

Simulations of Cosmic Structure Formation

MAUCA Meteor 2018 – Course Notes

Instructor:
Oliver HAHN
oliver.hahn@oca.eu



Contents

1	The Homogeneous Universe	2
1.1	The Expansion of the Universe	2
1.2	Comoving frames and peculiar velocities	6
1.3	The initial singularity and the CMB	7
1.4	Horizons	8
1.5	Some Thermodynamics	9
1.6	A primer on nucleosynthesis	10
1.7	Evidence for Dark Matter from nucleosynthesis	11
1.8	A picture of the Big Bang Universe	12
2	The Inhomogeneous Universe	14
2.1	The CMB: Temperature Fluctuations	14
2.2	Evolution of Small Perturbations – Eulerian Perturbation Theory	16
2.3	Origin of Inhomogeneities and Inflation	20
2.4	Properties of Inhomogeneities – Gaussian random fields	22
2.5	How to generate a realisation of the early Universe - Cosmological Initial Conditions	24
2.6	Lagrangian Perturbation Theory and Zel’dovich’s approximation	25
3	The Nonlinear Universe	28
3.1	Modelling Dark Matter	28
3.2	Collisionless simulations in one dimension	29
3.3	Collisionless simulations in three dimensions	33
3.4	Halo Finding and Merger Trees	38
3.5	Quantifying Large-Scale Structure	39
4	The Baryonic Universe	40
A	Hints for the computational exercises	41
A.1	Numerically integrating the Friedmann equation in PYTHON	41
A.2	A linear dark matter + baryon model in PYTHON	42
A.3	Generating a realisation of the young Universe	45
A.4	Plotting point distributions in projection and in 3D	48
A.5	Simulating Plane Wave Collapse	48

Chapter 1

The Homogeneous Universe

1.1 The Expansion of the Universe

The expansion of the Universe follows readily from the equations of general relativity, a fact that has been realised first by Einstein [3] and that he first considered a failure of his theory. To rectify it, he introduced the cosmological constant Λ , which allows for a static solution that is however unstable as he immediately realised. The discovery of the expansion by measurements of the recession of nearby galaxies by Lemaitre and Hubble then showed that the Universe is indeed expanding and the Λ -term has been no longer considered until more recent data indicated that it is indeed necessary to describe the expansion correctly – but this time not to allow for a static solution, but to match the accelerated expansion. We will not concern us here with relativistic theory – that is the subject of another course in MAUCA – instead we will stay mainly in the Newtonian picture.

1.1.1 Newtonian Cosmology

In fact, one can get the necessity for an expanding Universe already from Newtonian theory. We start with Poisson's equation which relates the local density ρ to the second derivative of the gravitational potential ϕ as

$$\nabla^2 \phi = 4\pi G \rho, \quad (1.1)$$

which can be rewritten in spherical coordinates (where $\nabla^2 \phi = r^{-2} \partial_r (r^2 \partial_r \phi)$) as

$$\partial_r \phi = \frac{4\pi G}{r^2} \int_0^r \rho r'^2 dr' \implies \ddot{r} = \frac{4\pi G}{3} \bar{\rho} r, \quad (1.2)$$

where we have used that $\ddot{r} = \nabla^2 \phi$ and assumed $\rho = \bar{\rho}$ to be constant in space. We can integrate once w.r.t. time (assuming $\bar{\rho}$ is also constant in time) to find

$$\dot{r}^2 - \frac{8\pi G}{3} \bar{\rho} r^2 = E, \quad (1.3)$$

where E is a constant of integration. If we write $r = a(t) \cdot r_0$, then we can express this equation as

$$\left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi G \bar{\rho}}{3} - \frac{K c^2}{a^2}, \quad (1.4)$$

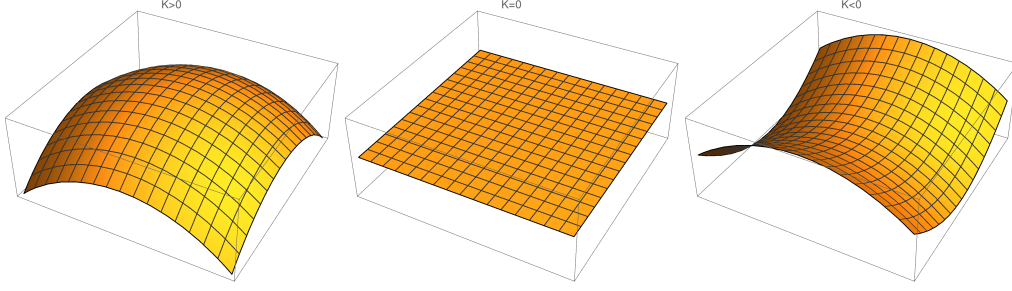


Figure 1.1: The influence of the total density on the curvature of the Universe. A universe in excess of the critical density $\bar{\rho} > \rho_{\text{crit}}$ is positively curved, a universe at the critical density is flat and a universe short of the critical density $\bar{\rho} < \rho_{\text{crit}}$ is negatively curved (hyperbolic).

where $K = -E/(cr_0)^2$. K is called the curvature and is zero only if $E = 0$, otherwise $K = \pm 1$ since we can choose r_0 accordingly so that this is the case. The function $a(t)$ that we have introduced is called the [scale factor of the Universe](#) and describes how length scales transform over time. We can thus say that $\dot{a} > 0$ corresponds to an expanding universe and $\dot{a} < 0$ to a contracting universe. Similarly, a universe in which $\ddot{a} > 0$ is accelerating, while one in which $\ddot{a} < 0$ is decelerating. A universe would be static if $\dot{a} = 0$ and only remain static if also $\ddot{a} = 0$.

