form.

A disturbance propagates between two points along the paths for which the travel time is stationary, i.e. for which the variation of integral (3.1) is zero:

$$\delta \int_A^B \frac{\mathrm{d}s}{v} = 0 \ . \tag{3.2}$$

The curve which satisfies this condition is called the extremal of Fermat's functional. This curve can be obtained as a solution of a system of differential equations known as Euler's equation. Let us give a general formulation of these equations.

Consider a function $F = F\left(x, y_j, y'_j\right)$ which is a function of an independent variable x and of certain functions $y_j = y_j(x)$ and of their derivatives $y'_j = \dfrac{\mathrm{d}y_j(x)}{\mathrm{d}x}$, where $j = 1, 2, \ldots, n$. For the functional

$$I = \int_A^B F\left(x, y_j, y'_j\right) \mathrm{d}x \ , \tag{3.3}$$

its extremal satisfies the following equations:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'_j}\right) - \frac{\partial F}{\partial y_j} = 0 \ . \tag{3.4}$$

These equations are referred to as *Euler's equations*. Before deriving this important theorem, let us remind certain analogies with analytical mechanics.

A very general principle of mechanics, *Hamilton's principle*, states that a system of particles moves in such a way that the corresponding action,

$$\int_{t_1}^{t_2} L(q_j, \dot{q}_j, t)\,\mathrm{d}t \;, \tag{3.5}$$

is stationary, i.e.

$$\delta \int_{t_1}^{t_2} L(q_j, \dot{q}_j, t)\,\mathrm{d}t = 0 \;; \tag{3.6}$$

here $L = T - V$ is the Lagrangian, T the kinetic energy, V the potential energy, $q_j = q_j(t)$ are generalised coordinates, $\dot{q}_j = \mathrm{d}q_j/\mathrm{d}t$ are generalised velocities, t is the time, and $j = 1, 2, \ldots, n$, n being the number of degrees of freedom. It is well known from analytical mechanics that the extremals of functional (3.5), i.e. the trajectories of particles, satisfy *Lagrange's equations* of the second kind:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0 \;. \tag{3.7}$$

Since functionals (3.3) and (3.5) are similar, the validity of Euler's equations (3.4) immediately follows from Lagrange's equations (3.7).

It should be noted that the analogy between Fermat's and Hamilton's principles is not only accidental. Indeed the apparatus of analytical mechanics can be extended to continuum mechanics, but we shall not use this complicated theory here.

In this chapter we shall draw main attention to the equations of seismic rays. For a given model of the medium, these equations can be obtained directly from Fermat's principle, using Euler's equations. This approach will be used systematically in the next chapter. However, for the special models of the medium which will be considered in this chapter, we shall also use a more elementary method. At first, we shall use Fermat's principle to derive Snell's law, and then we shall derive the equations of rays from Snell's law. This elementary approach does not require the knowledge of the calculus of variations, or analytical mechanics.

Nevertheless, the application of Euler's equations represents a general method of solving many problems. Therefore, we shall begin with their derivation.

3.3 Derivation of Euler's equation for an extremal

To simplify the problem which was formulated above, consider function $F = F(x, y, y')$ to be a function of an independent variable x, and only of one unknown function $y(x)$ and of its derivative $y'(x)=\frac{\mathrm{d}\,y(x)}{\mathrm{d}\,x}$. Let function F be continuous with all derivatives which will be needed.

Thus, consider the integral

$$I=\int_{x_1}^{x_2} F(x,y,y')\,\mathrm{d}x \ , \tag{3.8}$$

where F is a *known* function of the indicated variables x, y, and y', but the dependence on y is not fixed, i.e. function $y(x)$ is *unknown* (Arfken, 1970). This means that although the integral is from x_1 to x_2 the exact path of integration is not known (Fig 3.1).

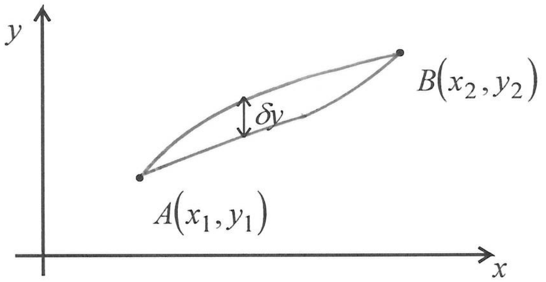


Fig. 3.1. A varied path.

We are seeking the path of integration between points A and B to minimise integral I. Strictly speaking, we shall determine extreme values of I, i.e. minima, maxima or saddle points. Denote by $y_0=y_0(x)$ the unknown path for which I is an extremum.

Compare the value of I for our (unknown) optimum path with the values obtained from neighbouring paths. Two possible paths (from an infinite number of possibilities) are shown in Fig. 3.1. The difference between two paths for a given x is called the variations of y, and denoted by δy.

Consider an arbitrary path between point A and B, and denote the difference between this path and extremal $y_0(x)$ by $\eta(x)$. Moreover, consider all paths with similar deviations from the extremal, which are described by the equation

$$y(x,\alpha)=y_0(x)+\alpha\eta(x) \ , \tag{3.9}$$

α being a parameter. We assume function $\eta(x)$ to be arbitrary except for two restrictions:

- $\eta(x)$ is differentiable;
- $\eta(x_1)=\eta(x_2)=0$. (3.10)

The extremal is the curve for $\alpha=0$, i.e. $y(x,\alpha=0)=y_0(x)$. Integrals (3.8) for these paths are now functions of our new parameter α:

$$I(\alpha)=\int_{x_1}^{x_2} F\left[x, y(x,\alpha), y'(x,\alpha)\right] \mathrm{d}x \ . \tag{3.11}$$

In other words, for a fixed function $\eta(x)$, curves (3.9) and integrals (3.11) are functions of parameter α only. Consequently, the condition for the extremal among these curves can be expressed as the condition for an extremum of function $I(\alpha)$ in differential calculus, i.e.

$$\left[\frac{\mathrm{d}\, I(\alpha)}{\mathrm{d}\,\alpha}\right]_{\alpha=0}=0 \ . \tag{3.12}$$

The dependence of integral (3.11) on α is contained in $y(x,\alpha)$ and $y'(x,\alpha)$; note that y and y' are treated as independent variables in function F. Therefore,

$$\frac{\mathrm{d}\, I(\alpha)}{\mathrm{d}\,\alpha}=\int_{x_1}^{x_2}\left[\frac{\partial F}{\partial y}\frac{\mathrm{d}\, y}{\mathrm{d}\,\alpha}+\frac{\partial F}{\partial y'}\frac{\mathrm{d}\, y'}{\mathrm{d}\,\alpha}\right]\mathrm{d}x \ . \tag{3.13}$$

It follows from Eq. (3.9) that

$$\frac{\mathrm{d}\, y}{\mathrm{d}\,\alpha}=\frac{\partial y}{\partial \alpha}=\eta(x) \, , \quad \frac{\mathrm{d}\, y'}{\mathrm{d}\,\alpha}=\frac{\mathrm{d}\,\eta(x)}{\mathrm{d}\, x} \ . \tag{3.14}$$

Equation (3.13) then becomes

$$\frac{\mathrm{d}\, I(\alpha)}{\mathrm{d}\,\alpha}=\int_{x_1}^{x_2}\left[\frac{\partial F}{\partial y}\eta(x)+\frac{\partial F}{\partial y'}\frac{\mathrm{d}\,\eta(x)}{\mathrm{d}\, x}\right]\mathrm{d}x \ . \tag{3.15}$$

Integrating the second term by parts, one gets

$$\int_{x_1}^{x_2}\frac{\mathrm{d}\,\eta(x)}{\mathrm{d}\, x}\frac{\partial F}{\partial y'}\mathrm{d}x=\eta(x)\frac{\partial F}{\partial y'}\bigg|_{x_1}^{x_2}-\int_{x_1}^{x_2}\eta(x)\frac{\mathrm{d}}{\mathrm{d}\, x}\frac{\partial F}{\partial y'}\mathrm{d}x \ . \tag{3.16}$$

The first term on the right-hand side vanishes by (3.10) and Eq. (3.12) becomes

$$\int_{x_1}^{x_2}\left(\frac{\partial F}{\partial y}-\frac{\mathrm{d}}{\mathrm{d}\, x}\frac{\partial F}{\partial y'}\right)\eta(x)\,\mathrm{d}x=0 \ . \tag{3.17}$$

In this form α has been set equal to zero.

The process just described can be repeated for other forms of function $\eta(x)$. Since $\eta(x)$ is arbitrary, Eq. (3.17) will be satisfied if the bracketed term itself is identically zero:

$$\boxed{\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'} = 0}\,. \tag{3.18}$$

Namely, assume the bracketed term to be non-zero, e.g. positive, at a point x^*. This term is then positive also in a certain vicinity of x^* as the consequence of the continuity of the corresponding functions. Choose function $\eta(x)$ to be positive in this interval and equal to zero outside. The left-hand side of Eq. (3.17) is then positive, which contradicts the right-hand side. Therefore, the bracketed term cannot be non-zero, but must satisfy Eq. (3.18). This equation represents the condition for our extremum. This is a partial differential equation, known as *Euler's equation.*

3.4 Derivation of Snell's Law from Fermat's Principle

At first we must derive the trivial fact that the rays in a homogeneous medium are straight lines. Thus, consider the velocity, v, to be constant. Fermat's principle (3.1) can then be expressed as

$$\frac{1}{v}\int_A^B \mathrm{d}s = \min. \tag{3.19}$$

It is evident that this integral is minimum if the curve connecting points A and B is an abscissa.

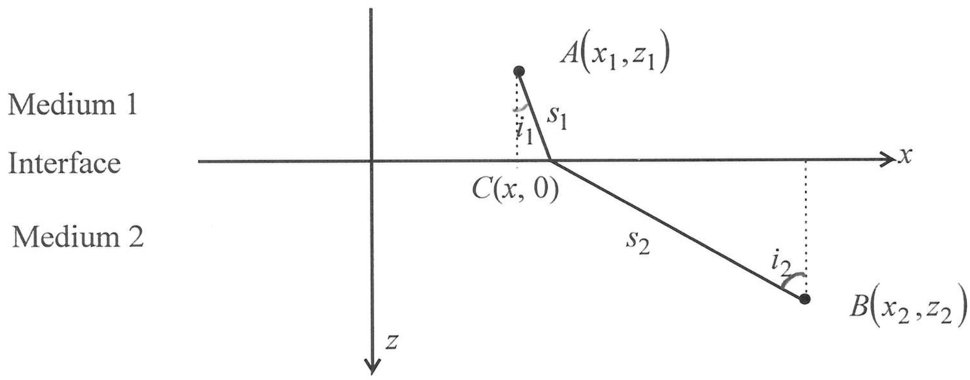


Fig. 3.2. A path connecting two points in different media.

Now, consider two homogeneous media separated by a plane interface (Fig 3.2). Denote the corresponding velocities by v_1 and v_2, respectively. Consider a fixed point $A(x_1, z_1)$ in the first medium (z_1 is negative), and a fixed point $B(x_2, z_2)$ in the second medium. Choose arbitrarily point $C(x, 0)$ at the interface, and calculate the travel time t along path ABC:

$$t = t(x) = \frac{s_1}{v_1} + \frac{s_2}{v_2} = \frac{\sqrt{(x - x_1)^2 + z_1^2}}{v_1} + \frac{\sqrt{(x - x_2)^2 + z_2^2}}{v_2} . \tag{3.20}$$

This time is minimum if its derivative is zero:

$$0 = \frac{\mathrm{d}t}{\mathrm{d}x} = \frac{x - x_1}{v_1 s_1} + \frac{x - x_2}{v_2 s_2} = \frac{\sin i_1}{v_1} - \frac{\sin i_2}{v_2} , \tag{3.21}$$

where we assume $x_1 \le x \le x_2$, as show in Fig. 3.2. Thus, it follows from Fermat's principle that Snell's law must hold true at an interface of two homogeneous media.

3.5 Seismic Rays and Travel Times in a Horizontally Layered Medium

Consider a medium which is composed of homogeneous, isotropic and parallel layers, as shown in Fig. 3.3. Denote the velocity and thickness of the m-th layer by v_m and d_m, respectively. The velocity v_m is the velocity of P waves or S waves according to the type of the corresponding seismic wave.

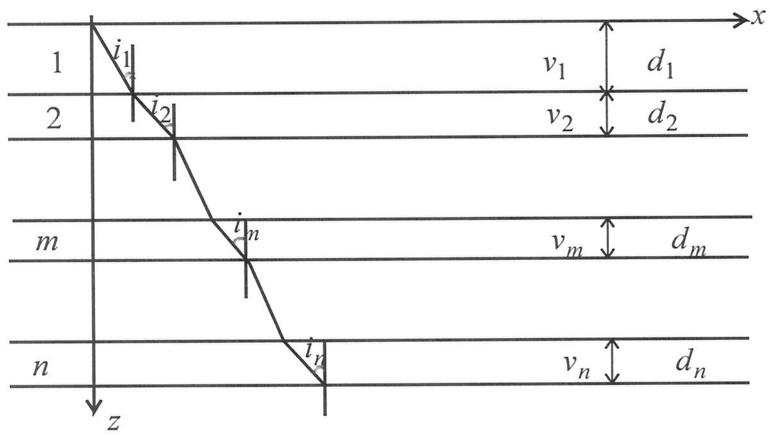


Fig. 3.3. Model of a layered medium and a seismic ray.

Let us now consider a wave without any reflection point (only transmissions). Denoting the angle of incidence in the m-th layer by i_m, we have from Snell's law that

$$\frac{\sin i_1}{v_1} = \frac{\sin i_2}{v_2} = \cdots = \frac{\sin i_m}{v_m} = \cdots = p . \tag{3.22}$$

The quantity p remains constant along the whole ray. We call it the *ray parameter*.

The inverse of p,

$$c = \frac{1}{p} = \frac{v_m}{\sin i_m} \tag{3.23}$$

is called the *apparent velocity*. It represents the velocity of propagation of the wavefront along the x-axis (horizontal axis); see more detailed discussions below. In the theory of surface waves, the velocity c represents the phase velocity of the corresponding surface wave.

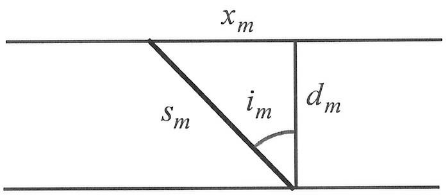


Fig. 3.4. Passage of a wave through the m-th layer.

Consider the passage of the wave through the m-th layer at an angle of incidence i_m. The corresponding increment of the epicentral distance, x_m, and of the travel time, τ_m, are as follows (Fig. 3.4):

$$x_m = d_m \tan i_m = d_m \frac{\sin i_m}{\cos i_m} = d_m \frac{p v_m}{\sqrt{1 - p^2 v_m^2}} ,$$

$$\tau_m = \frac{s_m}{v_m} = \frac{d_m}{v_m \cos i_m} = \frac{d_m}{v_m \sqrt{1 - p^2 v_m^2}} . \tag{3.24}$$

If the wave passes through the first to the n-th layer, we must sum up the individual contribution:

$$x(z_n, p) = \sum_{m=1}^{n} x_n = p \sum_{m=1}^{n} \frac{v_m d_m}{\sqrt{1 - p^2 v_m^2}} ,$$

$$\tau(z_n, p) = \sum_{m=1}^{n} \tau_m = \sum_{m=1}^{n} \frac{d_m}{v_m \sqrt{1 - p^2 v_m^2}} . \tag{3.25}$$

For the wave which is reflected at the bottom of the n-th layer back to the surface, if the source and receiver are on the surface, we obtain the value two times greater:

$$x(p) = 2p \sum_{m=1}^{n} \frac{v_m d_m}{\sqrt{1 - p^2 v_m^2}} , \quad \tau(p) = 2 \sum_{m=1}^{n} \frac{d_m}{v_m \sqrt{1 - p^2 v_m^2}} . \tag{3.26a,b}$$

The computation of the travel time of the reflected wave at a given epicentral distance x^* proceeds as follows. We choose an incidence angle i_1 in the first layer, compute the ray parameter p using (3.22), and insert into (3.26a). Since the computed epicentral distance $x(p)$ does not generally agree with the given value x^*, we must change the value of i_1 (or directly p). We repeat this process until $x(p)$ is close to x^* with the required accuracy. Finally, the value of p, found in this way, is inserted into (3.26b) in order to obtain the travel time.

This example demonstrates the usual situation which we encounter in computing theoretical travel-time curves. Let us, therefore, describe briefly the general situation.

For a given model of the medium, the travel-time curve is the dependence of travel time t on epicentral distance x:

$$t = t(x) \ . \tag{3.27}$$

However, this function can be expressed explicitly only in very simple cases. Usually, the equation of a travel-time curve is expressed in a parametric form as

$$t = \tau(p) \quad , \quad x = x(p) \ . \qquad \text{(3.28a,b)}$$

If we are to determine the travel time for a given distance x, we must first solve equation (3.28b), usually by some numerical method, i.e. to compute the inverse function $p = p(x)$. By inserting this value into (3.28a) we arrive at the desired travel time, $t = \tau(p(x))$. In this way we eliminate the auxiliary parameter p.

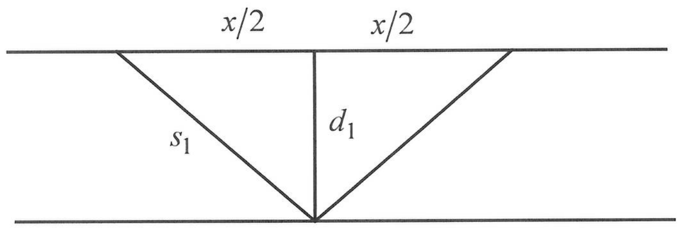


Fig. 3.5. Reflected wave in a single layer.

The travel-time curve for a reflected wave can be expressed in an explicit form only in the case of one layer, $n = 1$. In this case, parameter p can be expressed explicitly from Eq. (3.26a), and Eq. (3.26b) then yields

$$t = \frac{\sqrt{x^2 + 4d_1^2}}{v_1} \ . \tag{3.29}$$

Note that this equation also follows immediately from Fig. 3.5, because

$$t = \frac{2s_1}{v_1} = \frac{2}{v_1}\sqrt{(x/2)^2 + d_1^2} \ .$$

3.6 Seismic Rays and Travel Times in a Vertically Inhomogeneous Medium

In many applications, the Earth may be considered as a spherically symmetric body, i.e. as a body in which the physical parameters are dependent only on the radical distance from the centre; see the next chapter. If seismic waves do not penetrate to large depths, even the Earth's curvature may be neglected and the physical parameters may be considered as functions of one Cartesian coordinate only (depth). In this section we shall study seismic rays and travel-time curves for such a medium. A special case, when this medium is approximated by a system of homogeneous layers, has already been discussed above.

Hence, consider a medium in which the velocity is a function of only one Cartesian coordinate, $v = v(z)$. Since coordinate z will usually represent the depth, we shall speak of a vertically inhomogeneous medium.

3.6.1 Generalisation of formulae for a layered medium

A vertically inhomogeneous medium can be obtained as a limiting case of a layered medium, which was considered in Section 3.5, if the thicknesses of the homogeneous layers become infinitesimal. Consequently, all the results from Section 3.5 can be generalised to the vertically inhomogeneous medium. In particular, by generalising Eq. (3.22) we arrive at Snell's law for a vertically inhomogeneous medium in the form

$$\boxed{\frac{\sin i(z)}{v(z)} = p}\ , \qquad (3.30)$$

where the ray parameter p is a constant along the ray. Therefore, Eq. (3.30) represents the *equation of a seismic ray* in a vertically inhomogeneous medium. For a given initial point and a given value of p, this equation determines the form of the corresponding ray.

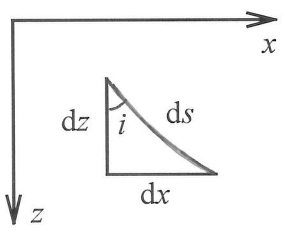


Fig. 3.6. Infinitesimal segment of a seismic ray.

Consider an infinitesimal part of a seismic ray in a vertically inhomogeneous medium (Fig. 3.6). When a wave has passed through a path ds, the corresponding increments in the epicentral distance, dx, and in the travel time, dτ, are

$$\mathrm{d}x = \mathrm{d}z \tan i(z) = \mathrm{d}z \frac{\sin i}{\cos i} = \mathrm{d}z \frac{pv}{\sqrt{1-p^2v^2}} \ ,$$

$$\mathrm{d}\tau = \frac{\mathrm{d}z}{v(z)\cos i(z)} = \frac{\mathrm{d}z}{v\sqrt{1-p^2v^2}} \ . \tag{3.31}$$

The other formulae are summarised in the left column of Tab 3.1 at the end of this chapter. The last formulae in the table are the parametric equations of the travel-time curve for the case of the source and receiver on the surface of the model:

$$x(p) = 2p \int_0^{z_p} \frac{v\,\mathrm{d}z}{\sqrt{1-p^2v^2}} \ , \quad \tau(p) = 2\int_0^{z_p} \frac{\mathrm{d}z}{v\sqrt{1-p^2v^2}} \ ; \tag{3.32}$$

depth z_p in these formulae is the largest depth of penetration of the ray, i.e. the depth of the turning point, where the ray starts to return to the surface. Since the descending and ascending segments of the ray are symmetrical, factor 2 stands in front of the integrals. Note that Eqs. (3.32) can be obtained immediately as the generalisation of Eqs. (3.26) by replacing the summation by integration.

3.6.2 Direct derivation of the equation of a ray from Fermat's principle

In Subsection 3.6.1, see also Tab. 3.1, we have derived the equation of a seismic ray in a vertically inhomogeneous medium using Snell's law. Here the same equation will be derived from Fermat's principle using the corresponding Euler's equation.

In the previous subsection we have used the z-coordinate as the independent variable (integration variable). Thus, let us express the path element ds as (Fig. 3.6)

$$\mathrm{d}s = \sqrt{\mathrm{d}x^2 + \mathrm{d}z^2} = \sqrt{(x')^2 + 1}\,\mathrm{d}z \ , \tag{3.33}$$

where $x' = \mathrm{d}x/\mathrm{d}z$. Fermant's functional can then be express as

$$\int_A^B \frac{\mathrm{d}s}{v(z)} = \int_A^B \frac{\sqrt{1+(x')^2}}{v(z)}\mathrm{d}z = \int_A^B F(z, x(z), x'(z))\,\mathrm{d}z \ , \tag{3.34}$$

where

$$F(z, x, x') = \frac{1}{v(z)}\sqrt{1+(x')^2} \ . \tag{3.35}$$

Since coordinate x does not enter explicitly into F (it is a cyclic coordinate in the terminology of mechanics), Euler's equation takes the following simple form

$$0=\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{\partial F}{\partial x'}\right)-\frac{\partial F}{\partial x}=\frac{\mathrm{d}}{\mathrm{d}z}\left[\frac{\partial F}{\partial x'}\right]. \tag{3.36}$$

Consequently, derivative $\partial F/\partial x'$ must be equal to a constant. Denoting this constant by p, we have

$$p=\frac{1}{v}\frac{x'}{\sqrt{1+(x')^2}}, \tag{3.37}$$

which is the *equation of a seismic ray*. This can be expressed as

$$\frac{\mathrm{d}x}{\mathrm{d}z}=\frac{pv}{\sqrt{1-p^2v^2}}. \tag{3.38}$$

Let us add a geometrical interpretation to Eq. (3.37). It follows from Fig. 3.6 that $x'=\mathrm{d}x/\mathrm{d}z=\tan i$. Thus, $\sqrt{1+(x')^2}=1/\cos i$, and

$$p=\frac{\sin i}{v}=\frac{\sin i(z)}{v(z)}. \tag{3.39}$$

Equations (3.38) and (3.39) are the same as the corresponding equations in Tab. 3.1.

3.7 Seismic Rays and Travel Times for a Medium with a Constant Velocity Gradient

Assume velocity v to be a linear function of depth z:

$$v(z)=v_0\left[1+\beta(z-z_0)\right] \tag{3.40}$$

for $z\geq z_0$, where v_0 and β are constants. Without loss of generality we may put $z_0=0$ and $x(z_0)=0$ in (3.40). Therefore, we assume

$$v(z)=v_0(1+\beta z) \tag{3.41}$$

for $z\geq 0$ (Fig. 3.7).

Calculate epicentral distance x_1 which corresponds to depth z_1. It follows from (3.31), and also from the integral expressions in Tab. 3.1, that

$$x_1 = x(z_1) = \int_0^{z_1} \frac{pv_0(1+\beta z)}{\sqrt{1-p^2v_0^2(1+\beta z)^2}} = \left[\frac{\sqrt{1-p^2v_0^2[1+\beta z]^2}}{-pv_0\beta} \right]_{z=o}^{z_1} . \qquad (3.42)$$

This equation can be expressed as

$$x_1 - \frac{1}{\beta}\sqrt{\frac{1}{p^2v^2} - 1} = -\sqrt{\frac{1}{p^2v_0^2\beta^2} - \left(\frac{1}{\beta} + z_1\right)^2} .$$

After squaring, we arrive at the equation of a circle,

$$(x_1 - x_c)^2 + (z_1 - z_c)^2 = R^2 , \qquad (3.43)$$

where the coordinates of its centre, x_c and z_c, and its radius R are given by

$$x_c = \frac{1}{\beta}\sqrt{\frac{1}{p^2v_0^2} - 1} , \; z_c = -\frac{1}{\beta} , \; R = \frac{1}{pv_0\beta} . \qquad (3.44)$$

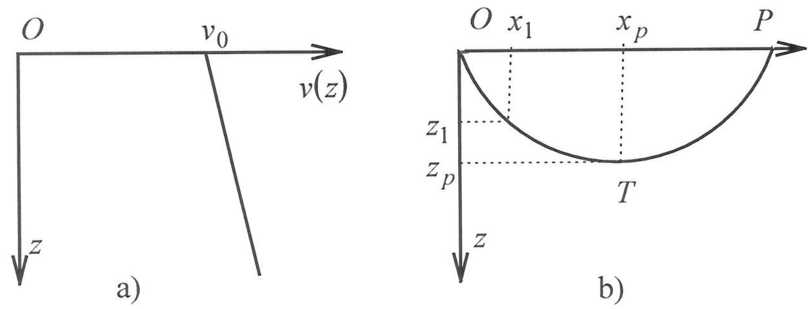


Fig. 3.7. Medium with a constant velocity gradient: a) velocity cross-section; b) seismic ray of a circular form.

Thus, the *ray* in a medium with a constant velocity gradient is an *arc of a circle*. In particular, the coordinates of the turning point $T(x_p, z_p)$ are

$$x_p = x_c = \frac{1}{\beta}\sqrt{\frac{1}{p^2v_0^2} - 1} , \; z_p = z_c + R = \frac{1}{\beta}\left(\frac{1}{pv_0} - 1\right) . \qquad (3.45)$$

The distance of point P where the ray again crosses the x-axis is

$$x(P) = 2x_c = \frac{2}{\beta}\sqrt{\frac{1}{p^2v_0^2} - 1} . \qquad (3.46)$$

Formulae (3.44) to (3.46) make it easy to construct the ray geometrically.

For the corresponding travel time we get from (3.31) that

$$\tau_1 = \tau(z_1) = \int_0^{z_1} \frac{dz}{v_0(1+\beta z)\sqrt{1-p^2 v_0^2 (1+\beta z)^2}} \ . \tag{3.47}$$

Using the usual substitution

$$\sqrt{1-p^2 v_0^2 (1+\beta z)^2} = 1-t \ , \tag{3.48}$$

we obtain

$$\tau_1 = \frac{1}{\beta v_0} \int_{t_0}^{t_1} \frac{dt}{t(2-t)} = \frac{1}{2\beta v_0} \int_{t_0}^{t_1} \left(\frac{1}{t} + \frac{1}{2-t} \right) dt = \frac{1}{2\beta v_0} \left[\ln \frac{t}{2-t} \right]_{t_0}^{t_1} , \tag{3.49}$$

where

$$t_0 = 1 - \sqrt{1-p^2 v_0^2} \ , \quad t_1 = 1 - \sqrt{1-p^2 v_0^2 (1+\beta z_1)^2} \ . \tag{3.50}$$

Note that the formulae derived above cannot be used for $\beta = 0$, i.e. for a homogeneous medium. In this case, expressions (3.44) become infinite, and (3.49) becomes indeterminate of the type $0/0$. Consequently, the expressions for x_1 and τ_1, given by (3.42) and (3.49), must be replaced by some expansions for $|\beta|$ close to zero.

3.8 Seismic Rays and Travel Times for a Medium with a Linear Quadratic Slowness

Another vertically inhomogeneous model, for which the rays and travel times can be calculated analytically, is described by the relation

$$\frac{1}{v^2(z)} = a - bz \ , \tag{3.51}$$

where v, is the velocity, z the depth, and a, b are constants. Since the reciprocal value of velocity, $1/v$, is called the slowness, relation (3.51) describes the quadratic slowness as a linear function of depth z. Relation (3.51) can also be expressed as

$$v(z) = \frac{v_0}{\sqrt{1-\gamma z}} \ , \tag{3.52}$$

where v_0 and γ are new constants.

We shall solve a similar problem as in the previous section, but instead of the velocity distribution (3.41) we shall consider relation (3.51). Assume again the ray to begin at the coordinate origin, and calculate epicentral distance x_1 and travel time τ_1 which correspond to depth z_1, i.e.

$$x_1 = \int_0^{z_1} \frac{pv\,\mathrm{d}z}{\sqrt{1-p^2v^2}} , \quad \tau_1 = \int_0^{z_1} \frac{\mathrm{d}z}{v\sqrt{1-p^2v^2}} , \tag{3.53}$$

where $v = v(z)$ is given by (3.51) or (3.52). Let us divide the numerator and denominator in the first integrand by v, in the second integrand by v^2, and substitute from (3.51):

$$x_1 = \int_0^{z_1} \frac{p\,\mathrm{d}z}{\sqrt{a-p^2-bz}} , \quad \tau_1 = \int_0^{z_1} \frac{(a-bz)\,\mathrm{d}z}{\sqrt{a-p^2-bz}} . \tag{3.54}$$

The integral for x_1 can immediately be calculated, and the integral for τ_1 can be obtained by integrating by parts. We shall arrive at the following simple formulae

$$x_1 = -\frac{2p}{b}\left[\sqrt{a-p^2-bz}\right]_{z=0}^{z_1} , \quad \tau_1 = \frac{2}{3b}\left[\left(bz-a-2p^2\right)\sqrt{a-p^2-bz}\right]_{z=0}^{z_1} . \tag{3.55}$$

These formulae are even simpler than the formulae (3.42) and (3.49) for a constant velocity gradient, as square roots only must be evaluated in (3.55). Consequently, many authors have approximated vertically inhomogeneous media by system of layers with linear quadratic slownesses.

3.9 Another Derivation of the Equation of a Seismic Ray in a Vertically Inhomogeneous Medium

We have already derived the corresponding equation of a seismic ray by two methods, namely from Snell's law in Subsection 3.6.1, and from Fermat's principle in Subsection 3.6.2. In the latter derivation we used the z-coordinate as an independent variable. It is quite natural, because the velocity is a function of the same variable, $v = v(z)$. In the case, Fermat's functional does not contain the dependent coordinate x, which leads to a substantial simplification of Euler's equations.

Here, we shall describe the third derivation of the same equation of a seismic ray by introducing the x-coordinate as an independent variable. Now, let us express the path element as

$$\mathrm{d}s = \sqrt{\mathrm{d}x^2 + \mathrm{d}z^2} = \sqrt{1+(z')^2}\,\mathrm{d}x , \tag{3.56}$$

where $z' = \mathrm{d}z/\mathrm{d}x$. Then

$$\int_A^B \frac{\mathrm{d}s}{v(z)} = \int_A^B \frac{\sqrt{1+(z')^2}}{v(z)}\mathrm{d}z = \int_A^B F\big(x, z(x), z'(x)\big)\mathrm{d}x \ , \tag{3.57}$$

where

$$F\left(x, z, z'\right) = \frac{1}{v(z)}\sqrt{1+(z')^2} \ . \tag{3.58}$$

Euler's equation now takes the form

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial z'}\right) - \frac{\partial F}{\partial z} = 0 \ . \tag{3.59}$$

However, this equation cannot be easily simplified although function F does not contain variable x explicitly. To explain this problem, let us remind a similar situation in mechanics.

If Lagrangian $L\big(q_j(t), \dot{q}_j(t), t\big)$ does not contain the time explicitly, it is better to consider the Hamiltonian. Introduce the generalised impulses

$$p_j = \frac{\partial L}{\partial \dot{q}_j} \ , \tag{3.60}$$

and the Hamiltonian

$$H = \sum_{j=1}^{n} p_j \dot{q}_j - L = \sum_{j=1}^{n} \frac{\partial L}{\partial \dot{q}_j}\dot{q}_j - L \ . \tag{3.61}$$

First, let us differentiate the Hamiltonian with respect to time:

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_j}\right)\dot{q}_j + \frac{\partial L}{\partial \dot{q}}\ddot{q}_j - \frac{\partial L}{\partial q_j}\dot{q}_j - \frac{\partial L}{\partial \dot{q}_j}\ddot{q}_j - \frac{\partial L}{\partial t} = -\frac{\partial L}{\partial t} \ ; \tag{3.62}$$

here we have used the Einstein summation convention and Lagrange's equations (3.7). Consequently, if the Lagrangian does not depend on time explicitly, i.e. $\partial L/\partial t = 0$, the Hamiltonian is constant, $H = 0$. (The Hamiltonian is equal to the mechanical energy, $H = T + V$, which is conserved in this case).

Since function F, given by (3.58), does not depend on x explicitly, it follows from the analogies from analytical mechanics that

$$\frac{\partial F}{\partial z'}z' - F \tag{3.63}$$

is equal to a constant. Denoting this constant by $(-p)$, we have

$$p = F - z'\frac{\partial F}{\partial z'} = \frac{1}{v(z)\sqrt{1+(z')^2}} \; . \tag{3.64}$$

This can be expressed as

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{\sqrt{1-p^2v^2}}{pv} \; , \tag{3.65}$$

which agrees with Eq. (3.38).

In Chapter 5 we shall consider further independent variables in Fermat's functionals, such as the length of the ray, s, or travel time τ. Wee shall not consider these formulations here, because in the special case of $v = v(z)$ they lead to more complicated equations.

It should also be noted that different types of differential equations are obtained in the individual cases. Euler's equation (3.59) is a second-order differential equation. Since function F is independent of variable x, it is possible to replace this equation by a first-order differential equation (3.65), i.e.

$$\frac{\mathrm{d}z(x)}{\mathrm{d}x} - \frac{\sqrt{1-p^2v^2(z)}}{pv(z)} = 0 \; . \tag{3.66}$$

Nevertheless, this differential equation is not easy to solve, because this is a non-linear equation (with the exception of the trivial case of $v(z) = \text{const.}$). However, if the last equation is reformulated for the inverse function $x = x(z)$, the situation simplifies substantially:

$$\frac{\mathrm{d}x(z)}{\mathrm{d}z} - \frac{pv(z)}{\sqrt{1-p^2v^2(z)}} = 0 \; ; \tag{3.67}$$

see Eq. (3.38). This is a very simple differential equation as the unknown function is not contained in the second term. The solution can be written immediately in the form of an integral:

$$x(z_1) = x(z_0) + \int_{z_0}^{z_1} \frac{pv}{\sqrt{1-p^2v^2}}\,\mathrm{d}z \; , \tag{3.68}$$

see Tab. 3.1. Therefore, for analytical or numerical computations, formulae (3.38) and (3.68) should be used.

3.10 A Review of the Formulae for Vertically Inhomogeneous Media and Spherically Symmetric Media

The basic formulae are contained in the following Tab. 3.1.

Table 3.1. Equations of rays and travel-time curves.

Vertically inhomogeneous	Spherically symmetric
Medium	
Velocity: $v = v(z)$	$v = v(r)$
Contributions in terms of i:	
$\mathrm{d}x = \mathrm{d}z \tan i(z)$	$r\,\mathrm{d}\Theta = \mathrm{d}r \tan i(r)$
$\mathrm{d}\tau = \dfrac{\mathrm{d}z}{v\cos i(z)}$	$\mathrm{d}\tau = \dfrac{\mathrm{d}r}{v\cos i(r)}$
Snell's law (equation of a seismic ray):	
$\dfrac{\sin i(z)}{v(z)} = p$	$\dfrac{r\sin i(r)}{v(r)} = p$
$\Rightarrow \sin i(z) = pv$	$\Rightarrow \sin i(r) = \dfrac{pv}{r}$
Contributions in terms of p:	
$\mathrm{d}x = \mathrm{d}z\dfrac{\sin i}{\cos i} = \mathrm{d}z\dfrac{pv}{\sqrt{1-p^2v^2}}$	$\mathrm{d}\Theta = \dfrac{\mathrm{d}r}{r}\dfrac{\sin i}{\cos i} = \dfrac{\mathrm{d}r}{r}\dfrac{pv}{\sqrt{r^2-p^2v^2}}$
$\mathrm{d}\tau = \dfrac{\mathrm{d}z}{v\sqrt{1-p^2v^2}}$	$\mathrm{d}\tau = \dfrac{r\,\mathrm{d}r}{v\sqrt{r^2-p^2v^2}}$
Integral expessions:	
$x(z,p) = x(z_0,p) + \displaystyle\int_{z_0}^{z}\frac{pv\,\mathrm{d}z}{\sqrt{1-p^2v^2}}$	$\Theta(r,p) = \Theta(r_0,p) + \displaystyle\int_{r_0}^{r}\frac{pv\,\mathrm{d}r}{r\sqrt{r^2-p^2v^2}}$
$\tau(z,p) = \tau(z_0,p) + \displaystyle\int_{z_0}^{z}\frac{\mathrm{d}z}{v\sqrt{1-p^2v^2}}$	$\tau(r,p) = \tau(r_0,p) + \displaystyle\int_{r_0}^{r}\frac{r\,\mathrm{d}r}{v\sqrt{r^2-p^2v^2}}$
Equation of the travel-time curve:	
$x(p) = 2p\displaystyle\int_{0}^{z_p}\frac{v\,\mathrm{d}z}{\sqrt{1-p^2v^2}}$	$\Theta(p) = 2p\displaystyle\int_{r_p}^{R}\frac{v\,\mathrm{d}r}{r\sqrt{r^2-p^2v^2}}$
$\tau(p) = 2\displaystyle\int_{0}^{z_p}\frac{\mathrm{d}z}{v\sqrt{1-p^2v^2}}$	$\tau(p) = 2\displaystyle\int_{r_p}^{R}\frac{r\,\mathrm{d}r}{v\sqrt{r^2-p^2v^2}}$

3.11 Appendix: Rays and Travel Times for a Convex Velocity-Depth Distribution

We have derived analytical formulae for seismic rays and travel times for two types of an inhomogeneous medium:

1) for the velocity as a linear function of depth, given by (3.41), see Section 3.7;
2) for a linear quadratic slowness, given by (3.51), see Section 3.8.

These velocity-depth distributions have been used by many authors. In the first case, the second derivative of velocity with respect to depth is zero, $v''(z) = 0$. In the second case, the velocity is a concave function of depth, $v''(z) > 0$.

However, the velocity-depth distribution at shallow depths is usually convex, $v''(z) < 0$, i.e. the velocity gradient near the Earth's surface is large, and then gradually decreases with depth. As a better approximation to this velocity-depth distribution we could consider, e.g., the following function:

$$v(z) = v_0\sqrt{1+\beta z}\ , \tag{3.69}$$

where v_0 and β are again constants, z is the depth. Let us derive the equations of rays and travel times for this velocity function.

Denote again the parameter of a ray by p, see (3.30). Introduce a new variable, namely velocity $v = v_0\sqrt{1+\beta z}$, instead of depth z. Since

$$\mathrm{d}v = \frac{v_0\beta}{2\sqrt{1+\beta z}} = \frac{v_0^2\beta}{2v}\mathrm{d}z\ , \tag{3.70}$$

the indefinite integrals for epicentral distance x and travel time τ become:

$$\begin{aligned} x &= \int\frac{pv}{\sqrt{1-p^2v^2}}\mathrm{d}z = -\frac{2}{pv_0^2\beta}\int v\frac{\left(-p^2v\right)}{\sqrt{1-p^2v^2}}\mathrm{d}v\ ,\\ \tau &= \int\frac{\mathrm{d}z}{v\sqrt{1-p^2v^2}} = \frac{2}{v_0^2\beta}\int\frac{\mathrm{d}v}{\sqrt{1-p^2v^2}}\ . \end{aligned} \tag{3.71}$$

Integrating by parts, the epicentral distance can be expressed as

$$\begin{aligned} x &= -\frac{2}{pv_0^2\beta}\left[v\sqrt{1-p^2v^2} - \int\sqrt{1-p^2v^2}\,\mathrm{d}v\right] =\\ &= -\frac{1}{pv_0^2\beta}\left[v\sqrt{1-p^2v^2} - \frac{1}{p}\arcsin(pv)\right]\ . \end{aligned} \tag{3.72}$$

For the travel time we immediately obtain

$$\tau = \frac{2}{p v_0^2 \beta} \arcsin(pv) \ . \tag{3.73}$$

Assume that the ray begins at the coordinate origin, where $x = z = \tau = 0$ and $v = v_0$. At a depth z_1, denote the velocity by

$$v_1 = v_0 \sqrt{1 + \beta z_1} \ . \tag{3.74}$$

It follows from (3.72) to (3.74) that the epicentral distance, x_1, and travel time, τ_1, corresponding to this depth are

$$x_1 = \frac{1}{p^2 v_0^2 \beta} \left[\arcsin(pv) - pv\sqrt{1 - p^2 v^2} \right]_{v_0}^{v_1} , \quad \tau_1 = \frac{2}{p v_0^2 \beta} \left[\arcsin(pv) \right]_{v_0}^{v_1} \ . \tag{3.75}$$

These relatively simple formulae solve our problem for the velocity distribution given by function (3.69).

Chapter 4

Seismic Rays and Travel Times in a Spherically Symmetric Medium

In this chapter we shall consider the Earth as a spherically symmetric body, i.e. as a body in which the physical parameters are dependent only on the radical distance from its centre.

4.1 Seismic Rays in a Medium Consisting of Concentric Spherical Layers

A simple model which takes the sphericity of the Earth into account is composed of spherical concentric layer. Here we shall assume each of the layers to be homogeneous and isotropic. The seismic rays in each layer are then straight lines. At the spherical interfaces we shall assume that the same laws of reflection and transmission hold as if the interface were approximated locally by the tangent plane. (It follows from more detailed considerations that the approximation by tangent planes is applicable if the radius of curvature is much larger than the wavelength). Seismic rays can then be constructed geometrically. We shall derive the corresponding formulae for this purpose.

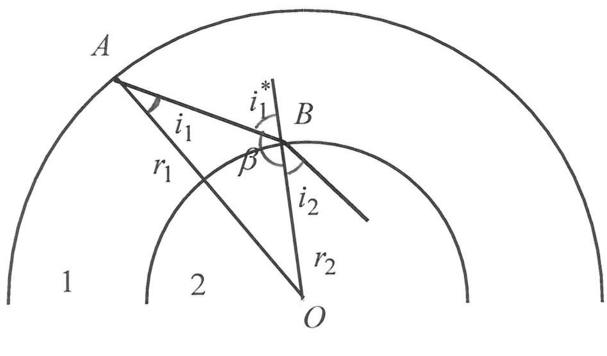


Fig. 4.1. A seismic ray in a spherical layer.

Consider the wave propagation in the first layer along abscissa AB connecting the outer and inner boundaries (Fig. 4.1). Their radii are denoted by r_1 and r_2, respectively. The angles made by the radius and the ray in the first layer are denoted by i_1 at point A, and i_1^* at point B. Since the layer is spherical, these angles are different, $i_1^* > i_1$. A simple relation between them follows from the sine theorem applied to triangle ABO:

$$\frac{\sin i_1}{r_2} = \frac{\sin \beta}{r_1} = \frac{\sin i_1^*}{r_1} , \tag{4.1}$$

β being the internal angle in triangle ABO at vertex B.

Snell's law, applied at point B (at the interface of the first and second layers), yields

$$\frac{\sin i_1^*}{v_1} = \frac{\sin i_2}{v_2} . \tag{4.2}$$

Combining (4.1) and (4.2), we arrive at the generalised Snell's law for a spherically layered medium:

$$\frac{r_1 \sin i_1}{v_1} = \frac{r_2 \sin i_2}{v_2} = \cdots = p . \tag{4.3}$$

Quantity p is called the *ray parameter* in a spherically layered medium. This is a constant along the whole ray in this medium. For a given model of the medium and a given initial angle i_1, Eq. (4.3) makes it possible to construct the whole ray. Consequently, this equation is also referred to as the equation of a ray in a spherically layered medium.

4.2 Equations of Seismic Rays and Travel Times in a Spherically Symmetric Medium

A medium in which the velocity is dependent only on the spherical coordinate r (distance from the Earth's centre) is called the spherically symmetric medium or radially symmetric medium. The velocity in this medium is thus $v = v(r)$.

4.2.1 Generalisation of formulae for a spherically layered medium

Snell's law (4.3) for homogeneous spherical layers can be generalised to a general spherically symmetric medium as follows:

$$\boxed{\frac{r \sin i(r)}{v(r)} = p} , \tag{4.4}$$

where p is a constant characterising the particular ray. This generalised Snell's law represents the *equation of a seismic ray* in a spherically symmetric medium.

Consider an infinitesimal part of a ray in a spherically symmetric medium (Fig. 4.1). When a wave has passed through a path $\mathrm{d}s$, the corresponding increments in the angular epicentral distance, $\mathrm{d}\Theta$, and in the travel time, $\mathrm{d}\tau$, are given by

$$\mathrm{d}\Theta = \frac{\mathrm{d}r}{r}\tan i(r) = \frac{\mathrm{d}r}{r}\frac{\sin i}{\cos i} = \frac{\mathrm{d}r}{r}\frac{pv}{\sqrt{r^2 - p^2v^2}} ,$$

$$\mathrm{d}\tau = \frac{\mathrm{d}r}{v\cos i(r)} = \frac{r\,\mathrm{d}r}{v\sqrt{r^2 - p^2v^2}} . \tag{4.5}$$

The other formulae are summarised in the right column of Tab. 3.1. The last equations in the table are the parametric equations of a seismic ray in a spherically symmetric medium for the source and receiver on the surface:

$$\Theta(p) = 2p\int_{r_p}^{R}\frac{v\,\mathrm{d}r}{r\sqrt{r^2 - p^2v^2}} , \quad \tau(p) = 2\int_{r_p}^{R}\frac{r\,\mathrm{d}r}{v\sqrt{r^2 - p^2v^2}} , \tag{4.6}$$

where R is the Earth's radius and r_p is the radial distance of the turning point from the Earth's centre (the deepest point of the ray); see Fig. 4.1.

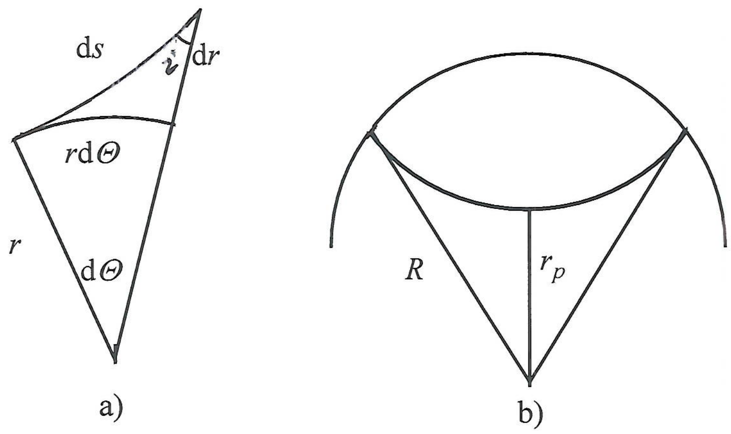


Fig. 4.1. Rays in a spherically symmetric medium: a) an infinitesimal segment; b) a whole ray.

4.2.2 Direct derivation of the equation of a ray from Fermat's principle

It follows from Fig. 4.1 that

$$\mathrm{d}s = \sqrt{(r\,\mathrm{d}\Theta)^2 + \mathrm{d}r^2} = \sqrt{r^2\Theta'^2 + 1}\,\mathrm{d}r , \tag{4.7}$$

where $\Theta' = \mathrm{d}\Theta/\mathrm{d}r$. Fermat's functional is

$$\int_A^B F(r, \Theta, \Theta')\,\mathrm{d}r = \int_A^B \frac{1}{v(r)}\sqrt{r^2\Theta'^2 + 1}\,\mathrm{d}r , \tag{4.8}$$

and the corresponding Euler's equation takes again a simple form:

$$0 = \frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{\partial F}{\partial \Theta'}\right) - \frac{\partial F}{\partial \Theta} = \frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{\partial F}{\partial \Theta'}\right) . \tag{4.9}$$

Therefore, the last derivative must be a constant,

$$p = \frac{\partial F}{\partial \Theta'} = \frac{1}{v}\frac{r^2 \Theta'}{\sqrt{r^2 \Theta'^2 + 1}} , \tag{4.10}$$

which may be expressed as

$$\frac{\mathrm{d}\Theta}{\mathrm{d}r} = \frac{pv}{r\sqrt{r^2 - p^2 v^2}} . \tag{4.11}$$

Again, let us express Eq. (4.10) in terms of angle i. It follows from Fig. 4.1 that

$$\tan i = \frac{r\,\mathrm{d}\Theta}{\mathrm{d}r} = r\Theta' .$$

Inserting it into Eq. (4.10) we get

$$p = \frac{r \tan i}{v\sqrt{\tan^2 i + 1}} = \frac{r \sin i}{v} . \tag{4.12}$$

Equations (4.11) and (4.12) agree with Eqs. (4.5) and (4.4).

4.3 Determination of the Ray Parameter from Observations. The Bendorf Equation

4.3.1 Determination from the travel-time curve

The ray parameter p can be determined from the travel-time curve of the seismic wave. To demonstrate it, let us consider the situation as show in Fig. 4.2 for a plane surface of the Eatrh.

Denote by v_0 the velocity near the Earth's surface and by i_0 the angle of incidence of a seismic wave at the surface. Further, denote by c the *apparent velocity* with which the wavefront propagates along the surface. Consider the wavefront *AB* which, after an infinitesimal time dt, is displaced by a distance $\mathrm{d}s = v_0\,\mathrm{d}t$ to point *C*. Its projection onto the Earth's surface moves from point *A* to point *C* by a distance $\mathrm{d}\Delta = c\,\mathrm{d}t$. From the triangle *ABC* we have

$$\sin i_0 = \frac{\mathrm{d}s}{\mathrm{d}\Delta} = \frac{v_0\,\mathrm{d}t}{\mathrm{d}\Delta} = \frac{v_0}{c} \,. \tag{4.13}$$

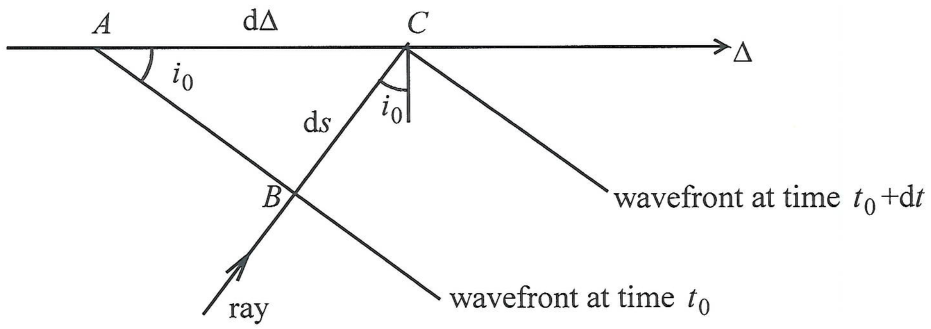


Fig 4.2. Situation for the derivation of the Bendorf equation.

Consequently, for the ray parameter $p = (\sin i_0)/v_0$ we get

$$\boxed{p = \frac{\mathrm{d}t}{\mathrm{d}\Delta}} \,. \tag{4.14}$$

This is the *Bendorf equation* for a *vertically inhomogeneous medium*. Thus, the ray parameter p can be obtained as the derivative of the travel-time curve at the corresponding epicentral distance Δ; see Fig. 4.3.

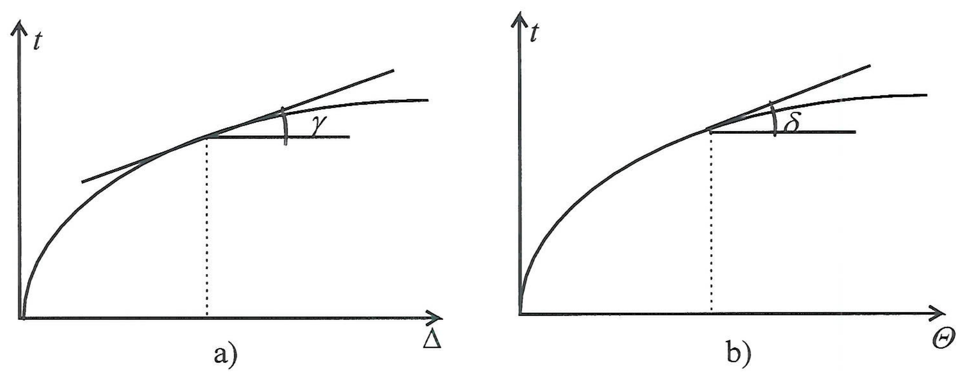


Fig. 4.3. Determination of the ray parameter p as the derivative of the travel-time curve: vertically inhomogeneous medium, $p = (\sin i_0)/v_0 = \tan\gamma$; b) spherically symmetric medium, $p = R(\sin i_0)/v_0 = \tan\delta$.

Moreover, it follows from (4.13) that

$$p = \frac{1}{c} \,. \tag{4.15}$$

Thus, the ray parameter p is the reciprocal value of the apparent velocity with which the wavefront propagates along the Earth's surface.

Now, let us consider a spherically symmetric Earth. In this case $\mathrm{d}\Delta = R\,\mathrm{d}\Theta$, where R is the Earth's radius and Θ is the angular epicentral distance. Inserting into (4.13) we get

$$\boxed{\frac{\mathrm{d}t}{\mathrm{d}\Theta} = \frac{R\sin i_0}{v_0} = p}\,, \tag{4.16}$$

where p is the ray parameter for a spherically symmetric model. Equation (4.16) is the *Bendorf equation* for a spherically symmetric model. It enables us to determine the ray parameter p again as the derivative of the travel time curve; see also Fig. 4.3b.

4.3.2 Direct determinations of the ray parameter

Instead of numerical differentiation of the travel-time curve, its derivative can be estimated from differences of the travel times observed at a group of closely distributed seismic stations, called a *seismic array*. Such measurements yield more accurate values of the ray parameter than the differentiation described above.

An estimate of the ray parameter can even be obtained from observations of one station only, if a three-component record is available. In this case the angle of incidence i_0 can be determined by means of polarisation analysis. If the surface velocity v_0 is known, we can then calculate the ray parameter p.

4.4 Interpretation of Travel-time Curves of Refracted Waves in a Spherically Symmetric Model. The Wiechert-Herglotz Method.

So far we have solved only so-called *forward seismic problems*, i.e. for a given model of the medium and a given source we have calculated seismic rays and other characteristics of the wave field. By *inverse seismic problems* we mean opposite processes, i.e. determinations of the parameters of the medium and of the source from observed data, such as travel-time curves, dispersion curves, seismic amplitudes or whole seismograms.

