COSMOLOGY

Lecturer: Professor Komissarov S.S
Room: 10.19 in Maths Satellite
email: S.S.Komissarov@leeds.ac.uk

(2012)
Contents

1 The Structure and Contents of Visible Universe 7
  1.1 Structure of visible Universe 7
  1.2 Expansion of visible Universe 10
  1.3 Contents of visible Universe 11

2 Gravity and Space-Time 13
  2.1 Space and time of Newtonian physics 13
     2.1.1 Space 13
     2.1.2 Time 15
     2.1.3 Galilean relativity 16
     2.1.4 The lack of speed limit 16
     2.1.5 Light 16
  2.2 Special relativity 17
  2.3 General Relativity 21

3 Friedmann’s equations 25
  3.1 Metric of the Universe 25
  3.2 Derivation of Friedmann’s equations in General Relativity 28

4 Basic Cosmological Models 31
  4.1 Einstein’s static Universe 31
  4.2 Key parameters of the Universe 32
     4.2.1 Expansion and deceleration parameters 32
     4.2.2 Critical density 34
  4.3 The Friedmann models 36
     4.3.1 Flat Universe 36
     4.3.2 Closed Universe 37
     4.3.3 Open Universe 38
  4.4 The Big Bang 39
  4.5 Advanced topic: Cosmic microwave background radiation 40
     4.5.1 Relic radiation 40
     4.5.2 Black body spectrum 41
     4.5.3 History of CMB radiation 42
     4.5.4 Evolution of CMB radiation after decoupling with matter 43
  4.6 Some other stages of Big Bang 44
     4.6.1 Nucleosynthesis 44
     4.6.2 Structure formation 44
### Table 1: Physical and Astronomical Units

<table>
<thead>
<tr>
<th>Unit</th>
<th>Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>erg</td>
<td>1 g cm² s⁻¹</td>
</tr>
<tr>
<td>eV</td>
<td>1.602 × 10⁻¹² erg</td>
</tr>
<tr>
<td>keV</td>
<td>10³ eV</td>
</tr>
<tr>
<td>MeV</td>
<td>10⁶ eV</td>
</tr>
<tr>
<td>AU</td>
<td>1.496 × 10¹³ cm</td>
</tr>
<tr>
<td>ly</td>
<td>9.463 × 10¹⁷ cm</td>
</tr>
<tr>
<td>pc</td>
<td>3.086 × 10¹⁸ cm</td>
</tr>
<tr>
<td>kpc</td>
<td>10³ pc</td>
</tr>
<tr>
<td>Mpc</td>
<td>10⁶ pc</td>
</tr>
<tr>
<td>yr</td>
<td>3.156 × 10⁷ s</td>
</tr>
<tr>
<td>M☉</td>
<td>1.99 × 10³³ g</td>
</tr>
<tr>
<td>L☉</td>
<td>3.9 × 10³³ erg cm⁻² s⁻¹</td>
</tr>
</tbody>
</table>

### Table 2: Fundamental Physical Constants

<table>
<thead>
<tr>
<th>Physical Constant</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>speed of light</td>
<td>$c = 2.998 \times 10^{10}$ cm s⁻¹</td>
</tr>
<tr>
<td>gravitational constant</td>
<td>$G = 6.672 \times 10^{-8}$ cm³ g⁻¹ s⁻²</td>
</tr>
<tr>
<td>Plank constant</td>
<td>$h = 6.626 \times 10^{-27}$ cm² gs⁻¹</td>
</tr>
<tr>
<td>Boltzmann constant</td>
<td>$k = 1.381 \times 10^{-16}$ erg K⁻¹</td>
</tr>
<tr>
<td>Thompson cross-section</td>
<td>$\sigma_T = 6.652 \times 10^{-23}$ cm²</td>
</tr>
<tr>
<td>Radiation constant</td>
<td>$a_r = 7.565 \times 10^{-15}$ erg cm⁻³ K⁻⁴</td>
</tr>
<tr>
<td>electron mass energy</td>
<td>$m_e c^2 = 0.511$ MeV</td>
</tr>
<tr>
<td>proton mass energy</td>
<td>$m_p c^2 = 938.3$ MeV</td>
</tr>
<tr>
<td>neutron mass energy</td>
<td>$m_n c^2 = 939.6$ MeV</td>
</tr>
</tbody>
</table>
Chapter 1

The Structure and Contents of Visible Universe

1.1 Structure of visible Universe

The nature and structure of the Universe has been debated during the whole history of Humanity. Before the advance of scientific method, our understanding of the Universe was rather primitive and based on descriptive interpretation of everyday life experiences. The human habitat was placed at the center of the Universe.

With the advance of science, our understanding of the Universe was revolutionized many times. Ancient Greeks already deduced that the Earth was not flat but more like a ball. According to the theory developed by Ptolemy, and adopted by Christian Church during the Middle Ages, it was in the center of The Universe with the Moon, the Sun, and other planets revolving around it. Stars were located further out and fixed to the Celestial Sphere. Only in 16th century, due to the efforts of Copernicus, a different other idea, also originated in ancient Greece (Aristarchus), began gradually to take over. Now, the Sun was in the center and the Earth and other planets were orbiting it. Stars were still attached to the Celestial Sphere and were considered as things very different from the Sun and the planets. Newtonian Mechanics and his Theory of Gravity provided an excellent mathematical description of the planetary motion and laid a solid foundation to the Copernicus Universe. After the invention of optical telescopes, it gradually became clear that the stars are objects like our Sun, but located very far away. This discovery had almost completely rejected the anthropocentric view of the Universe (where Humanity is placed into its center).
Figure 1.1: Open stellar clusters (left image) contain hundreds of young massive stars, whereas globular clusters (right image) contain around a million of old stars with mass similar to our Sun.

However, stars are not uniformly scattered in space. Often they come in clusters like those shown in Fig.1.1. These star clusters are held together by the force of gravitational attraction - they are gravitationally bound systems.

Moreover, on the sky the stars and stellar clusters concentrate towards the Milky Way, suggesting that there exist a larger stellar structure in the form of a disk, and that the Sun is near the disk plane. This structure was called the Galaxy and for a while it was believed to be the whole Universe, surrounded by empty space. Then came the realization that many diffuse objects on the sky were in fact similarly large collections of stars outside of the Galaxy. These galaxies come in different shapes and sizes, and our Galaxy is a rather typical representative. From a distance it should look like the one shown in Fig.1.2. It contains around $10^{11}$ stars (one hundred billions). Galaxies are also gravitationally bound systems.

Figure 1.2: A spiral galaxy which very similar to our own Galaxy.

The complex hierarchical organization of the Universe does not stop at galaxies. Galaxies themselves combine both into small groups with only few members and in very large clusters which may contain tens of
thousands of members (see Figure 1.3). These clusters are the largest gravitationally bound systems in the Universe. However, they are still not uniformly distributed in space but grouped into super-clusters, joined by “filaments” and “walls” of galaxies. The result is a foam-like structure with large “voids” (see Figure 1.4).

Figure 1.3: A typical cluster of galaxies.

Figure 1.4: The “foam-like” large scale structure of the Universe revealed as the result of a redshift survey of galaxies.

On the background of all this visible structure we something like an opaque screen at the temperature of $\sim 2.7$ K. This is the boundary of visible Universe. It produces thermal electromagnetic emission, most of it as microwaves. This emission is remarkably the same in all directions on the sky, it exhibits only really tiny fluctuations which could not be detected only until very recently. This emission is called the Cosmic Microwave Background (CMB).

Table 1.1 shows the typical length scales of the structures described above. The most convenient unit of length for the Solar system is the astronomical unit (AU), this is the mean radius of the Earth’s orbit ($1\text{AU} = 1.496 \times 10^{13}$ cm). The traditional unit of stellar astronomy is parsec (pc) – this the distance from which the angular size of the Earth’s orbit is exactly one arcsecond ($1\text{pc} = 3.086 \times 10^{18}$ cm $\approx 2.062 \times 10^5$ AU). Other traditional units of stellar and extra-galactic astronomy are kiloparsec ($1\text{kpc}=1000\text{pc}$) and megaparsec
Figure 1.5: Fluctuations of the Cosmic Microwave Background as measured during the recent NASA’s space mission WMAP.

Table 1.1: Characteristic length scales in the Universe

<table>
<thead>
<tr>
<th>Distance between stars</th>
<th>1 pc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size of a galaxy</td>
<td>30 kpc</td>
</tr>
<tr>
<td>Distance between galaxies</td>
<td>1 Mpc</td>
</tr>
<tr>
<td>Size of voids</td>
<td>50 Mpc</td>
</tr>
<tr>
<td>Size of observable Universe</td>
<td>$10^4$ Mpc</td>
</tr>
</tbody>
</table>

(1Mpc=1000 kpc).

1.2 Expansion of visible Universe

Soon after the discovery of galaxies it was noticed that lines in their spectra were systematically shifted towards red. That is their observed wavelength $\lambda_r$ was larger than the one measured in laboratory, $\lambda_e$. The relative magnitude of this cosmological redshift

$$z = \frac{\lambda_r - \lambda_e}{\lambda_e}$$  \hspace{1cm} (1.1)

is normally used in astronomy to describe this effect and it is often called the cosmological redshift too. Moreover, the American astronomer Edwin Hubble (see fig. 1.6) discovered that this redshift was stronger for distant galaxies, indicating the simple law

$$z = \frac{H_0}{c} r.$$  \hspace{1cm} (1.2)

Hubble explained his observations in terms of the Doppler effect. When a source of light is moving relative to the observer the wavelength of received light, $\lambda_r$, is different from the wavelength of the emitted light, $\lambda_e$, and depends on the velocity of the source. For velocities much smaller compared to the speed of light we have

$$\lambda_r = \lambda_e(1 + V/c),$$  \hspace{1cm} (1.3)
and

$$z = \frac{V}{c} \quad \text{(1.4)}$$

where $c$ is the speed of light and $V$ is the radial component of velocity (positive when the distance increases and negative otherwise). This is known as the Doppler effect. It can be used to measure the source motion when most other methods fail and it is particularly useful in astronomy, dealing with very distant sources. (In fact, when the spectrum of the source radiation is a featureless continuum this effect is not easy to use either. Fortunately, in many cases the spectrum contains very fine features, emission and absorption lines, reflecting the quantum nature of atoms, and this allows to make very accurate speed measurements.) One can see that if the source is moving away than $\lambda_r > \lambda_\nu$, and thus the optical lines are shifted towards longer wavelengths, or towards the red end of the spectrum. So Hubble concluded that distant galaxies move away from us. Combining Eq.1.2 with Eq.1.4 we find that

$$V = H_0 r \quad \text{(1.5)}$$

This is what is usually called the Hubble law. However, modern Cosmology, based on General Relativity, interprets the cosmological redshift not in terms of the relative motion of galaxies but rather as an expansion of the Universe as a whole. The difference between the two viewpoints can be illustrated by this analogy. The Hubble interpretation corresponds to bugs on a balloon all running away from a given point. The modern interpretation corresponds, to bugs sitting quietly on the balloon, but the balloon itself being inflated instead.

![Figure 1.6: Edwin Hubble and the telescope (“Hooker”, Mount Wilson observatory, USA) with which he discovered the expansion of the Universe.](image)

### 1.3 Contents of visible Universe

Stars, galaxies, interstellar, and intergalactic gas constitute the visible matter in the Universe, traditionally called just matter. Thermal energy of these objects is very small compared to their rest-mass energy and can be ignored when their gravitational interactions are considered.

Observations also indicate the presence of invisible, or dark matter, which gives itself away via gravitational interactions with visible matter. Paradoxically, it dominates the visible matter in total mass. The nature of dark matter is still a mystery. Most likely it is made of some exotic particles which do not produce electromagnetic radiation of any type.
Then there is electromagnetic radiation (photons), Cosmic Rays and other relativistic particles which are similar in some important respects to photons. To be more specific, their rest mass-energy is only a small fraction of their total mass-energy, and therefore can ignored in gravitational interactions. We will explain the meaning of rest mass-energy later on. In Cosmology, all these relativistic components are generally referred to as radiation.

Finally, there seems to exist another exotic component which dominates even the dark matter in terms of its effect on the evolution of the Universe at present time – the so called dark energy. In contrast to all other components it’s gravity is repulsive. We will discuss both the dark matter and dark energy later in the course.
Chapter 2

Gravity and Space-Time

The evolution of the Universe is described by the Einstein’s General Relativity. This is a rather complicated theory and requires a high level of mathematical preparation. Therefore, only a very brief description of the Theory of Relativity and its application to Cosmology will be given here, with most results presented without derivation.

2.1 Space and time of Newtonian physics

2.1.1 Space

The abstract notion of physical space reflects the properties of physical objects to have sizes and physical events to be located at different places relative to each other. In Newtonian physics, the physical space was considered as a fundamental component of the world around us, which exists by itself independently of other physical bodies and normal matter of any kind. It was assumed that 1) one could “interact” with this space and unambiguously determine the motion of objects in this space, in addition to the easily observed motion of physical bodies relative to each other, 2) that one may introduce points of this space, and determine at which point any particular event took place. The actual ways of doing this remained mysterious though. It was often thought that the space is filled with a primordial substance, called “ether” or “plenum”, which can be detected one way or another, and that “atoms” of ether correspond to points of physical space and that motion relative to these atoms is the motion in physical space. This idea of physical space was often called “the absolute space” and the motion in this space “the absolute motion”.

There also was a consensus that the best mathematical model for the absolute space was the 3-dimensional Euclidean space. By definition, in such space one can construct cuboids, rectangular parallelepipeds, such that the lengths of their edges, \(a\), \(b\), and \(c\), and the diagonal \(l\) satisfy the following equation

\[ l^2 = a^2 + b^2 + c^2, \]  

(2.1)

no matter how big the cuboid is. This was strongly supported by the results of practical geometry.

Given this property on can construct a set of Cartesian coordinates, \(\{x_1, x_2, x_3\}\) (the same meaning as \(\{x, y, z\}\)). These coordinates are distances between the origin and the point along the coordinate axes. In Cartesian coordinates, the distance between point A and point B with coordinates \(\{x_{a1}, x_{a2}, x_{a3}\}\) and \(\{x_{b1}, x_{b2}, x_{b3}\}\) respectively is

\[ \Delta l_{ab}^2 = (\Delta x_1)^2 + (\Delta x_2)^2 + (\Delta x_3)^2, \]  

(2.2)

where \(\Delta x^i = x^i_a - x^i_b\).

For infinitesimally close points this becomes

\[ dl^2 = (dx_1)^2 + (dx_2)^2 + (dx_3)^2, \]  

(2.3)

where \(dx^i\) are infinitesimally small differences between Cartesian coordinates of these points. This equation allows us to find distances along curved lines by means of integration.
Figure 2.1: Cartesian coordinates

Other types of coordinates can also be used in Euclidean space. One example is the spherical coordinates, \( \{r, \theta, \phi\} \), defined via

\[
\begin{align*}
    x^1 &= r \sin \theta \cos \phi; \\
    x^2 &= r \sin \theta \sin \phi; \\
    x^3 &= r \cos \theta.
\end{align*}
\]

Here \( r \) is the distance from the origin, \( \theta \in [0, \pi] \) is the polar angle, and \( \phi \in [0, 2\pi) \) is the azimuthal angle.

Figure 2.2: Spherical coordinates

The coordinate lines of these coordinates are not straight lines but curves. Such coordinate systems are called curvilinear. The coordinate lines of spherical coordinates are perpendicular to each other at every point. Such coordinate systems are called orthogonal (there are non-orthogonal curvilinear coordinates). In spherical coordinates, the distance between infinitesimally close points is

\[
dl^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2.
\]

Since, the coefficients of \( d\theta^2 \) and \( d\phi^2 \) vary in space

\[
\Delta l^2_{ab} \neq (\Delta r)^2 + r^2 (\Delta \theta)^2 + r^2 \sin^2 \theta (\Delta \phi)^2
\]

for the distance between points with finite separations \( \Delta r, \Delta \theta, \Delta \phi \). In order to find this distance one has to integrate

\[
l_{ab} = \int_a^b \dl
\]
along the line connecting these points. For example the circumference of a circle of radius $r_0$ is

$$\Delta l = \int dl = 2 \int_0^\pi r_0 d\theta = 2\pi r_0.$$  \hspace{1cm} (2.10)

(Notice that we selected such coordinates that the circle is centered on the origin, $r = r_0$, and it is in a meridional plane, $\phi = \text{const.}$ As the result, along the circle $dl = r_0 d\theta$.)

In the generic case of curvilinear coordinates, the distance between infinitesimally close points is given by the positive-definite quadratic form

$$dl^2 = \sum_{i=1}^{3} \sum_{j=1}^{3} g_{ij} dx^i dx^j,$$  \hspace{1cm} (2.11)

where the coefficients $g_{ij} = g_{ji}$ are normally functions of the coordinates $x^i$. Such quadratic forms are called metric forms. $g_{ij}$ are in fact the components of so-called metric tensor in the coordinate basis of utilized coordinates. It is easy to see that in a Cartesian basis

$$g_{ij} = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$  \hspace{1cm} (2.12)

Not all positive definite metric forms correspond to Euclidean space. If there does not exist a coordinate transformation which reduces a given metric form to that of Eq.2.3 then the space with such metric form is not Euclidean. This mathematical result is utilized in General Relativity.

### 2.1.2 Time

The notion of time reflects our everyday-life observation that all events can be placed in a particular order reflecting their causal connection. In this order, event A appears before event B if A caused B or could cause B. This causal order seems to be completely independent on the individual analyzing these events. This notion also reflects the obvious fact that one event can last longer then another one. In Newtonian physics, time was considered as a kind of fundamental ever going process, presumably periodic, so that one can compare the rate of this process to rates of all other processes. Although the nature of this process remained mysterious it was assumed that all other periodic processes, like the Earth rotation, reflected it. Given the fundamental nature of time it would be natural to assume that this process occurs in ether.

This understanding of time has lead naturally to the absolute meaning of simultaneity. That is one could unambiguously decide whether two events were simultaneous or not. Similarly, physical events could be placed in only one particular order, so that if event A precedes event B according to the observation of some observer, this has to be the same for all other observers, unless a mistake is made. Similarly, any event could be described by only one duration, when the same unit of time is used to quantify it. These are the reasons for the time of Newtonian physics to be called the “absolute time”. There is only one time for everyone.

Both in theoretical and practical terms, a unique temporal order of events could only be established if there existed signals propagating with infinite speed. In this case, when an event occurs in a remote place everyone can become aware of it instantaneously by means of such “super-signals”. Then all events immediately divide into three groups with respect to this event: (i) The events simultaneous with it – they occur at the same instant as the arrival of the super-signal generated by the event; (ii) The events preceding it – they occur before the arrival of this super-signal and could not be caused by it. But they could have caused the original event; (iii) The events following it – they occur after the arrival of this super-signal and can be caused by it. But they cannot cause the original event.

If, however, there are no such super-signals, things become highly complicated as one needs to know not only the distances to the events but also the motion of the observer and how exactly the signals propagated through the space separating the observer and these events. Newtonian physics assumes that such infinite speed signals do exist and they play fundamental role in interaction between physical bodies. This is how in the Newtonian theory of gravity, the gravity force depends only on the current location of the interacting masses.
2.1.3 Galilean relativity

Galileo, who is regarded to be the first true natural scientist, made a simple observation which turned out to have far reaching consequences for modern physics. He noticed that it was difficult to tell whether a ship was anchored or coasting at sea by means of mechanical experiments carried out on board of this ship.

