Chapter 1

The Earth in the Solar System

1.1 Solar System Formation, Accretion, and the Early Thermal State of the Earth

To understand the composition and early evolution of the Earth it is necessary to consider as far back as the formation of our solar system. Solar system formation was a complex process that is not well understood because of the lack of data and the vast physical and chemical complexities of the process. However, there are certain key parameters that we do know. As discussed in a later lecture, we know from the study of meteorites the age of the solar system and its initial composition. And, comparatively speaking, we know much about the nature of the present-day solar system. In addition, we have observations of old and young stars that inform us about the life cycle of the sun. The goal is to use all the information in combination with the laws of physics and chemistry to fill in the blanks between the initial state and the present state of the solar system, and to consider what this means for the constitution and initial state of the Earth.

1.2 Rotation and Angular Momentum

All the planets revolve in the same direction around the sun, and in practically the same plane. For the most part they also rotate in the same direction about their own axes, although there are notable exceptions, such as Venus. The gravitational collapse of molecular clouds is widely believed to lead to star formation and it is likely that our solar system condensed from a collapsed, rotating cloud of gas and dust. Rotating disks of material are ubiquitous in space, occurring all the way from planetary to galactic scales. A rotating disk is the signature of a self-gravitating system that has contracted in radius and amplified its angular velocity in order to preserve its total angular momentum. In a rotating protostar the gravitational attraction everywhere will be towards the center of mass. But the centrifugal force will be directed normal to the axis of rotation. The resolved force vector will move gas and dust nearer to the median plane as the cloud contracts. This process leads to the disk shape, which dissipates energy and minimizes collisions.

One of the more interesting boundary conditions is the present distribution of angular momentum. Consider a planet of mass m that orbits a central body of mass M, whose position with respect to the central body can be described by a vector **r**.

The orbital angular momentum (L) of the planet can be written

$$L = mr^2\omega = mr^2\frac{d\theta}{dt}, \tag{1.1}$$

where r is distance, m is the mass, ω is the angular velocity (=$d\theta/dt$), and θ is the angle with respect to a fixed direction in the orbit plane. It can be shown that

$$r^2\frac{d\theta}{dt} = \frac{L}{m} = 2\frac{dS}{dt}, \tag{1.2}$$

where S is the area swept out by **r**. Then

$$\frac{dS}{dt} = \frac{L}{2m}, \tag{1.3}$$

which is a statement of **Kepler's second Law of Motion**: the line between a planet and the sun sweeps out equal areas in equal periods of time. Equation (1.3) is a statement of **conservation of angular momentum**. The total planetary energy (E), which is the sum of the kinetic and potential contributions,

$$E = \frac{1}{2}mv^2 + \frac{mGM}{r} = \text{constant} \tag{1.4}$$

where

$$v^2 = \left(\frac{dr}{dt}\right)^2 + \left(r\frac{d\theta}{dt}\right)^2 \tag{1.5}$$

is the planetary velocity, is also conserved. It is possible to rewrite (1.1) in terms of the mass of the sun and doing so yields

$$L = mr^2\omega = (GM_s)^{1/2}mr^{1/2}. \tag{1.6}$$

By integrating (1.6) over all the planets we find that while the sun contains 99.8% of the mass of the solar system, it has only about 1% of the angular momentum. About 60% of the angular momentum of the solar system is associated with the orbit of Jupiter alone. Most models suggest that the protosun was rotating more rapidly than at present. Helioseismological results show that deeper parts of the sun rotate faster than the surface. The deep solar interior, which has not yet been probed, may hold the record of that body's relic rotation. Solar system evolution models must show how the protosun's angular momentum gets transported outward. Most models invoke magnetic and gravitational torques that spin down the sun and spin up the planets. Magnetizations of meteorites are consistent with this idea. The transfer of angular momentum could have contributed to the chemical fractionation of the solar system, since an outwardly migrating magnetic field would affect the ionized plasma but not condensed particles, which couple to the field only by viscous drag. Thus higher temperature condensates would remain in the inner part of the solar system and more volatile constituents would be transferred outward. In fact this is observed.

1.3 The Sun

Stellar Evolution: The Hertzsprung-Russell Diagram

A common method of characterizing stars is the **Hertzsprung-Russell (H-R) diagram**, which is a plot of absolute magnitude or luminosity versus effective (blackbody) temperature. It is traditional to plot effective temperature from high to low on the abscissa, and luminosity from dim to bright along the ordinate. For two stars with the same effective temperature, more light will come from the larger star than the smaller; hence the largest stars are at the top of an H-R diagram.

As each star proceeds through its life cycle it moves around on the H-R diagram. While we can't observe the life cycle of a single star, we can search through the current "snapshot" of our galaxy and find stars at all stages of evolution. Making an ensemble H-R plot reveals that many stars fall along a single line called the **main sequence**. Stars on the main sequence are in a relatively steady state of hydrogen burning in their cores, as is the present sun. An average G-type (yellow) star like our sun is thought to have a lifetime (i.e. a residence time on the main sequence) of about 10 billion years.

The T-Tauri Stage

The conspicuous absence of gas between the planets in the solar system must be explained in any model of solar system formation. Before a new star reaches the main sequence it goes through a pre-main sequence evolution of gravitational collapse from a protostellar nebula. Our best information about this stage comes from studying a class of young stars called **T Tauri stars**. T Tauri stars are thought to be still contracting and evolving, and typically less than one million years old. They are typically 0.2 to 2 solar masses in size, and they show evidence of strong magnetic activity. Some T Tauri stars have spectra that include "forbidden lines", which occur in low-density gas and are the signature of a gaseous nebula. Rapid fluctuations in ultraviolet and x-ray emissions are common. They also tend to show strong infrared emission and have spectra with silicon lines indicating that they are surrounded by dust clouds.

T Tauri stars are associated with strong solar winds and high luminosities. It is thought that our sun probably passed through a T Tauri stage in its early evolution, and that the volatile elements in the inner solar system were blown away during this stage.

1.4 Planetary Formation

Condensation and Cooling

The most widely accepted cosmogonical (formation) theory is that of V. Safronov, who was the first to hypothesize that the solar system initially accreted from a nebular cloud that evolved from a sphere to a disk. While details of solar system formation models differ, a common premise is that the planets formed from particle growth in an initially tenuous dust-gas nebula. The mechanism to trigger the initial collapse of the nebula has been argued and hypotheses range from uniform gravitational collapse, to galactic spiral density waves, to catastrophic suggestions such as a supernova in the solar neighborhood. A supernova, though a low probability event, is supported by the discovery of micro-diamonds in cosmic dust. These imply that the solar system environs achieved high pressures due to passage of severe shock waves that would accompany only an event of this intensity. The problem with the supernova hypothesis is that it would imply that solar system formation is not a common phenomenon, which runs contrary to current thought.

There are a number of scenarios for planetary growth. It is possible that the planets accumulated from small moon-sized bodies, called **planetesimals**, by infrequent encounters. Or instead accumulation may have occurred from groups of bodies that collectively became gravitationally unstable. It is not clear whether planetary accumulation occurred in a gaseous or gas-free environment. In a gaseous nebula temperatures tend to be homogeneous, but as gas clears due to the solar wind and condensation into dust grains the opacity of the nebula decreases significantly. During this time the system establishes a large temperature gradient.

It is generally accepted that the planets accreted from a nebula with a composition similar to that of the sun, i.e., made mostly of hydrogen. The slowly-rotating nebula had a pressure and temperature distribution that decreased radially outward. The density of the nebula was probably not very great. Model estimates of typical pressures in the vicinity of Earth's orbit generally fall in the range of 10-100 Pa but these are not very well constrained. The disk must have cooled primarily by radiation, condensing out dust particles that were initially of composed of **refractory** elements. These high temperature condensates first appear at temperatures of $1600^\circ - 1750^\circ$K and consist of silicates oxides and titanates of calcium and aluminum, such as $\mathrm{Al_2O_3}$, $\mathrm{CaTiO_3}$, and $\mathrm{Ca_2Al_2Si_2O_7}$ and refractory metals such as those in the platinum group. These minerals are found in white inclusions in the most primitive class of meteorites, the Type III carbonaceous chondrites, discussed in a later lecture. Metallic iron condenses out next, followed by the common silicate materials forsterite (an olivine) and enstatite (a pyroxene). Iron sulfide (troilite; FeS) and hydrous minerals condense at temperatures of 700°-800°K. **Volatile** materials, most notably $\mathrm{H_2O}$ and $\mathrm{CO_2}$ condense out at 300°-400°K. Planets that contain these substances were accumulated from material that condensed in this temperature range, which provides some clue about the early thermal structure of the solar nebula.

