Contents

Statement

The writing of these lecture notes mostly followed:

- D. Baumann, *Cosmology*, Cambridge University Press (7, 2022)

Inspiration was also taken from:

- S. Dodelson, *Modern Cosmology*, Academic Press, Amsterdam (2003)
- A. Riotto, *Lecture Notes in Cosmology*, Prague (2013)
- D. Tong, *Cosmology*, Cambridge (2019)

1 Observations, or the birth of Cosmology

The study of the Universe, and the subject of Cosmology with it, started with some incredible observations, that we wish to summarize in this first chapter. Unfortunately, it is not as simple as "observations drive the theory", but rather it has been, historically, a constant feedback relationship. For instance, we will see how some preliminary observations drove the community to devise a theory that, however, was later shown to be missing some components. In any case, in this chapter we will list the observations that allowed cosmologists to create a framework that describes of our Universe with very good precision.

1.1 Scales in the Universe

The length scales in our Universe can get incredibly big, so big in fact that there is no obvious way to get used to them unfortunately. Humbly, to get a sense of the distances in cosmology (if one can), we start out with the smaller scales, and then slowly build up the larger scales, up to the biggest ones we know.
Firstly, everyone knows that Earth is part of the *Solar System*, where it is the third planet from the Sun in terms of distance. In fact, we are 150 million kilometers away from it (about 11780 Earths stacked up side by side), such that light itself takes about 8 minutes to travel that distance.
Some time ago (not too far back) we also found out that our Solar System is not the end of the story, but much more can be found outside it. Peering out at the sky reveals the existence of a multitude of objects, of which the closest ones are stars, the little specks of light one sees at night (if you're lucky). The star nearest to us, *Proxima Centauri*, lies "just" 4.2 light years away. Since distances tend to get incredibly large outside of our Solar System, Cosmology, as any other framework in physics, decides to use its own length (energy) units, based on the scales at which it operates. (Think about particle or atomic physics. There the scales are much smaller, i.e. the energies at play are much bigger, so that it's convenient to use *electronvolts* eVs as a unit.) In particular, in Cosmology the *parsec* is the unit of choice:

$$1 \text{ pc} \approx 3.26 \text{ light years}$$

At the next level of scale we find galaxies. As it appears, the Universe contains roughly 100 billion galaxies, each of which hosts 100 billion stars; in fact, all the stars that we can see with our naked eye reside in our own galaxy, the *Milky Way*, a spiral galaxy around 30 kpc in length, with the Sun located at the edge of a spiral arm, 8 kpc away from its center. The Sun moves very slowly though, taking about 250 million years to complete a full orbit around the Milky Way.

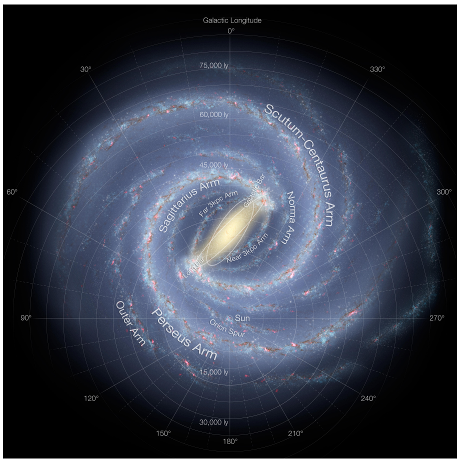

Figure 1: A representation of our galaxy, the Milky Way

The nearest galaxy to ours is the famous *Andromeda Galaxy* or *M31*, a spiral galaxy about 2 million light years away. It is actually one of about 50 other galaxies that are gravitationally bound together: this arrangement (including our own) is what's called the *Local Group*.
At even bigger scales, galaxies organize themselves in *galaxy clusters* and *superclusters*, with filamentary structures and gigantic voids in between them. The scales of such clusters, like our own *Local Supercluster*, are in the order of about 500 million light years. These last objects can be very difficult to imagine, so here below we show images taken from the *Millennium Simulation*[1], which aimed to simulate the larges scales known in the Universe:

[1]More information can be found at the website https://wwwmpa.mpa-garching.mpg.de/galform/virgo/millennium/

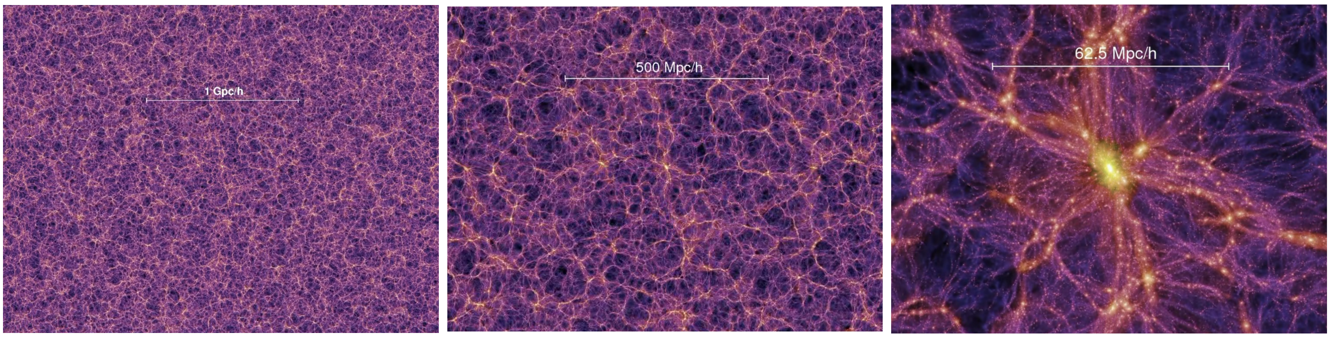

Figure 2: The largest scales in the Universe, from the Millennium Simulation. The figures show the dark matter distribution at present time, zooming in on a galaxy supercluster.

Finally, since light has only a finite speed, the largest scale we can probe is the *observable Universe*, i.e. the distance light could have travelled since the beginning of time, which has a radius of:

$$46.5 \text{ billion years} \approx 14 \text{ Gpc} \approx 4.4 \times 10^{26} \text{ m}$$

We have here discussed the observations regarding the scales of the Universe, to get a sense of how big these actually are. Now we turn to the observations regarding the content of the Universe. What is it made of?

1.2 The Content of the Universe

The Universe appears to strikingly simple. Observations suggest that the mass-energy content is divided into just four main components. The first, and one of the least relevant, is the "ordinary" *baryonic matter*, the protons, neutrons and electrons, the stuff we can actually see but makes up only a small percentage of the Universe; another component is the *radiation*, which is mostly made of photons and neutrinos; this, although important, takes up only a very small fraction of the budget.
However, things are still very mysterious when it comes to the most common types of mass-energy. In fact, we now are pretty certain that the majority of matter is in the form of an elusive *dark matter*, taking up roughly 27% of the energy pie. The final component, and the most incomprehensible, consists of what is usually called *dark energy*, which strikingly makes up the remaining 70% of the pie. Here we will summarize how observations managed to show what the Universe is made of.

1.2.1 Matter

While we will clarify what we mean by "matter" only later on, we can say here that cosmologists distinguish between:

- Baryonic matter: ordinary matter, such as nuclei and electrons. Clearly, this isn't strictly correct, as electrons are leptons, but their mass so small compared to the nuclei, that most of the mass is in the actual baryons anyway. This type of matter composes stars, planets and every living organism, and makes up only 5% of the energy content;
- Dark matter: a type of matter of which the nature still remains a mystery. There are many candidates for dark matter (such as right handed neutrinos, axions, primordial black holes etc...), but we don't even know if it is just one thing or multiple. Luckily, the large scale description of our Universe doesn't really depend on the small scale nature of dark matter, so that in the end we are able to create theories which then make predictions. This type of matter makes up 27% of the matter content of the Universe.

Altogether, as we can tell, the matter content only makes up roughly 31% of the Universe.
It is interesting now to review all the evidence there is for dark matter, since one might ask how we could know of the existence of such a component, given that we don't even know what it is. The key to the answer lies in the large scale properties of dark matter, as we alluded to before.

The first evidence for dark matter was found by Zwicky around 1933 while studying the properties of the galaxies in the Coma cluster (a galaxy cluster), of which a picture is presented here below:

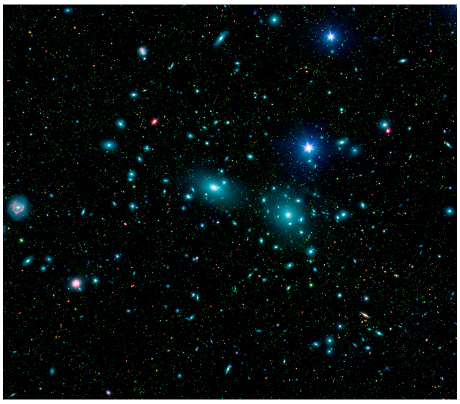

Figure 3: The Coma cluster

It turns out that there are multiple ways of estimating the mass of a galaxy cluster. One way is obvious. When looking with the telescope, we see a certain amount of light, which comes from galaxies, that are themselves made out of baryonic matter (stars are made of mostly hydrogen and helium, while bigger ones contain a small percentage of heavier nuclei. Altogether, however, we are still talking about "normal" matter). We then simply translate this total luminosity (energy of the photons per unit time) into the total baryonic mass of the cluster. There is also a way of calculating the mass that has to do with something called the *virial theorem* (the gravitational one). Let's suppose we have a system of N particles, with masses m_i and positions $\mathbf{x}_i$, that interact via a gravitational potential:

$$V = \sum_{i<j} V_{ij} = \sum_{i<j} -G \frac{m_i m_j}{|\mathbf{x}_i - \mathbf{x}_j|} \tag{1.1}$$

We also assume that the system is "bounded", i.e. all positions and velocities $\dot{\mathbf{x}}_i$ are bounded. The theorem then states that the average kinetic energy of the system T is proportional to the average potential energy V:

$$\langle T \rangle = -\frac{1}{2} \langle V \rangle \tag{1.2}$$

Let's apply all this to the Coma cluster. To make things simpler, we assume that all the galaxies all have the same mass m, so:

$$\langle T \rangle = \frac{1}{2N} \sum_i m v_i^2 \tag{1.3}$$

the virial theorem then states:

$$m \langle v^2 \rangle = \frac{1}{2} G m^2 N \left\langle \frac{1}{r} \right\rangle \tag{1.4}$$

where $\langle 1/r \rangle$ is the mean inverse distance of the galaxies, so that we have a rough estimate of the mass in the cluster:

$$Nm = \frac{2 \langle v^2 \rangle}{G \langle 1/r \rangle} \tag{1.5}$$

since all the quantities on the right are things we can measure. When comparing these two methods, the virial theorem and the luminosity conversion, we find that baryonic matter is not even close to the total amount of matter in the cluster.

Some time later, in the 1960s and 70s, Vera Rubin and her collaborator Kent Ford found further evidence for dark matter while studying the rotation curves of galaxies. What is a rotation curve? When talking about galaxies, a rotation curve describes how the velocity of objects, such as stars, changes as a function of their distance from the center of the galaxy. In order to get a feeling of what the problem seems to be, let's assume a spherically symmetric galaxy (although a bad approximation, this gets the idea across pretty well). If we consider a star, for instance, moving in a circular fashion inside the galaxy, we know that the centrifugal force is given by the gravitational attraction:

$$\frac{v^2}{R} = G \frac{M(R)}{R^2} \tag{1.6}$$

where $M(R) = \int_0^R dr 4\pi r^2 \rho(r)$ is the mass enclosed in a sphere of radius R. Clearly, this equation implies a certain expected function $v(R)$:

$$v(R) = \sqrt{\frac{GM(R)}{R}} \tag{1.7}$$

Now, in general $\rho(R)$ will be peaked near the center of the galaxy, so that when we get into the periphery, $M(R)$ will have stopped growing, i.e. it is a constant. This means that far from the center we have $v(R) \sim \sqrt{1/R}$. This theoretical prediction turns out to be basically always wrong for any given galaxy, for which the rotation curves can be determined by looking at the 21 cm line of hydrogen (converting from a blue/red-shift of the spectrum to the velocity of the gas). As we can see below, the velocity seems to saturate:

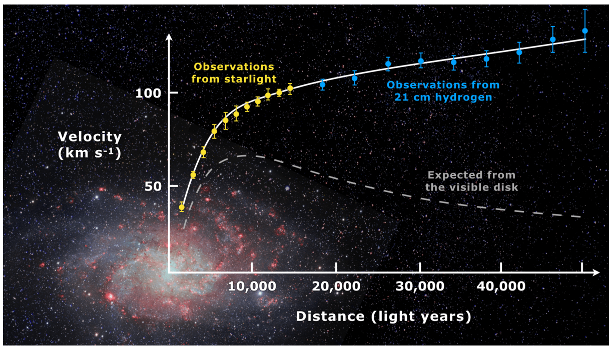

Figure 4: The rotation curve of the galaxy M33

What does this imply? Looking at (1.7), we see that a constant $v(R)$ far from the center is obtained if $M(R) \sim R$, that is the mass keeps growing. All in all we are starting to see a pattern: there is a certain mass that we can't seem to see. Since "seeing" implies some sort of interaction with matter or photons, this matter is labeled as non-interacting, or, better, "dark".

There is other stunning evidence for dark matter. This is seen through a phenomenon called gravitational lensing. It turns out that, in General Relativity, the trajectory of light (the geodesic) gets bent in the presence of massive objects, like cluster of galaxies; this allows us to see objects that are behind such clusters, and they will be more or less distorted, depending on how massive the lens is. Examples of gravitational lensing are shown below:

Figure 5: *Left*: the James Webb Space Telescope deep field. *Right*: the Bullet cluster

On the left we can see the first deep field of the James Webb Space Telescope, capturing thousands of galaxies, some of which are clearly distorted by the massive cluster in the middle of the picture. It should be intuitive that, given a lensed image of an object, like a galaxy or a group of galaxies, we can infer the mass of the lens. This is done for the Bullet cluster as we see on the right-hand side, formed by two clusters that are thought to have previously collided. Three types of matter are shown: the stars, the hot baryonic gas (baryonic matter not in the form of stars), shown in pink, which interacts rather strongly as the collision is taking place, and finally the dark matter, shown in purple, which clearly doesn't seem to have interacted with anything during the collision, since each cluster's halo looks basically unaffected. Again, this suggests that this type of matter is non-interacting.

All this evidence, for now, points towards a new type of non-interacting matter at macroscopic scales. There is, however, some last bit of evidence even at way smaller scales.
It turns out, in fact, that the abundances of light elements, like hydrogen and helium, that make up the baryonic matter in the Universe (combined, hydrogen and helium make up roughly 100%), can be estimated via an accurate

theory called *Big Bang Nucleosynthesis*. The baryon energy density that we get with this theory is, unsurprisingly, in contrast with mainly two observations. Unfortunately, the math being quite involved, we cannot dwell too much on these discrepancies, but rather we will give just a rough idea. For starters, it can be shown that, if dark matter did not exists, the amount of baryons would not be sufficient to form the galaxies we see today in the Universe. Basically, baryons need to be "gravitationally assisted" by dark matter. Finally, the total amount of matter (baryons and dark matter) influences the *cosmic microwave background* (CMB), a background radiation whose properties are known to incredible levels of precision. Again, if the baryons were all that existed, the CMB predictions would look rather wrong.

1.2.2 Radiation

Cosmologists usually label as radiation every particle whose momentum is much larger than its mass, an example clearly being photons, since they are massless. All the photons in the Universe are in the form of a background radiation, which we mentioned before.
Finally, in general, very light particles are considered radiation as well, like neutrinos, which turn out to have a non-negligible impact on the evolution of galaxies. In total, the radiation makes up roughly 0.01% of our Universe, a rather small percentage.

1.2.3 Dark Energy

To understand this last piece, we need to briefly explain the historical situation in cosmology at that time. Famously, Edwin Hubble in 1929 observed that galaxies moved away from our own at speeds proportional to their distance, i.e. the farther they were, the faster they moved. This is usually regarded as the first evidence of the fact that our Universe is *expanding* (actually, the expansion of the Universe was already predicted by Friedmann some years prior). Hubble presented his findings in a "Hubble diagram", as we see below (together with more recent measurements):

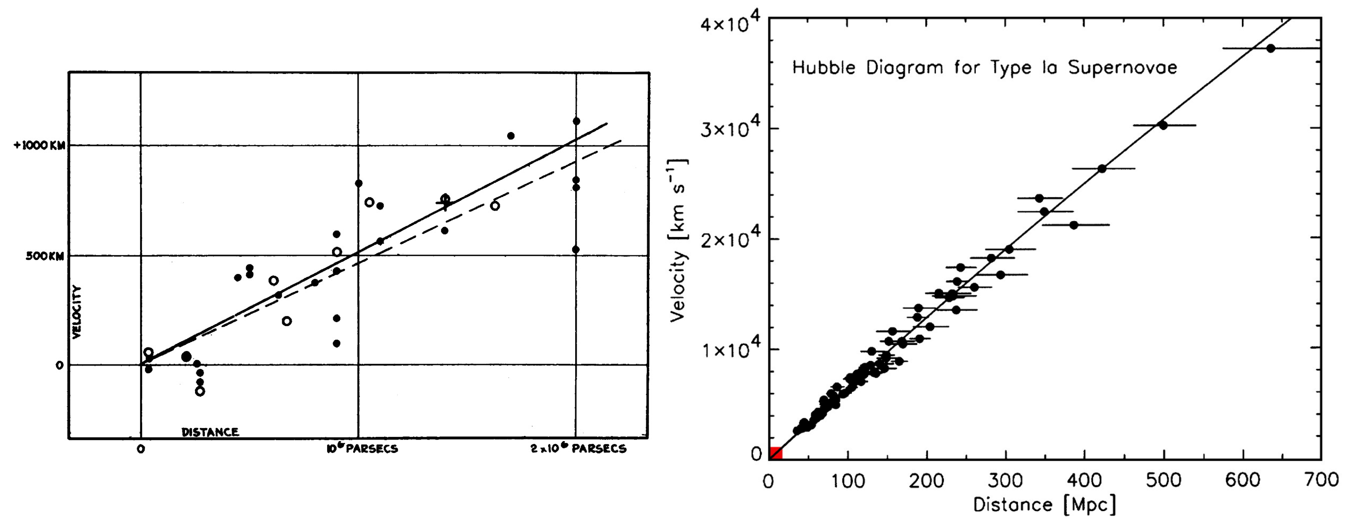

Figure 6: *Left*: Hubble's original results from his article in 1929. *Right*: the same diagram, but with more data points

In general, it is clear that such a diagram implies:

$$v_{\text{gal}} = H_0 D \tag{1.8}$$

where D is the distance of the galaxy, and H_0 is a constant, dubbed the *Hubble constant*, that we will encounter and explain in more details later on. What does this have to do with the rest of the energy content of the Universe? It turns out that this diagram was reasonably well fitted by a Universe with only matter and some radiation. However, at that time this "matter-only" model was in serious trouble, mainly for two reasons. First, it predicted the age of the Universe to be smaller than the age of some of the oldest stars within it; secondly, some other observations suggested that matter was only roughly 30% of the total content of the Universe. Incredibly, in the 1980s, very precise measurements of distant supernovae explosions, showed that they appeared fainter than a matter-only Universe predicted:

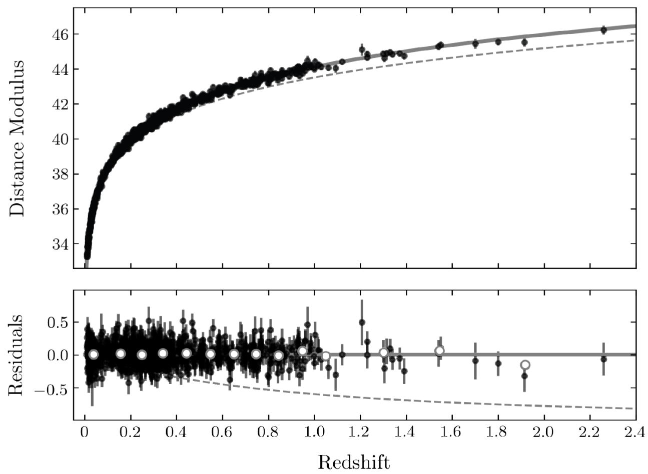

Figure 7: Distance of supernovae as a function of their redshift. The dashed line is the "matter-only" Universe, whereas the solid line introduces a dark energy component

The data could only be consistent if the Universe started *accelerating* sometime recently, and this is possible in General Relativity if one admits the existence of some sort of energy that doesn't dilute with the expansion of the Universe, some sort of *dark matter*. Fitting these new results, then, cosmologists deduced that this new component makes up a good 70% of our Universe. Its nature is still a mystery.

1.3 Large Scale Structure

We now Understand what the Universe is made of. To complete our observational tour-de-force we need to explain what the geometry of our Universe looks like and how the components are arranged in this geometry.
We start with the second topic. There is very good reason to believe that our Universe is *spatially homogeneous* and *isotropic* at very large scales (say, scales greater that 100 Mpc, since clearly it is not so locally). These are arguably the two most important attributes of the Universe, because they ensure that observations made from any vantage point (like where we are), are representative of the whole Universe. This is clearly very powerful when it comes to testing our theories, made from our standpoint, against the properties of the distant Universe. The most incredible evidence for the isotropy of the Universe is the background radiation we mentioned before, the CMB:

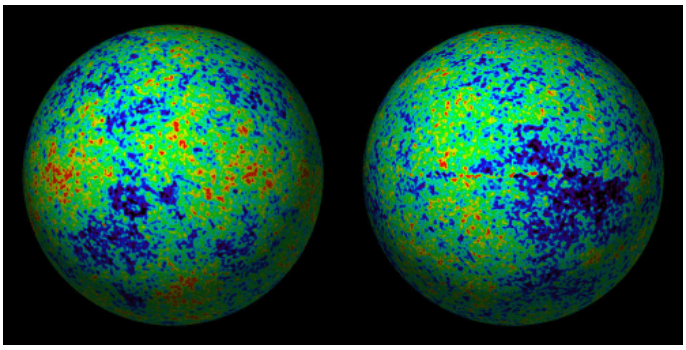

Figure 8: The cosmic microwave background as seen in the sky. The colormap distinguishes hot spots (red) from cold spots (blue). Note that the difference in temperature is of the order of 10^{-5} K.

Notwithstanding the origin of this background radiation, we only note that it is the most perfect black body we have in nature, at a temperature of $\bar{T}_0 = 2.73$ K. More importantly, this temperature is exactly the same (except minor fluctuations of the order of 10^{-5} K) in every direction we look.
In addition to the CMB, lately large galaxy surveys like the 2dF (2 degree field) survey, computed the position and distance of a huge number of galaxies. Strikingly, at scales $\gtrsim 100$ Mpc, they are distributed homogeneously and isotropically:

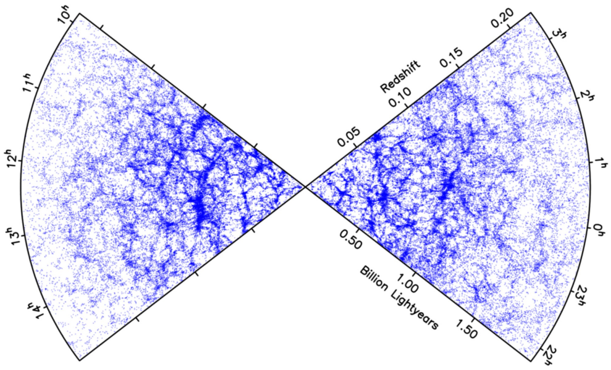

Figure 9: The 2dF galaxy survey

Finally, the last piece of observation regards the spatial geometry. It turns out that, by measuring the properties of the CMB (we are starting to see that this radiation is a rich source of information), we can distinguish between different geometries; this is roughly because the photons travel in very specific ways (they follow geodesics) in the respective geometries, leaving an imprint on the CMB. Very insipidly, our Universe looks very much *flat*, with no hint of any spatial curvature.

1.4 Where do we go from here?

Let's summarise what we know so far. Our Universe is mostly made of dark energy, a mysterious entity, dark matter, baryons and some radiation. Moreover, these components, if we zoom out enough, are distributed in an isotropic and homogeneous fashion throughout a space that looks flat. Finally, not only is the Universe expanding, but it is doing so while also accelerating.
What do we do now with these facts? Miraculously, cosmologists have devised a model, called the ΛCDM *model*, which explains to a high degree of precision every observation that we have listed. This model is the center of focus of the remaining lectures.