It is easy to determine where a body is moving through air – in the case of motion, it experiences the air resistance, the drag and lift forces. But here we are dealing only with a motion relative to air. What about the motion relative to the absolute space and the interaction with ether? If such an interaction occurred then one could measure the “absolute motion”. Galileo’s observation tells us that this must be at least a rather weak interaction. No other mechanical experiment, made after Galileo, has been able to detect such an interaction. Newtonian mechanics adopts the Galilean relativity via introducing the so-called inertial frames, which coast with constant speed through the absolute space. All laws of Newtonian mechanics have exactly the same form in all these frames. For example, the motion of a physical body which does not interact with other bodies remains unchanged. It moves with constant speed along straight line. This means that one cannot determine the motion through absolute space by mechanical means. Only the relative motion between physical bodies can be determined this way.

2.1.4 The lack of speed limit

Is there any speed limit a physical body can have in Newtonian mechanics? The answer to this question is No. To see this consider a particle of mass $m$ under the action of constant force $f$. According to the second law of Newton its speed then grows linearly,

$$v = v_0 + \frac{f}{m} t,$$

without a limit.

This conclusion also agrees with the Galilean principle of relativity. Indeed, suppose the is a maximum allowed speed, say $v_{\text{max}}$. According to this principle it must be the same for all inertial frames. Now consider a body moving with such a speed to the right of the frame $\tilde{S}$. This frame can also move with speed $v_{\text{max}}$ to the right relative to the frame $S$. Then according to the Galilean velocity addition this body moves relative to the frame $S$ with speed $2v_{\text{max}}$. This contradicts to our assumption that there exist a speed limit, and hence this assumption has to be discarded.

2.1.5 Light

The nature of light was a big mystery in Newtonian physics and a subject of heated debates between scientists. One point of view was that light is made by waves propagating in ether, by analogy with sound which is made by waves in air. The speed of light waves was a subject of great interest to scientists. The most natural expectation for waves in ether is to have infinite speed. Indeed, waves with infinite speed fit nicely the concept of absolute time, and if such waves exist then there is no more natural medium for such waves as the ether of absolute space. However, the light turned out to have finite speed. Dutch astronomer Roemer noticed that the motion of Jupiter’s moons had systematic variation, which could be easily explained only if one assumed that light had finite, though very large, speed. Since then, many other measurements have been made which all agree on the value for the speed of light

$$c \simeq 3 \times 10^{10} \text{cm/s}.$$

The development of mathematical theory of electromagnetism resulted in the notions of electric and magnetic fields, which exist around electrically charged bodies. These fields do not manifest themselves in any other ways but via forces acting on other electrically charged bodies. Attempts do describe the properties of these fields mathematically resulted in Maxwell’s equations, which agreed with experiments most perfectly.

What is the nature of electric and magnetic field? They could just reflect some internal properties of matter, like air, surrounding the electrically charged bodies. Indeed, it was found that the electric and magnetic fields depended on the chemical and physical state of surrounding matter. However, the experiments
clearly indicated that the electromagnetic fields could also happily exist in vacuum (empty space). This fact prompted suggestions that in electromagnetism we are dealing with ether. Analysis of Maxwell equations shows that electric and magnetic fields change via waves propagating with finite speed. In vacuum the speed of these waves is the same in all direction and equal to the known speed of light! When this had been discovered, Maxwell immediately interpreted light as electromagnetic waves or ether waves. Since according to the Galilean transformation the result of any speed measurement depends on the selection of inertial frame, the fact that Maxwell equations yielded a single speed could only mean that they are valid only in one particular frame, the rest frame of ether and absolute space. On the other hand, the fact that the astronomical observations and laboratory experiments did not find any variation of the speed of light as well seemed to indicate that Earth was almost at rest in the absolute space.

However, Newtonian mechanics clearly shows that Earth cannot be exactly at rest in absolute space all the time. Indeed, it orbits the Sun and even if at one point of this orbit the speed of Earth’s absolute motion is exactly zero it must be nonzero at all other points, reaching the maximum value equal to twice the orbital speed at the opposite point of the orbit. This simple argument shows that during one calendar year the speed of light should show variation of the order of the Earth orbital speed and that the speed of light should be different in different directions by at least the orbital speed. Provided the speed measurements are sufficiently accurate we must be able to see these effects. American physicists Mickelson and Morley were first to design experiments of such accuracy (by the year 1887) and to everyone’s amazement and disbelief their results were negative. Within their experimental errors, the speed of light was the same in all directions all the time! Since then, the accuracy of experiments has improved dramatically but the result is still the same, clearly indicating shortcomings of Newtonian physics with its absolute space and time. Moreover, no object has shown speed exceeding the speed of light. In his ground-braking work “On the electrodynamics of moving bodies”, published in 1905, Albert Einstein paved way to new physics with completely new ideas on the nature of physical space and time, the Theory of Relativity, which accommodates these remarkable experimental findings.

2.2 Special relativity

Special relativity can be described as a physical theory of space and time. It is called special because it is not general enough - it does not deal with gravity and it still operates with the inertial frames of reference. The space of each inertial frame is still assumed to be Euclidean.

Special Relativity generalizes the Galilean Relativity by stating that all inertial frames are equivalent as far as any law of Physics is concerned, not only Mechanics. This makes Absolute Space unobservable, and reduces it to a product of human imagination which has no place in reality. Instead of Absolute Space, Special Relativity associates each inertial frame with its own space, with its own distances between events and lengths of physical bodies.
As an example of application of this principle suppose that in a particular frame it is established that there is maximum possible speed that a physical object can have. Then in any other frame there must exist exactly the same maximum possible speed. In fact, Special Relativity yields such speed. It equals to the speed of light in vacuum, \( c \). Our every day experience speaks against this idea. For example the speed of a car moving along the road as measured relative to a stationary police patrol car is different from that measured relative to a moving patrol car. However, in our every day life we are dealing with speeds which are much less than \( c \). This is a singular limit where the true nature of motion is not revealed. On the other hand, the physical experiments designed to deal with very high speeds show that the speed of light is indeed invariant. Moreover, there has been no observations of speeds exceeding the speed of light so far.

The requirement for the speed of light to be the same in all inertial frames immediately leads to a number of spectacular results.

- **Relativity of simultaneity:** Events simultaneous in one frame may not be simultaneous in another. This implies that time interval between events depends on the inertial frame used for time measurements (recall that all frames are equipped with clocks manufactured to the same standards). That there is no unique temporal order of physical events and that each inertial frame has its own time.

- **Time dilation:** Consider a standard clock moving relative to some inertial frame with speed \( v \). Suppose that the time displayed by the clock’s screen increases from \( \tau \) to \( \tau + \Delta \tau \) (\( \tau \) is called the proper time of this clock). The corresponding time interval recorded by the clock grid of this frame, \( \Delta t \), is larger by the factor
  \[
  \gamma = \left(1 - \frac{v^2}{c^2}\right)^{-1/2},
  \]
  which is called the Lorentz factor. That is,
  \[
  \Delta t = \gamma \Delta \tau.
  \]  

- **Length contraction:** Consider a solid rod of length \( l_0 \) as measured using standard measuring tools moving together with the rod, this is called the proper length of the rod. When identical standard measuring tools are used by observers at rest in the inertial frame relative to which the rod is moving with speed \( v \) the result depend on the rod orientation. When the rod is positioned perpendicular to the direction of motion its length is the same,
  \[
  l = l_0.
  \]  
  When the rod is aligned with the direction of motion its length is smaller,
  \[
  l = \frac{l_0}{\gamma}.
  \]
  For intermediate angles \( l_0/\gamma < l < l_0 \).

- **Invariance of the spacetime interval.** Consider any two events and measure their separation in space and time, \( \Delta l \) and \( \Delta t \). The effects of time dilation and length contraction show that these quantities are different in different inertial frames. However, their combination
  \[
  \Delta s^2 = -c^2 \Delta t^2 + \Delta l^2.
  \]
  turns out to be the same! Surely, this remarkable result must be significant! In fact, it tells us that space and time can be united into a single 4-dimensional metric space with \( \Delta s \) being the generalized distance between points (events) in this space. This space is called Minkowskian spacetime, after the mathematician who introduced it.

Since \( \Delta l^2 = \Delta x^2 + \Delta y^2 + \Delta z^2 \) we can write
  \[
  \Delta s^2 = -c^2 \Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2.
  \]
  For infinitesimally closed points this becomes
  \[
  ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2.
  \]
This quadratic form is called the metric form of Minkowskian spacetime and $ds^2$ is called the spacetime interval. If we compare this metric form with the metric form of Euclidean space,

$$dl^2 = dx^2 + dy^2 + dz^2,$$

then we would immediately notice a number of important differences. First, there are fewer terms in the Euclidean metric. This is simply a reflection of the fact that the Euclidean space has three dimensions whereas the spacetime has four. In fact, mathematicians happily operate with higher dimensional spaces, including the four dimensional Euclidean space with the metric form

$$dl^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 + (dx^4)^2,$$

where $x^i$ are its four Cartesian coordinates. Second, there is a coefficient $c^2$ in front of $dt^2$ in Eq. 2.18. This difference, however, is also minor. Indeed, one can introduce new coordinate $x^0 = ct$ so that $c^2 dt^2$ becomes $(dx^0)^2$. Finally, there is the minus sign in front of $dt^2$ and this is what make Minkowskian spacetime qualitatively different from Euclidean one (Another name for Minkowskian spacetime is pseudo-Euclidean space.). In Euclidean space $dl^2$ is always non-negative, being zero only in the case of no separation between the end points. In Minkowskian space $ds^2$ can both positive and negative! This reflect the fact that the separation between events in the Universe can be of three distinctive types.

- First, there is space-like separation, where

$$ds^2 > 0. \quad (2.19)$$

For such events one can always find an inertial frame where these events are simultaneous, $dt = 0$, and hence separated only in space. In such a frame $ds^2$ reduces to $dl^2$:

$$ds^2 = dl^2 = dx^2 + dy^2 + dz^2. \quad (2.20)$$

Obviously, such events cannot be causally connected (Causal connection for such events would require a signal with infinite speed.).

- Second, there time-like separation, where

$$ds^2 < 0. \quad (2.21)$$
For such events one can always find an inertial frame where these events occur at the same place, $dl = 0$, and hence they are separated only in time. This time is the proper time of the standard clock of this frame, $d\tau$. Thus, we have

$$ds^2 = -c^2d\tau^2 \quad \text{or} \quad d\tau = \sqrt{-ds^2/c}. \quad (2.22)$$

Obviously, such events can be causally connected.

- Finally, for some events one can have

$$ds^2 = 0 \quad (2.23)$$

even if these events are different. This type of separation is called null. From Eq. 2.16 one can see that in this case

$$c^2dt^2 = dl^2 \quad \text{or} \quad dl/dt = c. \quad (2.24)$$

Thus, these two events can be events in the life of a photon, a particle moving with speed $c$.

The difference between the Newtonian and relativistic visions of space and time is nicely illustrated with the help of the Minkowski diagrams. The idea is to represent instantaneous localized events as points of a two dimensional graph. Surely, some information is lost because we really need four dimensions to fully represent spacetime events. So what is actually shown is time and one of the spatial coordinates (usually $x$) as measured in some inertial frame.

The left panel of Fig. 2.5 is a Newtonian version of Minkowski diagram. Along the horizontal axis we show the $x$ coordinate and along the vertical axis the time as measured in some inertial frame. Any two events, like A and B in this figure, that belong to a line parallel to the time axis, are simultaneous in this frame. The whole such line represents all real and potential events occurring at the same time and shows the spatial separation between them. One can say that it represents space at this particular time. The time axis represents its time. Now consider another inertial frame which has the same orientation of axes and moves along the $x$-axis of the first frame with constant speed $v$. For simplicity we also assume that the origins of both frames coincide at time $t' = t = 0$, where $t'$ is the time of the second frame. The line $t'$ shows the motion of the origin of the second frame relative to the first one – it can be considered as the time axis of the second frame. The events A and B occur at the same time $t' = t$ in this frame and, thus, its space at time $t'$ is the same as the space of the first frame at time $t = t'$. It is still represented by the line parallel to the $x$ axis as well.

![Diagram of Newtonian and Minkowskian visions of space and time](image)

Figure 2.5: Newtonian (left panel) and Minkowskian (right panel) visions of space and time. The angle $\alpha$ is given by $\tan(\alpha) = v/c$.

The right panel of Fig. 2.5 is a proper Minkowski diagram of spacetime in special relativity for the same combination of two inertial frames. Now events A and B are not simultaneous in the second frame. The set all events (real and potential) simultaneous in this frame cannot be represented by a line parallel to the $x$ axis. They are parallel to the $x'$ axis, which represents all events at $t' = 0$. The angle between the $x$ and $x'$ axes equals to the angle between the $ct$ and $ct'$ axes (see Figure 2.5). If we denote this angle as $\alpha$, then $\tan \alpha = v/c$. 

Thus, the space of the second frame, is different from the space of the first frame. The spacetime is the same, but for each inertial frame it splits into space and time specific to this frame.

The revolutionized notions of space and time in Special Relativity have dramatic implications for physics in general and particle dynamics in particular. One of the most important results is the equivalence of mass and energy,

$$E = mc^2,$$  

where $E$ is the total energy of the particle and $m$ is its inertial mass. One hand, this famous equation tells us that there is energy, $E_0$, associated with the rest mass of the particle, $m_0$, that is the inertial mass of this particle in the frame where it is at rest. In principle, this energy can be released and converted into other kinds of energy, for example in nuclear reactions. On the other hand, it tells us that the inertial mass, $m$, of moving particle is higher than its rest mass due to the contribution from its kinetic energy. In fact, one can show that

$$m^2 = m_0^2 + \frac{p^2}{c^2},$$  

where $p$ is the particle momentum, $p = mv$, or

$$m = m_0 \gamma.$$  

This has an important implication for the inertial mass-energy density of hot gas (fluid), as the particles of hot gas move around and hence have non-vanishing kinetic energy. Consider a frame where the macroscopic speed of gas is zero, the so-called comoving frame. In Newtonian physics, the inertial mass density of gas in this frame is

$$\rho = \rho_0 = m_0 n,$$

where $m_0$ is the rest mass of a single gas particle and $n$ is the number density of gas particles (the number of particles per unit volume). In relativity, we also need to take into account the contribution due to the thermal motion of these particles,

$$e_t = \frac{1}{\Gamma - 1} P,$$  

where $P$ is the gas thermal pressure and $1 < \Gamma < 2$ is the so-called ratio of specific heats. This leads to the result

$$\rho = \rho_0 + \frac{1}{\Gamma - 1} (P/c^2).$$  

The quantity given by Eq.2.29 is called the proper or comoving mass density\(^1\). For a non-relativistic gas, where the speed of thermal motion is low compared to the speed of light, $v_t \ll c$ and the corresponding Lorentz factor thermal motion $\gamma \simeq 1$, one has $e_t \ll \rho_0 c^2$, and $\rho = \rho_0$. For an ultra-relativistically hot gas, that is gas with the Lorentz factor of thermal motion, $\gamma \gg 1$, the second term on the right hand side of Eq.2.26 and hence the second term on the right hand side in Eq.2.29 dominate. Moreover, in this regime $\Gamma = 4/3$. Thus, one may ignore the contribution due to the rest mass and assume that

$$\rho = 3P/c^2.$$  

Photons, or the particles of light, move exactly with the speed of light. Then Eqs.2.25 and 2.27 imply that the photon’s rest mass density $m_0 = 0$ (otherwise its energy would be infinite). Thus, Eq.2.30, is exact for radiation, which can be considered as gas of photons.

### 2.3 General Relativity

Mass $m$ that appears in the second law of Newtonian mechanics (Eq.8.2) is the inertial mass. The same mass appears in the right hand side of the Newton’s law of gravity, Eq.8.1, where it determines the strength of gravitational force. There is no mathematical reason for this. In fact, by analogy with electromagnetism, the law of gravity could involve not the inertial mass but a different quantity, that could be called gravitational

\(^1\)In frames where the macroscopic speed of gas is not zero the inertial mass density is different because of the length contraction effect and the contribution to mass due to the kinetic energy of the macroscopic motion.
charge or gravitational mass, $m_g$. Some particles could have it, some not. Two particles with the same inertial mass could have different gravitational mass. However, in Nature

$$m_g = m. \quad (2.31)$$

Thus, Eq.2.31 is a physical law which is incorporated in the Newton’s law of gravity. It is called the Principle of Equivalence of inertial and gravitational mass. Because of this law “particles” of very different masses have exactly the same acceleration in a given gravitational field (see Eq.8.3) and provided they have the same initial velocity they exhibit identical motion. This suggests that particle motion in gravitational field has nothing to do with the properties of particles but is rather determined by the properties of the spacetime itself. In fact, Albert Einstein concluded that gravity makes itself felt via warping the spacetime.

![Figure 2.6: Left panel: Albert Einstein. Right panel: Naive illustration of the effect of the spacetime curvature on the propagation of light – the photon is compared with a ball running over a warped elastic sheet in laboratory under the action of downward gravity. The big orange ball in middle creates the curvature of the sheet.](image)

Indeed, the Minkowskian spacetime of Special Relativity is flat. By definition, a metric space is called flat if there exist such global coordinates, $x^\nu$, that its metric form is the sum of $\pm (dx^\nu)^2$ in every point of the space. Otherwise, it is called curved or warped. For example, a Euclidean plane is a flat two dimensional space. Indeed, in Cartesian coordinates its metric form is

$$dl^2 = dx^2 + dy^2 \quad (2.32)$$

everywhere. A sphere in Euclidean space is not flat (in fact, it is not even a space but a manifold). It is impossible to introduce such coordinates on the sphere such that the distance between any two of its infinitesimally close points is given by the above metric form. For example, in spherical coordinates, the metric form of a sphere of radius $r$ is

$$dl^2 = r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2,$$

which is different from that in Eq.2.32. However, for any point of the sphere one can introduce such coordinates, $x$ and $y$, that at this particular point the metric form is given by Eq.2.32. For this reason, the sphere is described as a locally Euclidean surface (manifold). In fact, one can do more. One can ensure that the metric form is the same as in Eq.2.32 along a curve on the sphere. But no more that this.

Similarly, the warped (or curved) spacetime of General Relativity is locally Minkowskian – for any point in spacetime one can introduce such coordinates that the metric form is Minkowskian (Eq.2.18) at this point. In fact, one can also ensure that the metric form is Minkowskian along a particular geodesic (generalized straight line) of spacetime. From the physical viewpoint, this corresponds to a locally inertial frame, which
you can imagine as a small free flying laboratory with its small Cartesian grid (confined within the laboratory’s walls) and standard clocks. Within the walls of this laboratory the spacetime curvature has a very little effect and all physical phenomena are not effected by gravity. However, it is generally impossible to extend this grid well outside of the laboratory without destroying its nice properties.