Time scales for the condensation of gas to dust, of accumulation of dust to planetesimals, and of accretion of planetesimals to planets and moons are also not well constrained. If cooling occurred slowly in comparison to other processes then planets would have formed during the cooling process and could have accreted inhomogeneously. If instead cooling occurred rapidly, then the planets would have formed from cold, generally homogeneous material. Homogeneous accretion models are favored, with planetary differentiation thought to be mostly accomplished in the early stages after accretion.

Accretion

The process or processes that were responsible for the accumulation of dust and small particles into planetesimals is a matter of debate. Sticking mechanisms such as electrostatic attraction and vacuum welding have been suggested. But as material accumulates, more planetesimal surface area is available for adding more material so the process accelerates. When planetesimals reach sizes of order 10^2 km gravitational attraction begins to dominate and accretion becomes dominated by that force. In the planetesimal accretion stage collisional velocities are a key consideration. If relative velocities between planetesimals are too low, then planetesimals will fall into nearly concentric orbits. Collisions will be low probability events and planets will not grow. Whereas if relative velocities between planetesimals are too high, fragmentation rather than accumulation will occur, and again planets won't grow. Safronov used scaling arguments concerning energy dissipation during collisions and an assumed size distribution of planetesimals to suggest that the mutual gravitation causes relative velocities to be somewhat less than the escape velocities of the largest bodies. By his estimation the system should regulate itself in a way to favor the growth of large planetesimals. If this idea holds in a general sense, then solar systems should form with a relatively small number of large planetary bodies rather than with many small bodies. Monte Carlo simulations bear this idea out.

1.5 Early Thermal State of the Earth

Accretional Heating

As planetesimals accrete into moons and planets a significant amount of energy must be released, much of which will be converted to heat. Various theories place the time for the accretion of Earth from 10^5-10^8 years, which was very rapid indeed in comparison to the age of the solar system. If accretion occurred rapidly, then not much cooling could have occurred between collisions.
To determine the amount of heating associated with accretion, it is necessary to take an inventory of the various sources of energy in the system. These include the kinetic energy of impacting projectiles, the potential energy of infalling material to the planetary surface and the thermal energy. For simplicity we will begin by assuming that accretion occurs sufficiently rapidly such that the process is adiabatic, i.e. with no heat lost from the accreting planet's surface. The total energy per unit mass of accreted material is simply a sum of the change in kinetic and potential contributions:

$$C_p \Delta T = \frac{1}{2}\left(v_\infty^2 - v_p^2\right) + \frac{GM}{R} \tag{1.7}$$

where ΔT is the temperature change, C_p is specific heat, v_∞ is the absolute velocity of the approaching projectile, v_p is the planetary velocity, $(\Delta v)^2 = v_\infty^2 - v_p^2$ is the relative impact velocity, G is the universal constant of gravitation, M is the mass of the planet, R is the planetary radius and

$$\frac{GM}{R} = gR, \tag{1.8}$$

where g is the gravitational acceleration at the planetary surface. It is reasonable to assume that the impact process is not perfectly efficient and that only a fraction h of the total energy will be converted to heat. Taking this into account and substituting (.5) we may write

$$C_p \Delta T = h\left[\frac{1}{2}\left(v_\infty^2 - v_p^2\right) + gR\right]. \tag{1.9}$$

This expression provides an upper limit of the increase in temperature that could occur during accretion. In practice the potential energy term dominates (1.9). But this expression isn't very realistic because it doesn't allow for cooling.
So we next consider the additional complication that heat is lost from the system by cooling at the surface. It is possible to write a balance between the gravitational potential energy of accretion, the heat lost by radiation, and the thermal energy associated with heating of the body. This causes the problem to become time dependent:

$$\rho \frac{GM(r)}{r} dr = \epsilon\sigma\left[T^4(r) - T_b^4\right] dt + \rho C_p \left[T(r) - T_b\right] dt \tag{1.10}$$

where $M(r)$ is the mass of the accumulating planet, ρ is the density of accreting material, ϵ is emissivity, σ is the Stefan-Boltzmann constant, T_b is the radiation equilibrium (blackbody) temperature, and t is time. In reality there will also be energy associated with latent heats of melting and vaporization that are ignored here. Temperature increases associated with the accretion of the the terrestrial planets from numerical solutions to (1.10) require rapid accretion times, 10^3 to 10^4 years for Earth, to exceed the melting temperature. These time scales are less than suggested by accretion models and would suggest that accretional heating is not very important for Earth or the other terrestrial planets. But it is necessary to consider in the most realistic sense possible the importance of radiation in ridding the planet of heat. Radiative temperature loss goes as T^4 and so is highly efficient in the sense that the planetary surface cools quickly. But if an impact site becomes buried by ejecta from fall-back or from nearby impacts, the surface would be covered. In this situation the outer part of the planet is hotter than the interior and thermal convection is prohibited. The only way to rid the planet of heat is to conduct it to the surface where it can be radiated away. Conduction is a much less efficient heat transport process and so accretional heat would be retained longer if that mechanism dominated. If accretional energy is buried deeply enough to prohibit thermal radiation from the surface, then temperature increases of order 2000° can be attained for planets that accrete in times suggested by models (10^6-10^7 years). But even if accretion did cause the near surface of the Earth to melt the process does not explain the earliest heating of the Earth's deep interior, which occurred through the process of differentiation.

Differentiation

From the Earth's moment of inertia (C/MR^2), which will be discussed later, we know that the Earth (and other terrestrial planets) have a radially stratified internal density structure. The implied increases in density with depth are greater than would be associated with simple self-compression due to an increase of pressure with depth. This leaves compositional changes, and to a lesser extent phase changes, to explain the observations. If the Earth accreted cold, then there must have been a process of internal differentiation to produce its radially stratified density structure. Differentiation from a homogeneous initial state to a structure with a distinct core and mantle involves a change in gravitational potential energy. The release of this energy was likely to have been an important source of heat in some planetary bodies. It is believed that differentiation would have occurred early in planetary evolution after a period of radioactive heating or in the last stages of impact accretion in which the temperature required to melt iron is achieved at shallow depth. Molten iron separates out from its silicate matrix and is denser than its surroundings and sinks by gravitational settling. It is reasonable to assume that the separation and sinking time is short compared to the time of heating. Also, the process is taking place in the interior so to first order surface heat loss may be neglected.