2 Towards a Mathematical Description of the Universe

In the previous lecture we learned that our Universe is expanding. Intuitively, then, at earlier times objects must have been close to one another and the far past conditions very hot and dense. During primordial times, particles collided frequently and they were all in thermal equilibrium at a temperature T. A useful relation, which we will not derive, relates the temperature of the Universe T to its age:

$$\frac{T}{1\ \text{MeV}} \approx \left(\frac{t}{1\ s}\right)^{-1/2} \tag{2.1}$$

This equation tells us that, for instance, the Universe had a temperature of about 1 MeV when it was just one second old. The rates of reactions were extremely high, and a lot of interesting phenomena took place in a short timeframe. Above 100 GeV, all the particles of the Standard Model, colliding very frequently, were in equilibrium and therefore their abundances were roughly equal to one another. This can be taken as the initial condition for our Universe. Then, in just 10^{-9} seconds, the Universe expanded by a factor of 10^4 and the temperature dropped rapidly. During this short time the Universe went through successive evolutionary stages, that we will now try to keep track of.
At around 100 GeV the electroweak (EW) symmetry of the Standard Model was broken in what is called the *EW phase transition*, where the weak and electromagnetic forces decoupled (above 100 GeV there is no distinction between the two forces) and particles acquired their mass. The detailed dynamics of this transition are still an open subject of research, even though the basics of it were confirmed by the discovery of the Higgs boson.
As the temperature drops the energy drops, thus after a while the particle-antiparticle annihilation reaction will be favored over the reverse process of particle creation. Clearly, since it requires less energy to form light particles rather than heavier ones, the first particles to disappear in this way were the most massive quarks, followed shortly after by the massive bosons W and Z, the tau lepton and the Higgs. Finally, at around 150 MeV, the remaining quarks, at this point having a very low kinetic energy, could condense into hadrons (mesons and protons/neutrons), resulting in the *QCD phase transition* (the name "QCD" stands for Quantum ChromoDynamics, the theory at the heart of strong interactions).

Particles can fall out of thermal equilibrium when their interaction rate drops below the expansion rate of the Universe, in other words the Universe expands way too rapidly for the particles to interact. At that "moment" those particles will stop interacting with the environment, they *decouple*, creating a relic abundance. One of the most important decoupling events is the *neutrino decoupling* around one second after the Big Bang, which produced the *cosmic neutrino background* (CνB).
About 1 minute after the Big Bang, the temperature dropped enough to allow the creation of the first light nuclei Helium-4 and Lithium-7. This Big Bang nucleosynthesis produced very few elements heavier than lithium, because there are no stable nuclei with 5 or 8 nucleons that would be required in order to sustain the reaction. Heavier elements instead formed only later on in the insides of stars, via *stellar nucleosynthesis*, where densities allow for the creation of, for instance, carbon out of alpha particles via the 3α channel. Later, with the explosion of stars, these heavier elements, up until iron, spread throughout the Universe; instead, the origin of elements that lie beyond iron is still debated, but they most likely formed via *explosive nucleosynthesis*, either with s-processes during the last phases of stellar evolution, or with r-processes in supernovae explosions or neutron star mergers.
About 370000 years after the Big Bang, the temperature dropped enough so that the first atoms could form. This process, known as *recombination*, when the photons stopped interacting with electrons (that in turn could begin to be captured by the nuclei) produced what we now see as a relic photon background, the cosmic microwave background (CMB), which we introduced in the last lecture. We have also mentioned how the CMB contains temperature fluctuations that, although very small ($\delta T/T \sim 10^{-5}$), contain very important information about the primordial Universe.

Now, if the cosmological principle held perfectly, if the distribution of matter in the Universe were perfectly homogeneous and isotropic, there would be no structures (galaxies, clusters etc...) now. Obviously, the inside of the Sun looks very different from the inside of my house, so at sufficiently small scales, inhomogeneities must exist. We then seem to need a mechanism which gives rise to these deviations from perfect uniformity. Where should we look for it? A clue comes from noticing the following thing. It is expected that a general relativistic description of the Universe breaks down at very early times when the Universe is so dense and hot that quantum effects become important (that's why our discussion began at 100 GeV. At higher energies we don't have the slightest of ideas of what happened or what theory we should even use); in fact, a model based only on GR has a number of conceptual problems when applied to the early Universe (as we will see). A possible solution to these problems requires extending the GR model by introducing density perturbations created via quantum fluctuations at early times, which are now believed to be responsible for the formation of all the large scale structure that we see. These fluctuations are also responsible for the temperature anisotropies of the CMB. A popular theory that accomplishes the generation of those fluctuation seeds is the *inflationary theory* in which the Universe is supposed to have gone through a rapid expansion over very little amount of time driven by one (or more) quantum field, the *inflaton*.
Our understanding of the Universe is still very far from complete and, despite the inflationary paradigm, we are still unable to predict the initial conditions for structure formation from first principles, forcing us to rely on a set of parameters taken from observational data.
However, once the initial conditions have been specified, it is very straightforward to compute the way in which density perturbations evolve in time in a homogeneous background, as we will do. When the universe is matter "dominated" (we will see what this means), the perturbations will grow, because a region where the initial density is slightly higher than the mean will attract its sorroundings more strongly than average. Therefore overdense regions become more overdense over time, and underdense regions become more rarefied as matter will flow away from them. The exact rate at which perturbations grow will depend on the cosmological model, and in particular this growth will stop when dark energy comes to dominate the Universe. In any case, once a certain region is overdense enough, it will stop growing and start to collapse, which is the birth of large objects. In general, as we will see, dark matter starts growing earlier than baryonic matter, and nonlinear, quasi equilibrium dark matter objects are called *dark matter halos*; they are the potential wells into which the baryons will eventually fall and form galaxies (this is what we meant when we said that baryons need to be "gravitationally assisted" by dark matter in order to form galaxies).

2.1 General Relativity

The study of Cosmology in general only requires a very basic level knowledge of general relativity. Practically, it is only sufficient to understand what the *metric* is and what it allows us to do, so that's what we will briefly recall here. Let's consider our usual Euclidian 3D space with coordinates $\vec{x} = (x, y, z)$, and two objects at locations $\vec{x}_1 = (x_1, y_1, z_1)$ and $\vec{x}_2 = (x_2, y_2, z_2) = \vec{x}_1 + d\vec{x} = (x_1 + dx, y_1 + dy, z_1 + dz)$ for a very small $d\vec{x}$. Clearly, their distance squared is:

$$ds^2 = dx^2 + dy^2 + dz^2 \tag{2.2}$$

which can be written introducing a *metric tensor*, in this case a bit uninterestingly:

$$ds^2 = \sum_{i,j=1}^{3} \delta_{ij} dx^i dx^j \tag{2.3}$$

where, confusingly, $(x_1, x_2, x_3) \equiv (x, y, z)$, and the fact that the indices are upstairs is just for convenience for now. More importantly, we have introduced:

$$\delta_{ij} = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \tag{2.4}$$

which is just a Kronecker delta, also written more handily as $\delta_{ij} = \text{diag}(1, 1, 1)$. Basically, we can see that the metric tensor δ_{ij} contains the same information as ds^2, and indeed specifies it. This means that a given space will have its own metric (this word is interchangeable with metric tensor), which can be used to determine lengths and angles in the said space. In fact, given two vectors $\vec{A}$ and $\vec{B}$, we can find their norm (the "lengths" part) and the angle between them (the "angles" part):

$$\begin{aligned} \|\vec{A}\|^2 &= \sum_{i,j=1}^{3} \delta_{ij} A^i A^j \\ \vec{A} \cdot \vec{B} &= \sum_{i,j}^{3} \delta_{ij} A^i B^j = \|\vec{A}\| \; \|\vec{B}\| \; \cos\theta \end{aligned} \tag{2.5}$$

Under a coordinate change, the metric can change. We can see this very clearly when we go from cartesian to spherical coordinates, where the distance between the same two objects $\vec{x}_1 = (r, \theta, \phi)$ and $\vec{x}_2 = (r + dr, \theta + d\theta, \phi + d\phi)$ becomes:

$$ds^2_{\text{spherical}} = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 = \sum_{i,j=1}^{3} g_{ij}(\vec{x}) dx^i dx^j \tag{2.6}$$

where we have assumed, as before, $(x_1, x_2, x_3) \equiv (r, \theta, \phi)$ for convenience. The point is that now the metric tensor changes its form a bit:

$$g_{ij} = \begin{pmatrix} 1 & & \\ & r^2 & \\ & & r^2 \sin^2\theta \end{pmatrix} \tag{2.7}$$

However, it should be intuitive that simply switching around some coordinates does not alter the physical world, so that the two objects remain at the same physical distance. Renaming $ds^2 \to ds^2_{\text{cartesian}}$ in (2.3), mathematically this observation translates to:

$$ds^2_{\text{cartesian}} = ds^2_{\text{spherical}} \tag{2.8}$$

The metric tensor is fundamental even when it comes to describing the motion of test particles in an arbitrary space. In fact, let's start with the motion of a free particle in flat space, described by the simple Lagrangian:

$$\mathcal{L} = \frac{1}{2} m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \frac{m}{2} \delta_{ij} \dot{x}^i \dot{x}^j \tag{2.9}$$

Here we introduced a handy notation, called the *Einstein notation*, where repeated indices are summed over, i.e. $\delta_{ij}\dot{x}^i\dot{x}^j \equiv \sum_{ij} \delta_{ij}\dot{x}^i\dot{x}^j$. Critically, this line of reasoning can be extended to any given space that has a metric tensor of the form $g_{ij}(\vec{x})$:

$$\mathcal{L} = \frac{m}{2} g_{ij}(\vec{x}) \dot{x}^i \dot{x}^j \tag{2.10}$$

Let's now try to obtain a general form of the equations of motion, which follow from solving the Euler-Lagrange equations:

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}^i} - \frac{\partial \mathcal{L}}{\partial x^i} = 0 \tag{2.11}$$

We then have:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \dot{x}^i} = m g_{ik} \dot{x}^k \implies \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}^i} &= m \frac{\partial g_{ik}}{\partial x^j} \dot{x}^j \dot{x}^k + m g_{ik} \ddot{x}^k \\ \frac{\partial \mathcal{L}}{\partial x^i} &= \frac{m}{2} \frac{\partial g_{jk}}{\partial x^i} \dot{x}^i \dot{x}^k \end{aligned} \tag{2.12}$$

Putting these together:

$$g_{ik} \ddot{x}^k + \left(\frac{\partial g_{ik}}{\partial x^j} - \frac{1}{2} \frac{\partial g_{jk}}{\partial x^i} \right) \dot{x}^j \dot{x}^k = 0 \tag{2.13}$$

but:

$$\frac{\partial g_{ik}}{\partial x^j} \dot{x}^j \dot{x}^k = \frac{1}{2} \left(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{ij}}{\partial x^k} \right) \dot{x}^j \dot{x}^k \tag{2.14}$$

Therefore (2.13) becomes:

$$g_{ik}\ddot{x}^k + \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{ij}}{\partial x^k} - \frac{\partial g_{jk}}{\partial x^i}\right)\dot{x}^j\dot{x}^k = 0 \tag{2.15}$$

Now we multiply the whole equation by the inverse of the metric tensor g^{-1}, which is defined by raising its indices. Also, by definition of the inverse, g^{-1} satisfies:

$$g_{ij}g^{jk} = \delta^i_k \tag{2.16}$$

Again, for now, don't worry about the positioning of the indices, since it has no real meaning, i.e. we can just assume $\delta^i_j \equiv \delta_{ij}$ (soon we will understand why). However, when we include the dimension of time, upper or lower indices will have their differences. In any case, the equations of motion become:

$$\ddot{x}^i + \Gamma^i_{jk}\dot{x}^j\dot{x}^k = 0 \tag{2.17}$$

where we have introduced the *Christoffel symbols*:

$$\Gamma^i_{jk}(\vec{x}) = \frac{1}{2}g^{il}\left(\frac{\partial g_{lj}}{\partial x^k} + \frac{\partial g_{lk}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^l}\right) \tag{2.18}$$

The $\vec{x}$ dependence is present because $g_{ij} = g_{ij}(\vec{x})$. Equation (2.18) is called the *geodesics equation* and it is fundamental, because if someone hands us an arbitrary metric tensor, we can find the free motion of a test particle simply by solving it (in many cases it is far from simple. Actually, it's very boring).

Now, in general relativity, the direction of time is added, so that the coordinates span 4 dimensions instead of 3. By convention, dimension number 0 is reserved for time, whereas 1,2,3 are the usual spatial dimensions; for this reason, a general 4-vector will be of the form $x^\mu = (ct, \vec{x}) \equiv (ct, x^i)$, where $\mu = 0, 1, 2, 3$. Having said this, everything that we have said before can easily be applied to this new situation. We start with the invariant ds^2:

$$ds^2 = g_{\mu\nu}(x)dx^\mu dx^\nu \tag{2.19}$$

of which the easiest example is flat space, or *Minkowski space*:

$$ds^2 = -c^2dt^2 + \delta_{ij}dx^idx^j \tag{2.20}$$

also referred to as $g_{\mu\nu} = \mathrm{diag}(-1, +1, +1, +1) \equiv (-, +, +, +)$. Given a certain $g_{\mu\nu}$ and two vectors A^μ and B^μ, we can use it to calculate lengths and angles as in (2.5):

$$\begin{aligned} A^2 &\equiv g_{\mu\nu}A^\mu A^\nu = A_\mu A^\mu \\ A \cdot B &\equiv g_{\mu\nu}A^\mu B^\nu = A^\mu B_\mu = A_\mu B^\mu = A^2\ B^2\ \cos\theta \end{aligned} \tag{2.21}$$

where now there is a difference between upper and lower indices, which are mapped from one to the other by the metric tensor:

$$\begin{aligned} A_\mu &= g_{\mu\nu}A^\nu \\ A^\mu &= g^{\mu\nu}A_\nu \end{aligned} \tag{2.22}$$

the inverse of $g_{\mu\nu}$ being defined as before:

$$g^{\mu\nu}g_{\nu\rho} = \delta^\mu_\rho \tag{2.23}$$

Now we are ready to understand why in Euclidian 3D space there is no distinction between upper and lower indices. It all has to do with the shape of the metric tensor, since, following the definition in (2.22), we have $A_i = \delta_{ij}A^j = A^i$. Following the same thread as before, the Lagrangian of a free particle can be written as:

$$\mathcal{L} = \frac{m}{2}g_{\mu\nu}(x)\dot{x}^\mu\dot{x}^\nu \tag{2.24}$$

and the geodesic equation as:

$$\ddot{x}^\mu + \Gamma^\mu_{\nu\rho}\dot{x}^\nu\dot{x}^\rho = 0 \tag{2.25}$$

with:

$$\Gamma^\mu_{\nu\rho} = \frac{1}{2}g^{\mu\lambda}\left(\frac{\partial g_{\rho\lambda}}{\partial x^\nu} + \frac{\partial g_{\nu\lambda}}{\partial x^\rho} - \frac{\partial g_{\nu\rho}}{\partial x^\lambda}\right) \tag{2.26}$$

Finally, any given spacetime $g_{\mu\nu}$ has to satisfy the Einstein equations:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} \tag{2.27}$$

where:

$$\begin{aligned} R &= R^{\mu\nu} R_{\mu\nu} \\ R_{\mu\nu} &= R^{\rho}_{\mu\rho\nu} \\ R^{\rho}_{\sigma\mu\nu} &= \partial_\mu \Gamma^{\rho}_{\nu\sigma} - \partial_\nu \Gamma^{\rho}_{\mu\sigma} + \Gamma^{\rho}_{\mu\lambda}\Gamma^{\lambda}_{\nu\sigma} - \Gamma^{\rho}_{\nu\lambda}\Gamma^{\lambda}_{\mu\sigma} \end{aligned} \tag{2.28}$$

and $T_{\mu\nu}$, the *energy-momentum tensor*, encapsulates all the forms of energy present in that spacetime.

As an aside, we note that in the geodesic equation (2.25) the dots mean a derivative with respect to *proper time* τ:

$$\frac{d^2 x^\mu}{d\tau^2} = -\Gamma^{\mu}_{\nu\rho} \frac{dx^\nu}{d\tau}\frac{dx^\rho}{d\tau} \tag{2.29}$$

The physical interpretation, and definition, of such an object, is the time measured by an observer at rest between any two events A and B. Clearly, in his reference frame, the events happen at the same spatial point, so that the distance ds^2 only regards a distance in time:

$$-c^2 d\tau^2 = ds^2 \tag{2.30}$$

This quantity is an invariant, and any given observer, who parametrizes the path of the observer in motion as $x(\sigma)$, calculates it in the following way:

$$\Delta\tau = \int_{\text{path}} \frac{1}{c}\sqrt{-ds^2} = \int_{\sigma_A}^{\sigma_B} \frac{d\sigma}{c}\sqrt{-g_{\mu\nu}\frac{dx^\mu}{d\sigma}\frac{dx^\nu}{d\sigma}} \tag{2.31}$$

Remember that in general relativity, we should be careful with assigning meaning to coordinates. The coordinate t should not be thought of as "time flowing", just as r is not a "distance from the center". The only physical objects in the theory are invariants, like ds^2, that everybody can agree on. As an example, we can think of the decay of a neutron. When physicists say, "a neutron decays on average in around 11 minutes", that lapse of time is the proper time $\Delta\tau$ of the neutron, that is the time measured in the position of the said particle. Any other observer whizzing by will measure their own *coordinate time* Δt, which itself depends on the speed and angle of the observer's velocity. Clearly, it doesn't make sense to compare the time Δt_1 and Δt_2 of two different observers, since one is not "more real" that the other. Time is in fact only a local concept. What the observers can agree on is that, calculating the proper time with equation (2.31), they will get the same result.

2.2 The Geometry of our Universe

We are now in a position to try to guess what the Universe's metric might look like. Even though this might seem like an impossible task, it turns out that the observations we listed in the first lecture decrease the number of possible alternatives rather drastically. Let's start with the most general metric we can write:

$$ds^2 = -c^2 g_{00}(t,\vec{x})dt^2 + 2g_{0i}(t,\vec{x})dt dx^i + g_{ij}(t,\vec{x})dx^i dx^j \tag{2.32}$$

Now, let's think for a moment about the $g_{00}(t,\vec{x})$ term. Since the Universe is homogeneous, it cannot possibly depend on $\vec{x}$, because in that case observers sitting in different points of space would measure time differently, violating homogeneity itself, so $g_{00} = g_{00}(t)$. Moreover, at this point we could define a new variable $dt' = \sqrt{-g_{00}(t)}dt$, so t is really just up to redefinition. Then, we know that all g_{0i} have to be zero, as a consequence of isotropy, because there cannot be a preferred spatial direction. So for now we can write (2.32) as:

$$ds^2 = -c^2 dt^2 + a^2(t) d\sigma^2 \tag{2.33}$$

where $a(t)$ is called the *scale factor* and it does not depend on r because, if that were the case, each observer would measure a different scale factor, violating homogeneity once again. We have also introduced the spatial metric:

$$d\sigma^2 = \gamma_{ij}(\vec{x}) dx^i dx^j \tag{2.34}$$

Now, how do we find γ_{ij}? It is very convenient that homogeneous and isotropic spaces must have constant intrinsic curvature (that means constant R as introduced in (2.28)), and in general relativity there exist only 3 options: the curvature can be zero, positive or negative. These correspond to flat, spherical and hyperbolic spaces respectively, depicted here below:

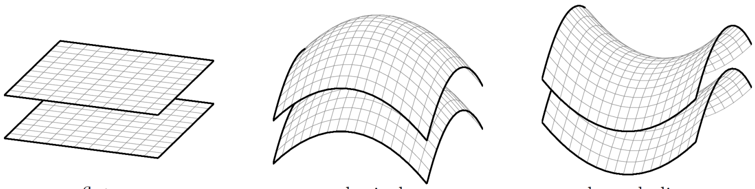

Figure 10: The only three spaces with constant curvature

Let's determine the metric for each of these cases:

- Flat space: this is simply Euclidian 3D space, the metric of which we defined before (2.2):

$$d\sigma^2 = d\vec{x}^2 = \delta_{ij} dx^i dx^j \tag{2.35}$$

  This type of space has constant curvature equal to 0.

- Spherical space: a 3-sphere with radius A is defined as an embedding in 4D Euclidian space:

$$d\sigma^2 = d\vec{x}^2 + du^2 \quad (x^1)^2 + (x^2)^2 + (x^3)^2 + u^2 = A^2 \tag{2.36}$$

  The 3-sphere has a positive constant curvature. It may be not that clear why (2.36) represents the metric on the sphere. To verify this, it is easier to operate in 2 spatial dimensions instead of 3, so that we start with the embedding condition:

$$(x^1)^2 + (x^2)^2 + (x^3)^2 = \tilde{A}^2 \tag{2.37}$$

  and the metric:

$$d\tilde{\sigma}^2 = d(x^1)^2 + d(x^2)^2 + d(x^3)^2 \tag{2.38}$$

  Since, for instance, the coordinate x^3 actually depends on the other two, we can rewrite $d\tilde{\sigma}^2$ as:

$$d\tilde{\sigma}^2 = d(x^1)^2 + d(x^2)^2 + \frac{(x^1 dx^1 + x^2 dx^2)^2}{\tilde{A}^2 - (x^1)^2 - (x^2)^2} \tag{2.39}$$

  Now introduce polar coordinates:

$$x^1 = \tilde{A} \sin\theta \cos\phi \quad x^2 = \tilde{A} \sin\theta \sin\phi \implies x^3 = \tilde{A} \cos\theta \tag{2.40}$$

  in terms of which we have:

$$d\tilde{\sigma}^2 = \tilde{A}^2 (d\theta^2 + \sin^2\theta d\phi^2) \tag{2.41}$$

  which is precisely the metric on the 2-sphere.

- Hyperbolic space: a 3-hyperboloid with curvature A is defined as an embedding in Minkowski space:

$$d\sigma^2 = d\vec{x}^2 - du^2 \quad (x^1)^2 + (x^2)^2 + (x^3)^2 - u^2 = -A^2 \tag{2.42}$$

  The 3-hyperboloid has constant negative curvature.

We can now combine the metric of the spherical and hyperbolic spaces like this:

$$d\sigma^2 = d\vec{x}^2 \pm du^2 \quad \vec{x}^2 \pm u^2 = \pm A^2 \tag{2.43}$$

Now, the embedding condition $\vec{x}^2 \pm u^2 = \pm A^2$ implies $udu = \mp \vec{x} \cdot d\vec{x}$, which we can substitute in the metric $d\sigma^2$:

$$d\sigma^2 = d\vec{x}^2 \pm \frac{(\vec{x} \cdot d\vec{x})^2}{A^2 \mp \vec{x}^2} \tag{2.44}$$

To combine this metric with flat space (2.35) we write all three of them in one go as:

$$d\sigma^2 = d\vec{x}^2 + k \frac{(\vec{x} \cdot d\vec{x})^2}{A^2 - k\vec{x}^2} \qquad k = \begin{cases} 0 & \text{flat} \\ +1 & \text{spherical} \\ -1 & \text{hyperbolic} \end{cases} \tag{2.45}$$

In general, it is convenient to write the metric in spherical polar coordinates (r, ϕ, θ) using the coordinate change:

$$\begin{aligned} d\vec{x}^2 &= dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \\ \vec{x} \cdot d\vec{x} &= r dr \\ \vec{x}^2 &= r^2 \end{aligned} \tag{2.46}$$

So finally (2.45) becomes:

$$d\sigma^2 = \frac{dr^2}{1 - kr^2/A^2} + r^2 d\Omega^2 \tag{2.47}$$

where $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$ is the 2-sphere metric. Putting this in (2.33) we find:

$$ds^2 = -c^2dt^2 + a^2(t)\left(\frac{dr^2}{1 - kr^2/A^2} + r^2 d\Omega^2\right) \tag{2.48}$$

However, by redefining $r/A = r'$, we have:

$$ds^2 = -c^2dt^2 + a^2(t)A^2\left(\frac{dr^2}{1 - kr^2} + r^2 d\Omega^2\right) \tag{2.49}$$

so that we can just decide to include A in the definition of $a(t)$, getting rid of it altogether. The following is then the metric of our Universe, or the so-called *Friedmann-Robertson-Walker metric* (FRW):

$$ds^2 = -c^2dt^2 + a^2(t)\left(\frac{dr^2}{1 - kr^2} + r^2 d\Omega^2\right) \tag{2.50}$$

This is the starting point in the study of cosmology, and in the next lecture we will analyze it in detail. We will understand what those coordinates mean and what can be deduced from them, but also how, by solving the Einstein equations, this metric governs the dynamical evolution of the Universe.

3 The Properties of the FRW Metric

Remember that in last lecture we introduced the *FRW metric*, which is supposed to represent the Universe, at least at large enough scales where it is homogeneous and isotropic:

$$ds^2 = -c^2dt^2 + a^2(t)\left(\frac{dr^2}{1 - kr^2} + r^2 d\Omega^2\right) \tag{3.1}$$

where we recall that $d\Omega^2 \equiv d\theta^2 + \sin^2\theta d\phi^2$ is the 2-sphere metric, and k controls the curvature(geometry) of the spatial slices:

$$k = \begin{cases} 0 & \text{Euclidian} \\ +1 & \text{Spherical} \\ -1 & \text{Hyperbolical} \end{cases} \tag{3.2}$$

This metric (3.1) is the crux of cosmology, and the goal of this lecture and the following is to understand it better.

3.1 What do the coordinates stand for?

As we have said, it is in general not obvious how to give meaning to the coordinates in General Relativity. However, in this case, the coordinates (r, θ, ϕ) are called *comoving coordinates*, and we can actually try to understand what that means and the reason for that name with the following example.