In most problems the metric form is not known before hand. So often it is written in the most general form

$$ds^2 = \sum_{\nu=0}^{3} \sum_{\mu=0}^{3} g_{\nu\mu} dx^{\nu} dx^{\mu}. \quad (2.33)$$

Here $x^{\nu}$ are some coordinates of spacetime (traditionally the indexes vary not from 1 to 4 but from 0 to 3), and the coefficients $g_{\nu\mu}$ are functions of these coordinates. In fact, these coefficients are the components of the so-called metric tensor

However, this general form hides the difference between spacetime, with its time-like and space-like directions, and other kinds of four dimensional metric manifolds. More transparent is the so-called $3+1$ representation

$$ds^2 = -\alpha c^2 dt^2 + \sum_{i=1}^{3} \beta_i dx^i dt + \sum_{i=1}^{3} \sum_{j=1}^{3} \gamma_{ij} dx^i dx^j. \quad (2.34)$$

Here $t$ is considered as the global time coordinate. The hyper-surface $t = t_0$ is considered as the global space at time $t_0$. In general it is not Euclidean but warped (see the right panel of Fig.2.6). Its metric form is

$$dt^2 = \sum_{i=1}^{3} \sum_{j=1}^{3} \gamma_{ij} dx^i dx^j, \quad (2.35)$$

with $x^i$ considered as coordinates of space ($i=1,2,3$). The coefficient $\alpha > 0$ is called the lapse function. In order to understand its meaning, consider a standard clock attached to the spatial grid and consider two events in the life of this clock separated by the proper time $d\tau$ (the time measured by this clock). For these events $ds^2 = -c^2 d\tau^2$ (see Eq.2.22). On the other hand $dx^i = 0$ and hence $ds^2 = -\alpha c^2 dt^2$. Thus, we have

$$-c^2 d\tau^2 = -\alpha c^2 dt^2 \quad \text{or} \quad \alpha = \left(\frac{d\tau}{dt}\right)^2. \quad (2.36)$$

Thus, the lapse function gives the rate of standard clocks attached to the spatial grid compared to the rate of global time.

At any point of the global space one can have many different local inertial frames (or fiducial inertial observers), each with their own local Euclidean Space. One of these frames can be considered as being at rest in the global space. Its local Euclidean space is essentially a part of the global space, meaning that events with $dt = 0$ appear simultaneous in this local inertial frame. Vector $\beta_i$, called the shift vector, gives the velocity $dx^i/dt$ of such inertial frame relative to the global spacial grid. In general, it is impossible to have $\beta = 0$ everywhere.

The key equation of General Relativity is the celebrated Einstein’s equation

$$R_{\nu\mu} - \frac{1}{2} R g_{\nu\mu} = \frac{8\pi G}{c^4} T_{\nu\mu}, \quad (2.37)$$

which Einstein discovered in 1915. Here $R_{\nu\mu}$ is the Ricci tensor, $R$ is the scalar curvature – they describe properties of spacetime. In Minkowskian spacetime $R_{\nu\mu} = 0$ and $R = 0$ everywhere. $T_{\nu\mu}$ is called the stress-energy-momentum tensor. It describes the distribution of energy, momentum, and stresses associated with matter, radiation, and all sorts of force fields.

The simple appearance of Einstein equation is deceptive. In fact, this is one of the most difficult equations of Mathematical Physics, which can be solved analytically only in very limited cases of highly symmetric problems. Only during the last decade mathematicians figured out how to solve this equation numerically. So here I only briefly discuss its nature.

---

2 In mathematics, the Euclidean space of this local inertial frame is called “tangent” to the hyper-surface $t = t_0$. 
Since \( \nu, \mu = 0 \ldots 3 \) we have sixteen equations in Eq.2.37. The Ricci tensor and the curvature scalar are functions of \( g_{\nu\mu} \), its first and second order derivatives with respect to all four coordinates. Thus, we are dealing not with one equation but with a system of sixteen simultaneous second order partial differential equations. The unknown functions of coordinates in these equations are \( g_{\nu\mu} \) and \( T_{\nu\mu} \) – the Einstein equation not only describes how matter warps spacetime but also how matter evolves in this warped spacetime.

If correct, the Einstein equation should also describe the Universe, which is filled with gravitationally interacting matter. In 1915 the Universe and the Milky Way (our Galaxy) were considered as the same thing, and the Milky Way appeared to be very much static. Einstein analyses models of static Universe and concluded that such Universe cannot be infinite. Instead, it must be finite but without boundaries - the space must be wrapped onto itself like a sphere in Euclidean geometry. Later however, he run into difficulty as his equation 2.37 did not allow such a static solution. So he modified this equation by adding the so called Cosmological term:

\[
R_{\nu\mu} - \frac{1}{2} R g_{\nu\mu} = 8 \pi G \frac{c^4}{c^4} T_{\nu\mu} - \frac{\Lambda}{c^2} g_{\nu\mu},
\]

(2.38)

where \( \Lambda \) is called the Cosmological constant. (In Newtonian Physics this is equivalent to introduction of a repulsive force to balance gravity.) After the Hubble’s discovery Einstein abandoned his work on the Cosmological term, and called it his “biggest blunder”. However, the modern data indicate that the Cosmological term is in fact needed to explain the observations.
Chapter 3

Friedmann’s equations

3.1 Metric of the Universe

The simplest from the mathematical view point model of the Universe is where it is uniform, that is where its geometry is the same at every point. From the observational prospective it is not obvious that our Universe is uniform. Indeed, the very existence of the boundary of visible Universe (the CMB “screen”) seem to indicate that it is not uniform. However, as we shall see later this observations are easily explained in the relativistic models of the Universe where it is assumed uniform. These models predict the observed Universe to appear isotropic (or almost isotropic) and this prediction agrees with the observations very well indeed.

What is the spacetime metric of such a Universe? As we have seen, metric of any spacetime can be written as

\[ ds^2 = -c^2 dt^2 + \sum_{i=1}^{3} \beta_i dx_i dt + dl^2, \]

where \( dl \) is the line element of space (see Sec.2.3). The uniformity of universe means that one can introduce such spacetime coordinates, \( t \) and \( x_i \), that neither \( \alpha \) nor \( \beta \) depend on the spatial coordinates, but only on time. Such a coordinate system is obviously preferable. In fact there are many such coordinate systems. Among them are those where \( \alpha = 1 \). Indeed, if \( \alpha \neq 1 \) we just redefine the time coordinate so that \( dt' = \sqrt{\alpha} dt \). Thus, one can prescribe the metric as

\[ ds^2 = -c^2 dt'^2 + \sum_{i=1}^{3} \beta_i' dx_i' dt' + dl^2. \]

There are still many different coordinate systems which allow such a metric form. The isotropy of universe tells us that among them there are systems where the metric form is isotropic, that is for \( dx_i = dx_i' \) and \( dx_i' = -dx_i \) we should have the same \( dl^2 \) and \( ds^2 \). This requires \( \beta_i = 0 \). The result is

\[ ds^2 = -c^2 dt^2 + dl^2. \]

Let us analyze this. Consider a standard clock with fixed spatial coordinates \( x_i' \). If \( \tau \) is the proper time of such a clock then the spacetime interval between any two events in its life is \( ds^2 = -c^2 d\tau^2 \) (see Eq.2.22). On the other hand Eq.3.16 gives \( ds^2 = -c^2 dt^2 \) for the same two events. Thus, \( dt = d\tau \), that the cosmological time \( t \) in Eq.3.16 “ticks” at the same rate as the time of standard clocks at rest relative to the spatial grid. Since the shift vector \( \beta_i \) is zero the spatial grid is at rest in space. In particular, if the universe is expanding then so does the grid, and at exactly the same rate. Just imagine inflating a spherical balloon with a coordinate grid printed on it and you will get the meaning of this. Our interpretation of the Hubble’s discovery will be that the distant galaxies are at rest in space, and hence have fixed spatial coordinates, but the distance between them is increasing.

Finally, we need to figure out the metric form of the space (the hyper-surface of the spacetime determined by the equation \( t = \text{const} \)). When Einstein was constructing his first model of the Universe he reasoned along the following lines. First, because of matter the spacetime and the space have to be curved. Uniformity of the Universe means that the curvature of space must be the same at every point. In order to guess what such
spaces could be, consider surfaces of 3-dimensional Euclidean space. The only two types of surfaces with uniform curvature are planes and spheres. Planes have to be excluded as they have zero curvature (they are flat). So the only relevant case is a sphere. This suggests that the geometry of the Universe must be the same as that of a three-dimensional hyper-sphere of four-dimensional Euclidean space!

What exactly is this hyper-sphere? If \( x_i \) where \( i = 1 \ldots 4 \) are the four Cartesian coordinates of the four-dimensional Euclidean space, centered on center of the hyper-sphere of radius \( R \), then its equation is

\[
(x^1)^2 + (x^2)^2 + (x^3)^2 + (x^4)^2 = R^2.
\]

Thus, the hyper-sphere is the set of all point located at the distance \( R \) from the origin. What is the metric form of such a hyper-surface? We can find it following the same steps we do in order to find the metric form of a normal sphere in three-dimensional Euclidean space.

![Figure 3.1: The generalized spherical coordinates \( \eta, \phi, \theta, \phi \) of 4-dimensional Euclidean space explained. Left panel: the radial coordinate \( \eta \), and “polar” angle \( \psi \). The blue plane represents the 3-dimensional hyper-plane normal to the \( x^4 \) axis. The other two coordinates of the point A are the same as the coordinates of projected point A’ in this hyper-plane. Right panel: the spherical coordinates of point A’ in the hyper-plane.](image)

First, we introduce the generalized spherical coordinates in the four-dimensional Euclidean space. Take a point in the space. Its distance from the origin denote as \( \eta \). This is the first generalized spherical coordinate. Denote the angle between the radius vector of this point and the \( x^4 \) axis as \( \psi \). This is the second coordinate. Project this point on to the hyper-plane \( x^4 = 0 \). This hyper-plane is a 3-dimensional Euclidean space, where we can introduce the usual spherical coordinates. The distance between the origin and the projected point in this space is \( \eta' = \eta \sin \psi \). Denote the angle between the radius vector of the projected point and the \( x^3 \) axis as \( \theta \). This is the third generalized spherical coordinate. Project this projected point onto the plane \( x^3 = 0 \). The distance of this second projected point from the origin is \( \eta'' = \eta' \sin \theta = \eta \sin \psi \sin \theta \). Denote the angle between the radius vector of the second projected point and the \( x^1 \) axis as \( \phi \). This is the fourth generalized spherical coordinate. By construction, \( \psi, \theta \in [0, \pi], \) and \( \phi \in [0, 2\pi) \). The transformation law between the generalized spherical and Cartesian coordinates is

\[
\begin{align*}
x^4 &= \eta \cos \psi, \\
x^3 &= \eta' \cos \theta = \eta \sin \psi \cos \theta, \\
x^2 &= \eta'' \sin \phi = \eta \sin \psi \sin \theta \sin \phi, \\
x^1 &= \eta''' \cos \phi = \eta \sin \psi \sin \theta \cos \phi.
\end{align*}
\] (3.4)

Next we find the metric form of the 4-dimensional Euclidean space in the generalized spherical coordinates. In fact, these coordinates are orthogonal, and hence the infinitesimally small parallelelepiped whose sides are made of the coordinate lines is rectangular. Its diagonal has the length given by the equation

\[
dl^2 = dl^2_\eta + dl^2_\psi + dl^2_\theta + dl^2_\phi,
\]

where \( dl_i \) are the lengths of the sides. It is easy to see that

\[
dl_\eta = d \eta,
\]
This looks familiar. In fact, for

In fact, for

Clearly,  \( \varepsilon \) can be written as

like surface in Euclidean space (see Figure 3.2).

\[ \chi \]

\( \psi \) where

where

\[ 0, \infty \), and Eq. 3.11 as

\[ d \varphi \]

\[ d \theta \]

\[ d \psi \]

Thus, the metric form is

\[ dl^2 = d \eta^2 + \eta^2 d \psi^2 + \eta^2 \sin^2 \psi (d \theta^2 + \sin^2 \theta d \phi^2). \]  

(3.5)

Now we can find the metric form of the hyper-sphere of radius \( R \). Since on the hyper-surface \( \eta = R \) and

\[ d \eta = 0, \]

we have

\[ dl^2 = R^2 (d \psi^2 + \sin^2 \psi (d \theta^2 + \sin^2 \theta d \phi^2)) . \]  

(3.6)

Clearly, \( R \psi \) is the distance from the north pole of the hyper-sphere along the \( \psi \) coordinate line. Denote this as \( r \) ( \( r \) varies from 0 to \( \pi R \)). Then the metric form (3.6) reads

\[ dl^2 = dr^2 + R^2 \sin^2 \left( \frac{r}{R} \right) (d \theta^2 + \sin^2 \theta d \phi^2). \]  

(3.7)

This looks familiar. In fact, for \( r \ll R \), we have \( \sin (r/R) \to r/R \) and the metric form (3.7) becomes

\[ dl^2 = dr^2 + r^2 (d \theta^2 + \sin^2 \theta d \phi^2), \]  

(3.8)

which is the metric form of 3-dimensional Euclidean space in spherical coordinates. Thus, our \( r, \theta \) and \( \phi \) are analogues of the spherical coordinates It is illuminating to compare (3.7) with the metric form of a normal sphere in the “polar” coordinates, which we derived earlier in class,

\[ dl^2 = d\rho^2 + R^2 \sin^2 \left( \frac{\rho}{R} \right) d\phi^2. \]  

(3.9)

For \( \rho \ll R \) this reduces to the metric form of 2-dimensional Euclidean space in polar coordinates. The similarity is striking, but expected.

The fact that we considered a hyper-sphere of a 4-dimensional Euclidean space does not necessarily mean that our physical space is embedded into some higher dimensional physical space. We simply used this approach to derive the metric of a 3-dimensional space (or rather a manifold) with constant curvature. No additional physical spatial dimensions are really needed to be introduced here. (Though they may be introduced, and they are introduced in modern physics, but for other reasons.)

Are there any other types of uniformly curved 3-dimensional manifolds? Yes. First, the trivial case of 3-dimensional Euclidean space, which has zero curvature everywhere. Its metric form is given by Eq. 3.8 and can also be written as

\[ dl^2 = R^2 (d \psi^2 + \psi^2 (d \theta^2 + \sin^2 \theta d \phi^2)) , \]  

(3.10)

where \( \psi = r/R \) varies from 0 to \( +\infty \). Finally, there is the so-called hyperbolic space with the metric

\[ dl^2 = R^2 (d \psi^2 + \sinh^2 \psi (d \theta^2 + \sin^2 \theta d \phi^2)) , \]  

(3.11)

where \( \Psi \) also varies from 0 to \( +\infty \). This also corresponds to a hyper-surface but now in the Minkowskian spacetime (also known as pseudo-Euclidean space). Its geometric properties are similar to those of a saddle-like surface in Euclidean space (see Figure 3.2).

In fact, all these three metric forms, Eqs. 3.6, 3.10, and 3.11, can be written in a very similar form. Indeed, Eq. 3.6 can be written as

\[ dl^2 = R^2 \left( \frac{d \chi^2}{1 - \chi^2} + \chi^2 (d \theta^2 + \sin^2 \theta d \phi^2) \right) , \]  

(3.12)

where \( \chi = \sin \psi \in [0, 1] \), Eq. 3.10 as

\[ dl^2 = R^2 \left( d \chi^2 + \chi^2 (d \theta^2 + \sin^2 \theta d \phi^2) \right) , \]  

(3.13)

where \( \chi = \psi \in [0, +\infty) \), and Eq. 3.11 as

\[ dl^2 = R^2 \left( \frac{d \chi^2}{1 + \chi^2} + \chi^2 (d \theta^2 + \sin^2 \theta d \phi^2) \right) , \]  

(3.14)
where $\chi = \sinh \psi \in [0, +\infty)$. Now one can see that all three metric form have the form

$$dl^2 = R^2 \left( \frac{d\chi^2}{1-k\chi^2} + \chi^2 (d\theta^2 + \sin^2 \theta d\phi^2) \right), \quad (3.15)$$

where $k$ is either 0, or $\pm 1$. A universe is called closed if $k = 1$, open if $k = -1$, and flat if $k = 0$. Open and flat universes have infinite volume, whereas close ones have finite volume.

Combining Eqs.3.3 and 3.15 the spacetime metric form of the Universe can be written as

$$ds^2 = -c^2 dt^2 + R^2 \left[ \frac{d\chi^2}{1-k\chi^2} + \chi^2 (d\theta^2 + \sin^2 \theta d\phi^2) \right]. \quad (3.16)$$

Metric Eq.3.16 is known as the Robertson-Walker metric.

### 3.2 Derivation of Friedmann’s equations in General Relativity

In the derivation we assume that the distribution of both matter and radiation can be described as continuous (on scales above the size of voids). That is both are treated as a uniform gas (compressible fluid) with the mass-energy density $\rho$ and pressure $P$.

Out of 16 coefficients $g_{\nu\mu}$ only four are non-vanishing in the Robertson-Walker metric: $g_{tt} = -c^2$, $g_{\chi\chi} = R^2(t)/(1-k\chi^2)$, $g_{\theta\theta} = R^2(t)\chi^2$ and $g_{\phi\phi} = R^2(t)\chi^2 \sin^2 \theta$ and they include only one unknown function, the scaling factor $R(t)$. After substituting these expressions into the Einstein’s equations (2.37) with the stress-energy-momentum tensor of ideal fluid one finds that out of these 16 equations only two are independent. They are

$$\frac{\ddot{R}}{R} \left( \frac{\dot{R}}{R} \right)^2 + \frac{k c^2}{R^2} = \frac{8\pi G}{3} \rho. \quad (3.17)$$

and

$$2 \frac{\dddot{R}}{R} + \left( \frac{\dot{R}}{R} \right)^2 + \frac{k c^2}{R^2} = -8\pi G \frac{P}{c^2}, \quad (3.18)$$

where $\rho$ is the mass-energy density and $P$ is the pressure. Equation 3.17 is known as the Friedmann equation, after the Russian mathematician Alexander Friedmann who first derived it back in 1922.

