Cosmology

Fall 2018

Syksy Räsänen

Preface

These are lecture notes for the course Cosmology at the University of Birzeit. They owe a heavy debt to lecture notes prepared by Hannu Kurki-Suonio, who has lectured a similar course at the University of Helsinki (as have I, following him). (Credit for some figures belongs to Elina Keihänen, Jussi Väliviita and Reijo Keskitalo.)

One problem in teaching cosmology is that some central aspects rely on rather advanced physics, such as quantum field theory in curved spacetime. Nevertheless, the main applications of these advanced concepts can be discussed in relatively simple terms, so first learning general relativity, quantum field theory and so on in detail is not necessary. Thus only standard bachelor level theoretical physics background (mechanics, special relativity, quantum mechanics, statistical physics) will be assumed in this course. The more advanced theories that cosmology relies on are reviewed as part of these notes to the required extent. This means that some results have to be accepted without proper derivation.

1 Introduction

1.1 Overview of modern cosmology

Cosmology is the study of the universe as a whole, its structure, origin and evolution. Cosmology is grounded on observations, many of them astronomical, and laws of physics measured on Earth. These lead naturally to the standard framework of modern cosmology, the hot big bang theory, according to which the universe expands and cools down.

As a science, cosmology has a rare, if not singular, restriction: we can observe only one universe. We cannot observationally make comparative or statistical studies among many universes (and there can even be difficulties with the concept of "many universes"). We also cannot move around our universe, but are restricted to (on cosmological scales) a single point in space and time. As a result, cosmology relies more on model-dependent interpretations than many other branches of physics.

Nevertheless, the last few decades have seen remarkable progress, as a significant body of observational data has become available with modern astronomical instruments. We now have a good observational handle of the overall history of the universe for all times between one second and the present time (the universe is today about 14 billion years old). Theoretically, we understand the evolution of the universe at all times between 10^{-11} s and a few billion years. Important questions remain, in particular about the nature of *dark matter*, *dark energy* and the processes of *inflation* and *baryogenesis*. In the first part of the course, we will consider dark energy and dark matter, and will look at inflation in the second part. Baryogenesis will be mentioned only in passing.

One historically important observation in support of the big bang theory is the *redshift* of distant galaxies. Their spectra are shifted towards longer wavelengths. The further out they are, the larger the shift. This implies that the galaxies are receding from us; their distance from us is increasing. According to general relativity, this is understood as the expansion of space itself (which is an aspect of the curvature of spacetime), not as motion of galaxies in space. As space expands, the wavelength of the light travelling through it is stretched.

The expansion appears to be uniform over large scales. While there are deviations of order unity in the expansion rate on small scales (for example, our galaxy does not expand), the average expansion rate on scales larger than galaxy clustering, 100 Mpc or so, is almost the same everywhere. In homogeneous and isotropic cosmological models, the expansion is simply described by a time-dependent *scale factor* $a(t)$. Starting from the observed present expansion rate, $H \equiv \dot{a}/a$, we can use general relativity to calculate $a(t)$ as a function of time, given our understanding of the physics of the matter content in the universe. (We will discuss this in detail later.) The result is that $a(t) \to 0$ and the density of the universe $\rho \to \infty$ about 14 billion years ago. At this *singularity*, time and space begin, and we can choose it as the origin of the time coordinate, $t = 0$. However, we do not expect general relativity, which governs the evolution of spacetime, nor the Standard Model of particle physics, which governs the behaviour of matter, to be applicable at extremely high energy densities. At the so called *Planck density*, $\rho_{\mathrm{Pl}} \sim 10^{97}$ kg/m^3, at the latest, quantum gravity effects are expected to be large. To describe the earliest times, the *Planck era*, we would need a theory of *quantum gravity*, which we do not have.[1] At present we don't know what happened at $t = 0$ or in its immediate vicinity at the Planck time 10^{-42} s. It seems likely that a major breakthrough in our understanding of physics is required before we can make any definite statements about that primordial era.

Thus, the big bang theory does not actually apply all the way to the beginning of time and a "big bang". Rather, it is a valid description of the history of the universe starting from some early time when the universe was very hot, very dense, and expanded rapidly. The universe was at early times filled with an almost homogeneous "soup" of particles which was in thermal equilibrium for a long time. We can therefore describe the state of the early universe with a small number of thermodynamic variables, and this makes the evolution of the universe calculable.

The fact that the scale factor tends to zero at early times does not imply that the universe would have started from a point. The part of the universe which we can observe today was indeed very small at very early times, possibly smaller than 1 mm in diameter at the earliest times that can be sensibly discussed within the big bang framework. And if the inflationary scenario is correct, it would have been much smaller than that before inflation. However, the universe extends beyond what can be observed today (beyond our *horizon*), and may even be infinite. In the current cosmological models, if the universe is infinite, it has always been infinite, except at the moment $t = 0$, when the size is not defined. (We do not know, theoretically or observationally, whether the universe is finite or infinite.)