Under these assumptions it is possible to estimate the increase in temperature associated with core formation. We may calculate the change in gravitational potential energy associated with the instantaneous differentiation of a planet from a homogeneous state to a final state with a core and mantle. We shall assume that the total mass in the system remains constant. In addition, we will neglect contributions from other effects such as phase changes, the latent heat of melting, rotational kinetic energy (due to the change in moment of inertia), and strain energy. The gravitational potential energy (Ω) for a spherical planet in hydrostatic equilibrium in which density is simply a function of radius may be written

$$\Omega = \int_0^M \frac{Gm}{r} dm \tag{1.11}$$

where $m = 4/3\pi r^3\rho$ is the mass of accreting spherical body, and $dm = 4\pi r^2\rho dr$. Substituting (1.8) we find

$$\Omega = \int_0^M g(r) r dm. \tag{1.12}$$

Table 1.5: Temperature increase due to core formation

Planet	Core radius (km)	Energy released (J) $\Delta\Omega$	Mean temp increase $\Delta T(^\circ K)$
Earth	3485	1.5×10^{31}	2300
Venus	?	?	?
Mars	$1400 - 2100$	$\approx 2 \times 10^{29}$	300-330
Mercury	1840	2×10^{29}	700
Moon	< 400	$\approx< 1 \times 10^{27}$	10

We then re-arrange once again to integrate over the radius so that

$$\Omega = 4\pi \int_0^R g(r)\rho(r)r^3 dr. \tag{1.13}$$

In practice $\rho = \rho(r)$ is determined from an empirically-derived equation of state that relates density to pressure (i.e. depth). Equation (1.13) must be evaluated numerically. Now assume that the change in gravitational potential energy will be fully converted to heat. Then

$$\Delta\Omega = C_p \Delta T \quad \text{or} \quad \Delta T = \frac{\Delta\Omega}{C_p}. \tag{1.14}$$

Table 1.5 shows the mean temperature increase associated with instantaneous core formation for the terrestrial planets based on (1.13) and (1.14). Note that for the Earth the increase in temperature is expected to have been great enough to have produced extensive melting. So shortly after accretion the Earth would have been largely molten and vigorously convecting in the interior as a consequence of differentiation. For Venus the size of the core isn't known but if it is similar to Earth (given that planet's similar radius and mass), then Venus also would have experienced significant early melting when it formed its core. Melting also probably occurred on Mercury. But for Mars and the Moon the temperature increase is not great enough for melt generation, even taking into account the considerable uncertainties in core radii. Core formation could not have been a significant heat source early in the evolution of these bodies.

Formation of the Moon

We discussed above the role of impacts in the Earth's early heat budget from the illustrative calculation of temperature increase due to accretional heating. But after accretion there will continue to be impact infall as the planets "sweep up" asteroidal debris. This is quite apparent from looking at the 4.6 BY-old lunar highlands, which are saturated with impact craters formed during the **terminal bombardment**. It is now thought that a massive post-accretional impact was responsible for the formation of the Moon. The origin of the Moon has been a long-debated topic. While moons around planets are common in the solar system, Earth's moon is somewhat unusual given its large size compared to the primary. One might wonder then, whether "special circumstances" were associated with lunar origin.

Traditional models for lunar formation included co-accretion (the Moon formed near the Earth), capture (the Moon strayed too near to Earth and became trapped in orbit), and fission (the Moon formed by spinning off the Earth during an early rapid rotational period). All of these models had serious problems in explaining important features like the Moon's bulk composition, the angular momentum of the Earth-Moon system, etc.

The theory that is currently is favored is the **giant impact hypothesis**, which has gained support from numerical simulations and is consistent with the features above. In this scenario, shortly after accretion the Earth received a glancing impact from a Mars-sized asteroidal body. Smoothed particle hydrodynamic simulations from independent groups at Harvard and the University of Arizona have the same general features: The mantles of both the early Earth and the impactor melted and vaporized and the core of the impacting body wrapped around Earth's core. Mantle material from Earth and the projectile that was ejected re-condensed in orbit to form the Moon. This hypothesis is able to explain the puzzling lack of iron in the Moon. If this event did indeed occur then the Earth would have been largely melted by the event. Such a catastrophic occurrence must factor in to scenarios for the post-accretional evolution of the Earth.

1.6 Radioactive Decay

Radioactivity was discovered by Henri Becquerel in 1896 and it ultimately had profound implications for the evolution of the Earth. By that time sedimentary layering in outcrops was somewhat understood, at least to the point where it was known that observed sedimentary strata must have taken hundreds of millions of years to accumulate. At that point the only known energy sources available to the sun and Earth, i.e. the energy associated with gravitational collapse, allowed a maximum age of 25 Ma. For about 3 decades geologists debated whether to accept this age. The argument became moot due to the discovery of radioactivity. From the point of view of the evolution of the Earth, the discovery of radioactivity had two effects:

(1) it removed the short-term age limit of the Earth by providing a mechanism for long-term heating (here we are referring to the internal heat that drives mantle convection and thus plate motions); and

(2) it provided a means of determining absolute dates for rocks.

The stability of elements with respect to decay is related to the relative numbers of protons and neutrons. If the numbers of these are not approximately equal then material is prone to decay. Elements with the same number of protons but a different number of neutrons are called **isotopes**.

Radioactive decay occurs because some of the mass of an atom is held in binding energy. If there is "too much" binding energy (and quantum mechanics is required to assess what exactly constitutes "too much") then the nucleus will decay spontaneously to lower the energy state. Induced nuclear reactions, say by bombarding large atoms with neutrons, may also be used to achieve the lower energy state. Radioactive decay can occur by three classes of mechanisms:

alpha decay – the escape of a helium nucleus,

beta decay – the escape of an electron or positron, or

gamma decay – the emission of gamma radiation.

Radioactive decay is described with a simple rate law. The change in the total number N of radioactive particles over the time interval dt is therefore:

$$\frac{dN}{dt} = -\lambda N, \tag{1.15}$$

where the minus sign indicates that activity decreases with time, and λN represents the average number of particles that decay per second. The constant λ is based on the probability of a particular decay mechanism operating in an atom of a given element. We may rewrite

$$\frac{dN}{N} = -\lambda\, dt, \tag{1.16}$$

and integrating both sides of (1.16) we find

$$\ln N = -\lambda t + c, \tag{1.17}$$

where the constant c is found in the limit where $t \to 0$ to be $\ln N_o$. Taking the exponential of both sides we may write

$$N(t) = N_o\, e^{-\lambda t}\,, \tag{1.18}$$

where N_o is the initial number of radioactive particles. Equation (1.3) is the rate law of radioactive decay. The **half-life**, $T_{1/2}$, which represents the time it takes for half of the number of particles to decay, is found by setting $N/N_o = 1/2$ such that

$$\frac{N}{N_o} = \frac{1}{2} = e^{-\lambda T_{1/2}} \quad \Rightarrow \quad T_{1/2} = \frac{\ln 2}{\lambda} \quad = \quad \frac{0.69315}{\lambda}\,. \tag{1.19}$$

Note that the half-life represents an alternative way of expressing the decay constant λ. Nuclear binding energies are very large and nuclei are so small that radioactive decay rates are not significantly affected by physical conditions on Earth such as pressure and temperature. However, half-lives can be changed slightly by changes in bonding energy. For example, solar wind studies have shown that radioactive beryllium decays at slightly different rates on the sun and Earth.

In principle, the experimentally-demonstrated accuracy of the simple expression (1.19) allows for the determination of the absolute ages of billion-year-old rocks. However, in practice the initial concentration of the radioactive **parent** element N_o is very often not known. We can more easily measure the concentration of the **daughter** product (D^*), which is simply

$$D^* = N_o - N\,. \tag{1.20}$$

We may substitute (1.3) for N to find

$$D^* = N_o - N_o e^{-\lambda t} = N_o\left(1 - e^{-\lambda t}\right). \tag{1.21}$$

We want to eliminate N_o so we divide by (1.3) which gives

$$\frac{D^*}{N} = \frac{N_o\left(1 - e^{-\lambda t}\right)}{N_o\, e^{-\lambda t}} = \frac{1 - e^{-\lambda t}}{e^{-\lambda t}}. \tag{1.22}$$

or

$$\frac{D^*}{N} = e^{\lambda t} - 1. \tag{1.23}$$

Equation (1.8) can be used directly in the determination of ages if there is no initial non-radiogenic daughter component, or if that initial component can be estimated.