Subtracting Eq.3.17 from Eq.3.18 one obtains the so-called acceleration equation

$$\frac{\ddot{R}}{R} = -\frac{4\pi G}{3} \left( \rho + \frac{3P}{c^2} \right). \quad (3.19)$$

Another useful equation, known as the fluid equation can be derived from Eqs.3.17 and 3.19 as follows. First, we differentiate Eq.3.17 to obtain

$$\frac{\dot{R}}{R} \left( \frac{\ddot{R} - \dot{R}^2}{R^2} \right) - \frac{2k c^2}{R^2} \frac{\dot{R}}{R} = \frac{8\pi G}{3} \dot{\rho}. \quad (3.19)$$
Then we substitute in the result the expressions for $\dot{R}$ from the acceleration equation and $kc^2/R^2$ from the Friedmann equation. This gives us the fluid equation

$$\dot{\rho} = -3 \left( \frac{\dot{R}}{R} \right) \left( \rho + \frac{p}{c^2} \right).$$

(3.20)

When the Einstein equation with the cosmological term (Eq.2.38) is used instead of the original Eq.2.37, the Friedmann and the acceleration equations become

$$\left( \frac{\dot{R}}{R} \right)^2 + \frac{k c^2}{R^2} = \frac{8\pi}{3} G \rho + \frac{\Lambda}{3},$$

(3.21)

$$\frac{\ddot{R}}{R} = -\frac{4\pi G}{3} \left( \rho + \frac{3p}{c^2} \right) + \frac{\Lambda}{3},$$

(3.22)

respectively, whereas the fluid remains unchanged.
Chapter 4

Basic Cosmological Models

4.1 Einstein’s static Universe

Back in 1917 Albert Einstein investigated whether his General Relativity can explain the observed stationary Universe, the contemporary astronomical data suggested that the Universe and the Milky Way were the same thing. For a stationary Universe the scaling factor $R$ must be constant and hence all its time derivatives must vanish. However, the acceleration equation (Eq.3.19) shows that $\ddot{R} < 0$, as $\rho, P > 0$. This was the reason that had led Einstein to introduce his “cosmological term”. With this term, the Friedmann equations read

$$\left(\frac{\dot{R}}{R}\right)^2 + \frac{kc^2}{R^2} = \frac{8\pi}{3} G \rho + \frac{\Lambda}{3}, \tag{4.1}$$

$$\frac{\ddot{R}}{R} = -\frac{4\pi G}{3} \left(\rho + \frac{3P}{c^2}\right) + \frac{\Lambda}{3}, \tag{4.2}$$

$$\dot{\rho} = -3 \left(\frac{\dot{R}}{R}\right) \rho + \frac{P}{c^2}. \tag{4.3}$$

Since the Milky is made of nonrelativistic matter (e.g. stars, planets, interstellar gas) $P/c^2 \ll \rho$ and can be ignored. Then Eq.4.3 becomes

$$\dot{\rho} = -3 \left(\frac{\dot{R}}{R}\right) \rho \quad \text{or} \quad \frac{d\rho}{\rho} = -3 \frac{dR}{R}.$$  

This separable equation is easily integrated:

$$\frac{4\pi}{3} \rho R^3 = M = \text{const.} \tag{4.4}$$

This result simply says that the mass of a sphere expanding with the Universe is constant. Given this we can rewrite Eqs.4.1 and 4.2 as

$$\left(\frac{\dot{R}}{R}\right)^2 + \frac{kc^2}{R^2} = \frac{2GM}{R^3} + \frac{\Lambda}{3}, \tag{4.5}$$

$$\frac{\ddot{R}}{R} = -\frac{GM}{R^3} + \frac{\Lambda}{3}. \tag{4.6}$$

Let us see if these equations allow the stationary solution $R(t) = R_0$.

When we substitute this into Eq.4.6 we find

$$\frac{GM}{R_0^3} = \frac{\Lambda}{3}. \tag{4.7}$$
Thus, the stationary solution is only possible for $\Lambda > 0$. Similarly, from Eq. 4.5 we find that

$$\frac{k c^2}{R_0^2} = \frac{2GM}{R_0^2}G + \frac{\Lambda}{3}$$

and using Eq. 4.7 obtain

$$\frac{k c^2}{R_0^2} = \Lambda.$$  
(4.8)

Thus, the stationary solution requires $k = 1$, which means that the Universe must be closed (see Sec. 3.2)!

However, this solution is unstable, as was first shown by Friedmann in 1922. To see this will use the technique of linear stability analysis. First, consider a small perturbation around this stationary solution $R(t) = R_0(1 + \varepsilon(t))$ where $\varepsilon \ll 1$.

Then Eq. 4.6 becomes the differential equation governing the evolution of $\varepsilon(t)$.

$$\frac{d^2}{dt^2} R_0(1 + \varepsilon) = -\frac{GM}{R_0^2}(1 + \varepsilon)^{-2} + \frac{\Lambda}{3} R_0(1 + \varepsilon).$$

Next expand $(1 + \varepsilon)^{-2}$ in Taylor series keeping only the first two terms

$$(1 + \varepsilon)^{-2} = 1 - 2\varepsilon.$$

This gives us

$$R_0\ddot{\varepsilon} = -\frac{GM}{R_0^2} + \frac{\Lambda}{3} R_0 + \frac{2GM}{R_0^2} \varepsilon + \frac{\Lambda}{3} R_0 \varepsilon.$$

The first two terms in the right hand side of this equation give zero (see Eq. 4.7) and so we end up with the evolution equation for $\varepsilon$

$$\ddot{\varepsilon} = a^2 \varepsilon,$$  
(4.9)

where

$$a^2 = \frac{2GM}{R_0^2} + \frac{\Lambda}{3} > 0.$$

The general solution of Eq. 4.9 is

$$\varepsilon = C_1 e^{at} + C_2 e^{-at},$$

where $C_1$ and $C_2$ are constants and without any loss of generality we may assume $a > 0$. From this solution we can see that at large time the perturbation grows exponentially, which shows the instability of the Einstein’s stationary solution.

### 4.2 Key parameters of the Universe.

#### 4.2.1 Expansion and deceleration parameters

For non-stationary Universe, it’s key kinematic parameters are the expansion (or contraction) and deceleration (or acceleration) rates. Using the Taylor expansion of $R(t)$ for $|t - t_0| \ll t_0$, where $t_0$ is the current time, one has

$$R(t) \simeq R(t_0) + \dot{R}_0(t - t_0) + \frac{1}{2} \ddot{R}_0(t - t_0)^2 = R_0 \left[ 1 + H_0(t - t_0) - \frac{q_0}{2} H_0^2 (t - t_0)^2 \right],$$  
(4.10)

where

$$H_0 = \frac{\dot{R}_0}{R_0},$$  
(4.11)

and

$$q_0 = -\frac{\ddot{R}_0}{R_0 H_0^2}.$$  
(4.12)
Obviously, $H_0 > 0$ corresponds to an expanding Universe and $H_0 < 0$ to contracting one. $q_0 > 0$ corresponds to decelerating expansion of the Universe and $q_0 < 0$ to its accelerating expansion. $q_0$ is called the deceleration parameter. For Einstein’s static Universe $H_0 = 0$ and $q_0 = 0$.

$q_0$ is a dimensionless parameter. The physical dimension of $H_0$ is $1/T$, the same as the dimension of the Hubble constant. In order to show this, consider two points in space participating in the expansion of the Universe. At one point we have a source of light and at the other an observer (us). Let us use spherical coordinates centered on the observer, so that the comoving coordinate of the observer is $\chi = 0$ and the comoving coordinate of the source is $\chi_c$. Let us analyze the propagation of a photon emitted by the source at time $t_e$ and received at time $t_0$. As the photon propagates radially towards the observer, along its trajectory in space, and hence along its worldline in spacetime we have $\theta = \text{const}$ and $\phi = \text{const}$. Moreover, along the worldlines of photons

$$ds^2 = 0$$

(see Sec.2.2). Then from the Robertson-Walker metric of the Universe, Eq.3.16, we have

$$-c^2 dt^2 + R^2(t) \frac{d\chi^2}{1 - k\chi^2} = 0.$$ 

Thus,

$$cdt = -\frac{R(t)d\chi}{\sqrt{1 - k\chi^2}}.$$ 

and

$$\int_{t_e}^{t_0} \frac{cdt}{R(t)} = \int_{0}^{\chi_c} \frac{d\chi}{\sqrt{1 - k\chi^2}}. \quad (4.13)$$

Now consider another photon, which is emitted a bit later, at time $t_e + dt_e$. It will be received a bit later too, at time $t_0 + dt_r$. Repeating the above calculations we then find

$$\int_{t_e + dt_e}^{t_0 + dt_r} \frac{cdt}{R(t)} = \int_{0}^{\chi_c} \frac{d\chi}{\sqrt{1 - k\chi^2}}. \quad (4.14)$$

From the last two equation we see that

$$\int_{t_e + dt_e}^{t_0 + dt_r} \frac{cdt}{R(t)} - \int_{t_e}^{t_0} \frac{cdt}{R(t)} = 0. \quad (4.15)$$

From the definition of definite integral, for any function $f(t)$

$$\int_{t_e + dt_e}^{t_0 + dt_r} f(t)dt = -f(t_e)dt_e + \int_{t_e}^{t_0 + dt_r} f(t)dt = +f(t_0)dt_r - f(t_e)dt_e + \int_{t_e}^{t_0} f(t)dt.$$ 

Thus, from Eq.4.15 we have

$$\int_{t_e + dt_e}^{t_0 + dt_r} \frac{cdt}{R(t)} = \frac{dt_r}{dt_e} = \frac{R_0}{R(t_e)}. \quad (4.16)$$

Now imagine that instead of two separate pulses of light we are dealing with two successive crests of a light wave with period at the source $T_e = dt_e$. At the receiver its period will be $T_r = dt_r$. Then Eq.4.16 becomes

$$\frac{T_r}{T_e} = \frac{R_0}{R(t_e)}. \quad (4.17)$$ 

Since the wavelength and period of a light wave are related via $T = \lambda/c$ this leads to

$$\frac{\lambda_r}{\lambda_e} = \frac{R_0}{R(t_e)}. \quad (4.17)$$
Because $t_0 > t_e$, for the expanding Universe we have $R_0/R(t_e) > 1$ and thus at the receiver the wavelength is longer, explaining the effect of cosmological redshift. Using the redshift parameter $z$ this result reads

$$\frac{R_0}{R(t_e)} = 1 + z.$$  \hspace{1cm} (4.18)

Remarkably, the curvature of space does not appear in the final result.

Equation 4.18 shows that the redshift of a distant source is a measure of the total expansion of the Universe that has occurred while the light was travelling between the source and the observer. It does not tell us the distance to the source or how long ago the light was emitted. These quantities depend on the precise nature of the Universe expansion between the instances of emission and observation. In different cosmological models we obtain different results for the same redshift. In the next section we give one particular example, which turns out to be very important.

Obviously, Eq.4.18 is different from Eq.1.4, which we naively used to explain the Hubble’s observations at first. No velocity is present here. However, Eq.1.2 also does not involve velocity and one can show that this equation is an approximation of Eq.4.18 for close sources, meaning that $R(t_e)$ is not much different from $R_0$, $t_e$ is close to $t_0$, and the redshift $z \ll 1$. Using the first two terms of the Taylor expansion (4.10) for $R(t)$ we have

$$\frac{R(t_e)}{R(t_0)} \simeq 1 + H_0(t_e - t_0),$$ \hspace{1cm} (4.19)

where $H_0 = \dot{R}_0/R_0$. Thus,

$$1 + z \simeq \frac{1}{1 + H_0(t_e - t_0)}.$$  

Since $z \ll 1$, this shows that $H_0(t_e - t_0) \ll 1$, and we can use the method of truncated Taylor series once more to write

$$1 + z \simeq 1 - H_0(t_e - t_0).$$

Finally, for close sources we can approximate $(t_e - t_0) \simeq -r/c$, which holds for all cosmological models, and obtain

$$z \simeq \frac{H_0}{c} r.$$ \hspace{1cm} (4.20)

This is indeed the Hubble law Eq.1.2.

The frequency of a photon is $f = c/\lambda$. From Eq.4.17 it follows that as a photon travels across our expanding Universe its wavelength increases and its frequency decreases as

$$\lambda(t) \propto R(t) \quad \text{and} \quad f(t) \propto 1/R(t).$$

Thus, the energy of the photon also decreases

$$E(t) = hf(t) \propto 1/R(t).$$ \hspace{1cm} (4.21)

The expression (1.4) and Hubble’s interpretation of the cosmological redshift based on the Doppler effect are inadequate. In General Relativity, only the relative velocity between close objects (essentially with the same spatial location) has well defined meaning. The new explanation, which we have presented in this section, does not appeal to the relative motion and hence is fully consistent with General Relativity. As we have already commented, in Hubble’s interpretation galaxies correspond to bugs on a balloon all running away from a given point. In the modern interpretation, these bugs are sitting quietly on the balloon surface, but the balloon itself is being inflated instead.

4.2.2 Critical density

Since the stationary Universe is physically impossible, and seems to be in conflict with the observations of distant galaxies, it makes sense to go back to the cosmological equations without the cosmological constant and explore their solutions. This is exactly where the attention of cosmologists turned to after the Hubble’s discovery.

Consider the Friedmann equation
\[
\left( \frac{\dot{R}}{R} \right)^2 = \frac{8\pi}{3} G \rho - \frac{k c^2}{R^2},
\]
and the acceleration equation
\[
\frac{\ddot{R}}{R} = -\frac{4\pi G}{3} \left( \rho + \frac{P}{c^2} \right)
\]
without the cosmological constant. Since both for matter and radiation \( \rho > 0 \) and \( P > 0 \), the acceleration equation shows that \( \ddot{R} < 0 \) \( (q_0 > 0) \). Hence, we conclude that, without the cosmological constant, *universe's expansion must be slowing down!*

What is even more interesting, the Friedmann equation shows that the actual geometry of the Universe without the cosmological constant can be deduced from its current density \( \rho_0 \) and the Hubble constant. Indeed, from Eq.4.22 we have that
\[
H_0^2 = \frac{8\pi}{3} G \rho_0 - \frac{k c^2}{R_0^2},
\]
or
\[
\rho_0 - \rho_c = \frac{3c^2}{8\pi G R_0^2} k,
\]
where
\[
\rho_c = \frac{3H_0^2}{8\pi G}.
\]
Now one can see that the Universe is closed if \( \rho_0 > \rho_c \), flat if \( \rho_0 = \rho_c \), and open if \( \rho_0 < \rho_c \). For this reason \( \rho_c \) is called *the critical density*. It is convenient to describe how close the current density in the Universe to the critical value by the dimensionless parameter
\[
\Omega_0 = \frac{\rho_0}{\rho_c}.
\]
It is called *the critical parameter*. If \( \Omega_0 < 1 \) then \( \rho_0 < \rho_c \) and \( k < 0 \) – the Universe is open. If \( \Omega_0 > 1 \) then \( \rho_0 > \rho_c \) and \( k > 0 \) – the Universe is closed. If \( \Omega_0 = 1 \) then the Universe is flat.

Let us see what are the numbers according to the astronomical observations. First, the Hubble constant. In the Hubble law, \( v = H_0 r \), the most convenient unit for the velocity is \( \text{km/s} \) and for the distance is \( \text{Mpc} \) (see the list of units before Chapter 1). This explains the rather peculiar traditional expression for the Hubble constant
\[
H = 100h \frac{\text{km/s}}{\text{Mpc}},
\]
where the dimensionless factor \( h \) is introduced to accommodate the current uncertainty in \( H \). The most accurate measurements to date give
\[
h = 0.72 \pm 0.08.
\]
The corresponding critical density in CGS units is
\[
\rho_c = 1.88h^2 10^{-29} \text{g/cm}^3.
\]
The number is startlingly small because the CGS units are tailored to our Earthly conditions rather than to the vastness of the Universe. More suitable is to measure mass in galaxies \( (10^{11} \text{ solar masses}) \) and length in \( \text{Mpc} \). Then
\[
\rho_c = 2.78h^2 \frac{\text{galaxies}}{\text{Mpc}^3}.
\]
From astronomical observations, the visible matter in the local Universe gives
\[
\Omega_{vm} \simeq 0.02h^{-2}.
\]
The contribution of radiation is even smaller,
\[
\Omega_r \simeq 4 \times 10^{-5} h^{-2}.
\]
Thus, based on these data alone one would conclude that our Universe is open. However, this is not the full story. In addition to the visible matter and radiation there are other gravitating elements in the Universe, the dark matter and the dark energy. Then they are accounted for, the critical parameter of the Universe becomes very close to unity, pointing towards our Universe being almost flat. We will come back to this issue later.

4.3 The Friedmann models

Will the deceleration be able eventually to stop the expansion and turn it into a contraction? At this point $\dot{R}$ would vanish. However, Eq.4.22 shows that this is not possible if $k \leq 0$. Thus, both the flat and the open Universes are destined to expand forever. For the closed Universe this equation suggests that it may be possible to reach the state where $\dot{R} = 0$. Since $\dot{R} < 0$ this is a maximum of $R(t)$ and the Universe will begin to contract after this point.

Since at present the radiation makes only a very small contribution to the energy-density, it makes sense to ignore it and consider the models with $P = 0$. This is exactly what was done by Alexander Friedmann in his study. In this case the Friedmann equations read

$$\left( \frac{\dot{R}}{R} \right)^2 = \frac{8\pi G}{3} \rho - \frac{k c^2}{R^2}, \quad (4.28)$$

$$\frac{\ddot{R}}{R} = -\frac{4\pi G}{3} \rho, \quad (4.29)$$

and

$$\dot{\rho} = -3 \frac{\dot{R}}{R} \rho. \quad (4.30)$$

Integrating Eq.4.30 we find that

$$\rho R^3 \rho_0 R_0^3 = \text{const.} \quad (4.31)$$

Substituting this result into Eq.4.28 we obtain the differential equation for $R$

$$\dot{R}^2 = c^2 \left( \alpha^2 \frac{R}{R} - k \right), \quad (4.32)$$

where

$$\alpha^2 = \frac{8\pi G}{3} \frac{\rho_0 R_0^3}{c^2}. \quad (4.33)$$

Below, we integrate separately for the all three possible values of $k$.

Before this we note that from Eqs.4.29 it follows that

$$\frac{\dot{R}_0}{R_0} = -\frac{4\pi G}{3} \rho_0. \quad (4.34)$$

When combined with the definition of the deceleration parameter, this result means that

$$\rho_0 = \frac{3H_0^2}{4\pi G} q_0 = 2q_0 \rho_c. \quad (4.35)$$

Thus, in all Friedmann’s models

$$\Omega_0 = 2q_0. \quad (4.36)$$

4.3.1 Flat Universe

In this case $k = 0$, $\Omega_0 = 1$, $q_0 = 1/2$, and Eq.4.32 reads

$$\dot{R}^2 = \frac{\alpha^2 c^2}{R}. \quad (4.37)$$
Integrating this equation we find
\[ R^{3/2} = at + C. \]
where \( a = \sqrt{\alpha^2 c^2} \) and \( C \) is the integration constant. From this it is clear that at some time \( t_{BB} \) we have \( R = 0 \). Resetting the clocks so this time becomes \( t = 0 \) (which is equivalent to imposing the boundary condition \( R(0) = 0 \)) we obtain
\[ R = At^{2/3}, \quad (4.36) \]
where \( A \) is constant. Thus, the Universe expands forever, as we have already argued. From the last equation we also find that
\[ \dot{R} = \frac{2R}{3t}. \]
Thus,
\[ H_0 = \frac{R_0}{t_0} = \frac{2}{3H_0}, \quad (4.37) \]
which shows that the current age of the Universe is
\[ t_0 = \frac{2}{3H_0} = 6.51 \times 10^9 h^{-1} \text{yr}. \quad (4.38) \]

### 4.3.2 Closed Universe

In this case \( k = 1, \rho_0 > \rho_c, \Omega_0 > 1, q_0 > 1/2 \), and Eq.4.32 reads
\[ \dot{R} = c \left( \frac{\alpha^2}{R} - 1 \right)^{1/2}. \quad (4.39) \]
Integrating this equation we find
\[ \int \frac{\sqrt{R} dR}{\sqrt{\alpha^2 - R}} = ct + C, \]
where \( C \) is the constant of integration. Via the substitution
\[ R = \alpha^2 \sin^2(x), \quad (4.40) \]
the integral on the left-hand side of this equation is reduced to
\[ 2\alpha^2 \int \sin^2(x) dx = \alpha^2 (x - \sin(2x)/2), \]
where we used
\[ \int \sin^2 x dx = (1/2)(x - \sin x \cos x), \]
Hence
\[ \alpha^2 (x - \sin(2x)/2) = ct + C. \]
Here again we see that \( \delta = 0 \) and hence \( R = 0 \) at time \( t = -C/c \). Resetting the clocks so that \( R(0) = 0 \) at \( t = 0 \) we obtain \( C = 0 \) and
\[ \alpha^2 (x - \sin(2x)/2) = ct. \quad (4.41) \]
Equations 4.40 and 4.41 determine \( R \) as an implicit function of \( t \). Let us analyze this result. Since
\[ \dot{x} = \frac{c}{\alpha^2 (1 - \cos 2x)} > 0 \]
\( x(t) \) is an increasing function of \( t \). Then Eq.4.40 shows that the initial expansion of the Universe terminates when \( x \) reaches \( \pi/2 \) and then the Universe begins to contract, eventually collapsing when \( x = \pi \), as we anticipated. The corresponding times are \( t_{max} = \pi \alpha^2/2c \) and \( t_{coll} = \pi \alpha^2/c \) respectively.
\( \alpha^2 \) can be expressed in terms of \( H_0 \) and \( q_0 \). Using Eqs. 4.28 we obtain

\[
H_0^2 = \frac{8\pi}{3} G \rho_0 - \frac{c^2}{R_0^2}.
\]

Combining this result with Eq. 4.33 we find that

\[
R_0 = \frac{c}{H_0(2q_0 - 1)^{1/2}},
\]

and

\[
\alpha^2 = \frac{2q_0 c}{H_0(2q_0 - 1)^{3/2}}.
\] (4.42)

Simple asymptotic analysis of the obtained solution shows that for \( t \ll \alpha^2/c \)

\[
R \propto t^{2/3},
\]

just like in the flat Universe at all times. This is illustrated in Figure 4.1.