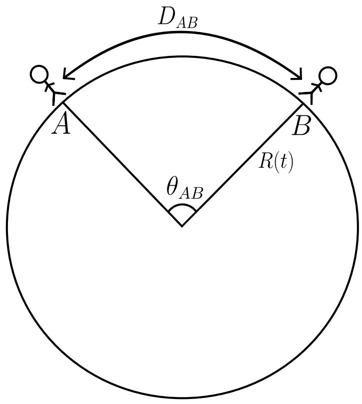

Figure 11: Two observers sitting on a circle

Consider a circle that is expanding, with a radius $R(t)$, and two observers fixed at positions A and B as in the figure on the left. These observers, then, are at a fixed *comoving* distance θ_{AB}, but their *physical* distance D_{AB} is instead increasing. The two are related by:

$$D_{AB}(t) = R(t) \times \theta_{AB} \tag{3.3}$$

We can actually go a bit further and ask what the *receding velocity* is, i.e. how fast A sees B going away from him or viceversa. Clearly, we just have to take a derivative:

$$\dot{D}_{AB}(t) = \dot{R}(t) \times \theta_{AB} = \frac{\dot{R}(t)}{R(t)} R(t) \times \theta_{AB} \equiv H(t) \times D_{AB} \tag{3.4}$$

We recognize this as Hubble's law! That is, the two observers see each other receding with a velocity that is proportional to their distance. We also have introduced the *Hubble constant* (even though it is not a constant, history has decided on this label) which, if evaluated at present time t_0, becomes H_0, the parameter we have introduced in the first lecture when we were talking about Hubble's discoveries.
Now, in this simplified model of the Universe, we can also understand how every observer sees everybody else receding with a velocity according to Hubble's law, like in the figure on the right below.

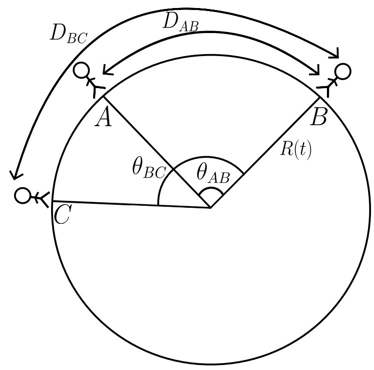

Figure 12: Multiple observers sitting on a circle

Clearly, A sees both observers B and C receding, but the same can be said from the point of view of observer C. This resembles a lot the picture of our Universe we have laid out in the first lecture, where every galaxy seems to be moving away from us independently of where we look. Also, homogeneity implies that this must be true for every galaxy, not just for us. We know, however, that our Universe is neither 2 dimensional nor a circle, so we need to generalize the discussion for the geometries that appear in the FRW metric. Consider a galaxy at a coordinate distance $\vec{r}$ and a scale factor $a(t)$, then its physical distance is $\vec{r}_{\rm phys} = a(t) \times \vec{r}$ and its velocity is:

$$\vec{v}_{\rm phys} = \dot{a}(t)\vec{r} + a(t)\dot{\vec{r}} \equiv H(t)\vec{r}_{\rm phys} + \vec{v}_{\rm pec} \tag{3.5}$$

We can see that the velocity has two contributions. First is the *Hubble flow* which, as we said before, results simply from the expansion of the Universe, and second is the *peculiar velocity*, because the galaxy might have its own local motion that may be due, for example, to the gravitational attraction of nearby galaxies.

We have gone through this example because it is applicable to our case, since the proper distance between two objects that are spatially separated by $d\sigma$ can be read from (3.1) by putting $dt = 0$:

$$d\ell(t) = a(t)d\sigma \implies d\dot{\ell}(t) = H(t)d\sigma \tag{3.6}$$

which is just the Hubble law, as before.
Basically, we have learned that the coordinates (r, θ, ϕ) are comoving, i.e. they are useful for describing observers moving with the Hubble flow. Moreover, this tells us that even the time t is the proper time of an observer moving with the Hubble flow itself, because if we place him at fixed comoving coordinates ($dr = d\theta = d\phi = 0$) we get from the metric:

$$-c^2 d\tau^2 = -c^2 dt^2 \tag{3.7}$$

Finally, let us think about the scale factor itself. We have seen that it is basically a conversion factor between physical and coordinate distances at a certain time t. Intuitively, it doesn't really matter what the actual values of $a(t)$ are, since the only important things are ratios, like $a(t_1)/a(t_2)$ at two different times. This is reflected in the fact that the metric (3.1) has a rescaling symmetry (as we have already seen with the overall factor A in the previous lecture):

$$r \to r/\lambda, \quad a \to \lambda a \tag{3.8}$$

This freedom allows us to set the scale factor equal to one at present time $a(t_0) = 1$, which is just a convenient value.

3.2 New coordinates and conformal time

In general, it can be a bit inconvenient to have a complicated g_{rr} component in the FRW metric (3.1), so that we define a new radial coordinate to get rid of it:

$$d\chi = \frac{dr}{\sqrt{1 - kr^2}} \tag{3.9}$$

With this change of coordinates, the FRW metric becomes:

$$ds^2 = -c^2 dt^2 + a^2(t)(d\chi^2 + S_k^2(\chi)d\Omega^2) \tag{3.10}$$

where:

$$S_k(\chi) \equiv \begin{cases} \sinh\chi & k = -1 \\ \chi & k = 0 \\ \sin\chi & k = +1 \end{cases} \tag{3.11}$$

Note how for a flat Universe ($k = 0$), there is no distinction between r and χ.
Finally, it is very convenient to introduce the notion of *conformal time*, which is just another time coordinate defined by:

$$d\eta = \frac{dt}{a(t)} \tag{3.12}$$

With this new time the metric (3.13) becomes:

$$ds^2 = a^2(\eta)(-c^2 d\eta^2 + (d\chi^2 + S_k^2(\chi)d\Omega^2)) \tag{3.13}$$

The obvious advantage of conformal time is that light rays ($ds^2 = 0$) now move at 45° in a spacetime diagram. However, it is a rather important concept altogether, as we will discover a bit later on in the lectures.

3.3 Geodesics, or the motion of objects

In the previous lecture we have seen how, given a certain metric, we can calculate how particles move by solving the geodesics equation. In this section we will do so for both massive and massless particles.
But first, since the geodesics equation requires the Christoffel's symbols:

$$\Gamma^\mu_{\alpha\beta} = \frac{1}{2}g^{\mu\lambda}(\partial_\alpha g_{\beta\lambda} + \partial_\beta g_{\alpha\gamma} - \partial_\gamma g_{\alpha\beta}), \quad \partial_\alpha \equiv \partial/\partial x^\alpha \tag{3.14}$$

we might as well calculate them now. Let's write the FRW metric as:

$$ds^2 = -c^2 dt^2 + g_{ij}dx^i dx^j = -c^2 dt^2 + a^2(t)\gamma_{ij}dx^i dx^j \tag{3.15}$$

Here we will calculate the Christoffel symbol $\Gamma^0_{\alpha\beta}$ and just list the other ones, since they are only a matter of calculations. Then:

$$\Gamma^0_{\alpha\beta} = \frac{1}{2} g^{0\lambda}(\partial_\alpha g_{\beta\lambda} + \partial_\beta g_{\alpha\lambda} - \partial_\lambda g_{\alpha\beta}) \tag{3.16}$$

Now, since $g^{0\lambda}$ is only nonzero if $\lambda = 0$ and $g_{00} = -1$, we have:

$$\Gamma^0_{\alpha\beta} = -\frac{1}{2}(\partial_\alpha g_{\beta 0} + \partial_\beta g_{\alpha 0} - \partial_0 g_{\alpha\beta}) \tag{3.17}$$

But $g_{00} = \text{const}$ and $g_{0i} = 0$, so that the first two terms vanish, leaving us with:

$$\Gamma^0_{\alpha\beta} = \frac{1}{2}\partial_0 g_{\alpha\beta} \tag{3.18}$$

clearly the temporal derivative is only nonzero for the spatial part of the metric, which contains the scale factor $a(t)$:

$$\Gamma^0_{ij} = c^{-1} a\dot{a}\gamma_{ij} \tag{3.19}$$

In all, the Christoffel's read:

$$\begin{aligned} \Gamma^\mu_{00} &= \Gamma^0_{\alpha 0} = \Gamma^\mu_{\alpha\beta} = 0 \\ \Gamma^0_{ij} &= c^{-1} a\dot{a}\gamma_{ij} = c^{-1}\frac{\dot{a}}{a} g_{ij} \\ \Gamma^i_{0j} &= c^{-1}\frac{\dot{a}}{a}\delta^i_j \\ \Gamma^i_{jk} &= \frac{1}{2}\gamma^{il}(\partial_j\gamma_{kl} + \partial_k\gamma_{jl} - \partial_l\gamma_{jk}) \end{aligned} \tag{3.20}$$

All the other ones are related to these by symmetries, because $\Gamma^\mu_{\alpha\beta} = \Gamma^\mu_{\beta\alpha}$ by virtue of the symmetry of the metric tensor $g_{\alpha\beta} = g_{\beta\alpha}$.
Having calculated them, we note one final thing. Consider the geodesics equation:

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} = 0 \tag{3.21}$$

Using the four momentum of the particle we are considering $P^\mu = m\ dx^\mu/d\tau$, and the fact that:

$$\frac{d}{d\tau}P^\mu(x^\alpha(\tau)) = \frac{dx^\alpha}{d\tau}\frac{\partial P^\mu}{\partial x^\alpha} = \frac{P^\mu}{m}\frac{\partial P^\mu}{\partial x^\alpha} \tag{3.22}$$

we can rewrite the geodesics equation as:

$$P^\alpha\frac{\partial P^\mu}{\partial x^\alpha} + \Gamma^\mu_{\alpha\beta}P^\alpha P^\beta = 0 \tag{3.23}$$

This equation is the one we'll use below, since it is pretty handy. However, there is one subtleness that we decided to just gloss over. In fact, for the whole discussion we have used proper time τ as a parameter, but this is exactly zero for massless particles, so it really can't be that useful for them. It turns out that we can just use (3.23) even for photons in the following way. Technically, no one stops us from using a parameter λ to parametrise their path in spacetime, and once we have chosen such a parameter, there is only one continuation of it such that $x(\lambda)$ satisfies the geodesics equation, provided that we interpret $P^\mu \equiv dx^\mu/d\lambda$ as the four momentum of the massless particle.
At last, we can move on to calculate how massive and massless particles move in spacetime. We work for the moment with $c = 1$.

- Massive particles: let's consider the $\mu = 0$ component of (3.23):

  $$P^0\frac{dP^0}{dt} + \frac{\dot{a}}{a}g_{ij}P^iP^j = 0 \tag{3.24}$$

  where we have used the Christoffel's symbols and $\partial_i P^0 = 0$ because of homogeneity. We also introduce the *physical momentum* $p^2 \equiv g_{ij}P^iP^j$ which obeys the on-shell relation:

  $$-(P^0)^2 + p^2 = -m^2 \tag{3.25}$$

  Differentiating (3.25) we get $c^2P^0dP^0 = pdp$, which, substituted in (3.24) gives:

  $$p\frac{dp}{dt} + \frac{\dot{a}}{a}p^2 = 0 \tag{3.26}$$

Finally, rearranging we have:

$$\frac{\dot{p}}{p} = -\frac{\dot{a}}{a} \tag{3.27}$$

This differential equation has a solution $p \sim a^{-1}$, meaning that the particles experience a *drag* given by the expansion of the Universe. This also reveals to us why the comoving coordinates are a natural reference frame, since we have:

$$p = \frac{mv}{\sqrt{1-v^2}} \sim \frac{1}{a} \tag{3.28}$$

where $v^i = dx^i/dt$ is the peculiar velocity of the particles (remember that the peculiar velocity of an object is the velocity relative to the comoving frame, i.e. how the object deviates from perfect expansion proportional to $a(t)$) and $v^2 = g_{ij}v^iv^j$ as with the momentum. The first equality in (3.28) is just the usual special relativity relationship $dx^\mu/d\tau = U^\mu = (U^0, U^i) = (\gamma, \gamma v^i)$ with γ the local Lorentz factor. Now, consider a particle with initial physical (peculiar) velocity $v(t_1) \equiv v_1$. At a later time t_2 it will have a velocity:

$$v_2 = v_1 \frac{a(t_1)}{a(t_2)} \tag{3.29}$$

Since the Universe is expanding, $a(t)$ is growing, so that $v_2 < v_1$, i.e. even if an observer has a nonzero initial velocity, he will come to rest in the comoving frame. This, after a moment's thought, should be intuitive. After a while the Universe expands so much (if it will do so for a long time), that the observer's velocity will be negligible.

- Massless particles: we can take (3.24) with the new constraint $m = 0$:

$$-(P^0)^2 + p^2 = 0 \tag{3.30}$$

Using $P^0 = E$, we find:

$$\frac{\dot{E}}{E} = \frac{\dot{a}}{a} \tag{3.31}$$

which again implies that the energy scales as $E \sim a^{-1}$ for photons. This has the physical interpretations that their wavelength gets stretch proportionally to their scale factor, giving rise to such a decrease in energy (given that $E \sim \lambda^{-1}$). Interestingly, this is also the phenomenon at the base of cosmological redshift, which we will touch on in the next section

3.4 Cosmological redshift

An important concept in cosmology is the *cosmological redshift*, since what we know of the properties of the Universe is a result of the observation of distant objects. In order to analyze these correctly, we need to know how light propagates as the Universe evolves in time. Consider the scenario in which a galaxy that sends light signals to us positioned at (r_e, θ_e, ϕ_e). Since the Universe is isotropic and homogeneous we can freely choose the propagation of light to follow a straight line $\theta = \phi =$ const along null geodesics $ds^2 = 0$ in the parametrization (3.13):

$$0 = a(\eta)(-c^2 d\eta^2 + d\chi^2) \to \Delta\chi(\eta) = \pm c\Delta\eta \tag{3.32}$$

where the $\pm$ corresponds to outgoing and ingoing photons. If a wave crest is emitted at time t_e from the galaxy, the time t_0 when it reaches us is given by:

$$c^2(\eta(t_0) - \eta(t_e)) = \chi(r_e) - \chi(0) = \chi(r_e) \tag{3.33}$$

Since the comoving distance of the galaxy does not change with time, a successive wave crest emitted shortly after, at $t_e + \delta t_e$, reaches the origin at time $t_0 + \delta t_0$:

$$c^2(\eta(t_0 + \delta t_0) - \eta(t_e + \delta t_e)) = \chi(r_e) \tag{3.34}$$

Combining these two equations and noticing that for real applications we can approximate $\delta t_e \ll t_e$ and $\delta t_0 \ll t_0$, we find:

$$\frac{\delta t_0}{a(t_0)} = \frac{\delta t_e}{a(t_e)} \tag{3.35}$$

where we have used the definition of conformal time. Thus the period of the wave, hence its wavelength, increases proportionally to the scale factor:

$$\frac{\lambda_0}{\lambda_e} = \frac{\nu_e}{\nu_0} = \frac{\delta t_0}{\delta t_e} = \frac{1}{a(t_e)} \tag{3.36}$$

where we remember that we have the freedom to choose $a(t_0) = 1$. Clearly, this is the same thing we found before with $E \sim a^{-1}$, but now in a more concise way. Defining the relative change of wavelength as the *redshift* z, we have:

$$a(t_e) = \frac{1}{1+z} \tag{3.37}$$

A galaxy at redshift $z = 1$ thus emitted the observed light when the Universe was half its current size, a galaxy at redshift 2 when the Universe was one third its size, and so on.
Interestingly, this ties to the observation Hubble was making in the 1920s. To see this, consider the expansion of the wave factor for nearby sources ($z \ll 1$) around $t = t_0$:

$$a(t_e) = 1 + (t_e - t_0)H_0 + ... \tag{3.38}$$

where $t_e - t_0$ is the *look-back time*. Using our new definition of the redshift (3.37) we have $z = H_0(t_0 - t_e) + ...$, but for close objects we can just use $d/c = t_0 - t_e$. Crucially, we also know that the non-relativistic limit of the formula for a longitudinal Doppler shift is:

$$z = \frac{\lambda_0}{\lambda_e} - 1 = \sqrt{\frac{1+\beta}{1-\beta}} - 1 = \sqrt{\frac{1+v/c}{1-v/c}} - 1 \approx \frac{v}{c} \tag{3.39}$$

Putting all these thigs together, we find Hubble law again:

$$v \approx cz \approx H_0 d \tag{3.40}$$

Basically, what Hubble found was a cosmological redshift that masqueraded as a Doppler redshift, given that he was investigating close-by sources. For sources that are very distant, we actually need to pay more attention to how we even define distances, as we will see in the next lecture.
Let's wrap up this discussion now. First of all, the cosmological redshift has nothing to do with the Doppler redshift, for a number of reasons. Note, in fact, that even galaxies that are moving with the Hubble flow will experience a redshift, so velocity has nothing to do with it. Moreover, even if the galaxies had some peculiar velocity, in General Relativity there is no notion of comparing velocities with each other anyway. Having said this, the two notions can be used interchangeably for close sources that are moving "slowly", which is what happened with the Hubble law.
Finally, today astronomers use redshift z to measure distances, instead of parsecs. To convert from a redshift to a distance we actually have to be careful, as distances in cosmology are a subtle concept, but for close enough sources we can just use, as we have said, $d = cz/H_0$.

3.5 Horizons in the Universe

Due to the speed of light being finite, there are regions of the Universe that are and will be unaccessible to us, even if we wait an infinite amount of time. These regions are defined by *horizons*. In cosmology, there are three of these boundaries, which define the causal structure of our Universe.
Before we continue, let's recall the FRW metric:

$$\begin{aligned} ds^2 &= -c^2dt^2 + a^2(t)\left(\frac{dr^2}{1-kr^2} + r^2 d\Omega^2\right) \\ &= ds^2 = a^2(\eta)(-c^2d\eta^2 + (d\chi^2 + S_k^2(\chi)d\Omega^2)) \end{aligned} \tag{3.41}$$

with the conformal time defined by:

$$d\eta = \frac{dt}{a(t)} \tag{3.42}$$

and the spatial geometry:

$$S_k(\chi) \equiv \begin{cases} \sinh\chi & k=-1 \\ \chi & k=0 \\ \sin\chi & k=+1 \end{cases} \tag{3.43}$$

The first of these horizons is what is called the *particle horizon* or the *cosmological horizon*. It is the maximum distance light could have traveled from the "beginning" of the Universe, i.e. the observable Universe. To be precise, we could ask the following question: given a comoving observer at coordinates $(\chi_0, \theta_0, \phi_0)$, for what values of (χ, θ, ϕ) would a light signal emitted at time $t = 0$ (the "beginning" of the Universe) reach the observer at time t? Given the FRW metric, we can answer this question rather easily. First of all, let's place the receiving observer at a convenient point in space (given our freedom of choice granted to us by homogeneity), setting $\chi_0 = 0$ (so the observer is us, Earth). Then, a light signal satisfies the geodesics equation with $ds^2 = 0$ and $d\phi = d\theta = 0$ (because the light moves on great circles). If this light signal reaches us at time t, then the comoving distance from the observer (who emitted the photon at the "beginning" of the Universe) is given by:

$$d_{h,comoving}(\eta) = \chi = \eta = c\int_0^t \frac{dt'}{a(t')} \tag{3.44}$$

In other words, the value of conformal time now η_0 is the observable Universe today. We can also define the *proper particle horizon*:

$$d_h(\eta) = a(\eta)c\int_0^t \frac{dt}{a(t)} \tag{3.45}$$

Notice that we have implicitly assumed $t = 0$ to be the beginning of the Universe. This notion, however, is not really defined, since we don't actually know what happened really far in the past above 100 GeV, as remarked two lectures ago. To put it another way, the equation (3.44) should really be written as:

$$\eta = c\int_0^{t_{100\text{ Gev}}} \frac{dt'}{a(t')} + c\int_{t_{100\text{ Gev}}}^{t} \frac{dt'}{a(t')} \tag{3.46}$$

Because the Universe was really small and really hot, we have no reason to trust classical general relativity as such high energies, so that the definition of "observable Universe" may make not much sense at all. In fact, the name "observable Universe" itself is usually meant to be the distance light could have traveled from some time $t_{100\text{ Gev}}$, where extrapolating the classical theory is still sensible.

Another type of horizon is the *event horizon*, which differs from the particle horizon, in that it is the maximum distance light, emitted at given time t, could travel to the infinite future. Again, the FRW metric allows us to easily calculate this distance, proceeding in the following way. Given a comoving observer at coordinates $(0, \theta_0, \phi_0)$ (again, homogeneity permits to set $\chi_0 = 0$) that emits a light signal at time t, the distance it travels up to $t = \infty$ is:

$$d_{e,comoving}(\eta) = \int_\eta^{\eta_f} d\eta = \eta_f - \eta = c\int_t^\infty \frac{dt'}{a(t')} \tag{3.47}$$

with its *proper* counterpart being:

$$d_e(\eta) = a(\eta)c\int_t^\infty \frac{dt'}{a(t')} \tag{3.48}$$

After a moment's thought, one notices that this definition really is just the "reverse" of the one for the particle horizon, in the sense that if we exchanged the beginning with the end of the Universe, the two concepts would switch. This is shown more clearly in the following figure, which depicts the two horizons:

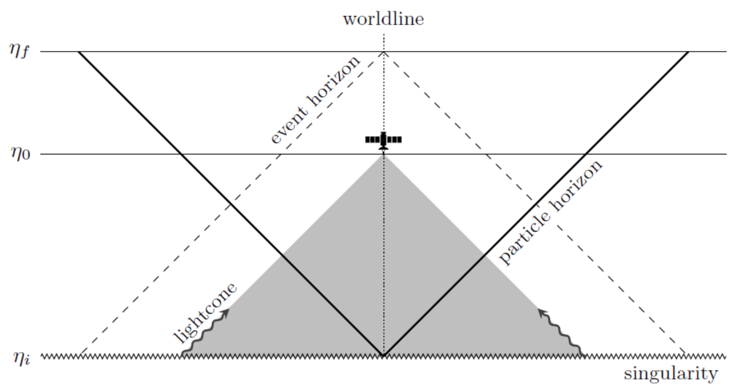

Figure 13: The particle and the event horizon, in conformal time.

We have, however, just glossed over some subtleties. For the moment, in fact, we have defined $\eta_i = \eta(t = 0) = 0$ and $\eta_f = \eta(t = \infty) \neq \infty$ (infinite time does not mean infinite conformal time). These two identities will be derived later on, where the dynamics of the Universe are discussed in detail. Moreover, notice the specularity of the definitions for $d_e(\eta)$ and $d_h(\eta)$: just as the particle horizon is sensitive to the initial conditions (the integral (3.45) goes down to $t = 0$), the event horizon is sensitive to the "final conditions" (the integral (3.47) goes until $t = \infty$). Practically, this is because we don't know what the fate of the Universe will be, this time not because of general relativity itself, but because the entity that controls its future is dark energy, whose nature is still really unresolved.

The final concept we linger on is the *Hubble horizon*. This is a confusing concept, for starters because it's not actually an horizon, so let's define it from the basics. As we have seen, Hubble's law dictates that the velocity of galaxies on the Hubble flow is proportional to their proper distance:

$$v_{\text{phys}} = H(t) r_{\text{phys}} \tag{3.49}$$

Replacing v_{phys} with the speed of light c, we find a proper distance d_{HR} (not comoving) above which galaxies (or objects in general) recede with a speed that's greater than that of light:

$$d_{HR} = \frac{c}{H(t)} \tag{3.50}$$

The *comoving Hubble horizon* is instead:

$$d_{HR} = \frac{c}{a(t)H(t)} \tag{3.51}$$

This is not in contrast with special relativity, as this velocity is not measured in any inertial frame. No observer is overtaking a light beam and locally, in the galaxy's position, observers measure the speed of light as c. Taking d_{HR}^3, we get what is called the *Hubble volume* or the *Hubble sphere*.
On the one hand, the Hubble sphere is not a measure of causality, because if two objects are separated by a distance greater than a Hubble length, it would be still possible for them to communicate. In fact, this is the case if the Universe is expanding and $\ddot{a} < 0$, i.e it is also *decelerating*, meaning that the Hubble sphere is actually increasing. After a while, the Hubble sphere will catch up with the light ray, which in turn will enter a zone where space is expanding slower than the speed of light, making it possible to freely flow to destination. If instead the Universe was expanding in an accelerating fashion, the Hubble sphere would decrease, and two objects separated by more than a Hubble length would forever be out of causal contact.
On the other hand, this discussion implies that the Hubble horizon has got at least something to do with causality, despite not really being an horizon. We can see how it is related to the particle horizon by rearranging the integral (3.44) starting at some time t_i:

$$d_h(\eta) = \int_{t_i}^{t} \frac{dt}{a} = \int_{a_i}^{a} \frac{da}{a\dot{a}} = \int_{\log a_i}^{\log a} (aH)^{-1} d\log a \tag{3.52}$$

Basically, the particle horizon is the logarithmic integral of the comoving Hubble horizon.