1.7 Radiometric Dating

The Rubidium-Strontium System

To illustrate the radiometric dating technique, consider the decay of the unstable isotope of rubidium, $^{87}\mathrm{Rb}$, into the stable isotope of strontium, $^{87}\mathrm{Sr}$. This system is particularly simple because the parent element only decays into one type of daughter element, unlike $^{40}\mathrm{K}$, say, which decays into both $^{40}\mathrm{Ar}$ and $^{40}\mathrm{Ca}$. The Rb-Sr system is useful for dating old rocks because the decay constant λ and half-life $T_{1/2}$ for the Rb-Sr system are well suited for the purpose:

$$\lambda = 1.42 \times 10^{-11}\,\mathrm{yr}^{-1}\,, \quad T_{1/2} = 48.8 \times 10^{9}\,\mathrm{yr}\,. \tag{1.24}$$

Only a fraction of the rubidium that was present in the solar nebula has so far decayed. If t is the time since some melting event reset the isotope ratios to their high-temperature values, then by (1.3), the current amount of $^{87}\mathrm{Rb}$ is reduced from its initial amount $^{87}\mathrm{Rb}_0$ by:

$$^{87}\mathrm{Rb} =^{87} \mathrm{Rb}_0\, e^{-\lambda_{87} t}\,. \tag{1.25}$$

The current amount of strontium, $^{87}\mathrm{Sr}$, is therefore increased from its initial amount, $^{87}\mathrm{Sr}_0$, by:

$$\begin{aligned} ^{87}\mathrm{Sr} &=^{87} \mathrm{Sr}_0 +^{87} \mathrm{Rb}_0 -^{87} \mathrm{Rb} && (1.26)\\ &=^{87} \mathrm{Sr}_0 +^{87} \mathrm{Rb}(e^{\lambda_{87} t} - 1)\,. \quad (1.9) && (1.27) \end{aligned}$$

To proceed with the dating, one uses a mass spectrometer to measure the amounts of $^{87}\mathrm{Sr}$ and $^{87}\mathrm{Rb}$ present in each sample. Since different parts of a rock will contain different concentrations of the unknown quantity $^{87}\mathrm{Sr}_0$, we must normalize against another stable isotope with similar chemistry that occurs in proportional concentrations, like $^{86}\mathrm{Sr}$. Dividing (1.9) by $^{86}\mathrm{Sr}$ yields:

$$\frac{^{87}\mathrm{Sr}}{^{86}\mathrm{Sr}} = \left(\frac{^{87}\mathrm{Sr}}{^{86}\mathrm{Sr}}\right)_0 + \frac{^{87}\mathrm{Rb}}{^{86}\mathrm{Sr}}(e^{\lambda_{87} t} - 1)\,. \tag{1.28}$$

The presence of initial daughter abundances also requires more than one measurement of the parent/daughter ratio to obtain an age. Samples that have different $^{87}\mathrm{Rb}/^{86}\mathrm{Sr}$ ratios can be plotted versus $^{87}\mathrm{Sr}/^{86}\mathrm{Sr}$ using (1.10). The $^{87}\mathrm{Rb}/^{86}\mathrm{Sr}$ ratio varies naturally from one mineral to another. For example it is typically higher in plagioclase than in pyroxene, so a spread in the samples is obtained by mineral separation. When plotted, the two ratios fall on a straight line called an **isochron** (meaning ”equal time”), which by (1.10) has a slope of $(e^{\lambda t} - 1) \approx \lambda t$ and a y-intercept of ($^{87}\mathrm{Sr}/^{86}\mathrm{Sr}_0$). If the the decay constant λ of the radioactive parent is known, the isochron yields the age, t, of the rock.

The most useful decay systems for radiometric dating are Rubidium-Strontium (Rb-Sr), Samarium-Neodymium (Sm-Nd), Potassium-Argon (K-Ar), Thorium-Lead (Th-Pb), and the two Uranium-Lead (U-Pb) systems. As illustrated above, in order for a parent-daughter system to be useful, a non-radiogenic reference isotope of the daughter must be present for comparison. In addition, the decay constant of the parent must be accurately known. The accuracy of radiometric dating also depends on the extent to which the rock under study has been a chemically closed system with respect to the parent and daughter elements. If it has not been a closed system then the daughter/parent ratio will not be solely due to radioactive decay, and the time information will be corrupted.

The Uranium-Lead System

The U-Pb system is especially useful because only measurements of Pb are required, and Pb tends to be reliable because it is not too mobile in rock. Also, because of decades of nuclear research the decay constants for uranium are very accurately known. Zircon crystals are resistant to uranium diffusion and are commonly used for this dating scheme. There are four isotopes of Pb: ^{204}Pb, ^{206}Pb, ^{207}Pb, ^{208}Pb. Only ^{204}Pb does not have a radioactive progenitor, and the decay schemes for the other three isotopes are:

$$^{238}\mathrm{U} \rightarrow^{206} \mathrm{Pb}\,, \quad \lambda_{238} \quad = 1.55 \times 10^{-10}\,\mathrm{yr}^{-1}\,, \quad T_{1/2} = 4.5\,\mathrm{By} \tag{1.29}$$

$$^{235}\mathrm{U} \rightarrow^{207} \mathrm{Pb}\,, \quad \lambda_{235} \quad = 9.85 \times 10^{-10}\,\mathrm{yr}^{-1}\,, \quad T_{1/2} = 0.7\,\mathrm{By} \tag{1.30}$$

$$^{232}\mathrm{Th} \rightarrow^{208} \mathrm{Pb}\,, \quad \lambda_{232} \quad = 4.95 \times 10^{-11}\,\mathrm{yr}^{-1}\,, \quad T_{1/2} = 14\,\mathrm{By}\,. \tag{1.31}$$

Using (1.8) and referencing to the non-radiogenic $^{204}\mathrm{Pb}$ we find

$$\frac{^{206}\mathrm{Pb}}{^{204}\mathrm{Pb}} = \left(\frac{^{206}\mathrm{Pb}}{^{204}\mathrm{Pb}}\right)_0 + \frac{^{238}\mathrm{U}}{^{204}\mathrm{Pb}}\left(e^{\lambda_{238}t} - 1\right)\,, \tag{1.32}$$

$$\frac{^{207}\mathrm{Pb}}{^{204}\mathrm{Pb}} = \left(\frac{^{207}\mathrm{Pb}}{^{204}\mathrm{Pb}}\right)_0 + \frac{^{235}\mathrm{U}}{^{204}\mathrm{Pb}}\left(e^{\lambda_{235}t} - 1\right)\,. \tag{1.33}$$

Now take the ratio of $^{207}\mathrm{Pb}/^{204}\mathrm{Pb}$ to $^{206}\mathrm{Pb}/^{204}\mathrm{Pb}$:

$$\frac{\frac{^{207}\mathrm{Pb}}{^{204}\mathrm{Pb}} - \left(\frac{^{207}\mathrm{Pb}}{^{204}\mathrm{Pb}}\right)_0}{\frac{^{206}\mathrm{Pb}}{^{204}\mathrm{Pb}} - \left(\frac{^{206}\mathrm{Pb}}{^{204}\mathrm{Pb}}\right)_0} = \frac{^{235}\mathrm{U}}{^{238}\mathrm{U}}\,\frac{e^{\lambda_{235}t} - 1}{e^{\lambda_{238}t} - 1}\,, \tag{1.34}$$

and rewrite into an isochron equation:

$$\frac{^{207}\mathrm{Pb}}{^{204}\mathrm{Pb}} = M\,\frac{^{206}\mathrm{Pb}}{^{204}\mathrm{Pb}} + B\,, \tag{1.35}$$

where the slope and y-intercept are:

$$M = \frac{^{235}\mathrm{U}}{^{238}\mathrm{U}}\,\frac{e^{\lambda_{235}t} - 1}{e^{\lambda_{238}t} - 1} = 0.613\,, \quad B = \left(\frac{^{207}\mathrm{Pb}}{^{204}\mathrm{Pb}}\right)_0 - M\left(\frac{^{206}\mathrm{Pb}}{^{204}\mathrm{Pb}}\right)_0 = 4.46\,. \tag{1.36}$$

The age information is contained in the slope, M, using only isotopes of Pb. The value of $^{235}\mathrm{U}$ $/^{238}\mathrm{U}$ is $1/137.88$, and this ratio is very nearly constant in all natural materials. To determine the initial lead ratios the standard practice is to look to meteorites. Iron meteorites have virtually no uranium. The least radiogenic lead found anywhere is in the Canyon Diablo meteorite. This is defined to be **primordial lead**, the initial lead ratio in the solar nebula.