![Figure 4.1: Evolution of closed, flat, and open Universes in the Friedmann solutions.](image)

### 4.3.3 Open Universe

In this case \( k = 1, \rho_0 < \rho_c, \Omega_0 < 1, q_0 < 1/2 \), and Eq. 4.32 reads

\[
\dot{R} = c \left( \frac{\alpha^2}{R} - 1 \right)^{1/2}.
\] (4.43)

Integrating this equation we find

\[
\int \frac{\sqrt{R} dR}{\sqrt{\alpha^2 + R}} = ct + C.
\]

Via the substitution

\[
R = \alpha^2 \sinh^2(x),
\] (4.44)

the integral on the left-hand side of this equation is reduced to

\[
2\alpha^2 \int \sinh^2(x) dx = \alpha^2 (\sinh(2x)/2 - x),
\]
where we used
\[ \int \sinh^2 x \, dx = (1/2) \sinh x \cosh x - x = (1/2) (\sinh(2x)/2 - x). \]

Here again we see that \( x = 0 \) and hence \( R = 0 \) at time \( t = -C/c \). Resetting the clocks so that \( R(0) = 0 \), we obtain
\[ \alpha^2 (\sinh(2x)/2 - x) = ct. \] (4.45)

Equations 4.44 and 4.45 determine \( R \) as an implicit function of \( t \). Let us analyze this result. Since
\[ \dot{x} = \frac{c}{\alpha (\cosh(2x) - 1)} > 0, \]
\( x(t) \) is an increasing function of \( t \). Then Eq.4.44 shows that the Universe expands forever.

Again, \( \alpha^2 \) can be expressed in terms of \( H_0 \) and \( q_0 \). Using Eqs.4.28 we obtain
\[ H_0^2 = \frac{8\pi}{3} G \rho_0 + \frac{c^2}{R_0^2}. \]

Combining these equations with the definition of the deceleration parameter we find that
\[ R_0 = \frac{c}{H_0 (1 - 2q_0)^{3/2}}, \]
and
\[ \alpha^2 = \frac{2q_0 c}{H_0 (1 - 2q_0)^{3/2}}. \] (4.46)

Simple asymptotic analysis of the obtained solution shows that for the open Universe we still have
\[ R \propto t^{2/3} \]
for \( t \ll \alpha^2/c \). Moreover, for \( t \gg \alpha^2/c \) the open Universe expands as
\[ R \propto t, \]
which is faster than in the flat Universe solution. This is illustrated in Figure 4.1.

4.4 The Big Bang

In all three models of Friedmann, the solutions continue into the past to the time where \( R = 0 \) (We have chosen \( t \) to be 0 when \( R = 0 \)). Thus, the whole of the Universe is created in a single explosion, which is called the Big Bang. As \( t \to 0 \) and \( R \to 0 \) the energy-density of matter \( \rho \to \infty \). The energy density of radiation grows even faster. Indeed, since for radiation \( P = \rho c^2/3 \), the fluid equation for radiation reads
\[ \dot{\rho}_r = -4 \frac{\dot{R}}{R} \rho_r. \] (4.47)

This equation is easily integrated, leading to
\[ \rho_r \propto R^{-4}, \] (4.48)
which grows faster than that of matter
\[ \rho_m \propto R^{-3} \] (4.49)
when we consider progressively younger Universe, \( R \to 0 \) (see Eq.4.31). In fact,
\[ \rho_r/\rho_m \propto R^{-1} \to \infty. \]

At present,
\[ \rho_m/\rho_r \simeq 6000. \]
Thus, we have equipartition, $\rho_m = \rho_r$, when $R = R_{eq}$, where
$$R_{eq} \simeq R_0/6000$$
and for $R < R_{eq}$ the Universe is dominated by radiation. Moreover, the temperature of matter
$$T \propto \frac{P}{\rho} \propto \rho^{\Gamma^{-1}} \quad (4.50)$$
also grows as $R \to 0$ since $\Gamma > 1$. At some point it becomes relativistic and the properties of matter become the same as those of radiation, with $P_m = \rho_m c^2 / 3$. Can the fact that the Universe becomes radiation-dominated change the conclusion about the Big Bang? It turns out that it does not. For example consider the case of flat Universe. Assuming that $\rho = \rho_r$ the Friedmann equation (Eq.4.28) reads
$$\frac{\dot{R}}{R} = \frac{a}{R^2},$$
where $a$ is constant. Given the initial condition $R(t_0) = R_0$ the solution of this equation is
$$R^2 - R_0^2 = 2a(t - t_0).$$
From this result one can see that at finite time
$$t_{BB} = t_0 - \frac{R_0^2}{2a}$$
the scaling factor vanishes,
$$R(t_{BB}) = 0.$$
Reseting the cosmological time so that $t_{BB} = 0$, we find
$$R \propto t^{1/2}, \quad (4.51)$$
$$\rho_r \propto R^{-4} \propto t^{-2}, \quad (4.52)$$
$$\rho_m \propto R^{-3} \propto t^{-3/2}, \quad (4.53)$$
$$H \equiv \frac{\dot{R}}{R} = \frac{1}{2t}. \quad (4.54)$$
(Notice that the Hubble constant is not really a constant – it varies with time!) Thus, the Universe again emerges from a singularity with infinite density and temperature. Immediately after the creation, we cannot possibly have individual stars or galaxies – all this structure must have been formed much much later. Initially, all there could be is just some very hot “sea” of matter and radiation, dominated by radiation. In fact, close to the Big Bang the conditions are so extreme that they are beyond the realm of modern Physics, making the Friedmann equations unjustified.

### 4.5 Advanced topic: Cosmic microwave background radiation

#### 4.5.1 Relic radiation

When you heat a piece of metal to high temperature it begins to glow, to emit light. In fact, baryonic (or just normal) matter emits thermal electromagnetic radiation at any temperature, but not always as visible light. In fact, higher temperature lead to higher frequency and higher intensity of the thermal emission. As the Universe must have been very hot in the past (see Eq.4.50), its matter was producing a lot of radiation. Where is it now? At present, the Universe is transparent to electromagnetic radiation, it is mostly empty. However, the opacity increases with density and in the past the Universe must have been opaque, with photons being absorbed soon after their emission. We should see the photons produced at the time of transition to transparency.
In 1965 two American radio astronomers, Penzias and Wilson, accidently discovered some strange microwave radiation, apparently coming to the Earth from every direction in the sky. Now, this radiation is called the cosmic microwave background (CMB) radiation. It has the classical black body thermal spectrum corresponding to a very low temperature, only \( \approx 2.7 \) K. There is no other plausible explanation for it, other than cosmological – that this is the radiation left from the early Universe, “the relic radiation”. But why is its temperature so low if the early Universe was very hot?

### 4.5.2 Black body spectrum

By the black body we mean here, we mean a body that does not reflect light. Instead it absorbs and reprocesses any incident electromagnetic radiation. The energy distribution of the black body radiation is given by

\[
\varepsilon(f) = \frac{8\pi h f^3}{c^3 \exp[f/kT] - 1},
\]

(4.55)

where \( f \) is the frequency of radiation, \( h \) is the Planck’s constant, \( k \) is the Boltzmann constant and \( T \) is the temperature. The meaning of this function is explained by this equation

\[
de = \varepsilon(f) dV df,
\]

where \( de \) is the total energy of photons with the frequencies from \( f \) to \( f + df \) within volume \( dV \). That is \( \varepsilon(f) \) is the radiation energy per unit volume and unit frequency range. The energy of an individual photon with frequency \( f \) is

\[
E = hf
\]

(4.56)

Equation 4.55 can be written as

\[
\varepsilon(f) = \frac{8\pi k^3 T^3}{c^3 h^2} F(x),
\]

(4.57)

where \( x = hf/kT \) and

\[
F(x) = \frac{x^3}{e^x - 1}.
\]

This function is shown in Figure 4.2. It has a maximum at \( x \approx 2.8 \), which corresponds to the photon energy \( E \approx 2.8kT \). For \( x \ll 1 \), \( F(x) \approx x^3 \), and for \( x \gg 1 \), \( F(x) \approx x^3 e^{-x} \).

The number of photons per unit volume and unit frequency range is

\[
n(f) = \frac{\varepsilon(f)}{hf} = \frac{8\pi}{c^3} \frac{f^2}{e^{hf/kT} - 1}.
\]

(4.58)

Integrating \( \varepsilon(f) \) over the whole range of frequencies, which requires a lot of skill, one finds the total energy and mass densities of the radiation

\[
e_r = \int_0^\infty \varepsilon(f) df = a_r T^4 \quad \text{and} \quad \rho_r = (a_r/c^2) T^4,
\]

(4.59)

where \( a_r = 8\pi^5 k^4 / 15 h^3 c^3 \simeq 7.565 \times 10^{-15} \text{erg cm}^{-3} \text{K}^{-4} \) is called the radiation constant. Similarly, one finds the number density of photons

\[
n_r = \int_0^\infty n(f) df = b_r T^3;
\]

(4.60)

where \( b_r = 20.3 \text{cm}^{-3} \text{K}^{-3} \). The mean energy of photons is

\[
\langle E \rangle = \frac{e_r}{n_r} \simeq 3kT.
\]

(4.61)
4.5.3 History of CMB radiation

Here we summarize the origin and evolution of CMB based on accurate and detailed calculations which involve a lot of physics.

- When the temperature of the Universe was $kT > 13.6\text{eV}$ (where eV is the most convenient unit of energy in atomic physics), atoms did not exist. Matter was fully ionized (plasma). The opacity was very high and the Universe was opaque to its thermal radiation.

- When $13.6\text{eV} < kT < 0.3\text{eV}$ the Universe was only partly ionized, the opacity was lower but the Universe was still opaque to its thermal radiation. Matter was coupled with the radiation, so that $T_{\text{matter}} = T_{\text{radiation}}$. The most important interaction is the so-called Compton scattering of photons by free electrons.

- When $kT$ had dropped below $0.3\text{eV}$ ($T \approx 3000\text{K}$), the ionization level had became too low (almost no ions). The Universe had become transparent to its thermal radiation. Matter had decoupled from its thermal radiation, so that $T_{\text{matter}} \neq T_{\text{radiation}}$. Firstly, the density of free electrons had dropped and hence the Compton scattering had become inefficient. Secondly, fewer and fewer photons could interact with the electrons bound in atoms. This is because the bound electrons can change their energy only in discrete portions, or quanta, and there is a minimum energy, which a photon must have in order to interact with a bound electron (Its value is not much lower than the ionization energy, $13.6\text{eV}$.)

- Today, for the CMB temperature of $T = 2.7K$, we have

$$\rho_{\text{CMB}} = 4.47 \times 10^{-34}\text{g cm}^{-3}, \quad n_{\text{CMB}} = 400\text{cm}^{-3}, \quad \langle E \rangle = 7 \times 10^{-4}\text{eV}.$$  

There are much more photons in the Universe compared to baryons, whose present number density is

$$n_b = 2.2 \times 10^{-5}\text{cm}^{-3}.$$  

When average over large volume, CMB radiation dominates all other kinds of electromagnetic radiation in the Universe. Though locally, close to strong sources of other emission, the relative CMB energy density can be low.
4.5.4 Evolution of CMB radiation after decoupling with matter

Here we show that in transparent expanding Universe the black body spectrum of CMB is preserved, but at a lower temperature. As we have seen in Sec. 4.4, for radiation

\[ \rho_r \propto R^{-4}. \]

Since photons are stable particles and do not decay spontaneously their number is conserved and hence

\[ n_r \propto R^{-3}. \]

Thus, the energy of a single photon

\[ E = hf \propto \frac{\rho_r}{n_r} \propto R^{-1} \quad \text{and} \quad f \propto R^{-1}. \]  

(4.62)

This is exactly the result we have already derived in Eq. 4.21. Suppose that at time \( t = t_1 \) the scale factor \( R = R_1 \) and the radiation has the black body spectrum

\[ n_1(f_1,t_1) = \frac{8\pi}{c^3} \frac{f_1^2}{e^{hf_1/kT_1} - 1} \]

with temperature \( T = T_1 \). Consider a small element of the cosmological fluid, expanding with the Universe, Denote as \( dV_1 \) its volume at time \( t_1 \). The total number of photons with frequency between \( f_1 \) and \( f_1 + df_1 \) in this volume is

\[ dN = n_1(f_1,t_1) dV_1 df_1. \]

As the Universe expands and its scaling factor increases to \( R_2 = \alpha R_1 \) (\( \alpha > 1 \)), the volume of the fluid element becomes

\[ dV_2 = dV_1 \alpha^3. \]

The frequency of photons decreases to

\[ f_2 = f_1 / \alpha, \]

so their move into the frequency “window” \( df_2 = df_1 / \alpha \). Thus, the total number of photons within the volume \( dV_2 \) and the frequency range \( (f_2, f_2 + df_2) \) at time \( t_2 \) is the same it was within the volume \( dV_1 \) and the frequency range \( (f_1, f_1 + df_1) \) at time \( t_1 \). That is

\[ n_1(f_1,t_1) dV_1 df_1 = n_2(f_2,t_2) dV_2 df_2. \]

Thus,

\[ n_2(f_2,t_2) = n_1(f_1,t_1) \frac{dV_1}{dV_2} \frac{df_1}{df_2} = n_1(f_1,t_1) \alpha^{-2}. \]

The next step is to use the black-body spectrum formula for \( n_1(f_1,t_1) \), which gives us

\[ n_2(f_2) = \frac{8\pi}{c^3} \frac{\alpha^{-2} f_1^2}{e^{hf_1/kT_1} - 1} = \frac{8\pi}{c^3} \frac{f_2^2}{e^{hf_2/\alpha kT_1} - 1}. \]

This, the black body spectrum for the temperature \( T_2 = T_1 / \alpha = T_1 R_1 / R_2 \). Thus, the black body spectrum of photons is preserved, but at a lower temperature, which depends on the scaling factor as

\[ T \propto R^{-1}. \]  

(4.63)

From Eq.4.63, one can conclude that at the time of decoupling the radius of the Universe was

\[ R_{dec} = \frac{2.7}{3000} R_0 \approx 10^{-3} R_0, \]

where \( R_0 \) is its current value. This is about six times larger compared to the scale of the equipartition and, thus, at this point the Universe is already matter-dominated.
4.6 Some other stages of Big Bang

4.6.1 Nucleosynthesis

The typical binding energy of complex atomic nuclei is around 1MeV. Thus, in a very hot young Universe, with $kT > 1\text{MeV}$, atomic nuclei could not survive the constant bombardment by extremely energetic particles and baryons, particles which assemble to make nuclei, were free. As the Universe expanded and cooled to $kT \lesssim 1\text{MeV}$, nuclei were formed. The detailed analysis of this primordial nucleosynthesis shows that only light elements were formed and the early Universe was made mainly of hydrogen. Heavier elements, including carbon, were produced much later by stars.

4.6.2 Structure formation

If soon after Big Bang the Universe was a uniform “sea” of hot plasma then how has all this presently observed structure, stars, galaxies etc., been formed? The basic picture is similar to that of the currently observed star formation in the galactic molecular clouds and the key process is the gravitational instability.

Consider the freely expanding Cosmological fluid. There is no gravity force acting on any fluid element. Now imagine that some fluid element is slightly compressed, and hence its mass density is slightly higher compared to its surrounding. The Universe is no longer exactly homogeneous and the gravitational force is no longer vanishing. The fluid element begins to feel its enhanced self-gravity. This force promotes further compression leading to even stronger self-gravity and so on. When the fluid is cold this instability always develops. However, in hot fluid it is stabilized by the pressure force. As the fluid element gets compressed, its pressure increases and tends to restore the initial volume.

Before the decoupling between matter and radiation in the Universe, the pressure was very high, dominated by the radiation pressure $P_r = e_r/3$, preventing the development of the gravitational instability. Instead, only sound waves were excited. At the decoupling, the pressure suddenly dropped, as the radiation was no longer contributing to it. After this, the gravitational instability was winning on many length scales, leading to the structure we see today.

The spectrum of sound waves generated before the time of decoupling can be determined rather accurately using well known physics. Because of these waves, the Universe was not completely homogeneous at the time of decoupling and hence the microwave background emission cannot be completely isotropic. Instead, there must be small fluctuations in its brightness, and hence temperature on the sky. Astronomers were looking for such fluctuations for many years, but all in vain. Only in 1992 they had been finally found with an instrument on board of the COBE satellite. This was a real triumph of modern Cosmology. The amplitude of the fluctuations is very small indeed,

$$\frac{\delta T}{T} \simeq 10^{-5},$$

where $T$ is the temperature.