3.6 Distances in cosmology

The measure of distances in cosmology is obviously fundamental for our understanding of the Universe's properties. There are, however, a lot of subtleties that one needs to consider when trying to formulate rigorous definitions.

3.6.1 Luminosity distance

First, let's review Hubble's argument briefly, as it is really the crux of the whole topic. By looking at relatively close galaxies, Hubble found them to be slightly redshifted. At the time, the Universe was thought to be static, so he interpreted the redshift as being caused by the peculiar motion of galaxies. Estimating the velocity of galaxies as cz (relativistic redshift) and measuring their distances, he found that the two were proportional, i.e. $v = cz = H_0 d$.
We have seen that, by interpreting the galaxies' movement as caused by an expansion of spacetime, the FRW metric actually predicts Hubble's law, as we did in (3.6). There is a problem here though. Hubble couldn't possibly have measured the proper distance, as it is *not* an observable, because it is the distance between two objects at a *fixed time* (in fact, we derived Hubble's law from the FRW metric by setting $dt = 0$), i.e. the distance someone would measure by taking a snapshot of the Universe at a certain time. On the other hand, we have also rightly seen how Hubble's law is just an approximation for very close objects. So what was Hubble actually measuring? Clearly, there is something deeper going on here.

A useful way of measuring distances in cosmology uses the so-called *standard candles*. These are objects of known *intrinsic luminosity*, so that, by measuring their *observed luminosity*, we can infer their distance, which is called the *luminosity distance*.
Hubble used *Cepheids* to discover his law. These are stars whose brightnesses vary periodically, with periods being a known function of intrinsic luminosity of the stars themselves. Therefore, by measuring their periods, Hubble could infer their distance rather easily.
Let's formalize this argument. Let's say we have identified a certain source with a known intrinsic luminosity L. The observed flux F (energy per unit time per unit area) can then be used to deduce the luminosity distance. Consider a source at redshift z, with comoving distance being:

$$\chi(z) = c\int_{t_1}^{t_0} \frac{dt}{a(t)} = c\int_0^z \frac{dz}{H(z)} \tag{3.53}$$

We assume that the source emits light isotropically, so that in a Euclidian static space, the energy would be spread on the surface of a sphere of radius $4\pi\chi^2$ around the source. Then, the relation between the absolute luminosity and the observed flux is:

$$F = \frac{L}{4\pi\chi^2} \quad \text{static Euclidian space} \tag{3.54}$$

In a more complex expanding spacetime, this formula is modified in three fundamental ways:

- First, the radius of the sphere is not χ, but rather $S_k(\chi)$, as we have written in the FRW metric. The sphere surface is then $4\pi S_k(\chi)^2$. Furthermore, when light reaches Earth at time t_0, the sphere has an actual area stretched by $a^2(t_0)$, because of expansion. In total, the denominator of (3.54) becomes:

$$4\pi a^2(t_0) S_k(\chi)^2 \tag{3.55}$$

- The photons reach Earth with a decreased rate, again because of expansion. This effect reduced the observed flux F in (3.54) by a factor of $a(t_1)/a(t_0) = 1/(1+z)$.
- The observed photons lose an energy proportional to $1/(1+z)$ during their trip to Earth, reducing the observed flux once again.

For these reasons, the correct formula relating observed flux and the intrinsic luminosity of a source at redshift z is (remember that $a(t_0) = 1$):

$$F = \frac{L}{4\pi S_k(\chi)^2(1+z)^2} \equiv \frac{L}{4\pi d_L^2(z)} \tag{3.56}$$

where we have defined the luminosity distance as:

$$d_L(z) = (1+z)S_k(\chi(z)) \tag{3.57}$$

Now, we can actually find a perturbative expression for $d_L(z)$, by finding an expansion for $\chi(z)$. We first expand $a(t)$ for $z \ll 1$:

$$a(t) = 1 + H_0(t-t_0) - \frac{1}{2}q_0 H_0^2 (t-t_0)^2 + ... \tag{3.58}$$

where we have defined the *deceleration parameter* at present time q_0:

$$q_0 \equiv -\frac{\ddot{a}}{aH^2}\Big|_{t_0} \tag{3.59}$$

Given the definition of cosmological redshift, we have that:

$$z = \frac{1}{a(t_1)} - 1 = H_0(t_0 - t_1) + \frac{1}{2}(2+q_0)H_0^2(t_0-t_1)^2 + ... \tag{3.60}$$

By inverting it:

$$H_0(t_0 - t_1) = z - \frac{1}{2}(2+q_0)z^2 + ... \tag{3.61}$$

We can then rewrite the comoving distance making use of (3.58) and (3.61):

$$\begin{aligned} \chi(z) = c\int_{t_1}^{t_0} \frac{dt}{a(t)} &= c\int_{t_1}^{t_0} dt(1 - H_0(t-t_0) + ...) \\ &= c(t_0 - t_1) + \frac{1}{2}\frac{H_0}{c}c^2(t_1 - t_0)^2 + ... \\ &= \frac{c}{H_0}\left(z - \frac{1}{2}(1+q_0)z^2 + ...\right) \end{aligned} \tag{3.62}$$

Knowing the geometry of the Universe, by computing $S_k(\chi(z))$ with (3.62), we can estimate the luminosity distance of a source at redshift z. For example, if the Universe is spatially flat $k=0$, then $S_k(\chi) = \chi$, and:

$$d_L(z) = \frac{c}{H_0}\left(z + \frac{1}{2}(1-q_0)z^2 + ...\right) \tag{3.63}$$

We see that the first order approximation is $d_L(z) \simeq cz/H_0$, which is exactly Hubble's law. Now, however, we can understand it better. When doing the expansion in the last lecture, we found $d = cz/H_0$, so we can now identify that d with d_L. Not only that, but our expression (3.63) is only true for Euclidian space. We avoided this fact in the previous lecture by assuming that $d/c = t_0 - t_e$, which is exactly the relation for Euclidian space. In general, however, the Hubble's law will be true, at first order, for all geometries, since the curvature of space will not be a factor for nearby sources. Finally, remember how we found a *non perturbative* Hubble law in (3.6). It happens that, for nearby objects, light takes a relatively short time to travel, so that the luminosity distance can roughly be taken as an instantaneous measure, making it very similar to the proper distance. In other words, Hubble's law is only $d_L(z) \simeq cz/H_0$ (for nearby sources); the relationship $\dot{d}_{proper} = H(t)d_{proper}$ is, on the other hand, a perfect relationship. However, for close sources, light takes a short amount of time to travel, making the two concepts interchangeable.

Summarizing, given an observed flux of photons from a certain source on our telescope, if we happen to precisely know its intrinsic luminosity, we could deduce its luminosity distance using (3.56). The important question here clearly is: how would we know the intrinsic luminosity of a source? As we have said above, in general cosmologists use standard candles, objects for which empirical relations between their properties and their intrinsic luminosities are known. Hubble used Cepheids, but they only work for relatively close sources. Instead, for more distant ones, we need brighter standard candles. Remarkably, type Ia supernovae are exactly the right candidates. These objects are white dwarfs (degenerate stars that are made of oxygen or carbon, for which the electron degeneracy pressure is keeping the star from collapsing on itself) in a binary system. It turns out that white dwarfs have a critical mass, the *Chandraseckhar mass*, of about 1.4 $M_\odot$, above which the star will collapse and explode, creating a supernova event. This critical mass could be reached, for instance, if the white dwarf accretes matter from the companion in the binary. The crucial point here is that the explosion will happen at the Chandraseckhar mass independently of where the binary system is located in the Universe. Therefore, because the same physical process underlies the explosions, the intrinsic luminosities of type Ia supernovae are well known.

3.6.2 Angular diameter distance

An alternative way to measure the distance of a given source is by measuring its angular amplitude. The angular size is linked with the distance because, for instance in Euclidian static space, farther objects appear smaller.
Let's follow the same steps as before. Given a source at comoving distance $\chi(z)$ which has emitted a photon at time t_1, and with a transverse size of D, its angular size is, in a static Euclidian space:

$$\delta\theta = \frac{D}{\chi} \quad \text{static Euclidian space} \tag{3.64}$$

where we assumed $\delta\theta \ll 1$, which is true for very distant objects, like those we deal with in cosmology. If light travels in a radial way, then from the FRW metric it follows that the angular diameter of the source is:

$$\delta\theta = \frac{D}{a(t_1)S_k(\chi)} \equiv \frac{D}{d_A(z)} \tag{3.65}$$

where we have defined the *angular diameter distance*:

$$d_A(z) = \frac{S_k(\chi)}{1+z} \tag{3.66}$$

Notice the difference with the definition of the luminosity distance (3.57). Whereas for d_L we were interested in the moment of observation $a(t_0)$, we are now taking the time of emission $a(t_1)$. Furthermore, we can work out a perturbative expression for $d_A(z)$, just by noticing its relationship with $d_L(z)$:

$$d_A = \frac{d_L}{(1+z)^2} \tag{3.67}$$

Therefore:

$$d_A(z) = \frac{c}{H_0}\left(z - \frac{1}{2}(1+q_0)z^2 + ...\right) \tag{3.68}$$

Notice how, for really close sources, the first order term gives $d_L = d_A = cz/H_0$, i.e. Hubble's law. This, as remarked above, is true because spacetime is locally flat, and in Euclidian static space we know that there is only one distance, as our daily-life experience tells us. The Euclidian static limit also reveals something deep. A rightful question is in fact the following: given all these distances (comoving, proper, angular diameter, luminosity), which is the "correct" one? Clearly, this question only makes sense because we are used to the Euclidian static limit, where all distances coincide. However, if space is curved and expanding, different definitions yield different results for very distant objects.

We have completed the overview of the properties of the FRW metric. Noticing that most of these actually depend on the function $a(t)$, the *dynamics* of the Universe, we now turn to the Einstein equations.

4 The dynamics of the Universe

Up until now we were only concerned with the properties of the Universe at large scales. Making use of symmetry arguments we determined the metric:

$$\begin{aligned} ds^2 &= -c^2dt^2 + a^2(t)\left(\frac{dr^2}{1-kr^2} + r^2d\Omega^2\right) \\ &= ds^2 = a^2(\eta)(-c^2d\eta^2 + (d\chi^2 + S_k^2(\chi)d\Omega^2)) \end{aligned} \tag{4.1}$$

We have also explored all these kind of spacetimes, but the scale factor $a(t)$, which determines the *dynamics* of the Universe, was never specified. It is in fact necessary to solve the Einstein equations to find how $a(t)$ evolves:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} \tag{4.2}$$

4.1 Curvature

In order to solve the Einstein equations, we have to calculate all the necessary tensors. We start with the left hand side first of (4.2). Unfortunately, although being quite straightforward, the calculation of the Ricci tensor and scalar is very tedious. Let's start by recalling the Christoffel symbols:

$$\begin{aligned} \Gamma^{\mu}_{00} &= \Gamma^{0}_{\alpha 0} = 0 \\ \Gamma^{0}_{ij} &= c^{-1}a\dot{a}\gamma_{ij} = c^{-1}\frac{\dot{a}}{a}g_{ij} \\ \Gamma^{i}_{0j} &= c^{-1}\frac{\dot{a}}{a}\delta^i_j \\ \Gamma^{i}_{jk} &= \frac{1}{2}\gamma^{il}(\partial_j\gamma_{kl} + \partial_k\gamma_{jl} - \partial_l\gamma_{jk}) \end{aligned} \tag{4.3}$$

Now we have, by definition:

$$R_{\mu\nu} = \partial_\lambda\Gamma^\lambda_{\mu\nu} - \partial_\nu\Gamma^\lambda_{\mu\lambda} + \Gamma^\lambda_{\lambda\rho}\Gamma^\rho_{\mu\nu} - \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\rho} \tag{4.4}$$

Setting $\mu = \nu = 0$, we can calculate R_{00}:

$$R_{00} = \partial_\lambda\Gamma^\lambda_{00} - \partial_0\Gamma^\lambda_{0\lambda} + \Gamma^\lambda_{\lambda\rho}\Gamma^\rho_{00} - \Gamma^\rho_{0\lambda}\Gamma^\lambda_{0\rho} \tag{4.5}$$

From (4.3) we can see that all the Christoffel's with two time indices vanish, so that we get:

$$R_{00} = -\partial_0\Gamma^i_{0i} - \Gamma^i_{0j}\Gamma^j_{0i} \tag{4.6}$$

But again from (4.3) we have $\Gamma^i_{0j} = c^{-1}\frac{\dot{a}}{a}\delta^i_j$, so we find:

$$R_{00} = -\frac{1}{c^2}\frac{d}{dt}\left(3\frac{\dot{a}}{a}\right) - \frac{3}{c^2}\left(\frac{\dot{a}}{a}\right)^2 = -\frac{3}{c^2}\frac{\ddot{a}}{a} \tag{4.7}$$

The other components of the Ricci tensor are only shown here below, because the necessary calculations are simply too long and uninteresting to be included:

$$\begin{aligned} R_{00} &= -\frac{3}{c^2}\frac{\ddot{a}}{a} \\ R_{0i} &= 0 \\ R_{ij} &= \frac{1}{c^2}\left(\frac{\ddot{a}}{a} + 2\left(\frac{\dot{a}}{a}\right)^2 + 2\frac{kc^2}{a^2}\right) \end{aligned} \tag{4.8}$$

Given the Ricci tensor, it is straightforward to compute the Ricci scalar:

$$\begin{aligned} R &= g^{\mu\nu}R_{\mu\nu} \\ &= -R_{00} + \frac{1}{a^2}\gamma^{ij}R_{ij} = \frac{3}{c^2}\frac{\ddot{a}}{a} + \frac{\delta^i_i}{c^2}\left(\frac{\ddot{a}}{a} + 2\left(\frac{\dot{a}}{a}\right)^2 + 2\frac{kc^2}{a^2}\right) \\ &= \frac{6}{c^2}\left(\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{kc^2}{a^2}\right) \end{aligned} \tag{4.9}$$

Finally, we state the components of the Einstein tensor $G_{\mu\nu} = R_{\mu\nu} - 1/2Rg_{\mu\nu}$:

$$
\begin{aligned}
G^0_0 &= -\frac{3}{c^2}\left(\left(\frac{\dot{a}}{a}\right)^2 + \frac{kc^2}{a^2}\right) \\
G^0_i &= 0 \\
G^i_j &= -\frac{1}{c^2}\left(2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{kc^2}{a^2}\right)\delta^i_j
\end{aligned}
\tag{4.10}
$$

4.2 The energy content

To complete the Einstein equations, we have to specify the form of the energy-momentum tensor $T_{\mu\nu}$ for all the fields (matter, radiation, dark energy) present in the Universe. Let us first recall the properties of the energy-momentum tensor itself and only later describe how it differs between the various species.

4.2.1 The energy-momentum tensor

The energy-momentum tensor (EMT) describes the presence and the dynamics of all the matter contained in spacetime. Let's start from scratch to understand what form of the EMT is necessary for our case, i.e. for a homogeneous and isotropic distribution of matter at large scales.
In Newtonian dynamics, many gravitational problems "only" present the interaction between a few particles or objects, such as the two body problem, from which we can deduce, for example, the trajectory of the Earth around the Sun. However, in many cases, where the number of interacting particles is especially large, describing each and every interaction becomes inappropriate or straight impossible. This is where the *fluid* approximation comes in handy. This approximation treats the system as a smooth continuum, which is now described only by averaged quantities, or fields (such as velocity, density, temperature, pressure etc...). That is, we assume that, in a sufficiently small neighborhood of every point $\vec{x}$, the system looks homogeneous, so that it can be specified by just a few variables. Basically, since the particles of the fluid will be moving randomly, this will give rise to a *pressure field*, a certain amount of *heat conduction* and *viscous forces* between neighboring fluid elements. In addition, viscous forces can result in *shearing* of the fluid.
One of the simplest type of fluids is the *perfect fluid*, and it turns out to be the one we are looking for in our description of the Universe at large scales. A fluid is said to be perfect if there exists a frame, the *comoving frame*, where there are no shear stresses nor heat conduction. As a consequence, in this reference frame, the local properties in the neighborhood of any point $\vec{x}$ are isotropic.
To express the fluid in this comoving frame, in special relativity, we need the energy-momentum tensor $T^{\mu\nu}$. Let us first recall what the various components mean (in a general reference frame):

- $T^{00}(x)$ is the energy density of the fluid element
- $cT^{0j}(x)$ is the energy flux of the fluid element in the j-th direction
- $\frac{1}{c}T^{i0}(x)$ is the i-th component of the momentum density of the fluid element
- $T^{ij}(x)$ is the flux of the i-th component of the momentum in the j-th direction. The momentum flux across two fluid elements indicates an exertion of force. If these forces are perpendicular to the interface between the fluid elements, then such forces are represented by a diagonal T^{ij}. If, on the other hand, the forces are parallel at the interface between the fluid elements, then these are represented by off-diagonal terms in T^{ij}.

Before moving on, note that the energy-momentum tensor is a symmetric quantity, i.e. $T^{\mu\nu} = T^{\nu\mu}$. This is a non-trivial statement. For instance, this implies that $T^{i0} = T^{0i}$, so (neglecting the c factors) the energy flux in the i-th direction is equal to component of momentum density in that same direction.
Now, if we want to describe the EMT of a perfect fluid, we can reason in the following way for the comoving frame, where there is no shear stress nor conductivity. First, since the conductivity is zero, we have $T^{0j} = 0 = T^{i0}$, since those components measure a flux (we could equally have said that these vectors have to be zero because of isotropy, as we did when arguing that $g_{0i} = 0$). Then, we have $T^{ij} = 0$ for $i \neq j$ because these terms correspond to viscous forces at the interfact between fluid elements. But T^{ij}, being a 3-tensor, must be diagonal in all reference frames connected to one another by a rotation. This necessarily implies that $T^{ij} \propto \delta^{ij}$. The proportionality constant is in this case the pressure $T^{ij} = P(x)\delta^{ij}$. Basically, we have found that in the comoving frame, the energy-momentum tensor reads:

$$
T^{\mu\nu}(x) = \begin{pmatrix} \rho(x)c^2 & & & \\ & P(x) & & \\ & & P(x) & \\ & & & P(x) \end{pmatrix}
\tag{4.11}
$$

Finally, we turn to general relativity. Clearly, a good model for energy-matter content of the Universe at large scales is the perfect fluid, because it incorporates isotropy by itself. The requirement of homogeneity then only takes out the $\vec{x}$ dependence, as seen below. We also have to notice that the metric is not $\eta^{\mu\nu}$, but the FRW one. For this reason, the convention is that, in the comoving frame, which in our case is just the one specified by the comoving coordinates, we have:

$$T^{\mu}_{\nu}(x) = \begin{pmatrix} -\rho(t)c^2 & & & \\ & P(t) & & \\ & & P(t) & \\ & & & P(t) \end{pmatrix} \tag{4.12}$$

where now the energy ρ and pressure P fields are dependent only on time, because of homogeneity. The energy-momentum tensor is then extended to the explicitly covariant form:

$$T_{\mu\nu} = \left(\rho + \frac{P}{c^2}\right) u_\mu u_\nu + P g_{\mu\nu} \tag{4.13}$$

where u^μ is the relative 4-velocity between the fluid and a comoving observer, i.e. the fluid's peculiar velocity. Notice that, as expected, in the comoving frame, with $u^\mu = (c, 0, 0, 0)$, we return to (4.11).
As a final note about the energy-momentum tensor, let's analyse the conservation equation it obeys:

$$\nabla_\mu T^{\mu}_{\nu} = 0 \tag{4.14}$$

Expanding it we have:

$$\nabla_\mu T^{\mu}_{\nu} = \partial_\mu T^{\mu}_{\nu} + \Gamma^{\mu}_{\mu\lambda} T^{\lambda}_{\nu} - \Gamma^{\lambda}_{\mu\nu} T^{\mu}_{\lambda} = 0 \tag{4.15}$$

These are four separate equations, one for every value of ν. The spatial components:

$$\nabla_\mu T^{\mu}_{i} = 0 \tag{4.16}$$

are identically satisfied. We actually expected this, as $\nabla_\mu T^{\mu}_{i}$ is a vector which, again because of isotropy, has to vanish. The time component of the conservation equation is instead more interesting:

$$\partial_\mu T^{\mu}_{0} + \Gamma^{\mu}_{\mu\lambda} T^{\lambda}_{0} - \Gamma^{\lambda}_{\mu 0} T^{\mu}_{\lambda} = 0 \tag{4.17}$$

This reduces to:

$$\frac{1}{c}\frac{d(\rho c^2)}{dt} + \Gamma^{\mu}_{\mu 0}(\rho c^2) - \Gamma^{\lambda}_{\mu 0} T^{\mu}_{\lambda} = 0 \tag{4.18}$$

From the Christoffel's (4.3) we see that $\Gamma^{\lambda}_{\mu 0}$ vanishes unless the indices are both spatial and equal to each other, in which case $\Gamma^{i}_{i0} = 3c^{-1}\dot{a}/a$. We then finally get:

$$\dot{\rho} + 3\frac{\dot{a}}{a}\left(\rho + \frac{P}{c^2}\right) = \dot{\rho} + 3H\left(\rho + \frac{P}{c^2}\right) = 0 \tag{4.19}$$

This equation is the continuity equation in an expanding spacetime. In order to solve (4.19), we need the last piece of the puzzle, the *equation of state*, which relates pressure to energy density:

$$P = P(\rho) \tag{4.20}$$

In general, most cosmological fluids are well described by a constant equation of state:

$$P = w\rho c^2 \tag{4.21}$$

With this in mind, the continuity equation reads:

$$\frac{\dot{\rho}}{\rho} = -3(1+w)\frac{\dot{a}}{a} \tag{4.22}$$

To which the solution obviously is:

$$\rho \propto a^{-3(1+w)} \tag{4.23}$$

In other words, the dilution of energy of a fluid in an expanding spacetime depends on the equation of state of the fluid itself.
In the first lecture we have described the types of fluids present in the Universe: matter (dark and baryonic), radiation and dark energy. What is the equation of state, i.e. the w, for each of these?

4.2.2 Equations of state

Let's think about what we have done so far. We have written down the Einstein equations (4.2) and calculated the left hand side, the "curvature part". To complete the equations, however, we also need to know the energy-momentum tensor of every species of the Universe. As we have argued in the above, each cosmological fluid, being isotropic and homogeneous (at large scales), is rather well described by a perfect fluid, whose energy-momentum tensor (in the comoving frame) is of the form (4.12). Then, solving the continuity equation using a constant equation of state, we have found the scaling of the energy density with the scale factor $a(t)$ (4.23). For this reason, to complete the Einstein equations, we need to find how $\rho(t)$ evolves (or $P(t)$, since they are proportional), which means we need to find w for matter, radiation and dark energy.

In order to do this, we need to turn to statistical mechanics. Consider a gas of particles that are at thermodynamic equilibrium with temperature T, inside a box of volume V. Then, the number density of particles with momentum (the absolute value) in the interval $(p, p+dp)$ is given by:

$$dn(p) = \frac{g}{(2\pi\hbar)^3} f(p,T) 4\pi p^2 dp \tag{4.24}$$

where $f(p,T)$ is the probability distribution function, i.e. the probability to find a randomly chosen particle with momentum p. For a classical gas, this is the Maxwell-Boltzmann distribution, whereas for a quantum gas, it's either the Fermi-Dirac distribution (if the particles are fermions), or the Bose-Einstein distribution (if the particles are bosons). The g factor in (4.24) is the spin degeneracy of the particles. This probability distribution is derived using basic quantum mechanics, but we will not do so here. Once we have $dn(p)$, we can calculate the thermodynamic variables:

$$\begin{aligned}
\frac{N}{V} &= \int_0^\infty dn(p) \\
\rho c^2 = \frac{E}{V} &= \int_0^\infty dn(p) E(p) \\
\frac{P}{V} &= \frac{1}{3}\int_0^\infty dn(p) p v(p) = \frac{1}{3}\frac{\int_0^\infty dn(p) p v(p)}{\int_0^\infty dn(p)} n = \frac{1}{3} n \langle p v(p) \rangle
\end{aligned} \tag{4.25}$$

where $E(p) = \sqrt{m^2c^4 + p^2c^2}$ is the particles' energy.

Now, let's evaluate P for every species that we know exists in the Universe. We start with matter, which is basically defined to be the opposite of radiation, that is:

$$E(p) = \sqrt{m^2c^4 + p^2c^2} \simeq mc^2 \qquad v(p) = \frac{p}{m} \tag{4.26}$$

This is the non-relativistic limit, and it applies to both dark energy and baryonic matter. The pressure then is:

$$P = \frac{1}{3} n \langle p v(p) \rangle = \frac{1}{3} n m \langle v(p)^2 \rangle = \frac{1}{3} n \frac{E}{c^2} \langle v(p)^2 \rangle \simeq 0 \tag{4.27}$$

where the last approximation is due to the non-relativistic nature of the particles, $\langle v^2 \rangle / c^2 \simeq 0$. For this reason we have:

$$P_{\text{matter}} = 0 \tag{4.28}$$

which implies $w_{\text{matter}} = 0$. The dilution of the energy density is then, from (4.23):

$$\rho_{\text{matter}} \propto a^{-3} \tag{4.29}$$

That is, the matter energy density scales like the inverse of the volume, with size a^{-3}.