As the universe expands, it is not expanding into some space "around" the universe. The universe by definition consists of all space: as the universe expands, the volume of space grows. (In the case of an infinite universe, it is meaningless to talk about change of the total volume. However, we can say that the volume of any sufficiently large finite portion of the universe grows in time. The disclaimer "sufficiently large" is needed because on small scales, space can collapse, as happens during structure formation.)

[1] *String theory* is a candidate for the theory of quantum gravity, and applying it to the early universe is an active area of research. However, we do not have a complete formulation of the theory, and it is not known whether string theory is correct, so its applications are necessarily rather speculative.

In order to describe behaviour at high temperatures and energies, we have to treat matter in terms of quantum fields rather than classical particles. (We will skip a lot of particle physics details.) The Standard Model of particle physics is based on the symmetry group $\mathrm{SU}(3)_C \otimes \mathrm{SU}(2)_W \otimes \mathrm{U}(1)_Y$. From the viewpoint of the Standard Model, we live today in a low-energy universe, where part of the symmetry of the theory is broken. The "natural" energy scale of the theory is reached when the temperature of the universe exceeds 100 GeV (about 10^{15} K), which was the case when the universe was younger than 10^{-11} s. Then the primordial soup of particles consisted of free massless fermions (quarks and leptons) and massless gauge bosons mediating the interactions (colour and electroweak) between these fermions. The standard model also includes a particle called the *Higgs boson*, which was discovered in 2012.[2]

The Higgs field is responsible for breaking the electroweak symmetry. In the electroweak (EW) transition the electroweak interaction splits into two separate interactions: 1) the weak interaction mediated by the massive gauge bosons $W^\pm$ and Z^0, and 2) the electromagnetic interaction mediated by the massless gauge boson γ, the photon.[3] Fermions and W and Z bosons get their masses in the EW transition (with the possible exception of the neutrinos, the origin of their masses is not clear). The mass is due to the interaction of the particle with the Higgs field. The EW transition took place when the universe cooled below the critical temperature $T_c \sim 100$ GeV of at $t \sim 10^{-11}$ s.

The order of the electroweak transition depends on the particle physics on those scales. In the Standard Model, the transition is a smooth crossover, not a phase transition. However, in extensions of the Standard Model, there can be a first order electroweak phase transition (so the two phases have different energy densities at the critical temperature), in which case it proceeds through the formation of bubbles of the new phase. The phase transition can then have interesting effects, and baryogenesis, the generation of the observed matter-antimatter asymmetry in baryons, may occur in that process. Observable amounts of gravitational waves may then also be generated.

Another transition (also a crossover, not a phase transition), the QCD transition, or the quark-hadron transition, took place at $t \sim 10^{-5}$ s. The critical temperature of the QCD transition is $T_c \sim 150$ MeV. Quarks, which had been free until this time, formed hadrons: baryons (which contain three quarks), e.g. the nucleons n and p, and mesons (which contain a quark and an antiquark), e.g. π and K. The matter filling the universe was converted from a quark-gluon plasma to a hadron gas. (All mesons and most baryons are unstable with rather short lifetimes, so only protons and neutrons remain at late times.)

For every type of particle there is a corresponding *antiparticle*, which has the same mass and quantum numbers as the particle, except for charges (like the electric charge or colour charge), which have the opposite sign. Particles that do not have any charges, such as photons, are their own antiparticles. At high temperatures,

[2]CERN announced the discovery of a new particle consistent with the Higgs boson on July 4, 2012. Many theoretical physicists felt confident already at the time that it is the Higgs particle, but it was not officially labeled the Higgs until March 14, 2013.

[3]More accurately, the interaction before the symmetry breaking had the two parts $\mathrm{SU}(2)_W \otimes \mathrm{U}(1)_Y$, and afterwards the broken weak interactions and the unbroken electromagnetic symmetry each have parts of both.

$T \gg m$, where m is the mass of the particle, particles and antiparticles are constantly created and annihilated in various reactions, and there is roughly the same number of particles and antiparticles. When $T \ll m$, particles and antiparticles may still annihilate each other (and decay, if they are unstable), but there is no more thermal production of particle-antiparticle pairs. As the universe cools, heavy particles and antiparticles therefore annihilate each other, so that their number density decreases rapidly. The annihilation reactions produce lighter particles and antiparticles. If the universe would originally have had an equal number of particles and antiparticles, only photons and neutrinos (of the known particles) would be left over today in any significant quantity. The presence of matter today indicates that in the early universe there must have been slightly more nucleons and electrons than antinucleons and positrons, and this excess was left over. The lightest known charged massive particle is the electron, so the last annihilation event was the electron-positron annihilation which took place when $T \sim m_e \sim 0.5$ MeV and $t \sim 1$ s.[4] After this the only remaining antiparticles were the antineutrinos, and the primordial soup consisted of a large number of photons (who are their own antiparticles) and neutrinos (and antineutrinos) and a smaller number of "left-over" protons, neutrons, and electrons. (Dark matter is left out of this story; in typical dark matter models, there are an equal number of dark matter particles and antiparticles in the late universe. We will come back to this later.)