The ratio $\dot{a}(t)/a(t)$ at the present time $t = t_0$ is called the [Hubble constant](#)

$$H_0 \equiv \dot{a}_0/a_0. \quad (1.5)$$

which is often also written in terms of the [Hubble parameter](#) h defined so that $H_0 = 100 h \text{ km/s/Mpc}$. The value of the curvature term K is then uniquely defined in terms of the density at the present time $\bar{\rho}_0$ and the Hubble constant. The curvature term is zero exactly in the case when the density is

$$\bar{\rho}_0 = \rho_{\text{crit},0} \equiv \frac{3H_0^2}{8\pi G}, \quad (1.6)$$

which we call the [critical density of the Universe](#). Given such a unique value for the density, we can express the mean density in units of this critical density as a [density parameter](#)

$$\Omega_0 \equiv \bar{\rho}_0/\rho_{\text{crit},0}. \quad (1.7)$$

This allows us to write the curvature in a particularly simple form as $K = H_0^2 a_0^2 (\Omega_0 - 1)$ so that

$$K = \begin{cases} -1 & : \text{negative curvature} \\ 0 & : \text{no curvature} \\ +1 & : \text{positive curvature} \end{cases} \quad \text{corresponds to} \quad \begin{cases} \Omega_0 < 1 \\ \Omega_0 = 1 \\ \Omega_0 > 1 \end{cases} \quad (1.8)$$

Finally, one customarily defines the [deceleration parameter](#) q_0

$$q_0 \equiv -\frac{\ddot{a}_0}{a_0} \left(\frac{\dot{a}_0}{a_0} \right)^{-2} = -\frac{\ddot{a}_0 a_0}{\dot{a}_0^2} \quad (1.9)$$

that characterises whether the expansion of the universe is currently accelerating ($q_0 < 0$) or decelerating ($q_0 > 0$).

We note that the expansion with time as $a(t)$ causes all length scales to expand, so also the wavelengths of photons. A photon emitted at a time t will thus experience a scaling of its wavelength after emission with λ_{em} so that $\lambda_{\text{obs}} = \lambda_{\text{em}} a_0 / a(t)$ and one typically defines this in term of the [cosmological redshift](#)

$$z \equiv \frac{\lambda_{\text{obs}}}{\lambda_{\text{em}}} - 1 = \frac{a_0}{a(t) - 1}. \quad (1.10)$$

If $a(t)$ is monotonically increasing, both a and z can thus be used to label time in an expanding Universe. Since photons travel with the speed of light, we will see distant photons coming from a time when the Universe was smaller (and thus younger) than today. The redshift of a photon thus indicates both its distance and the age of the Universe when it was emitted rendering it possible to peak into 'the past' when observing the high-redshift Universe.

In this derivation, we have required that $\bar{\rho}$ does not change with time. For ordinary matter that would dilute with the expansion of the universe, this means that $\bar{\rho} \propto a^{-3}$. We will investigate this next, when we consider the more general form of equation (1.4) next.

1.1.2 Friedmann's Equation

Our derivation above has only used Newtonian physics, however, the main result, eq. (1.4), is identical to a derivation from general relativity for a homogeneous and isotropic universe. In the relativistic derivation, one however sees more nicely that there can be multiple contributions to $\bar{\rho}$ whose physical density over time behaves differently with the expansion of the Universe.

- [nonrelativistic matter](#) is the ordinary matter (dominated by rest mass) we imagine to dilute as $\rho_m \propto a^{-3}$ with the expansion of the Universe
- [relativistic matter](#) (such as photons) is matter dominated by their relativistic momentum and its energy density changes as $\rho_r \propto a^{-4}$ (the number of photons also dilutes as a^{-3} , but their energy is also redshifted as a^{-1}).
- [cosmological constant](#) is the beast first introduced by Einstein in 1917 and now necessary to explain the accelerated expansion of the Universe. Its density $\rho_\Lambda = c^2 \Lambda / 8\pi G = \text{const.}$ does not dilute with the expansion of the Universe.
- [curvature and other topological terms](#) we have already seen above that the curvature appears as a term $\propto a^{-2}$ which means that as the Universe expands, also any non-zero curvature will be stretched out and thus reduced. In more exotic models, one could also construct other topological contributions that would contribute e.g. like $\propto a^{-1}$ (domain walls).

All these components can be conveniently expressed in terms of their density parameters Ω and contribute jointly to the [Friedmann equation](#), first derived in 1922 by the Russian physicist Alexander Friedmann [4],

$$H^2(t) = \left(\frac{\dot{a}}{a} \right)^2 = H_0^2 [\Omega_{r,0} a^{-4} + \Omega_{m,0} a^{-3} + \Omega_{k,0} a^{-2} + \Omega_{\Lambda,0}], \quad (1.11)$$

$$= H_0^2 [\Omega_r(a) + \Omega_m(a) + \Omega_k(a) + \Omega_\Lambda] \quad (1.12)$$

where one for convenience defines $\Omega_{k,0} = 1 - \Omega_{r,0} - \Omega_{m,0} - \Omega_\Lambda$ as the contribution from curvature if the total density does not amount to the critical value, i.e. $\bar{\rho}_0 \neq \rho_{\text{crit},0}$ in which case $\Omega_{k,0} \neq 0$.

We have no fundamental theory that predicts values for the large amount of constants that appear in the Friedmann equation ($H_0, \Omega_m, \Omega_r, \Omega_\Lambda, \Omega_k$) and they instead have to be measured from cosmological

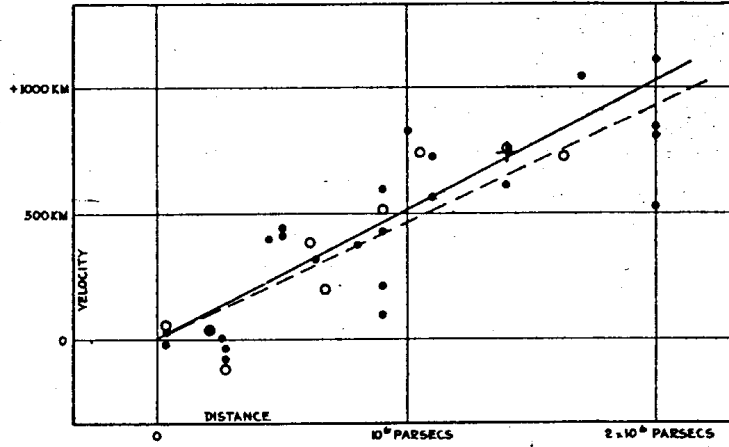


FIGURE 1

Figure 1.2: The original diagram that Edwin Hubble made in 1929 relating the distance and the recession velocity of nearby galaxies and the linear fit he provided of the form $v = H_0 d$. As you can see, his original value ($H_0 \sim 500$ km/s/Mpc) is quite different from the precision value we have today ($H_0 = 67.74 \pm 0.46$ km/s/Mpc). The discrepancy is mainly due to Hubble's determination of distances being wrong by about a factor of 7. However, the factor was roughly the same for all his galaxies, so that he indeed measured the expansion of the Universe.