Instead, for radiation (remember that by radiation we mean any ultra-relativistic species. At present time the only ones are photons and neutrinos, but every particle was once ultra-relativistic, as the Universe was very hot) we have the following properties:

$$E(p) = \sqrt{m^2c^4 + p^2c^2} \simeq pc \qquad v(p) = c \tag{4.30}$$

The pressure then becomes:

$$P = \frac{1}{3} n \langle p v(p) \rangle = \frac{1}{3} n \langle pc \rangle \tag{4.31}$$

whereas the energy density:

$$\rho c^2 = \frac{E}{V} = \int_0^\infty dn(p) pc = \frac{\int_0^\infty dn(p) pc}{\int_0^\infty dn(p)} n = n \langle pc \rangle \tag{4.32}$$

Putting (4.32) and (4.31) together, we find:

$$P_{\text{radiation}} = \frac{1}{3}\rho \tag{4.33}$$

which means $w_{\text{radiation}} = 1/3$, leading to an energy dilution of the form:

$$\rho_{\text{radiation}} \propto a^{-4} \tag{4.34}$$

Again, we find the scaling of the volume a^{-3}, but this time enhanced by another a^{-1}, which comes from the redshift. The last component we have to analyse is dark energy. As we have seen in the first lectures, it was clear that matter and radiation only could not describe our Universe. Because of this shortcoming, the concept of dark energy was introduced. To the best of our knowledge, this entity satisfies an equation of state of the following form:

$$P_\Lambda \simeq -\rho c^2 \tag{4.35}$$

that is $w_\Lambda = -1$, and its energy density stays constant:

$$\rho_\Lambda \propto a^0 \tag{4.36}$$

Since the Universe is expanding, energy has to be created continuously. A natural candidate that satisfies this behavior is *vacuum energy* because, as the Universe expands, more space is being created and therefore the energy increases in proportion to the volume, leading to a constant energy density. In quantum field theory, this vacuum energy is actually predicted, leading to an energy-momentum tensor of the form:

$$T^{\text{vac}}_{\mu\nu} = -\rho c^2 g_{\mu\nu} \tag{4.37}$$

Comparing this with (4.13), we see that this form implies $P = -\rho c^2$. This type of vacuum energy is actually also predicted by general relativity, through the cosmological constant. In fact, we could rewrite the Einstein equations (4.2) with a modified energy-momentum tensor:

$$G_{\mu\nu} = \frac{8\pi G}{c^4}(T_{\mu\nu} + T^\Lambda_{\mu\nu}) \tag{4.38}$$

where:

$$T^\Lambda_{\mu\nu} = -\frac{\Lambda c^4}{8\pi G} g_{\mu\nu} = -\rho_\Lambda c^2 g_{\mu\nu} \tag{4.39}$$

This clearly has the same form of (4.37). In general, the terms "vacuum energy" and "dark energy" are used intercheangably, to mean a certain entity that satisfies an equation of state with $w = -1$.

4.3 The Friedmann equations

We finally have all the pieces to solve, or at least rewrite, the Einstein equations. Let's summarise all the different items here. The non-zero components of the Einstein tensor $G_{\mu\nu}$ are:

$$\begin{aligned}
G^0_0 &= -\frac{3}{c^2}\left(\left(\frac{\dot{a}}{a}\right)^2 + \frac{kc^2}{a^2}\right) \\
G^0_i &= 0 \\
G^i_j &= -\frac{1}{c^2}\left(2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{kc^2}{a^2}\right)\delta^i_j
\end{aligned} \tag{4.40}$$

Then, the form of the energy-momentum tensor is that of a perfect fluid, which in the rest (comoving) frame reads:

$$T^\mu_\nu(x) = \begin{pmatrix} -\rho(t)c^2 & & & \\ & P(t) & & \\ & & P(t) & \\ & & & P(t) \end{pmatrix} \tag{4.41}$$

This has to satisfy the conservation equation, which yields a scaling of the energy density of:

$$\rho \propto a^{-3(1+w)} \tag{4.42}$$

The total energy density is the sum of the energy density of every component in the Universe:

$$\rho = \rho_{\text{matter}} + \rho_{\text{radiation}} + \rho_\Lambda \tag{4.43}$$

which scale as:

$$\rho_{\text{matter}} \propto a^{-3} \quad \rho_{\text{radiation}} \propto a^{-4} \quad \rho_\Lambda \propto a^0 \tag{4.44}$$

Now, let's pull everything together. There are only two Einstein equations, namely the 00 one, and any one of the spatial ones ij, as the G^i_j is proportional to δ^i_j, because of isotropy. The 00 equation reads:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} \tag{4.45}$$

This equation is called the *Friedmann equation* and is particularly important. Given initial conditions, from this equation it is rather straightforward to calculate $a(t)$, which controls the dynamics of the whole Universe. The ρ present in this equation is to be understood as a sum, namely the one found in (4.43). The spatial part of the Einstein equations yields the (second) Friedmann equation:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3P}{c^2}\right) \tag{4.46}$$

This equation actually follows from taking the time derivative of (4.45) and using the continuity equation (4.19). This shouldn't surprise us, because the Einstein equation implies the conservation equation for the energy-momentum tensor via the Bianchi identity $\nabla_\mu G^{\mu\nu} = 0$.
These two equations (4.45) and (4.46) control the large-scale dynamics of our Universe. In the next lecture we will explore them and analyse their fundamental implications.

5 A detailed study of the Friedmann equations

In the last lecture we introduced the Friedmann equations, which are just the Einstein equations that the FRW metric satisfies. The two equations read:

$$\begin{aligned}\left(\frac{\dot{a}}{a}\right)^2 &= \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2}\\ \frac{\ddot{a}}{a} &= -\frac{4\pi G}{3}\left(\rho + \frac{3P}{c^2}\right)\end{aligned} \tag{5.1}$$

They reveal everything we need to describe the evolution of the Universe, that is they are differential equations for the scale factor $a(t)$. Now, given that the scale factor controls the dynamics of the Universe, these equations give an explicit overview of the various stages the Universe itself went through, i.e. if it has been expanding forever and at which rate.
Remember another fundamental equation, the continuity equation:

$$\dot{\rho} + 3H\left(\rho + \frac{P}{c^2}\right) = 0 \tag{5.2}$$

through which we obtain the second Friedmann equation in (5.1) by differentiating with respect to time the first one.

5.1 Some properties

Before moving on to solve the equations themselves, it is interesting to see how a multitude of useful insights can be gained through some simple considerations that don't involve specific solutions.

5.1.1 The Newtonian perspective and the fate of the Universe

The main intuition behind the Friedmann equations is that they are simply a statement about conservation of energy in the Newtonian framework. To see this, consider a non-relativistic system containing a distribution of matter ($P = 0$) with density $\rho(t)$ and organized in a spherical fashion, where the sphere has radius $a(t)$. If a test particle is placed on the surface of the said sphere, by conservation of energy we have:

$$\frac{1}{2}m\dot{a}^2 - G\frac{mM}{a} = E \tag{5.3}$$

where E is the total energy and $M = 4/3\pi a^3(t)\rho(t)$ is the mass of the matter distribution. Rewriting this equation we have the following:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho(t) + \frac{2E}{ma^2} \tag{5.4}$$

Clearly, this matches the the first Friedmann equation (5.1) on two conditions. First, we need to introduce in the Newtonian theory both radiation and dark energy, so that the density ρ becomes the sum of all these terms. Secondly, the identification $2E/m$ with -2κ needs to be made. This means that the total energy of the system is related to the spatial curvature k. Since the energy is related to the concept of escape velocity, this implies that the curvature actually determines the fate of the Universe. If, for example, the energy of the system is 0, or equivalently the space is flat, then the sphere will expand forever, with a velocity that tends to 0 (note that, just because we are talking about a spherical matter distribution, it doesn't have any implication for the spatial geometry, which can be flat, open or closed). On the other hand, we can reason in the opposite way, because it is the density $\rho(t)$ that specifies the curvature of space in the first place, so that the two concepts are tightly linked.
To make even more explicitly the connection between the two, curvature and energy, we note that a flat Universe today ($t = t_0$) corresponds to the following *critical density*:

$$\rho_{\text{crit},0} = \frac{3H_0^2}{8\pi G} \simeq 1.26 \times 10^{11} M_\odot \text{Mpc}^{-3} \tag{5.5}$$

Now, in cosmology it is convenient to scale all the densities to the critical density, and work with the adimensional parameters Ω_i defined as:

$$\Omega_i = \frac{\rho_{i,0}}{\rho_{\text{crit},0}} \quad i \in \{\text{Matter, Radiation, Dark Energy}\} \tag{5.6}$$

In the last lecture we calculated the dependence of the different densities $\rho(t)$ on time for all the species, finding:

$$\rho_m(t) = \rho_{m,0}a^{-3} \quad \rho_r(t) = \rho_{r,0}a^{-4} \quad \rho_\Lambda(t) = \rho_{\Lambda,0}a^0 \tag{5.7}$$

Using these relations, the first Friedmann equation can then be rewritten as:

$$\frac{H^2}{H_0^2} = \Omega_r a^{-4} + \Omega_m a^{-3} + \Omega_k a^{-2} + \Omega_\Lambda \tag{5.8}$$

where we have defined the adimensional energy density parameter associated with curvature $\Omega_k = -kc^2/H_0^2$. Evaluating (5.8) at present time yields an important constraint:

$$1 = \sum_i \Omega_i \equiv \Omega_0 + \Omega_k \tag{5.9}$$

Where we have defined the total energy content $\Omega_0 \equiv \Omega_r + \Omega_m + \Omega_\Lambda$. This implies that the sign of $\Omega_0 - 1$ is related to that of Ω_k:

$$\Omega_0 - 1 = -\Omega_k = \frac{kc^2}{H_0^2} \tag{5.10}$$

In other words, if the energy content of the Universe happens to satisfy $\Omega - 1 = 0$, the spatial curvature will be zero, and the expansion will continue forever; this is the case of a *flat* Universe. The same line of reasoning can be carried out for the other cases, $\Omega - 1 < 0$ or $\Omega - 1 > 0$, which are the cases of *open* and *closed* Universes, where, respectively, the expansion will continue again forever or halt:

$$\begin{aligned} k = +1 &\longleftrightarrow \Omega_0 > 1 \qquad \text{Closed Universe} \\ k = 0 &\longleftrightarrow \Omega_0 = 1 \qquad \text{Flat Universe} \\ k = -1 &\longleftrightarrow \Omega_0 < 1 \qquad \text{Open Universe} \end{aligned} \tag{5.11}$$

Notice one final important consequence of the last form of the Friedmann equation (5.8). The different scalings of the various energy densities imply that the Universe underwent different phases in which only one of the species dominated. In particular, for very early times $a \ll 1$, the Universe underwent a period of *radiation domination* (RD), whereas only later and until close to present day was the evolution controlled by matter, in the period of *matter domination* (MD). Finally, we can see that in the far future (actually, even today) the dynamics are specified by dark energy in the period of *dark energy domination* (ΛD). This last period is still subject to uncertainty, as we remarked when we defined the cosmological event horizon. Since the nature of dark energy is still subject to intense study and research, it is not known if, for instance, its equation of state is really a constant $w_\Lambda = -1$ or if it depends on time $w_\Lambda = w_\Lambda(a)$. In this last case, the dark energy domination period will need to be revised.
Mathematically speaking, we have the following limits of (5.8):

$$\begin{aligned} \frac{H^2}{H_0^2} &\approx \Omega_r a^{-4} \longleftrightarrow \text{Early Universe (RD)} \\ \frac{H^2}{H_0^2} &\approx \Omega_m a^{-3} \longleftrightarrow \text{Late Universe (MD)} \\ \frac{H^2}{H_0^2} &\approx \Omega_\Lambda \longleftrightarrow \text{Today and the Future } (\Lambda\text{D}) \end{aligned} \tag{5.12}$$

5.1.2 Acceleration or deceleration?

During the study of distances in cosmology, we introduced the deceleration parameter q_0:

$$q_0 = -\frac{\ddot{a}}{aH^2}\Big|_{t_0} \tag{5.13}$$

Not surprisingly, the Friedmann equations allow us to find an analytical expression for this parameter. Using the second equation in (5.1) and the definition of critical density (5.5), we find:

$$q_0 = \frac{1}{2}\sum_i \Omega_i(1+3w_i) \tag{5.14}$$

where, as in (5.6), the index i spans all species. Now, using the various equations of state, this equation simplifies to:

$$q_0 = \frac{1}{2}(\Omega_m + 2\Omega_r - 2\Omega_\Lambda) \tag{5.15}$$

Therefore, by measuring the adimensional density parameters Ω_i, we can determine if the Universe today is accelerating or decelerating. Obviously, this calculation can be done for every instant of cosmic time, providing us even more insight into the dynamics of the Universe.

5.1.3 The age of the Universe

A final, useful application of the Friedmann equations is the calculation of the age of the Universe. To estimate this quantity, the following relationship can in fact be used:

$$\frac{da}{dt}\frac{1}{a} = H \tag{5.16}$$

This implies:

$$\int_0^{t_0} dt = \int_0^1 \frac{da}{a}\frac{1}{H} \tag{5.17}$$

where, again, the extrapolation to the past is intended to be valid up until (approximately) $t_{100\ \text{GeV}}$ and not $t = 0$. This is however just a nuisance (it techincally isn't) so we can go ahead and use (5.8) to find:

$$t_0 = \frac{1}{H_0}\int_0^1 da\ [\Omega_m a^{-1} + \Omega_r a^{-2} + \Omega_\Lambda a^{-2} + \Omega_k]^{-1/2} \tag{5.18}$$

The major takeaway from these properties derived from the Friedmann equations is that the measurement of the adimensional density parameters Ω_i is of fundamental importance. These parameters specify uniquely the scale factor $a(t)$ through the first Friedmann equation. In addition, they offer an expression for the deceleration parameter, making possible the measurement of distances of different sources. Finally, these same parameters determine the age of the Universe.

Having analysed the equations, we now turn to their solutions.

5.2 Solutions

In this section we give exact solutions to the first Friedmann equation:

$$\dot{a}^2 = H_0^2\left(\frac{\Omega_m}{a} + \frac{\Omega_r}{a^2} + \Omega_\Lambda a^2 + \Omega_k\right) \tag{5.19}$$

While a general analytical solution to this equation does not exist, we can find some relatively simple solutions if we restrict ourselves to special cases. For example, simple solutions exist for single-component Universes, which are a good approximation for periods in cosmic time where the equation of state was approximately constant, like the periods of matter, radiation and dark energy domination. Moreover, analytical solutions exist for two-component Universes, which are interesting because they incorporate the transition era between the two equations of state.

5.2.1 Matter Universes: $\Omega_m + \Omega_k = 1$

We first study the solutions to a Universe with no amount radiation and dark energy:

$$\dot{a}^2 = H_0^2 \left(\frac{\Omega_m}{a} + \Omega_k \right) \tag{5.20}$$

- Consider firstly a Universe where there is only matter, so that $\Omega_m = 1$ and $\Omega_k = 0$. In this case, the Friedmann equation simplifies down to the first in (5.12):

$$\dot{a}^2 = H_0^2 \Omega_m a^{-1} \tag{5.21}$$

  Imposing the condition $a(t_0) = 1$ we have the following solution:

$$a(t) = \left(\frac{t}{t_0} \right)^{2/3} \tag{5.22}$$

  This solution is historically interesting. It's called the *Einstein-de Sitter* (EdS) Universe, having been first derived by them. This connects back to the first lecture, where we hinted at the fact that a "matter-only" Universe (the one predicted by EdS) was having serious trouble explaining some observational facts, among which the predicted age of the Universe being shorter than some stars within it. Solving the integral (5.18), we find that the age of the Universe predicted by this model is:

$$t_0 = \frac{2}{3} \frac{1}{H_0} \tag{5.23}$$

  Therefore, given an estimate of the Hubble constant H_0, one can easily calculate the age of the Universe. In the next lecture we will give an overview of the various measurements that were carried out in the last few decades to form the ΛCDM model, and we will return to this issue.
  Although this solution doesn't completely describe the evolution of the Universe (we know there is more than just matter, like dark energy and radiation), it is a good approximation for the period of matter domination, as we said above.

- Consider now the case of $k = -1$, i.e. an open Universe. The solution can be expressed in parametric form:

$$a(\theta) = \frac{1}{2} \frac{\Omega_m}{|1 - \Omega_m|} (\cosh\theta - 1) \quad t(\theta) = \frac{1}{4c} \frac{\Omega_m}{|1 - \Omega_m|} (\sinh\theta - \theta) \tag{5.24}$$

  where $0 \leq \theta \leq \infty$. Note first that the solution reduces to $a(t) \propto t^{2/3}$ at early epochs $\theta \ll 1$: because of the difference in scaling between matter (a^{-3}) and curvature (a^{-2}), as long as $\Omega_m \neq 0$, there will always be an arbitrarily early time when matter dominates. More importantly, at later epochs $\theta \gg 1$ we have $a \propto t$, a phase of free expansion of the Universe, which will continue forever.

- Finally, we consider the solution for $k = +1$, given in parametric form by:

$$a(\theta) = \frac{1}{2} \frac{\Omega_m}{|1 - \Omega_m|} (1 - \cos\theta) \quad t(\theta) = \frac{1}{4c} \frac{\Omega_m}{|1 - \Omega_m|} (\theta - \sin\theta) \tag{5.25}$$

  where $0 \leq \theta \leq 2\pi$. We can see that the Universe reaches a maximum size $a_{\max}$ at time $t_{\max}$, given by:

$$a_{\max} = \frac{\Omega_m}{|1 - \Omega_m|} \quad t_{\max} = \frac{\pi}{4c} \frac{\Omega_m}{|1 - \Omega_m|} \tag{5.26}$$

  After reaching this maximum size, the Universe recollapses in a *big crunch* at $\theta = 2\pi$, where:

$$a_{\text{crunch}} = 0 \quad t_{\text{crunch}} = \frac{\pi}{2c} \frac{\Omega_m}{|1 - \Omega_m|} \tag{5.27}$$

  Again, note that at early epochs we have $a \propto t^{2/3}$, for the same reason as before.

The following plot summarises these three cases:

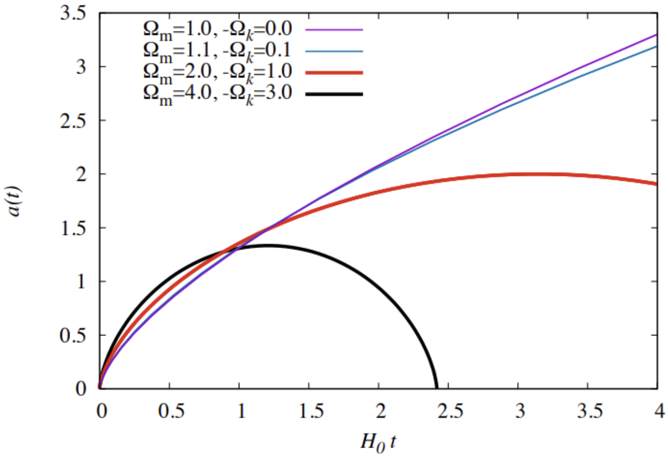

Figure 14: The three cosmologies where $\Omega_m + \Omega_k = 1$, corresponding to flat, open and closed Universes.

In the plot we can see the results we just derived. For a flat geometry ($\Omega_m = 1$ and $k = 0$), we get an ever expanding Universe with "escape velocity". On the other hand, an open Universe ($\Omega_m < 1$ and $k = -1$) still expands forever, whereas in a closed Universe ($\Omega_m > 1$ and $k = +1$) the matter density is high enough for it to recollapse on itself in a big crunch.

5.2.2 Radiation Universes: $\Omega_r + \Omega_k = 1$

Now we delve into the study of the solutions for a radiation Universe void of both matter and dark energy:

$$\dot{a}^2 = H_0^2 \left(\frac{\Omega_r}{a^2} + \Omega_k \right) \tag{5.28}$$

- As for the matter Universes, let's analyse first the case of zero spatial curvature $k = 0$:

$$\dot{a}^2 = H_0^2 \Omega_r a^{-2} \tag{5.29}$$

  Again, imposing the condition $a(t_0) = 1$, we have the following solution:

$$a(t) = \left(\frac{t}{t_0} \right)^{1/2} \tag{5.30}$$

  Calculating the age of the Universe with the integral (5.18), one gets:

$$t_0 = \frac{1}{2} \frac{1}{H_0} \tag{5.31}$$

  This solution is particularly interesting for the very early Universe. As said above, in fact, as long as $\Omega_r \neq 0$, there will always be an arbitrarily early time such that radiation dominates, because of the scaling with a^{-4}.

- Next, we add some amount of curvature, such that $k = +1$, i.e. a closed Universe. This case is very similar to the closed Universe with matter, in the sense that it has a big crunch at some finite time in the future:

$$a(\theta) = \sqrt{\frac{1}{2} \frac{\Omega_r}{|1 - \Omega_r|}} \sin\theta \quad t(\theta) = \frac{1}{2c} \sqrt{\frac{1}{2} \frac{\Omega_r}{|1 - \Omega_r|}} (1 - \cos\theta) \tag{5.32}$$

  Where, again, $0 \leq \theta \leq 2\pi$. Notice that the crunch time, and therefore the maximum, is reached at an earlier time compared to the matter Universe:

$$a_{\text{curnch}} = 0 \quad t_{\text{crunch}} = \frac{1}{c} \sqrt{\frac{1}{2} \frac{\Omega_r}{|1 - \Omega_r|}} \tag{5.33}$$

  Basically, in both models, the Universe expands with a cycloidal evolution of the scale factor, until a maximum size is obtained at $t_{\max}$, where the expansion is reversed. Then, the Universe starts to compress until it reaches the big crunch at t_{crunch}.

- Finally, consider a radiation dominated open Universe, with $k = -1$. In this case, the expansion will go on forever:

$$a(\theta) = \sqrt{\frac{1}{2}\frac{\Omega_m}{|1-\Omega_r|}}\sinh\theta \quad t(\theta) = \frac{1}{2c}\sqrt{\frac{1}{2}\frac{\Omega_r}{|1-\Omega_r|}}(\cosh\theta - 1) \tag{5.34}$$

where $0 \le \theta \le \infty$. The early Universe limit $\theta \ll 1$ correctly reproduces the zero curvature solution $a \propto t^{1/2}$, whereas the late Universe $\theta \gg 1$ sees an expansion $a \propto t$ that will continue forever, as we remarked.

5.2.3 Dark energy Universes: $\Omega_\Lambda + \Omega_k = 1$

Consider a Universe with a cosmological constant $\Lambda > 0$ and some curvature:

$$\dot{a}^2 = H_0^2\left(\Omega_\Lambda a^2 + \Omega_k\right) \tag{5.35}$$

This kind of Universe is a good approximation to present and late times, where dark energy starts to dominate all other kinds of energies. The solutions to this equation are:

$$a(t) = \sqrt{\frac{3}{\Lambda}}\begin{cases}\cosh(\sqrt{\Lambda c^2/3}t) & k = +1\\ \exp(\sqrt{\Lambda c^2/3}t) & k = 0\\ \sinh(\sqrt{\Lambda c^2/3}t) & k = -1\end{cases} \tag{5.36}$$

Note that the $k = +1$ solution doesn't have a singularity, whereas the scale factor vanishes at $t = -\infty$ and $t = 0$ for the $k = 0$ and $k = -1$ solutions respectively. These singularities are, however, only due to poor coordinate choices, since the three solutions (5.36) are only three different ways to slice the the same *de Sitter* space.