When the universe was a few minutes old, $T \sim 100$ keV, protons and neutrons formed nuclei of light elements. This process is known as Big Bang Nucleosynthesis (BBN), and it produced about 75% (of the total mass in ordinary matter) ^{1}H, 25% ^{4}He, 10^{-4} ^{2}H, 10^{-4} ^{3}He, and 10^{-9} ^{7}Li. (Heavier elements were formed much later, in stars.) At this time matter was completely ionized, all electrons were free. In this plasma the mean free path of a photon was short and the universe was opaque.

The universe became transparent when it was about 400 000 years old. At a temperature $T \sim 3000$ K (~ 0.25 eV), electrons and nuclei formed neutral atoms, and the photon mean free path became longer than the radius of the observable universe. This event is called *recombination*. (This name, taken from thermodynamics, is misleading, since this is actually the first time ever when electrons and nuclei combine.) Since recombination the primordial photons have been travelling through space almost without scattering. We can observe them today as the *cosmic microwave background* (CMB). The CMB is like a photograph of the universe at 380 000 years of age, modified by its passage to us through 14 billion years.

The CMB shows that the early universe was locally very homogeneous and isotropic, unlike the present universe, where matter has accumulated into stars, galaxies, clusters, filaments and superclusters. (The distribution of structures in the late time universe is *statistically* homogeneous and isotropic, but there are large local variations.) The density variations in baryonic matter and photons were of the order 10^{-5},[5] and we see these as small intensity variations of the CMB (the CMB *anisotropy*). Due to gravity, slight overdensities in the matter grew in time, and eventually they formed galaxies. This process is called cosmological *structure*

[4] Neutrinos also have very small masses. However, at temperatures less than the neutrino mass, the neutrino interactions are so weak that the neutrinos and antineutrinos cannot annihilate each other. It is also possible that neutrinos are their own antiparticles, like photons.

[5] There were larger density variations in dark matter. The amplitude depends on the scale, but for relevant cosmological scales they were of the order 10^{-3} – we will come to this later.

formation. Galaxies are not evenly distributed in space but form various structures, galaxy groups, clusters (large gravitationally bound groups), superclusters, filaments, and walls, separated by large, relatively empty voids. Observations of this present-day *large scale structure* of the universe form a significant body of data, which our cosmological theories are able to explain in detail.

There are two parts to structure formation:

1. The origin of the primordial density fluctuations, the “seeds of galaxies”. The scenario that has been most successful in explaining the observed properties of the primordial fluctuations is cosmic *inflation*, which probably occurred much before the EW phase transition. We will discuss inflation in the second part of the course and we will calculate how primordial perturbations are generated from quantum fluctuations.

2. The growth of fluctuations in time. The growth is due to gravity, and it depends on the composition and amount (average density) of matter, as well as the way the universe expands.

One of the main problems in cosmology today is that most of the matter and energy content of the universe appears to be in some unknown forms, called *dark matter* and *dark energy.* The dark matter issue dates back to 1930s, and the dark energy problem arose in the 1990s.

From the motions of galaxies we can deduce that the matter we can directly observe as stars, “luminous matter”, is just a small fraction of the total mass which affects the galaxy motions through gravity. The rest is dark matter, something that we observe only via its gravity. There are also many other lines of evidence for dark matter, including the structure of the CMB anisotropies, the pattern of large-scale structure and the motions of galaxies and gas in clusters and gravitational lensing.[6]

We do not know what most of this dark matter is. A small part of it is just ordinary, baryonic, matter, which consists of atoms just like stars, but does not shine enough for us to notice it. Possibilities include planet-like bodies in interstellar space, “failed” stars (too small, $m < 0.07 m_\odot$, to ignite thermonuclear fusion) called *brown dwarfs*, old white dwarf stars, and tenuous intergalactic gas. In large clusters of galaxies the intergalactic gas[7] is so hot that we can observe its radiation. Thus its mass can be estimated and it turns out to be several times larger than the total mass of the stars in the galaxies. We can infer that other parts of the universe, where this gas would be too cold to be observable from here, also contain significant amounts of thin gas, which is apparently the main component of *baryonic dark matter.*

However, we can estimate from BBN and the CMB anisotropies the total amount of baryonic matter, and it is not nearly enough to solve the whole dark matter problem. Most of the dark matter is non-baryonic, meaning that it is not made out of protons and neutrons[8]. The only non-baryonic particles in the standard model

[6] Since all evidence for dark matter comes from its gravitational effect, it might in principle be possible to explain the observations without dark matter by modifying the laws of gravity instead. However, fitting all the different observations requires resort to rather baroque models, whereas dark matter explains several observations via one simple hypothesis. Most of the focus is now on determining what kind of particles dark matter is made of, not whether it exists.

[7] This gas is ionized, so it should more properly be called *plasma*. Astronomers, however, often use the word “‘gas” also for ionized matter.