Moreover, the study of the anisotropy has proved to be crucial for determining the geometry of our Universe. The angular scale of the fluctuations strongly depends on the spatial curvature. The fitting of the observational data with theoretical models has allowed cosmologists to conclude that within observational errors our Universe is flat.
Chapter 5

Measuring the Universe

5.1 Distances to sources with given redshift

In the case of distant sources, the relation between the source redshift and its distance from us is different from that of Eq.1.2. Consider as an example the case of matter-dominated flat Universe. From the results obtained in Sec.4.4 we know that in this case

\[ R(t) = R_0 \left( \frac{t}{t_0} \right)^{2/3}, \]

where as usual \( t_0 \) and \( R_0 \) are the current time and scale factor of the Universe. Combining this result with Eq.4.18 we find the time of emission as a function of the redshift

\[ t_e = t_0 (1 + z)^{-3/2}. \]

Substituting \( R(t) \) from Eq.5.1 into Eq.4.13 and integrating, we find that

\[ \chi_e = 3c t_0 \frac{R_0}{R_e} \left( t_0^{1/3} - t_e^{1/3} \right). \]

Next we use the result Eq.5.2 in order to replace \( t_e \) with \( z \) and find that

\[ \chi_e = \frac{3c t_0}{R_0} (1 - (1 + z)^{-1/2}). \]

This equation tells us the comoving coordinate of the source with known redshift. In order to find the current metric distance to the source, we notice that the distance along the radial direction in flat Universe is simply \( dl = dr = R_0 d\chi \) and hence

\[ r(z) = R_0 \chi_e = 3c t_0 \left( 1 - (1 + z)^{-1/2} \right). \]

Combining this result with Eq.4.37 for the age of the matter-dominated flat Universe we finally find

\[ r(z) = \frac{2c}{H_0} \left( 1 - (1 + z)^{-1/2} \right), \]

which is indeed different from the Hubble law Eq.1.2. However, it is easy to check that the Hubble law is recovered from this result in the limit \( z \ll 1 \). Indeed, using the first two terms of Taylor’s expansion we obtain

\[ (1 - (1 + z)^{-1/2}) \sim 1 - (1 - (1/2)z) = (1/2)z, \]

and thus

\[ r(z) \sim \frac{e}{H_0} z. \]
5.2 The cosmological horizon

The Universe may be finite or infinite but can we actually verify this? How far can we “see”? Or to be more accurate, how far are the objects in the Universe which have had enough time since the birth of the Universe to communicate their presence to us, at least in principle? In order to answer this question, or even to see if this question makes sense, we need to find out how far the fastest possible signal can reach since the Big Bang. Obviously, this has to be a signal propagating with the speed of light along a “straight line” (a geodesic to be more accurate). In fact, we have already derived the required equation – this is Eq.4.13 with $t_c = 0$

\[ \int_0^t \frac{cdt}{R(t)} = \frac{z_0}{\sqrt{1-k\chi^2}}. \]  

(5.8)

Indeed, this equation defines the comoving radial coordinate $\chi_h(t)$ at time $t$ of the signal that was produced at time $t = 0$ at the origin of the coordinate grid. Obviously, the result depends on the geometry of the Universe, which we need to know in order to evaluate the integral on the right hand side of Eq.5.8, and on the history of the Universe, as we need to know $R(t)$ as well.

Consider, as an example, the model of flat matter-dominated Universe. Since $k = 0$, we have

\[ \int_0^t \frac{cdt}{R(t)} = \frac{3c}{A} \cdot t^{1/3}. \]

(5.9)

Thus, we have

\[ \chi_h(t) = \frac{3cA^{1/3}}{A}. \]

(5.10)

This result shows that we can explore only a finite part of the Universe. The rest is, so to speak, beyond the “horizon”. This is why $\chi_h(t)$ is called the comoving radius of the cosmological horizon. As the Universe gets older, the horizon expands, allowing us to see more of the Universe. However, this is still only an insignificant fraction as the Universe is infinite. The metric radius of the horizon in this model is

\[ r_h(t) = R(t)\chi_h(t) = 3ct. \]

(5.11)

At first glance this result appears rather odd as $r_h$ is larger than the speed of light times the propagation time, which seems to be in conflict with the Theory of Relativity. However, there is no conflict. Indeed, as the signal propagates the length of any given section of the path it has already made increases due to the expansion of the Universe.

In order to convert $\chi_h(t)$ into $r_h(t)$ for other models, we note that along the path of the signal both $\theta$ and $\phi$ are constant, and hence the line element of this path is

\[ dl = R(t)\frac{d\chi}{\sqrt{1-k\chi^2}}. \]

(5.12)

Thus,

\[ r_h(t) = R(t) \int_0^{\chi_h(t)} \frac{d\chi}{\sqrt{1-k\chi^2}}. \]

(5.13)

Thus, for the closed Universe ($k = 1$)

\[ r_h(t) = R(t)\sin^{-1}(\chi_h(t)); \]

(5.14)
and for the open Universe ( \( k = -1 \) )

\[
    r_h(t) = R(t) \sinh^{-1}(\chi_h(t)).
\]  

(5.16)

The horizon radius is very important for the causal structure of the Universe. Whatever events may have occurred in the Universe beyond the horizon centered on the Earth, they could not have had any effect on us. The Universe may have significant structure on scales \( \chi \gg \chi_h(t_0) \), contrary to the Cosmological Principle, and have evolved in a completely different manner beyond the horizon. However, we cannot tell if this true so or not in principle.

### 5.3 The angular size distance

Consider a rod of fixed metric length \( l \) placed at a distance \( r \gg l \) from us and perpendicular to our line of sight. In Euclidean geometry the angular size of such rod, as seen by us, is

\[
    \alpha = \frac{l}{r}.
\]  

(5.17)

Thus, given the angular size and the metric size (length) of the rod one can find the distance to the rod via

\[
    r_{\text{ang}} = \frac{l}{\alpha}.
\]  

(5.18)

This suggests a way of measuring distances in the Universe, provided it has objects of fixed metric size, a kind of “standard rods”. In Cosmology, the quantity determined this way is called in the angular diameter distance. Its meaning, however, is not that obvious. First of all, unless the Universe is flat its spatial geometry is not Euclidean. Secondly, because of the expansion of the Universe the observed angular size is different from the real one. Indeed, imagine two photons emitted from the ends of the rod towards us at time \( t_e \). Relative to the comoving spherical grid whose origin \( O \) is fixed at our location, these photons propagate radially towards the origin along the sides OA and OB of the triangle OAB whose points A and B mark the ends of the rod on this grid at the time of emission (see Fig.5.1). As the photons propagate towards us, the Universe expands and the metric distance between points A and B increases. However, the length of the rod itself is fixed. In other words, relative to the comoving grid the rod contracts – in Fig.5.1 its ends at a later time are shown as \( A' \) and \( B' \). For this reason, the actual angular size of the rod at this time, \( \beta \), is different from observed angular size \( \alpha \).

Now recall that the metric form of space in the Robertson-Walker coordinates is

\[
    dl^2 = R^2(t) \left[ \frac{d\chi^2}{1 - k\chi^2} + \chi^2 (d\theta^2 + \sin^2\theta d\phi^2) \right].
\]  

(5.19)

Suppose that the rod is located at \( \chi = \chi_e \) and aligned with the \( \theta \) direction. Then along the rod \( d\chi = d\phi = 0 \) and we have

\[
    dl = R(t) \chi_e d\theta.
\]  

(5.20)

At the emission time, \( t_e \), this yields the rod length

\[
    l = R(t_e) \chi_e \alpha.
\]  

(5.21)

Substituting this result into Eq.5.18 we obtain

\[
    r_{\text{ang}} = R(t_e) \chi_e.
\]  

(5.22)

For a source with redshift \( z \) we have \( R(t_e) = R_0/(1 + z) \) for all cosmological models (see Sec.5.1) and each models has its own function \( \chi_e(z) \). Thus, for such a source

\[
    r_{\text{ang}} = \frac{R_0 \chi_e(z)}{1 + z}.
\]  

(5.23)
To confirm our expectation that $r_{\text{ang}}$ is not the actual distance to the source, consider the model of flat Universe. In this case, the distance to the source with comoving radius $\chi_e$ at the time of observation is $R_0\chi_e$ and Eq.5.22 tells us that this is not the same as $r_{\text{ang}}$.

Although, $r_{\text{ang}}$ does not give us the actual distance to the source, this is still a very useful parameter. It is useful because 1) it is a directly measurable parameter, 2) it is a function of another observable parameter $z$, 3) for different cosmological models this is a different function, because of different $\chi_e(z)$. Thus, by determining the function $r_{\text{ang}}(z)$ from observations and comparing it with the theoretical predictions one may try to determine which cosmological model fits our own Universe.

As an example consider the model of matter-dominated flat Universe. We have already derive the expression for $\chi_e(z)$ in Sec.5.1 (Eq.5.3)

$$\chi_e(z) = \frac{3c t_0}{R_0} (1 - (1 + z)^{-1/2}). \quad (5.24)$$

Substituting this in Eq.5.22 we find

$$r_{\text{ang}}(z) = \frac{2c}{H_0} \frac{(1 - (1 + z)^{-1/2})}{1 + z}. \quad (5.25)$$

A rather interesting result is obtained when $r_{\text{ang}}$ from Eq.5.25 is substituted into Eq.5.18 in order to find $\alpha$ as a function of $z$. This gives us

$$\alpha(z) = \frac{lH_0}{2c} \frac{1 + z}{(1 - (1 + z)^{-1/2})}. \quad (5.26)$$

Figure 5.2 shows the behavior of the function

$$f(z) = \frac{1 + z}{(1 - (1 + z)^{-1/2})}.$$

One can see that for small $z$ its value decreases with $z$, in agreement with the expectation that the angular size of an object with fixed linear size decreases with the distance. However, for $z > 1.25$ we have the opposite trend - the angular size increases with the distance!
For $z \ll 1$ Eq.5.26 reduces to Eq.5.17.

Although, there are no very convenient “standard rods” in the Universe, the dependence of angular size on redshift turns out to be very important in determining the angular size of the fluctuations in the microwave background emission, which originates at $z \approx 1000$. In particular, the outcome of theoretical calculations is very sensitive to the curvature parameter $k$. The best fit to observations is given by the model with $k = 0$, implying that the critical parameter $\Omega_0 \simeq 1$ and much larger than that estimated on the basis of visible matter, $\Omega_{0,\text{vm}} \simeq 0.02h^{-2}$. This tells us that in addition to the visible matter, and radiation, there must be some invisible components in the Universe, which account for most of its mass!

### 5.4 The luminosity distance

By the source luminosity, $L$, we understand the total amount of energy it emits per unit time. If this energy is conserved, and the source is not variable, then the same amount of energy passes, per unit time, through a sphere of radius $r$ centered on the source. Denote as $S$ the energy crossing the unit surface element of this sphere. This quantity is called the flux density and it is directly measured in observations.

If the source is isotropic then

$$S = \frac{L}{A},$$

where $A = 4\pi r^2$ is the area of the sphere of radius $r$ centered on the source. If the source is a some sort of “standard candle”, so we know its luminosity before hand, then we can find the distance to this source via

$$r_L = \sqrt{\frac{L}{4\pi S}},$$

However, in these calculations we assume Euclidean geometry and ignore the expansion of the Universe. Thus in Cosmology, this cannot be the real distance to the source. Yet this is still a very useful parameter. Motivated by its origin, it is called the luminosity distance.

There are two main effects that require to modify Eq.5.27 in expanding Universe. First, as the photons propagate to us their energy decreases as

$$E_r = E_0/(1 + z),$$
Figure 5.3: Fluctuations of the microwave background radiation.

where $E_e$ is the emitted energy and $E_r$ is the received one (see Sec.5.1). Similarly, the photons arrive less frequently

$$dt_r = dt_e (1 + z)$$

(see Eq.4.16). Thus, we should instead of Eq.5.27 we should have

$$S = \frac{L}{A(1 + z)^2}.$$  \hfill (5.29)

As to the surface area,

$$A = 4\pi R_0^2 \chi_e^2,$$  \hfill (5.30)

where $\chi_e$ is the comoving radius of the source in the coordinate system centered on the Earth. Indeed, at present time, the scale factor is $R(t_0) = R_0$ and the space metric is

$$dl^2 = R_0^2 \left[ \frac{d\chi_e^2}{1 - k\chi_e^2} + \chi_e^2 (d\theta^2 + \sin^2 \theta d\phi^2) \right].$$  \hfill (5.31)

According to this metric, the distance along the $\theta$ direction on the sphere with comoving radius $\chi_e$ is

$$dl_\theta = R_0 \chi_e d\theta,$$

and the distance along the $\phi$ direction is

$$dl_\phi = R_0 \chi_e \sin \theta d\phi.$$

Thus, the area of the surface element with the angular sizes $d\theta$ and $d\phi$ is

$$dA = dl_\phi dl_\theta = R_0^2 \chi_e^2 \sin \theta d\theta d\phi.$$  

The total area of the sphere is found via integration

$$A = \int dA = R_0^2 \chi_e^2 \int_0^{2\pi} d\phi \int_0^\pi \sin \theta d\theta = 4\pi R_0^2 \chi_e^2.$$
Next, we substitute $A$ from Eq.5.30 into Eq.5.29 to find

$$S = \frac{L}{4\pi R_0^2 \chi_e^2 (1+z)^2}$$

and then substitute this expression into Eq.5.28. The result is

$$r_L = R_0 \chi_e (1+z). \quad (5.32)$$

Comparison with Eq.5.22 yields

$$r_L = r_{\text{ang}} (1+z)^2. \quad (5.33)$$

As we have already discussed in Sec.5.3 any cosmological model gives us $\chi_e(z)$ for objects with known redshift. For matter-dominated flat Universe this is Eq.5.24. Using this result we find

$$r_L = 3c t_0 (1 - (1+z)^{-1/2})(1+z). \quad (5.34)$$

Thus, Eq.5.32 provides us with another model-dependent relationship between directly observable parameters that can be used to test suitability of cosmological models. The only problem is to find suitable natural “standard candles” in the Universe. Astrophysicists seems to have found such objects, namely the supernovae of type Ia.

### 5.5 Supernovae and the deceleration parameter

Supernovae are powerful stellar explosions can be seen from huge distances. For a short period, about a day, the luminosity of such supernova is comparable to the combined luminosity of all stars in its parent galaxy. Moreover, for type Ia supernovae, the peak luminosity is very much the same every time, and this allows us to use them as standard candles. Such explosions are rather rare and constant monitoring of a very large number of galaxies is required in order to catch them. Only quite recently, in 1990s, the technological progress made such observations possible. Two groups of observers, the Supernova Cosmology Project and the High-z Supernova Search Team, were able to observe many of type Ia supernovae with redshift up to $z = 1$. These results have allowed to measure the deceleration parameter $q_0$. Let us see how this was done.

For simplicity, consider only the model of flat Universe (This case is particularly relevant because the observations of the CMB fluctuations already show that that $\Omega_0$ is very close to unity.). In this model, the propagation of light from the source to the observer is described by

$$\frac{d\chi}{dt} = -\frac{c}{R(t)}$$ \quad (5.35)

(see Sec.5.1). Now, instead of substituting $R(t)$ from the Friedmann solution for a flat Universe, we simply use the generic expression for $R(t)$ in terms of $H_0$ and $q_0$

$$R(t) = R_0 \left[1 + H_0 (t-t_0) - \frac{q_0}{2} H_0^2 (t-t_0)^2 \right]. \quad (5.36)$$

Then Eq.5.35 becomes

$$\frac{d\chi}{d\tau} = -\frac{c}{R_0} \left(1 + H_0 \tau - \frac{q_0}{2} H_0^2 \tau^2 \right)^{-1}. \quad (5.37)$$

where $\tau = t-t_0$. Integrating this equation, with the initial condition $\chi(0) = 0$ gives us the comoving radius of the observed source as

$$\chi_e = \frac{c}{H_0 R_0} F(H_0 \tau_e, q_0).$$

We do not specify the function $F$ as this is not crucial in our discussion. Substituting this result in Eq.5.32 for the luminosity distance we find

$$r_L = \frac{c (1+z)}{H_0} F(H_0 \tau_e, q_0). \quad (5.38)$$
Using the generic expression for $R(t)$ in the equation of the cosmological redshift, we also find that

$$(1+z)^{-1} = \frac{R_e}{R_0} = 1 + H_0 \tau_e - \frac{q_0}{2} H_0^2 \tau_e^2. \quad (5.39)$$

Solving this quadratic equation for $H_0 \tau_e$, we find

$$H_0 \tau_e = G(z, q_0), \quad (5.40)$$

where we do not specify the function $G$, and thus find that

$$r_L = r_L(z, H_0, q_0) = \frac{c(1+z)}{H_0} F(G(z, q_0), q_0). \quad (5.41)$$

For any single supernova astronomers can measure both $r_L$ and $z$. When two supernovae have been observed one can determine both $H_0$ and $q_0$ as there are now two equations 5.41, one per each supernovae. However, the result will be affected by various observational errors. In order to increase accuracy, it is best to have a large sample of supernovae and to determine $q_0$ via statistical fitting of the observational data with the law Eq.5.41.

The results of the supernova projects were very surprising indeed. They clearly indicated a negative value for the deceleration parameter, $q_0 \simeq -0.6$, showing that Universe was not decelerating but accelerating! Since none of the Friedmann models allows such a result, the Cosmological constant had to be reintroduced in modern Cosmology.
Chapter 6

Secret components of the Universe

As this has been noted in Sec.5.3, the observed fluctuations of the Cosmic microwave background radiation imply that the density parameter is very close to unity

$$\Omega_0 \simeq 1,$$

and the question arises if this result is consistent with the astronomical observations. Indeed, these observations allows us to measure the local density of the Universe directly.

6.1 Dark Matter

In galaxies, most of the visible mass is present in the form of stars. Many researchers looked into the issue of stellar contribution to the density parameter. They converged to the conclusion that it is very small, $$\Omega_{0,\text{stars}} = \rho_{\text{stars}}/\rho_c \simeq 0.01.$$ Spiral galaxies also have a substantial mass in the form of interstellar gas but this is still smaller compared to their stellar mass. A substantial amount of gas is also present in clusters of galaxies, the intergalactic gas, where its total mass can even exceed the stellar mass but not by much. In total, the visible matter in the Universe can only yield

$$\Omega_{0,\text{vm}} = 0.05.$$ (6.1)

If there is a massive invisible material component, the so-called dark matter, in the Universe then its presence will be felt by visible matter via the gravitational field of the dark matter. In particular the characteristic speeds of stars in galaxies, and galaxies in clusters of galaxies will be higher than expected. The gravity of this dark matter will also effect the quasi-hydrostatic equilibrium of the gaseous components. This is indeed what is observed.

In particular, the rotation of spiral galaxies is too fast to be explained by the gravity of the visible component alone. In fact, the motion of individual stars in galaxies is very accurately described by the Newtonian theory. According to this theory the centrifugal acceleration is provided by the gravitational pull so that

$$\frac{v^2}{r} = \frac{GM(r)}{r^2},$$ (6.2)

where $$v$$ is the orbital velocity, $$r$$ is the distance from the center of the galaxy, and $$M(r)$$ is the total mass contained within sphere of radius $$r$$ (see Sec.8.1). Thus by measuring the orbital motion of its stars, one can determine the distribution of mass in the galaxy. Studies conducted along these lines have concluded that galaxies must contain invisible matter in the form of extended and more or less spherical halos (see Fig.6.1). Moreover, its mass exceeds that of visible matter by a factor ranging from five to ten. Yet, the density parameter due to the contribution of dark galactic halos is still much lower then unity,

$$\Omega_{0,\text{halo}} \simeq 0.1.$$ (6.1)

Similar studies, where individual galaxies played the role of test particles, have deduced the presence of massive invisible component on the scale of galactic clusters. In fact, there are other methods that can be
used in the case of galactic clusters. For example, galactic clusters contain a large amount of intergalactic gas. This gas is very hot and is a strong source of thermal emission in the X-ray range of the electromagnetic spectrum. It is in the state of hydrostatic equilibrium, which is governed by the combined gravitational force of all gravitating components, including the dark matter.