observations. How this done is discussed in other MAUCA courses/meteors. The most recent values for these constants are found from the combination of various measurements [7] and are $H_0 = 67.74 \pm 0.46$ km/s/Mpc, $\Omega_m = 0.3089 \pm 0.0062$, $\Omega_\Lambda = 0.6911 \pm 0.0062$ and $\Omega_k \sim 0$. We will calculate the Ω_γ due to photons below. There is also a contribution Ω_ν from massive neutrinos which is estimated to be about 50% of Ω_γ and together they can be thought to make up $\Omega_r \simeq 8.4 \times 10^{-5}$ since the neutrinos have only recently become non-relativistic.

In Figure 1.2, you can see Hubble's original plot of the local expansion of the Universe ($v \sim H_0 d$), and his linear fit to the velocity/distance pairs [5]. Although his distance measurements were wrong by about a factor of seven, this is commonly seen as the discovery of the expansion of the Universe. Lemaître had similar measurements at the time [6], and in fact published them two years before Hubble in a Belgian journal in French (that apparently not many people read), but did not bother to re-publish them in English after Hubble's results came out.

Exercise 1 While today $\Omega_\Lambda > \Omega_m > \Omega_r$, this was not always the case. Determine the epoch (in terms of z and a) of equality between the cosmological constant and the matter energy density, as well as that of matter and radiation equality.

Exercise 2 Write a PYTHON script to numerically integrate the Friedmann equation (1.12) and plot the solution $a(t)$ for the best-fit model given above. Play with the parameters and discuss what changes. If one assumes a universe containing only matter ($\Omega_m = 1$), or only radiation ($\Omega_r = 1$), or only a cosmological constant ($\Omega_\Lambda = 1$), then analytic solutions for the Friedmann equation exist. Derive them and include them

in your plot. Compare your numerical solution against the analytic solutions for these cases. Discuss why Einstein's solution of fixing an Ω_Λ so that the Universe becomes static is unstable.

Hint for this exercise You can find a code snippet that shows how to carry out such a numerical integration in PYTHON in Appendix A.1

1.2 Comoving frames and peculiar velocities

In an expanding Universe, a trajectory $\vec{r}(t)$ of a particle can be separated into the part that is due to the expansion of the Universe and the so-called **co-moving** trajectory $\vec{x}(t)$ so that

$$\vec{r}(t) = a(t)\vec{x}(t) \quad (1.13)$$

the velocity is then of course given by

$$\vec{v}(t) = \dot{\vec{r}}(t) = a(t)\dot{\vec{x}}(t) + \dot{a}(t)\vec{x}(t) = a(t)\dot{\vec{x}}(t) + \frac{\dot{a}}{a}\vec{r}(t) = v_{\text{pec}} + v_{\text{exp}} \quad (1.14)$$

and we call $\vec{v}_c(t) = \dot{\vec{x}}(t)$ the **co-moving peculiar velocity** and $\vec{v}_{\text{pec}} = a(t)\vec{v}_c$ the **physical peculiar velocity** as they represent the genuine motion of the particle. For a particle at rest $\vec{x}(t) = \text{const}$, the first term vanishes and the second term shows a distance dependent motion that is due to the expansion alone. This second term is thus usually called the **Hubble flow** as it describes the increasing recession velocity of galaxies with distance from us measured by Hubble (and Lemaître) in 1929 and is the term responsible for the cosmological redshift.

Let us next consider how the expansion of the Universe acts on a freely moving particle. From the equation above, we find that the change in peculiar velocity must be

$$dv = -\frac{\dot{a}}{a}dr = -\frac{\dot{a}}{a}vdt. \quad (1.15)$$

We can rewrite this equation as $d \log v = -d \log a$ and integrating then yields

$$\frac{v_2}{v_1} = \frac{a(t_1)}{a(t_2)}, \quad (1.16)$$

which implies that peculiar velocities decay like $v_{\text{pec}} \propto a^{-1}(t)$. Note that this is true of course also for the momentum of nonrelativistic particles, i.e. $p \propto a^{-1}$ and it also turns out to be true for photons (this can also be seen from the $p = h\nu$ for photons and $\nu = c/\lambda$ where $\lambda \propto a$ due to the redshift).

Any peculiar motion along the line of sight (v_{\parallel} shall be the component of the physical peculiar velocity parallel to the line of sight) will of course introduce an additional redshift given by the relativistic Doppler formula

$$1 + z_v = \sqrt{\frac{1 + v_{\parallel}/c}{1 - v_{\parallel}/c}} \simeq 1 + \frac{v_{\parallel}}{c}, \quad (1.17)$$

where the final approximation is valid for $v \ll c$. The total observed redshift is then given by

$$1 + z_{\text{obs}} = (1 + z_v)(1 + z_{\text{cosm}}) \simeq (1 + z_{\text{cosm}})\left(1 + \frac{v_{\text{pec},\parallel}}{c}\right). \quad (1.18)$$

Since for distant galaxies, we cannot measure distance directly, but have to use the redshift as a distance estimator, peculiar motion will lead to what is called [redshift-space distortions](#).