Solutions of inverse problems are generally more complicated than solutions of forward problems. Inverse problems frequently lead to complicated systems of equations which are often non-linear, badly conditioned, etc. Only in exceptional cases, the solution of an inverse problem can be found in an analytical form. We shall describe here one of these exceptional cases. It will be the Wiechert-Herglotz method for interpreting travel-time curves of refracted waves in a spherically symmetric model. As a result of interpreting a travel-time curve by this method, we obtain a radial distribution of velocity, $v = v(r)$.

We shall restrict ourselves to the special case in which the ratio $r/v(r)$ decreases continuously with depth. The rays are then convex downward. Their parameters read Eq. (4.4).

Assume a travel-time curve of P or S waves to be known for epicentral distances from $\Theta = 0$ to a distance Θ^*. According to the Bendorf equation (4.16), we obtain the corresponding ray parameters p by differentiating the travel-time curve. Therefore, we assume that function $p = p(\Theta)$ is also known.

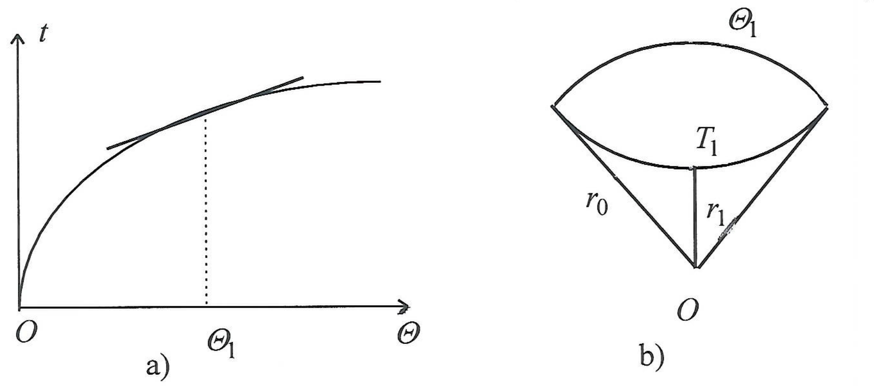


Fig 4.4. Graphs illustrating the Wiechert-Herglotz method: a) travel-time curve; b) seismic ray.

Consider an epicentral distance Θ_1, such that $0 < \Theta_1 \le \Theta^*$; see Fig. 4.4. The seismic ray arriving at this distance has its deepest point T_1 (turning point) at a radial distance r_1, which is unknown. Denote the velocity at this depth by $v_1 = v(r_1)$. Since the angle of incidence at this point is $90°$, for the ray parameter we have, see (4.4),

$$p(\Theta_1) = \frac{r_1}{v_1} . \tag{4.17}$$

Since $p(\Theta_1)$ can be determined from the Bendorf equation,

$$p(\Theta_1) = \left(\frac{\mathrm{d}t}{\mathrm{d}\Theta}\right)_{\Theta_1} , \tag{4.18}$$

the ratio r_1/v_1 is known. We need another equation to determine r_1 and v_1 separately. We shall derive such a formula, which will enable us to determine the radial distance r_1 independently.

Denote the Earth's radius by r_0 and the velocity at the surface by v_0. It will be convenient to introduce function $\eta(r)$, defined by the above-mentioned ratio

$$\eta(r) = r/v(r) . \tag{4.19}$$

At the surface this function attains value $\eta_0 = r_0/v_0$, and at the deepest point T_1 the value $\eta_1 = r_1/v_1$; see also the review of notations in Tab. 4.1.

Table 4.1. Notation used for different radial distances of the turning point, r_m. Function η is defined by (4.19), p is the ray parameter, Θ is the epicentral distance.

r_m	$\eta = r/v(r)$	$p = \eta(r_m)$	$\Theta(p)$
r_0	$\eta_0 = r_0/v_0$	η_0	0
r	η	η	Θ
r_1	$\eta_1 = r_1/v_1$	η_1	Θ_1

We assume that $\eta_0 > \eta_1$ and that $\eta(r)$ is a monotonously decreasing function when moving from the Earth's surface to the deepest point T_1 (Fig. 4.5).

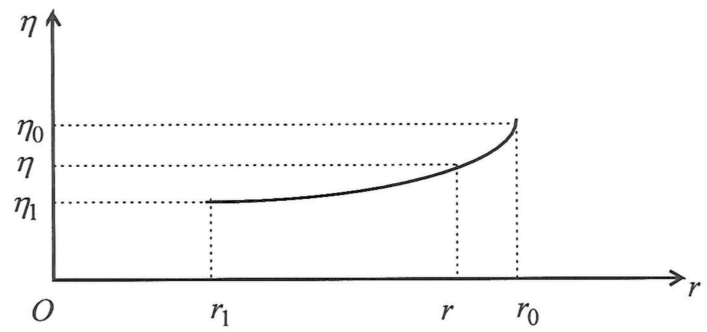


Fig 4.5. Function $\eta(r) = r/v(r)$.

It follows from (4.6) that

$$\Theta(p) = \int_{r_m}^{r_0} \frac{2p\,\mathrm{d}r}{r\sqrt{\eta^2 - p^2}}, \tag{4.20}$$

where we have denoted the radial coordinate of the turning point by r_m. Since we assume a unique correspondence between r and η, we may consider the inverse function, $r = r(\eta)$, and perform the integration with respect to η:

$$\Theta(p) = \int_{p}^{\eta_0} \frac{2p}{r\sqrt{\eta^2 - p^2}} \frac{\mathrm{d}r}{\mathrm{d}\eta}\mathrm{d}\eta, \tag{4.21}$$

where the lower bound of the integral is $\eta_m = r_m/v(r_m) = p$.

Equation (4.21) has the form of Abel's integral equation whose solution determines η as a function of r. There is a general theory of solving this type of integral equations. However, we shall not use this general theory here, but we shall derive the solution for our special case only.

Apply the operation

$$\int_{\eta_1}^{\eta_0} \left(p^2 - \eta_1^2\right)^{-1/2} \mathrm{d}\,p \tag{4.22}$$

to both sides of Eq. (4.21). This represents the integration over the seismic rays from the ray arriving at Θ_1 to shorter epicentral distances up to $\Theta = 0$. We obtain

$$\int_{\eta_1}^{\eta_0} \frac{\Theta \,\mathrm{d}\,p}{\sqrt{p^2 - \eta_1^2}} = \int_{\eta_1}^{\eta_0} \mathrm{d}\,p \int_{p}^{\eta_0} \frac{2p}{r\sqrt{p^2 - \eta_1^2}\sqrt{\eta^2 - p^2}} \frac{\mathrm{d}\,r}{\mathrm{d}\,\eta} \mathrm{d}\,\eta =$$
$$= \int_{\eta_1}^{\eta_0} \mathrm{d}\,\eta \int_{\eta_1}^{\eta} \frac{2p}{r\sqrt{p^2 - \eta_1^2}\sqrt{\eta^2 - p^2}} \frac{\mathrm{d}\,r}{\mathrm{d}\,\eta} \mathrm{d}\,p \;. \tag{4.23}$$

In the last step we interchanged the order of integration. Since we integrate over a triangular domain, the integration bounds must be modified according to Fig 4.6.

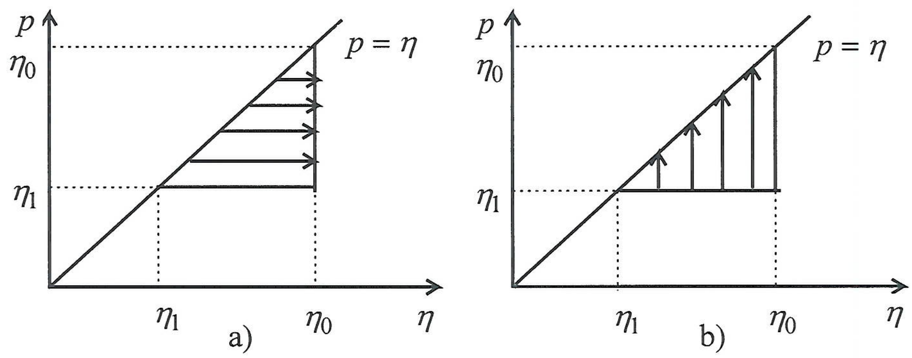


Fig. 4.6. Integration over the triangular domain in (4.23). The arrows indicate the first (inner) integration: a) the inner integration with respect to η; b) the inner integration with respect to p.

In integrating on the left-hand side of Eq. (4.23), we shall use the formula

$$\int \frac{\mathrm{d}\,x}{\sqrt{x^2 - 1}} = \cosh^{-1} x = \operatorname{ar}\cosh x = \ln\left(x + \sqrt{x^2 - 1}\right) \;. \tag{4.24}$$

Integrating by parts, we obtain

$$\left[\Theta \operatorname{ar}\cosh \frac{p}{\eta_1}\right]_{\eta_1}^{\eta_0} - \int_{\eta_1}^{\eta_0} \frac{\mathrm{d}\,\Theta}{\mathrm{d}\,p} \operatorname{ar}\cosh\left(\frac{p}{\eta_1}\right) \mathrm{d}\,p = \int_{\eta_1}^{\eta_0} \frac{\mathrm{d}\,r}{\mathrm{d}\,\eta} \frac{\mathrm{d}\,\eta}{r} \int_{\eta_1}^{\eta} \frac{2p\,\mathrm{d}\,p}{\sqrt{\left(p^2 - \eta_1^2\right)\left(\eta^2 - p^2\right)}} \;. \tag{4.25}$$

The first term on the left-hand side of Eq. (4.25) is zero, because $\Theta = 0$ for $p = \eta_0$ (see Tab. 4.1. and formula (4.21)), and $\operatorname{ar}\cosh(1) = 0$.

Now, calculate the last integral in (4.25). Since $2p$ in the numerator is the derivative of p^2, let us introduce the substitution $u = p^2$ and notation

$$N_1 = \eta_1^2 = \frac{r_1^2}{v_1^2}\ , \quad N = \eta^2 = \frac{r^2}{v^2}\ .$$

Then, since $\eta > \eta_1$,

$$I = \int_{\eta_1}^{\eta} \frac{2p\,\mathrm{d}p}{\sqrt{\left(p^2 - \eta_1^2\right)\left(\eta^2 - p^2\right)}} = \int_{N_1}^{N} \frac{\mathrm{d}u}{\sqrt{(u - N_1)(N - u)}} =$$
$$= \int_{N_1}^{N} \frac{\mathrm{d}u}{\sqrt{\left(\frac{N - N_1}{2}\right)^2 - \left(u - \frac{N + N_1}{2}\right)^2}}\ . \tag{4.26}$$

Denoting $a = (N - N_1)/2$ and $w = u - (N + N_1)/2$, we obtain

$$I = \int_{-a}^{a} \frac{\mathrm{d}w}{\sqrt{a^2 - w^2}} = \left[\arcsin\frac{w}{a}\right]_{w=-a}^{a} = \pi\ . \tag{4.27}$$

Consequently, Eq. (4.25) becomes (see the changes of the integration bounds in Tab. 4.1)

$$\int_0^{\Theta_1} \operatorname{ar}\cosh\frac{p}{\eta_1}\,\mathrm{d}\Theta = \pi \int_{r_1}^{r_0} \frac{\mathrm{d}r}{r} = \pi \ln\frac{r_0}{r_1}\ .$$

Finally, we arrive at

$$\boxed{\ln\frac{r_0}{r_1} = \frac{1}{\pi}\int_0^{\Theta_1} \operatorname{ar}\cosh\frac{p(\Theta)}{p(\Theta_1)}\,\mathrm{d}\Theta}\ . \tag{4.28}$$

This is the famous *Wiechert-Herglotz formula* for determining the *radial distance* r_1 of the turning point (deepest point) of the ray which arrives at epicentral distance Θ_1. Therefore, in order to calculate this radial distance, we must determine the derivatives of the travel-time curve, $p(\Theta)$, and to compute an integral over the epicentral distances from $\Theta = 0$ to Θ_1.

If r_1 is known, the corresponding velocity $v_1 = v(r_1)$ can be determined from (4.17). This process can be carried out for any r_1 in the range

$r_0 \geq r_1 \geq r^*$, where r^* corresponds to the maximum epicentral distance Θ^*. We have thus found a method for determining the velocity as a function of the radial distance.

Although the Wiechert-Herglotz method is a powerful method of studying the Earth's interior, it has some limitations, in particular:

1) Only spherically symmetric models can be obtained by this method (a modification for a vertically inhomogeneous medium also exist). Lateral inhomogeneities of the medium cannot be studied by this method.
2) The method fails at depths where no rays have the turning points, i.e. in low-velocity zones. Such problems occur in the asthenosphere at depths of about 200 km, and also in the transition zone between the outer and inner cores.

As opposed to it, the large decrease of velocities when passing from mantle to core does not cause principal problems. Namely, a travel-time curve reduced to the surface of the core can be constructed by subtracting the *PcP* travel times from the *PKP* travel times. The Wiechert-Herglotz method is then applied to the reduced travel-time curve to obtain the velocity structure in the Earth core. In spite of the limitations mentioned above, the Wiechert-Herglotz method has led to fairly precise determinations of seismic velocities throughout most of the Earth.

Chapter 5

Seismic Waves in More Complicated Models of the Medium

In this chapter we shall deals only with several selected problems of wave propagation in complicated media, which are closely related to the problems discussed in the previous chapter. A more detailed description of the wave propagation in complicated media can be found in specialised literature, e.g., in the lecture notes by Psencik (1994) and by Popov (1996).

5.1 Mathematical methods of studying the propagation of seismic waves

The theory of seismic waves is now an extensive discipline, in which various methods of mathematical physics are used. The mathematical methods in the theory of seismic waves can be divided into the following main groups:

a) numerical solution of the equation of motion;
b) wave methods;
c) approximate methods.

Each of these methods has certain advantages and disadvantages. A choice of the method for solving a particular problem depends on our computational possibilities, development of the theory, and required accuracy. Let us briefly discuss the main properties of the individual methods.

Direct numerical solutions of the elastodynamic equations are usually based on the *method of finite differences* or the *method of finite elements*. These methods make it possible to compute the complete wave field. However, only relatively small models can usually be considered, because the number of numerical operations increases enormously with the increase of the dimensions of the model. Moreover, various numerical instabilities of these methods represent also serious problems. We shall not deal with these methods here.

The *wave methods are* based on analytical solutions (the so-called formal solutions) of the elastodynamic equations. Such solutions can be found in the case of some simple models of the medium, e.g., for models in which the velocity is dependent on one coordinate (depth) only. In the case of a *point source*, the solution is usually given in the form of a line integral in the complex plane or along the real axis. The corresponding integration is very difficult to perform for two reasons:

1) the integrand is complicated, and has various singularities along the integration path;
2) the integrand is usually rapidly oscillating.

Consequently, the corresponding integrals have been calculated exactly only in exceptional cases. The wave methods do not provide us with expressions for individual waves, but yield the wave field as a whole. For this reason, they are especially convenient in the studies of wave phenomena where interference plays an important role, e.g., in studies of surface waves, waves in media containing thin layers, reflection and transmission at transition layers, etc. Typical representatives of the wave methods for layered media are so-called matrix methods. Their detailed description can be found, e.g., in the lecture notes by Novotny (1999).

The most important methods in seismic prospecting and in many other applications (see the previous chapter), are *approximate methods*, such as the ray method and its various modifications. In the ray methods, the individual waves (such as direct, reflected, refracted, head waves, etc.) are studied independently of each other. In spite of many limitations of these methods, they are usually very fast from the computational point of view, and applicable to very complicated models of the medium. At present, ray methods are practically the only methods which make it possible to compute the wave propagation in three-dimensional models of the medium.

We shall demonstrate below the possibilities of the ray method on several typical problems.

5.2 Seismic Rays in 2-D Media; the Lagrangian Approach

Consider an inhomogeneous medium where the velocity, v, is a function of two Cartesian coordinates, x and z, i.e. $v = v(x, z)$. We shall speak of a two-dimensional medium (2-D medium). As before, we consider the z-coordinate to be the depth, and the x-coordinate to be a horizontal coordinate.

We shall derive the equations of seismic rays again from Fermat's principle. In this case, Fermat's functional is of the form

$$\int_A^B \frac{\mathrm{d}s}{v(x, z)}\,, \tag{5.1}$$

where ds is a path element,

$$\mathrm{d}s = \sqrt{\mathrm{d}x^2 + \mathrm{d}z^2}\;; \tag{5.2}$$

see (3.1) and (3.33). As the independent variable we may choose varies quantities, and accordingly we may write

$$\mathrm{d}s = \sqrt{1 + \left(\frac{\mathrm{d}z}{\mathrm{d}x}\right)^2}\,\mathrm{d}x = \tag{5.3a}$$

$$= \sqrt{1 + \left(\frac{\mathrm{d}x}{\mathrm{d}z}\right)^2}\,\mathrm{d}z = \tag{5.3b}$$

$$= \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\tau}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}\tau}\right)^2}\,\mathrm{d}\tau = \tag{5.3c}$$

$$= \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}s}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}s}\right)^2}\,\mathrm{d}s = \tag{5.3d}$$

$$= \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\sigma}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}\sigma}\right)^2}\,\mathrm{d}\sigma\ , \tag{5.3e}$$

where τ is the travel time, s is the length along the ray, i.e. $\mathrm{d}s = v\,\mathrm{d}\tau$ and $s = \int v\,\mathrm{d}\tau$. Quantity σ is defined by $\sigma = \int v\,\mathrm{d}s$, $\mathrm{d}\sigma = v\,\mathrm{d}s = v^2\,\mathrm{d}\tau$. This quantity resembles the angular momentum in the mechanics of particles.

Let us derive the equations of seismic rays for the individual cases mentioned above. In this section we shall use Euler's equations, i.e. the analogues of Lagrange's equations of the second kind in analytical mechanics. We shall restrict ourselves only to the rays in the (x,z)-plane.

In order to compute the travel-time curves, the corresponding equation for the travel time must be added to the equations of seismic rays.

5.2.1 Independent variable *x*

If the x-coordinate is chosen as the independent variable, see (5.3a), the integrand in (5.1) takes the form

$$F\big(x, z(x), z'(x)\big) = \frac{\sqrt{1+(z')^2}}{v(x,z)}\ , \tag{5.4}$$

and Euler's equation reads

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial z'}\right) - \frac{\partial F}{\partial z} = 0\ . \tag{5.5}$$

This yields the equation of a seismic ray in the following form

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{z'}{v\sqrt{1+(z')^2}}\right) - \sqrt{1+(z')^2}\,\frac{\partial}{\partial z}\left(\frac{1}{v}\right) = 0\ . \tag{5.6}$$

We have obtained one ordinary differential equation of the second order for the unknown function $z = z(x)$. On integrating this equation we encounter problems with the rays which are close to the vertical, because very large increments Δz correspond to a given increment Δx.

5.2.2 Independent variable *z*

In this case, we can write the equations which are analogous to (5.4) and (5.5), but with the interchanged role of variables x and z. Consequently, we get

$$\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{x'}{v\sqrt{1+(x')^2}}\right)-\sqrt{1+(x')^2}\,\frac{\partial}{\partial x}\left(\frac{1}{v}\right)=0\ . \tag{5.7}$$

Problems now occur at places where the rays are close to the horizontal.

5.2.3 Independent variable τ

The problems with vertical or horizontal rays, mentioned above, can easily be removed if the travel time is used as the independent variable, see (5.3c). The integrand in (7.1) may then be expressed as

$$F\big(\tau, x(\tau), z(\tau), x'(\tau), z'(\tau)\big)=\frac{\sqrt{(x')^2+(z')^2}}{v\big(x,z\big)}\ , \tag{5.8}$$

where $x'=\mathrm{d}x/\mathrm{d}\tau$ and $z'=\mathrm{d}z/\mathrm{d}\tau$. For a seismic ray we now obtain two equations:

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{x'}{v\sqrt{(x')^2+(z')^2}}\right)-\sqrt{(x')^2+(z')^2}\,\frac{\partial}{\partial x}\left(\frac{1}{v}\right)=0\ ,\\ &\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{z'}{v\sqrt{(x')^2+(z')^2}}\right)-\sqrt{(x')^2+(z')^2}\,\frac{\partial}{\partial z}\left(\frac{1}{v}\right)=0\ .\end{aligned} \tag{5.9}$$

Since x' and z' are the x- and z- components of velocity , respectively, it holds that $\sqrt{(x')^2+(z')^2}=v$. Hence, Eqs. (5.9) may be simplified to read:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{1}{v^2}\frac{\mathrm{d}x}{\mathrm{d}\tau}\right)-v\frac{\partial}{\partial x}\left(\frac{1}{v}\right)=0,\quad \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{1}{v^2}\frac{\mathrm{d}z}{\mathrm{d}\tau}\right)-v\frac{\partial}{\partial z}\left(\frac{1}{v}\right)=0\ . \tag{5.10}$$

An equivalent form of these equations is as follows:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{1}{v^2}\frac{\mathrm{d}x}{\mathrm{d}\tau}\right)+\frac{1}{v}\frac{\partial v}{\partial x}=0,\quad \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{1}{v^2}\frac{\mathrm{d}z}{\mathrm{d}\tau}\right)+\frac{1}{v}\frac{\partial v}{\partial z}=0\ . \tag{5.11}$$

5.2.4 Independent variable *s*

Quite regular increments along the ray can be obtained if the length along the ray, *s*, is used as the independent variable; see (5.3d). Analogously to (5.9) we get

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{x'}{v\sqrt{(x')^2+(z')^2}}\right)-\sqrt{(x')^2+(z')^2}\,\frac{\partial}{\partial x}\left(\frac{1}{v}\right)=0\ ,$$

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{z'}{v\sqrt{(x')^2+(z')^2}}\right)-\sqrt{(x')^2+(z')^2}\,\frac{\partial}{\partial z}\left(\frac{1}{v}\right)=0\ ,$$

where $x'=\mathrm{d}x/\mathrm{d}s$ and $z'=\mathrm{d}z/\mathrm{d}s$ are the directional cosines. Consequently, $\sqrt{(x')^2+(z')^2}=1$, which follows also directly from (5.2) if both sides are divided by ds. Hence, the equations of a seismic ray simplify to read

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{v}\frac{\mathrm{d}x}{\mathrm{d}s}\right)-\frac{\partial}{\partial x}\left(\frac{1}{v}\right)=0\,,\quad \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{v}\frac{\mathrm{d}z}{\mathrm{d}s}\right)-\frac{\partial}{\partial z}\left(\frac{1}{v}\right)=0}\ . \tag{5.12}$$

This form of the equations of a seismic ray has frequently been used in numerical solutions. Note that Eqs. (5.12) also follow immediately from Eqs. (5.10); it is sufficient to divide Eqs. (5.10) by v and to write $v\,\mathrm{d}\tau=\mathrm{d}s$.

5.2.5 Independent variable σ

In the case of the independent variable σ, we could proceed analogously as in the previous cases, and take into account that

$$\sqrt{(x')^2+(z')^2}=\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\sigma}\right)^2+\left(\frac{\mathrm{d}z}{\mathrm{d}\sigma}\right)^2}=\frac{1}{v}\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}s}\right)^2+\left(\frac{\mathrm{d}z}{\mathrm{d}s}\right)^2}=\frac{1}{v}\ .$$

This leads to a simpler form of the final equations. We can obtain the corresponding equations directly by dividing Eqs. (5.12) by v and writing $\mathrm{d}\sigma=v\,\mathrm{d}s$. We get

$$\boxed{\frac{\mathrm{d}^2x}{\mathrm{d}\sigma^2}-\frac{1}{2}\frac{\partial}{\partial x}\left(\frac{1}{v^2}\right)=0\,,\quad \frac{\mathrm{d}^2z}{\mathrm{d}\sigma^2}-\frac{1}{2}\frac{\partial}{\partial z}\left(\frac{1}{v^2}\right)=0}\ . \tag{5.13}$$

If function $1/v^2$ (quadratic slowness) is, e.g., a linear function of coordinates, simple analytic solutions for a ray can be found. However, the usual way of solving the equations of a seismic ray is a numerical solution. Various standard methods, e.g., the Runge-Kutta method or Hamming predictor-corrector method can be used for this purpose.

Let us demonstrate a simple analytical solution of Eqs. (5.13) on the case of the vertically inhomogeneous medium which has already been considered in Section 3.8, i.e. a medium with a linear quadratic slowness. Thus, assume $1/v^2$ to be of the form (3.51), i.e.

$$\frac{1}{v^2}=a-bz\ , \tag{5.14}$$

where a and b are constants. The solutions of Eqs. (5.13) can then be expressed as

$$x = k_1\sigma + k_2 \ , \quad z = -\frac{b}{4}\sigma^2 + q_1\sigma + q_2 \ , \tag{5.15}$$

k_1, k_2, q_1 and q_2 being constants. Consider a ray passing through the coordinate origin $x = z = 0$. At this point, put $\sigma = 0$ and denote by δ_0 the angle which the ray makes with the z-axis. Under these initial conditions, we obtain $k_2 = q_2 = 0$, and

$$k_1 = \left(\frac{\mathrm{d}\,x}{\mathrm{d}\,\sigma}\right)_{\sigma=0} = \left(\frac{\mathrm{d}\,x}{v\,\mathrm{d}\,s}\right)_{z=0} = \frac{\sin\delta_0}{v_0}\ , \quad q_1 = \left(\frac{\mathrm{d}\,z}{\mathrm{d}\,\sigma}\right)_{\sigma=0} = \frac{\cos\delta_0}{v_0}\ , \tag{5.16}$$

where v_0 is the velocity at $z = 0$, i.e. $v_0 = 1/\sqrt{a}$. Eliminating parameter σ from Eqs. (5.15), we arrive at a very simple equation of a seismic ray:

$$z = -\frac{b}{4k_1^2}x^2 + \frac{q_1}{k_1}x \ , \tag{5.17}$$

or

$$z = -\frac{b}{4a\sin^2\delta_0}x^2 + \left(\cot\delta_0\right)x \ . \tag{5.18}$$

Compare these equations with Eqs. (3.55), where we calculated epicentral distance x as a function of depth z. If the role of these coordinates is interchanged, even a simpler function, namely quadratic function (5.17), is obtained. Hence, it is very convenient to approximate a vertically inhomogeneous medium by a system of layers with the velocity distribution of the form (5.14) in each layer. The computation of seismic rays is then very fast.

5.3 Other Descriptions of Seismic Rays in 2-D Media; the Hamiltonian Approach

The equations of rays, derived in the previous section, are differential equations of the second order. They may be replaced by a system of differential equations of the first order. At first, let us again remind some analogies with analytical mechanics.

Lagrange's equations of the second kind, see Section 3.2, represent a system of n differential equations of the second order, where n is the number of degrees of freedom. These equations can be replaced by a system of $2n$ equations of the first order, called Hamilton's equation, in the following way.

Introduce generalised momentum p_j by the relation

$$p_j = \frac{\partial L}{\partial \dot{q}_j} \tag{5.19}$$

instead of generalised velocity $\dot{q}_j$, and the Hamiltonian $H = H(q_j, p_j, t)$,

$$H = \sum_{j=1}^{n} p_j \dot{q}_j - L \ , \tag{5.20}$$

instead of the Lagrangian $L = L(q_j, \dot{q}_j, t)$. For the derivatives of the Hamiltonian we get

$$\frac{\partial H}{\partial q_j} = -\frac{\partial L}{\partial q_j} = -\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_j}\right) = -\frac{\mathrm{d}p_j}{\mathrm{d}t} = -\dot{p}_j \ ,$$

$$\frac{\partial H}{\partial p_j} = \dot{q}_j + p_j \frac{\partial \dot{q}_j}{\partial p_j} - \frac{\partial L}{\partial \dot{q}_j}\frac{\partial \dot{q}_j}{\partial p_j} = \dot{q}_j \ ,$$

where we have used Lagrange's equation (3.7) and definition (5.19). In this way we arrive at Hamilton's equations,

$$\dot{q}_j = \frac{\partial H}{\partial p_j}, \quad \dot{p}_j = -\frac{\partial H}{\partial q_j} \ . \tag{5.21}$$

Using the analogies from analytical mechanics, we shall rewrite the equations of seismic rays, derived in the previous section, in Hamiltonian forms.

5.3.1 Independent variable *x*

Using function F in the form (5.4), we shall introduce

$$p_z = \frac{\partial F}{\partial z'} = \frac{z'}{v\sqrt{1+(z')^2}} \ , \tag{5.22}$$

where $z' = \mathrm{d}z/\mathrm{d}x$. Equations (5.22) and (5.6) can then be expressed as

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{p_z v}{\sqrt{1-p_z^2 v^2}}, \quad \frac{\mathrm{d}p_z}{\mathrm{d}x} = \frac{1}{\sqrt{1-p_z^2 v^2}}\frac{\partial}{\partial z}\left(\frac{1}{v}\right) \ . \tag{5.23}$$

Denote by δ the angle which the ray makes with the positive part of the z-axis. This means that δ is the angle of incidence on the horizontal plane which passes through the corresponding point; we consider angle δ in the bounds

$-\pi \le \delta \le \pi$. Then $z' = \cot\delta$, $\sqrt{1+(z')^2} = 1/\sin\delta$, $p_z = (\cos\delta)/v$, and Eqs. (5.23) take the form

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \cot\delta\,, \quad \frac{\mathrm{d}\delta}{\mathrm{d}x} = \frac{1}{v}\left(-\cot\delta\frac{\mathrm{d}v}{\mathrm{d}x} + \frac{1}{\sin^2\delta}\frac{\partial v}{\partial z}\right) . \tag{5.24}$$

5.3.2 Independent variable z

Now, let us introduce

$$p_x = \frac{x'}{v\sqrt{1+(x')^2}} , \tag{5.25}$$

where $x' = \mathrm{d}x/\mathrm{d}z$. Equations (5.25) and (5.7) yield the following equations:

$$\frac{\mathrm{d}x}{\mathrm{d}z} = \frac{p_x v}{\sqrt{1-p_x^2 v^2}} , \quad \frac{\mathrm{d}p_x}{\mathrm{d}x} = \frac{1}{\sqrt{1-p_x^2 v^2}}\frac{\partial}{\partial x}\left(\frac{1}{v}\right) , \tag{5.26}$$

which are very similar to Eqs. (5.23).

Since $x' = \tan\delta$, $p_x = (\sin\delta)/v$, we may write Eq. (5.26) also in the form

$$\frac{\mathrm{d}x}{\mathrm{d}z} = \tan\delta\,, \quad \frac{\mathrm{d}\delta}{\mathrm{d}z} = \frac{1}{v}\left(-\frac{1}{\cos^2\delta}\frac{\partial v}{\partial x} + \tan\delta\frac{\partial v}{\partial z}\right) . \tag{5.27}$$

5.3.3 Independent variable τ

If travel time τ is used as the independent variable, Eqs. (5.11) can be expressed as

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}\tau} &= v^2 p_x\,, \quad \frac{\mathrm{d}p_x}{\mathrm{d}\tau} = -\frac{1}{v}\frac{\partial v}{\partial x}\,, \\ \frac{\mathrm{d}z}{\mathrm{d}\tau} &= v^2 p_z\,, \quad \frac{\mathrm{d}p_z}{\mathrm{d}\tau} = -\frac{1}{v}\frac{\partial v}{\partial z}\,. \end{aligned} \tag{5.28}$$

Denoting an element along the ray by $\mathrm{d}s = v\,\mathrm{d}\tau$, we have

$$\sin\delta = \frac{\mathrm{d}x}{\mathrm{d}z} = \frac{1}{v}\frac{\mathrm{d}x}{\mathrm{d}\tau}\,, \quad \cos\delta = \frac{\mathrm{d}z}{\mathrm{d}s} = \frac{1}{v}\frac{\mathrm{d}z}{\mathrm{d}\tau}\,, \quad p_x = \frac{\sin\delta}{v}\,, \quad p_z = \frac{\cos\delta}{v}\,.$$

Equations (5.28) then yield

$$\boxed{\frac{\mathrm{d}x}{\mathrm{d}\tau} = v\sin\delta\,, \quad \frac{\mathrm{d}\delta}{\mathrm{d}\tau} = -\cos\delta\frac{\partial v}{\partial x} + \sin\delta\frac{\partial v}{\partial z}\,, \quad \frac{\mathrm{d}z}{\mathrm{d}\tau} = v\cos\delta}\,, \tag{5.29}$$

where we have eliminated term $\partial v/\partial \tau$ in $\mathrm{d}p_x/\mathrm{d}\tau$ and $\mathrm{d}p_z/\mathrm{d}z$, which gives only one equation for $\mathrm{d}\delta/\mathrm{d}\tau$. System (5.29) has frequently been used to compute seismic rays in 2-D media. The initial conditions for this system are as follows: $x = x_0$, $z = z_0$, and $\delta = \delta_0$ for $\tau = \tau_0$. This system consists of only three equations, has no singularities, and works quite reliably. The computation of the trigonometric functions, however, extends the computer time, which represents the only disadvantage of this system of equations.

5.3.4 Independent variable *s*

Putting $p_x = \dfrac{1}{v}\dfrac{\mathrm{d}x}{\mathrm{d}s}$ and $p_z = \dfrac{1}{v}\dfrac{\mathrm{d}z}{\mathrm{d}s}$, equations (5.12) may be expressed as

$$\frac{\mathrm{d}x}{\mathrm{d}s} = vp_x , \quad \frac{\mathrm{d}p_x}{\mathrm{d}s} = \frac{\partial}{\partial x}\left(\frac{1}{v}\right) ,$$
$$\frac{\mathrm{d}z}{\mathrm{d}s} = vp_z , \quad \frac{\mathrm{d}p_z}{\mathrm{d}s} = \frac{\partial}{\partial z}\left(\frac{1}{v}\right) . \tag{5.30}$$

Note that p_x and p_z, introduced here, differ from those in Eq. (5.28). Equations (5.30) also follow immediately from (5.28) if we divide Eqs. (5.28) by v and put $\mathrm{d}s = v\,\mathrm{d}\tau$.

5.3.5 Independent variable σ

Equations (5.13) can also be expressed as

$$\frac{\mathrm{d}x}{\mathrm{d}\sigma} = p_x , \quad \frac{\mathrm{d}p_x}{\mathrm{d}\sigma} = \frac{1}{2}\frac{\partial}{\partial x}\left(\frac{1}{v^2}\right) ,$$
$$\frac{\mathrm{d}z}{\mathrm{d}\sigma} = p_z , \quad \frac{\mathrm{d}p_z}{\mathrm{d}\sigma} = \frac{1}{2}\frac{\partial}{\partial z}\left(\frac{1}{v^2}\right) . \tag{5.31}$$

This is also a convenient system for computing seismic rays. Since $\mathrm{d}\sigma = v\,\mathrm{d}s = v^2\,\mathrm{d}\tau$, in order to compute the travel time, we must add the equation

$$\frac{\mathrm{d}\tau}{\mathrm{d}\sigma} = \frac{1}{v^2} . \tag{5.32}$$

Chapter 6

Basic Concepts and Formulae of Continuum Mechanics

In this chapter we shall summarise the basic concepts and equations of the theory of elasticity which are required in the theory of seismic wave propagation. For details we refer the reader, e.g., to Fung (1965, 1969), and the lecture notes by Novotny (1999). This chapter represents a simplified version of the corresponding chapter from these lecture notes.

6.1 Mathematical Models in Physics

In order to simplify the mathematical and physical description of studied phenomena, various simplifications and models are used. The usual idealisations of material objects in mechanics are the mass point (particle), rigid body, and continuum. The model of a continuum is used in mechanics when the deformations of a body cannot be neglected.

The *continuum* in mechanics is a medium with a *continuous distribution of matter*. The molecular and atomic structures of matter are ignored in this model of the medium. The main *advantage* of the concept of a continuum consists in the possibility of applying the mathematical theory of *continuous functions*, and *differential* and *integral* calculi.

When the fine structure of matter attracts our attention, continuum mechanics cannot be used. In these cases we should use particle physics and statistical physics.

6.2 Displacement Vector

Real bodies are deformed by the action of forces. The description of the deformation is based on a comparison of the instantaneous state (volume and shape) of the body with some previous state, which will be regarded as an original state. In this section we shall study the corresponding displacements.

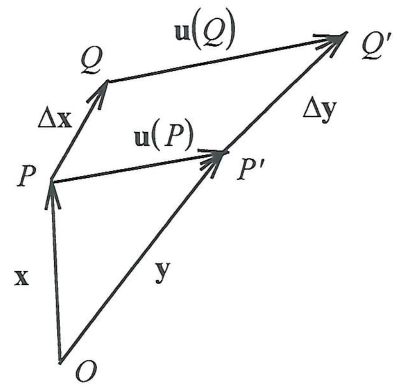


Fig. 6.1. Displacements of two neighbouring points, P and Q.

Therefore, we shall compare a continuum in two states, namely in the *original* (unstrained) state, and in the *deformed* (new, strained) state. Consider a particle at point P in the original state, which is moved to point P' in the deformed state (Fig. 6.1). Denote the radius vector of point P by $\mathbf{x} = (x_1, x_2, x_3)$, and of point P' by $\mathbf{y} = (y_1, y_2, y_3)$. The new position, given by vector $\mathbf{y}$, depends on the initial position $\mathbf{x}$, on the acting forces, physical properties of the continuum and the time between the original and new states.

The displacement of a particle from an original to a deformed position can be described by the corresponding *displacement vector* $\mathbf{u} = (u_1, u_2, u_3)$,

$$\mathbf{u} = \mathbf{y} - \mathbf{x} \ . \tag{6.1}$$

We shall usually consider the displacement vector as a function of the coordinates of the original state:

$$\mathbf{u} = \mathbf{y}(\mathbf{x}) - \mathbf{x}, \quad \text{i.e.} \quad \mathbf{u} = \mathbf{u}(\mathbf{x}) \ . \tag{6.2}$$

In this case we speak of the *Lagrangian description* of motion.

However, we can also express the displacement vector as a function of the coordinates of the deformed state, $\mathbf{u} = \mathbf{y} - \mathbf{x} = \mathbf{y} - \mathbf{x}(\mathbf{y})$, i.e. $\mathbf{u} = \mathbf{u}(\mathbf{y})$. In this case we speak of the *Eulerian description.* This description is frequently used in hydrodynamics. Here we shall use the Lagrangian description, with exceptions in Section 6.4.

In a neighbourhood of point P, let us consider another point, Q, which will be displaced to point Q' in the deformed state (Fig. 6.1). The radius vectors of points Q and Q' are $\mathbf{x} + \Delta\mathbf{x}$ and $\mathbf{y} + \Delta\mathbf{y}$, respectively. Using the Taylor expansion, we get

$$u_j(Q) = u_j(x_1 + \Delta x_1, x_2 + \Delta x_2, x_3 + \Delta x_3) =$$
$$= u_j(x_1, x_2, x_3) + \sum_{k=1}^{3} \left(\frac{\partial u_j}{\partial x_k} \right)_{(x_1, x_2, x_3)} \Delta x_k + \ldots = u_j(P) + \sum_{k=1}^{3} \left(\frac{\partial u_j}{\partial x_k} \right)_P \Delta x_k + \ldots \ , \tag{6.3}$$

where $j = 1, 2, 3$.

To simplify the formulae which follow, let us introduce Einstein's summation convention: If any suffix occurs twice in a single term, it is to be put equal to 1, 2 and 3 in turn and the results added.

Using this summation convention and neglecting the higher-order terms in (6.3), we get approximately

$$u_j(Q) = u_j(P) + \left(\frac{\partial u_j}{\partial x_k} \right)_P \Delta x_k \ . \tag{6.4}$$

We shall *assume* that the *displacement vector* and its *first derivatives are continuous* functions of coordinates. The continuity of displacement **u** guarantees that an originally continuous body will also remain continuous during the deformation. The continuity of $\partial u_j / \partial x_k$ guarantees the existence of the total differential of the displacement. Consequently, formula (6.4) can then be made as accurate as required by choosing point Q sufficiently close to P. This formula will play an important role in the theory which follows.

6.3 Strain Tensor

If the displacement is known for every particle in a body, we can construct the deformed body from the original. Hence, a deformation can be described by the displacement field. However, the displacement vector describes the translation, rotation and pure deformation (strain) of the medium. But we are not interested in translation and rotation; these motions are studied in detail in the mechanics of rigid bodies. We are only interested in those quantities which characterise the strain. There are two approaches to obtaining these characteristics:

1) subtracting the translation and rotation from the displacement;
2) considering changes in distances.

We shall use the second approach because this approach is more general.

6.3.1 Tensor of finite strain

It is evident that the change in the size and shape of a body will be determined in full if the changes in the distances of two arbitrary points are known. However, it will be more convenient to consider the squares of these distances instead of the distances themselves. Therefore, we shall characterise the deformations by quantities ε_{ij} which are defined by the relation

$$\overline{P'Q'}^2 - \overline{PQ}^2 = 2\varepsilon_{ij}\Delta x_i \Delta x_j \ , \tag{6.5}$$

where $\overline{PQ}$ is the distance between points P and Q in the original state, and $\overline{P'Q'}$ is the distance between the corresponding particles in the deformed state (Fig. 6.1).

The square of distance $\overline{PQ}$ can be expressed as (if the summation convention is used)

$$\overline{PQ}^2 = \Delta\mathbf{x} \cdot \Delta\mathbf{x} = \Delta x_i \Delta x_i \ . \tag{6.6}$$

It follows from the quadrangle $PP'Q'Q$ and Eq. (6.4) that

$$\mathbf{u}(P) + \Delta\mathbf{y} = \Delta\mathbf{x} + \mathbf{u}(Q) = \Delta\mathbf{x} + \mathbf{u}(P) + \left(\frac{\partial \mathbf{u}}{\partial x_i}\right)_P \Delta x_i \ .$$

By comparing the beginning and end of this equation, we see that

$$\Delta \mathbf{y} = \Delta \mathbf{x} + \left(\frac{\partial \mathbf{u}}{\partial x_i} \right)_P \Delta x_i \ . \tag{6.7}$$

We shall omit suffix P hereafter. Formula (6.7) can then be expressed in components as

$$\Delta y_i = \Delta x_i + \frac{\partial u_i}{\partial x_j} \Delta x_j \ . \tag{6.8}$$

Consequently,

$$\overline{P'Q'}^2 = \Delta \mathbf{y} \cdot \Delta \mathbf{y} = \Delta y_k \Delta y_k = \left(\Delta x_k + \frac{\partial u_k}{\partial x_i} \Delta x_i \right) \left(\Delta x_k + \frac{\partial u_k}{\partial x_j} \Delta x_j \right) . \tag{6.9}$$

Note that we have used different dummy indices, i and j, in the latter formula. Thus,

$$\overline{P'Q'}^2 - \overline{PQ}^2 = \frac{\partial u_k}{\partial x_i} \Delta x_i \Delta x_k + \frac{\partial u_k}{\partial x_j} \Delta x_k \Delta x_j + \frac{\partial u_k}{\partial x_i} \frac{\partial u_k}{\partial x_j} \Delta x_i \Delta x_j \ .$$

Change again the dummy indices to obtain the products $\Delta x_i \Delta x_j$ in all terms:

$$\overline{P'Q'}^2 - \overline{PQ}^2 = \frac{\partial u_j}{\partial x_i} \Delta x_i \Delta x_j + \frac{\partial u_i}{\partial x_j} \Delta x_i \Delta x_j + \frac{\partial u_k}{\partial x_i} \frac{\partial u_k}{\partial x_j} \Delta x_i \Delta x_j \ ;$$

By comparing this expression with (6.5), we arrive at

$$\varepsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} + \frac{\partial u_k}{\partial x_i} \frac{\partial u_k}{\partial x_j} \right) . \tag{6.10}$$

The array of nine quantities ε_{ij} is called the *tensor of finite strain*, and the individual quantities ε_{ij} are called the *components of the tensor of finite strain*. This tensor is symmetric, i.e. $\varepsilon_{ij} = \varepsilon_{ji}$.

Since the derivatives of the displacement vector have been calculated at point P, see (6.7), we shall also regard components ε_{ij} as defined at point P, and speak of the tensor of finite strain at point P.

6.3.2 Tensor of infinitesimal strain

The tensor of finite strain contains products of the derivatives of the displacement vector, $\partial u_i / \partial x_j$. These products represent non-linear terms, which complicate the solution of many problems. However, in many applications, these quadratic terms may be neglected.

We shall assume hereafter that the derivatives of the displacement are small, i.e.

$$\left|\frac{\partial u_i}{\partial x_j}\right| << 1 \ , \qquad (6.11)$$

so that their mutual products are small quantities of the second order, which may be neglected in comparison with the derivatives themselves. In this case, the tensor of finite strain ε_{ij} simplifies to yield the tensor

$$\boxed{e_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)} \ , \qquad (6.12)$$

which is called the *tensor of infinitesimal strain.* In speaking of the strain tensor only, we shall have in mind the tensor of infinitesimal strain (6.12).

6.3.3 Physical meaning of the components of the strain tensors

Consider an elementary abscissa, *PQ*, which is parallel to the x_1-axis in the original state, i.e. $\Delta\mathbf{x} = (\Delta x_1, 0, 0)$; see Fig. 6.2. As $\Delta x_2 = \Delta x_3 = 0$, Eq. (6.5) takes the simple form

$$|\Delta\mathbf{y}|^2 - |\Delta\mathbf{x}|^2 = 2\varepsilon_{11}(\Delta x_1)^2 \ .$$

Consequently,

$$|\Delta\mathbf{y}| = \sqrt{1 + 2\varepsilon_{11}}\,\Delta x_1 \ . \qquad (6.13)$$

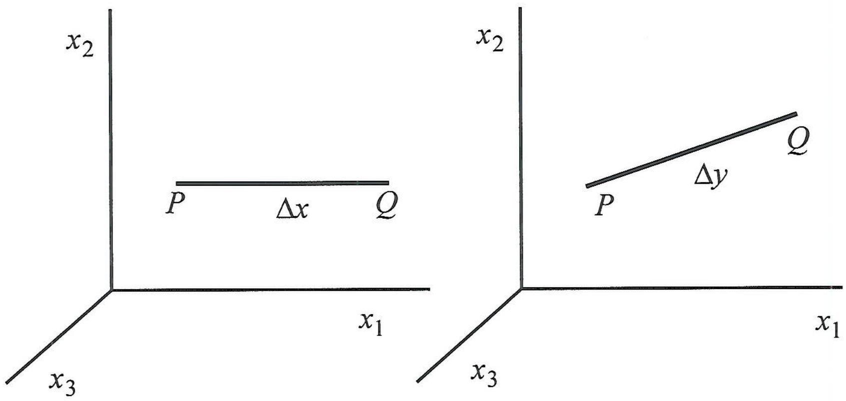


Fig. 6.2. Physical meaning of ε_{11}.

The relative extension of the abscissa *PQ* is defined by

$$E_1 = \frac{|\Delta\mathbf{y}| - |\Delta\mathbf{x}|}{|\Delta\mathbf{x}|} \ . \qquad (6.14)$$

Using (6.13), this extension can be expressed as

$$E_1 = \sqrt{1+2\varepsilon_{11}} - 1 \ . \tag{6.15}$$

Hence, component ε_{11} characterises the relative extension of an element which was originally parallel to the x_1-axis. Analogously, components ε_{22} and ε_{33} characterise the extensions along the second and third axes, respectively.

Furthermore, if ε_{11} is small and the higher-order terms are neglected, Eq. (6.15) simplifies to read

$$E_1 = \sqrt{1+2\varepsilon_{11}} - 1 \approx 1 + \varepsilon_{11} - 1 = \varepsilon_{11} \approx e_{11} \ . \tag{6.16}$$

Thus, in the case of small deformations, components e_{11}, e_{22} and e_{33} are equal to the relative extensions of the line elements which, in the original state, were parallel to the coordinate axes.

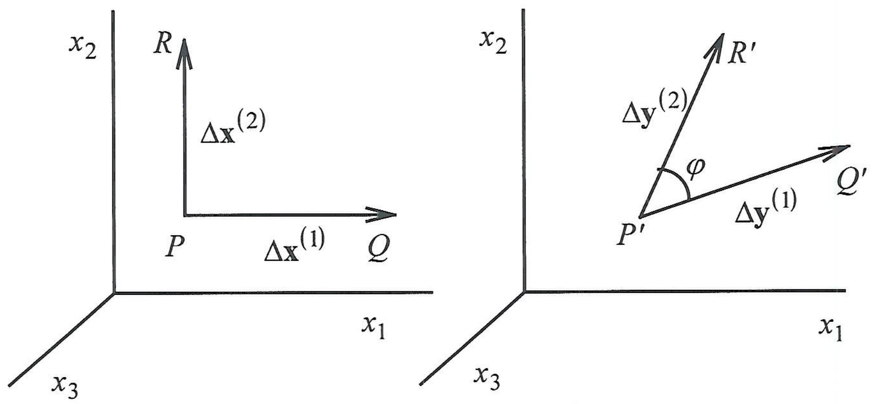


Fig. 6.3. Physical meaning of ε_{12}.

Now let us consider two perpendicular vectors in the original state, $\Delta\mathbf{x}^{(1)} = (\Delta x_1, 0, 0)$ and $\Delta\mathbf{x}^{(2)} = (0, \Delta x_2, 0)$; see Fig. 6.3. The components of the corresponding vectors $\Delta\mathbf{y}^{(1)}$ and $\Delta\mathbf{y}^{(2)}$ in the deformed state can easily be obtained from Eq. (6.8). For the scalar product of these vectors we then get

$$\Delta\mathbf{y}^{(1)} \cdot \Delta\mathbf{y}^{(2)} = \Delta y_i^{(1)} \Delta y_i^{(2)} = 2\varepsilon_{12} \Delta x_1 \Delta x_2 \ ;$$

a detailed derivation can be found in Novotny (1999). However, this scalar product can also be expressed as

$$\Delta\mathbf{y}^{(1)} \cdot \Delta\mathbf{y}^{(2)} = \left|\Delta\mathbf{y}^{(1)}\right| \left|\Delta\mathbf{y}^{(2)}\right| \cos\varphi \ ,$$

φ being the angle between vectors $\Delta\mathbf{y}^{(1)}$ and $\Delta\mathbf{y}^{(2)}$. Comparing both expressions for the scalar product, and using (6.13) to express $\left|\Delta\mathbf{y}^{(1)}\right|$ and $\left|\Delta\mathbf{y}^{(2)}\right|$, we arrive at

$$\cos\varphi = \frac{2\varepsilon_{12}}{\sqrt{1+2\varepsilon_{11}}\sqrt{1+2\varepsilon_{22}}} \ . \tag{6.17}$$

Introduce the angle $\alpha_{12} = 90^\circ - \varphi$, which describes the change of the right angle (decrease of the right angle) due to deformation. For small deformations, formula (6.17) then yields

$$\sin\alpha_{12} = \cos\varphi \approx 2\varepsilon_{12} \approx 2e_{12} \ .$$

Consequently, $\sin\alpha_{12}$ is small and may be approximated by α_{12}, so that

$$\alpha_{12} \approx 2e_{12} \ . \tag{6.18}$$

Thus, component e_{12} is equal to half the change of the right angle between two line elements, one of which was parallel in the original state to the x_1-axis, and the second was parallel to the x_2-axis. The physical meaning of the remaining components e_{13} and e_{23} is analogous.

6.3.4 Volume dilatation

Consider a small parallelepiped in the original state, the edges of which are parallel to the coordinate axes, and have lengths d_1, d_2, d_3, respectively. The volume of the parallelepiped is $V = d_1 d_2 d_3$. In the deformed state, these edges will have the lengths (neglecting higher-order terms)

$$d_1 + e_{11}d_1, \quad d_2 + e_{22}d_2, \quad d_3 + e_{33}d_3 \ ,$$

respectively. Therefore, the new volume will be

$$V' = d_1 d_2 d_3 (1+e_{11})(1+e_{22})(1+e_{33}) = V(1+e_{11}+e_{22}+e_{33}) \ . \tag{6.19}$$

The *volume dilatation* (cubical dilatation), defined by

$$\vartheta = \frac{V'-V}{V} \ , \tag{6.20}$$

then reads

$$\boxed{\vartheta = e_{11}+e_{22}+e_{33}} \ . \tag{6.21}$$

Strictly speaking, the new volume can be expressed by formula (6.19) only if the edges of the parallelepiped coincide with the so-called principal axes of strain (these axes remain perpendicular also after the deformation). However, it can be shown that the sum $e_{11} + e_{22} + e_{33}$ is an invariant, i.e. a quantity which is independent of the choice of the coordinate system. Consequently, quantity ϑ describes the relative change of an arbitrary infinitesimal volume which surrounds the considered point.

6.4 Stress Vector and Related Problems

6.4.1 Body forces and surface forces

In particle mechanics, we study two types of interactions between particles: by action at a distance and by collision. An analogous division of forces is convenient also in continuum mechanics. Therefore, we shall divide the forces acting in a continuum into two groups according to their "action radius":

1) *Body forces*, also called *voluminal forces*, which have a large action radius. Examples of body forces are gravitational forces, electromagnetic forces, inertial force (in dynamic problems), and also fictitious forces in non-inertial reference frames (Coriolis and centrifugal forces).
2) *Surface forces*, which have a small action radius. Examples of such forces are hydrostatic pressure, aerostatic pressure, and forces due to the mechanical contact of two bodies.

This separation of forces facilitates the formulation and solution of many problems because:

1) the effect of forces with a small action radius may be approximated by a surface integral (surface forces) instead of a more complicated volume integral;
2) body forces vanish in some limits, and may also be neglected in some problems, e.g., in many problems of elastic wave propagation.

6.4.2 Stress vector

A deformed continuum at rest resembles a rigid body. Therefore, we shall *assume* that some notions and equations from rigid-body mechanics can also be applied in continuum mechanics. However, these analogies will be no more than basic assumptions. This approach will only facilitate the formulation of the basic equations of continuum mechanics, but cannot be regarded as a derivation of these equations. Namely, the general equations of continuum mechanics cannot, in principle, be derived from more special equations for a rigid body or a mass point. The validity of the general equations can be verified only by comparing their solutions with experiments.

Let us start with the description of the stress state in a continuum. Consider a point, P, and an element of a surface, ΔS, drawn through this point (Fig. 6.4). Denote the normal to ΔS at point P by $\vec{\nu}$. Vector $\vec{\nu}$ enables us to define the positive and negative sides of the element ΔS (upper and lower sides in Fig. 6.4, respectively).

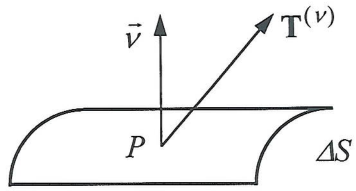


Fig. 6.4. Stress vector.

In analogy to the static equilibrium of a rigid body, we shall *assume* that, in a deformed continuum at rest, the effect of all surface forces exerted across the small element ΔS is statically equivalent to a single force $\Delta\mathbf{H}$, acting at point P in a definite direction, together with couple $\Delta\mathbf{G}$, acting also at P about a definite axis.

Let us indefinitely diminish surface element ΔS by any continuous process, always keeping point P within the element. From physical considerations it seems reasonable to *assume* that vector $\Delta\mathbf{H}/\Delta S$ tends to a non-zero limit,

$$\mathbf{T}^{(\nu)} = \lim_{\Delta S\to 0} \frac{\Delta\mathbf{H}}{\Delta S} = \frac{\mathrm{d}\,\mathbf{H}}{\mathrm{d}\,S} , \tag{6.22}$$

whereas vector $\Delta\mathbf{G}/\Delta S$ tends to the zero vector. Vector $\mathbf{T}^{(\nu)}$ is called *the stress vector* or traction at point P; see Fig. 6.4. Note that the direction of stress $\mathbf{T}^{(\nu)}$ need not coincide with the direction of normal $\vec{\nu}$. Vector $\mathbf{T}^{(\nu)}$ is the vector acting on the unit infinitesimal surface, the normal of which is $\vec{\nu}$. The stress at P varies, in general, with the direction of normal $\vec{\nu}$. Vector $\mathbf{T}^{(\nu)}$ can be decomposed into a normal component (in the direction of $\vec{\nu}$) and a tangential (shear) component, which is perpendicular to $\vec{\nu}$. We then speak of normal and tangential (shear) stresses, respectively.