[8] And electrons. Although electrons are not baryons (they are leptons), cosmologists refer to matter made out of protons, electrons, and neutrons as “baryonic”. Electrons are anyway so light that their contribution to the total mass is tiny.

of particle physics that could act as dark matter are neutrinos. If they had suitable mass, ~ 1 eV, neutrinos left over from the early universe would have sufficient mass density to be a significant dark matter component. However, producing the structures seen in the universe requires most of the dark matter to have different properties than neutrinos have. Technically, most of the dark matter must be "cold" or "warm", instead of "hot". These terms refer to the dynamics of the particles making up the matter, and do not specify the details of the particles. The difference between *hot dark matter* (HDM) and *cold dark matter* (CDM) is that HDM is made of particles whose velocities were large when structure formation began, but CDM particles had small velocities. Neutrinos with $m \sim 1$ eV would be HDM. Dark matter between the two alternatives is called *warm dark matter* (WDM), and it is still observationally allowed. As the Standard Model does not contain particles that could explain all of the dark matter, it appears that most of the matter in the universe is made out of some unknown particles.

Usually, the term "dark matter" is used to refer only to non-baryonic dark matter, and often neutrinos are also excluded, so that "dark matter" refers only to the unknown, exotic part of the matter that is not observed via light.

Particle physicists have independently come to the conclusion that the Standard Model is not the final word in particle physics. Many proposed extensions of the Standard Model contain suitable dark matter particle candidates (e.g. neutralinos, technibaryons, axions, right-handed neutrinos). Their interactions have to be rather weak to explain why they have not been detected so far, which imposes constraints on models of particle physics. Dark matter thus presents one area where the physics of the very small and the very large overlap.

The situation of *dark energy* is different. While cosmologists consider dark matter quite natural, many of them find dark energy puzzling. The increasing breadth and precision of cosmological observations has made it possible to determine the distance scales and the expansion of the universe accurately. In the context of the homogeneous and isotropic models based on general relativity, a form of matter called dark energy is required to fit these observations. Unlike dark matter, which is clustered, dark energy is relatively uniform in the observable universe. And while dark matter has negligible pressure, dark energy has a large *negative* pressure. The simplest possibility for dark energy is a *cosmological constant* or *vacuum energy*. High energy physics predicts that the the vacuum has an energy density, but it is difficult to understand the small energy scale $\sim$ meV that is required to explain the observations. Another possible explanation is modification of the law of gravity at large distances. In the dark energy case, this is less difficult than for dark matter, as the only observable effect of dark energy is to increase the expansion rate of the universe at late times, whereas the effect of dark matter is seen in various physical systems on different scales and in different eras of the universe. Nevertheless, constructing models that would explain the observations on large scales while being consistent with the precision tests of general relativity in the solar system has proven to be difficult (apart from the cosmological constant). This remains an active area of research. The third possibility is that the homogeneous and isotropic approximation is not good enough at late times due to the formation of non-linear structures. Studying this possibility is difficult, because it requires dealing with non-perturbative general relativity in the complex setting of cosmological structure formation.

1.2 Units and terminology

Natural units. We use natural units in which $c = \hbar = k_B = 1$.

$c = 1$

The theory of relativity unifies space and time into a single concept: spacetime. It is thus natural to use the same units for measuring spatial distance and time. Since the (vacuum) speed of light is $c = 299\ 792\ 458$ m/s, we set 1 s $\equiv$ 299 792 458 m, so that $c = 1$ and 1 second = 1 light second and 1 year = 1 light year. Velocity is thus a dimensionless quantity, and smaller than one for massive objects. Energy and mass have the same dimension and the relation between mass m and energy E for free particles $E^2 = m^2c^4 + p^2c^2$ is simply $E^2 = m^2 + p^2$, where p is particle momentum.

$k_B = 1$

Temperature T is a parameter that describes a thermal equilibrium distribution. The formula for the occupation number of energy level E includes the exponential form $e^{\beta E}$, where $\beta = 1/(k_B T)$. The only function of the Boltzmann constant, $k_B = 1.3805 \times 10^{-23}$ J/K, is to convert temperature into energy units. We decide to give temperature directly in energy units, so k_B becomes unnecessary. We define $1\ \text{K} = 1.3805 \times 10^{-23}$ J, or

$$1\ \text{eV} = 11600\ \text{K} = 1.78 \times 10^{-36}\ \text{kg} = 1.60 \times 10^{-19}\ \text{J}\ . \tag{1.1}$$

Thus $k_B = 1$, and the exponential form is just $e^{E/T}$.

$\hbar = 1$

The third simplification in the natural system of units is to set the reduced Planck constant to unity, $\hbar = h/2\pi = 1$. In SI units we have $\hbar = 1.054573 \times 10^{-34}$Js, so in the natural system of units the dimensions of mass and energy are equal to the dimension of 1/time or 1/distance. This is convenient, because the typical time and distance scales of quantum mechanics are associated with particle energy. For example, the energy of a photon $E = \hbar\omega = \omega$ is equal to its angular frequency. We have

$$1\,\text{eV} = 5.07 \times 10^6\,\text{m}^{-1} = 1.52 \times 10^{15}\,\text{s}^{-1}\ . \tag{1.2}$$

A useful relation to remember is

$$\hbar \approx 197\ \text{MeV}\,\text{fm} = 1\ , \tag{1.3}$$

where we have the energy scale ~ 100 MeV and length scale 1 fm of strong interactions.