The Age of Earth

Using the decay constants for uranium, the ^{235}U/^{238}U ratio given above, and the initial lead ratios from Canyon Diablo this gives a date of $(4.54 \pm 0.03) \times 10^9$ years. This is the time when isotopically homogeneous lead was isolated from the solar nebula in various bodies with different U/Pb ratios. This is the best estimate the age of the Earth. It is interesting that Earth and the meteorites fall on the same lead-lead isochron, which is evidence that lead and uranium were both isotopically homogeneous in the solar nebula before the accretion into the planets.
The radiometric dating of crustal and oceanic rocks from Earth's surface indicates that most of the surface is less than 100 million years old. The oldest rock was dated by Sam Bowring of MIT. It is an igneous rock and the age is 4.03×10^9 years. The surface of Earth is much younger than that of the other solid planets for several reasons.
Erosion is efficient on Earth because of the abundance of liquid water and because of the presence of a biosphere. Earth also renews its surface continuously through the action of plate tectonics, where new crustal material comes to the surface at mid-ocean ridges and old crustal material plunges below the surface at subduction zones. Given the great activity occurring on Earth's surface, it is not at all surprising that terrestrial rocks are not the oldest rocks in the solar system.

1.8 Radioactivity as a Heat Source

Radioactivity is an important source of heating in the early solar system and in addition represents the major source of long-term heating of the Earth (and other terrestrial planets). The important heat-producing radionuclides form two classes: long-lived nuclides and short-lived nuclides.

Long-Lived Nuclides

The **long-lived radionuclides**, of which the most important are ^{238}U, ^{235}U, ^{232}Th and ^{40}K, are a primary source of heat over the span of Earth history. They provide heat which drives present-day mantle convection. Long-lived radioactive elements have combinations of valence states and ionic radius that prevent them from being easily accommodated into the crystal lattices of the most common silicate rocks. They are examples of lithophile elements, which preferentially concentrate in the liquid phase; the majority of them are incorporated into the first few percent of a melt. For this reason, a significant fraction of the Earth's radioactivity is concentrated in the continental crust.

Short-Lived Nuclides

The **short-lived radionuclides** may have been an important source of heat responsible for the early melting of meteorites. They might also have provided an early heat source for planets, depending on the time between nucleosynthesis and planetary accretion. The most abundant of the short-lived radionuclides is ^{26}Al, which decays with a half-life of 720,000 years to ^{26}Mg. The evidence for ^{26}Al being an important source of heat in the early history of the solar system comes from excess amounts of ^{24}Mg found in CAI's in the Allende meteorite. The isotope ^{26}Mg was enriched relative to the most common isotope ^{24}Mg compared to solar abundance. The heat-producing ability of this isotope is such that solid objects a few km or greater would have been heated to melting if they formed with the ratio of ^{26}Al/^{27}Al implied to have been present in Allende.

A major question is: how could ^{26}Al be incorporated fast enough into early solar system objects to melt them? With such a short half-life, radioactive decay begins to produce heat after a cosmically short period of time. The isotope ^{26}Mg is only produced by decay of ^{26}Al, and ^{26}Al is only produced in **supernovae**. This suggests that our solar system might have formed close to a supernova. Another piece of information in support of the supernova hypothesis is the fact that very small diamonds have been found in some meteorites. On Earth diamond forms at great depths due to very high pressures which contract carbon to the closely-packed state, characterized by all covalent bonds. In space the pressures required to form diamond can only be achieved in a supernova. If the solar system formed near a supernova it would solve the problem of the mechanism that caused the protosolar cloud to collapse, as the shock waves that emanate from supernovae would provide a natural mechanism for compression. However, supernovae are rare events and if it is necessary to invoke the participation of one it would imply that the formation of our solar system was a chance event.

1.9 Meteorites and the Bulk Composition of the Earth

The chemical and mineralogical information contained in the various types of meteorites yield some of our most important clues about the nature of the early solar system and the early Earth. A wide range of information about the chemical and physical state derives from analysis of the various classes of meteorites. We review the various types of meteorites and their implications for the Earth's bulk composition.

1.10 Chondrites

Chondrites form the most abundant class of meteorites and as a group represent primordial objects that are chemically similar to the sun. They are so named because they contain **chondrules**[1], which are primitive, glassy, silicate globules up to a few millimeters in size. Chondrules were magnetized as independent grains, and have never been found in a terrestrial rock. It is believed that the chondrules condensed out of the protoplanetary nebula before being incorporated into a matrix that consists of crystalline silicate minerals and sometimes grains or filaments of iron-nickel alloy. The magnetization, spherical shape and fine crystalline structure indicate that they were melted and very rapidly cooled, perhaps on time scales as short as minutes. Chondrules have been preserved because they were incorporated into **parent bodies** that were not large enough to undergo dynamic processes that would modify or destroy them.

A comparison of elemental abundance in a chondrite versus elemental abundance in the sun's photosphere, as determined by spectroscopy, yields an astonishing correspondence. The only elements that don't match well are the most volatile elements, which tend to escape incorporation into a meteorite as it cools, and lithium, which is depleted in the sun due to nuclear reactions. The name chondrite has come to refer more broadly to any meteorite with a chemical composition that is similar to that of the sun, and there are a number of subclasses, defined on the basis of their chemistry and their degree of metamorphism, i.e., the modification of their structure and mineralogy due to temperature and pressure.

[1]From the Greek word $\chi o \nu \delta \rho o \varsigma$ or **chondros**, meaning "grain"

Carbonaceous Chondrites

An important fall occurred in Chihuahua, Mexico, in 1969, when a large meteor was observed to come into the atmosphere many pieces. The first piece was found near a house in the small village of Pueblito de Allende. Following standard practice, all of the meteorite fragments that were recovered from that fall are collectively named Allende. The **Allende** fall occurred just as the Apollo program was swinging into full gear, and it gave scientists who were preparing for the arrival of moon rocks an opportunity to practice techniques for analysis of chemical composition on an extraterrestrial sample. Because of its unique chemical constitution and the fact that there was plenty of it to go around, analysis of the Allende meteorite has taught us much about the early solar system.

Allende is a member of an important sub-class of chondrites referred to as the **carbonaceous chondrites**. It is a Class III carbonaceous chondrite that is the most primitive (i.e. least altered by heating or other metamorphism) of that class and much of what we know about the bulk composition of the Earth is based on that single meteorite.

1.11 Secondary Processing

If we want to use chondrites and other meteorites as a probe of the early solar system, we need to have a good idea of the types of processes that modified these meteorites from their original forms. Igneous meteorites, like the **howardite-eucrite-diogenite suite (HED)**, have definitely experienced secondary processing. However, they do provide unique information about the evolution of Earth and terrestrial planets, particularly basalt generation and core formation. Chondrites may be the least altered class of meteorites, but they too have undergone secondary processing. Most chondrites have experienced thermal metamorphism. The result is changes in texture and mineralogy, and possibly chemical composition. The temperatures necessary to cause metamorphism are in the 400° to 1000°C range for the relatively low pressures encountered in small parent bodies. Possibly important heat sources are the decay of short-lived radionuclides, electromagnetic induction, and accretion of material. The least metamorphosed type III chondrites (like Allende) probably carry the most information about the early solar system, but even these have been effected somewhat by secondary thermal processing. Along with heat, the chemical reactivity of water has played an important role in the secondary processing of some of the most compositionally primitive meteorites. This process, called aqueous alteration, tends to replace the preaccretionary lithology with new mineralizations, although the bulk chemistry is apparently preserved.

Along with internal processing, meteorites can be changed by exogenic processes like collisions. Violent impacts produce shock metamorphism of individual mineral grains, and also produce rocks called breccias that contain mixtures of different previous rocks, just like the breccias found on the lunar highlands. The study of breccias has provided information on the accretional growth and processing of parent bodies. The effects of shock metamorphism have been seen in all major groups of meteorites. It appears that collision-induced high speed impacts took place before, during, and after the initial accretion and differentiation of the parent bodies.