Analogously, we shall introduce the body force, $\mathbf{F}$, acting on the unit infinitesimal vicinity of point P.

6.4.3 Conditions of equilibrium in integral form

It is well-known from rigid-body mechanics that a rigid body is in static equilibrium if the total applied force and total applied torque are zero.

We shall *assume* that an arbitrary part of a continuum, in the deformed state at rest, is in equilibrium under the same conditions as if this part were a rigid body. This means that we shall express these conditions of equilibrium in the following form:

$$\iint_S \mathbf{T}^{(\nu)}\,\mathrm{d}\,S + \iiint_V \mathbf{F}\,\mathrm{d}\,V = 0 , \tag{6.23}$$

$$\iint_S \left(\mathbf{y} \times \mathbf{T}^{(\nu)}\right) \mathrm{d}S + \iiint_V (\mathbf{y} \times \mathbf{F}) \mathrm{d}V = 0 \ , \tag{6.24}$$

where V is the volume of the part of the continuum, S is its surface, $\mathbf{T}^{(\nu)}$ is the surface force acting from the side of the outward normal $\vec{\nu}$, $\mathbf{F}$ is the body force, and $\mathbf{y}$ is the radius vector of the point under consideration (in the deformed state). The first of these equations requires the resultant force to be equal to zero, and the second equation requires the resultant torque to be equal to zero.

6.4.4 Equations of motion in integral form

Using D'Alembert's principle, the equations of motion can easily be obtained from the conditions of equilibrium by adding the inertial forces.

Consider any portion of a material body. Let the volume of this portion at any time t be denoted by $V = V(t)$. Let $\mathbf{y}$ be the radius-vector of a particle, $\mathbf{v}$ be its velocity, and ρ be the density of the material at the corresponding point. Integral

$$\mathbf{P} = \iiint_V \rho \mathbf{v} \, \mathrm{d}V$$

is the linear momentum, and

$$\mathbf{L} = \iiint_V (\mathbf{y} \times \rho \mathbf{v}) \, \mathrm{d}V$$

is the angular momentum of this part of the body. Derivative $\mathrm{d}\mathbf{P}/\mathrm{d}t$ is the corresponding inertial force.

Hence, by adding the inertial terms on the right-hand sides of Eqs. (6.23) and (6.24), we arrive at the *equations of motion of a continuum* in the form

$$\iint_S \mathbf{T}^{(\nu)} \, \mathrm{d}S + \iiint_V \mathbf{F} \, \mathrm{d}V = \frac{\mathrm{d}}{\mathrm{d}t} \iiint_V \rho \mathbf{v} \, \mathrm{d}V \qquad , \tag{6.25}$$

$$\iint_S \left(\mathbf{y} \times \mathbf{T}^{(\nu)}\right) \mathrm{d}S + \iiint_V (\mathbf{y} \times \mathbf{F}) \mathrm{d}V = \frac{\mathrm{d}}{\mathrm{d}t} \iiint_V (\mathbf{y} \times \rho \mathbf{v}) \mathrm{d}V \, . \tag{6.26}$$

It should be noted that no demand was made on domain $V(t)$ other than that it must consist of the same material particles at all times. Equations (6.25) and (6.26) are applicable to any material body which may be considered as a continuum. Boundary surface S may coincide with the external boundary of the body, but it may also include only a small portion thereof..

Equations (6.25) and (6.26) represent the linear momentum theorem and the angular momentum theorem, respectively, applied to an arbitrary part of a continuum in the deformed configuration. These equations are also referred to as the *laws of motion of a continuum*, since they are considered to be valid generally.

6.5 Stress Tensor

6.5.1 Components of the stress tensor

In the previous section we introduced the basic assumption that the action of forces with a small "action radius" (surface forces) across any infinitesimal surface element can be described by a stress vector. Thus, to describe the stress state at a point, it is necessary to know the *stresses* acting *on all infinitesimal surfaces* drawn through this point. This means that surfaces of any shape should be considered.

To simplify the problem, we shall further assume that we can restrict ourselves to plane surfaces only. Thus, we adopt another assumption that the stress state at a point will be described if the *stresses* acting on *all plane infinitesimal surfaces* drawn through this point are known. We shall show that it will be even sufficient to know these stresses only on three perpendicular plane elements.

Consider plane element ΔS which is perpendicular to the i-th coordinate axis, so that its normal $\vec{\nu}$ is parallel to the i-th axis, and has the same orientation as this axis. Let $\mathbf{T}^{(i)} = \left(T_1^{(i)}, T_2^{(i)}, T_3^{(i)}\right)$ be the stress vector acting on this plane element, and introduce a new notation for its elements,

$$\boxed{\tau_{ij} = T_j^{(i)}}, \tag{6.27}$$

where $i, j = 1, 2, 3$; see Fig 6.5. The array of nine quantities τ_{ij} will be called the *stress tensor*, and the individual quantities τ_{ij} will be called the *components of the stress tensor*.

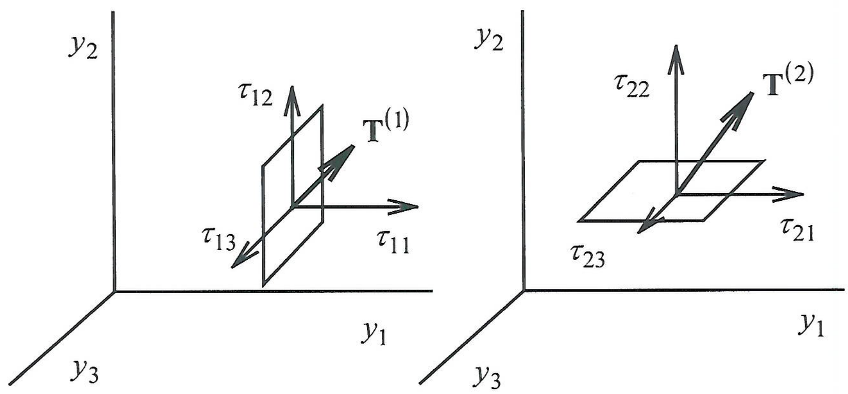


Fig. 6.5. Introduction of the components of the stress tensor.

Let us repeat the meaning of the individual subscripts in component τ_{ij}. Subscript i indicates that the corresponding plane element is perpendicular to the i-th axis, i.e. its normal $\vec{\nu}$ is parallel to the i-th axis. Subscript j denotes the j-th component of the corresponding force. For example, $\tau_{11,}$ τ_{12} and τ_{13} are the components of the force acting on a surface element which is perpendicular to the first axis.

6.5.2 Cauchy's formula

Now, we shall show that the nine components of stress tensor τ_{ij} are sufficient to describe the stress state at a particular point. Let us consider point P and an arbitrary, infinitesimal plane element drawn through this point (Fig. 6.6). Denote the unit vector which is normal to the element by $\vec{\nu} = (\nu_1, \nu_2, \nu_3)$, and the stress vector acting on the element by $\mathbf{T}^{(\nu)}$.

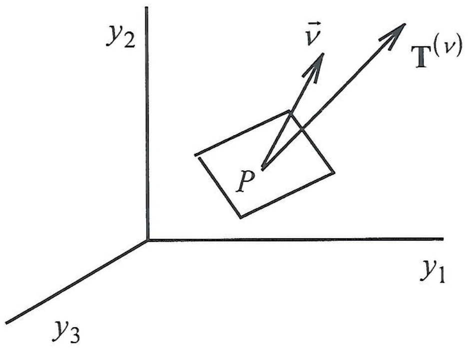


Fig. 6.6. Stress vector acting on a plane element.

If normal $\vec{\nu}$ coincides, e.g. , with the direction of the first axis, then

$$\mathbf{T}^{(\nu)}(P) = \mathbf{T}^{(1)}(P) = (\tau_{11}, \tau_{12}, \tau_{13})_P \;;$$

see definition (6.27). Thus, in this case, stress vector $\mathbf{T}^{(\nu)}$ is described by three components of vector $\mathbf{T}^{(1)}$. Analogously, if normal $\vec{\nu}$ coincides with the direction of the second or third axes, the corresponding vectors $\mathbf{T}^{(\nu)}$ are equal to $\mathbf{T}^{(2)}$ and $\mathbf{T}^{(3)}$, respectively.

For a general orientation of normal $\vec{\nu}$, it seems to be evident that the corresponding stress tensor $\mathbf{T}^{(\nu)}$ must be a combination of all three vectors $\mathbf{T}^{(1)}$,$\mathbf{T}^{(2)}$, $\mathbf{T}^{(3)}$, and that the contribution of these vectors should be weighted according to the projections of vector $\vec{\nu}$ onto the individual axes. Therefore,

$$\mathbf{T}^{(\nu)} = \mathbf{T}^{(1)}\nu_1 + \mathbf{T}^{(2)}\nu_2 + \mathbf{T}^{(3)}\nu_3 \;,$$

where we have omitted the letter P. For the i-th component of the stress vector we thus get

$$T_i^{(\nu)} = \tau_{1i}\nu_1 + \tau_{2i}\nu_2 + \tau_{3i}\nu_3 = \tau_{ji}\nu_j \ . \tag{6.28}$$

A more exact proof of this relation, based on the application of the condition of equilibrium (6.23), can be found in the literature mentioned at the beginning of this chapter.

Hence, the stress vector $\mathbf{T}^{(\nu)}$ acting on a surface element with unit normal $\vec{\nu}$ is completely determined by the components of stress tensor τ_{ji}. Consequently, the stress state at a point is described in full by the nine components τ_{ji}.

We shall see below that the stress tensor is symmetric (in usual materials), i.e. $\tau_{ij} = \tau_{ji}$. Consequently, Eq. (6.28) will usually be expressed as

$$T_i^{(\nu)} = \tau_{ij}\nu_j \ . \tag{6.29}$$

This formula is referred to as Cauchy's formula.

6.5.3 Conditions of equilibrium in differential form

The conditions of equilibrium in differential form can be derived from the integral conditions of equilibrium in several ways. In the elementary derivation, the equilibrium of an infinitesimal parallelepiped is usually considered (Fung, 1969). Here we shall give a shorter derivation which is based on the application of Gauss' theorem.

Gauss' theorem can be expressed as

$$\iiint\limits_V \operatorname{div}\mathbf{A}\,\mathrm{d}V = \iint\limits_S A_n\,\mathrm{d}S = \iint\limits_S \mathbf{A}\vec{\nu}\,\mathrm{d}S = \iint\limits_S A_j\nu_j\,\mathrm{d}S \ , \tag{6.30}$$

where $\mathbf{A}$ is a continuous vector with continuous derivatives, and $\vec{\nu}$ is the unit outward normal. Denoting the radius vector by $\mathbf{y} = (y_1, y_2, y_3)$ and using $\operatorname{div}\mathbf{A} = \partial A_j / \partial y_j$, we arrive at another form of Gauss' theorem:

$$\iiint\limits_V \frac{\partial A_j}{\partial y_j}\,\mathrm{d}V = \iint\limits_S A_j\nu_j\,\mathrm{d}S \ . \tag{6.31}$$

The i-th component of the integral condition of equilibrium (6.23) is

$$\iint\limits_S T_i^{(\nu)}\,\mathrm{d}S + \iiint\limits_V F_i\,\mathrm{d}V = 0 \tag{6.32}$$

or, using (6.28),

$$\iint\limits_S \tau_{ji}\nu_j\,\mathrm{d}S + \iiint\limits_V F_i\,\mathrm{d}V = 0 \ . \tag{6.33}$$

Putting $A_j = \tau_{ji}$ and using Gauss' theorem (6.31), the surface integral in (6.33) may be expressed as a volume integral:

$$\iiint_V \left(\frac{\partial \tau_{ji}}{\partial y_j} + F_i \right) \mathrm{d}V = 0 \ . \tag{6.34}$$

Since the integrand in (6.34) is assumed to be continuous, and volume V is arbitrary, integral (6.34) will be equal to zero only if the integrand is also equal to zero (see the explanation below). This yields the condition of equilibrium in the form

$$F_i + \frac{\partial \tau_{ji}}{\partial y_j} = 0 \ . \tag{6.35}$$

Let us explain how we have proceeded from Eq. (6.34) to Eq. (6.35) in greater detail. Assume that there is a point P^* where the integrand in (6.34) is non-zero, say, positive. Since we assume this integrand to be continuous, there is a vicinity, V^*, of point P^* where the integrand is also positive. The integral taken over V^* is then positive. This contradicts Eq.(6.34), which must be satisfied for any volume. This means that the integrand must be zero.

Now, let us consider the second integral condition of equilibrium, i.e. Eq. (6.24). For example, for the first component we have

$$\iint_S \left(y_2 T_3^{(\nu)} - y_3 T_2^{(\nu)} \right) \mathrm{d}S + \iiint_V \left(y_2 F_3 - y_3 F_2 \right) \mathrm{d}V = 0 \ . \tag{6.36}$$

Rearrange the surface integral in this equation by means of Cauchy's formula (6.29), Gauss' theorem (6.31) and the condition of equilibrium (6.35):

$$\iint_S \left(y_2 T_3^{(\nu)} - y_3 T_2^{(\nu)} \right) \mathrm{d}S = \iint_S \left(y_2 \tau_{j3} \nu_j - y_3 \tau_{j2} \nu_j \right) \mathrm{d}S =$$

$$= \iiint_V \left[\frac{\partial (y_2 \tau_{j3})}{\partial y_j} - \frac{\partial (y_3 \tau_{j2})}{\partial y_j} \right] = \iiint_V \left[\tau_{23} + y_2(-F_3) - \tau_{32} - y_3(-F_2) \right] \mathrm{d}V \ .$$

After inserting this expression into Eq. (6.36), since several terms vanish, we get

$$\iiint_V \left(\tau_{23} - \tau_{32} \right) \mathrm{d}V = 0 \ .$$

As the stress tensor is assumed to be continuous, we arrive at

$$\tau_{23} = \tau_{32} \ .$$

From the second and third components of Eq. (6.24), we would obtain $\tau_{13} = \tau_{31}$ and $\tau_{12} = \tau_{21}$, respectively. Consequently, we arrive at the condition of symmetry of the stress tensor,

$$\tau_{ij} = \tau_{ji} \ . \tag{6.37}$$

We have seen that the integral condition of equilibrium (6.24) does not yield a new differential equation, but only the condition of symmetry of the stress tensor. Using this symmetry, the conditions of equilibrium (6.35) can be expressed as

$$F_i + \frac{\partial \tau_{ij}}{\partial y_j} = 0 \ . \tag{6.38}$$

These conditions relate to the deformed state, i.e. to the Eulerian coordinates. If the difference between the Lagrangian and Eulerian coordinates may be neglected (see below), we can express the *conditions of equilibrium* also as

$$\boxed{F_i + \frac{\partial \tau_{ij}}{\partial x_j} = 0} \ . \tag{6.39}$$

The conditions of equilibrium are frequently used in this form.

Let us explain when the conditions of equilibrium (6.38) may be replaced by (6.39). Since $y_k = x_k + u_k$, the following relation holds between the derivatives of the stress tensor in Lagrangian and Eulerian coordinates:

$$\frac{\partial \tau_{mn}}{\partial x_j} = \frac{\partial \tau_{mn}}{\partial y_j} + \frac{\partial \tau_{mn}}{\partial y_k}\frac{\partial u_k}{\partial x_j} \ . \tag{6.40}$$

Assuming that the products of the derivatives are small and may be neglected, relation (6.40) simplifies to read

$$\frac{\partial \tau_{mn}}{\partial x_j} = \frac{\partial \tau_{mn}}{\partial y_j} \ . \tag{6.41}$$

This means that the difference between the Lagrangian and Eulerian descriptions vanishes in this case.

6.5.4 Equations of motion in differential form

The inertial force per unit volume is

$$\mathbf{F}_{iner} = -\frac{\mathrm{d}}{\mathrm{d}t}(\rho \mathbf{v}) \ , \tag{6.42}$$

ρ being the density, $\mathbf{v}$ the velocity, and t time. Assume that the time variations of density ρ may be neglected. Velocity $\mathbf{v}$, in Lagrangian coordinates, is a function of the form $\mathbf{v} = \mathbf{v}(x_1, x_2, x_3, t)$, where coordinates x_1, x_2, x_3 describe the original position, i.e. they are independent of time t. Consequently, the total derivative with respect to time is equal to the corresponding partial derivative:

$$\mathbf{F}_{iner} = -\rho \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\rho \frac{\partial \mathbf{v}}{\partial t} = -\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} , \tag{6.43}$$

where $\mathbf{u}$ is the displacement vector.

According to d'Alembert's principle, the equation of motion can be obtained from the condition of equilibrium, in our case from Eq. (6.39), by adding the inertial force. This yields the *equations of motion* of a continuum in the following final form:

$$\boxed{F_i + \frac{\partial \tau_{ij}}{\partial x_j} = \rho \frac{\partial^2 u_i}{\partial t^2}} . \tag{6.44}$$

This is one of the most important equations in continuum mechanics, and the basic equation in the theory of elastic waves.

To complete the description, let us also give the equations of motion in Eulerian coordinates. Since these coordinates are also functions of time, the inertial force contains further terms, and the equation of motion in these coordinates takes the form

$$F_i + \frac{\partial \tau_{ij}}{\partial y_j} = \rho \left(\frac{\partial v_i}{\partial t} + \frac{\partial v_i}{\partial y_j} v_j \right) . \tag{6.45}$$

The last term in this equation cannot be neglected in many problems of hydrodynamics, which causes the corresponding equations to be non-linear and, consequently, difficult to solve. From this point of view, Eqs. (6.44) in Lagrangian coordinates, which we shall use in these lecture notes, are simpler. Consequently, as opposed to hydrodynamics, we shall be able to solve many simple problems in analytical forms.

6.6 Stress-Strain Relations

6.6.1 Rheological classification of substances

The relation between strain and stress depends on the type of substance and on many other factors. This is different for gases, liquids and solids, but there are great differences even between substances of the same phase. The study of these relations is the subject of *rheology*. The relations between strain and stress in real substances may be very complicated, so that various simplified models are introduced in rheology.

Hereafter we shall restrict ourselves only to linear elastic substances. A substance is said to be *elastic* if the strain completely vanishes on removal of load. A special type of elastic substance is a linear elastic substance, in which the strain and stress are directly proportional.

6.6.2 Generalised Hooke's law

The classical Hooke's law describes deformation only in the direction of the acting force. However, we have seen that strain and stress are complicated quantities of a tensor character. Therefore, we shall generalise the classical Hooke's law by assuming that a general linear relation exists between the stress and strain tensors:

$$\boxed{\tau_{ij} = C_{ijkl} e_{kl}} \,. \tag{6.46}$$

This relation is referred to as the generalised Hooke's law, and quantities C_{ijkl} are called elastic coefficients. Relation (6.46) describes well the behaviour of many substances, such as crystals and many other anisotropic materials. As a special case, it also describes the properties of many isotropic substances.

The total number of coefficients C_{ijkl} is $3^4 = 81$. However, as a consequence of the symmetry of the stress and strain tensors, the number of independent elastic coefficients reduces to $6 \times 6 = 36$. Moreover, the elastic coefficients are also symmetric with respect to interchanging of the first and second pairs of the subscripts, i.e. $C_{ijkl} = C_{klij}$, which follows from energetic considerations. In this way, the number of independent elastic coefficients reduces to 21. This number of elastic coefficients appears in the triclinic crystallographic structure. For crystals of a higher symmetry, the number of independent coefficients reduces further, so that the monoclinic structure is characterised by 13 independent elastic coefficients, rhombic by 9, and cubic by 3 independent elastic coefficients.

An isotropic medium, which has the same properties in all directions, is characterised by 2 elastic coefficients. The Lamé coefficients, λ and μ, are usually used in theoretical papers as these two coefficients. The generalised Hooke's law for an isotropic medium then takes the form

$$\boxed{\tau_{ij} = \lambda \vartheta \delta_{ij} + 2\mu e_{ij}}, \tag{6.47}$$

where $\vartheta = \operatorname{div} \mathbf{u} = e_{11} + e_{22} + e_{33}$ is the volume dilatation. Coefficient μ can be identified as the *shear modulus* (rigidity), but coefficient λ has no immediate physical interpretation.

The independent elastic coefficients are usually sought by analysing the changes of these coefficients under various rotations of the coordinate frame (Fung, 1965; Pšencík, 1994). We shall not perform these tedious calculations here, but we shall only briefly derive Hooke's law for an isotropic medium in the form of (6.47). We shall start by assuming that the deformation of an isotropic body consists of two independent parts, namely of a dilatation part and a shearing part. This idea was adopted in the middle of the 19th century on the basis of extensive experiments.

Introduce the Kronecker symbol (Kronecker delta)

$$\delta_{ij} = \begin{cases} 1 & \text{for } i = j \,, \\ 0 & \text{for } i \neq j \,. \end{cases} \tag{6.48}$$

The obvious identity

$$e_{ij} = \frac{1}{3}\vartheta\delta_{ij} + \left(e_{ij} - \frac{1}{3}\vartheta\delta_{ij} \right) \tag{6.49}$$

allows us to divide the deformation into the voluminal and shape parts. Denote the individual terms on the right-hand side of (6.49) by

$$f_{ij} = \frac{1}{3}\vartheta\delta_{ij} \,, \quad g_{ij} = e_{ij} - \frac{1}{3}\vartheta\delta_{ij} \,. \tag{6.50}$$

These expressions yield

$$f_{ii} = f_{11} + f_{22} + f_{33} = \frac{1}{3}\vartheta\delta_{ii} = \vartheta \,, \tag{6.51}$$

and

$$g_{ii} = e_{ii} - \frac{1}{3}\vartheta\delta_{ii} = e_{ii} - \vartheta = 0 \,. \tag{6.52}$$

Thus, the voluminal changes are described by tensor f_{ij}. Tensor g_{ij} describes the changes when the volume does not change, i.e. this tensor describes the shape changes. Tensor g_{ij} is called the deviatoric (or distortional) strain tensor.

The identity analogous to (6.49) can also be applied to the stress tensor,

$$\tau_{ij} = p_{ij} + q_{ij} \,, \tag{6.53}$$

where

$$p_{ij} = \frac{1}{3}\kappa\delta_{ij} , \quad q_{ij} = \tau_{ij} - \frac{1}{3}\kappa\delta_{ij} , \quad \kappa = \tau_{11} + \tau_{22} + \tau_{33} \,. \tag{6.54}$$

According to these analogies we may expect stresses p_{ij} to produce changes of volume, and stresses q_{ij} to produce changes of shape. Therefore, we shall assume that two coefficients exist, k_1 and k_2, where k_1 expresses the proportionality between the dilatation parts of the stress and strain tensors, and k_2 expresses the proportionality between the shearing parts:

$$p_{ij} = k_1 f_{ij} , \quad q_{ij} = k_2 g_{ij} \,. \tag{6.55}$$

By inserting these expressions into Eq. (6.53) and using the definitions of f_{ij} and g_{ij} in Eq. (6.50), we obtain

$$\tau_{ij} = k_1 f_{ij} + k_2 g_{ij} = \frac{1}{3}\left(k_1 - k_2\right)\vartheta\delta_{ij} + k_2 e_{ij} \ .$$

Introducing a new notation for the elastic coefficients,

$$\lambda = \frac{1}{3}\left(k_1 - k_2\right), \quad 2\mu = k_2 \ ,$$

we immediately arrive at formula (6.47).

6.7 Equations of Motion for an Isotropic Medium

The general equation of motion (6.44) cannot be used in practice unless the relation between stress and strain is specified, e.g., in the form of the generalised Hooke's law (6.46). Here we shall specify the equations of motion for an isotropic medium.

Insert Hooke's law for an isotropic medium, i.e. Eq. (6.47), into the equation of motion (6.44):

$$F_i + \frac{\partial}{\partial x_j}\left[\lambda\vartheta\delta_{ij} + \mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)\right] = \rho\frac{\partial^2 u_i}{\partial t^2} \ . \tag{6.56}$$

This equation is sometimes called the Navier-Green equation.

For a homogeneous isotropic medium, i.e. assuming elastic coefficients λ and μ to be constant, we get

$$F_i + \lambda\frac{\partial\vartheta}{\partial x_i} + \mu\left(\frac{\partial^2 u_i}{\partial x_j^2} + \frac{\partial^2 u_j}{\partial x_i \partial x_j}\right) = \rho\frac{\partial^2 u_i}{\partial t^2} \ . \tag{6.57}$$

Remember the following notations:

$$\vartheta = \operatorname{div}\mathbf{u} = \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} = \frac{\partial u_j}{\partial x_j} \ ,$$

(6.58)

$$\frac{\partial^2 u_i}{\partial x_j^2} = \frac{\partial^2 u_i}{\partial x_1^2} + \frac{\partial^2 u_i}{\partial x_2^2} + \frac{\partial^2 u_i}{\partial x_3^2} = \nabla^2 u_i \ , \quad \operatorname{grad}\vartheta = \left(\frac{\partial\vartheta}{\partial x_1}, \frac{\partial\vartheta}{\partial x_2}, \frac{\partial\vartheta}{\partial x_3}\right) \ .$$

where ∇^2 is Laplace's operator. Equation (6.57) can now be expressed in terms of displacements as

$$F_i + (\lambda + \mu)\frac{\partial^2 u_j}{\partial x_i \partial x_j} + \mu \frac{\partial^2 u_i}{\partial x_j^2} = \rho \frac{\partial^2 u_i}{\partial t^2} . \tag{6.59}$$

This equation represents the i-th component $(i = 1, 2, 3)$ of the following vector equation,

$$\boxed{\mathbf{F} + (\lambda + \mu)\,\mathrm{grad}\,\mathrm{div}\,\mathbf{u} + \mu \nabla^2 \mathbf{u} = \rho \frac{\partial^2 \mathbf{u}}{\partial t^2}} . \tag{6.60}$$

Equations (6.59) and (6.60) are the desired *equations of motion for a homogeneous isotropic medium.*

Further, if $\mathbf{F}$ is replaced by a body force, $\mathbf{g}$, which is related to the unit mass, i.e.

$$\mathbf{F} = \rho \mathbf{g} , \tag{6.61}$$

equation (6.60) takes the form

$$\rho \mathbf{g} + (\lambda + \mu)\,\mathrm{grad}\,\mathrm{div}\,\mathbf{u} + \mu \nabla^2 \mathbf{u} = \rho \frac{\partial^2 \mathbf{u}}{\partial t^2} . \tag{6.62}$$

This form of the equation of motion is frequently used in the theory of seismic waves.

By the Laplacian of a vector we understand the application of the Laplacian to the individual components, i.e.

$$\nabla^2 \mathbf{u} = \left(\nabla^2 u_1, \nabla^2 u_2, \nabla^2 u_3\right) . \tag{6.63}$$

However, it should be noted that such a simple definition of $\nabla^2 \mathbf{u}$ may be introduced only in Cartesian coordinates. For example, in spherical or cylindrical coordinates it has a more complicated form.

6.8 Wave Equations

Let us derive two special forms of the equation of motion for a homogeneous isotropic medium, known as the wave equations. Neglect the body force $\mathbf{F}$ in Eq. (6.60), which is acceptable in many problems of wave propagation. Apply the divergence operator to this equation and change the order of the derivatives in the second and third terms:

$$(\lambda + \mu)\,\mathrm{div}\;\mathrm{grad}\;\mathrm{div}\,\mathbf{u} + \mu \nabla^2\,\mathrm{div}\,\mathbf{u} = \rho \frac{\partial^2 (\mathrm{div}\,\mathbf{u})}{\partial t^2} .$$

Since the Laplacian $\nabla^2 = \mathrm{div}\,\mathrm{grad}$, we arrive at a *scalar wave equation* for volume dilatation $\vartheta = \mathrm{div}\,\mathbf{u}$,

$$\boxed{\nabla^2 \vartheta = \frac{1}{\alpha^2}\frac{\partial^2 \vartheta}{\partial t^2}}, \tag{6.64}$$

where the velocity of propagation of dilatation changes (longitudinal waves, compressional waves) is

$$\boxed{\alpha = \sqrt{\frac{\lambda + 2\mu}{\rho}}}. \tag{6.65}$$

Similarly, again put $\mathbf{F} = 0$, denote $\Omega = \text{curl}\ \mathbf{u}$ and apply the operator curl to Eq. (6.60). We shall arrive at a *vector wave equation*,

$$\boxed{\nabla^2 \Omega = \frac{1}{\beta^2}\frac{\partial^2 \Omega}{\partial t^2}}, \tag{6.66}$$

where the velocity of the propagation of distortion changes (transverse waves, shear waves) is

$$\boxed{\beta = \sqrt{\frac{\mu}{\rho}}}. \tag{6.67}$$

It follows from these equations that two types of elastic waves can propagate in a homogeneous isotropic medium, namely longitudinal and transverse waves. We shall use wave equations (6.64) and (6.66) many times in the following chapters.

6.9 Equations of Motion for Anisotropic Media

Let us go back to the generalised Hooke's law (6.46), and express the strain tensor e_{kl} in terms of the displacement vector, i.e. in the form (6.12):

$$\tau_{ij} = \frac{1}{2} C_{ijkl}\left(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k}\right). \tag{6.68}$$

The equation of motion (6.44) then takes the form

$$F_i + \frac{1}{2}\frac{\partial}{\partial x_j}\left[C_{ijkl}\left(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k}\right)\right] = \rho\frac{\partial^2 u_i}{\partial t^2}.$$

We shall show that both terms in the square brackets are identical. Rearrange the second term by using the symmetry of the elastic coefficients, $C_{ijkl} = C_{ijlk}$, and exchanging the dummy indices k and l:

$$C_{ijkl}\frac{\partial u_l}{\partial x_k} = C_{ijlk}\frac{\partial u_l}{\partial x_k} = C_{ijkl}\frac{\partial u_k}{\partial x_l} .$$

The equation of motion may then be simplified to read:

$$F_i + \frac{\partial}{\partial x_j}\left[C_{ijkl}\frac{\partial u_k}{\partial x_l}\right] = \rho\frac{\partial^2 u_i}{\partial t^2} . \qquad (6.69)$$

This is the equation of motion for an inhomogeneous anisotropic medium. This is the most general equation of motion which will be considered in these lecture notes.

If the elastic coefficients are constant, Eq. (6.69) becomes

$$\boxed{F_i + C_{ijkl}\frac{\partial^2 u_k}{\partial x_j \partial x_l} = \rho\frac{\partial^2 u_i}{\partial t^2}} . \qquad (6.70)$$

This is the *equation of motion for a homogeneous anisotropic medium.*

6.10 A Review of the Most Important Formulae

From the seismological point of view, let us summarise the most important formulae which have been derived in this chapter:

- the expression for the tensor of infinitesimal strain e_{ij} in terms of displacement vector **u**,

$$e_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) ; \qquad (6.12)$$

- the equation of motion of a continuum,

$$F_i + \frac{\partial \tau_{ij}}{\partial x_j} = \rho\frac{\partial^2 u_i}{\partial t^2} ; \qquad (6.44)$$

- the generalised Hooke's law,

$$\tau_{ij} = C_{ijkl}e_{kl} ; \qquad (6.46)$$

- Hooke's law for an isotropic medium,

$$\tau_{ij} = \lambda\vartheta\delta_{ij} + 2\mu e_{ij} ; \qquad (6.47)$$

- the equation of motion for a homogeneous isotropic medium,

$$\mathbf{F} + (\lambda + \mu)\,\mathrm{grad}\,\mathrm{div}\,\mathbf{u} + \mu\nabla^2\mathbf{u} = \rho\frac{\partial^2\mathbf{u}}{\partial t^2}\ ; \qquad (6.60)$$

- the wave equations for a homogeneous isotropic medium,

$$\nabla^2\vartheta = \frac{1}{\alpha^2}\frac{\partial^2\vartheta}{\partial t^2}\,, \quad \alpha = \sqrt{\frac{\lambda + 2\mu}{\rho}}\ , \qquad (6.64,\ 6.65)$$

$$\nabla^2\Omega = \frac{1}{\beta^2}\frac{\partial^2\Omega}{\partial t^2}\,, \quad \beta = \sqrt{\frac{\mu}{\rho}}\ ; \qquad (6.66,\ 6.67)$$

- the equation of motion for a homogeneous anisotropic medium,

$$F_i + C_{ijkl}\frac{\partial^2 u_k}{\partial x_j \partial x_l} = \rho\frac{\partial^2 u_i}{\partial t^2}\ . \qquad (6.70)$$

Chapter 7

Special Forms of the Elastodynamic Equation

In this chapter we shall deal with simple forms of the elastodynamic equation, especially with its special forms for a homogeneous and isotropic medium. Even in this special case, the corresponding equation of motion, given by (6.60), is still rather complicated. This vector equation represents a system of three coupled partial differential equations. These equations are more complicated than the equations which are traditionally solved in the courses of mathematical physics. The standard methods of solving partial differential equations, such as the separation of the individual variables, cannot be immediately applied to solve Eq. (6.60). We shall, therefore, attempt to express its solution as a sum of solutions of simpler equations. This can be accomplished, e.g., by introducing suitable potentials.

Potentials are auxiliary functions which are frequently introduced in mathematics and physics to facilitate the solution of complicated problems. For example, the well-know gravitational and electrostatic potentials enable us to describe the corresponding fields by one scalar function instead of three components of intensity. The velocity potential in hydrodynamics, or the Lagrangian and Hamiltonian in analytical mechanics are examples of analogous auxiliary functions. Electromagnetic potentials make it possible to reduce Maxwell's equations to simpler equations in many problems. Similarly, we shall introduce elastodynamic potentials in order to reduce the equation of motion (6.60) for a homogeneous isotropic medium to two simpler wave equations.

7.1 Separation of the Elastodynamic Equation in a Homogeneous Isotropic Medium

Consider the equation of motion for a homogeneous isotropic medium without body forces:

$$(\lambda + \mu)\,\text{grad div}\,\mathbf{u} + \mu \nabla^2 \mathbf{u} = \rho \frac{\partial^2 \mathbf{u}}{\partial t^2} \ . \tag{7.1}$$

We have already found two special forms of this equation, namely the wave equations (6.64) and (6.66) for the quantities $\vartheta = \text{div}\,\mathbf{u}$ and $\Omega = \text{curl}\,\mathbf{u}$, respectively. Here we shall derive the same equations in terms of potentials.

7.1.1 Wave equations in terms of potentials

Assume that the displacement vector is continuous together with its first derivatives, and the vector and derivatives vanish at infinity. According to Helmholtz' theorem, this vector can then be decomposed into irrotational and solenoidal parts (Arfken, 1970),

$$\mathbf{u} = \operatorname{grad}\varphi + \operatorname{curl}\vec{\psi} \ , \tag{7.2}$$

where φ is a scalar potential and $\vec{\psi}$ is a vector potential. By inserting this expression into the equation of motion (7.1) and interchanging the order of some operations, we obtain

$$\operatorname{grad}\left[(\lambda + 2\mu)\nabla^2\varphi - \rho\frac{\partial^2\varphi}{\partial t^2}\right] + \operatorname{curl}\left[\mu\nabla^2\vec{\psi} - \rho\frac{\partial^2\vec{\psi}}{\partial t^2}\right] = 0 \ .$$

This equation will be satisfied if the expressions in the square brackets are constants. In a special case, when these constants are zero, we arrive at the wave equations

$$\boxed{\nabla^2\varphi = \frac{1}{\alpha^2}\frac{\partial^2\varphi}{\partial t^2}} \ , \tag{7.3}$$

$$\boxed{\nabla^2\vec{\psi} = \frac{1}{\beta^2}\frac{\partial^2\vec{\psi}}{\partial t^2}} \ , \tag{7.4}$$

where

$$\alpha = \sqrt{\frac{\lambda + 2\mu}{\rho}} \ , \quad \beta = \sqrt{\frac{\mu}{\rho}} \tag{7.5}$$

are the longitudinal and transverse wave velocities, respectively. We have arrived at the wave equations for potentials φ and $\vec{\psi}$. The scalar wave equation (7.3) describes longitudinal waves (compressional waves, *P* waves), and the vector wave equation (7.4) describes transverse waves (shear waves, *S* waves).

Note that non-zero constants, which we have omitted in Eqs. (7.3) and (7.4), would describe static deformations of the medium. Since we shall not solve static problems, we shall consider the wave equations without these terms.

Elastodynamic potentials are frequently used, e.g., in studying Rayleigh waves, since these waves contain both longitudinal and transverse components of motion; for details we refer the reader to the lecture notes by Novotny (1999). However, some simpler problems, such as the reflection and transmission of *SH* waves or the propagation of Love waves, are usually studied directly in terms of displacements. Some problems may be formulated without substantial differences both in displacements and in potentials (e.g., the reflection and transmission for the *P-SV* problem, see below).

7.1.2 Expressions for the displacement and stress in terms of potentials

In order to formulate boundary conditions, and for other purposes, it is also necessary to express the displacements and stresses in terms of the elastodynamic potentials. Here, we shall restrict ourselves to Cartesian coordinates only.

Let u, v and w be the Cartesian components of displacement vector $\mathbf{u}$, i.e. $\mathbf{u} = (u, v, w)$. According to (7.2), these components can be expressed in terms of potentials as

$$u = \frac{\partial \varphi}{\partial x} + \frac{\partial \psi_z}{\partial y} - \frac{\partial \psi_y}{\partial z}, \quad v = \frac{\partial \varphi}{\partial y} + \frac{\partial \psi_x}{\partial z} - \frac{\partial \psi_z}{\partial x}, \quad w = \frac{\partial \varphi}{\partial z} + \frac{\partial \psi_y}{\partial x} - \frac{\partial \psi_x}{\partial y}, \tag{7.6}$$

In an isotropic medium, the components of the stress tensor, τ_{ij}, are given by Hooke's law in the form of (6.47), i.e.

$$\tau_{ij} = \lambda \vartheta \delta_{ij} + 2\mu e_{ij} \ . \tag{7.7}$$

These stress components can be expressed in terms of potentials by inserting (7.6) into (7.7). However, we shall not need these general expressions, as we shall solve various special problems only.

7.2 Wave Motion Dependent on One Cartesian Coordinate only

Consider the equation of motion in the form (7.1), i.e. the equation of motion for a homogenous isotropic medium without body forces. We have shown that special cases of this equation are as follows:

a) the wave equation for dilatational waves, Eq. (6.64);
b) the wave equation for distortional waves, Eq. (6.66);
c) the wave equations for potentials, Eqs. (7.3) and (7.4).

We have identified the dilatational waves with longitudinal waves and the distortional waves with transverse waves but, in fact, we have not proved it yet. Here we shall prove it for special cases of plane waves.

We shall modify and solve the equation of motion (7.1) for the special case when the displacement vector is dependent, in addition to time t, on one Cartesian coordinate only, say coordinate x. Thus, consider the displacement vector $\mathbf{u} = (u, v, w)$, where $u = u(x, t)$, $v = v(x, t)$ and $w = w(x, t)$. This describes the propagation of plane waves along the x-axis; component u represents a longitudinal wave, and components v and w a transverse wave (or two transverse waves). Now we get

$$\operatorname{div} \mathbf{u} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = \frac{\partial u}{\partial x}, \quad \nabla^2 \mathbf{u} = \frac{\partial^2 \mathbf{u}}{\partial x^2} + \frac{\partial^2 \mathbf{u}}{\partial y^2} + \frac{\partial^2 \mathbf{u}}{\partial z^2} = \frac{\partial^2 \mathbf{u}}{\partial x^2},$$

and the equation of motion (7.1) yields the following equations for the individual components:

$$(\lambda + 2\mu)\frac{\partial^2 u}{\partial x^2} = \rho\frac{\partial^2 u}{\partial t^2}, \quad \mu\frac{\partial^2 v}{\partial x^2} = \rho\frac{\partial^2 v}{\partial x^2}, \quad \mu\frac{\partial^2 w}{\partial x^2} = \rho\frac{\partial^2 w}{\partial x^2}. \tag{7.8}$$

These equations can be expressed as

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{\alpha^2}\frac{\partial^2 u}{\partial t^2}, \quad \frac{\partial^2 v}{\partial x^2} = \frac{1}{\beta^2}\frac{\partial^2 v}{\partial x^2}, \quad \frac{\partial^2 w}{\partial x^2} = \frac{1}{\beta^2}\frac{\partial^2 w}{\partial x^2}. \tag{7.9a,b,c}$$

These are wave equations where velocities α and β are given by Eqs. (6.65) and (6.67), respectively. Hence, we have proved that the longitudinal wave propagates at the same velocity, α, as dilatational waves. Consequently, we may identity dilatational waves with longitudinal waves. In a similar way we may identity transverse waves with distortional waves.

We have not found yet general solutions of the one-dimensional equations (7.9). Consider Eq. (7.9a) only, because the solution of the remaining equations will be analogous. Since α is a constant, introduce a new variable, $\xi = \alpha t$, instead of time t. The corresponding equation then takes the form

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial \xi^2}. \tag{7.10}$$

Thus, the only condition imposed on the displacement component u is that the second derivatives with respect to x and ξ are identical. Consequently, component u must be a function of a linear combination of x and ξ, e.g., a function of a variable

$$s = x + k\xi + q, \tag{7.11}$$

where k and q are constants. We may then write

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial s}, \quad \frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial s^2}, \quad \frac{\partial u}{\partial \xi} = \frac{\partial u}{\partial s}k, \quad \frac{\partial^2 u}{\partial \xi^2} = \frac{\partial^2 u}{\partial s^2}k^2,$$

By inserting these expressions into Eq. (7.10), we get $k^2 = 1$, so that $k = \pm 1$. Therefore, the general solution of Eq. (7.9a) has the form

$$u(x,t) = F_1(x - \alpha t + q_1) + F_2(x + \alpha t + q_2), \tag{7.12}$$

where F_1 and F_2 are arbitrary functions, but functions of those variables as indicated. Dividing the arguments by α, we may express the general solution also in the form

$$u(x,t) = f_1\left(t - \frac{x}{\alpha} + t_1\right) + f_2\left(t + \frac{x}{\alpha} + t_2\right). \tag{7.13}$$

The first terms in (7.12) and (7.13) represent a plane wave propagating in the positive direction of the x-axis, and the second terms represent a plane wave propagating in the opposite direction. Both waves propagate at velocity α (Fig. 7.1).

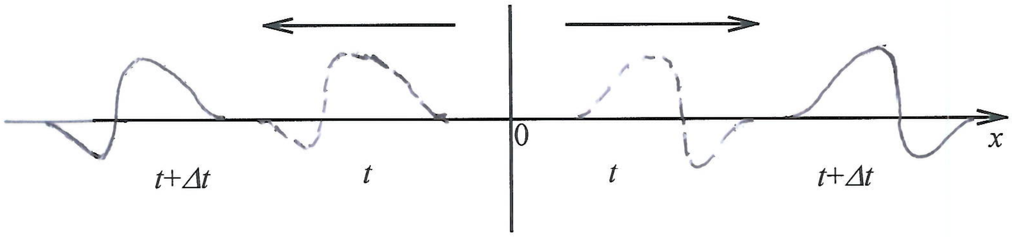


Fig. 7.1. Propagation of plane waves along the x-axis.

Very important special cases of the general solution (7.13) are *harmonic waves*:

$$u(x,t) = A_1 \cos\left[\omega_1\left(t - \frac{x}{\alpha} + t_1\right)\right] + A_2 \cos\left[\omega_2\left(t + \frac{x}{\alpha} + t_2\right)\right], \qquad (7.14)$$

where A_1 and A_2 are the amplitudes, and ω_1, ω_2 the angular frequencies. Instead of cosines we could also write sines. However, more frequently, we shall write the expressions for harmonic waves in exponential form:

$$u(x,t) = Ae^{i\omega_1(t - x/\alpha)} + Be^{i\omega_2(t + x/\alpha)}, \qquad (7.15)$$

where the complex constant $\exp(i\omega_1 t_1)$ and $\exp(i\omega_2 t_2)$ have been included into amplitudes A and B, respectively. It is one advantage of the exponential form (7.15). However, amplitudes A and B are generally complex. The second advantage, which is more important, consist in the fact that derivatives of (7.15) have a simpler form than derivatives of (7.14). However, real harmonic waves must be considered as real or imaginary parts of the corresponding complex wave.

7.3 Wave Motion Independent of One Cartesian Coordinate

In the previous section we studied the wave motions which are dependent on one Cartesian coordinate only, i.e. independent of two Cartesian coordinates. We formulated the problem in terms of the components of the displacement.

If the wave motion is a function of two or three Cartesian coordinates, it is usually more convenient to formulate the problem in terms of potentials, which are related to the displacement by Eqs. (7.2) and (7.6). We shall usually consider the Cartesian coordinate axes x and y to be horizontal, the z-axis to by vertical and positive downward (depth); see Fig. 7.2.

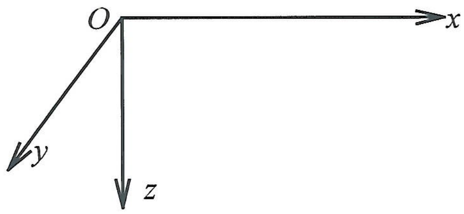


Fig. 7.2. The usual orientation of the Cartesian system.

In many problems, the wave propagation may be considered as independent of one Cartesian coordinate, say the y-coordinate. In this case, the wavefield is identical along the straight lines which are parallel to the y-axis. In other words, the derivatives of all quantities with respect to y are zero. The displacement vector $\mathbf{u}$ is now a function of the remaining two coordinates, x and z, and of time t:

$$\mathbf{u} = \mathbf{u}(x,z,t) \ . \tag{7.16}$$

Consequently, the potentials are functions of the same variables. Displacement components (7.6) now simplify to read

$$u = \frac{\partial \varphi}{\partial x} - \frac{\partial \psi_y}{\partial z} , \quad v = \frac{\partial \psi_x}{\partial z} - \frac{\partial \psi_z}{\partial x} , \quad w = \frac{\partial \varphi}{\partial z} + \frac{\partial \psi_y}{\partial x} \ . \tag{7.17}$$

It can be seen from these expressions that potentials φ, ψ_x, ψ_y and ψ_z, describing the wave propagation, are now separated into two groups. Namely, potentials φ and ψ_y appear only in the displacement components u and w, whereas ψ_x and ψ_z appear only in the displacement component v. Consequently, the wave motion described by components u and w, i.e. the motion in the (x, z)-plane, is now quite independent of the motion which is perpendicular to this plane. Hence, we can decompose this wave motion into the corresponding two parts and investigate them separately as two independent wave phenomena. Thus, let us write the vector potential $\vec{\psi}$ as the sum of two vectors,

$$\vec{\psi} = \vec{\psi}_{SV} + \vec{\psi}_{SH} \ , \tag{7.18}$$

where

$$\vec{\psi}_{SV} = (0, \psi_y, 0), \quad \vec{\psi}_{SH} = (\psi_x, 0, \psi_z) \ . \tag{7.19}$$

The displacement vector can then be expressed as

$$\mathbf{u} = \mathbf{u}_{P-SV} + \mathbf{u}_{SH} \ , \tag{7.20}$$

where

$$\mathbf{u}_{P-SV} = \operatorname{grad} \varphi + \operatorname{curl} \vec{\psi}_{SV} = (u, 0, w) \ , \quad \mathbf{u}_{SH} = \operatorname{curl} \vec{\psi}_{SH} = (0, v, 0) \ . \tag{7.21}$$

Vector $\mathbf{u}_{P-SV}$ represents the wave motion which is polarised in the (x, z)-plane. This motion consists of a longitudinal wave (P-wave, described by potential φ) and a transverse wave polarised in the (x, z)-plane (S-wave polarised in the vertical plane, thus denoted by SV). Vector $\mathbf{u}_{SH}$ represents a transverse wave which is polarised horizontally along the y-axis.

In solving P-SV problems, we usually use potentials φ and $\vec{\psi}_{SV} = (0, \psi_y, 0)$. Omitting the suffices SV and y, i.e. writing

$$\vec{\psi} = (0, \psi \;, 0) \;, \tag{7.22}$$

we arrive at

$$u = \frac{\partial \varphi}{\partial x} - \frac{\partial \psi}{\partial z} \;, \quad v = 0 \;, \quad w = \frac{\partial \varphi}{\partial z} + \frac{\partial \psi}{\partial x} \;. \tag{7.23}$$

Note that the vector potential is sometimes chosen with the opposite sign, i.e. $\vec{\psi} = (0, -\psi \;, 0)$, so that the corresponding terms in (7.23) have also the opposite signs.

In solving SH problems, we shall work directly with the displacement component v. The potentials will not be needed in these cases.

A special type of waves which are independent of the y-coordinate are plane waves, the rays of which are parallel to the (x, z)-plane, i.e. perpendicular to the y-axis. The wave surfaces of such waves are parallel to the y-axis (Fig. 7.3).

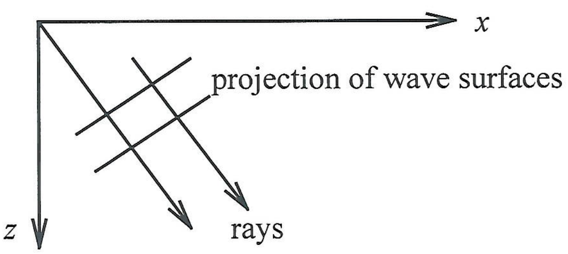


Fig. 7.3. A plane wave propagating in the (x, z)-plane.

A special type of these plane waves are the harmonic plane waves propagating in the (x, z)-plane. Let us derive the expressions, for example, for a longitudinal plane harmonic wave propagating in the (x, z)-plane. Choose the coordinate system in such a way that the ray at a point of observation, B, passes through the origin of the coordinate system (Fig. 7.4).

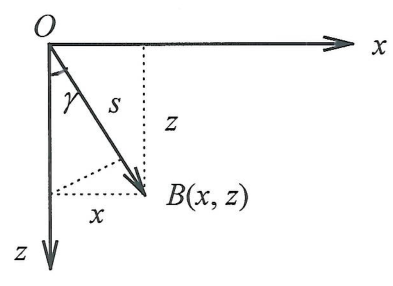


Fig. 7.4. Propagation of a wave in a positive direction of the x- and z-axes.

Denote by γ the angle between the ray and the z-axis; in cases of horizontal discontinuities it will be the angle between the ray and the normal, called the angle of incidence. The longitudinal potential (using analogies with the one-dimensional solution in Section 7.2) can then be expressed as

$$\varphi(x,z,t) = Ae^{i\omega(t-s/\alpha)} , \qquad (7.24)$$

where A is a constant amplitude, ω the angular frequency, α the longitudinal-wave velocity, and s the distance from the origin. It follows from Fig. 7.4 that

$$s = s_1 + s_2 = x\sin\gamma + z\cos\gamma , \qquad (7.25)$$

so that

$$\boxed{\varphi = Ae^{i\omega\left(t-\frac{x\sin\gamma+z\cos\gamma}{\alpha}\right)}} . \qquad (7.26)$$

Such expressions for plane harmonic waves are used in many applications, e.g., in the problems of reflection and refraction (see the corresponding chapter below). It can easily be verified that this potential satisfies the wave equation (7.3).

If the same angle γ is measured from the negative part of the z-axis, e.g., for a reflected wave, the coordinate z in (7.26) must be replaced by $(-z)$:

$$\varphi^* = A^*e^{i\omega\left(t-\frac{x\sin\gamma-z\cos\gamma}{\alpha}\right)} . \qquad (7.27)$$

This situation is shown in Fig. 7.5. Formula (7.27) can also be obtained immediately by replacing angle γ in (7.26) by $\gamma^* = 180°-\gamma$; then $\sin\gamma^* = \sin\gamma$ and $\cos\gamma^* = -\cos\gamma$.

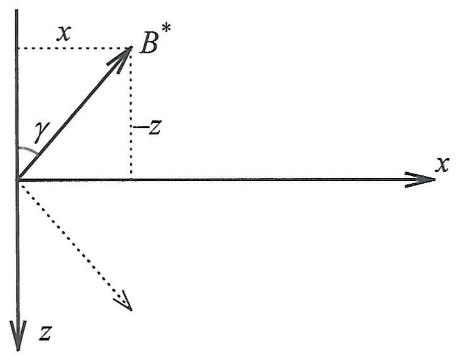


Fig. 7.5. Propagation in a direction having a positive component into the x-axis and a negative component into the z-axis.

In should be noted that the coefficients at the coordinates x and z in (7.26) are the directional cosines of the direction of propagation. Denote the angles between this direction and the coordinate axes by δ_x, δ_y, δ_z, respectively. In the case which is shown in Fig. 7.4, we have $\delta_x = 90° - \gamma$, $\delta_y = 0$ and $\delta_z = \gamma$. Consequently, the unit vector in the direction of propagation is

$$\mathbf{N} = \left(N_x, N_y, N_z\right) = \left(\cos\delta_x , \cos\delta_y , \cos\delta_z\right) = \left(\sin\gamma, 0, \cos\gamma\right) . \quad (7.28)$$

Using these directional cosines, formula (7.26) may be expressed as

$$\varphi = Ae^{i\omega\left(t - \frac{xN_x + zN_z}{\alpha}\right)} . \quad (7.29)$$

A further generalisation of this formula will be used in the next chapter.

Chapter 8

Plane Waves

In the previous chapter we derived the simplest solutions of the elastodynamic equation in special forms of plane waves. We considered the direction of propagation to have been parallel either with a Cartesian coordinate axis, or with a coordinate plane. As the next basic problem, which is traditionally considered in the textbooks of optics, acoustics and other wave phenomena, we should study the reflection and transmission of plane waves. However, before doing it, we shall study also other solutions of the elastodynamic equation for homogeneous media. These solutions are important not only in the theory of seismic waves, but some of them find direct applications in seismic prospecting. We shall deal with them in this chapter and in several chapter which follow. We shall begin with a more detailed treatise on plane waves.

Plane waves do not exist in real media but they are good approximations of waves generated by distant sources. One of the advantages of the plane-wave approximation consist in the fact that their source need not be considered, which simplifies the solution of many problems.

The study of plane waves is important for several reasons, in particular:

a) The behaviour of plane waves is simple, so that many properties of the wave propagation may be demonstrated on them.
b) More complicated waves, such as spherical or cylindrical waves, can also be expressed as superpositions of plane waves. The particular wave is expanded into plane waves, and each plane wave is propagated through the medium to the receiver. At the receiver, all the plane-wave contributions are again synthesised to form the solution of the original problem.
c) Many properties of plane waves can be used locally even in studies of more complex waves, such as waves in slightly inhomogeneous media. Studies of these waves by means of high-frequency asymptotic methods, such as the ray method, are based on many analogies with the plane-wave solutions.

Since a general time signal can be decomposed into individual harmonic components by means of the Fourier transform, we shall pay a special attention to the propagation of harmonic waves. The opposite approach, i.e. the investigation of plane waves in the time domain, has been described in detail by Psencik (1994). Each of these approaches has some advantages and disadvantages. The investigation in the frequency domain (harmonic waves) seems to be simpler from the pedagogical point of view, and represents the common approach also in other branches of physics, such as optics or electromagnetism in material media. This approach should be preferred in the situations when the wave velocity depends on frequency. In seismology, we encounter this situation in investigations of surface waves, in problems of the reflection and transmission at thin layers, and some others. This situation is quite common in optics, where the velocity of light in materials depends on frequency. On the other hand, the advantage of the description in the time domain (transient waves) consists in the fact that the results may immediately be compared with observed seismograms. In particular, this is convenient in studying over-critical reflections, when the reflected pulse changes its shape.