Equations become now simpler and the physical relations more transparent, since we do not have to include the above fundamental constants. However, still have to do conversions among different units because the preferred units used in particle units and cosmology (not to mention astrophysics) are different.

Astronomical units. A common unit of mass and energy is the solar mass, $m_\odot = 1.99 \times 10^{30}$ kg, and a common unit of length is one parsec, 1 pc = 3.26 light years = 3.09×10^{16} m. One parsec is defined as the distance from which 1 astronomical unit AU, the distance between the earth and the sun) forms an angle of one arcsecond, 1”.[9] (Astronomers and cosmologists only use light years when talking to outsiders.) A more common scale in cosmology is 1 Mpc = 10^6 pc, which is roughly the typical distance between galaxies at the present time.

1.3 The observable universe

The observations relevant to cosmology are mainly astronomical. The speed of light is finite, and therefore, when we look far away, we also look back in time. The universe has been transparent since the formation of atoms at about 380 000 years, so we can see more than 99.99% of the history of the universe.

The most important channel of observation is the electromagnetic radiation (light, radio waves, X-rays, etc.) coming from space. We also observe charged particles (protons, electrons, nuclei), called *cosmic rays*, as well as neutrinos. In addition, the composition of matter in the solar system has cosmological significance. An important channel is gravitational waves. The first direct detection of gravitational waves was made on September 14 2015 by the Advanced LIGO detector, opening the era of gravitational wave astronomy.

1.3.1 Redshift and the Hubble law

One of the starting points of modern observational cosmology was the discovery by Lemaître in 1927 (theoretically) and Hubble in 1929 (observationally) that the *redshifts* of galaxies are proportional to their distance. Redshift refers to the fact that the light is redder (has longer wavelength) when it arrives to us than when it was emitted. This redshift can be determined with high accuracy from spectral lines. These lines are caused by transitions between different energy states of atoms, and thus their original wavelengths λ_0 can be measured in the laboratory on Earth[10]. The redshift z is defined as

$$z \equiv \frac{\lambda - \lambda_0}{\lambda_0} \qquad \text{or} \qquad 1 + z = \frac{\lambda}{\lambda_0} \tag{1.4}$$

where λ is the observed wavelength. The redshift is observed to be independent of wavelength and follows the relation

$$z = H_0 d \,, \tag{1.5}$$

where d is the distance to the galaxy and z is its redshift. (The speed of light c is set to unity, as noted above.) This relation was first discovered by Lemaître, and it is called the *Hubble law*. The proportionality constant H_0 is correspondingly called

[9] One degree is divided into 60 arc minutes, denoted by 60’, and one arc minute is divided into 60 arc seconds, denoted by 60”.

[10] Assuming that the laws of physics are the same here and at the emission event. Put another way, spectral lines offer a sensitive test of the change of the laws of quantum electrodynamics and nuclear physics in time and space. No deviations from the laws observed on Earth have been found so far.

Hubble constant. It was introduced, its value was first determined from observations, and it was interpreted as the expansion rate by Lemaître.

While the redshift can be readily determined with high accuracy, it is more difficult to determine the distance d. Measurements of distances in general and the Hubble parameter in particular have been the subject of much work and controversy over decades. Distance determinations used to be exclusively based on the notion of the *cosmic distance ladder*. This refers to a series of relative distance determinations between more nearby and faraway objects. The first step of the ladder is made of nearby stars, whose absolute distance can be determined from their *parallax*, their apparent motion on the sky due to our motion around the Sun. The other steps require "standard candles", classes of objects with the same absolute luminosity (radiated power), so that their relative distances are inversely related to the square roots of their "brightness" or apparent luminosity (received flux density). Often several steps are needed, since objects that can be found close by are too faint to be observed from very far away, errors (inaccuracies) accumulate from step to step. Nowadays there are also measurements which do not use the distance ladder, and the value of H_0 is determined to a reasonable accuracy. The latest measurement of H_0 reports the value [**?**]

$$H_0 = 73.2 \pm 1.7 \text{ km/s/Mpc} .$$

The error represents the range where the real result is with a 68% probability. Doubling the error bar gives the 95% probability range. (Unless otherwise noted, all errors bars in the lecture notes correspond to the 68% range.)

There several observations pointing to a Hubble constant around this value, but there is still some discrepancy around 3 standard deviations. This uncertainty of the distance scale is reflected in many cosmological quantities. It is customary to give these quantities multiplied by the appropriate power of h, defined by

$$H_0 = h \cdot 100\text{km/s/Mpc}. \tag{1.6}$$

Very conservatively, we have $h = 0.6 \ldots 0.8$.

For small redshifts ($z \ll 1$), the redshift is often thought of as the Doppler effect due to the relative motion of the source and the observer. The distant galaxies are thus apparently receding from us with the velocity

$$v = z \ . \tag{1.7}$$

The further out the galaxies are, the faster they recede from us. Astronomers often report the redshift in units of velocity (by reintroducing c in (1.7), $z = v/c$).