5.2.4 Early Universe: matter and radiation

As we will learn in the next lecture, a flat ($k = 0$) Universe containing matter and radiation is very close to resembling our own, at least at early stages. The Friedmann equation we have to solve is:

$$\dot{a}^2 = H_0^2\left(\frac{\Omega_m}{a} + \frac{\Omega_r}{a^2}\right) \tag{5.37}$$

This model presents a transition from a period of radiation domination to a period of matter domination. The transition time, or scale factor $a_{\rm eq}$, called *matter-radiation equality* is defined by:

$$\frac{\Omega_m}{a} = \frac{\Omega_r}{a^2} \implies a_{\rm eq} = \frac{\Omega_r}{\Omega_m} \tag{5.38}$$

Rewriting (5.37) using this new definition simplifies things:

$$\dot{a}^2 = \frac{H_0^2\Omega_r}{a^2}\left(1 + \frac{a}{a_{\rm eq}}\right) \tag{5.39}$$

This implies the following differential equation:

$$H_0 dt = \frac{ada}{\Omega_r^{1/2}}\left(1 + \frac{a}{a_{\rm eq}}\right)^{-1/2} \tag{5.40}$$

Integrating it, we get the solution:

$$H_0 t = \frac{4a_{\rm eq}^2}{3\Omega_r^{1/2}}\left(1 - \left(1 - \frac{a}{2a_{\rm eq}}\right)\left(1 + \frac{a}{a_{\rm eq}}\right)^{1/2}\right) \tag{5.41}$$

As a sanity check, it is useful to compute the radiation and matter domination (RD and MD) limits and make sure they correspond to (5.30) and (5.22). First we impose the RD limit, obtained by $a/a_{\rm eq} \ll 1$, finding the following scaling:

$$a \simeq (3H_0\Omega_r^{1/2}t)^{1/2} \propto t^{1/2} \tag{5.42}$$

which is indeed the right proportionality. Now, the MD limit $a/a_{\rm eq} \gg 1$ implies:

$$a = \left(\frac{3\Omega_r H_0 t}{4a_{\rm eq}^{1/2}}\right)^{2/3} \propto t^{2/3} \tag{5.43}$$

Again, this coincides with the Einstein-de Sitter solution.
As a final application, one can also calculate the time of matter-radiation equality $t_{\rm eq}$, which is found by setting $a = a_{\rm eq}$ in (5.41):

$$t_{\rm eq} = \frac{4a_{\rm eq}^2}{3\Omega_r^{1/2}}\left(1 - \frac{1}{\sqrt{2}}\right) \tag{5.44}$$

5.2.5 Late Universe: matter and dark energy

Consider now a flat Universe containing matter and a positive cosmological constant, such that $\Omega_m + \Omega_\Lambda = 1$. This model, as we will see in the next lecture, describes rather well our own Universe at present and future times. The Friedmann equation to solve is, in this case:

$$\dot{a}^2 = H_0^2 \left(\frac{\Omega_m}{a} + \Omega_\Lambda a^2 \right) \tag{5.45}$$

The time of interest is now the moment of *matter-dark energy equality* $a_{\Lambda\mathrm{m}}$, defined by:

$$\frac{\Omega_m}{a} = \Omega_\Lambda a^2 \implies a_{\Lambda\mathrm{m}} = \left(\frac{\Omega_m}{\Omega_\Lambda} \right)^{1/3} \tag{5.46}$$

Rewriting the Friedmann equation in the following way:

$$\dot{a} = H_0 a \Omega_\Lambda^{1/2} \left(\frac{a_{\Lambda\mathrm{m}}^3}{a^3} + 1 \right)^{1/2} \tag{5.47}$$

yields the differential equation:

$$H_0 dt = \frac{a da}{\Omega_\Lambda^{1/2}} \left(\frac{a_{\Lambda\mathrm{m}}^3}{a^3} + 1 \right)^{-1/2} \tag{5.48}$$

The solution is the following:

$$H_0 t = \frac{2}{3\Omega_\Lambda^{1/2}} \log \left[\left(\frac{a}{a_{\Lambda\mathrm{m}}} \right)^{3/2} + \left(\frac{a^3}{a_{\Lambda\mathrm{m}}^3} + 1 \right)^{1/2} \right] \tag{5.49}$$

Let's check whether this solution gives the right scalings at early (5.22) and late (5.36) times. First, the limit $a/a_{\Lambda\mathrm{m}} \ll 1$ in (5.49) yields:

$$a \simeq \left(\frac{3H_0 t a_{\Lambda\mathrm{m}}^{3/2} \Omega_\Lambda^{1/2}}{2} t \right)^{2/3} \tag{5.50}$$

which is again the right proportionality for a matter dominated Universe. The late limit $a/a_{\Lambda\mathrm{m}} \gg 1$ instead results in:

$$a \simeq \frac{a_{\Lambda\mathrm{m}}}{2^{2/3}} \exp \left(\frac{3H_0 \Omega_\Lambda^{1/2} t}{2} \right) \tag{5.51}$$

which corresponds with the result of a flat de Sitter Universe.
As before, the time of matter-dark energy equality $t_{\Lambda\mathrm{m}}$ can be found by setting $a = a_{\Lambda\mathrm{m}}$ in (5.49):

$$t_{\Lambda\mathrm{m}} = \frac{2}{3H_0 \Omega_\Lambda^{1/2}} \log(1 + \sqrt{2}) \tag{5.52}$$

6 What the observations tell us

To finally complete our picture we have drawn of our Universe, we have to resort to the exploration of observational data and how it holds up when compared to our mathematical descriptions.
As we have seen throughout the lectures, the energy budget is subdivided into three components: matter (baryonic and dark), radiation and dark energy. Most importantly, however, at large enough scales (say $\gtrsim 100$ Mpc) the Universe appears isotropic and homogeneous, a fact which allowed us to simplify considerably the formalism. The metric is reduced down to the following:

$$ds^2 = -c^2 dt^2 + a^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2 d\Omega^2 \right] \tag{6.1}$$

In turn the Einstein equations can be grouped into only one differential equation, the (a-dimensional) Friedmann equation, for the scale factor $a(t)$, which governs the dynamics of the Universe:

$$H^2 = H_0^2 \left[\Omega_m a^{-3} + \Omega_r a^{-4} + \Omega_k a^{-2} + \Omega_\Lambda \right] \tag{6.2}$$

It is rather straightforward now to understand that, if we want to solve the equation here above, the next move is to find out what the Ωs are. Thank to modern observational techniques, these parameters are now known to a very high degree of precision. Unfortunately however, the observations are either based on data taken from the cosmic microwave background or Big Bang nucleosynthesis, both of which weren't discussed in these lectures. In fact, these last two topics, when combined with our large-scale description, make up the ΛCDM *model*, or the *standard cosmological model*, which is a complete and observationally robust description of the Universe. Nonetheless, we are going to give a summary of the relevant data that was extracted to find out the values of the density parameters.

6.1 The density parameters

Firstly, we recall the main events that happened during the very early Universe. As was elucidated in lecure 2, after a certain amount of time electrons and nuclei, mainly hydrogen, helium and lithium, find themselves in a plasma state. As the Universe expands, however, its temperature decreases, such that after roughly 370000 years after the Big Bang it is low enough for the first atoms to form, i.e. for the electrons to get captured by the nuclei. The photons now no longer interact and are free to stream away. This leftover radiation has been redshifting ever since due to the expansion of the Universe, so that now we see it as a microwave background. Indeed, the name we have for it is the Cosmic Microwave Background (CMB) and it is the most perfect black body we have measured in nature. It turns out that from this leftover radiation we can extract an incredible amount of information about the parameters that govern our Universe.

We start with radiation, which we recall to be any relativistic species today. The only known particles that satisfy this condition are photons and neutrinos (in reality, neutrinos are slightly massive, which makes them non-relativistic at late times). As far as photons are concerned, their density is easily recovered. The CMB, in fact, has a temperature simply by virtue of being a black body, which we recall to be:

$$\bar{T}_0 = 2.73 \text{ K} \tag{6.3}$$

Statistical mechanics provides a relationship between number/energy density of a system of photons in thermodynamical equilibrium and its temperature:

$$\begin{aligned} n_{\gamma,0} &= \frac{2\zeta(3)}{\pi^2}\left(\frac{\bar{T}_0^2}{\hbar c}\right)^3 \simeq 410 \text{ photons cm}^{-3} \\ \rho_{\gamma,0} &= \frac{\pi^2}{15}\left(\frac{\bar{T}_0^2}{\hbar c}\right)^4 \simeq 4.6\times 10^{-34} \text{ g cm}^{-3} \end{aligned} \tag{6.4}$$

(here $\zeta(s)$ is the Riemann zeta function, with $\zeta(3) \simeq 1.2$). Converting this energy density in units of critical density ρ_0, we have:

$$\Omega_\gamma = \frac{\rho_{\gamma,0}}{\rho_0} \simeq 5.4\times 10^{-5} \tag{6.5}$$

On the other hand, as far as neutrinos are concerned, as long as they are relativistic their energy density can be show to be roughly 68% of that of photons. If we extrapolate this assumption to the present time we get:

$$\Omega_\nu \simeq 3.6\times 10^{-5} \tag{6.6}$$

In total, therefore, the radiation energy density is:

$$\Omega_r = \Omega_\gamma + \Omega_\nu \simeq 9\times 10^{-5} \tag{6.7}$$

As we know, however, neutrinos are massive, despite their mass being substantially small. As a consequence, they actually become non-relativistic at late times, which increases their energy density contribution. Current observations constrain its range to be $0.0012 < \Omega_\nu < 0.003$.

Now we turn to matter, which we have said to be composed of a baryonic and a dark component. The baryonic energy density is known because of the theory of Big Bang nucleosynthesis, which predicts the abundances of light elements that were produced during the hot Big Bang. In addition, the Ω_b determines certain features of the CMB. Coupled together, these two observations yield:

$$\Omega_b \simeq 0.05 \tag{6.8}$$

Most of the matter in the Universe, however, is in the form of dark matter. From the first lecture, we recall that its effects are observed in the evolution and formation of large-scale structures (galaxies and clusters of galaxies) and in the CMB. The inferred energy density is:

$$\Omega_c \simeq 0.27 \tag{6.9}$$

where the subscript "c" indicates that we are assuming a *cold* form of dark matter, i.e. with equation of state $w_c = 0$. The total amount of matter in the Universe is thus:

$$\Omega_m = \Omega_b + \Omega_c \simeq 0.32 \tag{6.10}$$

Next, we note that the CMB actually gives precious information about the spatial curvature of the Universe. In fact, curvature changes the angular diameter distance to the surface of last-scattering, affecting the angle at which the scale is observed. These types of measurements give an upper bound to the energy density in the form of curvature:

$$|\Omega_k| < 0.005 \tag{6.11}$$

This number, therefore, suggests that the Universe is actually flat $k \simeq 0$. In other words, the spacetime metric is of the form:

$$ds^2 = -dt^2 + a^2(t)\delta_{ij}dx^i dx^j \tag{6.12}$$

Finally, the presence of a dark energy component, which makes the Universe accelerate, was inferred from the brightness of distant supernovae explosions. These appeared fainter than expected in a Universe only made up of matter and radiation. Assuming a flat Universe, the data could only be fit with the following dark energy density parameter:

$$\Omega_\Lambda \simeq 0.68 \tag{6.13}$$

A further point of interests regards the H_0 parameter, the Hubble constant. It is a fundamental value which, among other things, is instrumental in determining the age of the Universe. For this very reason, it is important to determine it accurately. Since measurements of this parameter used to come with very large uncertainties, it is conventional to define it as:

$$H_0 \equiv h \ 100 \text{ km s}^{-1} \text{ Mpc}^{-1} \tag{6.14}$$

However, the accuracy of H_0 measurements has improved drastically in the last decade, and it turns out that it gives rise to a fundamental statistical discrepancy. This constant can be measured directly from the cosmological model using *early* Universe data extracted from the CMB. The latest of these observations give:

$$h = 0.674 \pm 0.005 \tag{6.15}$$

On the other hand, the same constant can be measured with *local* Universe data, based on the distance-redshift relation. These methods are usually undertaken by building a "local distance ladder", which involves supernovae measurements; we won't explain here these methods. Interestingly, this H_0 measurement is in stark contrast with (6.15), since it gives:

$$h = 0.73 \pm 0.010 \tag{6.16}$$

The discrepancy between these two values has been dubbed the *Hubble tension*, and it is thought to be the most serious challenge to the standard cosmological model (ΛCDM). In fact, as can be seen in Figure 15, the measurements have been made more and more precise with the years, ruling out possible systematic uncertainties as being the solution. For this reason, at least part of the community thinks that the Hubble tension may be a gateway to new physics.

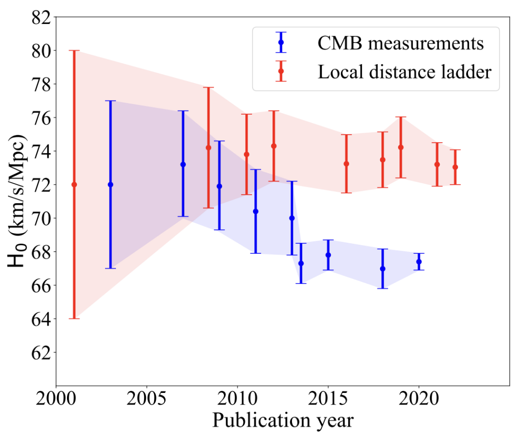

Figure 15: The Hubble tension

We have finally completed our history of the large-scale Universe. With these parameters, we can easily trace the various phases the Universe went through.
The scientific story starts at 100 GeV, where radiation is the dominant contribution to the energy content. After a

relatively brief span of time, the matter energy density starts to become non-negligible, until it starts to dominate at the "moment" of matter-radiation equality, which we have defined previously and can now calculate:

$$\begin{aligned} a_{\mathrm{eq}} &= \frac{\Omega_r}{\Omega_m} \simeq 2.9 \times 10^{-4} \\ z_{\mathrm{eq}} &\simeq 3400 \end{aligned} \tag{6.17}$$

Matter dominates the Universe for a long time, until close to now at matter-dark energy equality, when the energy provided by cosmological constant term gets the upper hand:

$$\begin{aligned} a_{\mathrm{m}\Lambda} &= \left(\frac{\Omega_m}{\Omega_\Lambda}\right)^{1/3} \simeq 0.77 \\ z_{\mathrm{m}\Lambda} &\simeq 0.3 \end{aligned} \tag{6.18}$$

Unless the nature of dark energy is wildly different than we think, the vacuum energy will dominate forever. Notice that there is no period where the energy from the spatial curvature is the leading term. This is due to the very negligible part of the energy budget made up by the Ω_k term. In fact, despite its scaling as a^{-2}, the time of matter-curvature equality $a_{\mathrm{mk}} \simeq 60$ is too far in the future, given that the energy from spatial curvature has dropped below the dark energy term already at $a_{\mathrm{k}\Lambda} \simeq 0.08$.

6.2 Derived properties

Once the density parameters Ω are known, it is not only possible to discern the complex dynamics of the Universe, but also to delve even deeper into the frameworks that we have introduced along the way, such as the deceleration parameter, the age of the Universe, the concept of horizons and the significance of distances in cosmology.

6.2.1 Acceleration or deceleration?

The deceleration parameter is defined as:

$$q \equiv -\frac{\ddot{a}}{aH^2} \tag{6.19}$$

Since the values at the denominator are strictly positive, if $q > 0$ the Universe is decelerating and viceversa, if $q < 0$. Using the second Friedmann equation and the definition of the critical density, we find the cleaner expression:

$$q(\{\Omega\}) = q(\{\Omega_r, \Omega_m, \Omega_\Lambda\}) = \frac{1}{2}\sum_i \Omega_i(1 + 3w_i) = \frac{1}{2}(\Omega_m + 2\Omega_r - 2\Omega_\Lambda) \tag{6.20}$$

Evaluating this parameter at present time, we have:

$$q_0 = q(\{2.9 \times 10^{-4}, 0.32, 0.7\}) \simeq \frac{1}{2}(0.32 - 2 \times 0.7) < 0 \tag{6.21}$$

In other words, at present time the Universe is accelerating. Since dark energy will dominate forever, in so far as our current scientific understanding is concerned, this acceleration will continue forever.
On the other hand, during the radiation (RD) and matter (MD) dominated periods, things were different. Since we only need to know the sign of q, we can draw conclusions without the need to calculate the Ωs for these two periods. In fact, we can just assign to these periods the following set of parameters: RD $\simeq \{1, 0, 0\}$ and MD $\simeq \{0, 1, 0\}$, since Ω_r and Ω_m respectively are the dominating terms. In other words, we have:

$$\begin{aligned} \mathrm{sgn}(q_{\mathrm{RD}}) &= \mathrm{sgn}(q(\{1, 0, 0\})) = +1 \\ \mathrm{sgn}(q_{\mathrm{MD}}) &= \mathrm{sgn}(q(\{0, 1, 0\})) = +1 \end{aligned} \tag{6.22}$$

Therefore, when the dominating energy was either that of radiation or matter, the Universe was decelerating.

6.2.2 The age of the Universe

The concept of acceleration and deceleration during cosmic expansion also affects the deduced age of the Universe. The function $t(a)$ was introduced last lecture, and it reads:

$$t(a, \{\Omega\}) = \frac{1}{H_0}\int_0^a \frac{da}{[\Omega_r a^{-2} + \Omega_m a^{-1} + \Omega_\Lambda a^2 + \Omega_k]^{1/2}} \tag{6.23}$$

Now, the age of the Universe is defined as $t_0 \equiv t(1)$. The calculation is practically carried out by setting some lower bound on a to when the hot Big Bang started at the temperature $t_{100\ \mathrm{GeV}}$. The calculation yields:

$$t_0 = t(\{2.9 \times 10^{-4}, 0.32, 0.7\}) \simeq 13.8 \times 10^9 \text{ years} \tag{6.24}$$

Interestingly, now it is also possible to give some perspective on the problems that a Universe without a cosmological constant term faces. In fact, in the first lecture we have alluded to a discrepancy that was found when calculating the age of a matter-only Universe (the Einstein-de Sitter model):

$$t_{\mathrm{Eds}} = t(\{0, 1, 0\}) = \frac{2}{3}\frac{1}{H_0} \simeq 9 \times 10^9 \text{ years} \tag{6.25}$$

As we have said, this number was in fact smaller than the age of some of the known stars in the Universe.

Finally, it is also interesting to calculate the times of matter-radiation and matter-dark energy equalities by setting $t_{\mathrm{eq}} = t(a_{\mathrm{eq}})$ and $t_{\mathrm{m}\Lambda} = t(a_{\mathrm{m}\Lambda})$. The integral gives:

$$\begin{aligned} t_{\mathrm{eq}} &\simeq 50000 \text{ years} \\ t_{\mathrm{m}\Lambda} &\simeq 10.2 \times 10^9 \text{ years} \end{aligned} \tag{6.26}$$

6.2.3 The Hubble horizon

The notion of a horizon, as we have discussed, is fundamental in cosmology. We have introduced the particle horizon $d_h(t)$, which is the maximum distance light could have traveled from the beginning of the Universe up until time t, and the event horizon $d_e(t)$, the maximum distance light could travel from time t to the infinite future. These two are defined in the following way:

$$\begin{aligned} d_{h,comoving}(\eta) &= \chi = \eta = c\int_0^t \frac{dt'}{a(t')} \\ d_{e,comoving}(\eta) &= \int_\eta^{\eta_f} d\eta = \eta_f - \eta = c\int_t^\infty \frac{dt'}{a(t')} \end{aligned} \tag{6.27}$$

In a diagram where the time t is replaced with the conformal time η, these two horizons are easy to draw, as they are straight lines. Finally, the notion of Hubble horizon was introduced as the distance at which objects recede at the speed of light. The comoving Hubble sphere is:

$$d_{HR}(t) = \frac{c}{a(t)H(t)} \tag{6.28}$$

Knowing the form of $a(t)$, we can now plot this horizon too. First, let's convert $d_{HR}(t)$ to conformal time, using the following relationships we derived:

$$a(t) \propto \begin{cases} t^{1/2} & \text{RD} \\ t^{2/3} & \text{MD} \\ e^{H_0\sqrt{\Omega_\Lambda}t} & \Lambda\text{D} \end{cases} \tag{6.29}$$

Using the relationship $\eta = \int_0^t dt/a$ we get:

$$a(\eta) \propto \begin{cases} \eta & \text{RD} \\ \eta^2 & \text{MD} \\ -\frac{1}{\eta} & \Lambda\text{D} \end{cases} \tag{6.30}$$

For this reason, the comoving Hubble horizon in conformal time is:

$$d_{HR}(\tau) = \frac{ca(\eta)}{a'(\eta)} \propto \begin{cases} \eta & \text{RD} \\ \eta & \text{MD} \\ -\eta & \Lambda\text{D} \end{cases} \tag{6.31}$$

This is the mathematical perspective of the arguments put forward some lectures ago. The Hubble sphere gets bigger if the Universe is expanding in a decelerating fashion, such as during radiation and matter dominated periods. This is because the distance at which objects recede at the speed of light grows, which is precisely the definition of an increasing Hubble horizon. On the other hand, an accelerated expansion, such as that caused by vacuum energy, shrinks the Hubble sphere, such that objects recede faster progressively close to us. The Hubble sphere is depicted in Figure 16, along with the particle and event horizons. Figure 16 also encapsulates neatly what we just said. There is a way of being *outside* the Hubble sphere and *inside* the past light cone, like at the bottom of the picture, when the

Universe is still decelerating. In this way an observer who founds himself outside the Hubble horizon can communicate with us in the future. Clearly, this is not possible when dark energy starts to dominate, as the top of the figure displays. The light cone closes up following the Hubble sphere, so that no observer outside it will be able to send us any information, forever.

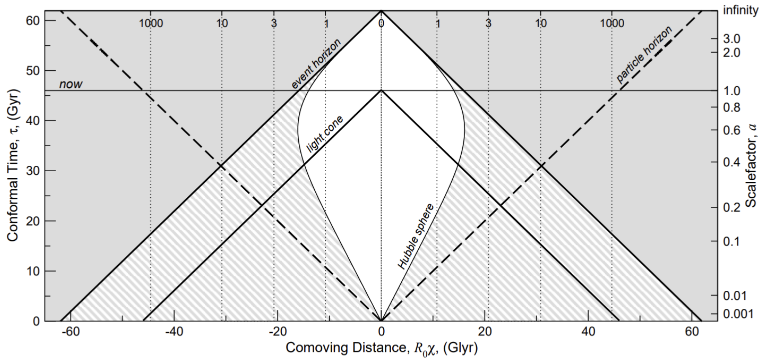

Figure 16: The conformal diagram showing the horizons.

6.2.4 Distances and the surface brightness

In cosmology, measuring the distance of an object can be quite a tricky business. We can make use of its apparent luminosity (the energy flux we detect on Earth), and then convert it to a distance using its intrinsic luminosity, provided we know it. This definition is the basis of the luminosity distance $d_L(z)$, which we found to be given by the following series:

$$d_L(z) = \frac{c}{H_0}\left[z + \frac{1}{2}(1-q_0)z^2 + \dots\right] \tag{6.32}$$

On the other hand, we can make use of the object's angular size to infer its distance. This line of reasoning brings us instead to the angular diameter distance $d_A(z)$:

$$d_A(z) = \frac{c}{H_0}\left[z - \frac{1}{2}(1+q_0)z^2 + \dots\right] \tag{6.33}$$

Using the values of the density parameters Ω, we can now plot these two functions:

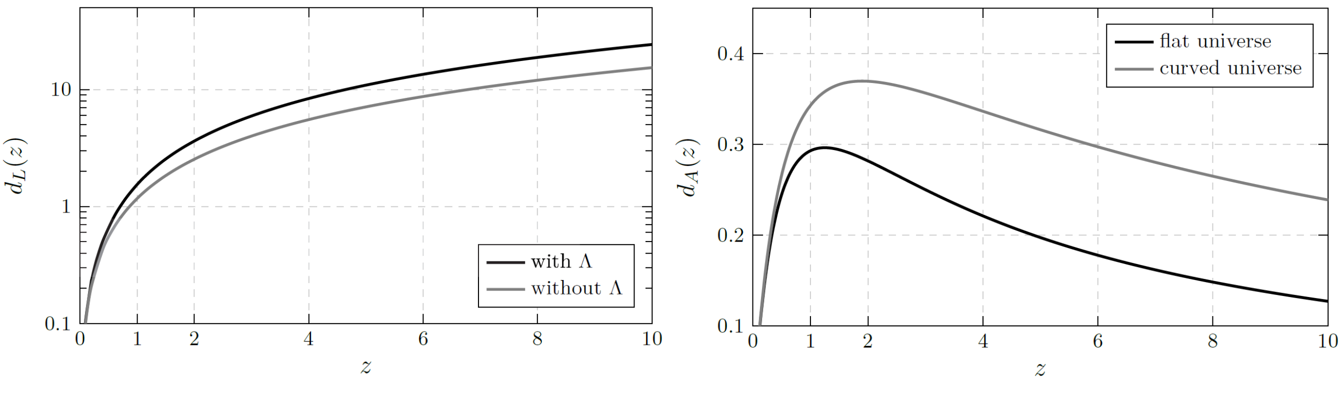

Figure 17: Distances in cosmology

Notice that the angular diameter distance exhibits apparently paradoxical behavior, that is there exists a turning point such that, beyond it, objects appear larger the more distant they are. In other words, if we look at galaxies at increasing redshifts, those that reside at greater redshifts appear bigger in the sky (remember that $\delta\theta \sim 1/d$). This turnover point occurs because the Universe is expanding. Objects that are *now* very distant actually were closer to us in the past. The light that we receive from these now-distant objects was emitted by them a long time ago, when they were nearer to us, and therefore spanned a larger angle in the sky.
Luckily, there is a way of distinguishing between the distance two objects that appear to be roughly the same dimension in the sky, but are at different redshifts. It has to do with a quantity called the *surface brightness* $\mathscr{B}$, which is defined (for extended objects) as the ratio of an object's apparent luminosity and its area:

$$\mathscr{B} \equiv \frac{L}{A} \tag{6.34}$$

In Euclidian optics, both quantities scale as d^{-2}, d being the distance to the object. For this reason we say that surface brightness is conserved. This does not happen in cosmology, where in fact we have:

$$\mathscr{B} \propto \frac{d_L^{-2}}{d_A^{-2}} \propto (1+z)^{-4} \tag{6.35}$$

This effect is called *cosmological dimming.* This solves the problem of the turnaround point of the angular diameter distance, since in reality objects actually appear fainter as their redshifts increase. This phenomenon is a big part of the reason why it is hard to see objects that are really young, that is at $z \gg 1$.