8.1 Plane Waves in a Homogeneous Isotropic Medium, Propagating in a General Direction

Instead of x, y, z, in the chapter we shall denote the Cartesian coordinates by x_1, x_2 and x_3, respectively. Consider a plane wave propagating in the direction which is given by a unit vector

$$\mathbf{N} = (N_1, N_2, N_3), \quad |\mathbf{N}| = \sqrt{N_1^2 + N_2^2 + N_3^2} = 1 . \tag{8.1}$$

This vector is parallel with the rays and positive in the direction of propagation; it is perpendicular to the wave surfaces.

Denote the velocity of propagation by c and the displacement vector by $\mathbf{u} = (u_1, u_2, u_3)$. If follows from the analogy with formula (7.29) that a plane wave must be a function of the variable

$$\tau = t - \frac{N_1 x_1 + N_2 x_2 + N_3 x_3}{c} = t - \frac{N_k x_k}{c} . \tag{8.2}$$

Therefore, let us seek the components of the displacement vector in the form

$$u_i(x_1, x_2, x_3, t) = U_i F\left(t - \frac{N_k x_k}{c}\right) , \tag{8.3}$$

where $\mathbf{N}$ is a given unit vector, c is an unknown velocity, U_i are amplitude factors ("amplitudes" of the individual components), and F is an arbitrary function describing the form of the wave, but identical for all there components.

Consider the elastodynamic equation without body forces in the form (7.1). Its i-th component may be expressed as

$$(\lambda + \mu)\frac{\partial^2 u_j}{\partial x_i \partial x_j} + \mu \frac{\partial^2 u_i}{\partial x_j^2} = \rho \frac{\partial^2 u_i}{\partial t^2} . \tag{8.4}$$

Insert the trial displacements (8.3) into this equation, and omit $\partial^2 F / \partial \tau^2$ in all terms (assuming this derivative to be non-zero). We get

$$(\lambda + \mu) U_j \frac{N_i N_j}{c^2} + \mu U_i \frac{N_1^2 + N_2^2 + N_3^2}{c^2} = \rho U_i .$$

Multiply this equation by c^2/ρ. Since $|\mathbf{N}| = 1$, we arrive at the following system of there equations $(i = 1, 2, 3)$:

$$\left[\frac{\lambda + \mu}{\rho} N_i N_j + \frac{\mu}{\rho}\delta_{ij}\right] U_j = c^2 \delta_{ij} U_j \ . \qquad (8.5)$$

However, we have four unknowns in this system, namely amplitudes U_1, U_2, U_3, and velocity c.

Denote the quantity in the square bracket of Eqs. (8.5) by

$$\Gamma_{ij} = \frac{\lambda + \mu}{\rho} N_i N_j + \frac{\mu}{\rho}\delta_{ij} \ . \qquad (8.6)$$

These quantities are the elements of a symmetric matrix, i.e. $\Gamma_{ij} = \Gamma_{ji}$, which is called the *Christoffel matrix for an isotropic medium.* Introduce the vector $\mathbf{U} = (U_1, U_2, U_3)$ and the scalar $\lambda = c^2$. Equations (8.5) can then be expressed in the following matrix form,

$$\Gamma\mathbf{U} = \lambda\mathbf{U} \ , \qquad (8.7)$$

where Γ is the Christoffel matrix. We have arrived at the so-called eigenvalue problem for matrix Γ.

The eigenvalue problems, studied in detail in linear algebra, consist in the following. For a given matrix, in our case for matrix Γ, we seek non-zero values of λ (eigenvalues) and the corresponding vectors $\mathbf{U}$ (eigenvectors) which satisfy the matrix equation of type (8.7).

8.1.1 Eigenvalues of the Christoffel matrix

Let us begin with the determination of the eigenvalues. Express Eqs. (8.5) as

$$\left(\Gamma_{ij} - c^2\delta_{ij}\right) U_j = 0 \ . \qquad (8.8)$$

This is a system of linear homogeneous equations (their right-hand sides are equal to zero). Such a system has a non-trivial solution if the corresponding determinant is equal to zero:

$$\begin{vmatrix} \Gamma_{11} - c^2 & \Gamma_{12} & \Gamma_{13} \\ \Gamma_{21} & \Gamma_{22} - c^2 & \Gamma_{23} \\ \Gamma_{31} & \Gamma_{32} & \Gamma_{33} - c^2 \end{vmatrix} = 0 \ . \qquad (8.9)$$

This determinant is called the Christoffel determinant, and Eq. (8.9) is called the Christoffel equation. This equation is a cubic equation for determining the unknown c^2.

The formulae just derived are also summarised below in Tab. 8.1 (only the summation index j, used here, is replaced by k in the table, in order to obtain a better correspondence to the case of anisotropic media).

Before calculating the Christoffel determinant, let us introduce the following abbreviations:

$$e = \frac{\lambda + \mu}{\rho}, \quad \beta^2 = \frac{\mu}{\rho} . \tag{8.10}$$

Elements (8.6) can then be written as

$$\Gamma_{ij} = eN_i N_j + \beta^2 \delta_{ij} . \tag{8.11}$$

The Christoffel equation (8.9) may now be expressed as

$$\begin{aligned}
&\left(eN_1^2 + \beta^2 - c^2\right)\left(eN_2^2 + \beta^2 - c^2\right)\left(eN_3^2 + \beta^2 - c^2\right) + 2e^3 N_1^2 N_2^2 N_3^2 - \\
&- \left(eN_2^2 + \beta^2 - c^2\right)e^2 N_1^2 N_3^2 - \left(eN_1^2 + \beta^2 - c^2\right)e^2 N_2^2 N_3^2 - \\
&- \left(eN_3^2 + \beta^2 - c^2\right)e^2 N_1^2 N_2^2 = 0 .
\end{aligned}$$

Since many terms cancel out, we get

$$\left(\beta^2 - c^2\right)^3 + \left(\beta^2 - c^2\right)^2 e\left(N_1^2 + N_2^2 + N_3^2\right) = 0 .$$

Considering again that **N** is a unit vector, and inserting from (8.10), we obtain the Christoffel equation in a very simple form:

$$\left(\frac{\lambda + 2\mu}{\rho} - c^2\right)\left(\frac{\mu}{\rho} - c^2\right)^2 = 0 . \tag{8.12}$$

It can be seen that all terms containing the components of vector **N** have dropped out in Eq. (8.12), so that velocity c is independent of the direction of propagation. We should expect such a result, since we consider an isotropic medium here.

We could even use this fact for a simple derivation of Eq. (8.12). Since the medium under consideration is isotropic, we can restrict ourselves to a special choice of vector **N**. For example, putting $\mathbf{N} = (1, 0, 0)$, elements (8.11) take the following simple form:

$$\Gamma_{11} = eN_1^2 = \frac{\lambda + 2\mu}{\rho}, \quad \Gamma_{22} = \Gamma_{33} = \frac{\mu}{\rho}, \quad \Gamma_{ij} = 0 \quad \text{for} \quad i \neq j .$$

The Christoffel equation (8.9) then immediately yields Eq. (8.12).

Equation (8.12) has one root

$$c_1^2 = \frac{\lambda + 2\mu}{\rho} \tag{8.13}$$

and one double root

$$c_{2,3}^2 = \frac{\mu}{\rho} \ . \tag{8.14}$$

These values are the desired eigenvalues of the Christoffel matrix (8.6). The corresponding velocities are

$$c_1 = \sqrt{\frac{\lambda + 2\mu}{\rho}} \ , \quad c_{2,3} = \sqrt{\frac{\mu}{\rho}} \ . \tag{8.15}$$

These velocities coincide with the velocities α and β for longitudinal and transverse velocities, discussed above. The negative velocities, i.e. $\left(-c_1\right)$, $\left(-c_{2,3}\right)$ which also satisfy Eqs. (8.12) to (8.14), represent only the propagation in the direction opposite to **N**.

Hence, we have proved that the velocities of plane waves in a homogeneous isotropic medium can attain two values given by formulae (8.15). This proof has been quite general, not restricted to special cases discussed in Chapter 7.

8.1.2 Eigenvectors of the Christoffel matrix

Now, let us proceed to the analysis of eigenvectors. It follows from the theory of linear operators that the eigenvectors are mutually perpendicular. Here we shall not give the algebraic proof of this general statement, but we shall explain it briefly on the basis of geometric considerations. An exact proof of the properties of eigenvectors will than be given for the special matrix (8.6).

Consider a linear transformation with a symmetric matrix. Such a transformation of a vector yields a new vector which has generally a different length and direction. Consider vectors **U** of a constant length, but of arbitrary orientation. Their endpoints form a spherical surface. By a linear transformation, such as $\Gamma\mathbf{U}$, the spherical surface transforms into an ellipsoidal surface. Along the axes of the ellipsoidal, the original vectors were only extended or contracted, but not rotated. Such vectors are the required eigenvectors, since they are transformed in to vectors of the same direction. It thus follows from the general theory that the polarisation of the wave propagating at velocity c_1 is perpendicular to the polarisation of the second wave propagating at velocity $c_{2,3}$.

Without using the general theory of linear transformations, we shall seek the eigenvectors directly for the special form of the Christoffel matrix (8.6).

Firstly, let us consider the solution in the form of a longitudinal wave, i.e. a wave polarised along the direction of propagation:

$$\mathbf{U} = k\mathbf{N} \ , \tag{8.16}$$

where k is a constant. Inserting $U_j = kN_j$ into Eq. (8.5) and omitting the constant k on both sides, we get

$$\frac{\lambda+\mu}{\rho} N_i N_j N_j + \frac{\mu}{\rho} N_i = c^2 N_i \ .$$

Since $N_j N_j = 1$, we arrive at the conclusion that the equation will be satisfied if c^2 satisfies Eq. (8.13). Thus, we have proved that the wave propagating at velocity $c_1 = \alpha$ is a longitudinal wave.

Secondly, assume the displacement to be perpendicular to the direction of propagation, i.e.

$$\mathbf{U} \bullet \mathbf{N} = U_j N_j = 0 \ . \tag{8.17}$$

In this case, Eq. (8.5) will be satisfied if c^2 satisfies Eq. (8.14), i.e. $c = \beta$. We have proved that the wave propagating at velocity $c_{2,3} = \beta$ is a transverse wave. The polarisation of the particle motion in this wave is not determined by Eq. (8.17), so that it may be arbitrary (linear polarisation, elliptic polarisation).

Table 8.1. A comparison of formulae for plane waves propagating in homogeneous isotropic and homogeneous anisotropic media. The formulae which have the same form for both cases are written in the middle the corresponding lines.

ISOTROPIC	ANISOTROPIC
Equation of motion (i-th component):	
$(\lambda+\mu)\dfrac{\partial^2 u_k}{\partial x_i \partial x_k} + \mu \dfrac{\partial^2 u_i}{\partial x_k^2} = \rho \dfrac{\partial^2 u_i}{\partial t^2}$,	$C_{ijkl} \dfrac{\partial^2 u_k}{\partial x_j \partial x_l} = \rho \dfrac{\partial^2 u_i}{\partial t^2}$
Form of the solution: $u_k = U_k F\left(t - \dfrac{N_m x_m}{c}\right)$	
Equation of motion for plane waves (i-th component): $\left(\Gamma_{ik} - c^2 \delta_{ik}\right) U_k = 0$	
Christoffel matrix:	
$\Gamma_{ik} = \dfrac{\lambda+\mu}{\rho} N_i N_k + \dfrac{\mu}{\rho} \delta_{ik}$,	$\Gamma_{ik} = \dfrac{1}{\rho} C_{ijkl} N_j N_l$
Christoffel equation: $\begin{vmatrix} \Gamma_{11} - c^2 & \Gamma_{12} & \Gamma_{13} \\ \Gamma_{21} & \Gamma_{22} - c^2 & \Gamma_{23} \\ \Gamma_{31} & \Gamma_{32} & \Gamma_{33} - c^2 \end{vmatrix} = 0$	

8.2 Plane Waves in a Homogeneous Anisotropic Medium

Now we shall solve the analogous problem as in the previous section, but for an anisotropic medium. The main steps of the derivation are briefly described in Tab. 8.1. Here we shall describe the derivation in detail.

Consider the equation of motion for a homogeneous anisotropic medium, i.e. Eq. (6.70), but neglect the body forces. The equation then becomes

$$C_{ijkl}\frac{\partial^2 u_k}{\partial x_j \partial x_l} = \rho\frac{\partial^2 u_i}{\partial t^2} , \tag{8.18}$$

where C_{ijkl} are the elastic coefficients in the generalised Hooke's law.

We shall seek a solution again in the form of a plane wave, see (8.3):

$$u_k = U_k F\left(t - \frac{N_m x_m}{c}\right) . \tag{8.19}$$

By inserting this form into the equation of motion (8.18) and multiplying by c^2/ρ, we get

$$\left(\Gamma_{ik} - c^2\delta_{ik}\right)U_k = 0 , \tag{8.20}$$

where the elements of the Christoffel matrix now read

$$\Gamma_{ik} = \rho^{-1}C_{ijkl}N_j N_l . \tag{8.21}$$

The Christoffel matrix is again symmetric, $\Gamma_{ik} = \Gamma_{ki}$. The Christoffel equation, i.e. the condition of solvability of Eqs. (8.20), is formally identical to Eq. (8.9):

$$\begin{vmatrix} \Gamma_{11} - c^2 & \Gamma_{12} & \Gamma_{13} \\ \Gamma_{21} & \Gamma_{22} - c^2 & \Gamma_{23} \\ \Gamma_{31} & \Gamma_{32} & \Gamma_{33} - c^2 \end{vmatrix} = 0 . \tag{8.22}$$

From the Christoffel equation (8.22) we can calculate the velocity of propagation, c, if the elastic parameters C_{ijkl}, density ρ and the direction of propagation $\mathbf{N}$ (described by the directional cosines N_m) are given. The Christoffel equation is a cubic equation for c^2,

$$c^6 - Pc^4 + Qc^2 - R = 0 , \tag{8.23}$$

where P, Q and R are the so-called invariants of matrix Γ. The formulae for P, Q and R could be obtained by rewriting Eq. (8.22) into the form (8.23).

The cubic equation (8.23) has generally three roots for c^2. It can be shown that all these roots are real (Psencik, 1994). Therefore, three different plane waves, with generally *three different velocities*, can propagate in an anisotropic homogeneous medium. Each root yields two velocities, $c = \pm\sqrt{c^2}$, which represent two plane waves propagating at velocity $|c|$ in two opposite directions. Since P, Q, R are functions of N_m, the velocities of the wave propagation are *dependent on the direction* of propagation.

Since the Christoffel matrix Γ is real and symmetric, its eigenvectors are mutually perpendicular. This means that the three different plane waves differ not only in their velocities but also in their polarisation, i.e. in the orientation of the eigenvectors, which specify the directions of the particle motion. In can be shown that the eigenvectors in an anisotropic medium (the polarisation of the waves) are generally neither parallel nor perpendicular to the direction of propagation $\mathbf{N}$.

Usually, the wave with the highest velocity, whose polarisation is closest to the direction of the phase normal $\mathbf{N}$, is called the *quasi-longitudinal*, *quasi-compressional* or *qP wave*. The other two waves are called the *quasi-transverse*, *quasi-shear* or *qS waves* (*qS1* and *qS2*).

Further details on the propagation of plane waves in homogeneous anisotropic media can be found in the lecture notes by Psencik (1994).

The main differences between the plane waves in isotropic and anisotropic media can be summarised as follows:

1) Two plane waves with different velocities can propagate in a homogeneous isotropic unlimited medium, namely the longitudinal wave and transverse wave. In an anisotropic medium, generally three independent plane waves with different velocities can propagate, one quasi-longitudinal and two quasi-transverse waves.
2) In an anisotropic medium, the velocities of the individual plane waves depend on the direction of propagation. In an isotropic medium, these velocities are independent of the direction of propagation.
3) In isotropic media, the polarisation direction (displacement vector, particle motion) is either parallel to the direction of propagation (in longitudinal waves), or perpendicular to this direction (in transverse waves). The polarisation of longitudinal waves is linear, that of transverse waves is generally elliptical. The polarisation of waves in anisotropic media is, in general, linear and has a general direction. The polarisation neither coincides with the direction of rays in the case of quasi-longitudinal waves, nor is perpendicular to the rays in the case of quasi-transverse waves.
4) The velocity of the propagation of the phase front (phase velocity, denoted by c above) and the velocity of the propagation of energy (group velocity) are equal in isotropic media. In anisotropic media, the phase and group velocities are generally different; the corresponding derivation can be found in Psencik (1994).

Chapter 9

Spherical and Cylindrical Waves

In this chapter we shall solve the wave equations in spherical and cylindrical coordinates. As the first problem, we must transform the Laplace operator into these coordinates. It will be the subject of Section 9.1. The reader who knows the corresponding formulae, in particular formula (9.28), may proceed to Section 9.2.

9.1 Basic Vector Operators in Orthogonal Curvilinear Coordinates

Many physical problems can better be solved in curvilinear coordinates than in Cartesian coordinates. We shall restrict ourselves to orthogonal curvilinear coordinates (their coordinate lines at any point are mutually perpendicular). The most important orthogonal curvilinear coordinates are spherical coordinates and cylindrical coordinates.

Denote Cartesian coordinates by x, y, z. Introduce spherical coordinates r, ϑ, λ, where r is the radial coordinate, ϑ the angular "polar" distance, and λ the longitude ("geographical" longitude); see Fig. 9.1a. The relation between these coordinates is given by the well-know formulae:

$$x = r\sin\vartheta\cos\lambda\ , \quad y = \sin\vartheta\sin\lambda\ , \quad z = r\cos\vartheta\ . \tag{9.1}$$

The coordinate lines of spherical coordinates are rays from the coordinate origin (along which the radial distance r varies only), "meridians" (variation of ϑ), and "parallels" (variation of λ).

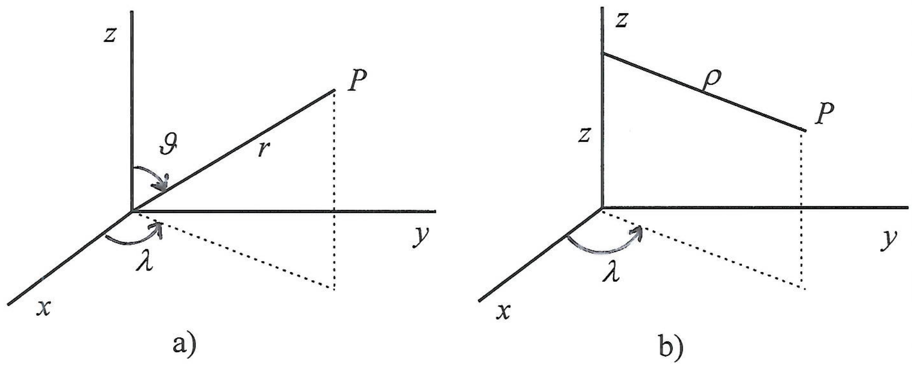


Fig. 9.1. Spherical coordinates (a), and cylindrical coordinates (b).

The relation between the Cartesian coordinates and cylindrical coordinates ρ, λ, z is given by the following formulae (Fig. 9.1b):

$$x = \rho\cos\lambda\ , \quad y = \rho\sin\lambda\ , \quad z = z\ ; \tag{9.2}$$

the identity $z = z$ means that we identify the Cartesian coordinate z on the left-hand side with the cylindrical coodinate z on the right-hand side.

9.1.1 Lamé's coefficients

We shall compare various quantities at a given point, P, and at neighbouring points. Therefore, let us consider an infinitesimal change of position, described by vector $\mathrm{d}\mathbf{s} = (\mathrm{d}x, \mathrm{d}y, \mathrm{d}z)$. Let us express this vector in orthogonal curvilinear coordinates. Denote these coordinates generally by q_1, q_2 and q_3.

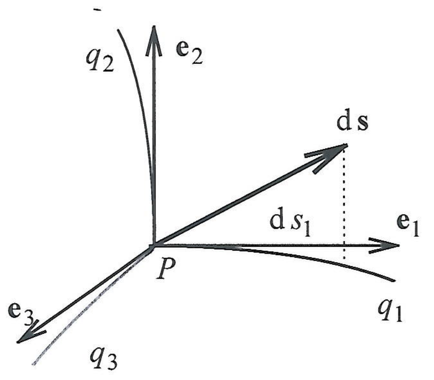


Fig. 9.2. Decomposition of an infinitesimal vector, **ds**, in an orthogonal curvilinear system (the projection onto the first coordinate line is shown only).

Introduce the unit vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ which are tangent to the curvilinear coordinate lines at the considered point (Fig. 9.2). Express vector **ds** in this local Cartesian system as

$$\mathrm{d}\mathbf{s} = \mathbf{e}_1\,\mathrm{d}s_1 + \mathbf{e}_2\,\mathrm{d}s_2 + \mathbf{e}_3\,\mathrm{d}s_3\,, \tag{9.3}$$

where coordinates $\mathrm{d}s_1$, $\mathrm{d}s_2$ and $\mathrm{d}s_3$ have the dimension of length. They represent the infinitesimal arcs (lengths) along the individual curvilinear coordinate lines. Denote the corresponding increments in the curvilinear coordinates by $\mathrm{d}q_1$, $\mathrm{d}q_2$ and $\mathrm{d}q_3$, respectively. It is evident that there must be a proportionality between these quantities:

$$\mathrm{d}s_1 = H_1\,\mathrm{d}q_1\,, \quad \mathrm{d}s_2 = H_2\,\mathrm{d}q_2\,, \quad \mathrm{d}s_3 = H_3\,\mathrm{d}q_3\,. \tag{9.4}$$

The coefficients H_1 to H_3 are referred to as *Lamé's coefficients* (they should not be confused with the elastic Lamé coefficients, λ and μ).

Omitting subscripts in (9.4), we may introduce the following definition of Lamé's coefficients. Lamé's coefficient H, corresponding to curvilinear coordinate q, is defined by the relation

$$\boxed{\mathrm{d}s = H\,\mathrm{d}q}\,, \tag{9.5}$$

where ds is the length of an element along the coordinate line q, and dq is the corresponding increment in this coordinate.

Hence, if coordinate q is a length, Lamé's coefficient is then equal to unity, $H=1$. This is the case of the spherical coordinate r, and of cylindrical coordinates ρ and z. However, if coordinate q is an angle, Lamé's coefficient is the radius of curvature (because $\mathrm{d}s = R\,\mathrm{d}\varphi$, R being the radius of curvature, and $\mathrm{d}\varphi$ the element of the angle).

Let as specify relation (9.5) for spherical coordinates. Put $q_1 = r$, $q_2 = \vartheta$, $q_3 = \lambda$, and write the arcs along the coordinate lines as

$$\mathrm{d}s_r = H_r\,\mathrm{d}r\,, \quad \mathrm{d}s_\vartheta = H_\vartheta\,\mathrm{d}\vartheta\,, \quad \mathrm{d}s_\lambda = H_\lambda\,\mathrm{d}\lambda\,. \tag{9.6}$$

It follows from geometrical considerations that

$\mathrm{d}s_r = \mathrm{d}r$ (an arc along a radial coordinate line);

$\mathrm{d}s_\vartheta = r\,\mathrm{d}\vartheta$ (an arc along a meridian, r being its radius);

$\mathrm{d}s_\lambda = r\sin\vartheta\,\mathrm{d}\lambda$ (an arc along a parallel, $r\sin\vartheta$ being its radius).

Comparing these relations with (9.6), we arrive at the following expressions for Lamé's coefficients for spherical coordinates:

$$\boxed{H_r = 1\,, \quad H_\vartheta = r\,, \quad H_\lambda = r\sin\vartheta}\,. \tag{9.7}$$

Using these Lamé's coefficients, we can express easily, e.g., a volume element in spherical coordinates, the edges of which are $\mathrm{d}s_r$, $\mathrm{d}s_\vartheta$, $\mathrm{d}s_\lambda$:

$$\mathrm{d}V = \mathrm{d}s_r\,\mathrm{d}s_\vartheta\,\mathrm{d}s_\lambda = H_r H_\vartheta H_\lambda\,\mathrm{d}r\,\mathrm{d}\vartheta\,\mathrm{d}\lambda = r^2\sin\vartheta\,\mathrm{d}r\,\mathrm{d}\vartheta\,\mathrm{d}\lambda\,. \tag{9.8}$$

The term $r^2\sin\vartheta$ is the well-known determinant of the transition from Cartesian to spherical coordinates.

Analogously, for cylindrical coordinates we get

$$\boxed{H_\rho = 1\,, \quad H_\lambda = \rho\,, \quad H_z = 1}\,. \tag{9.9}$$

For a volume element we obtain

$$\mathrm{d}V = H_\rho H_\lambda H_z\,\mathrm{d}\rho\,\mathrm{d}\lambda\,\mathrm{d}z = \rho\,\mathrm{d}\rho\,\mathrm{d}\lambda\,\mathrm{d}z\,. \tag{9.10}$$

Of course, we could obtain the expressions for Lamé's coefficients also quite formally from the transformation relations (9.1) and (9.2). A length element in Cartesian coordinates can be expressed as

$$\mathrm{d}s = \sqrt{(\mathrm{d}x)^2 + (\mathrm{d}y)^2 + (\mathrm{d}z)^2}\,. \tag{9.11}$$

Differentiating the transformation relations (9.1), we get

$$\begin{aligned}
\mathrm{d}x &= \mathrm{d}r \sin\vartheta \cos\lambda + r\cos\vartheta\, \mathrm{d}\vartheta \cos\lambda - r\sin\vartheta \sin\lambda\, \mathrm{d}\lambda \ , \\
\mathrm{d}y &= \mathrm{d}r \sin\vartheta \sin\lambda + r\cos\vartheta\, \mathrm{d}\vartheta \sin\lambda + r\sin\vartheta \cos\lambda\, \mathrm{d}\lambda \ , \\
\mathrm{d}z &= \mathrm{d}r \cos\vartheta - r\sin\vartheta\, \mathrm{d}\vartheta \ .
\end{aligned} \tag{9.12}$$

Insert these expressions into (9.11). Since many terms drop out, we obtain

$$\mathrm{d}s = \sqrt{(\mathrm{d}r)^2 + (r\,\mathrm{d}\vartheta)^2 + (r\sin\vartheta\,\mathrm{d}\lambda)^2} \ . \tag{9.13}$$

Writing this formula as

$$\mathrm{d}s = \sqrt{\left(H_r\,\mathrm{d}r\right)^2 + \left(H_\vartheta\,\mathrm{d}\vartheta\right)^2 + \left(H_\lambda\,\mathrm{d}\lambda\right)^2} \ , \tag{9.14}$$

we arrive at Lamé's coefficients (9.7).

The formulae for cylindrical coordinates are simpler:

$$\mathrm{d}x = \mathrm{d}\rho\cos\lambda - \rho\sin\lambda\,\mathrm{d}\lambda\ , \quad \mathrm{d}y = \mathrm{d}\rho\sin\lambda + \rho\cos\lambda\,\mathrm{d}\lambda\ , \quad \mathrm{d}z = \mathrm{d}z\ . \tag{9.15}$$

Formula (9.11) then takes the form

$$\mathrm{d}s = \sqrt{\left(\mathrm{d}\rho\right)^2 + \left(\rho\,\mathrm{d}\lambda\right)^2 + \left(\mathrm{d}z\right)^2} \ , \tag{9.16}$$

which yields Lamé's coefficients (9.9).

9.1.2 Gradient in orthogonal curvilinear coordinates

Consider a scalar function of coordinates, U, i.e. $U = U(x, y, z)$ or $U = U(q_1, q_2, q_3)$. The change of function U, corresponding to the change of position given by vector $\mathrm{d}\mathbf{s} = (\mathrm{d}x, \mathrm{d}y, \mathrm{d}z)$, can be expressed as

$$\mathrm{d}U = \frac{\partial U}{\partial x}\mathrm{d}x + \frac{\partial U}{\partial y}\mathrm{d}y + \frac{\partial U}{\partial z}\mathrm{d}z \ . \tag{9.17}$$

note that this formula is valid only if the derivatives $\partial U/\partial x$, $\partial U/\partial y$ and $\partial U/\partial z$ are continuous functions of coordinates which we shall assume here. The expression on the right-hand side of (9.17) is then referred to as the total differential. In the curvilinear coordinates we may write analogously

$$\mathrm{d}U = \frac{\partial U}{\partial q_1}\mathrm{d}q_1 + \frac{\partial U}{\partial q_2}\mathrm{d}q_2 + \frac{\partial U}{\partial q_3}\mathrm{d}q_3 \ . \tag{9.18}$$

The right-hand side of (9.17) represent the scalar product of two vectors,

$$\mathrm{d}U = \left(\frac{\partial U}{\partial x}, \frac{\partial U}{\partial y}, \frac{\partial U}{\partial z}\right)\left(\mathrm{d}x, \mathrm{d}y, \mathrm{d}z\right) = \left(\mathrm{grad}\,U\right).\mathrm{d}\mathbf{s} \ , \tag{9.19}$$

where the vector

$$\mathrm{grad}\,U = \left(\frac{\partial U}{\partial x}, \frac{\partial U}{\partial y}, \frac{\partial U}{\partial z}\right) \tag{9.20}$$

is called the gradient of scalar function U. Let us seek the analogous expressions for this vector in orthogonal curvilinear coordinates.

Multiply and divide the individual terms in formula (9.18) by the corresponding Lamé's coefficient. This yields the equivalent formula

$$\mathrm{d}U = \frac{1}{H_1}\frac{\partial U}{\partial q_1} H_1\,\mathrm{d}q_1 + \frac{1}{H_2}\frac{\partial U}{\partial q_2} H_2\,\mathrm{d}q_2 + \frac{1}{H_3}\frac{\partial U}{\partial q_3} H_3\,\mathrm{d}q_3 \ ,$$

which represents the scalar product of two vectors,

$$\mathrm{d}U = \left(\frac{1}{H_1}\frac{\partial U}{\partial q_1}, \frac{1}{H_2}\frac{\partial U}{\partial q_2}, \frac{1}{H_3}\frac{\partial U}{\partial q_3}\right)\left(H_1\,\mathrm{d}q_1, H_2\,\mathrm{d}q_2, H_3\,\mathrm{d}q_3\right) . \tag{9.21}$$

Thhe second vector on the right-hand side is vector d**s**, in view of (9.4) and (9.3). This vector is expressed in the coordinate system which is defined by vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ (Fig. 9.2). It then follows from (9.19) that the first vector in (9.21) must be the desired gradient. Hence, we have arrived at the conclusion that the gradient may be expressed in orthogonal curvilinear coordinates as

$$\boxed{\mathrm{grad}\,U = \frac{\mathbf{e}_1}{H_1}\frac{\partial U}{\partial q_1} + \frac{\mathbf{e}_2}{H_2}\frac{\partial U}{\partial q_2} + \frac{\mathbf{e}_3}{H_3}\frac{\partial U}{\partial q_3}} \ . \tag{9.22}$$

Omitting the subscripts, the component of the gradient into coordinate line q is thus

$$\left(\mathrm{grad}\,U\right)_q = \frac{1}{H}\frac{\partial U}{\partial q} \ . \tag{9.23}$$

9.1.3 Divergence in orthogonal curvilinear coordinates

We shall derive the formula for the divergence of a vector in orthogonal curvilinear coordinates by using the integral definition of the divergence, because this definition is not related to a concrete coordinate system.

Consider a point P and a small volume V which surrounds this point. Denote again the surface of this volume by S. Let $\mathbf{A}$ be a vector which is continuous together with its derivatives within V. Apply Gauss' theorem (divergence theorem) to this vector in volume V:

$$\iiint_V \operatorname{div}\mathbf{A}\,\mathrm{d}V = \iint_S \mathbf{A}.\vec{\nu}\,\mathrm{d}S\ , \tag{9.24}$$

where $\vec{\nu}$ is the unit outward normal to the surface element dS, see Eq. (6.30). According to the mean-value theorem, the volume integral in (9.24) may be expressed as

$$(\operatorname{div}\mathbf{A})_Q .V\ ,$$

where Q is some point in volume V. Pass to the limit for $V \to 0$, but keep point P still inside of V. Point Q then approaches P, and we arrive at the following integral definition of the divergence at point P:

$$\operatorname{div}\mathbf{A} = \lim_{V\to 0}\frac{1}{V}\iint_S \mathbf{A}.\vec{\nu}\,\mathrm{d}S\ . \tag{9.25}$$

Consider point P to be a vertex of an infinitesimal rectangular body ("parallelepiped"), the edges of which coincide with the curvilinear coordinate lines (Fig. 9.3). Let q_1, q_2, q_3 be the coordinates of point P, and Δq_1, Δq_2, Δq_3 the increments of coordinates along the edges of the body. We shall apply formula (9.25) to this body.

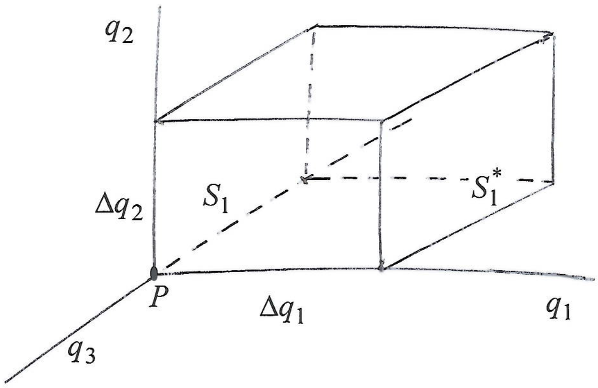


Fig 9.3. Volume element to which the divergence theorem is applied.

First, consider the contributions from the faces which are perpendicular to coordinate lines q_1. Face S_1 is characterised by coordinate q_1, and face S_1^* by coordinate $q_1 + \Delta q_1$. The outward normals to face S_1^* coincide with the positive directions of coordinate lines q_1, but the outward normals to face S_1 are of the opposite directions. Consequently, the contribution of these surfaces to the integral in (9.25) is

$$I_1 = \iint_{S_1^*} \mathbf{A}\vec{\nu}\,\mathrm{d}S + \iint_{S_1} \mathbf{A}\vec{\nu}\,\mathrm{d}S = \iint_{S_1^*} A_1\,\mathrm{d}S - \iint_{S_1} A_1\,\mathrm{d}S\ .$$

Consider a rectangular surface element dS which is perpendicular to coordinate line q_1 and its edges are $\mathrm{d}s_2$ and $\mathrm{d}s_3$. Using formulae (9.4), we get

$$\mathrm{d}S = \mathrm{d}s_2\,\mathrm{d}s_3 = H_2 H_3\,\mathrm{d}q_2\,\mathrm{d}q_3 \ .$$

Contribution I_1 then becomes

$$I_1 = \int_{q_2}^{q_2+\Delta q_2}\int_{q_3}^{q_3+\Delta q_3}\left[\left(A_1 H_2 H_3\right)_{q_1+\Delta q_1} - \left(A_1 H_2 H_3\right)_{q_1}\right]\mathrm{d}q_2\,\mathrm{d}q_3 \ ,$$

where subscripts $q_1 + \Delta q_1$ and q_1 indicate that the first term $A_1 H_2 H_3$ is taken at points of coordinates $\left(q_1 + \Delta q_1, q_2, q_3\right)$, whereas the second term at points of coordinates $\left(q_1, q_2, q_3\right)$. Using the mean-value theorem of differential calculus, the integrand in the square brackets may be expressed as

$$\frac{\partial\left(A_1 H_2 H_3\right)}{\partial q_1}\Delta q_1$$

where the derivative is taken at a point of coordinates $\left(q_1 + \vartheta\,\Delta q_1, q_2, q_3\right)$; the value of ϑ lies in the interval $0 \le \vartheta \le 1$, but varies with q_2 and q_3.

According to the mean-value theorem of integral calculus, a surface integral of a continuous function is equal to the function value at some internal point multiplied by the surface. Thus

$$I_1 = \left[\frac{\partial\left(A_1 H_2 H_3\right)}{\partial q_1}\right]_{P_1^*}\Delta q_1 \Delta q_2 \Delta q_3 \ ,$$

where P_1^* is some point in the body. Similar expressions may be obtained for the integrals over the remaining surfaces.

The volume element, in view of (9.4), is $\mathrm{d}V = \mathrm{d}s_1\,\mathrm{d}s_2\,\mathrm{d}s_3 =$ $= H_1 H_2 H_3\,\mathrm{d}q_1\,\mathrm{d}q_2\,\mathrm{d}q_3$. Using again the mean-value theorem, the volume of the whole body is $V = \left(H_1 H_2 H_3\right)_{P_V}\Delta q_1 \Delta q_2 \Delta q_3$, where P_V is some point of the body. Insert these expressions into formula (9.25), reduce quantities Δq_1, Δq_2, Δq_3, and shrink the body towards point P, still keeping this point as a vertex of the body. This yields the divergence at point P in the form

$$\boxed{\operatorname{div}\mathbf{A} = \frac{1}{H_1 H_2 H_3}\left[\frac{\partial}{\partial q_1}\left(H_2 H_3 A_1\right) + \frac{\partial}{\partial q_2}\left(H_1 H_3 A_2\right) + \frac{\partial}{\partial q_3}\left(H_1 H_2 A_3\right)\right]} \ . \qquad (9.26)$$

This is the general formula for the divergence in orthogonal curvilinear coordinates.

9.1.4 The Laplacian in orthogonal curvilinear coordinates

The Laplacian of a scalar U is defined by

$$\nabla^2 U = \operatorname{div} \operatorname{grad} U \ . \tag{9.27}$$

Putting $\mathbf{A} = \operatorname{grad} U$, and using the formulae (9.22) and (9.26) for the gradient and divergence, we arrive at

$$\boxed{\nabla^2 U = \frac{1}{H_1 H_2 H_3}\left[\frac{\partial}{\partial q_1}\left(\frac{H_2 H_3}{H_1}\frac{\partial U}{\partial q_1}\right) + \frac{\partial}{\partial q_2}\left(\frac{H_1 H_3}{H_2}\frac{\partial U}{\partial q_2}\right) + \frac{\partial}{\partial q_3}\left(\frac{H_1 H_2}{H_3}\frac{\partial U}{\partial q_3}\right)\right]} \tag{9.28}$$

This is a very important formula for the Laplacian in orthogonal curvilinear coordinates.

Let us specify the latter formula for spherical coordinates. Inserting Lamé's coefficients (9.7) into this formula, we get

$$\nabla^2 U = \frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2 \frac{\partial U}{\partial r}\right) + \frac{1}{\sin\vartheta}\frac{\partial}{\partial \vartheta}\left(\sin\vartheta \frac{\partial U}{\partial \vartheta}\right) + \frac{1}{\sin^2\vartheta}\frac{\partial^2 U}{\partial \lambda^2}\right] . \tag{9.29}$$

We shall also need a special form of this formula for the case when scalar U is a function of coordinate r only, i.e. $U = U(r)$. Then

$$\nabla^2 U = \frac{\partial^2 U}{\partial r^2} + \frac{2}{r}\frac{\partial U}{\partial r} \ . \tag{9.30}$$

For cylindrical coordinates, formula (9.28) takes the form

$$\nabla^2 U = \frac{\partial^2 U}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial U}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2 U}{\partial \lambda^2} + \frac{\partial^2 U}{\partial z^2} \ . \tag{9.31}$$

For the special case of the axial symmetry of function U, i.e. $U = U(\rho)$, one gets

$$\nabla^2 U = \frac{\partial^2 U}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial U}{\partial \rho} \ . \tag{9.32}$$

We should mention a similar form of expressions (9.32) and (9.30), with the exception of the coefficients of the last terms.

Let us add a remark to the special formulae (9.30) and (9.32). These formulae are very similar, with the exception of the coefficients with the last terms. This is a consequence of a more general formula, namely formula (9.42), which will be derived in the next subsection.

9.1.5 An elementary derivation of the Laplacian for spherically symmetric and axially symmetric functions

Here we shall perform a brief derivation of special formulae (9.30) and (9.32), without referring to the general formulae (9.28).

Consider a function U which depends on the spherical coordinate r only, $U = U(r)$. This coordinate is related to the Cartesian coordinates x, y, z by the formula

$$r = \sqrt{x^2 + y^2 + z^2} \ . \tag{9.33}$$

On differentiating this formula, one gets

$$\frac{\partial r}{\partial x} = \frac{x}{r}, \quad \frac{\partial r}{\partial y} = \frac{y}{r}, \quad \frac{\partial r}{\partial z} = \frac{z}{r} \ . \tag{9.34}$$

Since U is a function of r only, we have

$$\frac{\partial U}{\partial x} = \frac{\partial U}{\partial r}\frac{\partial r}{\partial x} = \frac{\partial U}{\partial r}\frac{x}{r} \ , \tag{9.35}$$

$$\frac{\partial^2 U}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial U}{\partial r}\frac{x}{r}\right) = \frac{\partial^2 U}{\partial r^2}\left(\frac{x}{r}\right)^2 + \frac{\partial U}{\partial r}\frac{\partial}{\partial x}\left(\frac{x}{r}\right) = \frac{\partial^2 U}{\partial r^2}\frac{x^2}{r^2} + \frac{\partial U}{\partial r}\left(\frac{1}{r} - \frac{x^2}{r^3}\right) \ . \tag{9.36}$$

The formulae for $\partial^2 U/\partial y^2$ and $\partial^2 U/\partial z^2$ are analogous. For the Laplacian,

$$\nabla^2 U = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} \ , \tag{9.37}$$

we then get formula (9.30).

In cylindrical coordinates we have

$$\rho = \sqrt{x^2 + y^2} \ , \tag{9.38}$$

which yields

$$\frac{\partial \rho}{\partial x} = \frac{x}{\rho}, \quad \frac{\partial \rho}{\partial y} = \frac{y}{\rho} \ . \tag{9.39}$$

For a function $U = U(\rho)$ we then get, in analogy with (9.36),

$$\frac{\partial^2 U}{\partial x^2} = \frac{\partial^2 U}{\partial \rho^2}\frac{x^2}{\rho^2} + \frac{\partial U}{\partial \rho}\left(\frac{1}{\rho} - \frac{x^2}{\rho^3}\right) \ , \tag{9.40}$$

and the similar expression for $\partial^2 U/\partial y^2$. Since $\partial^2 U/\partial z^2 = 0$ now, the Laplacian (9.37) yields formula (9.32).

The derivations just described may easily be generalised to the n-dimensional case. Consider a function $U = U(r)$, where r is the radius vector in the n–dimensional Euclidean space. Instead of (9.34) we now have

$$r = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} \; . \tag{9.41}$$

In view of formula (9.36), which holds true for all x_1 to x_n, the Laplacian takes the following form:

$$\nabla^2 U = \frac{\partial^2 U}{\partial x_1^2} + \ldots + \frac{\partial^2 U}{\partial x_n^2} = \frac{\partial^2 U}{\partial r^2}\frac{x_1^2 + \ldots + x_n^2}{r^2} + \frac{\partial U}{\partial r}\left(\frac{n}{r} - \frac{x_1^2 + \ldots + x_n^2}{r^3}\right) .$$

Taking into account (9.41), this yields

$$\nabla^2 U = \frac{\partial^2 U}{\partial r^2} + \frac{n-1}{r}\frac{\partial U}{\partial r} \; . \tag{9.42}$$

As the special case of this formula for $n = 3$, we get formula (9.30) in spherical coordinates, and for $n = 2$ we obtain formula (9.32) in cylindrical coordinates (if we write ρ instead of r). This can be explained in the following way. A spherically symmetric function is a function defined in a 3–dimensional space, and so $n = 3$. However, a function in cylindrical coordinates which is independent of coordinate z is, in fact, a function defined in a 2–dimensional space (in the plane $z = 0$). Thus, we must put $n = 2$.

9.2 Spherical Waves

Consider the wave equation for the scalar potential φ,

$$\nabla^2 \varphi = \frac{1}{\alpha^2}\frac{\partial^2 \varphi}{\partial t^2} \, , \tag{9.43}$$

where α is the longitudinal-wave velocity (a constant). Let us seek the solution of this equations in the special form $\varphi = \varphi(r, t)$, where r is the distance from the coordinate origin, and t is the time. Using formula (9.30) for the Laplacian, the wave equation (9.43) becomes

$$\frac{\partial^2 \varphi}{\partial r^2} + \frac{2}{r}\frac{\partial \varphi}{\partial r} = \frac{1}{\alpha^2}\frac{\partial^2 \varphi}{\partial t^2} \; . \tag{9.44}$$

Introduce a new function $\Phi = r\varphi$. Since

$$\frac{\partial \Phi}{\partial r} = \varphi + r\frac{\partial \varphi}{\partial r} , \quad \frac{\partial^2 \Phi}{\partial r^2} = 2\frac{\partial \varphi}{\partial r} + r\frac{\partial^2 \varphi}{\partial r^2} , \tag{9.45}$$

the wave equation (9.44) may be rewritten as

$$\frac{1}{r}\frac{\partial^2 (r\varphi)}{\partial r^2} = \frac{1}{\alpha^2}\frac{\partial^2 \varphi}{\partial t^2} . \tag{9.46}$$

Multiplying this equation by r (since the coordinate r is not a function of time), we finally get

$$\frac{\partial^2 (r\varphi)}{\partial r^2} = \frac{1}{\alpha^2}\frac{\partial^2 (r\varphi)}{\partial t^2} . \tag{9.47}$$

This equation for $r\varphi$ has the same form as the well-known wave equation (7.9a) for waves propagating along one Cartesian coordinate. Analogously to the solution (7.13) in the Cartesian coordinates, we can express the general solution of Eq. (9.47) in the form

$$\varphi(r,t) = \frac{1}{r} f_1\left(t - \frac{r}{\alpha} + t_1\right) + \frac{1}{r} f_2\left(t + \frac{r}{\alpha} + t_2\right) , \tag{9.48}$$

where f_1 and f_2 are arbitrary functions, characterising the form of the wave, and t_1, t_2 are arbitrary constants. For the special case of $t_1 = t_2 = 0$, we have

$$\varphi(r,t) = \frac{1}{r} f_1\left(t - \frac{r}{\alpha}\right) + \frac{1}{r} f_2\left(t + \frac{r}{\alpha}\right) . \tag{9.49}$$

It follows from the expression (9.49) that the surfaces of a constant phase $\tau = t \pm r/\alpha = \tau_0$ are spherical in this case, the centre of the sphere being at the origin of the coordinate system The wave with the sign "–" represents an expanding wave propagating in the direction from the origin. Its wave surface extends with increasing time. The wave with the sign "+" represents a contracting wave with a diminishing wave surface. In an unlimited homogenous medium, we shall consider only the first solution, i.e.

$$\varphi(r,t) = \frac{1}{r} f\left(t - \frac{r}{\alpha}\right) . \tag{9.50}$$

The solution of this form is called the *spherical wave*.

Analogously for a harmonic spherical wave, we have

$$\varphi(r, t, \omega) = \frac{A}{r} e^{i\omega(t - r/\alpha)} , \tag{9.51}$$

where ω is the angular frequency and A is a constant.

One important fact should be pointed out, namely the decrease of the amplitudes of spherical waves with distance. While the amplitudes of plane waves do not change during the propagation, the amplitudes of spherical waves decrease as $1/r$, see formulae (9.50) and (9.51). This fact is easily understandable, because the wave surfaces of a spherical wave are expanding, as opposed to the wave surfaces of plane waves.

In Chapter 2 we have derived the decrease of amplitudes of spherical waves (body waves) on the basis of certain considerations on the conservation of energy. Here, we have derived this decrease directly from the wave equation.

Spherical waves play a fundamental role in the theory of elastic waves and in seismic practice, too. Formula (9.50) gives the so-called fundamental solution of the wave equation. This solution is similar, e.g., to the function $1/r$, representing the fundamental solution of Laplace's equation in gravimetry. If the sources of seismic waves are located inside a certain homogeneous region, it is then relatively easy to express the source function as an integral superposition of spherical waves. Moreover, if we observe seismic waves at a large distance from a source, the dimensions of the source may usually be neglected, and a spherical wave itself yields a good approximation to the wave field.

The importance of spherical waves is also emphasised by Huygens' principle, which is frequently used not only in optics, but also in some methods of seismic prospecting. Huygens' principle states that every point of a medium, at which a disturbance has arrived, can be considered as a new source of a disturbance which propagates from this point in the form of a spherical wave. These waves are superposed in such a way that their envelope determines the resultant wave surface in any later time instant. The mathematical formulation of Huygens' principle is given by Kirchhoff's formula, which will be derived in the next chapter.

9.3 Cylindrical Waves

After discussing the simplest waves in Cartesian and spherical coordinates, let us consider the case of cylindrical coordinates. Denote these coordinates by ρ, λ, z; see Fig. 9.1b and formulae (9.2). We shall see that some properties of cylindrical waves are substantially different from the properties of plane or spherical waves. Consequently, we shall restrict ourselves to harmonic waves only. Moreover, we shall consider longitudinal waves only.

Let us, therefore, study whether a longitudinal, harmonic, cylindrical wave, dependent on coordinate ρ and time t only, may propagate in a homogeneous isotropic medium. Let us seek its potential in the form

$$\varphi(\rho, t, \omega) = \Phi(\rho, \omega) e^{i\omega t} , \tag{9.52}$$

where ω is the angular frequency.

Since the Laplacian is now given by (9.32), the wave equation (9.43) takes the form

$$\frac{\partial^2 \varphi}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial U}{\partial \rho} = \frac{1}{\alpha^2}\frac{\partial^2 \varphi}{\partial t^2} \; . \tag{9.53}$$

Inserting function (9.52) into this equation, we obtain the following *Helmholtz equation for cylindrical waves*:

$$\frac{\mathrm{d}^2 \Phi}{\mathrm{d}\rho^2} + \frac{1}{\rho}\frac{\mathrm{d}\Phi}{\mathrm{d}\rho} + k^2 \Phi = 0 \; , \tag{9.54}$$

where $k = \omega/\alpha$.

Modify the latter equation by introducing a new variable $\xi = k\rho$. Since

$$\frac{\mathrm{d}\Phi}{\mathrm{d}\rho} = \frac{\mathrm{d}\Phi}{\mathrm{d}\xi}\frac{\mathrm{d}\xi}{\mathrm{d}\rho} = k\frac{\mathrm{d}\Phi}{\mathrm{d}\xi} \; , \quad \frac{\mathrm{d}^2 \Phi}{\mathrm{d}\rho^2} = k^2 \frac{\mathrm{d}^2 \Phi}{\mathrm{d}\xi^2} \; ,$$

equation (9.54) yields (after dividing by k^2)

$$\frac{\mathrm{d}^2 \Phi}{\mathrm{d}\xi^2} + \frac{1}{\xi}\frac{\mathrm{d}\Phi}{\mathrm{d}\xi} + \Phi = 0 \; . \tag{9.55}$$

This is a special form of Bessel's equation.

9.3.1 Bessel functions

The famous *Bessel's equation* has the form (Abramowitz and Stegun, 1972)

$$\frac{\mathrm{d}^2 w}{\mathrm{d}x^2} + \frac{1}{x}\frac{\mathrm{d}w}{\mathrm{d}x} + \left(1 - \frac{n^2}{x^2}\right) w = 0 \tag{9.56a}$$

or, after multiplying by x^2,

$$x^2 \frac{\mathrm{d}^2 w}{\mathrm{d}x^2} + x\frac{\mathrm{d}w}{\mathrm{d}x} + \left(x^2 - n^2\right) w = 0 \; . \tag{9.56b}$$

Its solution, finite for $x = 0$, is the Bessel function of the n-th order,

$$w(x) = J_n(x) \; . \tag{9.57}$$

The graphs of several Bessel functions of low orders are shown in Fig. 9.4. The Bessel functions oscillate but are not periodic. The amplitude of $J_n(x)$ is not constant but decreases asymptotically as $1/\sqrt{x}$.

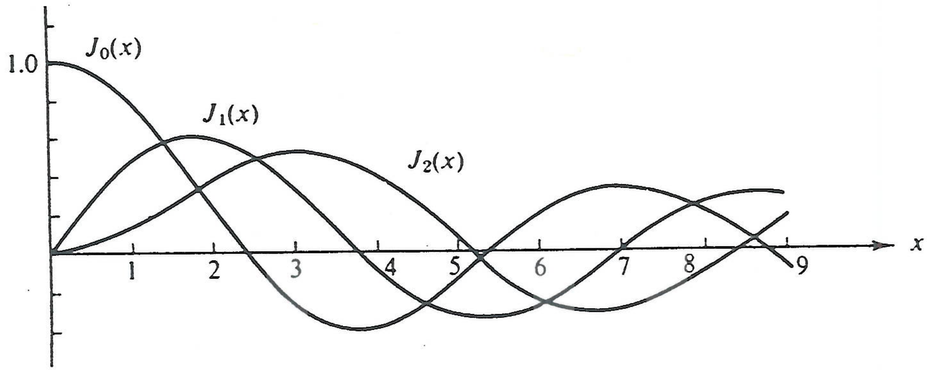


Fig. 9.4. Bessel functions of the first kind, $J_0(x)$, $J_1(x)$, and $J_2(x)$. (After Arfken (1970)).

Let us give several important formulae for Bessel functions:

a) Integral representation:

$$J_0(x) = \frac{1}{\pi}\int_0^{\pi} e^{ix\cos\lambda}\,\mathrm{d}\lambda = \frac{1}{2\pi}\int_0^{2\pi} e^{\pm ix\cos(\lambda-\chi)}\,\mathrm{d}\lambda , \tag{9.58}$$

where χ is an arbitrary constant.

b) Ascending series:

$$J_0(x) = 1 - \frac{\left(\frac{1}{4}x^2\right)}{(1!)^2} + \frac{\left(\frac{1}{4}x^2\right)^2}{(2!)^2} - \frac{\left(\frac{1}{4}x^2\right)^3}{(3!)^2} + \cdots . \tag{9.59}$$

c) Asymptotic expansion for large arguments $(x \to \infty)$:

$$J_n(x) \approx \sqrt{\frac{2}{\pi x}}\cos\left(x - n\frac{\pi}{2} - \frac{\pi}{4}\right) . \tag{9.60}$$

In particular for $n = 0$,

$$J_0(x) \approx \sqrt{\frac{2}{\pi x}}\cos\left(x - \frac{\pi}{4}\right) . \tag{9.61}$$

It follows from Fig. 9.4 and from the last formula that function $J_0(x)$ resembles a "damped" cosine function.

9.3.2 Harmonic cylindrical waves and their properties

It can be seem that Eq. (9.55) is a special form of Bessel's equation for $n = 0$. Consequently,

$$\Phi = J_0(k\rho) \ ,$$

and the potential of the corresponding cylindrical harmonic wave may be expressed as

$$\varphi(\rho, t, \omega) = AJ_0(k\rho)e^{i\omega t} \ , \tag{9.62}$$

where $k = \omega/\alpha$ and A is an arbitrary constant.

Since we assume that potential φ is independent of the z-coordinate, solution (9.62) may be regarded as the fundamental harmonic solution of the wave equation in the two-dimensional space. Note that we have considered only harmonic cylindrical waves, as opposed to a general form of plane and spherical waves, discussed above. This has deep physical reasons, which become evident from the comparison of the fundamental solutions in Tab. 9.1.

Table 9.1. Fundamental harmonic solutions of the wave equation in one-, two- and three-dimensional media: A is a constant, ω the angular frequency, c the velocity ($c = \alpha$ for longitudinal waves, $c = \beta$ for transverse waves), $k = \omega/c$ the wavenumber, and x, ρ, r are the distances from the source, respectively.