However, according to general relativity, movement in space is not the proper way understand the redshift. The galaxies are not actually moving, the distance between the galaxies increases because the space between the galaxies is expanding. We will later derive the redshift from general relativity. It turns out that equations (1.5) and (1.7) hold only at the limit $z \ll 1$, and the general result, $d(z)$, which relates distance d and redshift z is more complicated. In particular, the distance reaches a finite value as z goes to infinity – though we should be careful about what we mean by distance! We will look at this in detail in the next chapter. The redshift is directly related to the expansion. The easiest way to understand the cosmological redshift is that the wavelength of travelling light expands with the universe. Thus the universe has expanded by a factor $1+z$ during the time light travelled from an object with redshift z to us.

The largest observed redshift of a galaxy is $z = 11.1$ at present. Thus the universe has expanded by a factor of 12 while the observed light has been on its way. When the light left the galaxy, the age of the universe was only about 400 million years old.

The expansion rate H changes on the cosmological timescale. The time-dependent function $H(t)$ is called the Hubble parameter, and its present value is called the Hubble constant, H_0. In cosmology, it is customary to denote present values of quantities with the subscript 0. Thus $H_0 \equiv H(t_0)$.

The galaxies are not exactly at rest in the expanding space. Each galaxy has its own *peculiar motion* $v_{\rm gal}$, caused by the gravity of nearby mass concentrations, such as other galaxies. Neighbouring galaxies can fall towards each other, orbit each other and so on[11].

Thus the redshift of an individual galaxy is the sum of the cosmic and the peculiar redshift.

$$z = H_0 d + v_{\rm gal} \qquad (\text{when } z \ll 1) \ . \tag{1.8}$$

Usually only the redshift is known precisely. Typically $v_{\rm gal}$ is around 300 . . . 500 km/s. (In large galaxy clusters, where galaxies orbit each other, it can be several thousand km/s; but then one can take the average redshift of the cluster.) For faraway galaxies, $H_0 d \gg v_{\rm gal}$. The larger the redshift, the younger the universe was when the light left.

1.3.2 The horizon

Because of the finite speed of light and the finite age of the universe, only a finite part of the universe is observable. Our *horizon* is at the distance from which light has just had time to reach us during the entire age of the universe. Were it not for the expansion of the universe, the distance to this horizon $r_{\rm hor}$ would equal the age of the universe, around 14 billion light years (4300 Mpc). Expansion complicates the situation; we will calculate the horizon distance later. For large distances the redshift grows faster than (1.5). At the horizon $z \to \infty$, i.e. $d_{\rm hor} = d(z = \infty)$. The universe has been transparent only for $z < 1090$ (after recombination), so the "practical horizon", i.e. the limit to what we can see (at least using light; gravitational waves are not subject to this restriction), lies already at $z = 1090$. The distances $d(z = 1090)$ and $d(z = \infty)$ are very close to each other; $z = 4$ is about halfway from here to the horizon.

Therefore we can only observe a finite region of the universe, enclosed in the sphere with radius $d_{\rm hor}$. The universe can extend to large distances beyond that, and it may even be infinite. Sometimes the word "universe" is used to denote just this observable part of the "whole" universe. Then we can say that the universe contains some 10^{12} galaxies and 10^{23} stars. Over cosmological time scales the horizon recedes and parts of the universe that are beyond our present horizon become observable. (However, if the expansion continues to accelerate as it seems to have done during

[11] Alternatively, we may say that the galaxies remain in place, but there are local deviations in the expansion rate. This is the more natural interpretation from the point of view of general relativity. However, the idea of peculiar velocities is perhaps simpler to understand, as it is closer to Newtonian gravity in the way it is usually formulated.

the past few billion years, the observable region will not grow, and in the distant future galaxies that are now observable will disappear from sight.)

1.3.3 The electromagnetic channel

Consider first the electromagnetic channel of observation. Interstellar space is transparent, except for radio waves longer than 100 m, which are absorbed by interstellar ionized gas, short-wavelength ultraviolet radiation, which is absorbed by neutral gas and very high-energy gamma rays, which interact with the cosmic microwave background. However, Earth's atmosphere is opaque except for two wavelength ranges, the *optical window* ($\lambda = 300$–800 nm), which includes visible light, and the *radio window* ($\lambda = 1$ mm–20 m). The atmosphere is partially transparent to infrared radiation, which is absorbed by water molecules in the air; high altitude and dry air favours infrared astronomy. Accordingly, the traditional branches of astronomy are optical astronomy and radio astronomy. Observations at other wavelengths have become possible only during the past few decades, from space (with satellites), at very high altitude in the atmosphere (with planes, rockets and balloons) and in very dry regions (such as Antarctica).

From optical astronomy we know that there are *stars* in space. The stars are grouped into *galaxies*. There are different kinds of galaxies, such as irregular, elliptical, and flat disks or spirals. Our own galaxy (the Galaxy, or Milky Way) is a spiral galaxy. The plane of the disk can be seen (on a dark night) as a faint band –the Milky Way– across the sky.

Notable nearby galaxies and dwarf galaxies are the Andromeda galaxy (M31) and the Magellanic clouds (LMC, Large Magellanic Cloud, and SMC, Small Magellanic Cloud). These are the only other galaxies that are visible to the naked eye. (The Magellanic clouds and the center of the Milky Way lie too far south to be seen from Finland.) The number of galaxies that can be seen with powerful telescopes is many billions.