Another, related, important effect is the [finger-of-god effect](#). This appears for objects that are gravitationally bound and thus have an internal [velocity dispersion](#), such as galaxy clusters. Let's assume we can model this velocity dispersion along the line of sight with a single number σ_v . Then its apparent dispersion in redshift around the cosmological redshift will be

$$\sigma_z = \frac{\sigma_v}{c}(1 + z_{\text{cosm}}). \quad (1.19)$$

Exercise 3 Assume a gravitationally bound galaxy cluster of mass $M = 10^{15} M_\odot$ and radius $R = 2.5 \text{ Mpc}$ at $z = 0.1$ in virial equilibrium, i.e. $2E_{\text{kin}} + E_{\text{pot}} = 0$. What is its velocity dispersion, and what is its extent in redshift space? Do you understand where the name "finger-of-god" comes from?

1.3 The initial singularity and the CMB

If the Universe is expanding over time, it obviously contracts if we go back in time. If $a \rightarrow 0$ as $t \rightarrow 0$, then the Universe will have an initial singularity and a finite age. Clearly, we are entering a regime where we wouldn't expect our known physics to work just before the singularity would arise and this area is of much debate in the quantum gravity communities. For us it suffices to know that there is a finite time and that the energy density diverges for all components except the cosmological constant as $a \rightarrow 0$. Long before that regime, we would however expect that photons and atoms become so densely packed that they form a common plasma in equilibrium and that if such a plasma phase existed, we should see relic photons that were released when the plasma recombined to form neutral gas and the photons were free to travel without further scattering towards us. Since during such a plasma phase we imagine photons and ions to be in thermal equilibrium at the same temperature, we expect the photons to follow a blackbody spectrum at the time of recombination of the plasma. As the photons travel to us, this spectrum is only redshifted but not distorted. The energy density of radiation emitted by a black body of temperature T is

$$u_\lambda(T) = \frac{8\pi hc}{\lambda^5} \left[\exp\left(\frac{hc}{\lambda k_b T}\right) - 1 \right]^{-1}, \quad (1.20)$$

for which we note that the shape of the spectral energy distribution (the term in the square brackets) does not change its shape under red shifting if $\lambda \times T = \text{const}$ which means that we can always find a temperature $T \propto a^{-1}$ that describes the spectrum. The overall amplitude will however change and it should better agree with the scaling $\propto a^{-4}$ that we expect for photons! To see this, we integrate the spectrum over all wavelengths.

$$u(T) = \int_0^\infty u_\lambda(T) d\lambda = \frac{4\sigma}{c} T^4, \quad \text{where} \quad \sigma = \frac{2\pi^5 k^4}{15c^2 h^3} \quad (1.21)$$

is the Stefan-Boltzmann constant. And indeed we recover that $u \propto a^{-4}$ if $T \propto a^{-1}$, the black body spectrum will thus retain its spectral form and just appear redshifted to lower temperatures. This relic radiation, the [cosmic microwave background radiation \(CMB\)](#), has indeed been measured by Penzias & Wilson in 1964 (Nobel prize 1978) to be at a temperature $T \sim 3\text{K}$. Its spectral shape has later (most clearly by the COBE satellite) been determined to be a perfect black body at $T \sim 2.725\text{K}$. From this we can directly calculate $\Omega_\gamma = u(T = 2.725\text{K})/\rho_{\text{crit},0}c^2 = 2.47 \times 10^{-5} h^{-2}$. A calculation taking into

account ionisation equilibrium in a plasma shows that the plasma should combine to a neutral gas and the photons should stream freely when the temperature drops below ~ 3000 K at a redshift around $z = 1000$ (i.e. $a \sim 10^{-3}$).

Exercise 4 Calculate the age of the Universe for the best fit parameters given before. Calculate it also for a universe that would consist of normal matter only ($\Omega_m = 1$). We expect quantum gravity to become important when the Compton wavelength of a particle $\lambda_c = h/(mc)$ becomes comparable to its Schwarzschild radius $R_s = 2Gm/c^2$. From this one can write down an energy (mass) scale, a length scale and a time scale (all named after Max Planck). Write them down and give their values. At what redshift do these scales become important?

1.4 Horizons

The finite age of the Universe, along with the finite speed of light lead to light being able to travel only a finite distance. This gives rise to two concepts of horizon – meaning the maximum distance light can have travelled or will be able to travel.

1.4.1 Particle horizon

The finite age of the Universe implies that a ray of light can only have travelled a finite distance. In a short time interval dt , a light ray will have propagated a proper distance $a dx$, where x is a comoving scale, so $c dt = a dx$. The physical distance that a signal could have travelled since the big bang until time t is thus given by

$$d_H(t) = a(t) \int_0^t \frac{c dt'}{a(t')}, \quad (1.22)$$

which is the **particle horizon** at time t . So although the universe was very small just after the big bang, its expansion prevented a causal connection between points outside the particle horizon. If we find regions of the Universe that appear similar and that are separated by more than d_H (and we do), we have a causality problem. This is usually referred to as the **horizon problem**. We will discuss a possible solution to this problem (and others) in one of the next lectures.

1.4.2 Event horizons

Likewise, there may exist another horizon in the Universe. If there is a distance beyond which no light ray can travel for all future time ($t \rightarrow \infty$), we call this an event horizon. The rationale to calculate it is the same as for the particle horizon, we just integrate from t to infinity this time. The physical size of the event horizon is then

$$d_E(t) = a(t) \int_t^\infty \frac{c dt'}{a(t')}. \quad (1.23)$$

For an event horizon to exist, the universe needs to expand quickly enough. If an event horizon exists, it means that it will never be possible to have any knowledge about what's outside of it.

Exercise 5 Does our Universe, given the best fit cosmological parameters, have an event horizon? Does a flat $\Omega_m = 1$ universe have one? Which ingredient of the Friedmann equation is necessary to produce an even horizon? What do you think are the implications for the future of universes that have an event horizon?