The strength of gravitational field in galactic clusters can also be tested by bending of the trajectories of photons from distant sources of light, which pass through or near the cluster (see Fig. 2.6). The gravitational field works as a lens, magnifying, amplifying and distorting images of background sources. The arc-like features oriented perpendicular to the rays emanating from the center of the cluster shown in Fig. 1.3 are the images of background galaxies distorted by the cluster’s gravitational lens. All these observations are consistent with

$$\Omega_{0,\text{cluster}} \simeq 0.2 - 0.3.$$ 

A number of explanations have been proposed for the dark matter, from radiatively inefficient Jupiter-like objects and black holes to exotic types of elementary particles which do not produce electromagnetic radiation. The latter seems the most likely and the most exciting explanation. It is proposed that these particles are created during the extreme conditions of the Big Bang when they strongly interacted with each other and other particles. At present, however, they interact only by means of the gravitational force and for this reason they are invisible. In contrast, barons still interact electromagnetically with production of photons as a by-product. The preferred type of dark matter is the so-called cold dark matter, where the the rest mass-energy dominates over the thermal mass-energy density, in contrast to the case of hot dark matter, where this is the other way around.

In any case, the total contribution of matter, both visible and dark, to the critical parameter,

$$\Omega_{0,\text{m}} = \Omega_{0,\text{vm}} + \Omega_{0,\text{dm}} \simeq 0.27,$$ 

is still well below unity. Thus, more mass must be present in some other form in order to agree with $\Omega_0 \simeq 1$, which is deduced from the observations of the microwave background fluctuations.

### 6.2 Dark Energy

As this has already been mentioned in Sec. 5.4 the observations of type Ia supernovae cannot be explained with cosmological models without the cosmological constant. For this reason we return to the Friedmann equations with the cosmological constant retained
\[
\left( \frac{\ddot{R}}{R} \right)^2 + \frac{k c^2}{R^2} = \frac{8 \pi}{3} G \rho + \frac{\Lambda}{3}, \tag{6.4}
\]
and
\[
\frac{\ddot{R}}{R} = - \frac{4 \pi G}{3} \left( \rho + 3 \frac{p}{c^2} \right) + \frac{\Lambda}{3}, \tag{6.5}
\]
and analyses them in some details. From the last equation one see straight away that the Universe can be accelerating only if \(\Lambda > 0\).

As a matter of fact, one may describe the cosmological constant as a signature of fluid with rather peculiar properties. To see this we recast the terms involving \(\Lambda\) in Eqs.6.4 and 6.5 as contributions due to fluid with mass-energy density \(\rho_\Lambda\) and pressure \(p_\Lambda\). Start with Eq.6.4. Obviously, we should have
\[
\frac{8 \pi}{3} G \rho_\Lambda = \frac{\Lambda}{3}.
\]
This yields
\[
\rho_\Lambda = \frac{\Lambda}{8 \pi G}. \tag{6.6}
\]
Notice that, in spite of the expansion of the Universe, \(\rho_\Lambda\) is constant. This is the first peculiar property of the fluid. Now deal with Eq.6.5. It suggests that
\[
- \frac{4 \pi G}{3} \left( \rho_\Lambda + 3 \frac{p_\Lambda}{c^2} \right) = \frac{\Lambda}{3}.
\]
Substituting the expression for \(\rho_\Lambda\) from Eq.6.6 we then find
\[
P_\Lambda = - \frac{\Lambda c^2}{8 \pi G} = - \rho_\Lambda c^2. \tag{6.7}
\]
This result is consistent with the fluid equation (see Eq.3.20)
\[
\dot{\rho}_\Lambda = -3 \left( \frac{\ddot{R}}{R} \right) \left( \rho_\Lambda + 3 \frac{p_\Lambda}{c^2} \right). \tag{6.8}
\]
Indeed, since \(\dot{\rho}_\Lambda = 0\) the fluid equation immediately yields
\[
P_\Lambda = - \rho_\Lambda c^2. \tag{6.9}
\]
This equation can be considered as the equation of state of the cosmological fluid. Provided that \(\rho_\Lambda > 0\) it requires \(P_\Lambda < 0\). This is another peculiar property of the cosmological fluid.

More specifically, the cosmological fluid is sometimes thought as the energy associated with vacuum, the space absolutely void of particles. This could be a kind of “ground level energy”, which often appears in quantum physics. In fact, some theories of elementary particles predict cosmological constant but, unfortunately, of much higher value compared to the observations. Given this association, the term dark energy is often used to describe the origin of the cosmological constant.

By analogy with the density parameter of matter
\[
\Omega_m = \frac{\rho_m}{\rho_c}, \tag{6.10}
\]
where \(\rho_m\) is the total mass density of visible and dark matter, and
\[
\rho_c = \frac{3 H^2}{8 \pi G} \tag{6.11}
\]
is the critical density (see Sec.4.2), one can introduce the density parameter associated with the dark energy
\[
\Omega_\Lambda = \frac{\rho_\Lambda}{\rho_c} = \frac{\Lambda}{3 H^2}. \tag{6.12}
\]
Notice, that here we do not use the suffix "0", so that these equations can not only be applied to the current Universe but also to its past and future states. Then Eq.6.4 can be written as

\[ 1 + \frac{k c^2}{H^2 R^2} = \Omega, \]  

(6.13)

where

\[ \Omega = \Omega_m + \Omega_\Lambda. \]  

(6.14)

Thus, the dark energy can help to solve the missing mass problem. It allows to explain the observation of CMB fluctuations, which require \( k = 0 \), by increasing the current value of \( \Omega \) towards unity.

Figure 6.2: Constraints on the density parameters from the supernova searches and the studies of cosmic microwave background.

At present, the temperature of matter in the Universe is very low and its pressure can be ignored. If the dark matter is also cold then its pressure can be ignored as well and the acceleration equation becomes

\[ \frac{\ddot{R}}{R} = -\frac{4\pi G}{3} \rho + \Lambda = \frac{8\pi G}{3}(\rho_\Lambda - \frac{1}{2}\rho). \]  

(6.15)

Substituting this into the expression for the deceleration parameter one finds

\[ q_0 = -\frac{1}{H_0^2} \frac{\ddot{R}_0}{R_0} = -\frac{8\pi G}{3H_0^2}(\rho_\Lambda - \frac{1}{2}\rho_0), \]  

(6.16)

and thus

\[ q_0 = -\frac{1}{2} \Omega_{m,0} - \Omega_{\Lambda,0}, \]  

(6.17)

where, as usual, suffix 0 refers to quantities at present time, \( t = t_0 \). Combining this equation with

\[ \Omega_0 = \Omega_{m,0} + \Omega_{\Lambda,0}. \]  

(6.18)
we find that
\[ \Omega_{\Lambda,0} = \frac{1}{3} \Omega_0 - \frac{2}{3} q_0 \] (6.19)
and
\[ \Omega_{m,0} = \frac{2}{3} \Omega_0 + \frac{2}{3} q_0 . \] (6.20)
Thus, the observational data on \( q_0 \) and \( \Omega_0 \) can be used to find the values of \( \Omega_{m,0} \) and \( \Omega_{\Lambda,0} \). The current data show that
\[ \Omega_{m,0} \approx 0.27 \quad \text{and} \quad \Omega_{\Lambda,0} \approx 0.73 \]
(see Figure 6.2). This value of \( \Omega_{m,0} \) agrees very well with the astronomical observations of the local Universe.
Chapter 7

The phase of inflation

7.1 The problems of Friedmann Cosmology

There are a number of puzzling issues in the Friedmann Cosmology. Here we discuss only two of them, the flatness problem and the horizon problem. Other issues are a bit more complex and require a very good background in physics.

7.1.1 The flatness problem

According to the cosmological observations, the critical parameter is very close to unity. Why is this? Unless the Universe is exactly flat (\( k = 0 \)), we should have \( \Omega \neq 1 \). What was the value of \( \Omega \) in the past? Could it have evolved naturally towards unity when the Universe was very young? If it could then this observational result would have a simple explanation. If not, then this becomes a mystery.

From Eq.6.13 we have

\[ |1 - \Omega| = \frac{|k|c^2}{H^2R^2}. \]  

(7.1)

Thus, the evolution of \( \Omega \) is determined by the evolution of \( HR \).

For a young Universe, both the curvature term and the cosmological terms in the Friedmann equations can be ignored. Indeed, \( \Lambda \) is constant, whereas \( \rho \) evolves as \( R^{-3} \) in a matter-dominated Universe, and as \( R^{-4} \) in a radiation dominated Universe. Thus, for small \( R \) the Friedmann equation

\[ \left( \frac{\dot{R}}{R} \right)^2 + \frac{k c^2}{R^2} = \frac{8\pi}{3} G \rho + \frac{\Lambda}{3} \]  

(7.2)

reduces to

\[ \left( \frac{\dot{R}}{R} \right)^2 + \frac{k c^2}{R^2} = \frac{8\pi}{3} G \rho. \]  

(7.3)

The curvature term grows only as \( R^{-2} \) when \( R \) decreases and therefore for sufficiently small \( R \) it also becomes small compared to the matter/radiation term (the term on the right-hand side of this equation). Thus, the Friedmann equation for a young Universe is very simple

\[ \left( \frac{\dot{R}}{R} \right)^2 = \frac{8\pi}{3} G \rho. \]  

(7.4)

Finally, since \( \rho_{\text{radiation}} / \rho_{\text{matter}} \propto R^{-1} \), we may focus only on the case of radiation-dominated Universe, for which this equation gives the solution

\[ R \propto t^{1/2} \quad \text{and} \quad HR \propto t^{-1/2}. \]  

(7.5)

Substituting this result into Eq.7.1 we find that

\[ |1 - \Omega| \propto t. \]  

(7.6)
Thus, $\Omega$ does not evolve naturally towards unity as the Universe expands. On the contrary, it moves away from it and in order to have $\Omega_0 \simeq 1$ at present, $\Omega$ has to be extremely close to unity in the past. This is the flatness problem of Big Bang Cosmology.

### 7.1.2 The horizon problem

Why is the Universe uniform? Is there a natural explanation to this, hinted by the Friedmann solutions? Well, the uniformity could imply an efficient mechanism of erasing inhomogeneities in the early Universe. However, no matter what actual physical processes are involved in this, in order for them to work an effective communication between different parts of the now observable universe must have existed in the past.

In order to see this, consider a box, which has been filled with gas through a number of holes in its walls. Given time, the gas distribution in the box becomes uniform. It is the gas pressure which becomes uniform first and this process involves sound waves crossing the box from one end to another, back and forward. For the density to become uniform even more time is required because this involves slower processes, like diffusion and thermal conductivity. Thus, the time, which is required to reach a uniform state in the box is larger than the sound wave crossing time of the box. Since no wave can propagate faster than light, the light crossing time sets the lowest possible limit on the time of settling to a uniform state.

At present, we can observe the Universe up to the CMB sphere, beyond which the Universe is not transparent to the electromagnetic radiation (light). The CMB photons, which are reaching the Earth now, were produced at this sphere at the time of decoupling of matter and radiation. With time, the comoving radius of the CMB sphere increases, and larger part of the Universe becomes accessible to observational studies. The size of causally connected regions of the Universe is given by the radius of the cosmological horizon (see Sec.5.2). The fact that there are no significant variations of the CMB emission in the sky means that the comoving radius of the horizon at the time of decoupling was much larger compared to the current comoving radius of the CMB sphere.

One can formulate this condition in terms of metric radii rather than the comoving ones, which is in fact a bit more convenient. For this we simply need to convert the comoving radii into the metric ones using the same value of the scaling factor. For example, we may use the scaling factor at the time of decoupling. This will give us the metric radius of the horizon at the time of decoupling and the metric radius which the current CMB sphere had at that time. The isotropy of CMB requires the horizon radius to be much larger. It turns out that this is in conflict with the Friedmann models.

The comoving radius of the cosmological horizon $\chi_h(t)$ is given by the equation

$$
\int_0^t \frac{c dt}{R(t)} = \frac{\chi_h(t)}{\sqrt{1-k\chi^2}}
$$

(see Sec.5.2). The corresponding metric radius at this time is

$$
r_h(t) = R(t) \int_0^{\chi_h(t)} \frac{d\chi}{\sqrt{1-k\chi^2}}
$$

Thus, we have

$$
\int_0^t \frac{c dt}{R(t)} = \frac{r_h(t)}{R(t)}.
$$

Both the matter-dominated and the radiation-dominated solutions of the Friedmann equations give

$$
R(t) \simeq R_0 \left(\frac{t}{t_0}\right)^{2/3},
$$

with $1/2 < \alpha < 2/3$. This is because $\alpha = 2/3$ for the matter-dominated model and $\alpha = 1/2$ for the radiation-dominated model. In this case

$$
H(t) = \frac{\dot{R}}{R} = \frac{\alpha}{t}.
$$
Substituting $R(t)$ from Eq.7.10 into the left-hand side of Eq.7.9 and integrating, one finds that
\[ r_h(t) = \frac{ct}{1-\alpha} = \left( \frac{\alpha}{\alpha-1} \right) \frac{c}{H(t)} \simeq \frac{c}{H(t)}. \] (7.12)
Thus, at the time of decoupling
\[ r_h(t_{\text{dec}}) \simeq \frac{c}{H(t_{\text{dec}})}. \] (7.13)
The current comoving radius of the CMB sphere, $\chi_{\text{CMB}}$, is given by
\[ \int_{t_{\text{dec}}}^{t_0} \frac{c dt}{R(t)} = \int_{0}^{\chi_{\text{CMB}}} \frac{d\chi}{\sqrt{1-k\chi^2}}. \] (7.14)
The corresponding metric radius at the time of decoupling is
\[ r_{\text{CMB}}(t_{\text{dec}}) = R(t_{\text{dec}}) \int_{0}^{\chi_{\text{CMB}}} \frac{d\chi}{\sqrt{1-k\chi^2}}. \] (7.15)
Thus, we have
\[ \int_{t_{\text{dec}}}^{t_0} \frac{c dt}{R(t)} = \frac{r_{\text{CMB}}(t_{\text{dec}})}{R(t_{\text{dec}})}. \] (7.16)
Substituting into this equation $R(t)$ from Eq.7.10 and integrating, we obtain
\[ r_{\text{CMB}}(t_{\text{dec}}) = \frac{\alpha}{\alpha-1} \frac{R(t_{\text{dec}})}{R_0} \frac{c}{H_0} \simeq \frac{R(t_{\text{dec}})}{R_0} \frac{c}{H_0}. \] (7.17)
Thus,
\[ \frac{r_h}{r_{\text{CMB}}(t_{\text{dec}})} \simeq \frac{R_0}{R(t_{\text{dec}})} \frac{H_0}{H(t_{\text{dec}})}. \] (7.18)
Using Eqs.7.10 and 7.11
\[ \frac{H_0}{H(t_{\text{dec}})} = \frac{t_{\text{dec}}}{t_0} = \left( \frac{R(t_{\text{dec}})}{R_0} \right)^{1/\alpha} \]
and we finally obtain
\[ \frac{r_h}{r_{\text{CMB}}(t_{\text{dec}})} \simeq \left( \frac{R(t_{\text{dec}})}{R_0} \right)^{1/\alpha-1} \ll 1 \] (7.19)
because $R(t_{\text{dec}}) \ll R_0$ (see Sec.4.5). Thus, the observed CMB radiation is produced by causally disconnected parts of the Universe! This means that the observed uniformity of the Universe cannot be established naturally in the Friedmann cosmology! This result is known as the horizon problem.

### 7.2 Inflation

According to the Friedmann equation,
\[ \left( \frac{\dot{R}}{R} \right)^2 + k \frac{c^2}{R^2} = \frac{8\pi}{3} G \rho + \frac{\Lambda}{3}, \] (7.20)
the dynamics of the Universe will eventually be dominated by Cosmological constant. Indeed, both the curvature and the matter/radiation terms decrease with $R$, whereas the cosmological constant term remains constant. In this limit
\[ \left( \frac{\dot{R}}{R} \right)^2 = \frac{\Lambda}{3}, \] (7.21)
which yields the solution

\[ R = A \exp \left( \sqrt{\frac{\Lambda}{3}} t \right), \quad (7.22) \]

where \( A \) is the constant of integration. Thus, the Universe expands exponentially.

What is the evolution of the critical parameter during this expansion? The Hubble constant

\[ H = \frac{\dot{R}(t)}{R(t)} = \sqrt{\frac{\Lambda}{3}} \quad (7.23) \]

is constant, and

\[ HR \propto \exp \left( \sqrt{\frac{\Lambda}{3}} t \right). \]

This implies that

\[ |1 - \Omega| = \frac{|k|c^2}{H^2 R^2} \propto \exp \left( -2 \sqrt{\frac{\Lambda}{3}} t \right). \quad (7.24) \]

Thus, \( \Omega \) is driven towards unity. This analysis suggests that the flatness problem of the Friedmann Cosmology can be solved if soon after the Big Bang instead of the usual matter and radiation the Universe was filled with a “fluid” similar to the one associated with the Cosmological constant (see Sec.6.2). The corresponding phase of exponential expansion is called the inflation. At the end of this phase, a some kind of phase transition takes place, where this hypothetical exotic fluid is transformed into the usual mixture of radiation and matter.

What is about the horizon problem? First, we notice that from Eq.7.22 \( R(t) \to 0 \) only as \( t \to -\infty \), implying infinitely long phase of inflation. If the lifetime of the Universe is in fact finite, then the phase of inflation had a beginning, where it replaced some other previous phase in the evolution of the Universe. Without loss of generality we may put \( t = 0 \) at the start of the inflation. Assuming that the horizon radius at this point was small, \( \chi_h(0) \approx 0 \), we may compute its further evolution via

\[ \int_0^t \frac{c dt}{R} = \int_0^{\chi_h} \frac{d\chi}{\sqrt{1 - k\chi^2}} = \frac{r_h}{R} \quad (7.25) \]

as before (see Eq.7.9). Substituting into the left-hand side of this equation \( R(t) \) from Eq.7.22 and integrating we obtain

\[ r_h = R \frac{c}{A} \sqrt{\frac{3}{\Lambda}} \left[ 1 - \exp \left( -\sqrt{\frac{\Lambda}{3}} t \right) \right]. \quad (7.26) \]

For \( t \gg \sqrt{3/\Lambda} \) this gives us

\[ r_h = R \frac{c}{A} \sqrt{\frac{3}{\Lambda}} \approx \frac{c}{H} \exp \left( \sqrt{\frac{\Lambda}{3}} t \right) \gg \frac{c}{H}. \quad (7.27) \]

Comparing this result with Eq.7.12 we can see that the radius of the cosmological horizon at the end of inflation can easily exceed that of standard Friedmann cosmology by many orders of magnitude and hence can easily exceed the radius of the CMB sphere at the time of decoupling. This is simply because the Universe expands very slowly during the inflation, whereas the cosmological horizon keeps expanding with the speed of light as before. As the result it encloses much larger volume of the Universe.