Meteorites contain information related to their long exposures to galactic cosmic rays, solar radiation and the solar wind. It is possible to determine how long a meteorite existed free of its parent body before it impacted Earth by examining the cosmic ray damage. Noble gases are the most volatile elements in meteorites, but they are nonetheless present in measurable quantities in virtually all meteorites. Trapped noble gases are either “solar” or “planetary.” The solar noble gases are actually implanted solar-wind or solar-flare material, and provide relatively direct information about the sun. The planetary noble gases have elemental abundances similar to those found in Earth's atmosphere.

1.12 Achondrites

The **achondrites** are igneous meteorites that lack water-bearing (hydrous) or oxidized minerals. This class of meteorites includes the **eucrites**, **diogenites** and **howardites**. Because they are basaltic in composition they are believed to come from parent bodies that were large enough to have differentiated (melted) to produce a crust. The main belt asteroid **Vesta** is the best compositional analog for the eucrites.

Meteorites from Mars

Shergottites are named after a meteorite that fell in 1865 in Shergotty, which is located in the northeastern Indian state of Bihar, which borders Nepal. Unlike the eucrites, their pyroxene and plagioclase mineralogy is strikingly similar to terrestrial basalts. They also have small amounts of the hydrous mineral amphibole kaersutite, whereas eucrites show no evidence of water in their minerals, and they have some magnetite, which contains iron in oxidized form (Fe^{+3}), whereas eucrites contain only reduced iron. Their pyroxene crystals are elongated and arranged horizontally the way such crystals would accumulate after settling to the bottom of a magma chamber. Such igneous rocks are called **cumulates**. Two other types of cumulate meteorites have hydrous and oxidized minerals like the shergottites, these are the **nakhlites**, which contain the dark-green to black pyroxene mineral augite, and a unique meteorite that fell in Chassigny, France called **chassignite**, which contains mostly olivine. Together these three types of meteorites are referred to as the **SNC** (pronounced "snick") meteorites. The parent body for the SNC meteorites is conjectured to be Mars.

Meteorites from the Moon

Some meteorites have breccias with white clasts in darker matrix, like lunar rocks. These lunar origin of some meteorites was established without contention because we have many lunar samples which provided close geochemical and petrological matches to the meteorite samples. The identification of lunar meteorites opened the door to dynamical studies that subsequently established the possibility that meteorites can viably be deposited at Earth from Mars.

Ureilites

These meteorites contain olivine and pigeonite, and the matrix contains graphite or diamond. The carbon content suggests a link with carbonaceous chondrites.

1.13 Irons and Stony-Irons

Iron and stony iron meteorites make up several percent of the meteorite population. They constitute 4% of the meteorites that fall to Earth but are the most common find since they look so much different than the Earth's crustal rocks.

Stony iron meteorites consist of roughly equal parts of rock and iron, with the rock component consisting of olivine, the most common mantle material, and minor amounts of other silicate phases. Stony irons make up only about 1% of meteorites that fall to Earth.

The metallic component of these meteorites is predominantly iron with nickel in solid solution averaging usually about 10% but sometimes up to 20%. There are also smaller amounts of sulfide, graphite and occasionally silicate inclusions.

Within the iron there are two metal phases: the body-centered cubic (α) form kamacite (5.5% nickel) and the face-centered cubic (γ) taenite (variable, but usually > 27% nickel)

These phases occur because iron and nickel form a solid solution when mixed and are not completely miscible as they begin to cool. The iron and nickel are structurally similar but not identical. At high temperatures they exchange freely because the crystal lattice is expanded. But when cooling sets in their slight differences produce lattices with slightly different structures. At a point the total energy of the system is minimized by segregating the elements into 2 separate lattices: one rich in iron and the other poor. To minimize the mismatch where the lattices connect, newly formed lattices form in preferred orientations called exsolution lamellae.

Approximately 75% of iron meteorites exhibit a crystal pattern called **Widmanstatten structure**, which is the term used for these exsolution lamellae. The pattern is observed by taking a meteorite that is cut and polished and dipping it in acid. Because this pattern forms when an iron-nickel alloy crystallizes, it is an indication that some asteroids were at least partially melted after they formed. In fact, the details of the pattern tell the cooling history of the meteorite parent body from which it was derived. From the iron-nickel phase diagram we can see the evolution of the relative amounts of iron and nickel that crystallized as the iron cooled. And from the variations in composition across Widmanstatten structure boundaries it is possible to constrain the rate of cooling.

Iron Meteorites and Planetary Cores

In the iron-nickel system, equilibrium is maintained at temperatures above 650°C. Below 350°C crystal structures are frozen in. So the Widmanstatten structures yield the cooling rate in this temperature range. In the meteorites wide diffusion boundaries correspond to slow cooling and narrow diffusion boundaries correspond to rapid cooling. In the meteorites there are iron-nickel crystals that grown with lengths of up to several centimeters, which correspond to rates of $0.4 \rightarrow 40°$/Ma. So the Widmanstatten structures yield cooling rates of many millions of years. For a cooling periods in this range radii of meteorite parent bodies in the range 100–200 km are implied.

Identification of a significant metallic group within the meteorites made it the natural presumed component of dense planetary cores. It is known from the masses of the planets that the interiors are (in most cases), after correcting for the effect of self-compression, denser than rock. An iron-nickel alloy with approximate meteoritic proportions is the leading candidate for the dense component. Shock wave experiments indicate that seismic wave velocities in the Earth's deep interior are consistent with a predominantly iron composition and thus support the contention.

The other piece of evidence for iron planetary cores comes from the process of **nucleosynthesis**. Iron has the highest binding energy for nucleon and is thus highly stable and produced in abundance in stellar evolution. The **equilibrium process** (aka **e-process**) in stellar thermonuclear reactions breaks apart silicon atoms and re-arranges them to convert silicon to heavier and more stable nuclei. The most stable and thus most abundant element produced in the e-process is iron.

Table 1.14: Mean densities and moments of inertia for the terrestrial planets

Planet	Bulk Density (kg m^{-3})	C/MR^2
Mercury	5420	??
Venus	5250	0.34 (inferred)
Earth	5515	0.3335
Moon	3340	0.391
Mars	3940	0.366

1.14 The Terrestrial Planets

The table below shows a comparison of the mean densities and moment of inertia factors (C/MR^2) for the terrestrial planets. The mean densities are ”uncompressed” so they correct for self-compression and represent zero-pressure densities. The moment of inertia, which will be derived later in the semester, is a measure of the the extent to which mass is concentrated toward the center of a body with a value of 0.4 representing a uniform, homogeneous sphere and lesser values representing mass increasingly concentrated toward the center. This parameter reflects both self-compression and the presence of increasingly dense material toward the center.
Note that Venus and Earth are quite similar, consistent with their similarity in radius and mass. Mars has a lower bulk density and a higher moment of inertia core indicating less iron and a smaller core. Mercury is denser and while the moment of inertia hasn't been measured, it is thought based on the mass that its core is nearly 80% of the planet's radius. Mercury thus likely has a large iron core. The Moon has a bulk density similar to that of Earth's mantle and a moment of inertia just slightly less than that of a uniform sphere. The implied upper limit of core size is 350 km (out of a radius of 1738 km). Thus the Moon has very little iron in the interior.

1.15 One-dimensional Earth's structure

To a large extent, the radial stratification as inferred from seismology and constrained by astronomical and meteorite data represents the stable structure that results from differentiation processes as discussed by Maria Zuber. In addition to using this structure for the introduction of some useful terminology this radial structure serves as an important reference model for many geophysical investigations. The average properties are now fairly well known; it is fair to say that the uncertainties in the average values are insignificant as compared to the local and regional deviations from the reference value. Constraining and understanding the aspherical variations in physical properties is the objective of many geophysical studies. It involves the study of anomalies relative to a reference values (e.g., gravity and heat flow anomalies). The deviations from the reference values contain information pertinent to geodynamical processes.

Radial stratification of the Earth from seismic data

The essential data used to derive the depth variation of seismic wave speed in these early days of seismology were the travel times of different types of seismic waves. These can be determined from seismograms, which are records (analog or digital) of ground motion due to earthquakes or man-made explosions (e.g., nuclear tests).