1.5 Some Thermodynamics

As we have briefly mentioned above, the universe becomes hotter and denser as we go to earlier times. This leads to a heat bath in which many processes can occur in equilibrium. As the heat bath decreases in energy with the expansion of the Universe, these processes will freeze out and stop. We will now lay some foundations for understanding the physics behind this. In general, one express the statistical properties of a large number of particles in terms of their distribution function, which gives their density in six-dimensional phase space (3 positions \times 3 momentum coordinates). The distribution function $f(\mathbf{x}, \mathbf{p}, t)$ is then a positive function of these six coordinates and time. The number density, energy density and pressure which are functions only of the position can then be given by integrating out the velocity dimensions (marginalising over them). Since we assume that the Universe is homogeneous, they will be identical at every point in space and it suffices to consider them as functions of time only. Since the Universe is also isotropic, there will also be no directional dependence in the momenta \mathbf{p} , so that we can simply work with the modulus of the momentum p . We then have

$$n(t) = 4\pi \int f(p, t) p^2 dp \quad (1.24)$$

$$\rho(t) = \frac{4\pi}{c^2} \int E(p) f(p, t) p^2 dp \quad (1.25)$$

$$P(t) = 4\pi c^2 \int \frac{p^2}{3E(p)} f(p, t) p^2 dp. \quad (1.26)$$

Energy and momentum of particles are related through the relativistic momentum equation as $E^2 = p^2 c^2 + m^2 c^4$. The term in the expression for the pressure comes from the usual definition of the kinetic pressure of a gas $P = \frac{1}{3} n \langle pv \rangle$ and the relativistic momentum which gives $v = pc^2/E$.

If the particles are in a thermodynamic equilibrium state, they will follow either a Fermi-Dirac or a Bose-Einstein distribution function, depending on the species of the particles. These distributions are

$$f(p, t) = \frac{g}{(2\pi)^3} \left[\exp \left(\frac{E(p) - \mu}{kT(t)} \right) \pm 1 \right]^{-1}, \quad (1.27)$$

where the '+' version holds for fermions (spin-1/2) particles and the '-' version for bosons (integer spin), and μ is the chemical potential. The pre-factor g is a spin degeneracy factor and depends on the number of possible spin states for the particle type (e.g. electrons and photons have $g = 2$, quarks have $g = 6$, neutrinos have $g = 1$).

We will just compute a few limits of these equilibrium distributions. The particles are non-relativistic if the thermal energy is less than their rest mass energy $kT \ll mc^2$. In this case it is easy to carry out the integrals above and one finds

$$n(T) = g \left(\frac{mkT}{2\pi} \right)^{3/2} \exp \left(\frac{\mu - mc^2}{kT} \right), \quad \rho(t) = mn(t), \quad P(t) = nkT. \quad (1.28)$$

The other limit one can easily calculate is for the ultrarelativistic ($kT \gg mc^2$) non-degenerate ($\mu \ll kT$) case. In this limit one finds for fermions

$$n(t) = \frac{3}{4}(\zeta(3)\pi^2)gT^3, \quad \rho(t) = \frac{7}{8}(\pi^2/30)gT^4, \quad P = \rho/3 \quad (1.29)$$

and for bosons

$$n(t) = (\zeta(3)/\pi^2)gT^3, \quad \rho(t) = (\pi^2/30)gT^4, \quad P = \rho/3. \quad (1.30)$$

The evolution of the phase space distribution function follows the Boltzmann equation $Df/Dt = C[f]$, where C is the collision-term. In its absence, we immediately get the dilution of particles with cosmic expansion $n(t) \propto a^{-3}$.

$$\frac{1}{a^3} \frac{d}{dt}(a^3 n(t)) = 0. \quad (1.31)$$

If collisions (or reactions) are important, then the density of a species X evolves as

$$\frac{1}{a^3} \frac{d}{dt}(a^3 n(t)) = \frac{dn_X}{dt} + 3Hn_X = \int dp_X C[f_X]. \quad (1.32)$$

One can show¹, that for a reaction process $1 + 2 \leftrightarrow 3 + 4$, the right-hand side becomes

$$\frac{dn_1}{dt} + 3Hn_1 = n_1^{eq} n_2^{eq} \langle \sigma v \rangle \left[\frac{n_3 n_4}{n_3^{eq} n_4^{eq}} - \frac{n_1 n_2}{n_1^{eq} n_2^{eq}} \right], \quad (1.33)$$

where n_X^{eq} is the equilibrium density for species X and σ is the cross-section for the reaction.

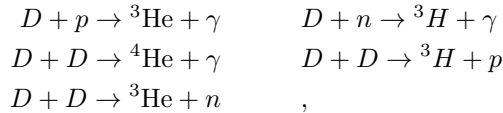
We abbreviate the reaction rate as $\Gamma = n \langle \sigma v \rangle$ and see that the left-hand side is on the order of $n_x H$, while the right-hand side of the order of $n_1 n_2 \langle \sigma v \rangle$. It immediately follows that if the reaction rate Γ is much larger than H , then the factor in brackets must approach zero. This gives us the equilibrium condition for the reaction

$$\frac{n_1 n_2}{n_3 n_4} = \frac{n_1^{eq} n_2^{eq}}{n_3^{eq} n_4^{eq}}, \quad (1.34)$$

or put the other way, in order to be in equilibrium, we need $\Gamma \gg H$. If the reaction rates drop due to the expansion of the universe, then the entire right-hand side will become unimportant and the number of particles will **freeze out** to a constant value. After freeze-out, the number density will simply scale as a^{-3} .