The theory of inflation successfully resolves many other problems of the Friedmann Cosmology, which we can not discuss within the limited scope of this module. In fact it remains a very hot topical issue in modern Cosmology.
Chapter 8

Supplemental material: Newtonian Gravity and Cosmology

It turns out that a number of key results can also be derived in the framework of Newton’s theory of gravity. A little bit of cheating is required but the mathematics is much simpler. The material of this chapter will not be examinable.

8.1 Newton’s Universal Gravity

8.1.1 The Newton’s law of gravity

Consider two particles, one with the inertial mass \( m \) and the other with the inertial mass \( m_1 \), separated by the distance \( r_1 \).

Then the gravitational force acting on the particle \( m \) is

\[
\mathbf{f}_1 = -\frac{Gm_1m}{r_1^2}\hat{r}_1,
\]

(8.1)

where \( \hat{r}_1 \) is the unit vector in the direction from \( m_1 \) to \( m \) (see Fig. 8.1). From the second law on Newtonian mechanics

\[
\mathbf{f}_1 = ma,
\]

(8.2)

where \( a \) is the acceleration of particle \( m \). Combining this with Eq. 8.1 one finds

\[
a = -\frac{Gm_1}{r_1^2}\hat{r}_1.
\]

(8.3)

(Notice that the acceleration does not depend on the mass of particle \( m \) – this fact was one of the Einstein’s motivations for developing General Relativity). It is easy to see that
\[ \frac{1}{r_1^2} \hat{r}_1 = -\nabla \left( \frac{1}{r_1} \right) \]  

(8.4)

and this allows us to rewrite Eq.8.3 as

\[ a = -\nabla \psi_1, \]  

(8.5)

where

\[ \psi_1 = -\frac{G m_1}{r_1} \]  

(8.6)

is called the potential of the gravitational field created by the particle \( m_1 \).

Next consider the case where the particle \( m \) interacts with more than one particle (see Fig.8.2).

Figure 8.2:

In this case the total force acting on \( m \) is simply the vector sum of all individual forces due to other particles as described by Eq.8.1. That is

\[ \mathbf{f}_{\text{tot}} = \sum_{i=1}^{n} \mathbf{f}_i = -m \sum_{i=1}^{n} \nabla \psi_i, \]  

(8.7)

where \( n \) is the number of particles and \( \psi_i = -\frac{G m_i}{r_i} \). This is known as the principle of superposition. In terms of the gravitational potential, this result reads

\[ \mathbf{f}_{\text{tot}} = -m \nabla \psi_{\text{tot}}, \quad \text{where} \quad \psi_{\text{tot}} = \sum_{i=1}^{n} \psi_i = -\sum_{i=1}^{n} \frac{G m_i}{r_i}. \]  

(8.8)

Thus, the potential of the total gravitational field created by these particles is the sum of the individual potentials.

If instead of a set of discrete particles we are dealing with a continuous distribution of matter with mass density \( \rho \) then the sum in Eq.8.8 should be replaced with integration over volume and the mass of individual particle with the mass of infinitesimal volume, \( dV \). That is

\[ \mathbf{f} = -m \nabla \psi, \quad \text{where} \quad \psi = -G \int_{V} \frac{\rho dV}{r}, \]  

(8.9)

where we drop suffix “tot” for brevity.

Now let us multiply both sides of the first equation in Eq.8.9 with \( \mathbf{v} = \frac{d\mathbf{r}}{dt} \), the velocity of particle \( m \), using the scalar product. For the left hand side we have

\[ (\mathbf{f} \cdot \mathbf{v}) = m \frac{d}{dt} \left( \frac{1}{2} m v^2 \right) = \frac{1}{2} \frac{d}{dt} \left( \mathbf{v} \cdot \mathbf{v} \right) = \frac{d}{dt} \frac{m v^2}{2}, \]
whereas for the right hand side we have

$$-m(v \cdot \nabla \psi) = -m(\frac{dr}{dt} \cdot \nabla \psi) = -m \frac{d\Psi}{dt} = -\frac{dm\Psi}{dt}.$$  

Thus, we obtain

$$\frac{d}{dt}(mv^2 + m\Psi) = 0,$$

or

$$E = \frac{mv^2}{2} + m\Psi = \text{const.}$$

The quantity

$$T = \frac{mv^2}{2}$$

is called the kinetic energy of particle $m$, the quantity

$$U = m\Psi$$

is called its gravitational potential energy, and the quantity

$$E = T + U$$

is its total energy. Thus, the result Eq.8.11 constitutes the energy conservation law.

### 8.1.2 Gravitational field of a spherically symmetric mass distribution

Here we derive the very important result which has direct application to the evolution of the Universe. We start with the problem of the gravitational field of a spherical shell with uniform surface mass density. It turns out that its potential is constant inside the shell, implying zero gravity force, whereas outside of the shell it is the same as the potential of a point-like particle with mass equal to that of the shell and located in its center.

![Diagram](image)

Figure 8.3: Left panel: The test particle is outside of the gravitating spherical shell. Here the distance $s$ varies between $r-a$ and $r+a$. Right panel: The test particle is inside the sphere. Here the distance $s$ varies between $a-r$ and $r+a$.

Let us show this. Denote the radius of the shell as $a$ and the distance from its center to the point where the gravitational potential is calculated as $r$. Because of the spherical symmetry the potential is a function of $r$ only. Suppose that $da$ is the infinitesimal thickness of the shell, the its potential is also infinitesimal and is given by the integral over the surface of the sphere,

$$d\Psi = -daG \int \frac{\rho dA}{s},$$
where \( s \) is the distance of the infinitesimal volume of the shell \( dV = da dA \) from the calculation point (see Fig.8.3). In terms of the polar angle, \( \theta \), and the azimuthal angle, \( \phi \), of the spherical coordinates with the origin at the center of the shell (see Fig.8.4)

\[
dA = (a \sin \theta d\phi)(a d\theta) = a^2 d\phi \sin \theta d\theta.
\]

From the triangle with sides \( a, r, \) and \( s \) (see Fig.8.3) we have

\[
s^2 = a^2 + r^2 - 2ar \cos \theta.
\]

This shows that when \( \theta \) varies

\[
2sd\theta = 2ar \sin \theta d\theta \quad \text{and thus} \quad \sin \theta d\theta = \frac{sd\theta}{ra}.
\]

Substituting this result in the above expression for \( dA \) we find that

\[
dA = \frac{as \phi ds}{r}.
\]

Thus, instead of integrating over \( \theta \) we can integrate over \( s \). However, the range of \( s \) depends on where \( r \) is larger or less than \( a \). (see Fig.8.3). First consider the case where \( r > a \). In this case \( s \in [r-a, r+a] \), and we have

\[
d\Psi = -daG \int_0^{2\pi} d\phi \int_{r-a}^{r+a} \frac{\rho a}{r} ds = -daG2\pi \frac{\rho a}{r} \int_{r-a}^{r+a} ds = -G\frac{4\pi \rho a^2 da}{r} = -G\frac{dM}{r},
\]

where \( dM = 4\pi \rho a^2 da \) is the total mass of the shell. The potential is the same as that of the point-like particle of mass \( dM \) located at the shell center. Next, consider the case where \( r < a \) (see the right panel of Fig.8.3). Now \( s \in [a-r, r+a] \) and repeating the above calculations we obtain.

\[
d\Psi = -daG2\pi \frac{\rho a}{r} \int_{a-r}^{r+a} ds = -G4\pi \rho ada = -G\frac{dM}{a} = \text{const}.
\]

Summarizing,

\[
d\Psi = \begin{cases} 
-GdM/a & \text{if } r \leq a, \\
-GdM/r & \text{if } r \geq a.
\end{cases} \tag{8.12}
\]
Figure 8.5: The gravitational potential of a uniform spherical shell of the mass $dM$ and radius $a$.

Since, the gravity force is determined by the gradient of the potential, this result shows that any particle placed inside the shell will not feel the shell’s gravity. A particle of mass $m$ placed outside of the shell will experience the force

$$df = m \frac{d^2r}{dt^2} = -G \frac{mdM}{r^2}. \quad (8.13)$$

Now imagine a particle between two such shells. It will feel only the presence of the inner shell.

Next consider a spherically symmetrical distribution of mass and a particle at the distance $r$ from its center. All the mass located further out will give no contribution to the gravity force experienced by the particle, only the mass located within the radius $r$ will. To find this force one has to integrate Eq.8.13

$$f = \int_{a<r} df = -G \frac{m}{r^2} \int_{a<r} dM = -G \frac{mM(r)}{r^2}, \quad (8.14)$$

where $M(r)$ is the total mass inside the radius $r$.

### 8.2 Derivation of Friedmann’s equations from Newton’s Theory of Gravity

A limited progress in Cosmology can also be made with Newtonian theory of gravity. However, the Newtonian Cosmology is plagued with problems. The following analysis is due to McCrea and Milne (1934).

#### 8.2.1 The Hubble law

First, we show that the Hubble law is a natural consequence of the homogeneity and isotropy of the Universe. Select a particular inertial frame coasting through the Euclidean space. One particular fluid element will be at rest in this frame. Select it as the origin of spatial grid. Select a particular radial direction. Due to the isotropy of the Universe the motion of the Cosmological fluid will be radial. At any particular time the fluid speed $v = f(r)$ is a function of the radial coordinate $r$ only. Then $f(dr)$ is the relative speed of two fluid elements separated by the distance $dr$. Moreover, $f(r+dr) - f(r)$ is the relative speed of two other fluid elements also separated by the distance $dr$. Since, the Universe is homogeneous both speeds must be the same,

$$f(r+dr) - f(r) = f(dr).$$

From this we see that

$$\frac{df}{dr}(r) = \frac{df}{dr}(0) = \text{const},$$
and hence \( f(r) \) is a linear function. Denoting the constant as \( H_0 \) and noticing that \( f(0)=0 \), we obtain the Hubble law

\[
v = H_0 r.
\]

Notice, that this argument allows \( H_0 \) to be a function of time. Integrating Eq.8.15 we find that

\[
r = S(t) r_0,
\]

where

\[
S(t) = \exp\left(\int _{t_0} ^t H_0 (t) dt\right).
\]

and \( r_0 = r(t_0) \). Thus, the distance between any two fluid elements grows in time as \( S(t) \). \( r_0 \) is an example of so-called Lagrangian coordinates of fluid elements, coordinates which do not change in time. \( S(t) \) is simply the scaling factor which relates this coordinate with the coordinate \( r \).

From the last equation it follows that

\[
H_0 = \frac{\dot{S}}{S}.
\]

### 8.2.2 The Friedmann equation

Since the Universe is uniform its mass distribution is obviously spherically symmetric with respect to any point, which can be considered as a center of symmetry. So if we take a fluid element of mass \( \delta m \) at the distance \( r \) from center, then it will feel only the gravity force due to the mass \( M = \frac{4\pi}{3} r^3 \rho \) inside the sphere of radius \( r \) (see Sec.8.1.2). Thus,

\[
\delta m \frac{d^2 r}{dt^2} = -G \frac{\delta m M}{r^2} \quad \text{or} \quad \ddot{r} = -G \frac{M}{r^2},
\]

where use the familiar notation \( \dot{r} = dr/dt \), \( \ddot{r} = d^2 r/dt^2 \), etc. Since this particle is simply participating in the global expansion of the cosmological fluid, the mass \( M \) remains constant even if \( r \) increases with time.\(^1\) Thus, we are dealing with a relatively simple second order ordinary differential equation (ODE). Repeating the calculation leading to the energy conservation law Eq.8.11 we find

\[
\frac{\dot{r}^2}{2} = \frac{GM}{r} - C,
\]

where \( C \) is a constant. Let us analyze this result. Suppose that \( \dot{r} \) is positive, which corresponds to the currently observed expansion of the Universe. Will it stay positive all the time or at some point the expansion will turn into a contraction? The answer depends of the sign of the constant \( C \). Indeed, at the turning point \( \dot{r} = 0 \) and thus we need to see if the right hand side of Eq.8.19 can vanish. Obviously, it cannot if \( C < 0 \) – in this case the Universe will expand forever. It will keep expanding even if \( C = 0 \), simply because \( GM/r > 0 \). If, however, \( C > 0 \), no matter by how much, there will be a turning point. Indeed, as \( r \) increases \( GM/r \) tends to zero and the right hand side of Eq.8.19 will eventually vanish. At this point, the Universe will begin to contract. In fact, the speed corresponding to \( C = 0 \),

\[
v = \sqrt{\frac{2GM}{r}},
\]

is known in the Newton’s theory of gravity as the \textit{escape speed}. This is the minimum speed in the radial direction a projectile must have in order to escape from the gravitational field of mass \( M \), whose center is located at distance \( r \) from the projectile.

We can rewrite Eq.8.19 in a somewhat more convenient form if we substitute the expression for \( M \) and multiply both sides by \( 2/r^2 \). This gives

\[
\left( \frac{\dot{r}}{r} \right)^2 = \frac{8\pi}{3} G \rho - \frac{2C}{r^2}.
\]

\(^1\)In fact, as we shall see later, the inclusion of mass-energy associated with heat allows \( M \) to vary, so here we are cheating a bit.
Substituting into this equation the expression for \( r \) from Eq.8.16 we obtain

\[
\left( \frac{\dot{S}}{S} \right)^2 = \frac{8\pi}{3} G \rho - \frac{2C}{r_0^2 S^2}.
\]

This shows that Eq.8.21 is consistent with the uniform expansion if \( C \propto r_0^2 \) (only in this case we have the same equation for the scaling factor for any \( r_0 \)). So we may put \( C = \frac{r_0^2}{A/2} \), where \( A \) is the same constant for all \( r_0 \), and obtain

\[
\left( \frac{\dot{S}}{S} \right)^2 = \frac{8\pi}{3} G \rho - \frac{A}{S^2}.
\]

This equation already has the form similar to that of the Friedmann equation. It can take exactly the same form after a simple rescaling. Indeed, introduce new scaling factor \( R(t) \) via

\[
R = \begin{cases} 
cS & \text{if } A = 0 \\
cS/\sqrt{A} & \text{if } A > 0 \\
cS/\sqrt{-A} & \text{if } A < 0
\end{cases}
\]

Then the above equation becomes

\[
\left( \frac{\dot{R}}{R} \right)^2 = \frac{8\pi}{3} G \rho - \frac{k c^2}{R^2},
\]

(8.22)

where \( k \) is either 0, +1, or -1. This change of scaling requires to change the Lagrangian coordinate to

\[
\chi = \begin{cases} 
\frac{r_0}{c} & \text{if } A = 0 \\
(\frac{r_0}{c})\sqrt{A} & \text{if } A > 0 \\
(\frac{r_0}{c})\sqrt{-A} & \text{if } A < 0
\end{cases}
\]

so that

\[
r = R(t) \chi.
\]

One can see that the Lagrangian coordinate \( \chi \) can be interpreted as the radial coordinate at the time \( t_0 \) when \( R = 1 \).

It is easy to see that

\[
\left( \frac{\dot{R}}{R} \right) = H_0
\]

for all three possible values of \( k \).

8.2.3 The fluid equation

The Friedmann equation can be solved for \( R(t) \) provided we know \( \rho(t) \). Thus we need an additional equation to determine the evolution of the inertial, and hence the gravitational, mass density of the cosmological fluid. According to the first law of thermodynamics, when a fluid is expanding adiabatically

\[
dE = -PdV
\]

(8.23)

where \( P \) is the fluid pressure, \( V \) is its volume, and \( E \) is its internal energy. In fact this is the energy conservation law: as the fluid expands it carries out work \( PdV \) on its surrounding, at the expense of its internal energy. For example, consider gas in a cylinder containing a piston with area \( A \). The pressure force acting on the piston is \( pA \). When the piston moves by \( dx \), the work done by the gas is \( PAdx = PdV \).

Equation 8.23 can be written in terms of time derivatives as

\[
\dot{E} = -PV.
\]

(8.24)

This law can be applied to the sphere of expanding cosmological fluid. Its volume is

\[
V = \frac{4}{3} \pi r^3 = \frac{4}{3} \pi R^3 \chi^3.
\]
At this point we cheat a bit again and use the relativistic mass-energy equation:

\[ E = Mc^2 = \rho c^2 V = \frac{4}{3} \pi R^3 \chi^3 \rho c^2. \]

Substituting these expressions into Eq.8.24 we obtain

\[ \dot{\rho} = -3 \left( \frac{\dot{R}}{R} \right) \left( \rho + \frac{P}{c^2} \right), \quad (8.25) \]

(recall that \( \chi = \text{const} \) and hence \( \dot{\chi} = 0 \)) which is exactly the same as Eq.3.20.

### 8.2.4 The acceleration equation

In order to obtain the acceleration equation we first differentiate the Friedmann equation (Eq.8.22) with respect to time

\[ \frac{2}{R} \left( R \dot{R} - \dot{R}^2 \right) - \frac{2k c^2}{R^2} \frac{\dot{R}}{R} = \frac{8 \pi}{3} G \dot{\rho}. \]

Next we substitute in this equation the expression for \( 2k c^2 / R^2 \) from the Friedmann equation and obtain

\[ 2 \frac{\dot{R}}{R} \left( \frac{R}{R} - \frac{8 \pi}{3} G \rho \right) = \frac{8 \pi}{3} G \dot{\rho}. \]

Finally, we substitute the expression for \( \dot{\rho} \) from the fluid equation and simplify the result by cancelling the common factor \( \dot{R}/R \) and rearranging the terms. This gives us

\[ \frac{\ddot{R}}{R} = -\frac{4 \pi G}{3} \left( \rho + \frac{P}{c^2} \right), \quad (8.26) \]

which is the acceleration equation Eq.3.19.

### 8.2.5 Newton’s law of gravity with cosmological constant

In order to derive the Friedmann equation with the cosmological term we need to modify the Newton’s law of gravity (Eq.8.1) by including a term with the cosmological constant. The modified equation is

\[ f_{1i} = -\frac{G m_1 m}{r_1^2} \dot{r}_1 + \frac{1}{3} \Lambda m r_1 \ddot{r}_1, \quad (8.27) \]

Notice that for \( \Lambda > 0 \) the new term corresponds to a repulsion force. Via repeating the calculations of Sections 8.1 and 8.2.2 one can derive Eq.3.21.

### 8.2.6 The key inconsistencies of the Newtonian Cosmology

The argument that one can choose any point in space as a center of spherical symmetry and then apply the results obtained in Sec.8.1.2 is wrong as it leads to conflict with the assumption of homogeneous Universe. Indeed, all gravity force vectors point towards one point in space, the selected center of spherical symmetry. This makes this point very special, which in clear conflict with the assumption of homogeneous Universe. Only the everywhere vanishing gravity force is consistent with the homogeneity and isotropy of the Universe. In Newtonian cosmology, equation 8.18 makes sense only if the Universe has a center (and the Earth is located close to it).

Second, although the theory of relativity is utilized in the derivation of the fluid equation (see Sec.8.2.3), the Newtonian cosmology is inconsistent with the speed of light being the highest possible speed. Indeed, Eq.8.15 shows that \( v \to \infty \) when \( r \to \infty \). (Notice that because the space of Newtonian theory is Euclidean \( r \) varies from 0 to \( \infty \).)