6.3 Problems on the horizon

We have finally concluded our description of the Universe at large scales. However, if the cosmological principle held perfectly, if the distribution of matter in the Universe were homogeneous and isotropic at all scales, there would be no structures (inhomogeneities) today, no galaxies and no clusters. We therefore clearly need a mechanism which gives rise to deviations from perfect uniformity, which cannot be found in the Friedmann Universe.
This problem is interestingly linked to some very serious problems that afflict this model that go by the names of *the horizon problem* and *the flatness problem.* It turns out, in fact, that trying to solve these two issues also introduces a framework with which to study the initial conditions of the Universe (what happened before $t_{100\text{ GeV}}$), and therefore also provides a way to generate seeds of initial inhomogeneities. This framework is classically called *inflation.*

7 Inflation

Large-scale homogeneity and flatness are the most fundamental and important concepts in cosmology. The implications of these simple concepts are however problematic.

7.1 The horizon problem

It is clear from the Cosmic Microwave Background that the observable components of the very early Universe, photons and baryons, constituted a system that was very uniform. The CMB is a perfect black body with temperature $\bar{T}_0 \sim 2.73$ K. How is it so? How was the Universe so uniform at such an early stage? How can two patches chosen at random in the sky possess the same temperature? The obvious explanation that comes to mind is thermalization. If we think of a gas of particles in a box with random initial conditions, the system will eventually relax to a state of maximum entropy with a shared temperature, that is the particles will find themselves in thermodynamical equilibrium. Applying the same idea to the Universe, however, fails miserably. Different parts of the CMB were so far apart at the time of recombination (when the CMB was formed) that they were not even in causal contact with each other. In other words, the particle horizon (the maximum distance light could have traveled from $t = 0$) at recombination was smaller than the distances of some patches with the same temperature in the sky. The *horizon problem* is then the following: how could causally disconnected patches have evolved to have the same exact temperature?
The problem can be quantified rather precisely. Recall the definition of the (comoving)particle horizon as the logarithmic integral of the (comoving)Hubble sphere:

$$d_h(\eta) = \int_{\log a_i}^{\log a} (aH)^{-1} d\log a \tag{7.1}$$

To solve the integral, we first derive the expression for the Hubble horizon in a flat Universe with only matter and radiation. The Friedmann equation gives us the form of the Hubble parameter:

$$H^2 = H_0^2[\Omega_m a^{-3} + \Omega_r a^{-4}] \tag{7.2}$$

where the time of equality is $a_{\text{eq}} = \Omega_r/\Omega_m \simeq 3400^{-1}$. The Hubble sphere is then given by:

$$(aH)^{-1} = \frac{1}{\sqrt{\Omega_m}} H_0^{-1} \frac{a}{\sqrt{a + a_{\text{eq}}}} \tag{7.3}$$

Therefore, the particle horizon is the following:

$$d_h(a) = \int_0^a \frac{d\log a}{aH} = \frac{2}{\sqrt{\Omega_m}} H_0^{-1}(\sqrt{a + a_{\text{eq}}} - \sqrt{a_{\text{eq}}}) \tag{7.4}$$

Evaluating the horizon today ($a_0 = 1$) and at recombination ($z_\star \simeq 1100$, $a_\star \simeq 1100^{-1}$):

$$\begin{aligned} d_{h,0} = \eta_0 &\simeq \frac{2}{\sqrt{\Omega_m}} H_0^{-1} \simeq 20\ 285 \text{ Mpc} \\ d_{h,\star} = \eta_\star &= \frac{2}{\sqrt{\Omega_m}} H_0^{-1} [\sqrt{1100^{-1} + 3400^{-1}} - \sqrt{3400^{-1}}] \simeq 265 \text{ Mpc} \end{aligned} \tag{7.5}$$

Using the comoving distance to the surface of last scattering (recombination) $\chi_\star = \eta_\star - \eta_0 \simeq 15.1$ Gpc, we can estimate easily the angle subtended by the particle horizon at recombination:

$$\theta_h = \frac{2 d_{h,\star}}{\chi_\star} \simeq 0.036 \text{ rad} \simeq 2.0^\circ \tag{7.6}$$

This neat results shows the extent of the horizon problem. Patches in the CMB that are more than 2° apart could not have interacted before recombination. Dividing the total square degrees on the sphere by 2° squared, we find that there are roughly 40000 causally disconnected patches in the CMB.
The following pictures depict the problem in a straightforward way:

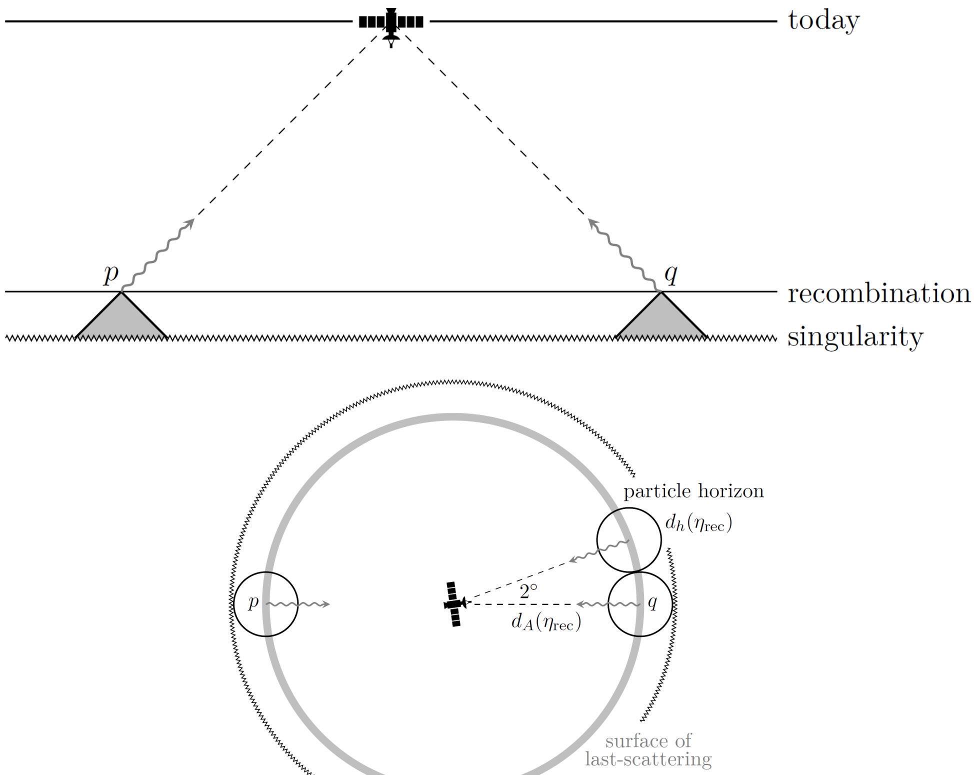

Figure 18: *Top*: Conformal diagram showing the light cones. *Bottom*: Pictorial representation of the particle horizons.

The picture on top shows two points in the sky, p and q, separated by 180 degrees. Since the particle horizon today η_0 (half of the base of the big triangle) is orders of magnitudes bigger than at recombination $\eta_\star$ (half of the base of the shaded triangle), the light cones of p and q do not overlap. They were never in causal contact before $\eta_\star$. Notice that we could have carried out the same argument if p and q were separated by only a little more than 2 degrees. Furthermore, the situation is even worse than depicted since $\eta_\star \simeq 10^{-2}\eta_0$, meaning that the shaded triangles should be even smaller. The picture on the bottom shows instead that CMB scales span much larger distances than the particle horizon at the surface of last scattering.

The horizon problem is exacerbated by the fact that the CMB actually shows small perturbations in its temperature, of the order of $\delta T/\bar{T}_0 \sim 10^{-5}$. Notwithstanding our origin of these perturbations, the fundamental problem resides in their correlations spanning acausal distances. To understand the problem, let's look at the figure below:

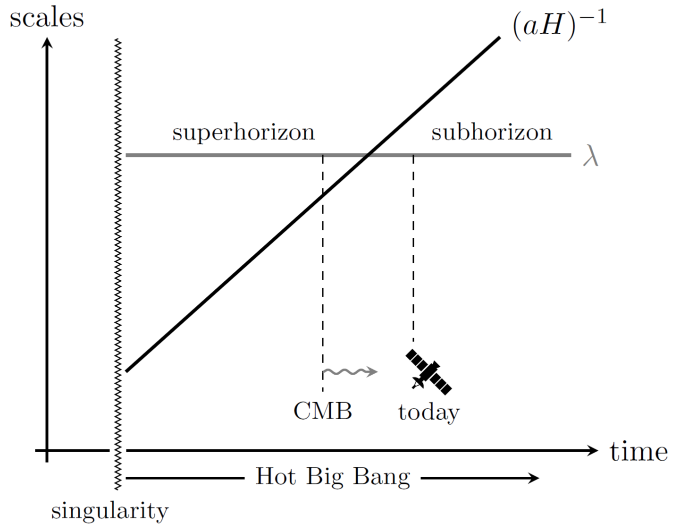

Figure 19: Comoving scales exiting the horizon at early times

The plot shows the evolution of a scale λ and the Hubble horizon $(aH)^{-1}$ in comoving coordinates. The scale, being comoving, stays the same, whereas the Hubble sphere grows, as should be clear from the derived expression (7.3). This implies that any scale that might reside inside the horizon today (subhorizon) will have found itself outside of it (superhorizon) at some early time. Then, the problem is the following. Although the horizon at recombination is about 265 Mpc, observations reveal that CMB fluctuations are correlated on scales larger that this value, even though they should not have been able to communicate in any way.
Therefore, not only is the CMB thermalised across acausal distances, but on these scales it also displays correlations between fluctuations.

7.1.1 Beware of the confusion

There is a subtle point that is worth discussing before moving forward. In fact, a fundamental concept in cosmology is that of (physical) scales entering and exiting the horizon, since only subhorizon modes are observable and are causally connected. In the discussion just above, this horizon was implied to be the Hubble sphere. But wait! Doesn't the particle horizon decide whether two points in spacetime are causally related or not? The resolution of this confusion lies in noticing that the particle and Hubble horizons are actually proportional to each other during both radiation and matter domination periods. For this reason, textbooks use the two concepts interchangeably.
To verify this claim, take the scale factor to be $a(t) \propto t^n$, where $n = 1/2, 2/3$ for radiation and matter domination respectively. The particle horizon then is:

$$d_h(t) = \int_0^t \frac{dt'}{a(t')} \propto \int_0^t dt' t'^{(-n)} = \frac{1}{1-n} t^{1-n} \tag{7.7}$$

On the other hand, the Hubble sphere grows like:

$$(aH)^{-1}(t) = \dot{a}^{-1}(t) \propto t^{1-n} \tag{7.8}$$

In this sense we see that up to numerical factors the two horizons are the same, that is $d_h(t) \propto (aH)^{-1}(t)$.
Since every (comoving)scale λ can be identified with its (comoving)wavenumber, $k = 2\pi/\lambda$, we have the following rule:

$$\begin{aligned} &k \ll aH \longleftrightarrow \text{The scale } \lambda \text{ is SUPERHORIZON} \\ &k \gg aH \longleftrightarrow \text{The scale } \lambda \text{ is SUBHORIZON} \end{aligned} \tag{7.9}$$

We say that a mode *enters the horizon* as it goes from $k \ll aH$ to $k > aH$, whereas it *leaves the horizon* if it goes from $k \gg aH$ to $k < aH$. Figuring out whether a certain scale is inside or outside the horizon is especially important since only inside the horizon can it become observable. Furthermore, this framework is critical in cosmological perturbation theory (the study of the evolution of the small perturbations around uniformity), because a given perturbation on a scale with wavenumber k is frozen in time if $k < aH$, i.e. if it lies outside the horizon, as causality prevents it from growing.

It is worth stressing that while textbooks replace the concept of particle horizon with that of the Hubble sphere, the latter is independently related to causality. In fact, the Hubble radius is the approximate distance over which light can travel in the course of one expansion time, i.e. the time it takes for the scale factor to increase by a factor

of e. In other words, the Hubble horizon provides a yardstick to measure whether two observers can interact within one expansion time.
There is, however, a fundamental difference between the two concepts. If two observers are separated by a distance greater than the Hubble sphere, they cannot communicate with each other *now* (that is in the course of one expansion time, as was explained above), but it is possible that they have communicated in *the past*. This could happen, for example, if the Hubble sphere had a period of rapid decrease, as is the case with inflation (see later). Instead, if the two observers are at some time t separated by a distance greater than the particle horizon, they could have *never* communicated before time t. This should be clear, since the particle horizon (the conformal time) traces the observer's light cone.

7.2 The flatness problem

Precise cosmic microwave background measurements have provided an upper bound on the value of today's curvature parameter $|\Omega_k| < 5 \times 10^{-3}$. This value, however, requires extremely fine-tuned initial conditions. To see why, consider the Friedmann equation for a Universe with unknown curvature:

$$H^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} \equiv \frac{8\pi G}{3}(\rho + \rho_k) \tag{7.10}$$

Using the definition of critical density $\rho_c(t) = 3H^2/8\pi G$, we get:

$$1 = \Omega(t) + \Omega_k(t) \tag{7.11}$$

The adimensional curvature parameter thus varies in time like:

$$\Omega_k(t) = -\frac{kc^2}{a^2 H^2} = \frac{H_0^2}{(aH)^2}\Omega_k \tag{7.12}$$

The *flatness problem* arises because, as one can see, $\Omega_k(t) = 0$ is a fixed point. In fact, using our solution for the Hubble radius in (7.3) one finds:

$$\Omega_k(t) = \frac{\Omega_k}{\Omega_m}\frac{a^2}{a + a_{\text{eq}}} \tag{7.13}$$

This expression implies that the curvature of the Universe decreases when going back in time. If $\Omega_k \simeq 0$ today, in the past it should have been extremely close to zero, but not exactly zero (otherwise it would still be zero today), which is a coincidence that appears rather unlikely. To quantify this, it's just sufficient to evaluate (7.13) at some early times. Choosing these to be $z_{\text{EW}} \simeq 10^{15}$, the time of the electroweak phase transition, and $z_{\text{BBN}} \simeq 10^8$, the time of Big Bang Nucleosynthesis, we have:

$$\begin{aligned} |\Omega_k(t_{\text{EW}})| &\lesssim 10^{-29} \\ |\Omega_k(t_{\text{BBN}})| &\lesssim 10^{-16} \end{aligned} \tag{7.14}$$

These extremely small values seem artificial enough that the flatness problem is also called the "fine-tuning" problem.

7.3 The solution: a period of accelerated expansion

The causality issues that arose in the standard Friedmann Universe point towards their own solution. The horizon problem, for instance, implies that the particle horizon should be orders of magnitude bigger than what we calculate it to be. Can't we then just think of a mechanism that would increase this quantity?
An obvious candidate that could generate such a mechanism is the Hubble sphere, since it is related to both the particle horizon and the causality structure. To build on this idea, recall the scaling of the Hubble horizon in figure 19. During both matter and radiation domination, the Hubble horizon's growth leads to subhorizon scales today necessarily being superhorizon at some point in the distant past, giving rise to the causality problems we discussed. In other words, this period in the Universe sees the Hubble sphere *growing* faster than physical scales:

$$\frac{d}{dt}\left(\frac{\lambda}{aH}\right) > 0 \tag{7.15}$$

It it thus possible to think of an era, sometime before radiation domination, when this condition was reversed, such that the Hubble sphere was *shrinking*. This period in the history of the Universe must surely have occurred before the era of Big Bang Nucleosynthesis, since the latter is the earliest era from which we have observational data agreeing rather well with our theory, which in turn presupposes the radiation equation of state. However, the thermal history

before nucleosynthesis is unknown.
The shrinking Hubble sphere solves the isotropy of the CMB rather straightforwardly. Scales that were outside of the horizon long ago were actually *inside* of it at the time of this primordial epoch. If this happens, the photons we see as a microwave background emitted from the last scattering surface by causally disconnected patches, actually have the same temperature because they were in causal contact (were inside the horizon) during this primordial era. Even the problem of correlated superhorizon correlations is now solved, since all scales could communicate. It is possible to visualize such a period in the Universe by extending figure 19 to the left in a symmetric fashion:

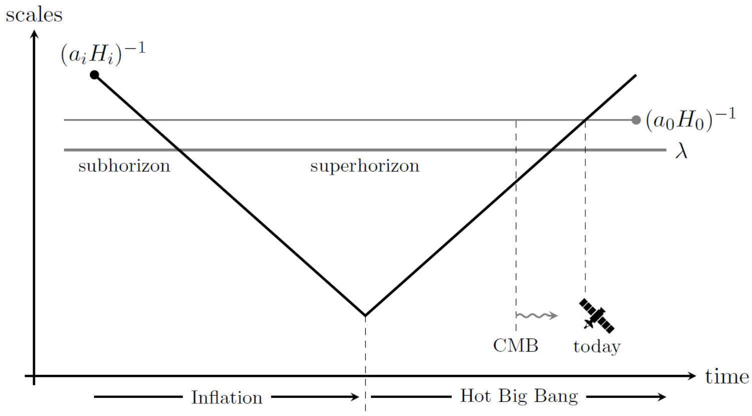

Figure 20: Scales re-entering the horizon during the period of a shrinking Hubble sphere

Notice that condition (7.15) is reversed in this new epoch:

$$\frac{d}{dt}\left(\frac{\lambda}{aH}\right) < 0 \tag{7.16}$$

implying (since λ is comoving):

$$\ddot{a} > 0 \tag{7.17}$$

This period is called *inflation*. It is a period of accelerated expansion that precedes the Big Bang and during which subhorizon scales were pushed outside the Hubble horizon. As we have changed the evolution of the Hubble horizon, now the particle horizon is way bigger than calculated in the Friedmann Universe, which was precisely our goal in the beginning.
An important application of the inflationary condition (7.17) is immediately found by inspection of the second Friedmann equation:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3P}{c^2}\right) \tag{7.18}$$

Imposing $\ddot{a} > 0$ gives:

$$\rho + \frac{3P}{c^2} < 0 \iff w_{\text{inflation}} < -\frac{1}{3} \tag{7.19}$$

The object that drives inflation needs to respect this condition on the equation of state. We already know from the study of dark energy that, in order to obtain an accelerated expansion, we need the pressure to be negative. It appears therefore that inflation was driven by a similar type of energy. It cannot be ordinary matter or radiation, since both have a positive equation of state, but the cosmological constant itself is also prohibited as it would give rise to a never ending inflationary period, while what is required is for inflation to end and transition to the radiation dominated phase. While the object driving inflation is still unknown, we will shortly provide the most commonly accepted framework.

7.3.1 Inflation and the horizon problem

From a conceptual standpoint, the horizon problem is solved. Superhorizon scales at the time of photon decoupling were subhorizon some time during the epoch of inflation. In this section we would like to offer a more detailed and quantitative description of the solution to the horizon problem.
Consider again the expression for the particle horizon:

$$d_h(\eta) = \int_{a_i}^{a} (aH)^{-1} d\log a \tag{7.20}$$

In the standard Friedmann Universe, as the Hubble horizon $(aH)^{-1}$ grows, the largest contribution to the integral comes from late times and we have $d_h \sim (aH)^{-1}$ as was previously argued. However, adding a period of a shrinking

Hubble sphere changes things dramatically, because the contribution close to early times in the integral dominates, causing the particle horizon to be much bigger $d_h \simeq (a_i H_i)^{-1} \gg (aH)^{-1}$.
To illustrate this concept, consider a flat Universe dominated by a fluid with a constant equation of state w, such that the Friedmann equation is:

$$H^2 = \frac{8\pi G}{3} a^{-3(1+w)} = a^{-3(1+w)} H_0^2 \tag{7.21}$$

where we have used the density scaling of the said fluid $\rho(t) = \rho(t_0) a^{-3(1+w)}$ we derived in (4.23) and $\Omega = 1$, since it is the only fluid in the Universe. Multiplying (7.21) by a^2, we see that the Hubble horizon then evolves like:

$$(aH)^{-1} = H_0^{-1} a^{\frac{1}{2}(1+3w)} \tag{7.22}$$

Plugging this into (7.20) yields the form of the particle horizon:

$$d_h(\eta) = \eta - \eta_i = \frac{2H_0^{-1}}{(1+3w)} \left[a^{\frac{1}{2}(1+3w)} - a_i^{\frac{1}{2}(1+3w)} \right] \tag{7.23}$$

It is crucial to analyze the initial conformal time η_i:

$$\eta_i = \frac{2H_0^{-1}}{(1+3w)} a_i^{\frac{1}{2}(1+3w)} \tag{7.24}$$

In the standard Friedmann cosmology, we have set the initial singularity at $a_i = 0$ which implies $\eta_i = 0$ because $1 + 3w > 0$. Adding inflation changes this fact entirely. We see that the initial singularity $a_i = 0$, for $1 + 3w < 0$, is pushed to (negative) infinite conformal time $\eta_i \to -\infty$, causing the particle horizon to be infinite. Clearly, this is an extrapolation due to having extended the integral (7.20) to $a_i = 0$, which is questionable at best. Today, we still don't know exactly when inflation happened and how long it lasted.
However, pushing the singularity to infinite conformal time solves the problem neatly as the following picture shows:

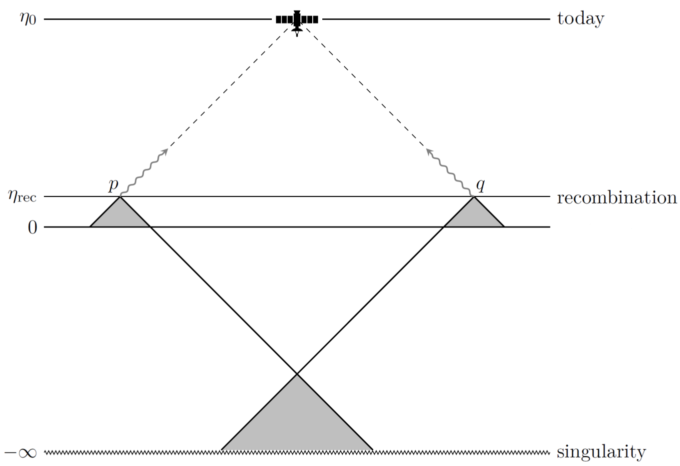

Figure 21: Conformal diagram with the inflationary period

At the top of figure 18 the conformal diagram in the Friedmann cosmology is shown. There, the spacetime points p and q were not in causal contact any time before recombination, making the correlations in temperature between the two difficult to explain. Extending the initial singularity to $-\infty$ obviously gets the job done, because now $\eta = 0$ is only the transition time between inflation and the Big Bang. The light cones of p and q can extend infinitely into the conformal past and intersect somewhere in the diagram. In other words, communication between the two was established during the period of inflation.

7.3.2 Inflation and the flatness problem

Until now a solution was only provided for the horizon problem. In this section we will see that inflation elegantly solves the flatness problem as well. Recall the evolution of the curvature parameter $\Omega_k(t)$ in (7.13), but now normalize it to some initial time t_i (before inflation):

$$\Omega_k(t) = \frac{(a_i H_i)^2}{(aH)^2} \Omega_k(t_i) \tag{7.25}$$

Given the decrease of the Hubble radius during inflation, any curvature will be quickly wiped out, leaving a Universe with zero spatial curvature. To see this more explicitly, consider a non-flat Universe dominated by a fluid with equation of state w. The Friedmann equation reads:

$$H^2 = (a_i H_i)^2 \left[(1 - \Omega_k(t_i)) \left(\frac{a}{a_i} \right)^{-3(1+w)} + \Omega_k(t_i) \right] \tag{7.26}$$

Multiplying by a^2 we get the evolution of the Hubble radius:

$$(aH)^2 = (a_i H_i)^2 \left[(1 - \Omega_k(t_i)) \left(\frac{a}{a_i} \right)^{-(1+3w)} + \Omega_k(t_i) \right] \tag{7.27}$$

Substituting in (7.25) we have:

$$\Omega_k(t) = \frac{\Omega_k(t_i) \left(\frac{a}{a_i} \right)^{(1+3w)}}{(1 - \Omega_k(t_i)) + \Omega_k(t_i) \left(\frac{a}{a_i} \right)^{(1+3w)}} \tag{7.28}$$

The solution has two fixed points, one being $\Omega_k(t_i) = 0$ as we know from before, and the other being $\Omega_k(t_i) = 1$. The qualitative behavior for a fluid with $1 + 3w > 0$, like matter and radiation, is similar: if $\Omega_k(t_i)$ is only slightly different from zero it grows rapidly away with time. For $\Omega_k(t_i) > 0$ the growth slows and converges to the asymptote $\Omega_k = 1$, an empty and Universe with negative curvature. On the other hand, in a positively curved Universe $\Omega_k(t_i) < 0$ the growth accelerates and the curvature diverges at the turnaround point $\dot{a} = 0$, when the Universe starts to fall back in on itself.
The behavior is instead crucially different for an energy density that satisfies $1 + 3w < 0$. In this case, if $\Omega_k(t_i) \neq 0$, the curvature very quickly approaches zero. Therefore the flatness problem is solved if inflation lasts long enough to wipe out any initial curvature. This has to happen to such an extent that the subsequent expansion during the Friedmann cosmology does not increase it above today's observational upper bound $\Omega_k(t_0) < 5 \times 10^{-3}$.