1.6 A primer on nucleosynthesis

The conditions for reactions in equilibrium with subsequent freeze-out of the abundances naturally produces different particles in different abundances (determined by their respective energies, and the cross-sections involved). A huge success of the big bang model was that it could explain the existence of light elements in the early universe in abundances consistent with observations. In the early Universe, energies are high enough for processes like deuterium formation $p + n \rightarrow D$ and the subsequent chains



¹See e.g. Section 3.1. in Dodelson "Modern Cosmology" [2]

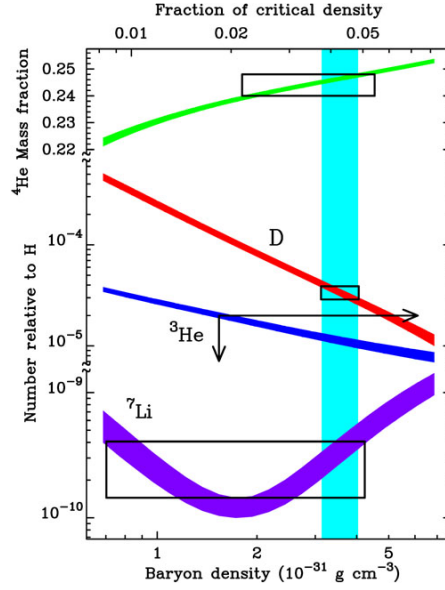
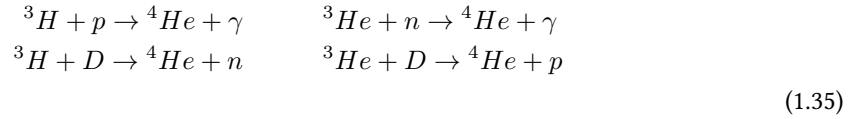
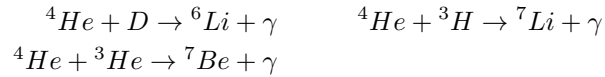


Figure 1.3: Abundance of light elements produced during Big Bang nucleosynthesis depends on the total density of baryons. The boxes indicate observational constraints on the abundance of these elements, the vertical stripe the region that one would get for the baryon density measured from the cosmic microwave background radiation (see also next chapter). [From [1]]

and the existence of Helium-3 allows the formation of Helium-4 by the processes



which will be followed by the formation of lithium and beryllium.



Heavier elements will not be formed since no stable elements exist with $A = 8$ and the time-scale is too short to circumvent this. All heavier elements thus must be produced at later times and this happens in stars (up to iron) and supernovae (also heavier than iron), as you might learn from other MAUCA courses.

1.7 Evidence for Dark Matter from nucleosynthesis

A more detailed calculation of the nucleosynthesis reactions show that the abundance of any of these light elements depends sensitively on the initial fraction of baryonic matter (i.e. initial abundance of protons and neutrons) as a fraction of the total matter in the Universe, see Figure 1.3. The baryonic density parameter is called Ω_b . The observed abundances of light elements are only consistent with the predicted abundances from big bang nucleosynthesis if $\Omega_b \sim \Omega_m/5$. This is very strong evidence that

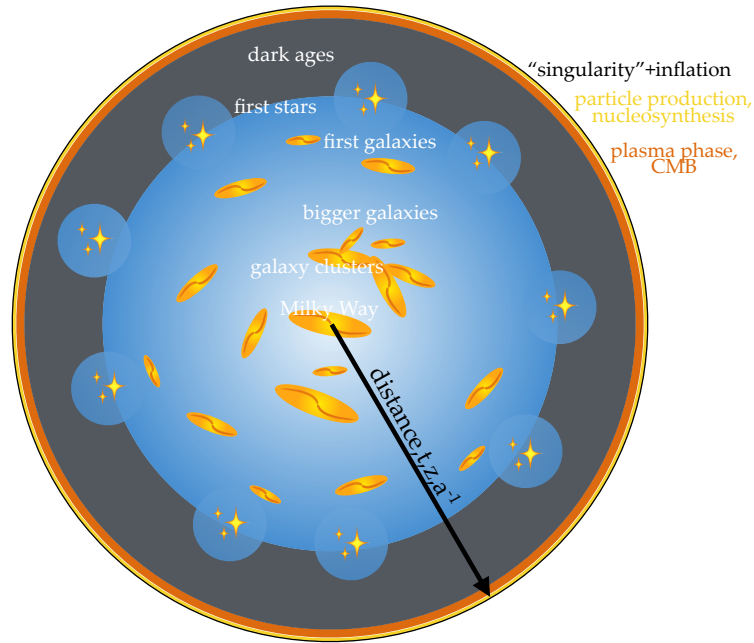


Figure 1.4: A cartoon picture of the full history of our Universe according to the Big Bang Theory. The distance from the centre can be expressed as either distance from the Milky Way, cosmological redshift, time or inverse of the cosmological expansion factor (although the relation between them is nonlinear). At a finite time, or distance, the formal 'singularity' occurred, followed by the phase of inflation (see next chapter), the phase of nucleosynthesis as well as the plasma phase in which nuclei, electrons and photons were in thermal equilibrium. At the end of it, the ions combine to neutral atoms, and the CMB radiation is able to freely propagate towards us. The release of the CMB is followed by the 'dark ages', in which no photons are produced and only terminates when the first stars, and then the first galaxies form and re-ionize the Universe. Over time, more massive galaxies form and most recently even clusters of hundreds and thousands of galaxies.

there must be other matter present in the universe that is not like ordinary baryons, but participates in gravity: we need enough matter to produce the correct expansion rate of the Universe, but not all of it can be in baryons in order to be consistent with the abundance of light elements. This is a first strong hint for the existence of **dark matter**, that we will encounter in many different ways over and over later. It will manifest itself always in this form: one needs more matter than just baryonic matter in the Universe for observations to be consistent with theory. The easiest way out is by assuming that there is a still undiscovered particle which does not participate in the strong and electromagnetic interactions and thus neither interacts with light (otherwise we could see it), nor plays a role in nucleosynthesis.

1.8 A picture of the Big Bang Universe

In Figure 1.4, we collect all that we know about the history of the Universe so far and put it all in perspective. All of the history can be drawn on a finite time-line due to the finite age of the Universe.

The caption explains what goes into this cartoon picture.