Medium	Fundamental solution
$1-\text{D}$	$A\,e^{i(\omega t-kx)}$
$2-\text{D}$	$A\,J_0(k\rho)\,e^{i\omega t}$
$3-\text{D}$	$\dfrac{A}{r}\,e^{i(\omega t-kr)}$

It can be seem from this table that the fundamental solution for a $2-\text{D}$ medium differs substantially from waves in $1-\text{D}$ and $3-\text{D}$ media. The fundamental waves in $1-\text{D}$ and $3-\text{D}$ media have the form

$$fe^{i(\omega t-kd)} \ ,$$

where d is the distance from the source ($d = x$ and $d = r$, respectively) and f is an amplitude ($f = A$ and $f = A/r$, respectively). In these cases, the angular frequency ω enters only the exponential factors, not amplitude f. As opposed to it, the variation of the amplitude in a $2-\text{D}$ medium depends on frequency. Consequently, the fundamental waves in $1-\text{D}$ and $3-\text{D}$ media (plane and spherical waves) are *non-dispersive*, whereas the waves in $2-\text{D}$ media (cylindrical waves) are *dispersive*. The dispersion causes that the wave changes its form during the propagation, e.g., a narrow impulse becomes broader with increasing distance. This property applies generally to all types of waves in $2-\text{D}$ media.

Assuming the conservation of the mechanical energy, in Chapter 2 we derived that the amplitudes of cylindrical waves (surface waves) decreased with distance from the source as $1/\sqrt{\rho}$. Formulae (9.62) and (9.61) give even a better estimate of this decrease, namely that the amplitude of a cylindrical harmonic wave, a, decreases at large distances as

$$a \sim 1/\sqrt{k\rho}\,, \tag{9.63}$$

$k = \omega/\alpha$ being the wavenumber. This confirms the decrease of a harmonic wave as $1/\sqrt{\rho}$ but, moreover, it yields a similar dependence on frequency. Therefore, since short-period waves decrease more rapidly than long-period ones, the wavefield at large distances becomes gradually long-periodic. (In real media, this decrease of the amplitudes of short-period waves is further emphasised by their higher attenuation).

Although formula (9.63) is valid for cylindrical waves in a homogeneous unlimited medium, to some extent this also applies to surface waves propagating in a homogeneous half-space or in a layered medium. For example, this explains the long-period character of surface waves at large epicentral distances.

Chapter 10

Solutions of the Wave Equation in Integral Forms

In the previous two chapters we studied the simplest solutions of the wave equation. Now we shall proceed to more complicated solutions of this equation.

In mathematics, complicated functions are frequently expressed in the form of integrals or infinite series. Both these representations are also used in the theory of elastic waves. In this chapter we shall deal with some integral forms of the solution of the wave equation. The solutions in the forms of infinite series will be studied in Chapter 11.

Our methods of solving the wave equation will be very similar to the methods in gravimetry, where simpler equations, Laplace's or Poisson's equations, are solved. Therefore, let us remind some of these gravimetric problems.

First, consider the gravitational field of a point mass m. According to Newton's gravitational law, its potential at a distance r from the source is $U(r) = -Gm/r$, where G is the gravitational constant. If the mass is distributed in a finite volume V with density ρ, the gravitational potential (according to the superposition principle) is

$$U(P) = -G \iiint_V \frac{\rho}{r} \mathrm{d}V \ , \tag{10.1}$$

where r is the distance of the mass element, $\mathrm{d}m = \rho \mathrm{d}V$, from the point of observation, P. If the distribution of density is known, the problem of determining the gravitational field reduces to the problem of computing the integral in (10.1).

However, we known the density within the Earth with a low accuracy only. Consequently, we cannot compute the interior field accurately. Nevertheless, the situation is better in studying the exterior field (at least in principle). Namely, we can avoid the unknown density and to use other data. There are two basic possibilities:

a) Using Green's second theorem, the exterior field may be expressed by means of two integrals over the surface of the Earth (instead of the volume integral (10.1)). For this representation we must know potential U and its normal derivative $\partial U / \partial n$ on the Earth's surface (see Section 10.2).
b) According to the multipole-expansion theorem, the field may be expressed in the form of a series, the coefficients of which are determined from surface or satellite measurements.

In this chapter we shall describe approach a). First, we shall deal with the corresponding mathematical theory, then we shall mention the gravimetric applications, and finally some analogous seismic problems will be solved.

10.1 Green's Theorems

First, we shall derive Green's first theorem, which will be used later on. We assume that the reader is familiar with Gauss' theorem. Therefore, we shall start with this theorem but we shall not prove it here.

A summary of the main equations, which will be derived here, is given in the left column of Tab. 10.1; see below.

10.1.1 Gauss' theorem

Let V be a finite volume, bounded by a surface S. The surface should satisfy certain properties, e.g., be composed of smooth parts. Consider a vector $\mathbf{A} = \mathbf{A}(x, y, z)$ which is continuous together with its first partial derivatives within V and on boundary S. *Gauss' theorem,* also called the *divergence theorem*, then states that

$$\boxed{\iiint_V \operatorname{div} \mathbf{A} \, \mathrm{d}V = \iint_S A_n \, \mathrm{d}S}\,, \tag{10.2}$$

where

$$\operatorname{div} \mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}\,, \tag{10.3}$$

and A_n is the component of vector $\mathbf{A}$ in the direction of the outward directed normal to surface S. Gauss' theorem can be expressed in words as follows: The integral of the divergence of a vector field over a region of space is equal to the integral over the surface of that region of the component of the field in the direction of the outward directed normal to the surface (Kellogg, 1967).

Gauss' theorem makes it possible to transform a certain volume integral into a surface integral.

10.1.2 Green's preliminary formula

Let us apply Gauss' theorem to a vector $a\mathbf{A}$ where $a = a(x, y, z)$ is a scalar function. First, let us calculate $\operatorname{div}(\mathrm{a}\,\mathbf{A})$:

$$\operatorname{div}(a\mathbf{A}) = \frac{\partial (aA_x)}{\partial x} + \frac{\partial (aA_y)}{\partial y} + \frac{\partial (aA_z)}{\partial z} =$$

$$= a\left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}\right) + \frac{\partial a}{\partial x} A_x + \frac{\partial a}{\partial y} A_y + \frac{\partial a}{\partial z} A_z\,.$$

We have arrive at the well-known vector identity,

$$\boxed{\operatorname{div}(a\mathbf{A}) = a \operatorname{div} \mathbf{A} + \mathbf{A} \cdot \operatorname{grad} a}\,. \tag{10.4}$$

Now, apply Gauss' theorem to the vector $(a\mathbf{A})$, and use (10.4):

$$\iiint_V (a \operatorname{div} \mathbf{A} + \mathbf{A} \cdot \operatorname{grad} a) \, \mathrm{d}V = \iint_S a A_n \, \mathrm{d}S \, . \tag{10.5}$$

We shall rewrite this theorem for a special vector field **A**, namely for a conservative field (described by a scalar potential *W*). Thus, consider scalar functions *U*, *V* such that

$$a = U \, , \quad \mathbf{A} = \operatorname{grad} W \, , \tag{10.6}$$

where *U*, *W* are uniquely defined and continuous together with their first and second derivatives. By inserting (10.6) into (10.5) we get

$$\boxed{\iiint_V U \nabla^2 W \, \mathrm{d}V + \iiint_V (\operatorname{grad} U) \cdot (\operatorname{grad} W) \, \mathrm{d}V = \iint_S U \frac{\partial W}{\partial n} \, \mathrm{d}S} \, . \tag{10.7}$$

This is the so-called *Green's preliminary formula.* This is also written in the following form (Pick et al., 1973):

$$\iiint_V \left[\frac{\partial U}{\partial x}\frac{\partial W}{\partial x} + \frac{\partial U}{\partial y}\frac{\partial W}{\partial y} + \frac{\partial U}{\partial z}\frac{\partial W}{\partial z} \right] \mathrm{d}V = \iint_S U \frac{\partial W}{\partial n} \, \mathrm{d}S - \iiint_V U \nabla^2 W \, \mathrm{d}V \, . \tag{10.8}$$

10.1.3 Green's first theorem

Interchange functions *U* and *W* in Green's preliminary formula (10.7), and subtract the corresponding equations. This yields *Green's first theorem*:

$$\boxed{\iiint_V \left(U \nabla^2 W - W \nabla^2 U \right) \mathrm{d}V = \iint_S \left(U \frac{\partial W}{\partial n} - W \frac{\partial U}{\partial n} \right) \mathrm{d}S} \, . \tag{10.9}$$

This theorem represents a generalisation of Gauss' theorem, but for conservative fields only. For $W = \text{const.}$, theorem (10.9) yields Gauss' theorem for a conservative field $\mathbf{A} = \operatorname{grad} U$:

$$\iiint_V \nabla^2 U \, \mathrm{d}V = \iint_S \frac{\partial U}{\partial n} \, \mathrm{d}S \, . \tag{10.10}$$

10.1.4 Green's second theorem

Let us choose function *W* in Green's first theorem as follows:

$$W = 1/r \, , \tag{10.11}$$

where r is the distance from a fixed point, P. Since function W is singular at point P, we must distinguish whether point P is inside the volume V, on its surface, or outside.

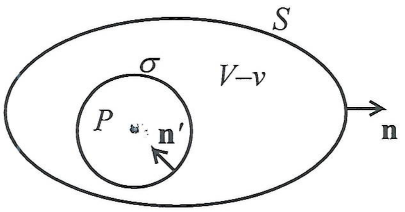


Fig 10.1. The region used to derive Green's second theorem.

First, consider point P inside V; see Fig. 10.1. Select a small sphere with its centre at P, and exclude it from region V. Denote the radius of the sphere by ε, its volume by v and surface by σ. Denote the volume V without the sphere by $V-v$, and apply Green's first theorem to it. The volume $V-v$ is bounded by two surfaces, namely by the surface S with its outward normal $\mathbf{n}$ and by the surface σ with normal $\mathbf{n}'$. Since the normal $\mathbf{n}'$ must be outward with respect to volume $V-v$, its orientation is inwards the sphere. Therefore, instead of one surface integral in Eq. (10.9), we must consider two integrals, one over S and the other over σ. As function $1/r$ satisfies Laplace's equation in $V-v$, i.e. $\nabla^2(1/r)=0$, we can write Green's first theorem in this special case in the following form:

$$-\iiint\limits_{V-v}\frac{1}{r}\nabla^2 U\,\mathrm{d}V-\iint\limits_{S}U\frac{\partial}{\partial n}\left(\frac{1}{r}\right)\mathrm{d}S+\iint\limits_{S}\frac{1}{r}\frac{\partial U}{\partial n}\mathrm{d}S=\iint\limits_{\sigma}U\frac{\partial}{\partial n'}\left(\frac{1}{r}\right)\mathrm{d}\sigma-\iint\limits_{\sigma}\frac{1}{r}\frac{\partial U}{\partial n'}\mathrm{d}\sigma \tag{10.12}$$

Let us rearrange the expressions on the right-hand side of the latter equation. Since normal $\mathbf{n}'$ has the direction opposite to the radial coordinate r, we have

$$\frac{\partial}{\partial n'}\left(\frac{1}{r}\right)=-\frac{\partial}{\partial r}\left(\frac{1}{r}\right)=\frac{1}{r^2}\,,\quad \frac{\partial U}{\partial n'}=-\frac{\partial U}{\partial r}\ .$$

Taking into account that $r=\varepsilon$ on σ, the right-hand side of Eq. (10.12), which we shall denote by I, may be expressed as

$$I=\frac{1}{\varepsilon^2}\iint\limits_{\sigma}U\,\mathrm{d}\sigma+\frac{1}{\varepsilon}\iint\limits_{\sigma}\frac{\partial U}{\partial r}\mathrm{d}\sigma\ . \tag{10.13}$$

Considering the continuity of U and $\partial U/\partial r$, we may apply the mean value theorem to the latter integrals, which yields

$$I = \frac{1}{\varepsilon^2} U(P_1) 4\pi\varepsilon^2 + \frac{1}{\varepsilon}\left(\frac{\partial U}{\partial r}\right)_{P_2} 4\pi\varepsilon^2 , \tag{10.14}$$

where P_1 and P_2 are some points on surface σ.

Now, let us shrink sphere v to point P, i.e. pass to the limit for $\varepsilon \to 0$. The second term in (10.14) vanishes, and so

$$I = 4\pi U(P) . \tag{10.15}$$

The first integral in Eq. (10.12) over volume $V - v$ then becomes the integral over the whole volume V. We can extend the integration over the whole volume without problems, although the integrand is singular. In order to show it, let us consider the corresponding integral over sphere v:

$$J = \iiint_v \frac{1}{r} \nabla^2 U \,\mathrm{d}v .$$

As the consequence of the continuity of the second derivatives of U, these derivatives are limited. Therefore, there is a finite constant M, such that $\left|\nabla^2 U\right| \le M$. Then

$$\left|J\right| \le \iiint_v \left|\frac{1}{r} \nabla^2 U\right| \mathrm{d}v \le M \iiint_v \frac{\mathrm{d}v}{r} .$$

Let us calculate the latter integral in spherical coordinates r, ϑ, λ. If holds that $\mathrm{d}v = r^2 \sin\vartheta \,\mathrm{d}r \,\mathrm{d}\vartheta \,\mathrm{d}\lambda$; see formula (9.8). We get

$$\iiint_v \frac{\mathrm{d}v}{r} = \int_0^{\varepsilon}\int_0^{\pi}\int_0^{2\pi} r \sin\vartheta \,\mathrm{d}r \,\mathrm{d}\vartheta \,\mathrm{d}\lambda = 2\pi\varepsilon^2 .$$

Therefore, $J \to 0$ for $\varepsilon \to 0$.

Consequently, for an internal point P, Eq. (10.12) may be expressed in the following final form:

$$\boxed{- \iiint_V \frac{1}{r} \nabla^2 U \,\mathrm{d}V - \iint_S U \frac{\partial}{\partial n}\left(\frac{1}{r}\right) \mathrm{d}S + \iint_S \frac{1}{r} \frac{\partial U}{\partial n} \mathrm{d}S = 4\pi U(P)} . \tag{10.16}$$

This is *Green's second theorem* for point P inside volume V.

If point P is located on the surface of volume V, i.e. on surface S, we must remove only a hemisphere from volume V. Consequently, we shall obtain $2\pi U(P)$ on the right-hand side of Eq. (10.16). Finally, it point P is outside V, we do not need to exclude any sphere or hemisphere, so that the right-hand side of Eq. (10.16) will be equal to zero.

Consequently, we arrive at the following complete form of Green's second theorem:

$$-\iiint_V \frac{1}{r}\nabla^2 U \,\mathrm{d}V - \iint_S U \frac{\partial}{\partial n}\left(\frac{1}{r}\right)\mathrm{d}S + \iint_S \frac{1}{r}\frac{\partial U}{\partial n}\mathrm{d}S = \begin{cases} 4\pi U(P), & P \text{ inside } V, \\ 2\pi U(P), & P \text{ on } S, \\ 0, & P \text{ outside } V. \end{cases} \tag{10.17}$$

Note that, from the point of view of applications, the most important of these forms is the form for the internal point P, i.e. Eq. (10.16).

It follows from (10.16) that a function which is continuous together with its first and second derivatives cannot take an arbitrary value at a given point P. Its value, $U(P)$, is closely related to the behaviour of this function in a certain vicinity of this point. Even, the integrals on the left-hand side of (10.16) may be interpreted as Newtonian potentials for a volume distribution, double layer and single layer, respectively.

10.1.5 Special forms of Green's second theorem for the gravitational field

Denote the intensity of the gravitational field by $\mathbf{E}$, and the gravitational potential by U, $\mathbf{E} = -\operatorname{grad} U$. Let us restrict ourselves only to volume distributions of mass. Assume that the density, ρ, satisfies certain conditions which are sufficient for the validity of Poisson's equation,

$$\nabla^2 U = 4\pi G\rho \, , \tag{10.18}$$

where G is the gravitational constant. For example, the density should have its second derivatives continuous, or to satisfy Hölder's conditions, which are weaker (Pick et al., 1973). Inserting Poisson's equation (10.18) into the volume integral of Green's second theorem for the internal point, i.e. into Eqs. (10.16) and (10.17), we arrive at

$$U(P) = -G\iiint_V \frac{\rho}{r}\mathrm{d}V - \frac{1}{4\pi}\iint_S U \frac{\partial}{\partial n}\left(\frac{1}{r}\right)\mathrm{d}S + \frac{1}{4\pi}\iint_S \frac{1}{r}\frac{\partial U}{\partial n}\mathrm{d}S \, . \tag{10.19}$$

This is Green's second theorem specified for the gravitational field. Note that the point P, where we calculate the gravitational field, must be located inside the region V. However, this region may be rather arbitrary, and so the masses generating the gravitational field may be all inside V, all outside V, or even partially inside and partially outside (Fig. 10.2). Let us discuss these cases separately.

a) *All masses inside V.* Assume all masses to be contained in a bounded region V_m which lies inside the integration domain V; sec Fig. 10.2a. In this case, point P may be either outside or inside V_m. Since density ρ is non-zero in volume V_m only, the volume integral over V in Eq. (10.19) reduces to the integration over V_m only. Let us extend the region V so that it becomes a sphere of a large radius R with the centre at P. The quantities on the surface of the sphere can be estimated as follows:

$$\frac{1}{r}=\frac{1}{R}, \quad \frac{\partial}{\partial n}\left(\frac{1}{r}\right)=\frac{\partial}{\partial R}\left(\frac{1}{R}\right)=-\frac{1}{R^2}, \quad U\sim-\frac{GM}{R}, \quad \frac{\partial U}{\partial n}=\frac{\partial U}{\partial R}\sim\frac{GM}{R^2},$$

where M is the total mass in V_m. Therefore, the integrands in both surface integrals in (10.19) can be approximated by GM/R^3, and the integrals themselves by $4\pi GM/R$. For R tending to infinity, these surface integrals vanish, and we arrive at

$$U(P)=-G\iiint_{V_m}\frac{\rho}{r}\,\mathrm{d}V\ . \tag{10.20}$$

We have obtained the well-known formula for calculating the gravitational potential if the distribution of density is known. Therefore, Green's second theorem (10.19) includes this important formula as a special case.

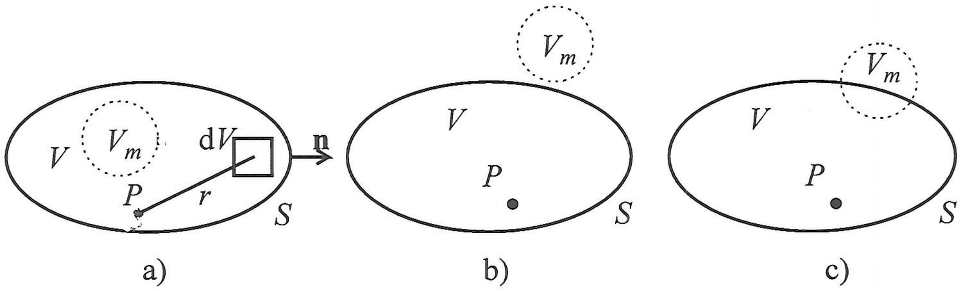


Fig. 10.2. Mutual position of the masses and the integration domain in Greens's second theorem (10.19). Volume V is the integration domain, S its surface, P a point inside V where the gravitational field is calculated, V_m the region where the masses are distributed.

b) *All masses outside V.* It happens very often that the distribution of masses is not known. The gravitational field inside the masses then cannot be calculated. However, the external field can be calculated if compensatory information is known. Namely, outside the masses, Poisson's equation (10.18) becomes Laplace's equation,

$$\nabla^2 U=0\ , \tag{10.21}$$

which can be solved if appropriate boundary conditions are given.

Green's second theorem makes it possible to solve this problem in an integral form. If $\rho=0$ inside V, see Fig. 10.2b, Eq. (10.19) yields

$$U(P) = -\frac{1}{4\pi}\iint_S U \frac{\partial}{\partial n}\left(\frac{1}{r}\right) \mathrm{d}S + \frac{1}{4\pi}\iint_S \frac{1}{r}\frac{\partial U}{\partial n} \mathrm{d}S \,. \qquad (10.22)$$

This solution is relatively simple; this is the sum of potentials of a double layer and of a single layer. However, two functions must be known on surface *S*, namely potential *U* and its normal derivative $\partial U/\partial n$. It can be seen that the position of point *P* influences these integrals through the reciprocal distance, $1/r$, and the normal derivative of $1/r$.

It is well-known that Laplace's equation can also be solved if only one of the above-mentioned functions is known on *S*. The formulation for a given *U* on surface *S* is called Dirichlet's problem, and its solution may be expressed as

$$U(P) = -\frac{1}{4\pi}\iint_S U \frac{\partial G_1}{\partial n} \mathrm{d}S \,, \qquad (10.23)$$

where function G_1 is called Green's function (sometimes it is also called Green's first function). Instead of two integrals in Eq. (10.22), here we have only one integral. However, this apparent simplification is compensated by a more complicated integrand; the simple function $1/r$ in Eq. (10.22) is replaced here by Green's function, which is more complicated and dependent also on the shape of surface *S*. Analytical expressions for Green's function G_1 are known only for very simple forms of surface *S*.

If $\partial U/\partial n$ is given on *S*, we speak of Neumann's problem. Its solution may be expressed as

$$U(P) = \frac{1}{4\pi}\iint_S G_2 \frac{\partial U}{\partial n} \mathrm{d}S + U_0 \,, \qquad (10.24)$$

where U_0 is a constant. Function G_2, called Neumann's function, is generally more complicated then $1/r$, and depends also on the shape of *S*. Since functions G_1 and G_2 generally differ from function $1/r$, potentials (10.23) and (10.24) cannot be interpreted as potentials of a double and single layer, respectively.

Formulae of type (10.22) to (10.24) play a fundamental role in the classical theory of the Earth's gravity field and shape of the Earth (Pick et al., 1973).

c) *Only a part of the masses inside V*. This situation is shown in Fig. 10.2c. The potential at point *P* can also be determined if density ρ is known inside region *V* only (and unknown outside), and if *U* and $\partial U/\partial n$ are known on surface *S*.

10.2 Inhomogeneous Wave Equations and Helmholtz Equations

In the following sections of this chapter, we shall solve the inhomogeneous equation of motion for a homogeneous isotropic medium, i.e. the equation containing body forces:

$$(\lambda + \mu)\,\mathrm{grad}\,\mathrm{div}\,\mathbf{u} + \mu\nabla^2\mathbf{u} + \rho\mathbf{g} = \rho\frac{\partial^2\mathbf{u}}{\partial t^2}\ . \tag{10.25}$$

Express the displacement **u** and body force **g** in terms of potentials φ, $\vec{\psi}$, q and **p** as follows:

$$\mathbf{u} = \mathrm{grad}\,\varphi + \mathrm{curl}\,\vec{\psi}\ ,\quad \mathbf{g} = \mathrm{grad}\,q + \mathrm{curl}\,\mathbf{p}\ . \tag{10.26}$$

The equation of motion (10.25) then takes the form

$$\mathrm{grad}\left[(\lambda + 2\mu)\nabla^2\varphi + \rho q - \rho\frac{\partial^2\varphi}{\partial t^2}\right] + \mathrm{curl}\left[\mu\nabla^2\vec{\psi} + \rho\mathbf{p} - \rho\frac{\partial^2\vec{\psi}}{\partial t^2}\right] = 0\ . \tag{10.27}$$

This equation will be satisfied if the expressions in the square brackets are equal to zero; see the analogous discussion in Subsection 7.1.1. This yields the following *inhomogeneous wave equations*:

$$\alpha^2\nabla^2\varphi + q = \frac{\partial^2\varphi}{\partial t^2}\ ,\quad \alpha = \sqrt{\frac{\lambda + 2\mu}{\rho}}\ , \tag{10.28}$$

$$\beta^2\nabla^2\vec{\psi} + \mathbf{p} = \frac{\partial^2\vec{\psi}}{\partial t^2}\ ,\quad \beta = \sqrt{\frac{\mu}{\rho}}\ . \tag{10.29}$$

Let us consider only the first of these equations, Eq. (10.28), and rewrite it in the frequency domain. Therefore, let functions φ and q be harmonic functions (harmonic components of a general signal):

$$\begin{aligned} \varphi(x_1, x_2, x_3, t, \omega) &= \Phi(x_1, x_2, x_3, \omega)e^{i\omega t}\ , \\ q(x_1, x_2, x_3, t, \omega) &= Q(x_1, x_2, x_3, \omega)e^{i\omega t}\ , \end{aligned} \tag{10.30}$$

where ω is the angular frequency, and functions Φ, Q depend on the coordinates and ω. By inserting these expressions into Eq. (10.28), we arrive at the *inhomogeneous Helmholtz equation*:

$$\alpha^2\nabla^2\Phi + Q + \omega^2\Phi = 0\ . \tag{10.31}$$

TABLE 10.1 APLICATIONS OF GREEN'S FIRST THEOREM (REVIEW OF FORMULAE).

GRAVIMETRY

Green's first theorem:

$$\iiint_V \left(U\nabla^2 W - W\nabla^2 U\right) \mathrm{d}V = \iint_S \left(U\frac{\partial W}{\partial n} - W\frac{\partial U}{\partial n}\right) \mathrm{d}S$$

Homogeneous equations for the second (auxiliary) functions:

$\nabla^2 W = 0 \ldots$ Laplace's equation

Fundamental solutions: $W = 1/r$

Special forms of Green's first theorem:

$$U(P) = -\frac{1}{4\pi}\iiint_V \frac{1}{r}\nabla^2 U \,\mathrm{d}U + \frac{1}{4\pi}\iint_S \left[\frac{1}{r}\frac{\partial U}{\partial n} - U\frac{\partial}{\partial n}\left(\frac{1}{r}\right)\right] \mathrm{d}S$$

Green's second theorem

Inhomogeneous equations for the first functions:

$\nabla^2 U = 4\pi G\rho \ldots$ Poisson's equation

General solutions for the first functions:

$$U(P) = -G\iiint_V \frac{\rho}{r}\mathrm{d}V + \frac{1}{4\pi}\iint_S \left[\frac{1}{r}\frac{\partial U}{\partial n} - U\frac{\partial}{\partial n}\left(\frac{1}{r}\right)\right] \mathrm{d}S$$

SEISMOLOGY

$$\iiint_V \left(\Phi\nabla^2 \Omega - \Omega\nabla^2 \Phi\right) \mathrm{d}V = \iint_S \left(\Phi\frac{\partial \Omega}{\partial n} - \Omega\frac{\partial \Phi}{\partial n}\right) \mathrm{d}S$$

$\nabla^2 \Omega + k^2 \Omega = 0 \ldots$ Helmholtz equation

$$\Omega = \frac{1}{r}e^{-ikr}$$

$$\Phi(P) = -\frac{1}{4\pi}\iiint_V \frac{e^{-ikr}}{r}\left(k^2\Phi + \nabla^2\Phi\right)\mathrm{d}V + $$
$$+\frac{1}{4\pi}\iint_S \left[\frac{1}{r}e^{-ikr}\frac{\partial \Phi}{\partial n} - \Phi\frac{\partial}{\partial n}\left(\frac{1}{r}e^{-ikr}\right)\right]\mathrm{d}S$$

$\nabla^2 \Phi + k^2 \Phi = -\dfrac{Q}{\alpha^2} \ldots$ inhomog. Helmholtz equation

$$\Phi(P) = \frac{1}{4\pi\alpha^2}\iiint_V \frac{e^{-ikr}}{r}Q\,\mathrm{d}V + $$
$$+\frac{1}{4\pi}\iint_S \left[\frac{1}{r}e^{-ikr}\frac{\partial \Phi}{\partial n} - \Phi\frac{\partial}{\partial n}\left(\frac{1}{r}e^{-ikr}\right)\right]\mathrm{d}S$$

Helmholtz inner formula

10.3 Helmholtz Formulae

In Section 10.1 we specified Green's first theorem for function $W = 1/r$, which represents the fundamental solution of Laplace's equation. This gave Green's second theorem, which yields a general solution of Poisson's equation. Thus, starting with a simple solution of Laplace's equation (a homogeneous equation), Green's first theorem enabled us to obtain a solution of Poisson's equation (an inhomogeneous equation).

In this section we shall derive analogous relations between the solutions of the homogeneous and inhomogeneous Helmholtz equations. Therefore, we shall seek an analogue of Green's second theorem for the Helmholtz equation. The main formulae, which will be derived in this section, are also summarised in the right column of Tab. 10.1.

10.3.1 Helmholtz inner formula

Consider again a closed surface S (Fig. 10.1). Denote the outward normal to this surface by $\mathbf{n}$ and the volume inside the surface by V. We shall calculate the wavefield at a point P. Assume the function Q, introduced in (10.30), to be limited at this point. Let some functions Φ and Ω be continuous in V together with their first and second derivatives. Green's first theorem for these functions may then be expressed as

$$\iiint_V \left(\Phi \nabla^2 \Omega - \Omega \nabla^2 \Phi \right) \mathrm{d}V = \iint_S \left(\Phi \frac{\partial \Omega}{\partial n} - \Omega \frac{\partial \Phi}{\partial n} \right) \mathrm{d}S \ . \qquad (10.32)$$

Function Φ will be the unknown solution of the inhomogeneous Helmholtz equation (10.31). As function Ω we shall choose the spatial part of the expression for a spherical wave, i.e.

$$\Omega = \frac{1}{r} e^{-ikr} \ , \qquad (10.33)$$

where $k = \omega/\alpha$ is the wavenumber, and r is the distance from point P. Using the Laplacian in spherical coordinates, see (9.30), it can easily be verified that function Ω satisfies the corresponding homogeneous Helmholtz equation (without body forces):

$$\nabla^2 \Omega + k^2 \Omega = 0 \ . \qquad (10.34)$$

Note that function W, given by (10.11), and this function Ω differ only by the factor $\exp(-ikr)$. These functions represent the fundamental solutions of the corresponding homogeneous equations.

It follows from Eqs. (10.31) and (10.34) that the integrand on the left-hand side of (10.32) is

$$\Phi\nabla^2\Omega - \Omega\nabla^2\Phi = Q\Omega/\alpha^2 \ . \tag{10.35}$$

Green's first theorem (10.32) then takes the form

$$\frac{1}{\alpha^2}\iiint_V Q\Omega \,\mathrm{d}V = \iint_S\left(\Phi\frac{\partial\Omega}{\partial n} - \Omega\frac{\partial\Phi}{\partial n}\right)\mathrm{d}S \ . \tag{10.36}$$

First, let us consider point P inside volume V. In this case, however, we cannot write Green's first theorem in the form (10.36), because function Ω is singular at point P. Therefore, we shall again exclude a small sphere with its centre at P from volume V, as shown in Fig. 10.1. Using the same notations as in Fig. 10.1, equation (10.36) now becomes

$$\frac{1}{\alpha^2}\iiint_{V-v} Q\Omega \,\mathrm{d}V - \iint_S\left(\Phi\frac{\partial\Omega}{\partial n} - \Omega\frac{\partial\Phi}{\partial n}\right)\mathrm{d}S = \iint_\sigma\left(\Phi\frac{\partial\Omega}{\partial n'} - \Omega\frac{\partial\Phi}{\partial n'}\right)\mathrm{d}\sigma \ . \tag{10.37}$$

Calculate the integral on the right-hand side of this equation, i.e. the integral over sphere σ. Since normal $\mathbf{n}'$ has the opposite direction to the radial coordinate r, one gets

$$\frac{\partial\Omega}{\partial n'} = -\frac{\partial\Omega}{\partial r} = \left[\frac{1}{r^2} + \frac{ik}{r}\right]e^{-ikr} \ , \quad \frac{\partial\Phi}{\partial n'} = -\frac{\partial\Phi}{\partial r} \ .$$

Moreover, for points on surface σ we have $r = \varepsilon$. The integral on the right-hand side of Eq. (10.37), denoted by I, then becomes

$$I = e^{-ik\varepsilon}\left[\frac{1}{\varepsilon^2}\iint_\sigma \Phi \,\mathrm{d}\sigma + \frac{ik}{\varepsilon}\iint_\sigma \Phi \,\mathrm{d}\sigma + \frac{1}{\varepsilon}\iint_\sigma \frac{\partial\Phi}{\partial r}\mathrm{d}\sigma\right] . \tag{10.38}$$

The following process is analogous to that in Subsection 10.1.4. Since we assume functions Φ and $\partial\Phi/\partial r$ to be continuous, and the area of the sphere is $\sigma = 4\pi\varepsilon^2$, the second and third integrals in (10.38) will vanish in the limit for $\varepsilon \to 0$. The first integral yields

$$I = 4\pi\Phi(P) \ . \tag{10.39}$$

Equation (10.37) then yields

$$\boxed{\Phi(P) = \frac{1}{4\pi\alpha^2}\iiint_V Q\frac{e^{-ikr}}{r}\mathrm{d}V + \frac{1}{4\pi}\iint_S\left[\frac{e^{-ikr}}{r}\frac{\partial\Phi}{\partial n} - \Phi\frac{\partial}{\partial n}\left(\frac{e^{-ikr}}{r}\right)\right]\mathrm{d}S} \ . \tag{10.40}$$

This formula is called the *Helmholtz inner formula*, sometimes also the Kirchhoff formula for harmonic waves. Let us remind that function Φ is the spatial part of the longitudinal-wave potential; see Eqs. (10.30) and (10.31). Hence, the Helmholtz inner formula makes it possible to calculate the potential function Φ at a point P inside a closed surface S if the body forces are known inside S and functions Φ and $\partial\Phi/\partial n$ on S (see Fig. 10.3a). The unknown distribution of body forces outside S is compensated here by the knowledge of Φ and $\partial\Phi/\partial n$ on S.

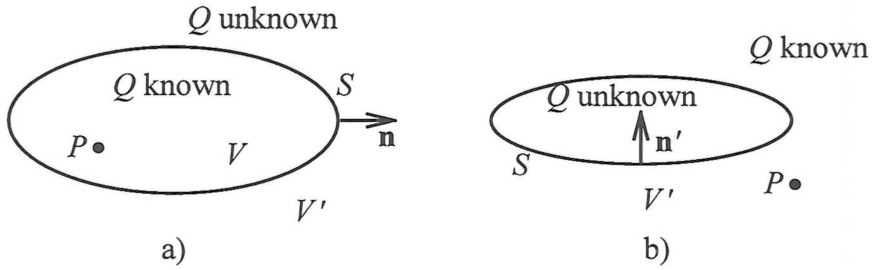


Fig. 10.3. Regions to which the Helmholtz formulae are applied: a) inner formula; b) outer formula.

10.3.2 Helmholtz outer formula

Now, consider point P to be located outside a closed surface S (Fig 10.3b). Let us apply Eq. (10.36) to the outer region. Analogously to the previous case, construct a spherical surface σ around point P, and let the radius of the sphere, ε, tend to zero. Moreover, construct another spherical surface, S_R, with its centre at P and of a radius R. Let radius R be large enough, so that the whole surface S lies inside S_R. Denote the volume between surface S and S_R by V'. Surface S is the inner boundary, and S_R the outer boundary of this volume. On the outer surface S_R we have $\partial/\partial n = \partial/\partial r$, where r is the distance from P. On the inner surface S we must take the inward normal, $\mathbf{n}'$, because this is the outer normal with respect to volume V'. Formula (10.40) will be preserved also in this case if V is replaced by V', normal $\mathbf{n}$ by $\mathbf{n}'$, and the corresponding integral over surface S_R is added, i.e. the integral

$$I_R = \frac{1}{4\pi}\iint_{S_R}\left[\frac{e^{-ikr}}{r}\frac{\partial\Phi}{\partial r} - \Phi\frac{\partial}{\partial r}\left(\frac{e^{-ikr}}{r}\right)\right]\mathrm{d}S = \frac{1}{4\pi}\iint_{S_R}\left[\left(\frac{\partial\Phi}{\partial r} + ik\Phi\right)\frac{1}{r} - \frac{\Phi}{r^2}\right]e^{-ikr}\,\mathrm{d}S \quad . \tag{10.41}$$

Since the area of sphere S_R is $4\pi R^2$, it can be seen that none of the terms in the latter square brackets vanishes generally when $R \to \infty$. However, these terms will vanish if we assume that the following two conditions hold true:

$$\lim_{r\to\infty} r\left(\frac{\partial \Phi}{\partial r} + ik\Phi\right) = 0, \quad |r\Phi| < M \ , \tag{10.42}$$

where M is a real constant. These two conditions play a fundamental role in the theory of propagation of harmonic waves in unlimited media. They are called the *Sommerfeld radiation conditions* or simply the radiation conditions. We shall return to their discussion at the end of this chapter.

Provided the Sommerfeld radiation conditions are valid, integral I_R will vanish for $R \to \infty$. From Eq. (10.40) we then arrive at the resultant formula for the case of point P outside S:

$$\Phi(P) = \frac{1}{4\pi\alpha^2}\iiint_{V'} Q\frac{e^{-ikr}}{r}\,\mathrm{d}V + \frac{1}{4\pi}\iint_S \left[\frac{e^{-ikr}}{r}\frac{\partial \Phi}{\partial n'} - \Phi\frac{\partial}{\partial n'}\left(\frac{e^{-ikr}}{r}\right)\right]\mathrm{d}S \ . \tag{10.43}$$

This formula is called the *Helmholtz outer formula*. Let us repeat that V' is the volume outside surface S, and $\mathbf{n}'$ is the inward normal to S; see Fig. 10.3b.

10.4 Kirchhoff Formulae

Let us rewrite the Helmholtz formulae into the time domain. Since they are formally identical (with the exception of V' and n'), we shall rewrite only the inner formula (10.40).

First, modify Eq. (10.40) by performing the differentiation in its last term:

$$\Phi(P) = \frac{1}{4\pi\alpha^2}\iiint_V Q\frac{e^{-ikr}}{r}\,\mathrm{d}V + \frac{1}{4\pi}\iint_S e^{-ikr}\left[\frac{1}{r}\frac{\partial \Phi}{\partial n} - \Phi\frac{\partial}{\partial n}\left(\frac{1}{r}\right) + \frac{ik\Phi}{r}\frac{\partial r}{\partial n}\right]\mathrm{d}S \ . \tag{10.44}$$

Multiply both sides of the latter equation by $\exp(i\omega t)$, and perform the Fourier integration with respect to ω from $(-\infty)$ to $(+\infty)$, ω being the angular frequency. In fact, the integration from 0 to $+\infty$ is sufficient as we assume the seismic signals to be real functions; see below. Before interpreting the results, let us repeat some basic formulae of the Fourier transform.

10.4.1 Basic formulae of the Fourier transform

Consider a function $f(t)$ satisfying, e.g., the following conditions (called the Dirichlet conditions):

1) The function is absolutely integrable, i.e. $\int_{-\infty}^{+\infty} |f(t)|\,\mathrm{d}t < M$.
2) Within each finite interval, the function has only a finite number of maxima, minima and discontinuities of the first order (the discontinuities where the limits from the left and right are finite).

Seismic signals usually satisfy these conditions. Under these conditions, the following integral exists:

$$F(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-i\omega t}\, \mathrm{d}t \ , \tag{10.45}$$

This function $F(\omega)$ is called the Fourier spectrum of function $f(t)$.

The inverse Fourier transform has the form

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{i\omega t}\, \mathrm{d}\omega \ . \tag{10.46}$$

If function $f(t)$ is real, which we shall assume here, its spectrum for negative ω is complex conjugate, $F(-\omega) = F^*(\omega)$, and the inverse Fourier transform may be expressed as

$$f(t) = \frac{1}{\pi} \mathrm{Re} \int_{0}^{+\infty} F(\omega) e^{i\omega t}\, \mathrm{d}\omega \ . \tag{10.47}$$

Two basic properties of the Fourier transform, which we shall need below, are given in Tab. 10.2.

Table 10.2. Some properties of the Fourier transform.

Function	Spectrum
$f(t)$	$F(\omega)$
$f(t-\tau)$	$e^{-i\omega\tau} F(\omega)$
$\dfrac{\mathrm{d} f(t)}{\mathrm{d}t}$	$i\omega F(\omega)$

10.4.2 Integration of the Helmholtz formulae

Let us go back to the integration of (10.44). According to Tab. 10.2, the multiplication of a spectrum by $\exp(-ikr) = \exp[-i\omega(r/\alpha)]$ corresponds to the time shift by r/α in the time domain, and the multiplication by $i\omega$ in the last term corresponds to the differentiation with respect to time (note that $ik = i\omega/\alpha$). Consequently, in view of relations (10.30), Eq. (10.44) can be expressed in the time domain as

$$\boxed{\varphi(P,t) = \frac{1}{4\pi\alpha^2} \iiint_V \frac{1}{r}[q]\,\mathrm{d}V + \frac{1}{4\pi} \iint_S \left\{ \frac{1}{r}\left[\frac{\partial \varphi}{\partial n}\right] - [\varphi]\frac{\partial}{\partial n}\left(\frac{1}{r}\right) + \frac{1}{\alpha r}\left[\frac{\partial \varphi}{\partial t}\right]\frac{\partial r}{\partial n} \right\} \mathrm{d}S} \ . \tag{10.48}$$

The quantities in the square brackets denote that time t, after carrying out the corresponding operation, has been replaced by $t - r/\alpha$. Thus,

$$[q] = q\left(x_i, t - \frac{r}{\alpha}\right), \quad \left[\frac{\partial \varphi}{\partial n}\right] = \left(\frac{\partial \varphi(x_i, \tau)}{\partial n}\right)_{\tau = t - r/\alpha},$$

$$[\varphi] = \varphi\left(x_i, t - \frac{r}{\alpha}\right), \quad \left[\frac{\partial \varphi}{\partial t}\right] = \left(\frac{\partial \varphi(x_i, \tau)}{\partial \tau}\right)_{\tau = t - r/\alpha}. \tag{10.49}$$

As the consequence of the time shift $t - r/\alpha$, the latter quantities in the square brackets are often called the retarded potentials. Hence, the potential at point P and time t is not determined by the quantities in volume V and on surface S at the same time t, but at some previous times. The corresponding time delay is equal to the travel time from the particular point to the point of observation, P.

Formula (10.48) represents the most common form of the so-called *Kirchhoff formulae*, which are widely used in methods of seismic prospecting, especially in migration methods.

When applying the Kirchhoff formulae we must take into account the important fact that these formulae are based on the wave equation and on the assumption of a constant velocity, α. In other words, these formulae have been derived under the assumption that the medium is homogeneous and isotropic. In inhomogeneous or anisotropic media, these formulae are not valid, but must be generalised. Instead of the Kirchhoff formulae, the so-called *elastodynamic representation theorems* are then obtained. Of course, such theorems have much more complicated form (Aki and Richards, 1980; Psencik, 1994).

10.5 Sommerfeld Radiation Conditions

Let us return to a brief discussion of the radiation conditions (10.42). The first condition, i.e.

$$\lim_{r \to \infty} r\left(\frac{\partial \Phi}{\partial r} + ik\Phi\right) = 0, \tag{10.50}$$

says that sources in a bounded region can generate only divergent waves outside this region, not convergent waves. Let us demonstrate this fact on the case of spherical waves. Consider a divergent spherical wave with potential

$$\varphi_d = \Phi e^{i\omega t} = \frac{1}{r} e^{i(\omega t - kr)}.$$

For this wave we get

$$r\left(\frac{\partial \Phi}{\partial r} + ik\Phi\right) = -\frac{1}{r} e^{-ikr},$$

which approaches zero as $1/r$ for $r \to \infty$. Contrary to it, for a convergent spherical wave,

$$\varphi_c = \widetilde{\Phi} e^{i\omega t} = \frac{1}{r} e^{i(\omega t + kr)} \ ,$$

we get

$$r\left(\frac{\partial \widetilde{\Phi}}{\partial r} + ik\widetilde{\Phi}\right) = \left(2ik - \frac{1}{r}\right) e^{ikr} \ ,$$

which does not approach zero for $r \to \infty$.

The second radiation condition, $\left| r\Phi \right| < M$, requires the amplitudes of the waves to decrease at least as $1/r$ for $r \to \infty$.

10.6 Advantages and Disadvantages of Solutions in Integral Forms

In Section 10.1 we found the solutions of Poisson's and Laplace's equations by means of integral representations which were based on Green's first theorem. A similar approach was used in Sections 10.3 and 10.4 to solve the inhomogeneous Helmholtz and wave equations, respectively. These solutions have the following advantages:

- From the formal point of view, the corresponding integral solutions are relatively simple and elegant.
- The boundary surface S may be of a general shape.

On the other hand, this method has the following disadvantages and limitations:

- The shape of surface S, over which the integration is performed, must be known (this is not satisfied, e.g., in some gravimetric problems).
- Both the potential and its normal derivative must be known on the boundary of the integration domain, S. This limits the applicability of the method. For example, the classical boundary problems for Laplace's equation, the Dirichlet and Neumann problems, cannot be solved immediately by this method, because only one of the above-mentioned functions is given on the boundary.
- Numerical integration, which must generally be used to evaluate the corresponding integrals, is not very convenient from the computational point of view. This integration may be rather time consuming, and the corresponding algorithms may be complicated.

Since these integral methods solve only a certain category of problems, also other methods must be considered.

Chapter 11

Solution of the Helmholtz Equation by the Separation of Variables

In the previous chapter we found the solutions of the corresponding equations in the integral forms which were based on Green's theorem. Another general method of solving partial differential equations is the Fourier method of separation of variables. In this method we attempt to solve a partial differential equation under the assumption that the solution may be expressed as a product of several functions, each of which depends on one coordinate only (or on a restricted number of coordinates). In this way we obtain particular solutions only. However, by the summation or integration of these particular solutions, we may obtain general solutions.

The separation of variables is frequently used when the solution is sought in an infinite region or in a region which is bounded by surfaces which coincide with coordinate surfaces of an orthogonal coordinate system.

In this chapter we shall apply the method of separation of variables to the Helmholtz equation in Cartesian, cylindrical and spherical coordinates. A brief review of the main formulae for Cartesian and cylindrical coordinates is given in Tab. 11.1.

We shall restrict ourselves to the propagation of longitudinal waves in a homogeneous isotropic medium. The corresponding potential φ satisfies the wave equation

$$\nabla^2 \varphi = \frac{1}{\alpha^2} \frac{\partial^2 \varphi}{\partial t^2} , \tag{11.1}$$

where α is the velocity of longitudinal waves. Consider only harmonic waves, i.e. the potential in the form

$$\varphi = \Phi e^{i\omega t} , \tag{11.2}$$

where ω is the angular frequency, and Φ is a function of coordinates only. The wave equation for longitudinal waves then simplifies to the following *Helmholtz equation*:

$$\boxed{\nabla^2 \Phi + k_\alpha^2 \Phi = 0} , \tag{11.3}$$

where $k_\alpha = \omega/\alpha$ is the wavenumber.

11.1 Separation of Variables in Cartesian Coordinates

Consider the Helmholtz equation in Cartesian coordinates (for longitudinal waves):

$$\frac{\partial^2 \Phi}{\partial x^2}+\frac{\partial^2 \Phi}{\partial y^2}+\frac{\partial^2 \Phi}{\partial z^2}+k_\alpha^2 \Phi=0 \ , \tag{11.4}$$

and solve it by the separation of variables. Hence, let us seek function Φ in the form

$$\Phi(x, y, z, \omega)=X(x)\, Y(y)\, Z(z) \ , \tag{11.5}$$

where function X depends on coordinate x only, Y on coordinate y only, and Z on coordinate z only. Moreover, functions X, Y, Z depend also on parameter ω and on some other constants of separation, which will be specified later. By inserting expression (11.5) into the Helmholtz equation (11.4), we get

$$YZ\frac{\mathrm{d}^2 X}{\mathrm{d}x^2}+XZ\frac{\mathrm{d}^2 Y}{\mathrm{d}y^2}+XY\frac{\mathrm{d}^2 Z}{\mathrm{d}z^2}+k_\alpha^2 XYZ=0 \ ; \tag{11.6}$$

note that here we could replace the partial derivatives by the normal derivatives, because each of functions X, Y, Z depends on one coordinate only. Dividing this equation by the product XYZ, one gets

$$\frac{1}{X}\frac{\mathrm{d}^2 X}{\mathrm{d}x^2}+\frac{1}{Y}\frac{\mathrm{d}^2 Y}{\mathrm{d}y^2}+\frac{1}{Z}\frac{\mathrm{d}^2 Z}{\mathrm{d}z^2}+k_\alpha^2=0 \ . \tag{11.7}$$

The first term of the latter equation is a function of x only, the second of y only, and the third of z only (k_α is a constant). Thus, we can introduce three constants of separation, k_x, k_y and k_z, in such a way that

$$\frac{1}{X}\frac{\mathrm{d}^2 X}{\mathrm{d}x^2}=-k_x^2 \ , \quad \frac{1}{Y}\frac{\mathrm{d}^2 Y}{\mathrm{d}y^2}=-k_y^2 \ , \quad \frac{1}{Z}\frac{\mathrm{d}^2 Z}{\mathrm{d}z^2}=-k_z^2 \ . \tag{11.8}$$

It follows from Eq. (11.7) that these constants must satisfy the condition

$$k_\alpha^2=k_x^2+k_y^2+k_z^2 \ , \tag{11.9}$$

and so only two of them are independent.

Instead of one partial differential equation (11.4) we have obtained three ordinary differential equations (11.8). The latter equations represent the equations of harmonic oscillators:

$$\frac{\mathrm{d}^2 X}{\mathrm{d}x^2}+k_x^2 X=0 \ , \quad \frac{\mathrm{d}^2 Y}{\mathrm{d}y^2}+k_y^2 Y=0 \ , \quad \frac{\mathrm{d}^2 Z}{\mathrm{d}z^2}+k_z^2 Z=0 \ . \tag{11.10}$$

Their solutions are very simple:

$$
\begin{aligned}
X(x,k_x) &= Ce^{ik_x x} + De^{-ik_x x} , \\
Y(y,k_y) &= Ee^{ik_y y} + Fe^{-ik_y y} , \\
Z(z,k_z) &= Ge^{ik_z z} + He^{-ik_z z} ,
\end{aligned}
\tag{11.11}
$$

where C to H are arbitrary constants.

As the independent constants, let us choose constants k_x and k_y. Constant k_z is then dependent on them:

$$
k_z = \sqrt{k_\alpha^2 - k_x^2 - k_y^2} = \sqrt{(\omega/\alpha)^2 - k_x^2 - k_y^2} . \tag{11.12}
$$

Formulae (11.5) and (11.11) yield particular solutions of the Helmholtz equation. The general solution of the Helmholtz equation (11.4) can be constructed as the superposition of all particular solutions over all possible values of k_x and k_y:

$$
\begin{aligned}
\Phi(x,y,z,\omega) = \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} A(k_x,k_y,\omega) e^{ik_x x + ik_y y + ik_z z} \, \mathrm{d}k_x \, \mathrm{d}k_y \, + \\
+ \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} B(k_x,k_y,\omega) e^{ik_x x + ik_y y - ik_z z} \, \mathrm{d}k_x \, \mathrm{d}k_y ;
\end{aligned}
\tag{11.13}
$$

the constants $(2\pi)^{-2}$ could be included into functions A and B, but we write them in front of the integrals in order to obtain expressions which are similar to the two-dimensional inverse Fourier transform.

The latter integrals play an important role in some procedures in seismic prospecting. The following properties of these integrals should be pointed out:

1) Integrals (11.13) represent the expansion of a harmonic wavefield into elementary plane waves. If the z-axis is oriented downwards (z = depth), then the first integral, after multiplying by $\exp(i\omega t)$, will be composed of plane waves propagating obliquely upwards, and the second integral of plane waves propagating obliquely downwards.
2) As mentioned above, the integral representation (11.13) has the form of the $2-\mathrm{D}$ Fourier integral, where x, y are the original variables, and k_x, k_y are the transform variables. Note that k_z is not a transform variable, as it is related to k_x and k_y by (11.12).
3) The signs with $i\omega t$, $ik_x x$ and $ik_y y$ are arbitrary. We have chosen them "+", but they may be chosen "–", or mixed, e.g. $i\omega t - ik_x x - ik_y y$. Since the corresponding integrals run from $-\infty$ to $+\infty$, all negative and positive values are always included.

In Sections 11.2 and 11.3 we shall describe two important applications of the integral representation (11.13).

TABLE 11.1. SOLUTION OF THE HELMHOLTZ EQUATION BY THE SEPARATION OF VARIABLES

CARTESIAN COORDINATES (x, y, z)

Helmholtz equation:

$$\frac{\partial^2 \Phi}{\partial x^2}+\frac{\partial^2 \Phi}{\partial y^2}+\frac{\partial^2 \Phi}{\partial z^2}+k_\alpha^2 \Phi=0$$

Form of the solution required:

$$\Phi(x, y, z)=X(x)Y(y)Z(z)$$

Special form of the Helmholtz equation:

$$\underbrace{\frac{1}{X}\frac{\mathrm{d}^2 X}{\mathrm{d} x^2}}_{-k_x^2}+\underbrace{\frac{1}{Y}\frac{\mathrm{d}^2 Y}{\mathrm{d} y^2}}_{-k_y^2}+\underbrace{\frac{1}{Z}\frac{\mathrm{d}^2 Z}{\mathrm{d} z^2}}_{-k_z^2}+k_\alpha^2=0$$

$$\left(k_z=\sqrt{k_\alpha^2-k_x^2-k_y^2}\right)$$

Ordinary differential equations:

$$\frac{\mathrm{d}^2 X}{\mathrm{d} x^2}+k_x^2 X=0, \text{ etc.}$$

Their solutions:

$$X(x)=Ce^{ik_x x}+De^{-ik_x x}$$

$$Y(y)=Ee^{ik_y y}+Fe^{-ik_y y}$$

$$Z(z)=Ge^{ik_z z}+He^{-ik_z z}$$

General solution of the Helmholtz equation:

$$\Phi(x, y, z)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} A(k_x, k_y)e^{ik_x x+ik_y y+ik_z z}\,\mathrm{d}k_x\,\mathrm{d}k_y+$$

$$+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} B(k_x, k_y)e^{ik_x x+ik_y y-ik_z z}\,\mathrm{d}k_x\,\mathrm{d}k_y$$

CYLINDRICAL COORDINATES (ρ, λ, z)

$$\frac{\partial^2 \Phi}{\partial \rho^2}+\frac{1}{\rho}\frac{\partial \Phi}{\partial \rho}+\frac{1}{\rho^2}\frac{\partial^2 \Phi}{\partial \lambda^2}+\frac{\partial^2 \Phi}{\partial z^2}+k_\alpha^2 \Phi=0$$

$$\Phi(\rho, \lambda, z)=R(\rho)\Lambda(\lambda)Z(z)$$

$$\underbrace{\frac{1}{R}\left(\frac{\mathrm{d}^2 R}{\mathrm{d}\rho^2}+\frac{1}{\rho}\frac{\mathrm{d}R}{\mathrm{d}\rho}\right)+\frac{1}{\rho^2}\underbrace{\frac{1}{\Lambda}\frac{\mathrm{d}^2 \Lambda}{\mathrm{d}\lambda^2}}_{-l^2}}_{-k^2}+\underbrace{\frac{1}{Z}\frac{\mathrm{d}^2 Z}{\mathrm{d} z^2}+k_\alpha^2}_{k^2}=0$$

$$\left(k_z=\sqrt{k_\alpha^2-k^2}\right)$$

$$\frac{\mathrm{d}^2 \Lambda}{\mathrm{d}\lambda^2}+l^2\Lambda=0, \text{ etc.}$$

$$\Lambda(\lambda)=C\cos l\lambda+D\sin l\lambda=C_1e^{il\lambda}+C_2e^{-il\lambda}$$

$R(\rho)=J_l(k\rho)\ldots$ Bessel function of the l – th order

$$Z(z)=Ge^{ik_z z}+He^{-ik_z z}$$

$$\Phi(\rho, \lambda, z)=\int_0^\infty\int_0^{2\pi} J_l(k\rho)\left[A(k, l)e^{ik_z z}+B(k, l)e^{-ik_z z}\right]\begin{Bmatrix}\cos l\lambda\\ \sin l\lambda\end{Bmatrix}\mathrm{d}k\,\mathrm{d}l$$

11.2 The ω - k Migration

Integral expression (11.13) is frequently used in the downward and upward continuations of the wavefield. Since the downward and upward continuations form an important part of a migration procedure, we speak of the $\omega - k$ migration (if frequency $f = \omega/(2\pi)$ is used). First, let us describe the migration in the frequency domain.

