Chapter 11

Origin of the Solar System

In this chapter we examine the origin of the solar system. A successful model for solar system formation must account for a large number of observed features, such as the nature of planetary orbits, the age of the Sun and the planets and their composition. Furthermore we should be able to apply this model to other stars and planetary systems currently being born.

The current view is that the solar system originated from the collapse of a large gas cloud some 4.6 billion years ago. This *nebular hypothesis* was first put forward by Kant in 1755 and later developed by Laplace and others.

Other models that have since been excluded include so-called catastrophe theories, which were widely discussed in the 19th century. In these models the initially isolated Sun had a close encounter with another star. Material from this star would be pulled into orbit around the Sun and then condense into planets. But given the observed number density of stars in the galaxy, the probability of such a close encounter is very low. Furthermore one would not expect such an event to leave enough matter orbiting around the Sun to form the planets we observe today.

Another possibility is that the Sun and planets formed separately but that the planets at some point wandered close to the Sun and were captured. An important difficulty with this theory is that it does not explain why the planets are all found orbiting in roughly the same plane, and all in the same direction.

We will not go into the details of any of the alternative models here but rather will concentrate on the nebular hypothesis.

11.1 Overview of solar system formation

Below we outline some of the main events in the formation of the solar system.

- A cold gas cloud is perturbed, perhaps by a nearby supernova explosion, and collapses under its own gravity.
- As the cloud collapses it heats up, and energy is radiated away as light. The loss of energy together with conservation of angular momentum force the material of the solar system into a rotating disc. The time scale for the initial collapse is roughly 10^6 years.

- The gas cools enough for condensation of metals, dust particles.
- Dust particles collide, form larger particles up to boulder size.
- Some boulders grow large enough for their own gravity to play a significant role in attracting more material. This accelerates their growth (runaway growth) and planetesimals are formed.
- At around 10^7 years after the initial collapse, the sun ignites and the solar wind expels the remaining gas.
- Planetesimals collide and coalesce, forming protoplanets.
- Some 10^8 years after the start the solar system has roughly its present form. The planets often carry a characteristic of their last large collision.
- The accretion process continues as asteroids and comets collide with planets leaving impact craters. This was particularly intense for the first 500 million years, called the heavy bombardment era, but asteroid impacts continue today.

We will now look at these steps in greater detail starting with the nature of the gas clouds from which the solar system is thought to have formed.

11.2 The solar nebula

We see many thousands of clouds of gas and dust in the galaxy showing up as dark patches blocking light from a field of more distant stars. An example is the molecular cloud Barnard 68, shown in Fig. 11.1(a). The temperature of the clouds can be estimated from the thermal broadening of spectral lines. Typical values are in the range from 10 to 50 K, mostly consisting of molecular hydrogen (H_2), and a typical size is on the order of 10^5 AU. The density of the gas can be estimated from the strength of the cloud's radio emission, with typical values of 10^{10} molecules per m^3.

One of the strongest pieces of evidence supporting the nebular hypothesis is the observation of stars currently being formed from the collapse of gas clouds. An example is the Eagle Nebula, part of the open cluster M16, shown in Fig. 11.1(b). The long pillars are clouds of gas and dust, and the bright spots at the tips are newly formed stars.

Although the overall composition of our solar system is dominated by hydrogen and helium, it contains significant amounts of heavier elements, such as silicon, iron, etc. In the first several minutes after the Big Bang, however, only hydrogen and helium were formed, with very small quantities of slightly heavier elements like lithium (see e.g. [41]). Essentially all of the heavier elements formed much later in the interiors of stars. This means that the solar nebula must have contained products from earlier generations of stars.

We will now discuss the conditions under which a cloud is expected to collapse. We will use an important result called the *virial theorem*, which relates the time average kinetic and potential energies, $\langle K\rangle$ and $\langle U\rangle$, of a gravitationally bound system in equilibrium. The theorem states that for such a system,

Figure 11.1: (a) The molecular cloud Banard 68 [39]. (b) Star formation in the Eagle nebula [40].

$$\langle K \rangle = -\frac{1}{2}\langle U \rangle \ . \tag{11.1}$$

A derivation of the virial theorem is given in Appendix B. Here we are using the convention that the potential energy of a system is zero when its constituents are separated to infinite distance, so for a bound system the potential energy is negative.