Other observable objects include dust clouds, which hide stars behind them, and gas clouds. Gas clouds absorb starlight at certain frequencies, which excite the gas atoms to higher energy states. As the atoms return to lower energy states they emit photons at the corresponding wavelength. Thus we can determine from the spectrum of light what elements the gas cloud is made of. In the same way the composition of stellar surfaces can be determined.

The earliest "cosmological observation" was that the night sky is dark. If the universe were eternal and infinitely large, unchanging and similar everywhere, our eye would eventually meet the surface of a star in every direction, and the entire night sky would be as bright as the Sun. This is called *Olbers' paradox.*

Optical astronomy and the large scale structure. There is a large body of data relevant to cosmology from optical astronomy. By counting the number of stars and galaxies we can estimate the matter density they contribute to the universe. From the different redshifts of galaxies within the same galaxy cluster we obtain their relative motions, which reflect the gravitating mass within the system. The mass estimates for galaxy clusters obtained this way are much larger than those obtained by counting the visible stars and galaxies in the cluster, pointing to the existence of dark matter.

From the spectral lines of stars and gas clouds we can determine the relative amounts of different elements and their isotopes in the universe.

The distribution of galaxies in space and their relative velocities tell us about the *large scale structure* of the universe. The galaxies are not distributed uniformly. There are galaxy groups and clusters. Our own galaxy belongs to a small group of galaxies called the Local Group. The Local Group consists of three large spiral galaxies: M31 (the Andromeda galaxy), M33, and the Milky Way, and about 60 smaller galaxies and dwarf galaxies. The local group's diameter is around 3 Mpc. The nearest large cluster is the Virgo Cluster. The grouping of galaxies into clusters is not as strong as the grouping of stars into galaxies. Rather, galaxies are distributed in a complex pattern called the *cosmic web*, which consists of walls, filaments, clusters and voids (low-density regions). Most galaxies are not part of any well defined cluster.

Radio astronomy. The sky looks very different on radio wavelengths than to the naked eye. There are many strong radio sources very far away. These are galaxies which are optically barely observable. They are distributed isotropically, i.e. there are equal numbers of them in every direction, but there are more far away (at $z > 1$) than close by ($z < 1$). The isotropy is evidence of the homogeneity of the universe at the largest scales—there is structure only at smaller scales. The dependence on distance is a time evolution effect in two ways: the radio sources evolve and the volume of the universe evolves. In general, there are more objects at larger distances simply because there is more volume there. However, the evolution in the number counts is not explained only by the change in the volume of the universe, it shows that the radio sources themselves evolve. Some galaxies are strong radio sources when they are young, but become weaker with age by a factor of more than 1000.

Cold gas clouds can be mapped using the 21 cm spectral line of hydrogen. The ground state ($n = 1$) of hydrogen is split into two very close energy levels depending on whether the proton and electron spins are parallel or antiparallel (the hyperfine structure). The separation of these energy levels, the hyperfine structure constant, is 5.9 μeV, corresponding to a photon wavelength of 21 cm, i.e. radio waves. The redshift of this spectral line shows that redshift is independent of wavelength (it is the same for radio waves and visible light), as it should if it is due to the expansion of space.

Cosmic microwave background. At microwave frequencies the sky is dominated by the *cosmic microwave background* (CMB), which is highly isotropic, i.e. the microwave sky glows almost uniformly, with only very small contrasts. The electromagnetic spectrum of the CMB is the black body spectrum with a temperature of $T_0 = 2.72548 \pm 0.00057$ K [2]. It follows the theoretical black body spectrum better than anything else we have observed or produced. It is the remnant of a hot state in the early universe, when matter and light were almost homogeneously and isotropically distributed and in thermal equilibrium. The temperature of the CMB falls as $(1+z)^{-1}$ due to photon redshift, so as the CMB redshift is about 1090, the original temperature was about 3000 K.

The state of a system in thermal equilibrium is determined by a small number of thermodynamic variables, in this case the temperature and chemical potentials (for particles with conserved quantum numbers). The observed temperature of the

CMB and the observed density of the present universe allows us to fix the evolution of the temperature and the density of the universe, which then allows us to calculate the sequence of events in the early universe. That the early universe was hot and in thermal equilibrium is a central part of the big gang paradigm, and it is often called the hot dig bang theory to spell this out.

With sensitive instruments a small anisotropy can be observed in the microwave sky. This is dominated by the *dipole anisotropy* (one side of the sky is slightly hotter and the other side colder), with an amplitude of 3.346 ± 0.017 mK, or $\Delta T/T_0 = 0.0012$. This is a Doppler effect due to the motion of the observer, i.e. the motion of our Solar System with respect to the radiating matter at our horizon. The velocity of this motion is $v = (\Delta T/T_0)\, c = 369$ km/s, or $v = 0.00123$ and it is directed towards the constellation of Leo (R.A. $11^h 8^m 50^s$, Dec. $-6°37'$), near the autumnal equinox (where the ecliptic and the equator cross on the sky) [3]. It is due to two components, the motion of the Sun around the center of the Galaxy, and the peculiar motion of the Galaxy due to the gravitational pull of matter concentrations up to 100–200 Mpc away[12].