Assume, for example, that the spectrum of a measured wavefield at the Earth's surface $(z = 0)$ corresponds to a wave propagating upwards, i.e. from below. This may be, e.g., a wavefield reflected from some discontinuity. We then have to put $B(k_x, k_y, \omega) = 0$, because this term represents downgoing waves. Formula (11.13) then becomes

$$\Phi(x, y, z, \omega) = \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} A(k_x, k_y, \omega) e^{ik_x x + ik_y y + ik_z z} \, \mathrm{d}k_x \, \mathrm{d}k_y \, . \quad (11.14)$$

For the spectrum of the recorded field at $z = 0$ we thus get

$$\Phi(x, y, 0, \omega) = \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} A(k_x, k_y, \omega) e^{ik_x x + ik_y y} \, \mathrm{d}k_x \, \mathrm{d}k_y \, . \quad (11.15)$$

We have arrived at a formula which appears in the theory of the two-dimensional Fourier transform. Let as remind the basic formulae of this transform.

Consider a function of two Cartesian variables, $f(x, y)$, satisfying certain general conditions (seismic signals usually satisfy these conditions).The $2 - \mathrm{D}$ Fourier spectrum of this function is defined by

$$F(k_x, k_y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) e^{-ik_x x - ik_y y} \, \mathrm{d}x \, \mathrm{d}y \, . \quad (11.16)$$

The inverse transform is of the form

$$f(x, y) = \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(k_x, k_y) e^{ik_x x + ik_y y} \, \mathrm{d}k_x \, \mathrm{d}k_y \, . \quad (11.17)$$

It follows from the comparison of formulae (11.17) and (11.15) that function A is the $2 - \mathrm{D}$ spectrum of function Φ. The definition (11.16) then yields

$$A(k_x, k_y, \omega) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Phi(x, y, 0, \omega) e^{-ik_x x - ik_y y} \, \mathrm{d}x \, \mathrm{d}y \, . \quad (11.18))$$

Since the function $\Phi(x,y,0,\omega)$ at the Earth's surface is known, we can use (11.18) to determine $A(k_x,k_y,\omega)$. Inserting this function into (11.14) we obtain the integral expression for the downward continuation of the wavefield spectrum.

In the time domain, we must add the corresponding Fourier transform for computing the spectrum at the surface,

$$\Phi(x,y,0,\omega) = \int_{-\infty}^{+\infty} \varphi(x,y,0,t)\, e^{-i\omega t}\, \mathrm{d}t \ , \tag{11.19}$$

and the inverse transform for computing the time signal at depth z,

$$\varphi(x,y,z,t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Phi(x,y,z,\omega)\, e^{i\omega t}\, \mathrm{d}\omega \ . \tag{11.20}$$

The sequence of the computations is, therefore, as follows. From the wavefield at the surface, $\varphi(x,y,0,t)$, we compute the corresponding spectrum (11.19), the $2-\mathrm{D}$ spectrum (11.18), then spectral amplitude (11.14), and finally (11.20), which yields the desirable downward continuation of the wave field.

11.3 Expansion of a Spherical Wave into Plane Waves. The Weyl Integral

Using the integral expression for the wavefield derived in Section 11.1, we can also derive an expansion formula of a spherical wave into plane waves.

Consider a spherical harmonic longitudinal wave propagating in a homogeneous medium at a constant velocity α (for transverse waves we would take velocity β):

$$\varphi(r,t,\omega) = \frac{1}{r} e^{i\omega(t-r/\alpha)} \ , \tag{11.21}$$

see Eq. (9.51) where we have put $A=1$. Denote the expression without the time term by

$$\Phi(r,\omega) = \frac{1}{r} e^{-ik_\alpha r} \ , \tag{11.22}$$

where $k_\alpha = \omega/\alpha$, and r is the distance of the observation point from the point source. To find the expansion formula, we shall use the general expression (11.13), where k_z is given by (11.12).

Now we shall consider a point source at the origin of the coordinate system, $x=y=z=0$, and the receiver in the region $z \geq 0$. As the z-axis is oriented downwards, the wave propagating in the half-space $z > 0$ due to a point source at $z = 0$ is downgoing and, consequently, we shall put $A(k_x, k_y, \omega) = 0$. Thus,

$$\Phi(x, y, z, \omega) = \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} B(k_x, k_y, \omega) e^{ik_x x + ik_y y - ik_z z} \, \mathrm{d}k_x \, \mathrm{d}k_y \ . \tag{11.23}$$

For $z = 0$, we get

$$\Phi(x, y, 0, \omega) = \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} B(k_x, k_y, \omega) e^{ik_x x + ik_y y} \, \mathrm{d}k_x \, \mathrm{d}k_y \ . \tag{11.24}$$

This expression has again the form of a $2-\mathrm{D}$ inverse Fourier transform. Consequently, function B can be expressed as

$$B(k_x, k_y, \omega) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Phi(x, y, 0, \omega) e^{-ik_x x - ik_y y} \, \mathrm{d}x \, \mathrm{d}y \ . \tag{11.25}$$

Function Φ in the latter integral is defined by (11.22), where $r = \sqrt{x^2 + y^2 + z^2}$. For $z = 0$, this function becomes

$$\Phi(x, y, 0, \omega) = \frac{1}{\sqrt{x^2 + y^2}} e^{-ik_\alpha \sqrt{x^2 + y^2}} \ . \tag{11.26}$$

The corresponding function $B(k_x, k_y, \omega)$, given by (11.25), now takes the form

$$B(k_x, k_y, \omega) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{x^2 + y^2}} e^{-ik_\alpha \sqrt{x^2 + y^2} - ik_x x - ik_y y} \, \mathrm{d}x \, \mathrm{d}y \ . \tag{11.27}$$

Formulae (11.27) and (11.23) solve formally the problem of the decomposition of a spherical wave into plane waves, but it is desirable to find a simpler expression for $B(k_x, k_y, \omega)$.

Introduce polar coordinates as follows:

$$x = \rho \cos \lambda \ , \quad y = \rho \sin \lambda \ , \tag{11.28}$$
$$k_x = k \cos \psi \ , \quad k_y = k \sin \psi \ , \quad k^2 = k_x^2 + k_y^2 \ .$$

Since the surface element in polar coordinates is $\mathrm{d}S = \rho \, \mathrm{d}\rho \, \mathrm{d}\lambda$, integral (11.27) takes the form

$$B(k_x,k_y,\omega)=\int_0^{\infty}\int_0^{2\pi} e^{-i\rho\left[k_\alpha+k\cos(\lambda-\psi)\right]}\,\mathrm{d}\rho\,\mathrm{d}\lambda\ . \tag{11.29}$$

First, we shall perform the integration with respect to ρ:

$$B(k_x,k_y,\omega)=\int_0^{2\pi}\mathrm{d}\lambda\left[\frac{e^{-i\rho\left[k_\alpha+k\cos(\lambda-\psi)\right]}}{-i\left[k_\alpha+k\cos(\lambda-\psi)\right]}\right]_{\rho=0}^{\infty}\ . \tag{11.30}$$

However, the integrand in the latter formula does not tend to a limit for $\rho\to\infty$ if the expression $k_\alpha+k\cos(\lambda-\psi)$ is real, because the exponential is oscillating. To avoid this problem, the following approaches may be used:

1) We may assume that the medium is very slightly absorbing, i.e. velocity α is complex-valued and $1/\alpha$ has a small negative imaginary part. The integrand in (11.30) then vanishes for $\rho\to\infty$, and we get

$$B(k_x,k_y,\omega)=\frac{1}{i}\int_0^{2\pi}\frac{\mathrm{d}\lambda}{k_\alpha+k\cos(\lambda-\psi)}=-i\int_0^{2\pi}\frac{\mathrm{d}\eta}{k_\alpha+k\cos\eta}\ . \tag{11.31}$$

2) For a pure real α, the Cauchy theorem may be applied to transform the path of integration into the complex plane. If $k_\alpha+k\cos(\lambda-\psi)$ is positive, the path of integration is transformed from the real axis to the infinite arc and imaginary axis of the fourth quadrant. The contribution from the infinite arc vanishes, and the integration along the negative part of the imaginary axis yields again formula (11.31). If $k_\alpha+k\cos(\lambda-\psi)$ is negative, we shall transform the path of integration into the first quadrant.

Integral (11.31) can be found in tables of integrals (or, e.g., the substitution $\xi=\cos(\eta/2)$ can be used to transform the integrand into a rational function). One gets

$$\int_0^{2\pi}\frac{\mathrm{d}\eta}{k_\alpha+k\cos\eta}=\frac{2\pi}{\sqrt{k_\alpha^2-k^2}}\ . \tag{11.32}$$

Further details of this derivation can be found in the book by Tygel and Hubral (1987) and in the lecture notes by Psencik (1994). Taking into account that

$$k_z=\sqrt{(\omega/\alpha)^2-k_x^2-k_y^2}=\sqrt{k_\alpha^2-k_x^2-k_y^2}\ , \tag{11.33}$$

we arrive at

$$B(k_x,k_y,\omega)=-2\pi i/k_z\ . \tag{11.34}$$

Thus, formulae (11.22) and (11.23) finally yield

$$\frac{e^{-ik_\alpha r}}{r} = \frac{1}{2\pi i} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{k_z} e^{-ik_x x - ik_y y - ik_z z} \, \mathrm{d}k_x \, \mathrm{d}k_y \ , \tag{11.35}$$

where we have taken the opposite signs in the terms $ik_x x$ and $ik_y y$ in the exponents. In fact, we have replaced k_x by $(-k_x)$ and k_y by $(-k_y)$, which does not change the value of the integral. Formula (11.35) represents the expansion of a spherical wave into plane waves. This is known as the *Weyl integral.*

Since k_z may attain real or pure imaginary values, the superposition (11.35) contains two types of plane waves, namely the so-called homogeneous and inhomogeneous waves. We shall discuss them in connection with the Sommerfeld integral in Section 11.5.

11.4 Separation of Variables in Cylindrical Coordinates

Consider the Helmholtz equation for longitudinal waves, i.e. Eq. (11.3), in cylindrical coordinates ρ, λ, z. Using formula (9.31) for the Laplacian in cylindrical coordinates, we get

$$\frac{\partial^2 \Phi}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial \Phi}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2 \Phi}{\partial \lambda^2} + \frac{\partial^2 \Phi}{\partial z^2} + k_\alpha^2 \Phi = 0 \ . \tag{11.36}$$

A general solution of this equation is briefly described in Tab. 11.1. Here we shall restrict ourselves only to the solutions which are independent of coordinate λ, i.e. to the solutions which are axially symmetrical with respect to the z-axis. The Helmholtz equation then simplifies to read

$$\frac{\partial^2 \Phi}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial \Phi}{\partial \rho} + \frac{\partial^2 \Phi}{\partial z^2} + k_\alpha^2 \Phi = 0 \ . \tag{11.37}$$

Let us solve the latter equation by the separation of variables. We shall seek the solution in the form of the product

$$\Phi(\rho, z) = R(\rho) Z(z) \ , \tag{11.38}$$

where function R depends on ρ only and function Z on z only. Moreover, all the functions depend also on the angular frequency ω and on a certain parameter of separation. Insert this form of Φ into Eq. (11.37), and divide the equation by RZ. We get

$$\underbrace{\frac{1}{R}\left(\frac{\mathrm{d}^2 R}{\mathrm{d}\rho^2} + \frac{1}{\rho}\frac{\mathrm{d}R}{\mathrm{d}\rho}\right)}_{-k^2} + \underbrace{\frac{1}{Z}\frac{\mathrm{d}^2 Z}{\mathrm{d}z^2}}_{k^2} + k_\alpha^2 = 0 \ . \tag{11.39}$$

Introduce a separation constant k^2 as indicated. The solution of the partial differential equation (11.37) then reduces to two ordinary differential equations:

$$\frac{d^2 R}{d\rho^2} + \frac{1}{\rho}\frac{dR}{d\rho} + k^2 R = 0 \ , \tag{11.40}$$

$$\frac{d^2 Z}{dz^2} - \left(k^2 - k_\alpha^2\right) Z = 0 \ . \tag{11.41}$$

The first equation, according to the analogy with Eq. (9.54), has a particular solution

$$R = J_0(k\rho) \ , \tag{11.42}$$

and the second equation has a particular solution

$$Z = A(k)e^{\nu z} + B(k)e^{-\nu z} \ , \tag{11.43}$$

where

$$\nu = \sqrt{k^2 - k_\alpha^2} \ . \tag{11.44}$$

The general solution of the Helmholtz equation (11.37) can then be expressed as

$$\Phi(\rho, z, \omega) = \int_0^\infty J_0(k\rho)\left\{A(k,\omega)e^{\nu z} + B(k,\omega)e^{-\nu z}\right\} dk \ , \tag{11.45}$$

where $A(k, \omega)$ and $B(k, \omega)$ are arbitrary functions. They must be determined from boundary conditions. Note that we could introduce $\nu' = \sqrt{k_\alpha^2 - k^2}$, so that $\nu = i\nu'$ and the exponentials in (11.45) would be replaced by $\exp(i\nu' z)$ and $\exp(-i\nu' z)$, respectively.

11.5 The Sommerfeld Integral

The Weyl integral represents the expansion of a spherical wave into plane waves. Similarly, we can expand a spherical wave into cylindrical waves. This expansion into cylindrical waves is given by the Sommerfeld integral. In applications, the Sommerfeld integral has been used even more frequently than the Weyl integral.

In order to derive the Sommerfeld integral, the general solution (11.45) in cylindrical coordinates is usually used. We can again consider a point source at the origin of the coordinate system, $\rho = z = 0$, and the receiver in the region $z \geq 0$. For large k the square root $\sqrt{k^2 - k_\alpha^2}$ is real. In order to obtain a limited integrand for $z > 0$, we must put $A(k, \omega) = 0$. Analogously to the Fourier transform, which we used to determine the function $B(k_x, k_y, \omega)$ in Section 11.3, now we could determine the function $B(k,\omega)$ by means of the Fourier-Bessel transform. The following procedure would be very similar to that used in Section 11.3 in deriving the Weyl integral. However, we shall not describe this derivation here. Since we have already derived the Weyl integral, it will be simpler to use the known form of this integral, and transform it into cylindrical coodinates.

Therefore, introduce polar coordinates (11.28) into the Weyl integral (11.35) and express k_z in the form (11.33):

$$\frac{e^{-ik_\alpha r}}{r} = \frac{1}{2\pi i}\int_0^\infty \int_0^{2\pi} \frac{1}{\sqrt{k_\alpha^2 - k^2}} e^{-ik\rho\cos(\psi-\lambda) - iz\sqrt{k_\alpha^2 - k^2}}\, k\,\mathrm{d}k\,\mathrm{d}\psi \ . \qquad (11.46)$$

Using the integral definition (9.58) of the Bessel function J_0, i.e.

$$J_0(x) = \frac{1}{2\pi}\int_0^{2\pi} e^{\pm ix\cos(\psi-\chi)}\,\mathrm{d}\psi \ , \qquad (11.47)$$

the Weyl integral can be expressed as

$$\frac{e^{-ik_\alpha r}}{r} = \frac{1}{i}\int_0^\infty J_0(k\rho)\frac{1}{\sqrt{k_\alpha^2 - k^2}} e^{-iz\sqrt{k_\alpha^2 - k^2}}\, k\,\mathrm{d}k \ . \qquad (11.48)$$

We have performed the derivation for $z \geq 0$. Since the derivation for $z \leq 0$ would be analogous, we shall replace z by $|z|$ in the latter formula. Moreover, let us introduce

$$\nu = \sqrt{k^2 - k_\alpha^2} = i\sqrt{k_\alpha^2 - k^2} \ . \qquad (11.49)$$

We then arrive at the final expression for a spherical wave in the form

$$\frac{1}{r}e^{-ik_\alpha r} = \int_0^\infty J_0(k\rho)\, e^{-\nu|z|}\,\frac{k\,\mathrm{d}k}{\nu} \ . \qquad (11.50)$$

This is the famous *Sommerfeld integral*. Let us repeat that r is the spherical coordinate, ρ and z are cylindrical coordinates, $r = \sqrt{\rho^2 + z^2}$, J_0 the zero-order Bessel function, k the integration variable, ν is given by (11.49), and $k_\alpha = \omega/\alpha$.

We have already mentioned that the superpositions given by the Weyl and Sommerfeld integrals contain two types of waves. For $k < k_\alpha$ the exponential in (11.50) is oscillatory, which corresponds to usual waves, called homogeneous waves. However, for $k > k_\alpha$, the waves decay exponentially with increasing $|z|$. These latter waves are called inhomogeneous waves. The superposition of homogeneous waves alone would give only a finite value of the wavefield at the source of spherical waves. The presence of inhomogeneous waves produces the necessary singularity at the source.

Hence, we have found that inhomogeneous waves constitute an important part of spherical waves. We shall see in the next chapter that inhomogeneous waves are also produced from homogeneous plane waves at the total reflection.

11.6 Separation of Variables in Spherical Coordinates

In spherical coordinates, the Helmholtz equation (11.3) reads

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\Phi}{\partial r}\right)+\frac{1}{r^2\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial\Phi}{\partial\vartheta}\right)+\frac{1}{r^2\sin^2\vartheta}\frac{\partial^2\Phi}{\partial\lambda^2}+k_\alpha^2\Phi=0\ ;\tag{11.51}$$

see formula (9.29) for the Laplacian. We shall seek a separated solution in the form

$$\Phi(r,\vartheta,\lambda)=R(r)\,\Theta(\vartheta)\,\Lambda(\lambda)\ .\tag{11.52}$$

Inserting this form into Eq. (11.51) and multiplying by r^2/Φ yields

$$\frac{1}{R}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right)+k_\alpha^2r^2+\frac{1}{\Theta\sin\vartheta}\frac{\mathrm{d}}{\mathrm{d}\vartheta}\left(\sin\vartheta\frac{\mathrm{d}\Theta}{\mathrm{d}\vartheta}\right)+\frac{1}{\Lambda\sin^2\vartheta}\frac{\mathrm{d}^2\Lambda}{\mathrm{d}\lambda^2}=0\ .\tag{11.53}$$

Since coordinate λ is contained only in the term $\Lambda^{-1}\left(\mathrm{d}^2\Lambda/\mathrm{d}\lambda^2\right)$, this term must be constant. Denoting this constant by $\left(-m^2\right)$, we arrive at the first ordinary differential equation

$$\frac{\mathrm{d}^2\Lambda}{\mathrm{d}\lambda^2}+m^2\Lambda=0\ .\tag{11.54}$$

Its particular solution is

$$\Lambda(\lambda)=Ae^{im\lambda}+Be^{-im\lambda}\ ,\tag{11.55}$$

A and B being constants. From the natural requirement on function Λ that it must be periodic,

$$\Lambda(\lambda+2\pi)=\Lambda(\lambda)\ ,\tag{11.56}$$

we can deduce that m must be an integer.

Introduce a new variable $\mu=\cos\vartheta$ instead of ϑ. Since $\mathrm{d}\mu=-\sin\vartheta\,\mathrm{d}\vartheta$, equation (11.53) may now be expressed as

$$\frac{1}{R}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right)+k_\alpha^2 r^2+\frac{1}{\Theta}\frac{\mathrm{d}}{\mathrm{d}\mu}\left[\left(1-\mu^2\right)\frac{\mathrm{d}\Theta}{\mathrm{d}\mu}\right]-\frac{m^2}{1-\mu^2}=0\ . \tag{11.57}$$

It is now evident that the first two terms of this equation must be equal to a constant, K, and the remaining two terms must be equal to the constant $(-K)$. If we put $K=l(l+1)$, l being an integer, the remaining ordinary differential equations take the form

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right)+\left[k_\alpha^2 r^2-l(l+1)\right]R=0\ , \tag{11.58}$$

$$\frac{\mathrm{d}}{\mathrm{d}\mu}\left[\left(1-\mu^2\right)\frac{\mathrm{d}\Theta}{\mathrm{d}\mu}\right]+\left[l(l+1)-\frac{m^2}{1-\mu^2}\right]\Theta=0\ . \tag{11.59}$$

The latter equation is well known, e.g., from gravimetry. Its solutions are *associated Legendre functions*,

$$\Theta(\mu)=P_l^m(\mu)=P_l^m(\cos\vartheta)\ . \tag{11.60}$$

Note that these functions are usually defined by the relation

$$P_n^m(x)=\left(1-x^2\right)^{m/2}\frac{\mathrm{d}^m P_n(x)}{\mathrm{d}x^m}\ , \tag{11.61}$$

where

$$P_n(x)=\frac{1}{2^n n!}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left[\left(x^2-1\right)^n\right] \tag{11.62}$$

are Legendre polynomials. The products $P_l^m(\cos\vartheta)e^{im\lambda}$ and $P_l^m(\cos\vartheta)e^{-im\lambda}$ are called the spherical surface harmonic functions.

Let us modify Eq. (11.58) by introducing a new variable $\zeta=k_\alpha r$. We obtain

$$\zeta^2\frac{\mathrm{d}^2 R}{\mathrm{d}\zeta^2}+2\zeta\frac{\mathrm{d}R}{\mathrm{d}\zeta}+\left[\zeta^2-l(l+1)\right]R=0 \tag{11.63a}$$

or

$$\frac{\mathrm{d}^2 R}{\mathrm{d}\zeta^2}+\frac{2}{\zeta}\frac{\mathrm{d}R}{\mathrm{d}\zeta}+\left[1-\frac{l(l+1)}{\zeta^2}\right]R=0\ . \tag{11.63b}$$

This is again a well-known differential equation in mathematical physics. Its solutions are *spherical Bessel functions*:

$$j_l(\zeta) = \sqrt{\frac{\pi}{2\zeta}} J_{l+1/2}(\zeta), \quad n_l(\zeta) = \sqrt{\frac{\pi}{2\zeta}} N_{l+1/2}(\zeta) . \tag{11.64}$$

Another possible pair of linearly independent solutions is as follows:

$$h_l^{(1)}(\zeta) = \sqrt{\frac{\pi}{2\zeta}} H_{l+1/2}^{(1)}(\zeta), \quad h_l^{(2)}(\zeta) = \sqrt{\frac{\pi}{2\zeta}} H_{l+1/2}^{(2)}(\zeta) . \tag{11.65}$$

Here J_n and N_n are the Bessel functions of the first and second kind (N_n is also called the Neumann function), $H_n^{(1)}$ and $H_n^{(2)}$ are the Hankel functions of the first and second kind.

Thus, the elementary solutions of the Helmholtz equation (11.3) in spherical coordinates may be expressed as the following combinations:

$$\begin{Bmatrix} j_l(k_\alpha r) \\ n_l(k_\alpha r) \end{Bmatrix} P_l^m(\cos\vartheta) \begin{Bmatrix} e^{im\lambda} \\ e^{-im\lambda} \end{Bmatrix} . \tag{11.66}$$

Such solution are widely used in seismology (e.g., in the theory of the free oscillations of the Earth). The constants of separation are l and m. The sum over l and m is again a solution of the Helmholtz equation.

For further details we refer the reader to the lecture notes by Pšencík (1994).

Chapter 12

Reflection and Transmission of Plane Elastic Waves at a Plane Interface

12.1 Reflection and Transmission as a Special Problem of Wave Propagation

In the preceding chapters we investigated waves which propagated in homogeneous media. If a wave strikes a boundary or interface, new conditions, so-called *boundary conditions*, must be added to the differential equations of motion. These boundary conditions must be determined from observations or from the integral equations of motion, which are more general than the differential equations of motion.

From the mathematical point of view, the problem consists in the following. In Chapter 6 we derived the differential equations of motion under the assumptions that the displacements and stresses were continuous together with their first partial derivatives. Moreover, we assumed the validity of Hooke's law. However, these requirements are in contradiction at places where the elastic coefficients are discontinuous. For example, assuming the first derivatives of the displacement to be continuous, and applying Hooke's law at a material discontinuity, we obtain discontinuous stresses. Or conversely, assuming continuous stresses at such discontinuities and the validity of Hooke's law, we must admit discontinuous derivatives of the displacement. Therefore, some of the continuity conditions (or the validity of Hooke's law) must be abandoned at the discontinuities. We shall usually keep the continuity of the displacement and stress, but abandon the continuity of their derivatives. However, this means that the differential equations of motion do not hold at discontinuities of elastic parameters. Consequently, we can solve these equations in the regions without discontinuities, but at a discontinuity we must sew the particular solutions together on the basis of boundary conditions.

It is well-established fact that a disturbance of any kind propagating in one medium and impinging upon an interface gives rise, in general, to reflected and refracted (transmitted) waves. Since the term "refracted wave" is frequently used to denote another type of waves (see Chapters 3 and 4), here we shall prefer the term "transmitted wave". Therefore, we shall speak of the reflection and transmission of waves at interfaces. The problem of the reflection and transmission of waves belongs to the basic problems in all branches of physics which deal with wave phenomena.

Knott (1899) seems to have been the first to derive the general equations for the reflection and transmission of elastic waves at plane interfaces. Another formulation of this problem was developed by Zoeppritz in 1907, but partly because of his death in 1908, his paper was not published until 1919; see Zoeppritz (1919). The Zoeppritz equations then became very popular and have been frequently used.

12.2 Model of the Medium and Boundary Conditions

Consider a medium which consist of two homogeneous and isotropic half-spaces in contact. Assume the media to be perfectly elastic and the contact to be welded, i.e. the displacement is continuous across the interface. Denote by α_1, β_1 and ρ_1 the compressional wave velocity, shear wave velocity and density in the first half-space, respectively, and by α_2, β_2 and ρ_2 the corresponding parameters in the second half-space.

Consider an incident plane wave propagating in the first half-space. The point where a selected ray strikes the interface, choose as the origin of a Cartesian coordinate system (x, y, z). Let the x- and y-axes be in the plane of the interface, and the z-axis be perpendicular to this interface and directed into the second half-space (Fig. 12.1). The incident ray and the z-axis determine a plane which is called the plane of incidence. Choose the x-axis in this plane.

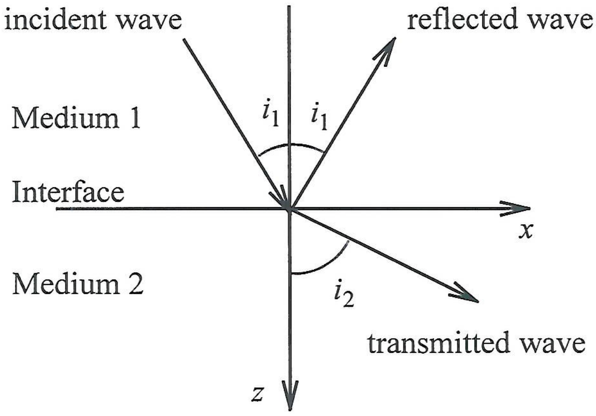


Fig. 12.1. The coordinate system and seismic rays.

As usual, denote the displacement vector by $\mathbf{u} = (u, v, w)$ and the stress tensor by τ (with its components τ_{xx}, τ_{xy}, etc.). The following subscripts will be attached to these quantities in order to distinguish the individual wavefields:

- no subscript or subscript 0 for the incident wave;
- R for the reflected wave;
- T for the transmitted wave;
- 1 for the resultant wavefield in the first half-space;
- 2 for the resultant wavefield in the second half-space.

At the interface we shall generally require the followings boundary conditions to be satisfied:

1) the displacement vector to be continuous across the interface, i.e.

$$\mathbf{u}_1 = \mathbf{u}_2 \quad \text{for} \quad z = 0 \; ; \tag{12.1}$$

2) the stresses acting at the interface (at the plane $z = 0$) to be continuous:

$$(\tau_{zx})_1 = (\tau_{zx})_2 \, , \quad (\tau_{zy})_1 = (\tau_{zy})_2 \, , \quad (\tau_{zz})_1 = (\tau_{zz})_2 \quad \text{for} \quad z = 0 \; . \tag{12.2}$$

12.3 Reflection and Transmission of *SH* Waves

Consider a plane harmonic wave of the *SH* type which is impinging on the interface from the first half-space (Fig. 12.1). Denote its angular frequency by ω and the angle of incidence by i_1. Without loss of generality we may put its amplitude to unity. Therefore, let the displacement in the *incident wave* be

$$\mathbf{u} = (0, v, 0) \,, \tag{12.3}$$

where

$$v = v(x,z,t) = \exp\left[i\omega \left(t - \frac{x \sin i_1 + z \cos i_1}{\beta_1} \right) \right] . \tag{12.4}$$

12.3.1 Expected forms of the reflected and transmitted waves

In order to satisfy the boundary conditions we shall assume that, in agreement with observations, an incident *SH* wave gives rise to a reflected wave and a transmitted wave of the same *SH* type, and that the rays of the reflected and transmitted waves remain in the plane of incidence. If we are able to satisfy the wave equations and all boundary conditions, it will confirm our assumption that such waves really exist.

Therefore, we shall introduce two new waves, generated at the interface:

1) a reflected plane harmonic wave, $\mathbf{u}_R = (0, v_R, 0)$, where

$$v_R = A \exp\left[i\omega_R \left(t - \frac{x \sin_1^R - z \cos_1^R}{\beta_1} \right) \right] ; \tag{12.5}$$

2) a transmitted plane harmonic wave, $\mathbf{u}_T = (0, v_T, 0)$, where

$$v_T = B \exp\left[i\omega_T \left(t - \frac{x \sin_2 + z \cos_2}{\beta_2} \right) \right] . \tag{12.6}$$

In expressions (12.5) and (12.6) for the reflected and transmitted waves, we have denoted the angular frequencies by ω_R and ω_T, the angles of reflection and transmission by i_1^R and i_2, and the amplitudes by A and B, respectively. All these quantities must be determined. Remind that the *reflection coefficient* is defined by the ratio of the amplitudes of the reflected and incident waves. Analogously, the *transmission coefficient* is the ratio of the amplitudes of the transmitted and incident waves. Since we assume the unit amplitude of the incident wave, amplitude A is directly equal to the reflection coefficient, and B is the transmission coefficient.

12.3.2 Application of the boundary condition on the continuity of displacement

The wavefield in the first half-space is composed of the incident and reflected waves, i.e. $\mathbf{u}_1 = \mathbf{u} + \mathbf{u}_R = (0, v_1, 0)$, where

$$v_1 = v + v_R = \exp\left[i\omega\left(t - \frac{x\sin i_1 + z\cos i_1}{\beta_1}\right)\right] + A\exp\left[i\omega_R\left(t - \frac{x\sin_1^R - z\cos_1^R}{\beta_1}\right)\right] \tag{12.7}$$

The wavefield in the second half-space is formed by the transmitted wave only, $\mathbf{u}_2 = \mathbf{u}_T = (0, v_2, 0)$, where

$$v_2 = v_T = B\exp\left[i\omega_T\left(t - \frac{x\sin_2 + z\cos_2}{\beta_2}\right)\right] . \tag{12.8}$$

The vector boundary condition (12.1) now reduces to one scalar condition,

$$v_1 = v_2 \quad \text{for} \quad z = 0 , \tag{12.9}$$

because $u_1 = u_2$ and $w_1 = w_2$ are satisfied identically. Inserting (12.7) and (12.8) into (12.9) yields

$$\exp\left[i\omega\left(t - \frac{x\sin i_1}{\beta_1}\right)\right] + A\exp\left[i\omega_R\left(t - \frac{x\sin_1^R}{\beta_1}\right)\right] = B\exp\left[i\omega_T\left(t - \frac{x\sin_2}{\beta_2}\right)\right] . \tag{12.10}$$

This boundary condition must be satisfied at any time and at any place along the x-axis. This will be satisfied if the corresponding exponential terms are identical, i.e. their arguments are identical. The independence from time yields

$$\omega_R = \omega_T = \omega , \tag{12.11}$$

which represents the well-known fact that the reflected and transmitted waves have the same frequencies as the incident wave. The independence from coordinate x yields

$$\frac{\sin i_1}{\beta_1} = \frac{\sin i_1^R}{\beta_1} = \frac{\sin i_2}{\beta_2} . \tag{12.12}$$

Consequently, the angle of reflection is equal to the angle of incidence, $i_1^R = i_1$, and instead of i_1^R we shall write only i_1. The condition for the angle i_2 is the well-known Snell's law.

The results, at which we have just arrived, may be summarised as the so-called reflection/transmission laws (R/T laws):

I) The rays of the *reflected and transmitted waves* remain *in the plane of incidence*. (Here we have taken this property as an empirical fact; we have not derived it. Its derivation from the equations of motion and boundary conditions would require a more detailed analysis of the problem).

II) *The angle of reflection equals the angle of incidence*, $i_1^R = i_1$, see Fig. 12.1.

III) *The angle of transmission satisfies Snell's law*:

$$\frac{\sin i_1}{\beta_1} = \frac{\sin i_2}{\beta_2} \ . \tag{12.13}$$

Note that these reflection/transmission laws have been specified here for *SH* waves, but they have a more general validity.

Since all exponential terms in the boundary condition (12.10) are now identical, this condition reduces to

$$1 + A = B \ . \tag{12.14}$$

This is one equation for the unknown reflection and transmission coefficients. The second equation for these coefficients will follow from the continuity of stress.

12.3.3 Application of the boundary condition on the continuity of stress

Hooke's law for an isotropic medium, i.e.

$$\tau_{ij} = \lambda \vartheta \delta_{ij} + \mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) , \tag{12.15}$$

now yields the stresses in (12.2) in the form

$$\tau_{zx} = \mu \left(\frac{\partial w}{\partial x} + \frac{\partial u}{\partial z} \right) = 0 , \quad \tau_{zy} = \mu \left(\frac{\partial w}{\partial y} + \frac{\partial v}{\partial z} \right) = \mu \frac{\partial v}{\partial z} ,$$

$$\tau_{zz} = \lambda \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \right) + 2\mu \frac{\partial w}{\partial z} = 0 , \tag{12.16}$$

where we have omitted the subscripts 1 and 2. In the last expression we have taken into account that the displacement v is independent of the y-coordinate, so that $\partial v / \partial y = 0$. Since τ_{zx} and τ_{zz} are identically equal to zero, three boundary conditions (12.2) reduce to one condition, namely

$$\mu_1 \frac{\partial v_1}{\partial z} = \mu_2 \frac{\partial v_2}{\partial z} \quad \text{for} \quad z = 0 \; . \tag{12.17}$$

Inserting displacements (12.7) and (12.8) into this condition, and omitting the identical exponential terms, we get

$$\frac{\mu_1}{\beta_1} \cos i_1 (1 - A) = B \frac{\mu_2}{\beta_2} \cos i_2 \; . \tag{12.18}$$

We shall rewrite this equation in two forms. First, inserting $\mu_1 = \rho_1 \beta_1^2$ and $\mu_2 = \rho_2 \beta_2^2$, we obtain

$$1 - A = B \frac{\rho_2 \beta_2 \cos i_2}{\rho_1 \beta_1 \cos i_1} \; . \tag{12.19}$$

The solution of the system of Eqs. (12.14) and (12.19) can then be expressed in a "symmetrical" form as

$$A = \frac{\rho_1 \beta_1 \cos i_1 - \rho_2 \beta_2 \cos i_2}{\rho_1 \beta_1 \cos i_1 + \rho_2 \beta_2 \cos i_2} \, , \quad B = \frac{2 \rho_1 \beta_1 \cos i_1}{\rho_1 \beta_1 \cos i_1 + \rho_2 \beta_2 \cos i_2} \; . \tag{12.20}$$

Second, introduce the index of transmission for shear waves, $n = \beta_1 / \beta_2$. Equation (12.18) then gives

$$1 - A = B \frac{n \mu_2 \cos i_2}{\mu_1 \cos i_1} \; . \tag{12.21}$$

Since $\sin i_2 = (\sin i_1)/n$ according to Snell's law, the solution of Eqs. (12.14) and (12.21) can now be expressed in terms of the angle of incidence as follows:

$$\boxed{\begin{aligned} A &= \frac{\mu_1 \sqrt{1 - \sin^2 i_1} - \mu_2 \sqrt{n^2 - \sin i_1}}{\mu_1 \sqrt{1 - \sin^2 i_1} + \mu_2 \sqrt{n^2 - \sin^2 i_1}} \, , \\ B &= \frac{2 \mu_1 \sqrt{1 - \sin^2 i_1}}{\mu_1 \sqrt{1 - \sin^2 i_1} + \mu_2 \sqrt{n^2 - \sin^2 i_1}} \; . \end{aligned}} \tag{12.22}$$

If $n > 1$, i.e. in the case of the transmission towards the normal, formulae (12.22) give the reflection and transmission coefficients for any angle of incidence from the interval $\langle 0, 90° \rangle$. If $n < 1$ (transmission away from the normal), these formulae can be used only for the case of $\sin i_1 \le n$. Their modification for larger angles of incidence will be described in the next section.

A very important case is the so-called normal incidence when the angle of incidence is equal to zero. Formulae (12.20) then attain a very simple form:

$$A = \frac{\rho_1\beta_1 - \rho_2\beta_2}{\rho_1\beta_1 + \rho_2\beta_2}, \quad B = \frac{2\rho_1\beta_1}{\rho_1\beta_1 + \rho_2\beta_2} . \tag{12.23}$$

It should be noted that the R/T coefficients, derived here, are independent of frequency. Therefore, we could also solve this problem in the time domain (Psencik, 1994). However, the solution in the frequency domain has been simpler (the exponentials can easily be differentiated, etc.).

It can be shown that the reflection and transmission on more complicated structures, such as a thin layer or a transition zone, depend also on frequency. These more complicated problems are usually solved by matrix methods; see Novotny (1999).

12.4 Total Reflection of *SH* Waves

Consider the case of $\beta_1 < \beta_2$, i.e. $n < 1$. It follows from Snell's law that

$$\sin i_2 = \frac{\beta_2}{\beta_1} \sin i_1 , \tag{12.24}$$

so that the transmission occurs away from the normal, $i_2 > i_1$. In this case, there is the so-called *critical angle* of incidence, i_1^*, for which the angle of transmission i_2 is equal to $90°$. Putting $\sin i_2 = 1$ in (12.24) we get the following formula for the critical angle:

$$\sin i_1^* = \frac{\beta_1}{\beta_2} = n . \tag{12.25}$$

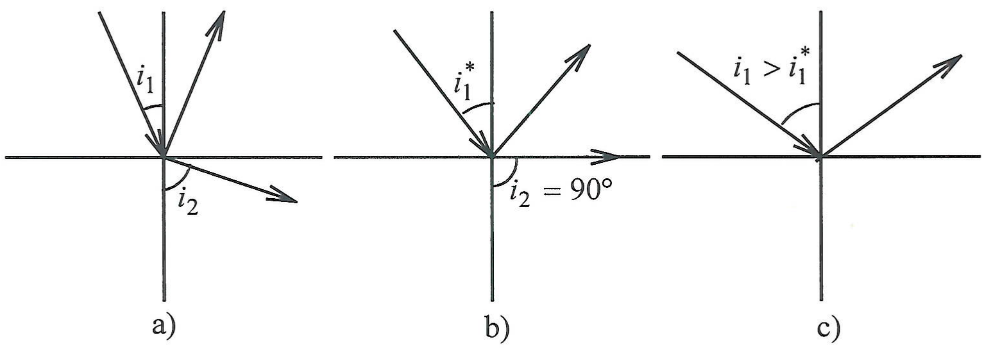


Fig. 12.2. Reflection and transmission at different angles of incidence: a) subcritical; b) critical; c) overcritical.

If we vary the angle of incidence from small values to large values, three specific situations occur (Fig. 12.2):

a) *Subcritical incidence*, $i_1 < i_1^*$. In this case, the angle of transmission i_2 satisfies the condition $i_1 < i_2 < 90°$; see Fig 12.2a. A usual transmitted wave, called *homogeneous wave*, propagates in the second medium.

b) *Critical incidence*, $i_1 = i_1^*$. In this case, the angle of transmission $i_2 = 90°$, which means that the transmitted wave in the second medium is parallel to the interface; see Fig 12.2b.

c) *Overcritical incidence*, $i_1 > i_1^*$. In this case we have $\sin i_1 > n$; see Fig. 12.2c. Let us discuss this situation in detail.

For an overcritical angle of incidence, formula (12.24) yields $\sin i_2 > 1$, and so the angle of transmission i_2 must be complex-valued. This also leads to an imaginary value of the square root $\sqrt{n^2 - \sin^2 i_1}$ in formulae (12.22). Therefore, we must put there

$$\sqrt{n^2 - \sin^2 i_1} = \pm i\sqrt{\sin^2 i_1 - n^2} \ . \tag{12.26}$$

The correct sign in the latter expression must be selected on the basis of physical considerations as follows.

Consider expression (12.8) for the transmitted wave for the case of $i_1 > i_1^*$. Since

$$\cos i_2 = \sqrt{1 - (\sin^2 i_1)/n} = \frac{1}{n}\sqrt{n^2 - \sin^2 i_1} = \pm\frac{i}{n}\sqrt{\sin^2 i_1 - n^2} \ ,$$

the displacement v_T is

$$v_T = B\exp\left[i\omega_T\left(t - \frac{x\sin i_1}{\beta_1}\right)\right]\cdot\exp\left[-i\omega\frac{z}{\beta_1}\left(\pm i\sqrt{\sin^2 i_1 - n^2}\right)\right]. \tag{12.27}$$

Since coordinate z is positive in the second medium, the sign "+" would yield an exponentially increasing wave for an increasing distance from the interface. This is physically inadmissible (for $z \to \infty$ the wave would have infinite amplitudes and, consequently, an infinite energy), and so we must choose the sign "–". Thus

$$\sqrt{n^2 - \sin^2 i_1} = -i\sqrt{\sin^2 i_1 - n^2} \ . \tag{12.28}$$

(Note that the sign "+" had to be chosen in the case of the opposite orientation of the z-axis). The amplitude of the wave in the second medium is now exponentially decreasing with the distance from the interface. Such waves are called *inhomogeneous waves* or evanescent waves. It follows from (12.27) that the corresponding wave propagates in the direction of the x-axis with the velocity $c = \beta_1/\sin i_1$. Therefore, the inhomogeneous transmitted wave (12.27) may be expressed as

$$v_T = B\exp\left[-\omega\frac{z}{\beta_1}\sqrt{\sin^2 i_1 - n^2}\right]\exp\left[i\omega\left(t - \frac{x}{c}\right)\right] \ . \tag{12.29}$$

For the overcritical incidence, the reflection and transmission coefficients (12.22) now take the form

$$\boxed{\begin{aligned} A &= \frac{\mu_1\sqrt{1-\sin^2 i_1} + i\mu_2\sqrt{\sin^2 i_1 - n^2}}{\mu_1\sqrt{1-\sin^2 i_1} - i\mu_2\sqrt{\sin^2 i_1 - n^2}} \,, \\ B &= \frac{2\mu_1\sqrt{1-\sin^2 i_1}}{\mu_1\sqrt{1-\sin^2 i_1} - i\mu_2\sqrt{\sin^2 i_1 - n^2}} \,. \end{aligned}} \tag{12.30}$$

Denoting $a = \mu_1\sqrt{1-\sin^2 i_1}$, $b = \mu_2\sqrt{\sin^2 i_1 - n^2}$, we may write the reflection coefficient as

$$A = \frac{a+ib}{a-ib} \,. \tag{12.31}$$

Since this is the ratio of a complex number to the complex conjugate number, it follows that the absolute value of A is equal to unity, $|A| = 1$. Since the amplitude of the reflected wave (its absolute value) is equal to the amplitude of the incident wave, i.e. the whole energy of the incident wave is returned back into the first medium, we speak of the *total reflection.*

Writing the reflection coefficient also as

$$A = |A|e^{i\varphi} \,,$$

we get

$$|A| = 1 \,, \quad \tan\varphi = \frac{2ab}{a^2 - b^2} \,,$$

where a, b have been defined above. Thus, the total reflection does not change the amplitude (the absolute value) of the reflected wave, but causes its phase shift with respect to the incident wave.

The expression for the phase shift may even be simplified. Write the numerator and denominator of (12.31) as

$$a + ib = ce^{i\gamma} \,, \quad a - ib = ce^{-i\gamma} \,, \tag{12.32}$$

where $c = \sqrt{a^2 + b^2}$ and

$$\tan\gamma = \frac{b}{a} = \frac{\mu_2\sqrt{\sin^2 i_1 - n^2}}{\mu_1\sqrt{1-\sin^2 i_1}} \,. \tag{12.33}$$

Inserting (12.32) into (12.31) yields

$$A = e^{2i\gamma} \,, \tag{12.34}$$

which means that $|A| = 1$ and the phase shift is $\varphi = 2\gamma$, where γ is given by (12.33). Formulae (12.33) and (12.34) are used, e.g., in the theory of Love waves in deriving dispersion equations from the condition of constructive interference (Savarensky, 1975; Novotny, 1999).

Note that the total reflection is frequently explained in the textbooks of elementary physics as a phenomenon when no waves propagate in the second medium. This is an oversimplified description, used in geometrical optics and analogous geometrical theories. In physical theories we must admit the existence of waves also in the second medium in order to be able to satisfy the boundary conditions. Namely, if the first medium is in a harmonic motion, it is not possible for the second medium to be completely at rest. In popular words, the incident wave must always penetrate into the second medium, at least a little, in order to "find out" how to propagate further. However, as follows from (12.29), the exponential decay of inhomogeneous waves depends on frequency. Consequently, the "depth" of penetration of high-frequency inhomogeneous waves into the second medium may be very small. This substantiates the fact that, in many problems, these waves may be neglected.

12.5 Reflection of *SH* Waves at a Free Surface

A very important problem is the reflection of seismic waves at the Earth's surface. Therefore, consider a homogeneous and isotropic elastic half-space. In Fig. 12.3 this medium is represented by the lower half-space, $z \geq 0$. Denote the shear wave velocity, density and shear modulus in this half-space by β, ρ and μ, respectively. The upper half space is the vacuum. Note that the same results would be obtained for the reflection of *SH* waves if the upper half-space were the air or another fluid, because shear waves do not propagate in fluids.

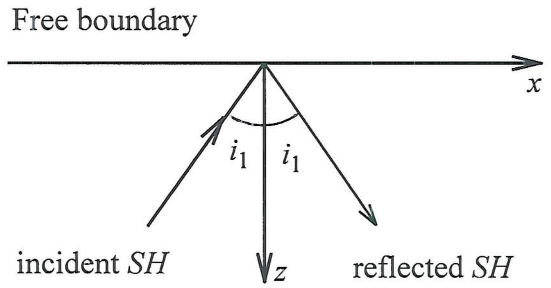


Fig. 12.3. Reflection of *SH* waves at a free surface of an elastic solid.

Let a plane harmonic *SH* wave of an angular frequency ω be impinging on the free boundary at an angle i_1 (Fig. 12.3). The reflection coefficient A could be obtained immediately from formulae (12.20) or (12.22) as their special case when the second medium is the vacuum. By putting $\beta_2 = 0$ and $\mu_2 = 0$ in these formulae, we arrive at the result that the reflection coefficient is equal to unity,

$$A = 1 \ . \tag{12.35}$$

Nevertheless, it will be instructive to derive this simple result from the beginning.

Assume again that the plane of incidence coincides with the (x,z)-plane. The displacement vector for the incident *SH* wave is then $\mathbf{u} = (0, v, 0)$, where

$$v = \exp\left[i\omega\left(t - \frac{x\sin i_1 - z\cos i_1}{\beta} \right) \right] . \tag{12.36}$$

Let us seek the reflected *SH* wave in the form $\mathbf{u}_R = (0, v_R, 0)$, where

$$v_R = A\exp\left[i\omega\left(t - \frac{x\sin i_1 + z\cos i_1}{\beta} \right) \right] ; \tag{12.37}$$

here we have already taken into account that the angular frequency of the reflected wave and the angle of reflection are equal to the corresponding quantities for the incident wave. This follows from the same considerations as in Section 12.3.

The resultant displacement in the half-space, which we shall denote by $\mathbf{u}_1 = (0, v_1, 0)$, is composed of the incident and reflected waves, $\mathbf{u}_1 = \mathbf{u} + \mathbf{u}_R$. Thus

$$v_1 = v + v_R . \tag{12.38}$$

Since the boundary of the half-space is free, we have no constraint on the displacement there. In other words, as opposed to the problem in Section 12.3, here we shall not consider any condition of type (12.1), concerning the displacements. The only boundary condition at the free surface will be the requirement that the corresponding stress component should be zero:

$$\tau_{zy} = \mu\frac{\partial v_1}{\partial z} = 0 \quad \text{for} \quad z = 0 . \tag{12.39}$$

This one boundary condition corresponds to one wave which is generated at the boundary, namely the reflected *SH* wave. Inserting (12.36) to (12.38) into (12.39), we arrive again at the result that $A = 1$.

Since the amplitude of the reflected wave is equal to the amplitude of the incident wave, the motion on the surface is

$$v_1 = 2\exp\left[i\omega\left(t - \frac{x\sin i_1}{\beta} \right) \right] \quad \text{for} \quad z = 0 . \tag{12.40}$$

Hence, the amplitude of the motion at the free surface is twice larger than the amplitude of the incident *SH* wave. This fact must be taken into account in interpreting seismic observations.

12.6 Reflection and Transmission of *P* Waves at an Interface between Two Liquids

In the section we shall discuss another simple problem of reflection and transmission, namely the reflection and transmission of plane harmonic waves at a plane interface between two homogeneous liquids. Since no shear waves can propagate in liquids, we shall consider an incident *P* wave, and only two waves generated at the interface, i.e. a reflected *P* wave and a transmitted *P* wave (Fig. 12.4).

The elastic properties of the media under consideration are described by Lamé's coefficients λ_1, λ_2, and densities ρ_1, ρ_2, respectively. Since the shear moduli are now equal to zero, $\mu_1 = \mu_2 = 0$, the velocities of *P* waves are $\alpha_1 = \sqrt{\lambda_1/\rho_1}$ and $\alpha_2 = \sqrt{\lambda_2/\rho_2}$.

12.6.1 Expressions for displacements

Two approaches have been used in the literature to study the reflection and transmission of *P* waves, namely the formulation in terms of potentials (Chapter 7), and the formulation directly in displacements. The former approach seems to be a little simpler, because each wave is described by one potential only, whereas two displacement components must be considered in the latter approach. However, the former approach yields only the reflection and transmission coefficients for potentials, which must further be transformed into the coefficients for displacements. Here we shall use the second approach, which leads directly to the reflection and transmission coefficients for displacements.

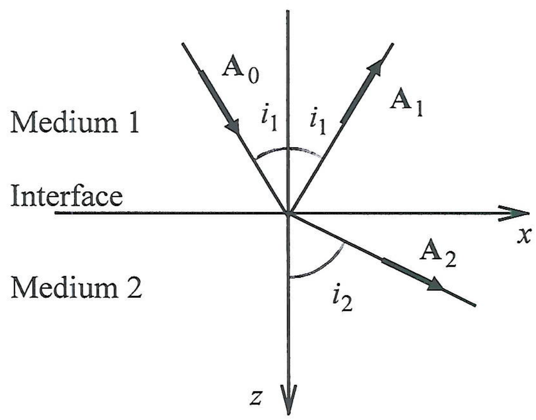


Fig. 12.4. Reflection and transmission of a *P* wave at an interface of two liquids.

Denote the angle of incidence by i_1, the angle of reflection by i_1^R, and the angle of transmission by i_2. Further, denote the "vector" amplitudes for the incident, reflected and transmitted waves by $\mathbf{A}_0$, $\mathbf{A}_1$ and $\mathbf{A}_2$, respectively (Fig. 12.4). This means that $A_0 = |\mathbf{A}_0|$ is the amplitude, and the direction of $\mathbf{A}_0$ determines the polarisation of the incident wave (particle motion in the incident wave). Vectors $\mathbf{A}_1$ and $\mathbf{A}_2$ have analogous meanings.

The analogous discussion as in Section 12.3 would yield that the frequencies of the reflected and transmitted waves are equal to the frequency of the incident wave, that the angle of reflection is equal to the angle of incidence, $i_1^R = i_1$, and the angles of incidence and of transmission are related by Snell's law:

$$\frac{\sin i_1}{\alpha_1} = \frac{\sin i_2}{\alpha_2} \; . \tag{12.41}$$

Therefore, assume the displacements in the individual waves in the following forms (Fig. 12.4):
Incident wave:

$$\mathbf{u}_0 = \mathbf{A}_0 \exp\left[i\omega\left(t - \frac{x \sin i_1 + z \cos i_1}{\alpha_1} \right) \right] , \quad \mathbf{A}_0 = \left(A_0 \sin i_1 , 0, A_0 \cos i_1 \right) ,$$

Reflected wave:

$$\mathbf{u}_R = \mathbf{A}_1 \exp\left[i\omega\left(t - \frac{x \sin i_1 - z \cos i_1}{\alpha_1} \right) \right] , \quad \mathbf{A}_1 = \left(A_1 \sin i_1 , 0, - A_1 \cos i_1 \right) ,$$

Transmitted wave:

$$\mathbf{u}_T = \mathbf{A}_2 \exp\left[i\omega\left(t - \frac{x \sin i_2 + z \cos i_2}{\alpha_2} \right) \right] , \quad \mathbf{A}_2 = \left(A_2 \sin i_2 , 0, A_2 \cos i_2 \right) . \tag{12.42}$$

The resultant displacement in the first half-space, which we shall denote by $\mathbf{u}_1 = (u_1, 0, w_1)$, is composed of the incident and reflected waves, whereas the displacement in the second half-space, $\mathbf{u}_2 = (u_2, 0, w_2)$, is formed by the transmitted wave only:

$$\mathbf{u}_1 = \mathbf{u}_0 + \mathbf{u}_R \, , \quad \mathbf{u}_2 = \mathbf{u}_T \; . \tag{12.43}$$

12.6.2 Boundary conditions at a liquid-liquid interface

As boundary conditions we shall require the continuity of the normal components (z-components) of the displacement and stress:

$$w_1 = w_2 \quad \text{for} \quad z = 0 \; , \tag{12.44}$$

$$\left(\tau_{zz}\right)_1 = \left(\tau_{zz}\right)_2 \quad \text{for} \quad z = 0 \; . \tag{12.45}$$

As opposed to the boundary conditions (12.1) for a welded contact of two solids, here we do not require the continuity of the x-components of displacements, since liquids may slide along the interface. Thus, u_1 may differ from u_2 at $z = 0$. Analogously, we would not require the y-components of displacement to be continuous although, in this case, we assume $v_1 = v_2 = 0$ everywhere. Moreover, since $\mu = 0$ in liquids, shear stresses vanish there, and the first and second boundary conditions in (12.2) are satisfied identically. Consequently, only the boundary conditions (12.44) and (12.45) must be taken into account at the liquid-liquid interface.