For a cloud of a given mass M, temperature T, and composition we can work out the kinetic and potential energies, K and U. Here the time averaging is not important since there are so many particles in the system that both K and U are individually very close to constant. For the cloud to be gravitationally stable, equation (11.1) must hold. Ignoring the time averaging means we will have $K = -\frac{1}{2}U$. If we find that the kinetic energy is smaller than $-\frac{1}{2}U$, then this means the cloud cannot be in equilibrium, but rather it will collapse to a smaller size.

Consider a spherical cloud of radius R and constant density ρ. Of course a real cloud will not have a uniform density but this model will nevertheless reveal the main issues involved. The mass of the cloud is

$$M = \frac{4}{3}\pi R^3 \rho \ . \tag{11.2}$$

From dimensional arguments we know that the gravitational potential energy of the cloud must be $U \sim -GM^2/R$. The exact value will depend on how the mass is distributed inside the volume. If we assume a spherical cloud with uniform density we can show

$$U = -\frac{3}{5}\frac{GM^2}{R} \ . \tag{11.3}$$

Now to find the kinetic energy for a cloud of a given composition and temperature we use the equipartition theorem. If we ignore rotational and vibrational motion of the molecules we have

$$K = \frac{3}{2}kTN \;, \tag{11.4}$$

where k is Boltzmann's constant and N is the number of molecules. This can be written

$$N = \frac{M}{m_{\text{molecule}}} = \frac{M}{\mu m_{\text{H}}} \;, \tag{11.5}$$

where here M is the mass of the cloud and the average molecular mass m_{molecule} can be expressed as the mean molecular weight μ times the mass of a hydrogen atom m_{H}.

Substituting these ingredients into the virial theorem, equation (11.1), we find

$$\frac{3}{2}\frac{MkT}{\mu m_{\text{H}}} = \frac{3}{10}\frac{GM^2}{R} \;. \tag{11.6}$$

The kinetic term on the left is proportional to $M \propto R^3$, while the potential term on the right goes as $M^2/R \propto R^5$. Thus for a cloud of a given density and a sufficiently large size, the potential term will dominate and the cloud will collapse. The radius at which the equilibrium condition (11.6) just holds is called the *Jeans radius* R_{J}. Solving equation (11.6) for the radius and using also (11.2) for the mass we find

$$R_{\text{J}} = \sqrt{\frac{15kT}{4\pi G \mu m_{\text{H}} \rho}} \;. \tag{11.7}$$

Equivalently we can use equation (11.2) to solve for the mass at which the equilibrium condition holds, which gives the *Jeans mass*,

$$M_{\text{J}} = \sqrt{\frac{3}{4\pi\rho}} \left(\frac{5kT}{\mu m_{\text{H}} G} \right)^{3/2} \;. \tag{11.8}$$

By making different assumptions about the initial shape and mass distribution of the cloud we would find expressions for R_{J} and M_{J} similar to these but which could contain slightly different numerical factors.

We can now plug in some numbers and estimate the Jeans radius for a typical gas cloud. Let us take a temperature of $T = 20\,\text{K}$ and a number density of $n = 10^{10}$ molecules per m^3. For molecular hydrogen we have a molecular weight of $\mu = 2$, so the density is

$$\rho = n\mu m_{\text{H}} = 10^{10}\,\text{m}^{-3} \times 2 \times 1.67 \times 10^{-27}\,\text{kg} = 3.3 \times 10^{-17}\,\text{kg/m}^3 \;. \tag{11.9}$$

Using this in equation (11.7), we find

$$R_{\text{J}} = 6.7 \times 10^{15}\,\text{m} = 0.22\,\text{pc} = 4.5 \times 10^4\ \text{AU} \;, \tag{11.10}$$

and from (11.8) we have the corresponding Jeans Mass,

$$M_{\text{J}} = 4 \times 10^{31}\,\text{kg} \approx 20 M_{\text{sun}} \;. \tag{11.11}$$

Only some fraction of the cloud's mass will wind up in the star, so $20M_{\rm sun}$ for the initial mass of a collapsing gas cloud is not unreasonable as an order of magnitude estimate.