When we subtract the effect of this motion from the observations (and look away from the plane of the Galaxy—our Galaxy also emits microwave radiation, but with a non-thermal spectrum) the true anisotropy of the CMB remains, with an amplitude of about 3×10^{-5}, or 80 μK.[13] This anisotropy gives a picture of the small density variations in the early universe, the "seeds" of galaxies. Theories of structure formation have to match both the small inhomogeneity of the order 10^{-5} at $z = 1090$ and the structure observed today ($z = 0$).

Gamma ray bursts and quasars. The highest energy region of the electromagnetic spectrum is occupied by γ rays. Space-based γ-ray observatories have discovered powerful Gamma Ray Bursts (GRB) on the sky. These are short events lasting from a fraction of a second to a few minutes. They are observed about once per day, and are distributed isotropically on the sky. The isotropic distribution suggests that they are at cosmological distances (further out than our own or nearby galaxies). This has been confirmed by the identification of some GRB's with galaxies with high redshifts ($z > 1$). This means that the bursts must have extremely high energies. The longer duration (longer than a second) GRB's appear to be related to particularly powerful supernova events. The shorter duration (less than a second) are possibly due to collisions of neutron stars with each other or with black holes.

[12] Sometimes it is asked whether there is a contradiction with special relativity here—doesn't CMB provide an absolute reference frame? There is no contradiction. The relativity principle just says that the *laws of physics* are the same in the different reference frames. It does not say that *systems* cannot have reference frames which are particularly natural for that system, e.g. the center-of-mass frame or the laboratory frame. For road transportation, the surface of the earth is a natural reference frame. In cosmology, the CMB gives us a good "natural" reference frame—it is closely related to the center-of-mass frame of the observable part of the universe, or rather, a part of it which is close to the horizon (the *last scattering surface*). The different parts of the plasma from which the CMB originates are moving with different velocities (part of the 10^{-5} anisotropy is due to these velocity variations). If there is something surprising here, it is that these relative velocities are so small, of the order of just a few km/s, reflecting the astonishing homogeneity of the early universe over large scales. We will return to the question of whether these are natural initial conditions when we discuss inflation in the second part of the course.

[13] These numbers refer to the standard deviation of the CMB temperature on the sky. The hottest and coldest spots deviate some 4 or 5 times this amount from the average temperature.

Quasars (Quasistellar Objects, QSOs) are the most powerful continuously radiating objects in the universe. Thus the most-distant (earliest) objects observed in the universe are mostly quasars. The highest observed power is about 10^{41} W, which can be compared to the total luminosity of the Milky Way of 5×10^{36} W. At first quasars were considered different from galaxies since they looked like point-like objects. In photographs they looked like stars, but their redshifts revealed their huge distances and thus their huge power outputs. Now better observations have revealed "host" galaxies around several quasars. It has been concluded that quasars are powerfully radiating galactic nuclei, and are related to some more close-by galaxies (Seyfert galaxies), whose nuclei are also fairly powerful sources of radiation. Together these objects are called Active Galactic Nuclei (AGN). Quasars are powerful sources at many different wavelengths (radio, optical, X-ray). Some of them belong to the radio sources mentioned earlier, others are radio quiet. There are more quasars at large distances (in the past, $z > 1$) than nearer to us (later, $z < 1$, because quasars grow fainter as they age; they become normal quiet galaxies.

The power source of an AGN is thought to be a very large black hole (with $m = 10^8 M_\odot$ or so) at the center of the galaxy, into which surrounding matter is falling. As it approaches the hole, this matter is heated up and begins to radiate. AGN's quiet down over cosmological time scales as the black hole gradually cleans up the surrounding regions.

1.3.4 Cosmic rays

Cosmic rays are electrons, positrons, protons, antiprotons and nuclei coming from space. Some of them have extremely high energies, some above 10^{20} eV (in the laboratory frame; in the center-of-mass frame the energy is of the order 10^{15} eV). These energies are higher than what can be reached in particle accelerators (LHC reaches $\sim 10^{13}$ eV). It is thought that cosmic rays originate from supernovae (exploding stars). Since they are charged particles their paths are warped by galactic magnetic fields, so their arrival direction does not point towards their origin. The cosmic rays are about 90% protons, 10% other nuclei, 1% electrons and less than 1% antiprotons and positrons. All elements up to uranium are represented.

References

[1] A. G. Riess *et al.*, Astrophys. J. **826** (2016) no.1, 56 doi:10.3847/0004-637X/826/1/56 [arXiv:1604.01424 [astro-ph.CO]].

[2] D.J. Fixsen, Astrophys. J. **707** (2009) 916 [arXiv:0911.1955 [astro-ph.CO]]

[3] N. Aghanim *et al.* [Planck Collaboration], Astron. Astrophys. **571** (2014) A27 [arXiv:1303.5087 [astro-ph.CO]] (conversion to equatorial coordinates: `www.astro.utu.fi/EGal/CooC/CooC6.html`)