As one final note, notice that inflation does not change the global geometric properties of the Universe. Whether it is open or closed, it will remain so independently of inflation. What inflation does is it smooths out the local spacetime so that it looks spatially flat with great precision.

7.3.3 The duration of inflation

Before turning to the physical models that realize inflation, let us understand just how much accelerated expansion is needed to solve the problems that arise within the Friedmann cosmology.

Recall that, to solve the horizon problem, the increase of the Hubble radius during matter and radiation domination has to be compensated by a period of inflation where it is decreasing. To provide a lower bound on the duration of inflation, we require that all the fluctuations inside the horizon today were inside the horizon at early times as well. This, we remember, is because the CMB displays a uniform temperature with small correlated perturbations even on the largest scales, i.e. those that have just entered the horizon today $(a_0 H_0)^{-1}$. This condition is expressed as follows:

$$(a_0 H_0)^{-1} < (a_i H_i)^{-1} \tag{7.29}$$

where we denote with i the beginning of inflation. To quantify the amount by which the Hubble sphere decreases during inflation we keep track of the increase of the scale factor in terms of the total number of "e-foldings" (the number of factors of e) between the time when inflation begins (i) and ends (e):

$$N_{\text{tot}} \equiv \log \left(\frac{a_e}{a_i} \right) \tag{7.30}$$

From this perspective, the goal is to find the minimum number of e-foldings that are required to solve the horizon problem. To get a rough estimate, we will assume that the temperature at the end of inflation (the *reheating temperature*, we will have more to say on this later) was approximately $T_e \simeq 10^{15}$ GeV, and we will ignore the recent periods of matter and dark energy domination. Since during radiation domination we have $H^2 \propto a^{-2}$ we can write the following:

$$\frac{a_0 H_0}{a_e H_e} = \frac{a_0}{a_e} \left(\frac{a_e}{a_0} \right)^2 = \frac{a_e}{a_0} \sim \frac{T_0}{T_e} \sim 10^{-28} \left(\frac{10^{15} \text{ GeV}}{T_e} \right) \tag{7.31}$$

where we have used $T_0 \sim 10^{-13}$ GeV and $T(a) \sim 1/a$. Plugging the above relation into the condition (7.29) yields:

$$(a_i H_i)^{-1} > (a_0 H_0)^{-1} \sim 10^{-28} \left(\frac{10^{15} \text{ GeV}}{T_e} \right) (a_e H_e)^{-1} \tag{7.32}$$

In the next section we will argue that, during inflation, the Hubble parameter is approximately constant. This being the case, we have that $H_e = H_i$ and therefore:

$$N_{\text{tot}} > 64 + \log\left(\frac{T_e}{10^{15}\ \text{GeV}}\right) \tag{7.33}$$

Given the result, we say that the Universe had to expand for roughly more than 60 e-folds for the largest scales to have been in causal contact in the past.

It is also interesting to figure out the number of e-folds necessary to solve the flatness problem. In that case, the condition is that the inflationary period has to reduce the curvature to such an extent as to compensate for its increase during radiation and matter domination. In particular, the curvature today cannot be above the observational bound:

$$|\Omega_k| < 5 \times 10^{-3} \tag{7.34}$$

Recall that this value for the curvature today implies the values (7.14) at the epoch of nucleosynthesis and electroweak phase transition. In the same manner we could extrapolate Ω_k at the start of radiation domination, which corresponds to the end of inflation. Taking the temperature to be $T_e \simeq 10^{15}$ GeV, equation (7.13) gives:

$$\Omega_k(t_e) \simeq 10^{-56} \tag{7.35}$$

Finally, recalling (7.25) we have that:

$$\frac{\Omega_k(t_e)}{\Omega_k(t_i)} = \frac{(a_i H_i)^2}{(a_e H_e)^2} = \left(\frac{a_i}{a_e}\right)^2 = e^{-2N_{\text{tot}}} \tag{7.36}$$

where we have used $H_e = H_i$. Again, to get a rough estimate of the number of e-folds, taking $\Omega_k(t_i)$ of order unity we get:

$$N_{\text{tot}} \simeq 64 \tag{7.37}$$

In other words, it is enough to require 64 e-folds to solve the flatness problem. Obviously, if inflation lasted for even more e-folds the curvature today would be very close to zero with great precision. One can then say that a generic prediction of inflation is that, at present time:

$$\Omega_k = 1 \iff \text{INFLATION} \tag{7.38}$$

7.4 The physics of inflation

The advantages of a period of accelerated expansion, or equivalently of a shrinking Hubble sphere, should now be clear enough. However, we have yet to propose a physical candidate that realizes inflation. As far as we understand it, this candidate has to obey the following conditions:

(i) Its equation of state w has to satisfy $w < -1/3$ as required by the second Friedmann equation;

(ii) The period of inflation it gives rise to has to last for a sufficiently long time in order to solve both the horizon and flatness problems. As we laid out, roughly 60 e-folds are sufficient;

(iii) It should be dynamic enough such that inflation ends successfully and continuously connects to the Big Bang (the beginning of radiation domination).

7.4.1 A real scalar field

Let us focus for now on condition (i). This condition can be attained by means of a simple real scalar field, which has been dubbed the *inflaton* in modern literature. The action of a scalar field in curved spacetime reads:

$$S = \int d^4x\ \sqrt{-g}\mathcal{L} = \int d^4x\ \sqrt{-g}\left[-\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi)\right] \tag{7.39}$$

where $\sqrt{-g} = a^3$ for the FRW metric and $V(\phi)$ is the inflaton's potential which we leave unspecified for the time being. Solving the Euler-Lagrange equations

$$\partial^\mu \frac{\delta(\sqrt{-g}\mathcal{L})}{\delta(\partial^\mu\phi)} - \frac{\delta(\sqrt{-g}\mathcal{L})}{\delta\phi} = 0 \tag{7.40}$$

we obtain:

$$\ddot{\phi} + 3H\dot{\phi} - \frac{\nabla^2\phi}{a^2} + \frac{dV}{d\phi} = 0 \tag{7.41}$$

Interesting is the appearance of the friction term $3H\dot{\phi}$ due to the expansion of the Universe. The field, as it moves in its potential, suffers a friction that is not present in a static spacetime. The energy-momentum tensor of a scalar field reads:

$$T^\mu_\nu = \partial^\mu\phi\partial_\nu\phi - \delta^\mu_\nu\left[\frac{1}{2}\partial^\alpha\phi\partial_\alpha\phi + V(\phi)\right] \tag{7.42}$$

The time component T^0_0 is equal to $-\rho_\phi$ and the pressure is $P_\phi = 1/3T^i_i$:

$$\begin{aligned}\rho_\phi &= \frac{\dot{\phi}^2}{2} + V(\phi) + \frac{(\nabla\phi)^2}{2a^2} \\ P_\phi &= \frac{\dot{\phi}^2}{2} - V(\phi) - \frac{(\nabla\phi)^2}{6a^2}\end{aligned} \tag{7.43}$$

If the gradient term $\nabla\phi$ were dominant we would obtain $P_\phi = -\rho_\phi/3$, which is not enough to power inflation. We will assume that the field is homogeneous to zeroth order, consisting of a zeroth order infinite wavelength part and a perturbation which depends on the position:

$$\phi(x^\mu) = \phi_0(t) + \delta\phi(t, \vec{x}) \tag{7.44}$$

In this section we will be concerned with the homogeneous part of the field, justified by the fact that the fluctuations $\delta\phi$ are much smaller. This way we can neglect the gradient term and write the equation of state of the field as (we set $c = 1$ from here on):

$$w_\phi = \frac{P_\phi}{\rho_\phi} = \frac{\frac{\dot{\phi}_0^2}{2} - V(\phi_0)}{\frac{\dot{\phi}_0^2}{2} + V(\phi_0)} \tag{7.45}$$

Crucially, if we impose the condition:

$$V(\phi_0) \gg \dot{\phi}_0^2 \tag{7.46}$$

we obtain:

$$P_\phi = -\rho_\phi \iff w_\phi = -1 \tag{7.47}$$

From this straightforward calculation we realize that a configuration of a scalar field whose energy dominates the Universe and whose potential dominates the kinetic term can drive inflation.

7.4.2 A slowly rolling field

Let us now dwell on condition (ii), which requires inflation to last long enough. To quantify this idea, we will introduce the so-called *slow-roll parameters* and impose that condition on them. A key characteristic of inflation is that physical quantities are slowly varying, despite the very rapid expansion of space. We remember the inflationary condition to be:

$$\frac{d}{dt}(aH)^{-1} = -\frac{\dot{a}H + a\dot{H}}{(aH)^2} \equiv -\frac{1}{a}(1+\epsilon) \tag{7.48}$$

where we have introduced the slow-roll parameter ϵ:

$$\epsilon = -\frac{\dot{H}}{H^2} = -\frac{d\log H}{dN} \tag{7.49}$$

where $dN = d\log a = Hdt$. Since the Hubble radius shrinks, this parameter has to satisfy $\epsilon < 1$, meaning that the fractional change in Hubble parameter per e-fold of expansion is small. It can be showed that the *scale invariance* of primordial fluctuations generated by inflation actually requires $\epsilon \ll 1$, implying that $H =$ const. during inflation, a fact we used in the previous section.
Since it is also necessary for inflation to last sufficiently long, ϵ has to be small for a sufficiently large number of e-foldings. This condition is measured by a second slow-roll parameter κ:

$$\kappa \equiv \frac{d\log\epsilon}{dN} = \frac{\dot{\epsilon}}{H\epsilon} \tag{7.50}$$

Since we need $\epsilon \ll 1$ to last, we require also that $|\kappa| \ll 1$. These conditions can be imposed directly on the field in the following way. The inflaton being the dominating energy, the complete action during inflation is:

$$S = \int d^4x\sqrt{-g}\left(R - \frac{1}{2}\partial^\mu\phi_0\partial_\mu\phi_0 - V(\phi_0)\right) \tag{7.51}$$

Varying with respect to $g_{\mu\nu}$ and ϕ_0 yields the equations that govern the dynamics of the Universe:

$$\begin{aligned} H^2 &= \frac{1}{3M_{\rm Pl}^2}\left(\frac{1}{2}\dot{\phi}_0^2 + V(\phi)\right) \\ \ddot{\phi}_0 + 3H\dot{\phi}_0 &= -\frac{dV(\phi_0)}{dt} \end{aligned} \tag{7.52}$$

where we have introduced the reduced Planck mass $M_{\rm Pl}^2 = (8\pi G)^{-1}$. We immediately note that the two equations can be combined to find the evolution of the Hubble parameter:

$$\dot{H} = -\frac{1}{2}\frac{\dot{\phi}_0^2}{M_{\rm Pl}^2} \tag{7.53}$$

Taking the ratio of the Friedmann equation (7.52) and (7.53) yields an expression for the slow roll parameter ϵ:

$$\epsilon = -\frac{\dot{H}}{H^2} = \frac{1/2\dot{\phi}_0^2}{M_{\rm Pl}H^2} = \frac{3/2\dot{\phi}_0^2}{1/2\dot{\phi}_0^2 + V(\phi_0)} \tag{7.54}$$

Inflation therefore occurs if $\dot{\phi}_0^2 \ll V(\phi_0)$, a condition we already found before. Due to the kinetic energy of the field being small, this situation is called *slow-roll inflation*, and sees the inflaton slowly-rolling on the potential. As explained, in order for this slow-roll to persist, we require that $|\kappa| \ll 1$. This is actually equivalent to imposing the inflaton's acceleration to be small too, as we can see by introducing a new slow-roll parameter δ:

$$\delta \equiv -\frac{\ddot{\phi}_0}{H\dot{\phi}_0} \tag{7.55}$$

Taking the time derivative of (7.54) we get:

$$\dot{\epsilon} = -\frac{\dot{\phi}_0\ddot{\phi}_0}{M_{\rm Pl}^2H^2} - \frac{\dot{\phi}_0^2\dot{H}}{M_{\rm Pl}^2H^3} \tag{7.56}$$

Plugging (7.56) into the definition of κ (7.50):

$$\kappa = \frac{\dot{\epsilon}}{H\epsilon} = 2\frac{\ddot{\phi}_0}{H\dot{\phi}_0} - 2\frac{\dot{H}}{H^2} = 2(\epsilon - \delta) \tag{7.57}$$

Therefore we see that $\{\epsilon, |\kappa|\} \ll 1$ implies $\{\epsilon, |\delta|\} \ll 1$ so that the two sets of parameters are interchangeable. This line of reasoning shows that if both the speed and the acceleration of the inflaton field are small, then the accelerated expansion will continue for a long time. This is slow-roll inflation.
The conditions for prolonged inflation can equivalently be cast on the form of the potential in the following way. First, note that the condition $\epsilon \ll 1$, or $\dot{\phi}_0^2 \ll V(\phi_0)$, implies the simplification of the Friedmann equation in (7.52):

$$H^2 \simeq \frac{V(\phi_0)}{3M_{\rm Pl}^2} \tag{7.58}$$

The Hubble expansion rate is therefore strictly determined by the inflaton's potential V. Then, the condition $|\delta| \ll 1$ simplifies the inflaton's dynamics:

$$3H\dot{\phi}_0 \simeq -\frac{dV(\phi_0)}{d\phi_0} \tag{7.59}$$

This equation provides a simple relation between the slope of the potential and the field's kinetic energy, such that requiring a small kinetic energy is equivalent to requiring some degree of flatness in the potential. Plugging both simplifications (7.58) and (7.59) into (7.54) yields:

$$\epsilon = \frac{1/2\dot{\phi}_0^2}{M_{\rm Pl}^2H^2} \simeq \frac{M_{\rm Pl}^2}{2}\left(\frac{V_{,\phi_0}}{V}\right)^2 \equiv \epsilon_V \tag{7.60}$$

where we use the subscript V when a certain parameter is specified in terms of the potential. The second condition on the potential is conveniently found by combining the parameters ϵ and δ. First take the derivative of the second equation in (7.52):

$$3\dot{H}\dot{\phi}_0 + 3H\ddot{\phi}_0 = -V_{,\phi_0\phi_0}\dot{\phi}_0 \tag{7.61}$$

Now define the new parameter η_V as:

$$\delta + \epsilon = -\frac{\ddot{\phi}_0}{H\dot{\phi}} - \frac{\dot{H}}{H^2} \simeq M_{\rm Pl}^2\frac{V_{,\phi_0\phi_0}}{V} \equiv \eta_V \tag{7.62}$$

Hence, a convenient and straightforward way to judge whether a given potential $V(\phi_0)$ leads to slow-roll inflation is to check that $\{\epsilon_V, \eta_V\} \ll 1$.
Summarizing this discussion, a period of inflation driven by a real scalar field is obtained if the following field-slow-roll parameters are small:

$$\begin{aligned} \epsilon &= \frac{3/2\dot{\phi}_0^2}{1/2\dot{\phi}_0^2 + V(\phi_0)} \\ \delta &= -\frac{\ddot{\phi}_0}{H\dot{\phi}_0} \end{aligned} \tag{7.63}$$

Equivalently, the potential-slow-roll parameters have to be small:

$$\begin{aligned} \epsilon_V &= \frac{M_{\rm Pl}^2}{2}\left(\frac{V_{,\phi_0}}{V}\right)^2 \\ \eta_V &= M_{\rm Pl}^2 \frac{V_{,\phi_0\phi_0}}{V} \end{aligned} \tag{7.64}$$

The prototype for the potential $V(\phi_0)$ is plotted in the following figure:

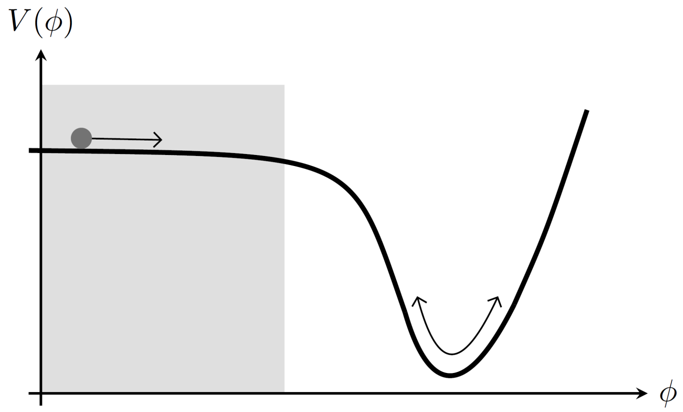

Figure 22: Form of the potential $V(\phi_0)$: slow-roll is obtained on top of the hill

As long as the field is slowly rolling on top of the hill (the shaded region), the physical quantities are slowly varying and accelerated expansion is sustained.

As a final application of inflation, we note that within the slow-roll approximations it is easy to compute the the number of e-foldings between the start and end of inflation. Denoting by $\phi_{0,i}$ and $\phi_{0,e}$ the values of the inflaton at the beginning and end of inflation, $N_{\rm tot}$ is calculated as:

$$\begin{aligned} N &= \int_{t_i}^{t_e} H dt \simeq \int_{\phi_{0,i}}^{\phi_{0,e}} d\phi_0 \, \frac{H}{\dot{\phi}_0} \simeq -3 \int_{\phi_{0,i}}^{\phi_{0,e}} d\phi_0 \, \frac{H^2}{V'} \\ &= -\frac{1}{M_{\rm Pl}^2} \int_{\phi_{0,i}}^{\phi_{0,e}} d\phi_0 \, \frac{V}{V'} \end{aligned} \tag{7.65}$$

The field values $\phi_{0,i}$ and $\phi_{0,e}$ are calculated at the boundary of the interval where $\epsilon_V < 1$, in the sense that the beginning and end of inflation are defined by $\epsilon(t_i) = \epsilon(t_e) = 1$, which implies that the accelerated expansion is over:

$$\epsilon = -\frac{\dot{H}}{H^2} = 1 - \frac{\ddot{a}}{H^2 a^2} = 1 \iff \ddot{a} = 0 \tag{7.66}$$

Since the first in (7.63) relates ϵ to the inflaton, we see that $\epsilon = 1$ is satisfied if $\dot{\phi}_0^2 = V(\phi_0)$, that is the field starts to acquire enough kinetic energy such that it becomes of the same order of the potential energy. This happens, for instance, when the field starts rolling down the hill in figure (22). This discussion segways nicely into the final section about the complex topic of reheating.

7.4.3 The end of inflation and reheating

We now finally turn to the condition (iii) which regards the end of the inflationary era.
The inflaton, being a scalar field, is a dynamic enough object that it accommodates neatly for the end of accelerated

expansion. In fact, as was explained just above, inflation ends when the energy associated with the inflaton becomes smaller than the kinetic energy of the field. Usually this happens near the minimum of the potential, as in figure (22). However, at the end of inflation the Universe is typically in a highly non-thermal state, due to inflation's ability to homogeneise the Universe and thus to leave it at effectively zero temperature. For this reason, any successful theory of inflation must explain not only how inflation ends, but how the cosmos was *reheated* to the high temperatures required for the radiation dominated epoch to begin. At the very least primordial nucleosynthesis requires the Universe to be close to thermal equilibrium at a temperature around 1 MeV. Remarkably, the scalar field framework can also account for this phenomenon. Since the theories complicate substantially, in this section we will only illustrate the main ideas. The "old" theory of reheating, developed shortly after the first inflationary theories, was based on single particle decays. The energy density of the inflaton is converted to ordinary particles which quickly thermalize. Since the said energy density was the same everywhere (except for small perturbations) in the Universe, the temperature was likewise uniform. This *reheating temperature* $T_{\rm RH}$ can be calculated by assuming an instantaneous conversion of the inflaton's energy density into radiation, when the decay width of the inflaton Γ_ϕ is equal to H (otherwise the expansion prevents the decay). As we said, the inflaton field, when inflation ends, is near the minimum of the potential at some $\phi_0 \simeq \phi_m$. Here, it executes oscillations and the potential can be approximated by:

$$V(\phi_0) \simeq \frac{1}{2}V''(\phi_m)(\phi_0-\phi_m)^2 \equiv \frac{1}{2}m^2(\phi_0-\phi_m)^2 \tag{7.67}$$

The equation of motion for ϕ_0 becomes:

$$\ddot{\phi}_0 + 3H\dot{\phi}_0 + m^2(\phi_0-\phi_m) = 0 \tag{7.68}$$

The solution is:

$$\phi_0(t) = \phi_i\left(\frac{a_i}{a}\right)^3\cos(m(t-t_i)) \tag{7.69}$$

where now the subscript i denotes the beginning of oscillations. The equation satisfied by the energy density in the oscillating field over many oscillations is:

$$\begin{aligned}\langle\dot{\rho}_\phi\rangle &= \left\langle\frac{d}{dt}\left(\frac{1}{2}\dot{\phi}_0+V(\phi_0)\right)\right\rangle = \left\langle\dot{\phi}_0\left(\ddot{\phi}_0+V'(\phi_0)\right)\right\rangle \\ &= \left\langle\dot{\phi}_0\left(-3H\dot{\phi}_0\right)\right\rangle = -3H\left\langle\dot{\phi}_0^2\right\rangle = -3H\left\langle\rho_\phi\right\rangle\end{aligned} \tag{7.70}$$

where we used the equipartition of the energy density during oscillations $\langle\dot{\phi}_0^2/2\rangle = \langle V(\phi_0)\rangle = \langle\rho_\phi/2\rangle$. The solution is (removing averaging symbols):

$$\rho_\phi = (\rho_\phi)_i\left(\frac{a_i}{a}\right)^3 \tag{7.71}$$

The Hubble parameter is then:

$$H^2 = \frac{1}{M_{\rm Pl}^2}\rho_\phi = \frac{(\rho_\phi)_i}{M_{\rm Pl}^2}\left(\frac{a_i}{a}\right)^3 \tag{7.72}$$

Equating H to Γ_ϕ allows to find an expression for $(a_{\rm RH}/a_i)$, and therefore for the reheating temperature. As we argued, we could simplify matters by assuming that all the available energy density is instantaneously converted into radiation at the time $a_{\rm RH}$. Follwing this logic, we set ρ_ϕ equal to the radiation energy density:

$$\rho_r = \frac{\pi^2}{30}g_\star T_{\rm RH}^4 \tag{7.73}$$

where $g_\star$ is the effective number of relativistic degrees of freedom at the temperature $T_{\rm RH}$ (see the theory on Big Bang nucleosynthesis). Equating H to Γ_ϕ results in:

$$T_{\rm RH} = \left(\frac{90}{g_\star\pi^2}\right)^{1/4}\sqrt{\Gamma_\phi M_{\rm Pl}} \tag{7.74}$$

7.4.4 The inflaton and large-scale structure

The realization of inflation via a real scalar field is a remarkable framework. Not only does it solve elegantly the horizon and flatness problems, but it is also dynamical enough to last a long time and end by reheating the Universe, transitioning continuously into the radiation dominated epoch.

However, the inflaton also realizes a natural mechanism that gives rise to the deviations from inhomogeneity that are observed in the Universe on scales $\lesssim$ 100 Mpc (even intuitively, at a local level galaxies are distributed far from homogeneously). Remember the inflaton's decomposition:

$$\phi(x^\mu) = \phi_0 + \delta\phi(t,\vec{x}) \tag{7.75}$$

It turns out that treating the perturbations $\delta\phi$ quantum mechanically sets the initial conditions for the Universe: before inflation it is riddled with inhomogeneities due to the uncertainty principle. These are pushed outside the horizon during inflation and start to grow only when they re-enter during radiation and matter domination. In particular, it can be shown that when the universe is matter dominated the perturbations will grow because, roughly, a region whose initial density is slightly higher than the mean will attract its sorroundings more strongly than average. Therefore overdense regions become more overdense over time, and underdense regions become more rarefied as matter will flow away from them.

The exact rate at which perturbations grow will depend on the cosmological model, and in particular this growth will stop when dark matter comes to dominate the Universe. In any case, once a certain region is overdense enough, it will stop growing and start to collapse, which is the start of the virialization of large objects. In general, dark matter starts growing earlier than baryonic matter, and nonlinear, quasi equilibrium dark matter objects are called dark matter halos; they are the potential wells into which the baryons will eventually fall and form galaxies. Over time, these galaxies form clusters and superclusters until dark matter starts to dominate the Universe.

Inflation provides a beautiful and natural link between the quantum mechanical microphysics and the macrophysics of the Universe. This is one of the few synergies that have been found between high energy physics and gravity.