The value of $R_{\rm J}$ from (11.10) also seems reasonable in that we see many thousands of clouds with sizes in this range in our galaxy. This then presents a problem in that one would naively expect them to collapse in a time of roughly 10^6 years (see Section 11.3 below). Given that our galaxy formed more than 10^{10} years ago it seems that these clouds should have collapsed long ago. It is generally assumed that other important effects come into play, in particular that of magnetic fields. These can act to stabilize a cloud that otherwise would collapse under its own gravity.

In many galaxies, however, we find a very high rate of formation of new stars, perhaps 10^6 stars created in a timescale of 10^7 to 10^8 years. This rate is so high that it could not possibly be sustained throughout the galaxy's lifetime and therefore it is assumed to be a transient phenomenon. This is called a *starburst galaxy*, where the collapse of a very large region of gas is triggered by an encounter with another galaxy.

So although we have a partially coherent picture of how a gas cloud can begin to collapse, the details remain a bit murky. The situation improves when we refine the model and include magnetic effects but still it is difficult to make accurate predictions for the rate of star formation. The Jeans radius (or mass) nevertheless remains the fundamental quantity to consider, even though the naive calculation above fails to explain the long lifetimes of large, cold clouds. These quantities are also used when we try to understand the collapse of the gas in the early universe into galaxies (see, for example, [42]).

11.3 Contraction of the solar nebula

If the gas cloud were to consist of noninteracting particles with zero initial kinetic energy, they would all simply fall toward the centre. The time that it takes to do so is called the *free-fall time*, $t_{\rm ff}$. A particle at the outer edge of the cloud at a radius R would fall as if all of the rest of the cloud's mass M were concentrated at the centre. The trajectory would be the limiting case of the elliptical Kepler orbit, here with zero angular momentum. The distance to the centre, R, is simply twice the semi-major axis a of an ellipse with an eccentricity of unity and the centre-of-mass (the eventual star) at one focus. The time to cover this distance is one half of the period, P. Kepler's third law tells us $P^2 = \frac{4\pi^2}{GM}a^3$, and we can also use $M = (4/3)\pi R^3 \rho$, where here R and ρ are the initial radius and density. We therefore have for the free-fall time

$$t_{\rm ff} = \frac{P}{2} = \frac{1}{2} \cdot \frac{2\pi}{\sqrt{GM}} a^{3/2} = \frac{\pi}{\sqrt{GM}} \left(\frac{R}{2}\right)^{3/2} = \left(\frac{3\pi}{32G\rho}\right)^{1/2} . \tag{11.12}$$

Notice that for a given density, the free-fall time is independent of the mass or radius of the cloud. Substituting numbers from the example above gives $t_{\rm ff} \approx 400\,000$ years.

The free-fall time characterizes the *dynamical timescale*, the timescale for motion in the absence of pressure. But we know the interior temperature and pressure must eventually rise, and this slows the time for collapse. Initially, dust particles collide and radiate in the infrared, and for sufficiently low densities this radiation can escape and the cloud remains cool. Eventually, however, the optical thickness of the cloud exceeds unity and the radiation cannot escape without further interaction, and thus the cloud heats up.

After roughly one million years, the solar nebula had collapsed to protostar. The core temperature was still too low to allow nuclear burning, and the energy source was still based on the conversion of gravitational potential radiation into thermal energy, which is then radiated away as light. This light and the solar wind expelled gas and dust from the surrounding region and the Sun became optically visible. This is what is called the T-Tauri phase of stellar formation.