Some history: Until late last century most research relevant to seismology was in fact done by physicists and mathematicians who loved to study elastic wave propagation (famous names that also contributed to seismology are **Navier**, **Poisson**, **Gauss**, **Rayleigh**). Therefore, most of the theory was already available at around the time of birth of observational seismology (late last century; installation of the first seismometers for systematic earthquake monitoring).
Following a rapid succession of discoveries in the early years of this century, the major subdivision of the Earth in concentric shells was established about 60 years ago, in the mid- to late thirties, by the pioneering work of **Jeffreys** and **Bullen** and by **Gutenberg** and **Richter**. The following dates give some idea about the pace of developments:

- 1892- First record of earthquake (Japan)
- 1906- Oldham demonstrates existence of core from seismic data
- 1909- Mohorovicic discovers seismic interface that marks the crust-mantle boundary (MOHO)
- 1912- Gutenberg estimates depth to core mantle boundary
- 1936- Lehman discovers existence of Inner Core
- 1939- early knowledge summarized in first 1D Earth models, the famous Jeffreys-Bullen tables (which are still surprisingly accurate)
- 1948- Bullen uses wave speed information + estimates of Earth's moment of inertia to determine average density as a function of increasing depth. After that, main focus on determining 3D structure. Mid seventies: pioneering work at MIT (K. Aki, M. N. Toksöz) and Harvard (A. Dziewonski) results in first realistic 3D models of seismic propagation speed in Earth's mantle.

How does this work? Just a brief account will do here: modern seismometers measure the three spatial components of this ground motion over in a wide frequency band and with a large dynamic range. What is important in these records is (1) the large variation in frequency (body waves and surface waves) and (2) the arrival of distinct seismic phases (such as P and S waves). For imaging purposes the difference in frequency controls the amount of detail that can be investigated with certain seismic data (resolution) whereas the different phases sample different parts of the Earth's interior. What's important for now is how the travel times measured from such records can tell us something about Earth's radial structure.
Let us assume for the time being that we know where and when an earthquake occurred (for instance from local damage reports). One can then construct so-called record sections, in which the seismograms are sorted according to the distance of the station to the earthquake epicentre (location at the surface). This distance (epicentral distance) is typically denoted by a Δ [°]. If many data are available one can determine best-fitting travel-time curves for different seismic phases. (Note: in practice we do not know the earthquake location and origin time (hypocentre) and the determination of source location and spatial variations in wave speed are intimately connected, with some nasty trade-offs).
The variation of the travel time as a function of distance (T(Δ) [s]) can be used to construct models of the variation of wave speed with depth. Using some simple physics one can also show that the knowledge of the wavespeed as function of depth (in combination with constraints such as the

average density and the moment of inertia (I 0.33 MR^2 for the Earth) can be used to deduce the radial variation in density. The relevant equation, the Adams-Williamson equation, is valid in regions where density is controlled by adiabatic compression and will be derived later.

In general, the wave speed increases with increasing depth in the mantle due to the effect of increasing pressure on the bulk modulus (κ) and rigidity (μ). (Note: *P* wave speed $\alpha = \sqrt{\kappa + \frac{4}{3}\mu/\rho}$; *S* wave speed $\beta = \sqrt{\mu/\rho}$).

An abrupt increase in both *P* and *S* wave speed occurs at a depth of about 10-40 km. This seismic discontinuity marks the crust-mantle interface and was discovered by Mohorovicic (hence "Moho" discontinuity). The definition (and thus the mapping) of the seismological Moho is primarily based on seismic wave speeds and does not necessarily coincide with a petrologic moho (See Anderson, 1988, for discussion). This leads to a subdivision in terms of crust-mantle-core; the subdivision in terms of crust-lithosphere-asthenosphere-mantle is based on thermal or mechanical properties. The latter nomenclature is, unfortunately, not unique and different scientists may mean different things, so beware!)

Between about 400 and 1000 km in depth the increase in wave speeds is larger than expected from adiabatic compression. This observation (first by Birch, 1952) has major implications for mantle dynamics and will be discussed in more detail later. The increase in now generally believed to be due to isochemical phase changes in the mantle silicates.

In the outer core, the P-wave speed decreases abruptly and this, in fact, causes the "shadow zone" for P- waves that was used as one of the major arguments by Oldham (1906) and Gutenberg (1912) to infer from seismological data the existence of a core and the depth to the core-mantle-boundary (CMB). The core behaves as a fluid (pointed out by Jeffreys in 1926 on the basis of tidal data, i.e. later than the discovery of the CMB!) even on short time scales so that shear waves cannot propagate through the (outer) core. The inner core (discovered by Inge Lehman (Denmark) in 1936) is solid; the seismological evidence for this will be discussed later. (Question: temperature increases monotonically with increasing depth, so why is there an alternation between depth intervals where rock is solid and where it behaves as a liquid?)

Composition

Despite centuries of geologic prospecting we have limited access to most parts of the Earth. consequently our understanding of its bulk composition must come from inference based on remote observation as well as the record from meteorites and the solar atmosphere. The dominant non-volatile constitutents in the sun are silicon [Si], magnesium [Mg] and iron [Fe]. The meteorites similarly are dominated by these elements and their oxides.

Some of the internal boundaries of the Earth represent phase changes that separate regions of the same bulk composition, others coincide with a change in chemistry. Seismology helps but can often not uniquely determine variations in temperature, composition, or density. There are often several mineral assemblages that have the same density or elastic properties. Using the cosmic abundance of elements and results of laboratory experiments at high pressures and temperatures with rock forming minerals and their analogs we have, however, arrived at the following broad picture of the average composition in each of the concentric shells:

(Recall that the most abundant non-volatile elements in the solar system are magnesium [Mg], iron [Fe], and silicon [Si].)

The **crust**, which is only 0.5% of the volume of the mantle, is rich in SiO_2 and Al_2O_3 (hence the old name SiAl) + CaO, Na_2O.

The composition of the **mantle** is roughly given by the following series of solid solutions, with primary constituents SiO_2 and MgO (hence the old name SiMa; also MaFic, i.e. magnesium and iron (Fe) rich):
Important are the following two pairs:

$$
\begin{array}{rcll}
(\mathrm{Mg}_\alpha, \mathrm{Fe}_{1-\alpha})\mathrm{O} + \mathrm{SiO}_2 & \rightarrow & (\mathrm{Mg}_\alpha, \mathrm{Fe}_{1-\alpha})\mathrm{SiO}_3 & \text{(orthopyroxene)} \\
2(\mathrm{Mg}_\alpha, \mathrm{Fe}_{1-\alpha})\mathrm{O} + \mathrm{SiO}_2 & \rightarrow & (\mathrm{Mg}_\alpha, \mathrm{Fe}_{1-\alpha})_2\mathrm{SiO}_4 & \text{(olivine)}
\end{array}
$$

with $\alpha = \mathrm{Mg}/(\mathrm{Mg} + \mathrm{Fe})$ approximately 0.9 (names used for olivine: $\alpha = 1$: forsterite; $\alpha = 0$: fayalite). The *magnesium number* is defined as $100 \times \mathrm{Mg}/(\mathrm{Mg} + \mathrm{Fe})$ and is often used in discussions of mantle rheology (a typical value is 90). Another important upper mantle mineral is garnet, also a (Mg,Fe) silicate. With increasing pressure in the mantle olivine transforms to spinel (same composition) and post-spinel. Pyroxene + garnet transform to perovskite in the lower mantle.
There is some debate whether the relative amount of iron increases in the **lower mantle** (perhaps with a small, 2%, increase in intrinsic density (i.e. in addition to the effect of adiabatic compression): such a small density contrast may not be significant from a dynamic point of view) and there is a growing consensus that the composition of the lower mantle is similar to that of the upper mantle and transition zone. However, due to the higher pressures the crystallographic structure is different, more compact. The most important constituents in the lower mantle are:

$$
\begin{array}{rl}
(\mathrm{Mg}, \mathrm{Fe})\mathrm{O} & \text{(Magnesio)wüstite} \\
(\mathrm{Mg}, \mathrm{Fe})\mathrm{SiO}_3 & \text{(Perovskite)}
\end{array}
$$