Chapter 2

The Inhomogeneous Universe

2.1 The CMB: Temperature Fluctuations

The tiny inhomogeneities in the Universe during the plasma phase become imprinted into the CMB radiation. Along different lines-of-sight, one thus expects minute temperature fluctuations. Temperature fluctuation means that the best-fit temperature changes from direction to direction. The temperature-shift is related to fluctuations in the gravitational potential from which the photons come. If they originate from an overdense region, they will lose a little bit of energy compared to the mean, and vice versa, if they originate from an underdense region, they will gain a little bit of energy. The temperature fluctuation is related to the potential fluctuation through the [Sachs-Wolfe effect](#) that also takes into account the time-dilation effect due to the change in gravitational potential and amounts to

$$\frac{\delta T}{T} = \frac{\delta \phi}{3c^2} \quad (2.1)$$

in a matter-dominated Universe where $\delta \phi$ is the fluctuation of the gravitational potential.

The CMB temperature fluctuations have characteristic amplitudes on different scales. This scale dependent amplitude of fluctuations is commonly analysed in terms of a [power spectrum of the fluctuations](#). To understand what this means consider first a Fourier cosine decomposition of a one-dimensional function $f(x)$. If the function is sufficiently regular and periodic on a domain of length D , we can always write

$$f(x) = \sum_{i=0}^{\infty} b_i \cos(k_i x + \phi_i), \quad (2.2)$$

where $k_i = 2\pi i/D$ is the wave number, ϕ_i is a phase, and b_i is the amplitude of mode k_i . The amplitudes b_i carry information about how much of mode k_i is in the signal, compared to other modes. Since the CMB arrives to us on a spherical surface, it decomposes its temperature fluctuations of course into two-dimensional functions, the spherical harmonics $Y_{\ell m}$, so that

$$T(\theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\theta, \phi), \quad (2.3)$$

where $Y_{\ell m} = \exp(im\phi)P_{\ell}^m(\cos \theta)$, where the P_{ℓ}^m are Legendre polynomials. One can then define an average at fixed ℓ over azimuthal values m to assess the amplitude contribution of fluctuations of longitudinal mode ℓ to the signal. We want to square the amplitudes in order to have identical contributions

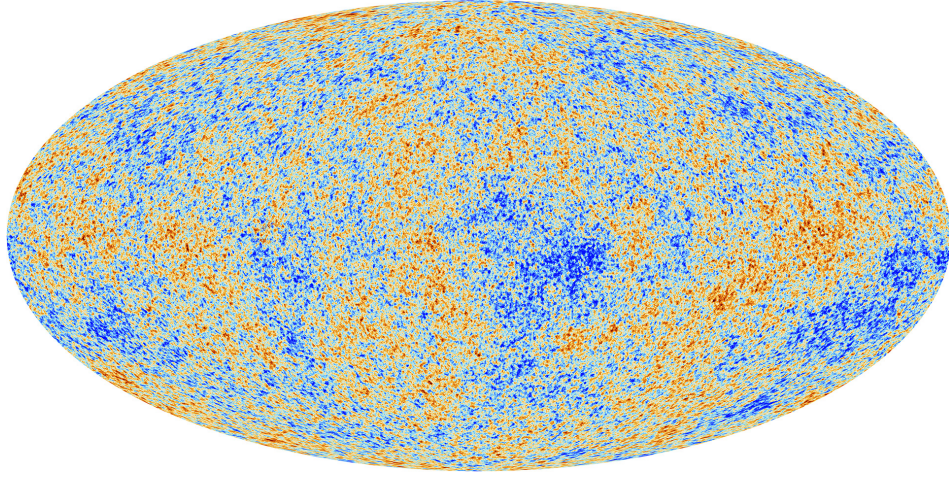


Figure 2.1: The CMB temperature fluctuations around the $T \sim 2.7\text{K}$ black body spectrum. This map has been obtained by the PLANCK satellite mission and the most accurate map to-date. Fluctuations have an amplitude $\delta T/T \sim 10^{-5}$. [Copyright 2013, ESA and the Planck Collaboration]

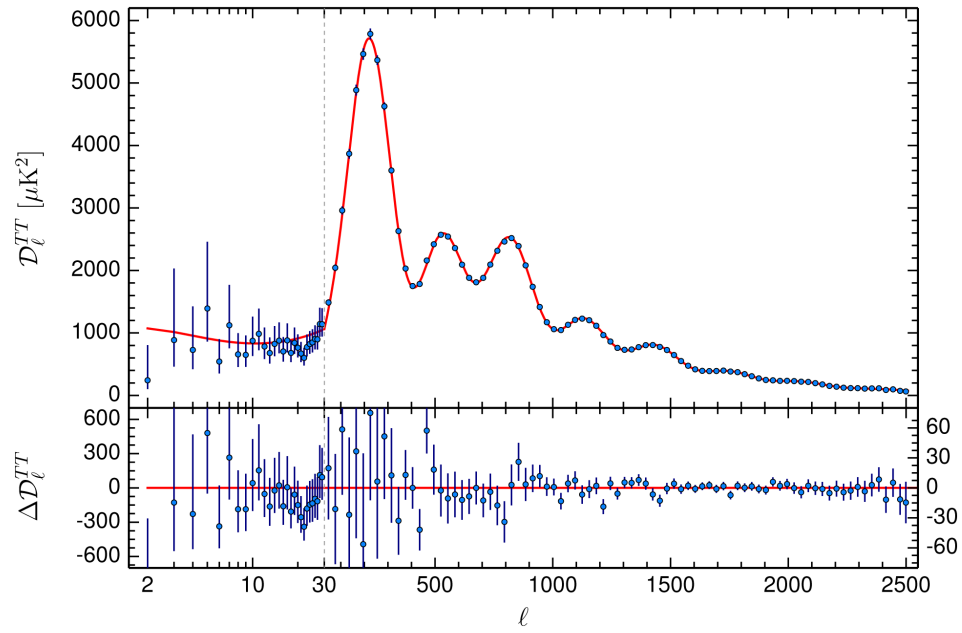


Figure 2.2: The power spectrum of CMB temperature fluctuations. Error bars show $\pm 1\sigma$ uncertainties and are on large scales (small ℓ) dominated by cosmic variance. The red line is our currently best-fit cosmological model. Note that there are no significant discrepancies for large enough ℓ where cosmic variance is unimportant. The bottom panel shows the difference between measurement and model. [Copyright 2015, by the Planck Collaboration]

from modes of amplitude $-A$ and $+A$. This gives the **angular power spectrum**

$$C_\ell := \langle |a_{\ell m}|^2 \rangle. \quad (2.4)$$

Note that for every ℓ there are $2\ell + 1$ values of m over which we average so that at small values of ℓ , one only has very few independent measurements (e.g. only 7 for $\ell = 3$). The measurement of C_ℓ for small ℓ will thus be limited by our inability to assess larger volumes, and will have large error bars due to what we call **cosmic variance**. In Figure 2.2 you see the power spectrum of the temperature fluctuation map from Figure 2.1.