If gravitational energy were to be the only source of light from the sun, then the timescale on which the Sun would radiate away all of the initial potential energy, called the *Kelvin-Helmholtz timescale*, would be

$$t_{\mathrm{KH}} \approx \frac{U}{L} \,, \tag{11.13}$$

where U is the available potential energy and L is the average luminosity, i.e., the power output of the Sun. If we plug in numbers based on the current size and luminosity of the Sun, we find $t_{\mathrm{KH}} \approx 3 \times 10^7$ years. When this was first worked out by Kelvin in the 19th century, however, there was already geological evidence that showed the age of the earth to be significantly older than this. So one was confronted with a sort of 'age crisis', whose resolution required an understanding of nuclear physics that came only many decades later.

Eventually the temperature in the core exceeded the nuclear ignition temperature of around 10^7 K and nuclear fusion took over as the energy source for the Sun's light. The timescale for our Sun to reach this stage was on the order of 10^7 years. The timescale for the Sun to exhaust its supply of hydrogen through nuclear burning is more like 10^{10} years, which is compatible with its currently estimated age of 4.6×10^9 years.

11.4 Disc formation

During the collapse of the molecular cloud the solar system flattened to a disc. Initially the cloud will have had some small rotational motion, e.g., from the turbulence of the interstellar medium. At the very least some rotation of the gas cloud would arise from the fact that the parts of the cloud closer to the galaxy's centre have a shorter period of galactic rotation. As the cloud contracts, angular momentum is conserved and so therefore the rate of angular rotation must increase. If particles move towards the midplane (perpendicular to the cloud's axis of rotation), then they can collide inelastically whereby kinetic energy is converted to heat and eventually released as blackbody radiation. Such a process, shown in Fig. 11.2(a), involves no change in the total angular momentum of the system.

Let us now view the gas cloud along its axis of rotation as in Fig. 11.2(b). If two particles collide inelastically, kinetic energy is converted to heat, but the total angular momentum must be conserved. For this to happen the particles will move into orbits with the largest angular momentum for a given kinetic energy — this is a circular orbit. So the net result of angular momentum conservation and inelastic collisions is to move the particles into circular orbits lying in a central plane.

Notice that if there was no energy loss mechanism, then the flattening would not happen. If the collision in Fig. 11.2(a) were to be elastic, then the outgoing momentum components along the axis of rotation would be equal to those incoming, and the cloud would remain spherical.

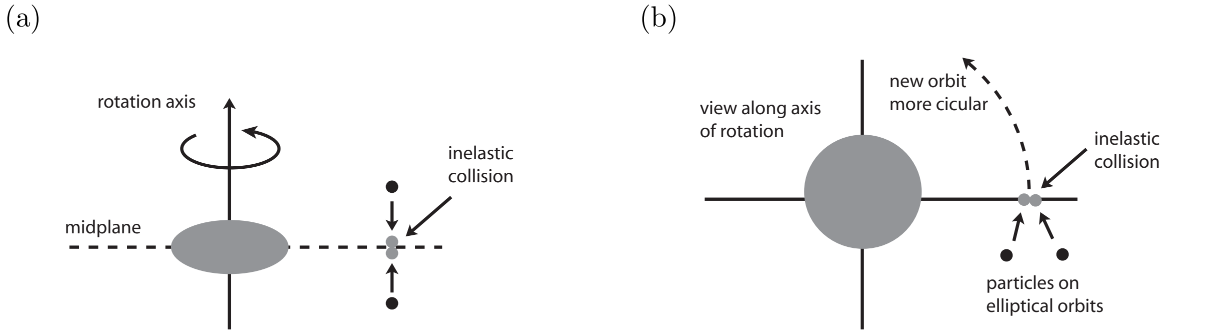

Figure 11.2: Two views of the contracting solar nebula, (a) side view and (b) top view, along the rotational axis. (See text.)

This is similar to what 'dark matter' particles (called WIMPS) are believed to do in galaxy formation.

11.5 Formation of planets

In preparation