There is also some debate about the exact ratio of pyroxene to olivine. The solar abundance would favor a Mg/Si ratio close to 1, which would predict a predominance of pyroxene, but the presence of Fe allows for more olivine and the ratio that matters is the (Mg+Fe)/Si ratio. The thermodynamics of the phase transformations in the two silicate systems (olivine and pyroxene) are very important for our understanding of mantle dynamics, in particular for the question as to whether material can flow across the transition zone into the lower mantle. Because of their importance for mantel dynamics, phase changes will be discussed in more detail later in this course.
The major outstanding problems pertinent to the mantle include the scale of mantle convection, the effectiveness of mixing, and the survival of separate reservoirs of compositional heterogeneity as inferred from isotope data. We will come back to these exciting topics towards the end of the course.
The **core** (32% of the mass of the Earth; how much of its volume?) largely consist of iron. Fe is the only heavy element with large enough solar abundance to be a suitable candidate to form the heavy element required to explain the large density of the Earth's core. However, if the core would consist entirely of metallic iron, its density would be higher than required from the moment of inertia and the mean density of the Earth. There must, therefore, be a light alloying element. Several of the light elements that are abundant in the solar system are either too volatile or are insoluble with metallic iron. The most likely candidate is oxygen (O) (which is insoluble with metallic iron at low pressures and is thus not found in iron meteorites), although some would argue for Sulfur (S) (which is found in iron meteorites), or Silicon (Si). (Stacey uses 80% O

+ 20% S, but *S* does not seem to be required). This is still an area of active research; because of the extreme physical conditions there are hardly any experimental data $\rightarrow$ lot of theoretical (thermodynamics) studies of Equations Of State (EOS) (See Anderson, 1988, for an introduction). The **inner core** (IC) has a radius of about 1220 km the IC comprises less than 1%(!) of the volume of the Earth, but it represents about 1.7% of the Earth's mass. Because of the small size of the IC, the uncertainty of its density is relatively large. But within the uncertainty, the IC may simply be a "frozen" version of the OC, and thus heavier since the lighter elements are selectively rejected, with as major constituents Fe_2O, FeNiO, pure Fe or a Fe-Ni alloy. Some compositional layering of the core is probable. (Anderson (1988) is a good reference for core composition).
Because of the recent developments in understanding the structure of the inner core (anisotropy, net rotation relative to the Earth's mantle) and the importance of the core for the Earth's magnetic field, the core will be discussed in detail during this course (including reading assignments). (Some outstanding problems: what maintains the IC-OC boundary? OC-lower mantle boundary? What produces the energy to drive core convection that produces the magnetic field? Is there convection in the solid IC?).

1.16 Lateral heterogeneity in the mantle

Introduction

The outer core behaves as a liquid and its low viscosity cannot sustain lateral variations in density and can thus be regarded as homogeneous for many practical purposes. The mantle, however, is probably heterogeneous at all length scales. To a large extent this heterogeneity can be attributed to convective circulation in the mantle and the recycling of (oceanic) lithosphere. The most important dynamic processes will be discussed later. This heterogeneity causes observed seismic data to differ from predictions from a simple radially stratified reference model. Later in this course I will show how certain imaging techniques can be used to interpret these travel time residuals to map the aspherical structure of the Earth's interior.
Plate Tectonics forms an integral aspect of this convective system and a relevant discussion of plate tectonics thus involves more than the essentially kinematic concepts that were outlined in the 1960-ies. Therefore, plate tectonics, and its relationship with mantle convection, will be discussed in the second half of this course. However, to facilitate communication it is useful to introduce some jargon and define important concepts, processes, and their resulting structures.

Plate boundaries as belts of high seismic activity

Once travel time table are established from a redundant seismological data set they can be used to locate earthquakes. The pattern of seismicity is now well established and defines the locations of the plate boundaries. Note that the distribution of earthquakes within plate boundaries (or plate boundary zones) varies, which gives important information about the deformation in certain regions. On the basis of seismic belts about 12 major plates have been recognized. Important aspects of kinematic plate-tectonic theory are that (1) the deformation within the plates (*intraplate*) is neglected; all deformation is assumed to take place between plates (*interplate*) (if need be plates are subdivided into smaller units until this condition is met). In other words, plates act as stress guides, which is important for the understanding of the driving forces of plate tectonics; (2) once new oceanic lithosphere is created it forms part of a rigid plate; the plate may or may not contain continents; (3) in order to conserve the total area of the Earth's surface: the rate of lithosphere creation equals rate of lithosphere destruction. The seismograms also contain information about the physical rupture mechanism of earthquakes and can thus be used in the classification of different types of plate boundary.

divergent plate boundaries Associated with "normal" faults (plates move away from each other). Characteristic in Mid Ocean Ridge (MOR) systems; locations of sea floor spreading where new oceanic crust is being created rather passively by means of decompression melting. Such boundaries are sometimes referred to as *accreting margins* but this is confusing since an important mechanism of continental growth is by means of accretion of allochthonous terranes (accretionary wedges, accreted terranes) onto Precambrian shields, for instance the western cordillera in Canada and eastern Australia.

convergent plate boundaries Associated with "reverse" faults (plates move towards each other). Sometimes referred to as destructive margins. Characteristic in circum-Pacific belt. Often accompanied with mountain building, subduction, deep seismicity, and arc volcanism. Diagnostic bathymetric feature: deep sea trenches.

transcurrent plate boundaries Associated with strike slip faults. Important in particular in association with ocean floor spreading → transform faults. There is a separate classification of transform faults based on whether they connect ridges or trenches, but that's not so exciting. Most common are ridge-ridge transforms. These are locations of many earthquakes since slip rate is twice as fast as spreading rate!

Lithosphere, asthenosphere, subduction zones

We will use the term **lithosphere**[2] for the mechanically strong outer shell of the earth, which contains the crust and part of the mantle. It is a thermal boundary layer (TBL) through which heat is lost to the surface by conduction. There is also a thermal definition, in which the base of the lithosphere coincides with the 1300°C isotherm (up to 100 km deep for oceans; at least twice that much for some (parts of) continents), but the mechanically strong (elastic) part of the lithosphere that can transmit stresses and support surface loads (e.g., sea mounts) on geologic time scales is about half the TBL thickness (coinciding with the 650°C isotherm). The cooling of oceanic lithosphere (+ relationship between lithospheric thickness and heat flow with square root of its age since formation at the MOR) is one of the "classical" aspects of plate tectonics that will be discussed in the course of this term. In contrast, deformation in the **asthenosphere**[3] occurs more freely by means ductile creep on geologic time scales. The asthenosphere may "lubricate" but does not participate directly in tectonic motion, although it could accommodate some of the return flow. Heat loss in asthenosphere and deeper mantle is controlled by convection.
As oceanic lithosphere moves away from the MOR it will cool and thicken and become more dense due to thermal contraction. Eventually, it will become gravitationally unstable and can, in principle, sink into the Earth's interior at ocean trenches (note: initiation of subduction is not trivial since even gravitationally unstable parts of the lithosphere can be retained at the surface by the strength of the lithosphere). We will use the term **subduction zone** rather loosely for the geographical region where plate convergence results in the descent of one plate beneath the other,

[2] $\lambda\iota\theta o\varsigma$ or *lithos*=stone [Greek]
[3] $\alpha\sigma\theta\epsilon\nu\eta\varsigma$ or *a-sthenes*= not-strong [Greek]

typically the more **m**a**f**ic (**M**g+**F**e rich) oceanic beneath the more acid, granitic continental plate. The result of this process is a **slab of subducted former oceanic lithosphere** that sinks into the mantle. The negative buoyancy of the downgoing slab is, in fact, one of the most important forces of plate motion. The depth extent of slabs is still debated, but that does not — according to these definitions — influence the meaning of a *subduction zone*. In the upper mantle and transition zone the slabs can be delineated by earthquakes (the so called Wadati-Benioff zones), the slab is said to be seismogenic down to about 670 km depth (the actual depth to the deepest earthquakes can be smaller depending on the thermal structure of the slab). There is now convincing evidence to support the aseismic continuation of slabs into the lower mantle.