Since the boundary condition (12.44) must be satisfied for any coordinate x, all exponential terms in (12.42) must be identical for $z = 0$, which yields Snell's law (12.41). This boundary condition then immediately yields the equation

$$A_0 \cos i_1 - A_1 \cos i_1 = A_2 \cos i_2 \ , \tag{12.46}$$

It follows from Hooke's law that

$$\tau_{zz} = \lambda \operatorname{div} \mathbf{u} + 2\mu \frac{\partial w}{\partial z} \ . \tag{12.47}$$

In liquids, where $\mu = 0$, we have

$$\tau_{zz} = \lambda \operatorname{div} \mathbf{u} \ . \tag{12.48}$$

The boundary condition (12.45) can then be expressed as

$$\lambda_1 \left[\operatorname{div} \mathbf{u}_0 + \operatorname{div} \mathbf{u}_R \right] = \lambda_2 \operatorname{div} \mathbf{u}_T \quad \text{for} \quad z = 0 \ . \tag{12.49}$$

It holds that

$$\operatorname{div} \mathbf{u}_0 = \frac{\partial u_0}{\partial x} + \frac{\partial w_0}{\partial z} = A_0 \left(-\frac{i\omega}{\alpha_1} \right) \left(\sin^2 i_1 + \cos^2 i_1 \right) \ , \tag{12.50}$$

where we have omitted the exponential term at the last expression. Since $\lambda_1 = \rho_1 \alpha_1^2$, we get

$$\lambda_1 \operatorname{div} \mathbf{u}_0 = -i\omega \rho_1 \alpha_1 A_0 \ , \tag{12.51}$$

where the corresponding exponential term should again be added. Considering similar expressions for $\operatorname{div} \mathbf{u}_R$ and $\operatorname{div} \mathbf{u}_T$, Eq. (12.49) yields

$$\rho_1 \alpha_1 \left(A_0 + A_1 \right) = \rho_2 \alpha_2 A_2 \ . \tag{12.52}$$

12.6.3 Reflection and transmission coeficients

Equations (12.46) and (12.52) are the desired equations for computing the amplitudes A_1 and A_2. Introducing the reflection coefficient $R_1 = A_1 / A_0$ and the transmission coefficient $R_2 = A_2 / A_0$, these equations may be expressed as

$$\begin{aligned} \left(1 - R_1\right) \cos i_1 &= R_2 \cos i_2 \ , \\ \left(1 + R_1\right) \rho_1 \alpha_1 &= R_2 \rho_2 \alpha_2 \ . \end{aligned} \tag{12.53}$$

The solution of this system of equations is

$$\boxed{R_1 = \frac{\rho_2\alpha_2\cos i_1 - \rho_1\alpha_1\cos i_2}{\rho_2\alpha_2\cos i_1 + \rho_1\alpha_1\cos i_2}\,, \quad R_2 = \frac{2\rho_1\alpha_1\cos i_1}{\rho_2\alpha_2\cos i_1 + \rho_1\alpha_1\cos i_2}}\,. \tag{12.54}$$

Introducing so-called acoustic impedances, $Z_1 = \rho_1\alpha_1$ and $Z_2 = \rho_2\alpha_2$, the resultant formulae are also written as

$$R_1 = \frac{Z_2\cos i_1 - Z_1\cos i_2}{Z_2\cos i_1 + Z_1\cos i_2}\,, \quad R_2 = \frac{2Z_1\cos i_1}{Z_2\cos i_1 + Z_1\cos i_2}\,. \tag{12.55}$$

It should be pointed out that coefficients (12.54) and the coefficients (12.20) for *SH* waves have rather similar forms. It can be seen that the positions of the terms $\cos i_1$ and $\cos i_2$ in (12.20) and in (12.54) are identical. However, apart from the different velocities, formulae (12.20) contain the combinations $\rho_1\beta_1\cos i_1$ and $\rho_2\beta_2\cos i_2$ only, whereas formulae (12.54) contain $Z_1\cos i_1$, $Z_1\cos i_2$ and $Z_2\cos i_1$.

12.6.4 Reflection and transmission for the normal incidence

For the special case of the normal incidence, when $i_1 = i_2 = 0$, formulae (12.54) become

$$R_1 = \frac{\rho_2\alpha_2 - \rho_1\alpha_1}{\rho_2\alpha_2 + \rho_1\alpha_1}\,, \quad R_2 = \frac{2\rho_1\alpha_1}{\rho_2\alpha_2 + \rho_1\alpha_1}\,. \tag{12.56}$$

These formulae are similar to the formulae (12.23) for the normal incidence of *SH* waves, but the signs of the reflection coefficients are opposite. Therefore, we may say that the processes of the normal incidence of *SH* waves on a solid-solid interface, and of *P* waves on a liquid-liquid interface represent two analogous phenomena. The sign difference may be ascribed to the different orientations of reflected waves. Namely, the orientation of the displacement vectors for the incident and reflected *SH* waves was identical, coinciding with the positive direction of the *y*-axis (Section 12.3). However, for the normal incidence $(i_1 = 0)$, the vector $\mathbf{A}_0$ for the incident *P* wave is directed along the *z*-axis downwards, whereas the orientation of vector $\mathbf{A}_1$ is opposite (Fig. 12.4).

12.6.5 Reflection of *P* waves at a free surface of a liquid

Consider a similar problem as shown in Fig. 12.3 (see Section 12.5), but for *P* waves propagating in a homogeneous liquid. Since now the second medium is the vacuum, the reflection coefficient for *P* waves at a free surface of a liquid can be obtained from (12.54) by putting $\alpha_2 = 0$. We get

$$R_1 = -1\,. \tag{12.57}$$

Note again that this reflection coefficient is analogous to the corresponding reflection coefficient for *SH* waves, see (12.35), but its sign is opposite.

Finally, let us study the motion of the free surface. It follows from (12.57) that $A_1 = -A_0$, and formulae (12.42) and (12.43) yield the displacement vector for $z = 0$ in the form $\mathbf{u}_1 = (u_1, 0, w_1)$, where

$$u_1 = 0\,, \quad w_1 = 2A_0(\cos i_1)\exp\left[i\omega\left(t - \frac{x\sin i_1}{\alpha_1}\right)\right]. \tag{12.58}$$

Hence, due to the superposition of the incident and reflected *P* waves, the motion of the free surface of a liquid has only a vertical component.

12.7 Reflection and Transmission of *P* waves. The Zoeppritz Equations

Now we shall consider a more complicated problem than in the previous section. We shall return back to the model consisting of two homogeneous and isotropic solid half-spaces in a welded contact (Section 12.2).

Consider a plane harmonic *P* wave incident on the interface from the first half-space (Fig. 12.5). Choose again the plane $z = 0$ to coincide with the interface and the (x,z)-plane to coincide with the plane of incidence.

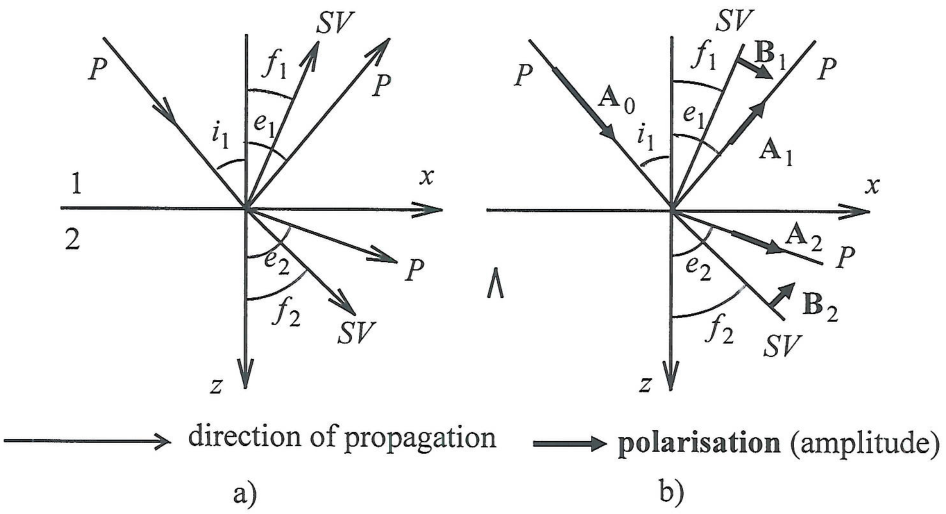


Fig. 12.5. The coordinate system and the seismic rays for the case of an incident *P* wave. The arrows indicate: a) the directions of propagation; b) the polarisation.

12.7.1 Displacements and boundary conditions

We shall again denote the displacement vector by $\mathbf{u} = (u,v,w)$, to which the corresponding subscripts and superscripts will be added for the individual waves. In particular, denote the displacement vector for the incident wave by

$$\mathbf{u}_0 = (u_0, 0, w_0) , \qquad (12.59)$$

by $\mathbf{u}_1$ the resultant displacement in the first half-space, which is composed of the displacements for the incident and reflected waves. Analogously, the resultant displacement in the second half-space, composed of transmitted waves, will be denoted by $\mathbf{u}_2$.

Since the incident *P* wave has non-zero displacement components into the *x*- and *z*-axes, also the resultant displacements u_1, w_1, u_2 and w_2 will be generally non-zero. Consequently, the stress components $(\tau_{zx})_1$, $(\tau_{zz})_1$ $(\tau_{zx})_2$ and $(\tau_{zz})_2$ will also be non-zero. We may expect that only the following boundary conditions will be satisfied identically:

$$v_1 = v_2 = 0 , \quad (\tau_{zy})_1 = (\tau_{zy})_2 = 0 . \qquad (12.60)$$

Therefore, the following boundary conditions are to be satisfied for $z = 0$:

$$u_1 = u_2 , \qquad (12.61a)$$
$$w_1 = w_2 , \qquad (12.61b)$$
$$(\tau_{zx})_1 = (\tau_{zx})_2 , \qquad (12.61c)$$
$$(\tau_{zz})_1 = (\tau_{zz})_2 . \qquad (12.61d)$$

In order to satisfy these four boundary conditions, we must assume that four waves are generated at the interface, namely a reflected *P*, transmitted *P*, but also a reflected *SV* and transmitted *SV*. For the incident *P* wave, denote its angle of incidence by i_1 and the "vector" amplitude by $\mathbf{A}_0$ (Fig. 12.5b). This means that $A_0 = |\mathbf{A}_0|$ is the amplitude, and the direction of $\mathbf{A}_0$ determines the direction of the particle motion (see also Section 12.6). Introduce analogous notations for the reflected and transmitted waves: e_1 and $\mathbf{A}_1$ for the reflected *P*, f_1 and $\mathbf{B}_1$ for the reflected *SV*, e_2 and $\mathbf{A}_2$ for the transmitted *P*, f_2 and $\mathbf{B}_2$ for the transmitted *SV* (Fowler, 1990). Each of these vectors could be chosen as shown in Fig. 12.5b or in the opposite direction. We have chosen their orientations so that each vector would have a positive component into the *x*-axis.

Considerations, similar to those in Section 12.3, lead to the generalised *Snell's law*:

$$\frac{\sin i_1}{\alpha_1} = \frac{\sin e_1}{\alpha_1} = \frac{\sin f_1}{\beta_1} = \frac{\sin e_2}{\alpha_2} = \frac{\sin f_2}{\beta_2} . \qquad (12.62)$$

Although $e_1 = i_1$, we shall use both angles in order to distinguish terms belonging to the incident and reflected *P* waves.

As in the previous section, we shall solve the problem directly in terms of displacements, not in terms of elastodynamic potentials. Therefore, assume the displacements of the individual waves in the following forms (Fig. 12.5b):
Incident P:

$$\mathbf{u}_0 = \mathbf{A}_0 \exp\left[i\omega\left(t - \frac{x \sin i_1 + z \cos i_1}{\alpha_1}\right)\right], \quad \mathbf{A}_0 = \left(A_0 \sin i_1 , 0, A_0 \cos i_1\right) .$$

Reflected P:

$$\mathbf{u}_R^P = \mathbf{A}_1 \exp\left[i\omega\left(t - \frac{x \sin e_1 - z \cos e_1}{\alpha_1}\right)\right], \quad \mathbf{A}_1 = \left(A_1 \sin e_1 , 0, - A_1 \cos e_1\right) .$$

Reflected SV:

$$\mathbf{u}_R^{SV} = \mathbf{B}_1 \exp\left[i\omega\left(t - \frac{x \sin f_1 - z \cos f_1}{\beta_1}\right)\right], \quad \mathbf{B}_1 = \left(B_1 \cos f_1 , 0, B_1 \sin f_1\right) .$$

Transmitted P:

$$\mathbf{u}_T^P = \mathbf{A}_2 \exp\left[i\omega\left(t - \frac{x \sin e_2 + z \cos e_2}{\alpha_2}\right)\right], \quad \mathbf{A}_2 = \left(A_2 \sin e_2 , 0, A_2 \cos e_2\right) .$$

Transmitted SV:

$$\mathbf{u}_T^{SV} = \mathbf{B}_2 \exp\left[i\omega\left(t - \frac{x \sin f_2 + z \cos f_2}{\beta_2}\right)\right], \quad \mathbf{B}_2 = \left(B_2 \cos f_2 , 0, - B_2 \sin f_2\right). \tag{12.63}$$

12.7.2 Continuity of displacements

The resultant displacement in the first half-space is then

$$\mathbf{u}_1 = \mathbf{u}_0 + \mathbf{u}_R^P + \mathbf{u}_R^{SV} , \tag{12.64}$$

and in the second half-space,

$$\mathbf{u}_2 = \mathbf{u}_T^P + \mathbf{u}_T^{SV} . \tag{12.65}$$

The boundary conditions (12.61a) and (12.61b) for $z = 0$, after omitting the identical exponential terms, now yield the vector equation

$$\mathbf{A}_0 + \mathbf{A}_1 + \mathbf{B}_1 = \mathbf{A}_2 + \mathbf{B}_2 . \tag{12.66}$$

For the x- and z-components of this equation we get from (12.63) that

$$A_0 \sin i_1 + A_1 \sin e_1 + B_1 \cos f_1 = A_2 \sin e_2 + B_2 \cos f_2 , \tag{12.67}$$
$$A_0 \cos i_1 - A_1 \cos e_1 + B_1 \sin f_1 = A_2 \cos e_2 - B_2 \sin f_2 . \tag{12.68}$$

12.7.3 Continuity of stress

As introduced above, see (12.59), the x- and z-components of vector $\mathbf{u}_0$ have been denoted by u_0 and w_0, respectively. We shall use similar notations also for the components of the other vectors $\mathbf{u}_R^P$, $\mathbf{u}_R^{SV}$, $\mathbf{u}_T^P$, $\mathbf{u}_T^{SV}$, $\mathbf{u}_1$ and $\mathbf{u}_2$. The boundary condition (12.61c) may then be expressed as

$$\mu_1\left[\frac{\partial}{\partial x}\left(w_0+w_R^P+w_R^{SV}\right)+\frac{\partial}{\partial z}\left(u_0+u_R^P+u_R^{SV}\right)\right]=$$
$$=\mu_2\left[\frac{\partial}{\partial x}\left(w_T^P+w_T^{SV}\right)+\frac{\partial}{\partial z}\left(u_T^P+u_T^{SV}\right)\right] \quad \text{for} \quad z=0 \; . \qquad (12.69)$$

Inserting expressions (12.63) into this condition, we obtain

$$\frac{\mu_1}{\alpha_1}\left(-A_0\sin 2i_1+A_1\sin 2e_1\right)+\frac{\mu_1}{\beta_1}B_1\cos 2f_1=-\frac{\mu_2}{\alpha_2}A_2\sin 2e_2-\frac{\mu_2}{\beta_2}B_2\cos 2f_2 \qquad (12.70)$$

The last boundary condition (12.61d) can be expressed as

$$\lambda_1\,\mathrm{div}\,\mathbf{u}_1+2\mu_1\frac{\partial w_1}{\partial z}=\lambda_2\,\mathrm{div}\,\mathbf{u}_2+2\mu_2\frac{\partial w_2}{\partial z} \qquad (12.71)$$

for $z=0$, where λ_i, μ_i $(i=1,2)$ are Lamé's elastic parameters. Analogously to Subsection 12.6.2, calculate the divergences of the displacements for the individual waves. For $z=0$, omitting the identical exponential terms on the right-hand sides of the following expressions, we get

$$\mathrm{div}\,\mathbf{u}_0=\frac{\partial u_0}{\partial x}+\frac{\partial w_0}{\partial z}=-\frac{i\omega}{\alpha_1}A_0\,,\quad \mathrm{div}\,\mathbf{u}_R^P=-\frac{i\omega}{\alpha_1}A_1\,,\quad \mathrm{div}\,\mathbf{u}_R^{SV}=0\,,$$
$$\mathrm{div}\,\mathbf{u}_T^P=-\frac{i\omega}{\alpha_2}A_2\,,\quad \mathrm{div}\,\mathbf{u}_T^{SV}=0\;. \qquad (12.72)$$

We could expect the zero values of the divergence for *SV* waves, because div $\mathbf{u}$ is equal to the volume dilatation, but shear waves are not connected with volume changes. Now, denote the contribution of the incident wave to the expression in (12.71) by T_0, i.e.

$$T_0=\lambda_1\,\mathrm{div}\,\mathbf{u}_0+2\mu_1\frac{\partial w_0}{\partial z}\;. \qquad (12.73)$$

This may be expressed as (we still put $z=0$ and omit the exponential terms)

$$T_0=\frac{-i\omega}{\alpha_1}A_0\left[\lambda_1+2\mu_1\cos^2 i_1\right]=-\frac{i\omega}{\alpha_1}A_0\left[\lambda_1+2\mu_1-2\mu_1\sin^2 i_1\right]\;.$$

Since $(\lambda_1 + 2\mu_1) = \rho_1\alpha_1^2$, $\mu_1 = \rho_1\beta_1^2$ and $\sin i_1 = (\alpha_1/\beta_1)\sin f_1$ according to Snell's law (12.62), we get the following simple expression:

$$T_0 = -i\omega\rho_1\alpha_1 A_0 \cos 2f_1 \ . \tag{12.74a}$$

Similarly

$$T_R^P = -i\omega\rho_1\alpha_1 A_1 \cos 2f_1 \ , \quad T_T^P = -i\omega\rho_2\alpha_2 A_2 \cos 2f_2 \ , \tag{12.74b,c}$$

For the contributions of the shear waves we have

$$T_R^{SV} = 2\mu_1 \frac{\partial w_R^{SV}}{\partial z} = 2\mu_1 B_1 \sin f_1 \frac{i\omega}{\beta_1} \cos f_1 = i\omega\rho_1\beta_1 B_1 \sin 2f_1 \ , \tag{12.74d}$$

$$T_T^{SV} = i\omega\rho_2\beta_2 B_2 \sin 2f_2 \ . \tag{12.74e}$$

Consequently, the boundary condition (12.71) yields

$$\begin{aligned} -\rho_1\alpha_1 A_0 \cos 2f_1 - \rho_1\alpha_1 A_1 \cos 2f_1 + \rho_1\beta_1 B_1 \sin 2f_1 = \\ = -\rho_2\alpha_2 A_2 \cos 2f_2 + \rho_2\beta_2 B_2 \sin 2f_2 \ . \end{aligned} \tag{12.75}$$

12.7.4 The Zoeppritz equations

Finally, boundary conditions (12.61), i.e. Eqs. (12.67), (12.68), (12.70) and (12.75), may be expressed as (Fowler, 1990):

$$A_1 \sin e_1 + B_1 \cos f_1 - A_2 \sin e_2 - B_2 \cos f_2 = -A_0 \sin i_1 \ , \tag{12.76a}$$

$$A_1 \cos e_1 - B_1 \sin f_1 + A_2 \cos e_2 - B_2 \sin f_2 = A_0 \cos i_1 \ , \tag{12.76b}$$

$$A_1\gamma_1 W_1 \sin 2e_1 + B_1 W_1 \cos 2f_1 + A_2\gamma_2 W_2 \sin 2e_2 + B_2 W_2 \cos 2f_2 = A_0\gamma_1 W_1 \sin 2i_1 \tag{12.76c}$$

$$A_1 Z_1 \cos 2f_1 - B_1 W_1 \sin 2f_1 - A_2 Z_2 \cos 2f_2 + B_2 W_2 \sin 2f_2 = -A_0 Z_1 \cos 2f_1 \tag{12.76d}$$

where

$$\begin{aligned} Z_1 = \rho_1\alpha_1 \ , \quad Z_2 = \rho_2\alpha_2 \ , \\ W_1 = \rho_1\beta_1 \ , \quad W_2 = \rho_2\beta_2 \ , \\ \gamma_1 = \beta_1/\alpha_1 \ , \quad \gamma_2 = \beta_2/\alpha_2 \ . \end{aligned} \tag{12.77}$$

Equations (12.76) are referred to as the *Zoeppritz equations* for an incident *P* wave. They can be used to compute the amplitudes of reflected and transmitted waves if the angle of incidence i_1 and amplitude A_0 of the incident *P* wave are given, and the remaining angles are determined from Snell's law (12.62). Generally, these equations must be solved numerically.

Let us remind that Eqs. (12.76a) and (12.76b) have come from the conditions of the continuity of the horizontal and vertical displacements, and Eqs. (12.76c) and (12.76d) from the continuity of the horizontal and vertical stresses at the interface.

Let us specify the reflection and transmission coefficients for the important case of normal incidence on the interface, i.e. for the angle of incidence $i_1 = 0$. In this case, the other angles are also equal to zero, and the Zoeppritz equations (12.76) reduce to

$$\begin{aligned} B_1 - B_2 &= 0 \ , \\ A_1 + A_2 &= A_0 \ , \\ B_1 W_1 + B_2 W_2 &= 0 \ , \\ A_1 Z_1 - A_2 Z_2 &= -A_0 Z_1 \ . \end{aligned} \tag{12.78}$$

It follows that no shear waves are generated in this case, $B_1 = B_2 = 0$. The reflection coefficient A_1/A_0 and transmission coefficient A_2/A_0 for P waves now read

$$\begin{aligned} \frac{A_1}{A_0} &= \frac{Z_2 - Z_1}{Z_2 + Z_1} = \frac{\rho_2\alpha_2 - \rho_1\alpha_1}{\rho_2\alpha_2 + \rho_1\alpha_1} \ , \\ \frac{A_2}{A_0} &= \frac{2Z_1}{Z_2 + Z_1} = \frac{2\rho_1\alpha_1}{\rho_2\alpha_2 + \rho_1\alpha_1} \ . \end{aligned} \tag{12.79}$$

Note that the last formulae are identical with the formulae (12.56) for the normal incidence of P waves at the interface of liquid media.

Let us discuss in greater detail how the special equations for a liquid-liquid interface, derived in Section 12.6, follow from the Zoeppritz equations (12.76) for a solid-solid interface. Firstly, omit the first equation (12.76a), which is not valid for a liquid-liquid interface (displacement u may be discontinuous there). Secondly, omit the third equation (12.76c), which is satisfied identically since $\tau_{zx} = 0$ in liquids. Then, omit the terms corresponding to SV waves in the remaining two Zoeppritz equations. Taking into account different notations, i.e. replacing e_1 by i_1 and e_2 by i_2, Eq. (12.76b) becomes Eq. (12.46). The last equation (12.76d) requires certain modifications because it contains angles f_1 and f_2 for SV waves. However,

$$\cos 2f_1 = 1 - 2\sin^2 f_1 = 1 - 2(\beta_1/\alpha_1)^2 \sin i_1 \ ,$$

and a similar formula holds for $\cos 2f_2$. Therefore, for $\beta_1 = \beta_2 = 0$ we must put $\cos 2f_1 = \cos 2f_2 = 1$, and Eq. (12.76d) yields Eq. (12.52).

12.7.5 Numerical examples and their discussion

There are extensive tables and graphs of various reflection and transmission coefficients; see, e.g., the graphs in Ewing et al. (1957). Here we shall consider only the coefficients of the PP type (a P wave is incident, and the reflected P wave is considered). In seismic prospecting, these very coefficients are most important. Let us consider only the case when the velocity below the interface is higher than the velocity above it, i.e. if $n = \alpha_1/\alpha_2 < 1$.

The reflection coefficients of the PP type have a different character for interfaces with a weak velocity differentiation $(n \sim 0.6 - 1.0)$ and for interfaces with a strong velocity differentiation $(n \le 0.6)$. Two typical examples of the PP reflection coefficients are shown in Fig. 12.6 (Cerveny, 1976). In both cases, the shear wave velocities and densities satisfied the following conditions:

$$\frac{\alpha_1}{\beta_1} = \frac{\alpha_2}{\beta_2} = \sqrt{3}\,, \quad \frac{\rho_1}{\rho_2} = 1\,.$$

For a comparison, the reflection coefficients for a liquid-liquid interface (and again for $\rho_1/\rho_2 = 1$) are shown in the figure by dashed lines; see formulae (12.54). It can be seen that for the normal incidence $(i_1 = 0)$, the reflection coefficients for solid and liquid media are identical; see (12.56) and (12.79).

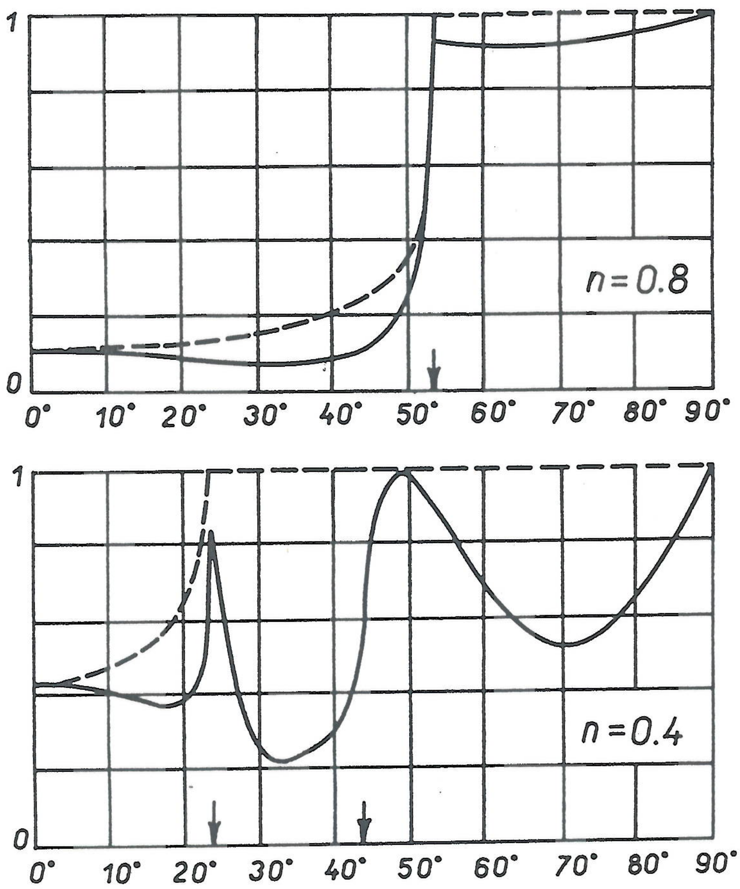


Fig. 12.6. Reflection coefficients of the PP type for the coefficients of transmission $n = 0.8$ and $n = 0.4$. (After Cerveny (1976)).

For $i_1 > 0$, the reflection coefficient for solid media is always smaller than that for liquid media. This difference is connected with the generation of *SV* waves in solid media, which carry away a part of energy. This difference is pronounced especially in the case small indices of transmission.

At first, let us discuss the reflection coefficient for a weak velocity differentiation $(n = 0.8)$. Typical examples of such discontinuities are the discontinuities within the Earth's crust, e.g., the Mohorovicic discontinuity and others. In our case, the critical angle is $i_1^* \sim 53°$; see the arrow in Fig. 12.6. Undercritical reflections $\left(i_1 < i_1^*\right)$ show a different behaviour from overcritical reflections $\left(i_1 > i_1^*\right)$, the overcritical reflections being considerably stronger than the undercritical ones. This fact has been well confirmed in seismic practice, i.e. in studies of the reflections from the Mohorovicic discontinuity. These reflections are usually observed at larger epicentral distances only, from $50 - 70$ km farther.

In the case of a strong velocity differentiation ($n = 0.4$ in Fig. 12.6), even the undercritical reflections are relatively strong. For larger angles of incidence, the situation is rather complicated. If $\beta_2 > \alpha_1$, in addition to the so-called first critical angle i_1^*, there is also the second critical angle $i_1^{**} = \arcsin\left(\alpha_1/\beta_2\right)$. Their values in our case are $i_1^* = 23.5°$ and $i_1^{**} \sim 44°$; see the arrows in the figure. The reflection coefficient has usually a deep minimum between these critical angles, and a broad maximum is formed immediately behind the second critical angle.

12.8 Reflection and Transmission of *SV* Waves

Let us consider a similar problem as in the previous section, but for an incident *SV* wave. Denote its angle of incidence by j_1 and the vector amplitude (polarisation vector) by $\mathbf{B}_0$ as shown in Fig. 12.7. For the reflected and transmitted waves we shall use the same notations as in Fig. 12.5.

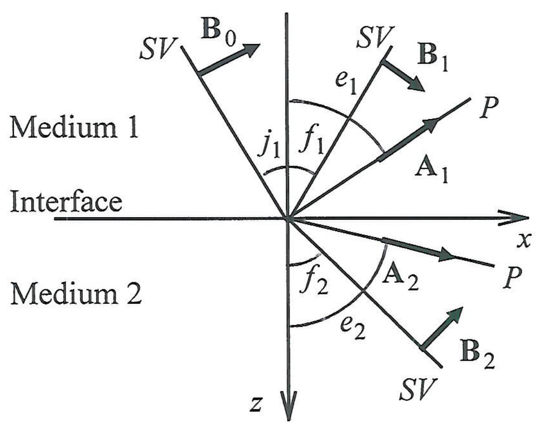


Fig. 12.7. Reflection and transmission of an incident *SV* wave.

Snell's law now reads

$$\frac{\sin j_1}{\beta_1} = \frac{\sin e_1}{\alpha_1} = \frac{\sin f_1}{\beta_1} = \frac{\sin e_2}{\alpha_2} = \frac{\sin f_2}{\beta_2} \ . \tag{12.80}$$

The displacement vector for the incident *SV* wave now takes the form

$$\mathbf{u}_0 = \mathbf{B}_0 \exp\left[i\omega\left(t - \frac{x \sin j_1 + z \cos j_1}{\beta_1}\right)\right], \quad \mathbf{B}_0 = \left(B_0 \cos j_1, 0, -B_0 \sin j_1\right) \ , \tag{12.81}$$

but the expressions for the reflected and transmitted waves are identical with those in the previous section; see (12.63). Consequently, the left-hand sides of the Zoeppritz equations (12.76) will remain without any changes, but the right-hand sides will be different. Let us attempt to determine these right-hand sides without performing the corresponding calculations.

It follows from the comparison of vector $\mathbf{A}_0$ with vector $\mathbf{B}_0$ that the expression $\left(-A_0 \sin i_1\right)$ on the right-hand side of (12.76a) must now be replaced by $\left(-B_0 \cos j_1\right)$, and $A_0 \cos i_1$ in (12.76b) by $\left(-B_0 \sin j_1\right)$. The new right-hand sides of (12.76c) and (12.76d) would follow from more complicated calculations. However, the right-hand sides of the Zoeppritz equations may also be determined directly on the basis of the following similarities.

Firstly, write the left-hand sides of the Zoeppritz equations (12.76). The vector $\mathbf{A}_0$ for the incident *P* wave has a similar direction as vector $\mathbf{A}_2$ for the transmitted *P* wave (positive components into the *x*- and *z*-axes). Consequently, the right-hand sides of (12.76) are analogous to the terms for the transmitted *P* wave on the left-hand sides (to obtain the right-hand sides, it is sufficient to replace e_2, $\mathbf{A}_2$, etc., by i_1, $\mathbf{A}_0$, etc.).

In a similar way we can obtain the *Zoeppritz equations* for an *incident SV wave*. It follows from Fig. 12.7 that the right-hand sides must be analogous to the terms for the transmitted *SV* wave. Therefore, the term $\left(-B_2 \cos f_2\right)$ on the left-hand side of (12.76a) indicates that the right-hand side should be $\left(-B_0 \cos j_1\right)$, etc. Finally, we get (Richter, 1958)

$$\boxed{\begin{aligned}
&A_1 \sin e_1 + B_1 \cos f_1 - A_2 \sin e_2 - B_2 \cos f_2 = -B_0 \cos j_1 \\
&A_1 \cos e_1 - B_1 \sin f_1 + A_2 \cos e_2 - B_2 \sin f_2 = -B_0 \sin j_1 \qquad (12.82) \\
&A_1 \gamma_1 W_1 \sin 2e_1 + B_1 W_1 \cos 2f_1 + A_2 \gamma_2 W_2 \sin 2e_2 + B_2 W_2 \cos 2f_2 = B_0 W_1 \cos 2j_1 \\
&A_1 Z_1 \cos 2f_1 - B_1 W_1 \sin 2f_1 - A_2 Z_2 \cos 2f_2 + B_2 W_2 \sin 2f_2 = B_0 W_1 \sin 2j_1
\end{aligned}}$$

where we have used notations (12.77), and the angles of reflection and transmission are given by Snell's law (12.80).

12.9 Reflection of *P* and *SV* Waves at a Free Surface

Very important special cases of the Zoeppritz equations (12.76) and (12.82) are the equations for the reflection of *P* and *SV* waves at the free surface of the Earth (see the analogous problem for *SH* waves in Section 12.5). In these cases, in general, an incident *P* or *SV* wave produces both reflected *P* and reflected *SV* waves.

Many authors have solved the problem of the reflection of plane waves at a free surface. For example, the reflection coefficients for the elastodynamic potentials and for the energy of waves can be found in Ewing et al. (1957). However, for practical applications, the most important reflection coefficients are those for displacements. Here we shall derive these coefficients from special forms of the Zoeppritz equations given above.

It should be noted that no constraints are now imposed on the displacements at the surface. The only two boundary conditions are the requirements that the stresses at the surface be zero.

12.9.1 Reflection of *P* waves

Consider a homogeneous and isotropic half-space with a free boundary (without stresses). Denote the compressional wave velocity, shear wave velocity and density in this half-space by α_1, β_1 and ρ_1, respectively.

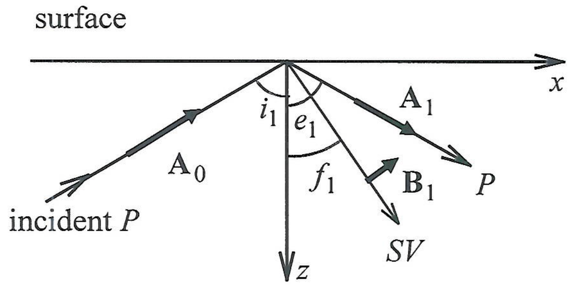


Fig. 12.8. Reflection of a *P* wave at a free surface.

Let a plane harmonic *P* wave, propagating in the half-space obliquely upwards, be incident at the free surface (Fig. 12.8). Denote again the angle of incidence by i_1, the angle of reflection for the *P* wave by e_1, and for the *SV* wave by f_1. Introduce the polarisation vectors of the individual waves, $\mathbf{A}_0$, $\mathbf{A}_1$ and $\mathbf{A}_2$, as show in the figure. The displacement vector $\mathbf{u}_0$ for the incident wave, $\mathbf{u}_R^P$ for the reflected *P* wave and $\mathbf{u}_R^{SV}$ for the reflected *SV* wave can then be expressed as

$$\mathbf{u}_0 = \mathbf{A}_0 \exp\left[i\omega\left(t - \frac{x\sin i_1 - z\cos i_1}{\alpha_1}\right)\right], \quad \mathbf{A}_0 = \left(A_0 \sin i_1, 0, -A_0 \cos i_1\right),$$

$$\mathbf{u}_R^P = \mathbf{A}_1 \exp\left[i\omega \left(t - \frac{x \sin e_1 + z \cos e_1}{\alpha_1} \right) \right], \quad \mathbf{A}_1 = \left(A_1 \sin e_1, 0, A_1 \cos e_1 \right),$$

$$\mathbf{u}_R^{SV} = \mathbf{B}_1 \exp\left[i\omega \left(t - \frac{x \sin f_1 + z \cos f_1}{\beta_1} \right) \right], \quad \mathbf{B}_1 = \left(B_1 \cos f_1, 0, -B_1 \sin f_1 \right). \tag{12.83}$$

This problem can be considered as the special case of the problem from Section 12.7 if the second medium is the vacuum, and the half-spaces in Fig. 12.5 are reversed (however, the orientation of the *z*-axis is still downwards). Therefore, omit the first two Zoeppritz equations (12.76a,b), because they concern the displacements, and put $A_2 = B_2 = 0$ in the remaining two equations. After dividing Eq. (12.76c) by W_1, and Eq. (12.76d) by Z_1, we arrive at the following equations for our problem:

$$A_1 \gamma_1 \sin 2e_1 + B_1 \cos 2f_1 = A_0 \gamma_1 \sin 2i_1,$$

$$A_1 \cos 2f_1 - B_1 \gamma_1 \sin 2f_1 = -A_0 \cos 2f_1, \tag{12.84}$$

where $e_1 = i_1$, angle f_1 is determined by Snell's law (12.62) and $\gamma_1 = \beta_1/\alpha_1$. Note that the orientation of the *z*-axis has not influenced the form of the final equations (12.84), although the *z*-components of the displacements in (12.63) and (12.83) are of the opposite signs.

The solution of Eqs. (12.84) may be expressed in the following simple form:

$$\frac{A_1}{A_0} = \frac{\gamma_1^2 \sin 2i_1 \sin 2f_1 - \cos^2 2f_1}{\gamma_1^2 \sin 2i_1 \sin 2f_1 + \cos^2 2f_1},$$

$$\frac{B_1}{A_0} = \frac{2\gamma_1 \sin 2i_1 \cos 2f_1}{\gamma_1^2 \sin 2i_1 \sin 2f_1 + \cos^2 2f_1}. \tag{12.85}$$

These are the desired formulae for the reflection coefficients when a *P* wave is incident at a free surface.

12.9.2 Reflection of *SV* waves

In a similar wave as in the preceding subsection, we can solve the problem for an incident *SV* wave (Fig. 12.9).

Denote the angle of incidence of the *SV* wave by j_1 and the corresponding displacement vector by

$$\mathbf{u}_0 = \mathbf{B}_0 \exp\left[i\omega \left(t - \frac{x \sin j_1 - z \cos j_1}{\beta_1} \right) \right], \quad \mathbf{B}_0 = \left(B_0 \cos j_1, 0, B_0 \sin j_1 \right). \tag{12.86}$$

The displacements for the reflected waves, $\mathbf{u}_R^P$ and $\mathbf{u}_R^{SV}$, are given by the same expressions as in (12.83).

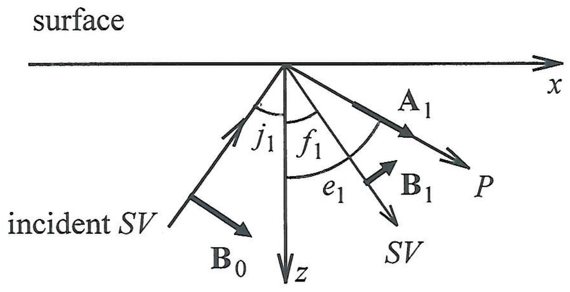


Fig. 12.9. Reflection of an *SV* wave at a free surface.

Now we shall simplify the Zoeppritz equations (12.82) for an incident *SV* wave. Omit again the first two equations, and put $A_2 = B_2 = 0$ in the remaining two equations. This yields the two following equations,

$$A_1\gamma_1 \sin 2e_1 + B_1 \cos 2f_1 = B_0 \cos 2j_1 \ ,$$
$$A_1 \cos 2f_1 - B_1\gamma_1 \sin 2f_1 = B_0\gamma_1 \sin 2j_1 \ , \qquad (12.87)$$

where $f_1 = j_1$, angle e_1 is determined by Snell's law (12.80) and $\gamma_1 = \beta_1/\alpha_1$.

The solution of Eqs. (12.87) may be expressed in the following simple form:

$$\frac{A_1}{B_0} = \frac{\gamma_1 \sin 4j_1}{\cos^2 2j_1 + \gamma_1^2 \sin 2j_1 \sin 2e_1} \ ,$$
$$\frac{B_1}{B_0} = \frac{\cos^2 2j_1 - \gamma_1^2 \sin 2j_1 \sin 2e_1}{\cos^2 2j_1 + \gamma_1^2 \sin 2j_1 \sin 2e_1} \ . \qquad (12.88)$$

These are the reflection coefficients for the case of an incident *SV* wave.

12.9.3 Motion at the free surface

We derived in Section 12.5 that an incident *SH* wave produced the motion of the free surface with the double amplitude.

In the case of an incident *P* or *SV* waves, the motion of the free surface also differs from the motion in the incident wave due to the contribution of the reflected waves. However, now the situation is more complicated, because two reflected waves contribute to the surface motion. Moreover, as opposed to the simple reflection coefficient $A = 1$ for an *SH* wave, the reflection coefficients (12.85) and (12.88) depend also on the angle of incidence.

In order to obtain the components of the displacement for a given superposition of waves at the Earth's surface, the incident wave must be multiplied by certain functions, called the *conversion coefficients*.

Incident *P* waves. For the case of an incident *P* wave, it follows from (12.83) that the resultant displacement at the surface $(z = 0)$ is

$$(\mathbf{u}_1)_{z=0} = \mathbf{U}_1 \exp\left[i\omega\left(t - \frac{x \sin i_1}{\alpha_1}\right)\right], \quad \mathbf{U}_1 = (U_1, 0, W_1), \tag{12.89}$$

where

$$\begin{aligned} U_1 &= A_0 \sin i_1 + A_1 \sin e_1 + B_1 \cos f_1 , \\ W_1 &= -A_0 \cos i_1 + A_1 \cos e_1 - B_1 \sin f_1 . \end{aligned} \tag{12.90}$$

Denote the denominators in the reflection coefficients (12.85) by

$$D = \gamma_1^2 \sin 2i_1 \sin 2f_1 + \cos^2 2f_1 , \tag{12.91}$$

and insert these coefficients into (12.90). After rearrangements, using Snell's law (12.62), we obtain

$$U_1 = \frac{2A_0 \cos i_1}{D} \sin 2f_1 , \quad W_1 = -\frac{2A_0 \cos i_1}{D} \cos 2f_1 . \tag{12.92}$$

Express the displacement vector $\mathbf{u}_0$ for the incident wave as $\mathbf{u}_0 = (u_0, 0, w_0)$. In view of (12.83), it then holds for the angle of incidence i_1 that

$$\tan i_1 = -\frac{u_0}{w_0} . \tag{12.93}$$

Angle i_1 is also called the actual angle of incidence. The analogous ratio for the resultant components of the surface motion determines another angle, ψ, which is called the *apparent angle of incidence*. This angle is thus defined by the relation

$$\tan \psi = -\frac{U_1}{W_1} . \tag{12.94}$$

By inserting (12.92) into this formula, we arrive at the simple relation $\tan \psi =$ $= \tan 2f_1$, i.e.

$$\psi = 2f_1 . \tag{12.95}$$

Consequently,

$$\cos \psi = \cos 2f_1 = 1 - 2\sin^2 f_1 = 1 - 2\left(\frac{\beta_1}{\alpha_1}\right)^2 \sin^2 i_1 .$$

This yields the important relation

$$\sin i_1 = \frac{\alpha_1}{\beta_1}\sqrt{\frac{1-\cos\psi}{2}} \ , \tag{12.96}$$

which is sometimes referred to as Wiechert's relation; see the references in Ewing et al. (1957). Hence, the actual angle of incidence i_1 can be determined from the observed displacements by means of (12.94) and (12.96).

Note that Wiechert's relation may be expressed in the following simpler form as

$$\sin i_1 = \frac{\alpha_1}{\beta_1}\sin\frac{\psi}{2} \ . \tag{12.97}$$

We could obtain this formula immediately from (12.95) as follows:

$$\sin\frac{\psi}{2} = \sin f_1 = \frac{\beta_1}{\alpha_1}\sin i_1 \ ,$$

Incident *SV* waves. For an incident *SV* wave, we get the resultant displacement at the free surface in the form

$$(\mathbf{u}_1)_{z=0} = \mathbf{U}_1 \exp\left[i\omega\left(t - \frac{x\sin j_1}{\beta_1}\right)\right], \quad \mathbf{U}_1 = (U_1, 0, W_1) \ , \tag{12.98}$$

where

$$\begin{aligned} U_1 &= B_0\cos j_1 + A_1\sin e_1 + B_1\cos f_1 \ , \\ W_1 &= B_0\sin j_1 + A_1\cos e_1 - B_1\sin f_1 \ , \end{aligned} \tag{12.99}$$

and A_1, B_1 are given by (12.88). After rearrangements one gets

$$U_1 = \frac{2B_0\cos j_1}{D}\cos 2j_1 \ , \quad W_1 = \frac{2B_0\cos j_1}{D}\gamma_1^2\sin 2e_1 \ , \tag{12.100}$$

where D is the denominator in the reflection coefficients (12.88), i.e.

$$D = \cos^2 2j_1 + \gamma_1^2\sin 2j_1\sin 2e_1 \ . \tag{12.101}$$

The apparent angle of incidence, in view of (12.86), is now

$$\tan\psi = \frac{W_1}{U_1} = \gamma_1^2\frac{\sin 2e_1}{\cos 2j_1} \ . \tag{12.102}$$

Note that instead of the angles of incidence, i_1 for *P* waves and j_1 for *SV* waves, many authors have used their complements, $I_1 = 90°-i_1$ and $J_1 = 90°-j_1$; see Bullen (1965), Ewing et al. (1957), Savarensky (1975). Angles I_1 and J_1 are called the *angles of emergence*.

Chapter 13

Ray Methods

A comprehensive description of the contemporary formulations of the ray method in seismology can be found in the lecture notes by Psencik (1994). More advanced methods, in particular the Gaussian beam method, are described in the lecture notes by Popov (1996). Consequently, in this chapter we shall discuss the ray method only briefly, emphasising its physical foundations, advantages, disadvantages, and some interesting aspect from the historical development of this method. Since the methods of computing seismic rays and travel times have already been described in Chapters 3 to 5, here we shall pay the main attention to the ray method for amplitudes.

13.1 Methods of Solving the Elastodynamic Equations for Inhomogeneous Media

The properties of elastic waves are described by the equations of motion of an elastic continuum (Chapter 6). These equations of motion are also called the *elastodynamic equations*. In the previous chapters we derived various exact solutions of these equations for homogeneous media, including the contact of homogeneous half-spaces. The solutions were expressed in the form of simple analytical formulae, in integral forms or in the form of infinite series.

Now we shall deal with the solution of the elastodynamic equations for inhomogeneous media. However, for general inhomogeneous media, analogous formulae for exact solutions are not know. In these cases we can use two approaches:

- To solve the elastodynamic equations *numerically*, e.g., by the finite difference (FD) or finite element (FE) methods. In principle, these numerical methods can yields accurate results, but cannot be applied to large models of the medium. For models whose dimensions exceed several wavelengths, these methods become extremely time consuming.
- To solve the elastodynamic equations *approximately*, e.g., by *ray methods*. We shall deal with these methods in this chapter.

However, we shall not consider quite general inhomogeneous media, but we shall restrict ourselves to media with slowly varying elastic parameters and with smooth interfaces. Therefore, we shall not consider diffraction phenomena at edges, corrugated boundaries, etc., or interference phenomena in thin layers, although the corresponding theories also exist. Under the words "slow variation" and "smooth interface" we shall mean certain comparisons with the prevailing wavelength. Namely, if the elastic parameters do not vary very much within a distance of one wavelength, we intuitively expect that such a medium may be considered as nearly homogeneous (quasi-homogeneous, locally homogeneous). The wavefield will then have similar properties as in homogeneous media, e.g., it will be approximately separated into the waves of longitudinal and transverse types. Analogously, if the wavelength is much shorter than the radius of curvature of an interface, in studying the reflection and transmission of waves it will be possible to approximate the curved interface by the tangent plane.

We shall thus restrict ourselves to short-wavelength approximations or, which is the same, to high-frequency approximations of wavefields. The ray method yields this type of approximate solutions.

The ray method was developed originally for the applications in optics, acoustics and radiophysics. Only later it was applied and developed for the purposes of seismology. This method is known under various names, such as the geometrical seismics (derived from "geometrical optics"), asymptotic ray method, the ray series method.

The ray method represents a powerful method of solving wave propagation problems in rather general inhomogeneous isotropic and/or anisotropic media. As the main *advantages* of the ray method we should point out the following properties:

- The ray method is applicable to inhomogeneous media.
- It is computationally effective and fast. For example, no evaluation of integrals is contained in the method (such as the complicated integrals of the Sommerfeld type).
- The wavefield is separated into individual waves, which increases the physical insight into the wave propagation process and facilitates its deeper understanding.
- It represents the basis for other related, more sophisticated methods, such as the Gaussian beam method, the paraxial ray method, the Maslov method, etc.

On the other hand, the ray method has also some *disadvantages* and limitations, in particular:

- It is only approximate. Moreover, the ray method itself does not allow to estimate the accuracy of the results, which represents a serious drawback of this method. Consequently, the results of the ray method must be compared with various "standard" solutions, i.e. with analytical solutions of simpler problems, or with the solutions obtained by the FD or FE methods.
- The ray method yields inaccurate results, or even physically implausible results, in some special regions, called singular regions (regions of caustics, of critical points, in transition zones between illuminated and shadow regions).
- It is applicable only to smooth media, in which the characteristic dimensions of inhomogeneities are considerably larger than the prevailing wavelength of the considered waves. Thus it cannot be used to study the diffraction of waves at edges, small inclusions in the medium, or corrugated boundaries.
- Since the wavefield in the ray method is separated into individual waves, the method cannot also be used to study interference waves (surface waves, guided waves, waves in thin layers).
- Some other types of waves cannot also be studied by the ray method. For example, head waves do not exist in the zero-approximation of the ray theory, and so higher approximations are needed for their study.

In order to increase the accuracy of the ray method and to overcome some of its limitations, more sophisticated methods, mentioned above, have been developed. These methods do not describe the wave propagation solely along an isolated ray, but also in a small vicinity of the ray. This corresponds better to the physical reality, e.g., to the Huygens principle.

13.2 Conservation Laws in the Physical World

The first formulations of the ray method did not proceed from the elastodynamic equations, but were based on other general laws of physics. A question therefore arises what are the relations between these different descriptions of the same wave phenomena. An excellent analysis of these relations in physical problems was given in the textbook by Kittel et al (1962), pp. 133-134, from which we reproduce the following passages:

"In the physical word there exist a number of conservation laws, some exact and some approximate. A conservation law is usually the consequence of some underlying symmetry in the universe. There are conservation laws relating to energy, linear momentum, angular momentum, charge, number of baryons ..., strangeness, and various other quantities ...

If all the forces in a problem are known, and if we are clever enough and have computers of adequate speed and capacity to solve for trajectories of all the particles, then the conservation laws give us no additional information. But they are very powerful tools which a physicist uses every day. Why are conservation laws powerful tools?

1. Conservation laws are independent of the details of the trajectory and, often, of the details of the particular force. The laws are therefore a way of stating very general and significant consequences of the equations of motion. A conservation law can sometimes assure us that something is impossible. Thus we do not waste time analyzing an alleged perpetual motion device ...

2. Conservation laws have been used even when the force is unknown; this applies particularly in the physics of elementary particles.

3. Conservation laws have an intimate connection with invariance. In the exploration of new and not yet understood phenomena the conservation laws are often the most striking physical fact we know. They may suggest appropriate invariance concepts ...

4. Even when the force is known exactly, a conservation law may be a convenient aid in solving for the motion of a particle. Many physicists have a regular routine for solving unknown problems: First we use the relevant conservation laws one by one; only after this, if there is anything left to the problem, will we get down to real work with differential equations, variational and perturbation methods, computers, intuition, and the other tools at our disposal."

We have paid much attention to these general problems, because the historical development of the ray method also followed some of these trends. We shall attempt to show it briefly below.

In order to demonstrate these relations, let us mention an example from mechanics. It is well known that the motion of a particle is described in full by Newton's Second Law,

$$m\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2}=\mathbf{F}\ , \tag{13.1}$$

if the acting force **F**, initial position and initial velocity are given. However, in solving many problems we use also the law of conservation of mechanical energy,

$$E_k+E_p=\text{const.}\ , \tag{13.2}$$

E_k being the kinetic energy and E_p the potential energy of the particle. Let us remind that the conservation law (13.2) represents a consequence of the equation of motion (13.1), i.e. it may be derived from this equation, if the force field is conservative (Kittel et al., 1962). Hence, Newton's Second Law is a more general law than the conservation of mechanical energy. The scalar equation (13.2) for the conservation of energy cannot compete with the vectorial equation of motion (13.1), which represents three scalar equations. Only in the cases of one-dimensional motions, these equations may be equivalent. In more complicated problems, the law of the conservation of energy can yield only the absolute value of the particle velocity, but not its direction.

We shall encounter a very similar situation in computing the amplitudes of seismic waves by the ray method.

13.3 Ray Approximations

Various approximate expressions for the displacement vector of elastic waves, which are used in the ray method, are called *ray approximations*. The ray approximations are used both in the frequency domain (for harmonic waves) and in the time domain.

The approximation in the frequency domain is more obvious physically. This corresponds to the assumption of a high frequency, $\omega\to\infty$. This approximation has widely been used in optics, acoustics, to study the electromagnetic waves in the ionosphere, etc. However, for direct applications in seismic prospecting, the ray approximation in the time domain is more important (Cerveny, 1978; Psencik, 1994). Nevertheless, here we shall pay attention especially to the ray approximation in the frequency domain, as it is simpler from the theoretical point of view (see also Popov (1996)). The ray approximation in the frequency domain was used for the first time by Sommerfeld and Rung in 1911 to study the propagation of electromagnetic waves.

As mentioned above, we shall restrict ourselves to the waves in inhomogeneous media which, at high frequencies, resemble waves in locally homogeneous media. Therefore, let us recapitulate the formulae for the main types of waves in homogeneous media, i.e. for plane, spherical and cylindrical waves. For example, the potential of longitudinal harmonic waves may be expressed as follows (see also Tab. 9.1 in Section 9.3):

1) For a plane wave propagating in the (x, z)-plane,

$$\varphi = Ae^{i\omega\left(t - \frac{x\sin\gamma + z\cos\gamma}{\alpha}\right)} . \tag{13.3}$$

2) For a spherical wave,

$$\varphi = \frac{A}{r} e^{i\omega\left(t - \frac{r}{\alpha}\right)} . \tag{13.4}$$

3) For a cylindrical wave,

$$\varphi = AJ_0(k\rho)e^{i\omega t} . \tag{13.5.a}$$

Replacing the Bessel function by the Hankel function and using its asymptotic expression for large arguments, we get

$$\varphi = A\sqrt{\frac{2}{\pi k\rho}} e^{i\omega\left(t - \frac{\rho}{\alpha} - \frac{\pi}{4}\right)} . \tag{13.5b}$$

Note that the factors A in these expressions are constants or functions of ω only.

It can be seen from these formulae that the displacement vector for harmonic P waves, but also for harmonic S waves, may be expressed in these cases in the following common form:

$$\mathbf{u}_H(x_m, t, \omega) = \mathbf{A}(x_m, \omega)e^{i\omega(t - \tau(x_m))} , \tag{13.6}$$

where the vector function $\mathbf{A}(x_m, \omega)$ is still dependent on ω. However, for large ω this function varies more slowly with the variations of coordinates than the corresponding exponential term. We can thus assume that this function tends to a finite limit for $\omega \to \infty$. Consequently, for large ω we may write approximately

$$\boxed{\mathbf{u}_H(x_m, t, \omega) = \mathbf{U}^0(x_m)e^{i\omega(t - \tau(x_m))}} . \tag{13.7}$$

This is the final form for the *displacement vector in the ray approximation* in the frequency domain.

It is evident that many objections could be risen against the considerations given above. Nevertheless, many important conclusions, with valuable practical applications, may be derived from formula (13.7). Moreover, this formula represents the better approximation, the higher is the frequency. In order to increase the accuracy of this formula, we could express function $\mathbf{A}(x_m, \omega)$ in the form of the asymptotic expansion in the powers of ω^{-1}:

$$\mathbf{A}(x_m, \omega) \sim \mathbf{U}^0(x_m) + \frac{1}{i\omega}\mathbf{U}^1(x_m) + \frac{1}{(i\omega)^2}\mathbf{U}^2(x_m) + \dots \,. \tag{13.8}$$

Analogous asymptotic series will be considered below. In the following two sections we shall use only the zero term, i.e. approximation (13.7). We shall speak of the *leading term of the ray series*, or of the *zero-order ray approximation*. Here we use the zero superscript with $\mathbf{U}^0$ only to obtain the consistent notation with the notation which will be introduced below.

We shall assume the vector $\mathbf{U}^0$ to be generally complex, but function τ to be real. Note that function τ is also assumed to be complex in some formulations of the ray method (in the ray theory with a complex eikonal), but we shall not consider these formulations here.

Vector $\mathbf{U}^0$ is called the *ray vector complex amplitude* or *ray vector complex amplitude coefficient (factor)*. Its absolute value, $|\mathbf{U}^0|$, is called the *ray amplitude*. The *m*-th component of vector $\mathbf{U}^0$, i.e. U_m^0, is referred to as the *ray complex amplitude* of the *m*-th component of displacement.

Function $\tau(x_1, x_2, x_3)$ is called the *eikonal* or the *phase function*. This is the travel time which the wavefront, passing through the coordinate origin, needs for arriving at the point with coordinates x_1, x_2, x_3.

We may expect the ray approximation to be a convenient approximation to the waves which are close to plane or spherical waves. This follows from the comparison of the ray approximation (13.7) with formulae (13.3) and (13.4). However, the application of (13.7) to approximate cylindrical waves may be more problematic. Namely, the amplitude in the cylindrical wave (13.5b) tends to zero for $\omega \to \infty$ $(k \to \infty)$, but we are interested in a non-zero limit $\mathbf{U}^0$.

13.4 Energy of Elastic Waves

The density of mechanical energy is defined by

$$E = E_k + E_p \,, \tag{13.9}$$

where E_k is the density of kinetic energy and E_p the density of potential energy. For small deformations of a continuum, these densities read

$$E_k = \frac{1}{2}\rho v_i v_i \,, \quad E_p = \frac{1}{2}\sigma_{ij} e_{ij} \,, \tag{13.10}$$

where ρ is the density of the medium, $v_i = \dot{u}_i$ are the components of velocity, σ_{ij} the components of the stress tensor and e_{ij} the components of the tensor of infinitesimal strain. Here we have denoted the stress tensor by σ_{ij}, in order to avoid the confusion of the previous notation τ_{ij} with the eikonal τ. We shall not derive the above-mentioned expression for E_p; this derivation is not elementary, but may be found in many textbooks on continuum mechanics, see also Aki and Richards (1980).

Let us calculate the mechanical energy of a plane harmonic wave propagating in a homogeneous medium (see the analogous problem in Section 1.3). Assume that the wave propagates along the x-axis. The displacement may then be expressed as

$$u = A\sin(\omega t - kx) \ , \tag{13.10}$$

where the usual notations have been used. Since the particle velocity is

$$v = \dot{u} = \omega A\cos(\omega t - kx) \ , \tag{13.11}$$

we get the instantaneous density of the kinetic energy in the form

$$E_k = \frac{1}{2}\rho\omega^2 A^2 \cos^2(\omega t - kx) \ . \tag{13.12}$$

Instead of calculating the potential energy according to the formula given above, we shall apply the law of the conservation of mechanical energy. This energy should be equal to the maximum kinetic energy (because $E_p = 0$ for $E_k = \max.$). This yields the following formula for the energy density in a plane harmonic wave:

$$E = \frac{1}{2}\rho\omega^2 A^2 \ . \tag{13.13}$$

The latter formula has frequently been used generally, without a deeper analysis, also for arbitrary waves in arbitrary inhomogeneous media. However, for the ray approximation in an inhomogeneous medium, an analogous formula really holds true (Cerveny, 1978; Psencik,1994). Therefore, in the ray approximation we may write

$$E = \frac{1}{2}\rho\omega^2 \left|\mathbf{U}^0\right|^2 \ . \tag{13.14}$$

In Section 13.6 we shall only need that the energy density is proportional to the density of the medium and to the quadratic amplitude, $E \sim \rho\left|\mathbf{U}^0\right|^2$.

13.5 Ray Coordinates and Ray Tubes

In many problems of the ray theory, it is convenient to introduce so-called *ray coordinates*,

$$(s, \gamma_1, \gamma_2) . \tag{13.15}$$

Parameter s specifies the position of a point on a selected ray. This may be an arbitrary parameter along the ray. Usually this is the length of the ray (length of the arc) measured from some reference point, where we put $s = 0$. In the case of a point source, coordinate s may be the length of the ray from this source. Instead of a length we may also use the travel time τ, and then we write the ray coordinates as

$$(\tau, \gamma_1, \gamma_2) . \tag{13.16}$$

The remaining two ray coordinates, γ_1 and γ_2, specify the selected ray. They are usually introduced in two ways:

a) In the case of a point source, we may use the spherical coordinates δ_0 and φ_0 which determine the direction of the ray at the source (Fig. 13.1). Angle δ_0 is usually measured from the vertical towards the ray, and so $0 \le \delta_0 \le \pi$. Angle φ_0 is then the angle in the horizontal plane ("geographical longitude"), varying in the interval $0 \le \varphi_0 \le 2\pi$.

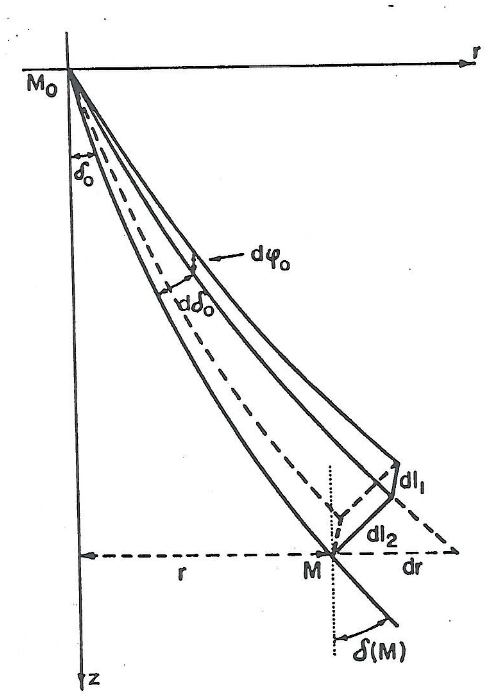


Fig. 13.1. Ray coordinates for a point source and a ray tube. (After Cerveny (1978)).

b) We may select an arbitrary wavefront (for some reference time τ_0), and choose γ_1 and γ_2 as curvilinear coordinates on this wavefront (Fig. 13.2). Quantities q_1 and q_2 are also called the *ray parameters*.

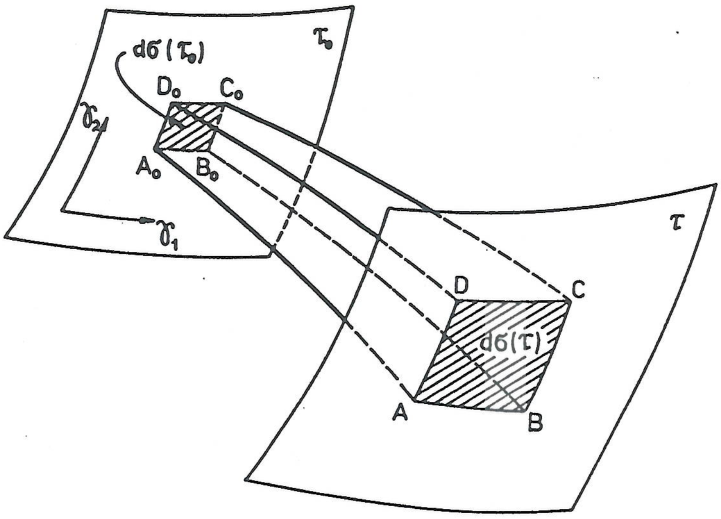


Fig. 13.2. Ray parameters on a references wavefront and a ray tube. (After Cerveny (1978)).

A very important quantity in the ray theory is also the Jacobian of the transformation from Cartesian coordinates (x, y, z) to ray coordinates (s, γ_1, γ_2). We shall denote this determinant by J,

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial s} & \dfrac{\partial y}{\partial s} & \dfrac{\partial z}{\partial s} \\ \dfrac{\partial x}{\partial \gamma_1} & \dfrac{\partial y}{\partial \gamma_1} & \dfrac{\partial z}{\partial \gamma_1} \\ \dfrac{\partial x}{\partial \gamma_2} & \dfrac{\partial y}{\partial \gamma_2} & \dfrac{\partial z}{\partial \gamma_2} \end{vmatrix} . \tag{13.17}$$

We shall speak of a regular ray field in a region if $J \neq 0$ at every point of the region. The ray field will be called irregular at a point M if $J(M) = 0$. In the region where $J \neq 0$, the relation between coordinates (s, γ_1, γ_2) and (x, y, z) is unique. In this region, just one ray passes through a given point. In other words, the rays in this region neither cross one another, nor a "ray shadow" is formed.

Note that the well-known Jacobian for the transformation between Cartesian and spherical coordinates has the form $J = r^2 \sin\vartheta$; see Chapter 9. The reader may verify this result by substituting (9.1) into the determinant analogous to (13.17). For an infinitesimal volume element in spherical coordinates we then have $\mathrm{d}V = J \,\mathrm{d}r\, \mathrm{d}\vartheta\, \mathrm{d}\lambda$.

Another important notion in the ray theory is the *ray tube*. We shall define the ray tube as the set of rays, the parameters of which are within the intervals

$$\left(\gamma_1, \gamma_1 + \mathrm{d}\gamma_1\right), \quad \left(\gamma_2, \gamma_2 + \mathrm{d}\gamma_2\right), \tag{13.18}$$

see Fig. 13.2. Since we consider infinitesimal quantities $\mathrm{d}\gamma_1$ and $\mathrm{d}\gamma_2$, we also speak of an elementary ray tube. Ray tubes are used in studying ray amplitudes.

13.6 Ray Method for Amplitudes – Energetic Approach

Preliminary information about seismic amplitudes can be obtained from ray diagrams. For example, consider the model of a vertically inhomogeneous medium as shown in Fig. 13.3. The corresponding ray diagrams, computed with a constant step of the angle at the source, are shown in Fig. 13.4.

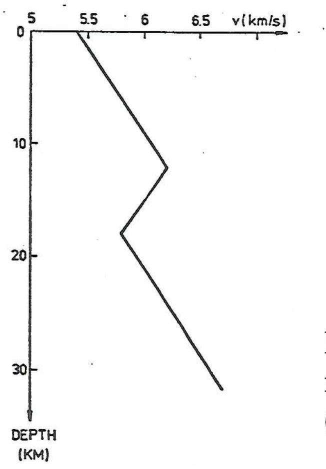


Fig. 13.3. Velocity cross-section for a model with a low-velocity channel. (After Cerveny (1978)).

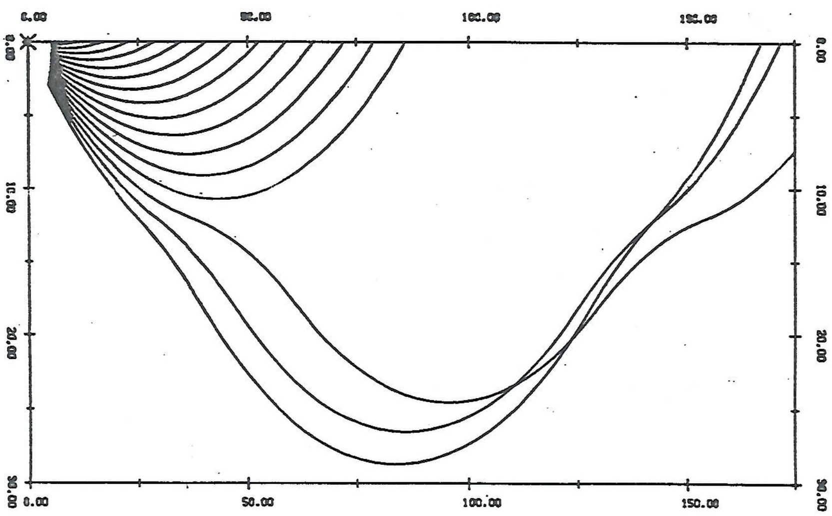


Fig. 13.4a. Ray diagram for the model in Fig. 13.3 and a surface source. (After Cerveny (1978)).

The regions of concentrated seismic rays indicate the places of the concentration of seismic energy and, consequently, of large amplitudes.

Conversely, small amplitudes of seismic waves may be expected in the regions where the seismic rays are sparse.

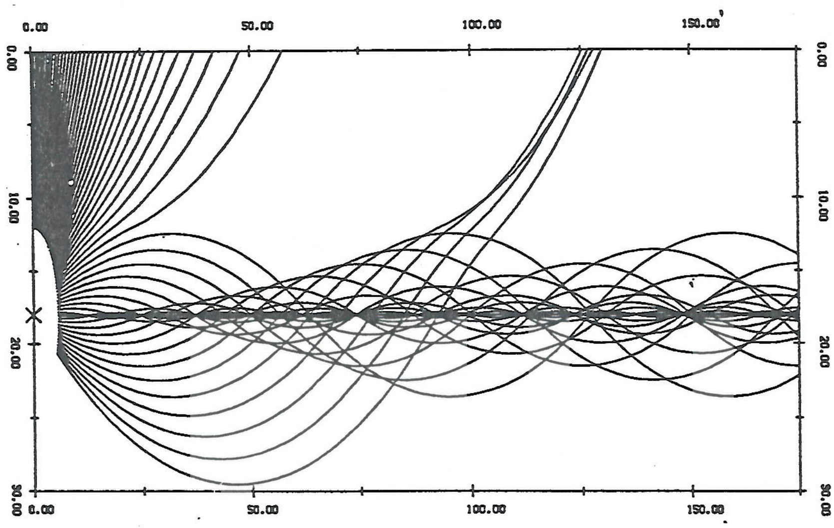


Fig. 13.4b. Ray diagram for the model in Fig. 13.3 and a source in the low-velocity channel. (After Cerveny (1978)).

The previous considerations on seismic amplitudes were only qualitative, but we need quantitative estimates. For this reason, let us consider the propagation of the mechanical energy of elastic waves.

13.6.1 Basic assumptions of the propagation of energy

Let us adopt the following assumptions concerning the wave propagation in the ray approximation:

I) For high frequencies, the wavefield in inhomogeneous media is *separated into P and S waves*, propagating at velocities $\alpha = \alpha(x, y, z)$ and $\beta = \beta(x, y, z)$, respectively. We thus assume, in the ray approximation, an analogous separation of the wavefield as in homogeneous media.

II) *Mechanical energy propagates along the rays.* Energy thus does not penetrate through the surface of a ray tube. The part of energy, which was radiated by the source into a wave tube, remains in this tube.

The application of the above-mentioned postulate of the independent propagation of *P* and *S* waves, together with the postulate on the energy propagation, makes it possible to calculate the amplitudes of seismic waves in inhomogeneous media, including general three-dimensional media.

Nevertheless, the energetic considerations are not sufficient to answer all questions about the propagation of seismic waves. If we are interested in amplitudes only, we shall obtain the comprehensive answer, at least for *P* waves. However, in constructing ray synthetic seismograms, these results are not sufficient. One encounters the following limitations of the energetic approach:

a) Amplitudes must be considered as complex quantities (they contain, e.g., phase shifts due to total reflections). In computing the form of a signal, we need to know both the real and imaginary parts of the complex amplitude. The energetic approach will yield only the absolute value of this complex amplitude.
b) The amplitude factor is generally a vector quantity. In the case of *P* waves we know that the displacement has the direction of the ray. However, in the case of *S* waves, the direction is not known in full, we know only that this direction is perpendicular to the ray. In other words, the energetic approach does not allow us to determine the polarisation of *S* waves.

In spite of these disadvantages, the energetic approach yields satisfactory result in many problems. Moreover, this also gives a physical insight into more complicated results obtained from the elastodynamic equations.

13.6.2 Energy flux and a general formula for amplitudes

We assume that seismic energy neither leaks through the surfaces of ray tubes, nor is transformed into other forms of energy. For harmonic waves, which we consider here, this means that the energy flux through various cross-sections of the same ray tube is identical.

Consider an arbitrary ray, e.g., the ray which passes through points A_0 and A in Fig. 13.2. Consider further an elementary tube which surrounds this ray, or is tangent to it. Let $\mathrm{d}\sigma$ be an arbitrary section of the ray tube which is perpendicular to the rays. The energy flux through this cross-section is then

$$\widetilde{E} = Ev\,\mathrm{d}\sigma, \tag{13.19}$$

where E is the energy density, and v the wave velocity. We must thus put $v = \alpha$ for P waves, and $v = \beta$ for S waves. Let us restrict ourselves to P waves. The conservation of the energy flux mentioned above, together with formula (13.14) for the energy density, then yield the conclusion that

$$\rho\alpha\,\mathrm{d}\sigma\left|\mathbf{U}^0\right|^2 = \mathrm{const} \tag{13.20}$$

along the ray tube.

Let us specify the latter equation for points A_0 and A with ray coordinates s_0 and s, respectively (see Fig. 13.2 where coordinates τ_0 and τ are indicated). We obtain the following general formula

$$\boxed{\left|\mathbf{U}^0(s)\right| = \sqrt{\left|\frac{\mathrm{d}\sigma(s_0)}{\mathrm{d}\sigma(s)}\right| \frac{\alpha(s_0)\rho(s_0)}{\alpha(s)\rho(s)}}\left|\mathbf{U}^0(s_0)\right|} \,. \tag{13.21}$$

For S waves we must replace α by β in Eqs. (13.20) and (13.21).

Note that the cross-section $d\sigma$ is sometimes introduced in such a way that it may attain not only positive, but also negative values. For this reason, in formula (13.21) we have introduced the absolute values of these quantities.

If the amplitude of a harmonic wave is known for some reference ray coordinate s_0 (e.g., close to the seismic source), formula (13.21) makes it possible to calculate the amplitudes along the corresponding ray.

The calculation of the ratio $d\sigma(s_0)/d\sigma(s)$ usually represents the most complicated step in determining seismic amplitudes by means of formula (13.21). The simplest way is, in principle, to estimate this ratio numerically, i.e. by replacing infinitesimal surface elements by finite ones. However, one ray is not sufficient for this purpose, but at least three rays are needed in the case of a 3–D medium, and at least two rays in a 2–D medium. In a 3–D medium we must determine three rays for parameters (γ_1, γ_2), $(\gamma_1 + \Delta\gamma_1, \gamma_2)$ and $(\gamma_1, \gamma_2 + \Delta\gamma_2)$, where $\Delta\gamma_1$ and $\Delta\gamma_2$ are small increments. In a 2–D medium, two rays for parameters γ_1 and $\gamma_1 + \Delta\gamma_1$ are sufficient.

Other methods of computing the ratio of these elementary surfaces will be mentioned below. For a vertically inhomogeneous medium and a point source, this ratio can be determined from simple geometrical considerations (Section 13.7).

13.6.3 Another form of the general formula for amplitudes. The Jacobian *J*

A volume element in ray coordinates can be expressed as

$$dV = J\, ds\, d\gamma_1\, d\gamma_2 \, , \tag{13.22}$$

where J is the determinant of the transformation (the Jacobian) from Cartesian to ray coordinates; see the analogous case of spherical coordinates, mentioned in Section 13.5. Determinant J is given by formula (13.17).

A volume element in ray coordinates can also be expressed as $dV = d\sigma\, ds$, where $d\sigma$ is the area of the corresponding cross-section. By comparing this expression with (13.22) we obtain

$$d\sigma = J\, d\gamma_1\, d\gamma_2 \, . \tag{13.23}$$

Since increments $d\gamma_1$ and $d\gamma_2$ are constant along the ray tube, we get

$$d\sigma(s_0) = J(s_0)\, d\gamma_1\, d\gamma_2 \, , \quad d\sigma(s) = J(s)\, d\gamma_1\, d\gamma_2 \, . \tag{13.24}$$

This makes it possible to replace the ratio of the infinitesimal surface elements in (13.21) by the ratio of the Jacobians at the corresponding points:

$$\left|\mathbf{U}^0(s)\right| = \sqrt{\left|\frac{J(s_0)}{J(s)}\right| \frac{\alpha(s_0)\rho(s_0)}{\alpha(s)\rho(s)}} \left|\mathbf{U}^0(s_0)\right| . \tag{13.25}$$

In addition to the definition (13.17), many methods have been suggested for computing the Jacobian J, such as the computation of the radii of curvature of the ray, or solving systems of additional ordinary differential equations. A special formula for function J will be derived in the next section.

13.7 Ray Amplitudes of Waves Generated by a Point Source

In deriving formulae (13.21) and (13.25), no assumptions of the source and medium were introduced. Both the source and the medium might be quite arbitrary. In this section we shall specify these formulae for a point source in a general inhomogeneous medium, and for a point source in a vertically inhomogeneous medium.

13.7.1 Point source in a general medium. The geometrical spreading

For a point source, it is convenient to express the formulae for amplitudes in another form. To simplify the problem, we shall consider the following, slightly idealised situation.

Consider a point source which generates, say, P waves. Surround the source with a spherical surface of a sufficiently large radius, so that the whole surface will lie in the "elastic zone", where the elasticity theory may be applied. Assume that this vicinity of the source may be regarded as *locally homogeneous* (although the source may be located in a general inhomogeneous medium). Without loss of generality we may put the radius of the sphere to be equal to unity. Any point of the surface is described by two spherical coordinates, δ_0 and φ_0; see Fig. 13.1. Angle δ_0 is usually measured from the vertical towards the ray, and so it lies within the interval $0 \le \delta_0 \le \pi$. Angle φ_0 is the angle in the horizontal plane ("geographical longitude"), and so $0 \le \varphi_0 \le 2\pi$. Since a surface element on the sphere of radius r is $\mathrm{d}\sigma = r^2 \sin\delta_0 \,\mathrm{d}\delta_0 \,\mathrm{d}\varphi_0$, for the unit sphere we get

$$\mathrm{d}\sigma_0 = \sin\delta_0 \,\mathrm{d}\delta_0 \,\mathrm{d}\varphi_0 . \tag{13.26}$$

It then follows from (13.24) that $J(s_0) = \sin\delta_0$ on the unit sphere.

Moreover, assume that the distribution of the ray vector amplitude $\mathbf{U}^0$ on the unit sphere is known as a function of δ_0 and φ_0. Denote this distribution by function $\mathbf{g}_P(\delta_0, \varphi_0)$, called *complex vector directional radiation characteristics* of P waves. Formula (13.25) can then be expressed as

$$\left|\mathbf{U}^0(s)\right| = \sqrt{\frac{\sin\delta_0\, \alpha(s_0)\rho(s_0)}{|J(s)|\alpha(s)\rho(s)}}\left|\mathbf{g}_P(\delta_0, \varphi_0)\right| . \tag{13.27}$$

If the reference wavefront is the unit sphere $(s_0 = 1)$, the following function is frequently introduced:

$$L(s) = \sqrt{\frac{\mathrm{d}\sigma(s)}{\mathrm{d}\sigma(s_0)}} = \sqrt{\frac{J(s)}{J(s_0)}} = \sqrt{\frac{J(s)}{\sin\delta_0}} . \tag{13.28}$$

Function L is referred to as the *geometrical spreading of a wavefront*. Using this notation, formula (13.27) takes the form

$$\boxed{\left|\mathbf{U}^0(s)\right| = \frac{\left|\mathbf{g}_P(\delta_0, \varphi_0)\right|}{|L(s)|}\sqrt{\frac{\alpha(s_0)\rho(s_0)}{\alpha(s)\rho(s)}}} . \tag{13.29}$$

This is the final formula for the amplitudes of P waves in an inhomogeneous medium, in the case of a point source.

The final formula for S waves is similar to (13.29), only instead of velocity α we must write β, and instead of $\mathbf{g}_P$ write the directional characteristics for S waves, $\mathbf{g}_S(\delta_0, \varphi_0)$. Note that vector $\mathbf{g}_P$ is perpendicular to the unit sphere, whereas vector $\mathbf{g}_S$ is tangent to this surface.

Note that the definition of the geometrical spreading must be modified if there are interfaces in the medium.

As a very simple example, consider the body waves generated by a point source in a homogeneous medium. At a distance R from the source, one gets $\mathrm{d}\sigma = R^2 \sin\delta_0\, \mathrm{d}\delta_0\, \mathrm{d}\varphi_0$, i.e. $J = R^2 \sin\delta_0$. From (13.28) we then get $L = R$, and formula (13.29) yields

$$\left|\mathbf{U}^0(s)\right| = \frac{1}{R}\left|\mathbf{g}_P(\delta_0, \varphi_0)\right| . \tag{13.30}$$

This is the well-known formula for the amplitudes of spherical waves in a homogeneous medium.

Another special case will be considered in the following subsection.

13.7.2 Point source in a vertically inhomogeneous medium

Consider the situation as shown in Fig. 13.1. For the surface element at point M we have $\mathrm{d}\sigma(M) = \mathrm{d}l_1 . \mathrm{d}l_2$, where the length of the horizontal edge is

$$\mathrm{d}l_1 = r\, \mathrm{d}\varphi_0 ,$$

and the length of the edge in the vertical plane is

$$\mathrm{d}l_2 = \mathrm{d}r \cos\delta(M) = \frac{\partial r}{\partial \delta_0} \cos\delta(M)\, \mathrm{d}\delta_0 \, .$$

Note that here we must write the partial derivative $\partial r / \partial \delta_0$, because the horizontal distance depends not only on angle δ_0, but also on depth z (or ray coordinate s). Formulae (13.23) and (13.28) then yield

$$J = r \frac{\partial r}{\partial \delta_0} \cos\delta(M) \, , \tag{13.31}$$

and

$$L = \sqrt{r \frac{\partial r}{\partial \delta_0} \frac{\cos\delta(M)}{\sin\delta_0}} \, . \tag{13.32}$$

By inserting the latter formula into (13.29), we arrive at the formula for the ray amplitude at an arbitrary point of the ray. Angle $\delta(M)$ in formula (13.32) can be determined from Snell's law, and formulae for the horizontal distance r were derived in Chapter 3.

The derivative $\partial r / \partial \delta_0$ can be estimated by numerical differentiation, or by analytically differentiating the corresponding formulae in Chapter 3. Note that the expression (13.32) for the geometrical spreading becomes indeterminate at the turning point (deepest point) of a ray. Namely, at this point we have $\delta(M) = 90°$, $\cos\delta(M) = 0$, but the derivative $\partial r / \partial \delta_0$ becomes infinite. Hence, a modified formula for the geometrical spreading must be used in a vicinity of the turning point.

13.8 Principles of the Method of Ray Series

In this section we shall briefly describe a more advanced method of solving the wave propagation problems in inhomogeneous media which differs from the previous approaches in the following substantial aspects:

1) the method proceeds directly from the elastodynamic equations;
2) the solution is not sought in the form of one term, but in the form of a series (although only the leading term of the series is predominantly used in practice).

We may expect that this method will be more general and more accurate than the methods described above. In particular, the main principles and assumptions on which the previous approaches were based, such as Fermat's principle, should follow from this more general method as its consequences.

The ray series solutions of the elastodynamic equations were first suggested by Babich and his colleagues in Russia (see, e.g., Babich and Alekseev (1958)). Nearly at the same time, an analogous approach was used by Karal and Keller (1959) in the USA.

13.8.1 Ray series in the frequency domain

We shall seek the displacement vector for a harmonic wave in the form of the so-called *deformed plane wave*,

$$\mathbf{u}(x_m, t, \omega) = \mathbf{U}(x_m, \omega) e^{-i\omega(t - \tau(x_m))} , \tag{13.33}$$

which is practically identical with the form (13.6). As opposed to a plane wave, the amplitude coefficient $\mathbf{U}$ is not constant now, but depends on coordinates, and the phase function τ is not a linear function of coordinates, but is more complicated. Moreover, the amplitude coefficient may be a function of angular frequency ω. Note that, as opposed to (13.6), here we shall use the negative sign in the argument of the exponential (in order to obtain the positive sign in the term $e^{i\omega\tau}$). The choice of this sign is not substantial, but the change of the sign leads to the complex conjugate results, which must be taken into consideration.

In the previous sections we approximated the amplitude coefficient by a high-frequency limit (we used the zero-order ray approximation). The approximation by a series, which will be used here, should yield more accurate results.

Let a function $f(\omega)$ be continuous together with its higher derivatives in a vicinity of the origin. The corresponding Taylor expansion of this function,

$$f(\omega) = f(0) + \frac{f'(0)}{1!}\omega + \frac{f''(0)}{2!}\omega^2 + \dots ,$$

is then convergent in the mentioned vicinity. If ω is the angular frequency, this formula represents the low-frequency expansion of function *f*. Replacing ω by $1/\omega$ we get the corresponding high-frequency expansion (for large ω),

$$f(\omega) = f(\infty) + \frac{f'(\infty)}{1!}\frac{1}{\omega} + \frac{f''(\infty)}{2!}\frac{1}{\omega^2} + \dots , \tag{13.34}$$

where $f(\infty)$, $f'(\infty)$, etc., are the limits for $\omega \to \infty$.

Analogously to (13.34), we shall seek the amplitude coefficient in (13.33) in the form

$$\mathbf{U}(x_m, \omega) = \mathbf{U}^0(x_m) + \frac{1}{(-i\omega)}\mathbf{U}^1(x_m) + \frac{1}{(-i\omega)^2}\mathbf{U}^2(x_m) + \dots , \tag{13.35}$$

where $\mathbf{U}^0$, $\mathbf{U}^1$, etc., are unknown functions which are to be determined.

However, there is a substantial difference between formulae (13.34) and (13.35). Namely, the coefficients in series (13.34) are expressed in terms of the derivatives of function *f*. Since this series is convergent, the more terms of the series are used, the more accurate approximation will be obtained. As opposed to this, function **U** is unknown, and we do not know in advance whether the unknown functions $\mathbf{U}^k \left(k = 0, 1, \ldots\right)$ will be related to the corresponding derivatives of **U**. Consequently, such a series may not be convergent, and rather has the character of the so-called asymptotic series (Popov, 1996). In summing up such a series, the accuracy generally increases only to some term, and by adding further terms, the approximation becomes even worse. This situation is schematically shown in Fig. 13.5. The accuracy of the series may only be increased be passing to higher values of ω.

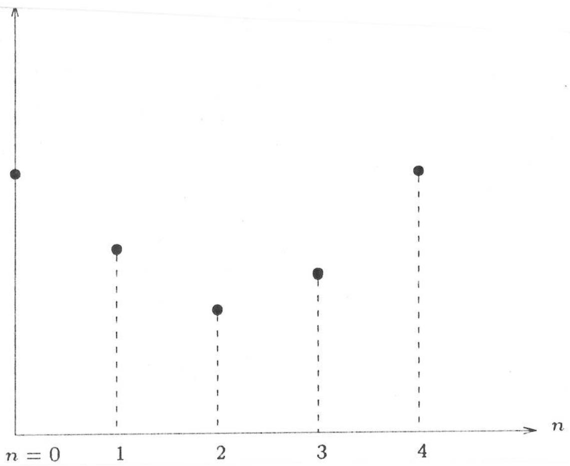


Fig. 13.5. Typical behaviour of errors if a function is approximated by an asymptotic series: n is the number of terms of the series, and the errors are plotted on the vertical axis. (After Popov (1996)).

Since we shall write the elastodynamic equations in components, we shall also express the displacements in components. We shall thus express the i-th component of the displacement for a harmonic wave in the form of the following *ray series*:

$$\boxed{u_i\left(x_m, t, \omega\right) = e^{-i\omega\left(t - \tau\left(x_m\right)\right)} \sum_{n=0}^{\infty} \frac{U_i^n\left(x_m\right)}{\left(-i\omega\right)^n}} . \qquad (13.36)$$

Since the ray series generally has the character of an asymptotic series, only a few terms of the series are usually used. The higher will be the angular frequency ω, the fewer terms will be sufficient. For high ω we may neglect all terms with the exception of the first term, i.e. the term for $n = 0$. The ray series then reduces to the ray approximation (we also speak of the zero-approximation of the ray theory).

13.8.2 Equations of motion for an inhomogeneous medium

The equations of motion for an elastic medium are also called the elastodynamic equations. We shall use the equations of motion of a continuum without body forces, i.e.

$$\frac{\partial \sigma_{ij}}{\partial x_j} = \rho \frac{\partial^2 u_i}{\partial t^2} , \qquad (13.37)$$

where σ_{ij} are the components of the stress tensor. Let us restrict ourselves to isotropic media described by Hooke's law

$$\sigma_{ij} = \lambda \vartheta \delta_{ij} + 2\mu e_{ij} , \qquad (13.38)$$

where

$$\vartheta = \operatorname{div} \mathbf{u} = \frac{\partial u_k}{\partial x_k}, \quad e_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) . \qquad (13.39)$$

Inserting these expressions into (13.37) yields

$$(\lambda + \mu)\frac{\partial^2 u_j}{\partial x_i \partial x_j} + \mu \frac{\partial^2 u_i}{\partial x_j^2} + \frac{\partial \lambda}{\partial x_i}\frac{\partial u_j}{\partial x_j} + \frac{\partial \mu}{\partial x_j}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) = \rho \frac{\partial^2 u_i}{\partial t^2} . \qquad (13.40)$$

This is the equation of motion for an isotropic inhomogeneous medium.

Let us derive also the simpler equations for the acoustic case. Introduce the pressure $p = -\lambda \vartheta$, and put $\mu = 0$. Hooke's law then yields the well-know formula for liquids, $\sigma_{ij} = -p\delta_{ij}$, which is also known as Pascal's law. Consequently, the equation of motion (13.37) simplifies to read

$$-\frac{\partial p}{\partial x_i} = \rho \frac{\partial^2 u_i}{\partial t^2} . \qquad (13.41)$$

Differentiate this equation with respect to x_i, and neglect the derivatives of density ρ:

$$-\frac{\partial^2 p}{\partial x_i^2} = \rho \frac{\partial^2 \vartheta}{\partial t^2} .$$

By inserting $\vartheta = -p/\lambda$, we arrive at the wave equation for pressure p,

$$\nabla^2 p = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} , \qquad (13.42)$$

where the velocity $c = \sqrt{\lambda/\rho}$ is a function of coordinates, $c = c(x_m)$. For harmonic waves, the latter equation yields the Helmholtz equation,

$$\nabla^2 p + k^2 p = 0 , \qquad (13.43)$$

where the wavenumber $k = \omega/c$ is again a function of coordinates (and of ω).

After neglecting the derivatives of density, we have arrived at the wave equation for an inhomogeneous medium. This equation was also considered by Popov (1996). The more general equation, including the derivatives of density, was considered by Psencik (1994). However, in solving the corresponding equations by the ray method, both results are very similar.

13.9 Equations of the Ray Methods for the Acoustic Case

We shall solve the wave equation (13.42) approximately by the method of ray series in the frequency domain. Analogously to the ray series (13.36), we shall seek the solution in the form

$$p(x_m, t, \omega) = e^{-i\omega(t-\tau(x_m))} \sum_{n=0}^{\infty} \frac{P^n(x_m)}{(-i\omega)^n} ; \tag{13.44}$$

note that the letter n with P^n means again a superscript, not a power. In this way, our task is reduced to the determination of the eikonal $\tau(x_m)$ and the amplitude coefficients $P^n(x_m)$.

We shall insert the trial solution (13.44) into the wave equation (13.42), or into the Helmholtz equation (13.43). First, calculate the derivative

$$\frac{\partial p}{\partial x_i} = e^{-i\omega(t-\tau)} \sum_{n=0}^{\infty} \frac{1}{(-i\omega)^n} \left[i\omega \frac{\partial \tau}{\partial x_i} P^n + \frac{\partial P^n}{\partial x_i} \right] . \tag{13.45}$$

Differentiating this expression with respect to x_i once more, one gets $\nabla^2 p$. The Helmholtz equation, after omitting the exponential term, then yields

$$\sum_{n=0}^{\infty} \frac{1}{(-i\omega)^n} \left[i\omega \frac{\partial \tau}{\partial x_i} \left(i\omega \frac{\partial \tau}{\partial x_i} P^n + \frac{\partial P^n}{\partial x_i} \right) + \right.$$
$$\left. + i\omega \left(\frac{\partial^2 \tau}{\partial x_i^2} P^n + \frac{\partial \tau}{\partial x_i} \frac{\partial P^n}{\partial x_i} \right) + \frac{\partial^2 P^n}{\partial x_i} + \frac{\omega^2}{c^2} P^n \right] = 0 . \tag{13.46}$$

This equation can be rewritten as

$$\sum_{n=0}^{\infty} \frac{1}{(-i\omega)^n} \left\{ (i\omega)^2 \left[\frac{\partial \tau}{\partial x_i} \frac{\partial \tau}{\partial x_i} - \frac{1}{c^2} \right] P^n + \right.$$
$$\left. + (i\omega) \left[2 \frac{\partial \tau}{\partial x_i} \frac{\partial P^n}{\partial x_i} + \frac{\partial^2 \tau}{\partial x_i^2} P^n \right] + \frac{\partial^2 P^n}{\partial x_i^2} \right\} = 0 . \tag{13.47}$$

13.9.1 Eikonal equation

In order to satisfy Eq. (13.47), the coefficients with the individual powers of ω must be zero. The highest power appearing in the series is ω^2. The term with this power yields the *eikonal equation*,

$$\boxed{\frac{\partial \tau}{\partial x_i}\frac{\partial \tau}{\partial x_i}=\frac{1}{c^2}}, \tag{13.48}$$

where $c = c(x_1, x_2, x_3)$. Other forms of this equation are

$$\left(\frac{\partial \tau}{\partial x_1}\right)^2+\left(\frac{\partial \tau}{\partial x_2}\right)^2+\left(\frac{\partial \tau}{\partial x_3}\right)^2=\frac{1}{c^2}, \tag{13.49}$$

or

$$(\mathrm{grad}\,\tau)^2=\frac{1}{\mathrm{c}^2}. \tag{13.50}$$

From the mathematical point of view, the eikonal equation is a nonlinear, first-order partial differential equation for $\tau = \tau(x_m)$. This can be solved, e.g., by using rays. As opposed to the simple case of a homogeneous medium, here the velocity c may vary with coordinates.

As a special case, let us solve the eikonal equation for a homogeneous medium. It now follows from Eq. (13.49) that the sum of squares of the first derivatives of τ must be equal to a constant. This equation can thus be satisfied, e.g., if eikonal τ is a linear function of Cartesian coordinates:

$$\tau = a_1x_1 + a_2x_2 + a_3x_3 ,$$

where

$$a_1^2 + a_2^2 + a_3^2 = c^{-2} .$$

These equations will be satisfied if τ is of the form

$$\tau=\frac{x_1\cos\alpha_1+x_2\cos\alpha_2+x_3\cos\alpha_3}{c}, \tag{13.51}$$

where α_1, α_2, α_3 are directional cosines, satisfying the condition

$$\cos^2\alpha_1+\cos^2\alpha_2+\cos^2\alpha_3=1 .$$

The expression (13.51) for τ describes a plane wave. We have thus arrived at the result that a plane wave represents one solution of the eikonal equation for a homogeneous medium.

13.9.2 Transport equations

Let us go back to Eq. (13.47). In view of the eikonal equation, the term in the first square bracket in (13.47) vanishes, which yields

$$\sum_{n=0}^{\infty} \frac{1}{(-i\omega)^n} \left\{ (i\omega) \left[2 \frac{\partial \tau}{\partial x_i} \frac{\partial P^n}{\partial x_i} + \frac{\partial^2 \tau}{\partial x_i^2} P^n \right] + \frac{\partial^2 P^n}{\partial x_i^2} \right\} = 0 \; . \tag{13.52}$$

Assume this series to be only finite, from $n = 0$ to $n = N$. Compare the coefficients with the lower powers of ω, i.e. with ω^1, ω^0 up to ω^{-N}. This yields the following system of equations:

$$\begin{aligned}
2\frac{\partial \tau}{\partial x_i}\frac{\partial P^0}{\partial x_i} &+ \frac{\partial^2 \tau}{\partial x_i^2} P^0 & & = 0 \, , \\
2\frac{\partial \tau}{\partial x_i}\frac{\partial P^1}{\partial x_i} &+ \frac{\partial^2 \tau}{\partial x_i^2} P^1 &- \frac{\partial^2 P^0}{\partial x_i^2} & = 0 \, , \\
&\vdots & & \\
2\frac{\partial \tau}{\partial x_i}\frac{\partial P^N}{\partial x_i} &+ \frac{\partial^2 \tau}{\partial x_i^2} P^N &- \frac{\partial^2 P^{N-1}}{\partial x_i^2} & = 0 \, , \\
& & \frac{\partial^2 P^N}{\partial x_i^2} & = 0 \, .
\end{aligned} \tag{13.53}$$

These are the *transport equations* for determining the amplitude coefficients $P^0(x_m)$ to $P^N(x_m)$. The first of these equations is a linear partial differential equation of the first order, the remaining equations are linear partial differential equations of the second order. They can be solved along rays. In that case they reduce to ordinary differential equations.

System (13.53) represents $N+2$ equations for $N+1$ unknown functions P^0 to P^N. Consequently, the last equation in (13.53) is superfluous. This equations can be used to check the accuracy of the approximate solution.

Very frequently we use the ray approximation only $(N = 0)$. In this case, system (13.53) reduces to two equations:

$$\begin{aligned}
2\frac{\partial \tau}{\partial x_i}\frac{\partial P^0}{\partial x_i} + \frac{\partial^2 \tau}{\partial x_i^2} P^0 & = 0 \, , \\
\frac{\partial^2 P^0}{\partial x_i^2} & = 0 \, .
\end{aligned} \tag{13.54}$$

The second of these equations is again superfluous.

13.10 Equations of the Ray Method for an Isotropic Inhomogeneous Medium

Let us consider the elastodynamic equations (13.40) for an isotropic inhomogeneous medium. We shall solve them again by the ray method in the frequency domain. However, in the order to simplify the problem, let us restrict ourselves to the ray approximation only, i.e. to the leading term for $n = 0$ in the ray series (13.36). Therefore, assume the displacements to be of the form

$$u_i(x_m, t, \omega) = U_i(x_m) e^{-i\omega(t-\tau(x_m))} , \tag{13.55}$$

where we write simply U_i instead of U_i^0. For the first and second derivatives of u_i with respect to coordinates we get

$$\begin{aligned} \frac{\partial u_i}{\partial x_j} &= \left[\frac{\partial U_i}{\partial x_j} + i\omega U_i \frac{\partial \tau}{\partial x_j} \right] e^{-i\omega(t-\tau)} , \\ \frac{\partial^2 u_i}{\partial x_j \partial x_k} &= \left[\frac{\partial^2 U_i}{\partial x_j \partial x_k} + i\omega L^i_{jk} - \omega^2 U_i \frac{\partial \tau}{\partial x_j} \frac{\partial \tau}{\partial x_k} \right] e^{-i\omega(t-\tau)} , \end{aligned} \tag{13.56}$$

where we have denoted

$$L^i_{jk} = \frac{\partial U_i}{\partial x_j} \frac{\partial \tau}{\partial x_k} + \frac{\partial U_i}{\partial x_k} \frac{\partial \tau}{\partial x_j} + U_i \frac{\partial^2 \tau}{\partial x_j \partial x_k} . \tag{13.57}$$

By inserting (13.55) into (13.40), and omitting the common exponential term $e^{-i\omega(t-\tau)}$, we then obtain

$$\begin{aligned} &(\lambda + \mu)\left[\frac{\partial^2 U_j}{\partial x_i \partial x_j} + i\omega L^j_{ij} - \omega^2 U_j \frac{\partial \tau}{\partial x_i} \frac{\partial \tau}{\partial x_j} \right] + \\ &+ \mu \left[\frac{\partial^2 U_i}{\partial x_j^2} + i\omega L^i_{jj} - \omega^2 U_i \frac{\partial \tau}{\partial x_j} \frac{\partial \tau}{\partial x_j} \right] + \\ &+ \frac{\partial \lambda}{\partial x_i} \left[\frac{\partial U_j}{\partial x_j} + i\omega U_j \frac{\partial \tau}{\partial x_j} \right] + \\ &+ \frac{\partial \mu}{\partial x_j} \left[\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i} + i\omega \left(U_i \frac{\partial \tau}{\partial x_j} + U_j \frac{\partial \tau}{\partial x_i} \right) \right] = -\rho\omega^2 U_i . \end{aligned} \tag{13.58}$$

From the same reasons as in the acoustic case, we shall put the coefficients with ω^2, ω^1 and ω^0 equal zero. The coefficient with ω^2 yields the equation

$$\left(\frac{\lambda+\mu}{\rho}\frac{\partial\tau}{\partial x_i}\frac{\partial\tau}{\partial x_j}+\frac{\mu}{\rho}\frac{\partial\tau}{\partial x_k}\frac{\partial\tau}{\partial x_k}\delta_{ij}-\delta_{ij}\right)U_j=0\ . \tag{13.59}$$

This equation is similar to that which we obtained in Chapter 8 when we studied the propagation of plane waves. However, in general, the elastic parameters and density are not constant now. Consequently, the decomposition of the wavefield into plane waves will hold locally. Equations (13.59) for $i=1,2,3$ can thus be identified as the *local Christoffel equations for an isotropic medium*,

$$\left(\Gamma_{ij}-\delta_{ij}\right)U_j=0\ , \tag{13.60}$$

where the elements of the Christoffel matrix Γ are

$$\Gamma_{ij}=\frac{\lambda+\mu}{\rho}\frac{\partial\tau}{\partial x_i}\frac{\partial\tau}{\partial x_j}+\frac{\mu}{\rho}\frac{\partial\tau}{\partial x_k}\frac{\partial\tau}{\partial x_k}\delta_{ij}\ . \tag{13.61}$$

The corresponding Christoffel determinant must be equal to zero, which again yields a cubic equation for determining the eigenvalues. The eigenvectors are again mutually perpendicular; see the analogous derivation in Section 8.1. Here we shall use another simple derivation.

The equation $\tau = const$ is the equation of the surface with a constant phase, i.e. the equation of a wavefront. The gradient of function τ has the direction perpendicular to the wavefront, i.e. the direction of a ray. Therefore, decompose the vector $\mathbf{U}=(U_1,U_2,U_3)$ into the component along the ray, and the component perpendicular to it. Let us study these components as two new amplitude vectors.

Firstly, assume the amplitude vector $\mathbf{U}$ to be parallel with the gradient of τ, i.e.

$$U_j=K\frac{\partial\tau}{\partial x_j}\ , \tag{13.62}$$

$K=K(x_m)$ being a coefficient. Inserting (13.62) into (13.59) yields

$$K\frac{\partial\tau}{\partial x_i}\left(\frac{\lambda+\mu}{\rho}\frac{\partial\tau}{\partial x_j}\frac{\partial\tau}{\partial x_j}+\frac{\mu}{\rho}\frac{\partial\tau}{\partial x_k}\frac{\partial\tau}{\partial x_k}-1\right)=0\ .$$

In order to satisfy this equation, the term in the square brackets must be equal to zero. Replacing the dummy index *k* by *j*, we arrive at the eikonal equation

$$\boxed{\frac{\partial\tau}{\partial x_j}\frac{\partial\tau}{\partial x_j}=\frac{1}{\alpha^2}}, \qquad (13.63)$$

where $\alpha=\sqrt{(\lambda+2\mu)/\rho}$ is the local velocity of P waves (it varies with coordinates).

Secondly, assume the vector **U** to be perpendicular to the gradient of τ, $\mathbf{U}.\text{grad}\,\tau=0$, i.e.

$$U_j\frac{\partial\tau}{\partial x_j}=0\ . \qquad (13.64)$$

Under this condition, the first term in Eq. (13.59) vanishes, and we arrive at the eikonal equation in the form

$$\boxed{\frac{\partial\tau}{\partial x_j}\frac{\partial\tau}{\partial x_j}=\frac{1}{\beta^2}}, \qquad (13.65)$$

where $\beta=\sqrt{\mu/\rho}$ is the local velocity of S waves.

Hence, we have arrived at the important conclusion that, under the assumption of high frequencies, the wavefield separates into two independent wave processes even in slightly inhomogeneous media.

The transport equations can easily be obtained as the conditions that the coefficients with ω^1 and ω^0 in Eqs. (13.58) should be zero. We shall not present these equations here, but refer the reader to the lecture notes by Psencik (1994) and Popov (1996). As opposed to the simple methods which were based on energy considerations (Section 13.6), the solution of the transport equations yields not only the absolute values of wave amplitudes, but also the polarisation of S waves.

13.11 Relations between the Eikonal Equation and Fermat's Principle

In Chapters 3 to 5 we postulated the validity of Fermat's principle, and then we used this principle to calculate seismic rays. In the method of ray series, we build the theory on the elastodynamic equations. Therefore, we should clarify the relation between Fermat's principle and the solutions of the elastodynamic equations. We can, e.g., compare the corresponding differential equations which follow from these approaches.

The eikonal equation is a nonlinear first-order partial differential equation, which can be solved by the method of characteristics (Psencik, 1994). In this case, the characteristics are determined by the system of ordinary differential equations which are identical with the equations following from Fermat's principle (Chapter 5). This proves the equivalence of the eikonal equation and Fermat's principle for the purposes of computing seismic rays. Hence, instead of solving the eikonal equations, we could calculate seismic rays on the basis of Fermat's principle. The advantage of applying this principle consisted in the fact that we could calculate many kinematic characteristics of seismic waves without knowledge of the elasticity theory. However, at last the conclusions following from Fermat's principle should be verified and compared with the more exact solutions of the elastodynamic equations.

13.12 Other Approaches

In this chapter we considered only harmonic waves, and studied them by the method of ray series. Let us briefly mention other approaches.

13.12.1 Description in the time domain

The description of wave phenomena in the frequency domain is convenient in many branches of physics (optics, acoustics), but less convenient in seismology, where we usually encounter waves of a short time duration (transient waves). Consequently, seismic waves are frequently studied in the time domain (Psencik, 1994). Two important notions are used for these purposes, namely the *Hilbert transform* and the *analytic signal.*

The Hilbert transform of a function $g(t)$ is defined by the formula

$$h(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{g(\xi)}{\xi - t} \mathrm{d}\xi \ . \tag{13.66}$$

Functions $g(t)$ and $h(t)$ form the so-called Hilbert transform pair.

The analytic signal $F(t)$, corresponding to function $g(t)$, is defined by

$$F(t) = g(t) + ih(t) \ ,$$

where $h(t)$ is the Hilbert transform of $g(t)$.

It can be shown that functions $g(t)$, $h(t)$ and $F(t)$ behave like $\cos t$, $-\sin t$ and e^{-it}. Since it is convenient to describe harmonic waves in terms of exponentials, in the time domain it is analogously convenient to use analytic signals.

Thus, we usually work with the analytic signal $F(t)$, but the physically meaningful signal is $g(t)$. For further details we refer the reader to the lecture notes by Psencik (1994).

13.12.2 Other methods

The ray methods, described above, consider the propagation of energy only along individual rays. Consequently, these high-frequency approximations yield unreal results in some regions (transitions into shadow zones, caustics, etc.). More accurate approximations are, for example, the paraxial approximation or the method of Gaussian beams, which consider the wavefield also in a certain vicinity of the ray. Detailed descriptions of these method can be found in Popov (1996).

A comprehensive description of the contemporary ray methods in seismology should soon be available in the monograph by Cerveny (in press).

References

Aki K., and Richards P. G. (1980): *Quantitative Seismology*. W. H. Freeman and Company, San Francisco.

Abramowitz M., and Stegun I. A. (Eds.) (1972): *Handbook of Mathematical Functions*. Dover Publications, New York.

Arfken G. (1970): *Mathematical Methods for Physicists*. (Second Edition). Academic Press, New York.

Babich V. M., and Alekseev A. S. (1958): O luchevom metode vychislenia intensivnosti volnovykh frontov. Izv. AN SSSR, ser. geof., No. 1, 9-15. (On the ray method of computing the intensity of wavefronts; in Russian).

Båth M. (1979): *Introduction to Seismology*. Birkhäuser Verlag, Basel. (Second, revised edition).

Bolt B. A. (1988): Earthquakes. W.H. Freeman and Company, New York

Bullen K. E. (1965): *An Introduction to the Theory of Seismology*. Cambridge University Press, Cambridge. (Third edition).

Bullen K. E., and Bolt B. A. (1993): *An Introduction to the Theory of Seismology*. Cambridge University Press, Cambridge. (Fourth edition).

Cerveny V. (1976): *Sireni elastickych vln*. MFF, Universita Karlova, Praha. (Propagation of Elastic Waves; in Czech).

Cerveny V. (1978): *Paprskove metody v seismologii*. MFF, Universita Karlova, Praha. (Ray Methods in Seismology; in Czech).

Ewing W. M., Jardetzky W. S., and Press F. (1957): *Elastic Waves in Layered Media*. McGraw-Hill Book Company, New York.

Fowler C. M. R (1990): *The Solid Earth - An Introduction to Global Geophysics*. Cambridge University Press, Cambridge.

Grünthal G. (Ed.) (1993): *European Macroseismic Scale 1992 (up-dated MSK-scale)*. European Seismological Commission, Luxembourg.

Heiskanen W. A., and Vening Meinesz F. A. (1958): *The Earth and Its Gravity Field*. McGraw-Hill, New York.

Karal F. C., and Keller J. B. (1959): Elastic wave propagation in homogeneous and inhomogeneous media. J. Acoust. Soc. Am. **31**, 694-705.

Kittel Ch., Knight W. D., and Ruderman M. A. (1962). *Mechanics*. (Berkeley Physics Course, Vol. 1). McGraw-Hill, New York.

Knott C. G. (1899): Reflection and refraction of elastic waves with seismological applications. Phil. Mag., ser. 5, **48**, 64-97.

Novotny O. (1998): *Motions, Gravity Field and Figure of the Earth*. (Lecture notes). CPGG/UFBA, Salvador.

Novotny O. (1999): *Seismic Surface Waves*. (Lecture notes). CPGG/UFBA, Salvador.

Popov M. M (1996): *Ray Theory and Gaussian Beam Method for Geophysicists*. (Lecture notes). PPPG/UFBA, Salvador.

Pšencík I. (1994): *Introduction to Seismic Methods*. (Lecture notes). PPPG/UFBA, Salvador.

Richter Ch. F. (1958): *Elementary Seismology*. W. H. Freeman and Company, San Francisco.

Savarensky E. (1975): *Seismic Waves*. Mir Publishers, Moscow.

Stacey F. D. (1992): *Physics of the Earth*. 3rd Ed. Brookfield Press, Brisbane.
Tygel M., and Hubral P. (1987): *Transient Waves in Layered Media*. Elsevier, Amsterdam.
Zoeppritz K. (1919): Über Erdbebenwellen VIIb. Göttingen Nachrichten, 66-84.