This power spectrum contains a wealth of information about the physics of our Universe! The location of the first peak depends sensitively on the total density (and thus the geometry) of the Universe and thus sensitively constrains Ω_k . The ratio of the first to the second peak sensitively constrains the amount of baryons in the Universe as the sequence of peaks seen is a result of oscillations of the baryon fluid in the dark matter density perturbations. During the time of the primordial plasma, the baryons are in thermodynamic equilibrium with the photons meaning that the sound speed in the gas is essentially $c/3$ and the baryons are oscillating in and out of the dark matter density perturbations. This effect is called **baryon acoustic oscillation**, and we will revisit it again later. There is a wealth of other information in the CMB spectrum, some of which we will have time to touch upon, most of it is however beyond what we can do in this class.

2.2 Evolution of Small Perturbations – Eulerian Perturbation Theory

All structure in the Universe is thought to arise by gravitational collapse from the tiny fluctuations we see in the CMB. We will in this Section now study how density fluctuations grow in an expanding Universe.

2.2.1 The fluid equations in an expanding universe

Eulerian perturbation theory considers itself with solving equations for perturbations around the mean density of the Universe. This is done in a coordinate frame that is fixed in the Universe (either expanding or co-moving with the expansion) so that the fluid of dark matter and baryons has a non-zero velocity with which it moves. We first write down the standard equations for a fluid with density ρ and velocity \mathbf{u} in coordinates co-moving with the expansion of the universe. These are

$$\frac{\partial \rho}{\partial t} + a^{-1} \nabla \cdot (\rho \mathbf{u}) = 0 \quad (2.5)$$

$$\frac{\partial \mathbf{u}}{\partial t} + a^{-1} \mathbf{u} \cdot \nabla \mathbf{u} + H \mathbf{u} = -a^{-1} \nabla P - a^{-1} \nabla \phi \quad (2.6)$$

$$a^{-2} \nabla^2 \phi = 4\pi G (\rho - \bar{\rho}). \quad (2.7)$$

While they may look overwhelmingly complicated, they are actually quite simple. The first equation is just the continuity equation, rewritten for an expanding universe. The second is the Euler equation including the **Hubble drag** that slows down peculiar velocities ($H\mathbf{u}$), with a pressure term on the right hand side as well as the gravitational force. The third is simply Poisson's equation that encountered at the very beginning. But now we have subtracted the expansion of the universe out and are left with gravitational forces only if $\rho \neq \bar{\rho}$, i.e. the *inhomogeneous* case.

2.2.2 Looking at small perturbations

The next step to make this analytically tractable is to consider fluctuations only, i.e. let's write

$$\rho \equiv \bar{\rho}(1 + \delta), \quad \mathbf{u} = \bar{\mathbf{u}} + \mathbf{v}, \quad (2.8)$$

where we consider both δ and \mathbf{v} to be small. Note that in the co-moving frame, the gradients and time derivatives of the mean quantities will vanish! Also, we are free to choose $\bar{\mathbf{u}} = 0$ without loss of generality. If we insert them into the full set of equations, we find

$$\frac{\partial \delta}{\partial t} + a^{-1} \nabla \cdot ((1 + \delta) \mathbf{v}) = 0 \quad (2.9)$$

$$\frac{\partial \mathbf{v}}{\partial t} + a^{-1} \mathbf{v} \cdot \nabla \mathbf{v} + H \mathbf{v} = -a^{-1} \nabla P - a^{-1} \nabla \phi \quad (2.10)$$

$$a^{-2} \nabla^2 \phi = 4\pi G \bar{\rho} \delta \quad (2.11)$$

Let's now use $\bar{\rho} = \Omega_m a^{-3} \rho_c = a^{-3} \Omega_m 3H_0^2 / 8\pi G$ and let's only keep terms that are linear in the perturbation variables (if δ and \mathbf{v} are small, any of their products will be much smaller). We shall also make use of the fact that $\nabla P = \frac{\partial P}{\partial \rho} \nabla \rho = c_s^2$, where c_s is the speed of sound. Then we find

$$\frac{\partial \delta}{\partial t} + a^{-1} \nabla \cdot \mathbf{v} = 0 \quad (2.12)$$

$$\frac{\partial \mathbf{v}}{\partial t} + H \mathbf{v} = -\frac{c_s^2}{a} \nabla \delta - \frac{1}{a} \nabla \phi \quad (2.13)$$

$$\nabla^2 \phi = \frac{3}{2} H_0^2 \Omega_m a^{-1} \delta \quad (2.14)$$

Note that the advection term for the velocities has completely dropped out. We are left with a set of coupled *linear* partial differential equations. This does not look so much better than the set of coupled *nonlinear* partial differential equations that we started out with, but it is!

2.2.3 Linear partial differential equations become ordinary ones in Fourier space

The key is that the equations are now linear. *For a linear partial differential equations, the method of choice for their solution is Fourier transformation!* This is because Fourier transformation turns differential operators into multiplications.

To see this, let's consider the Fourier representation of a function $f(x)$

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \hat{f}(k) \exp[ikx] dk \quad (2.15)$$

$$\frac{\partial}{\partial x} f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \hat{f}(k) \frac{\partial}{\partial x} \exp[ikx] dk = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} ik \hat{f}(k) \exp[ikx] dk, \quad (2.16)$$

i.e. the pair $f(x) \leftrightarrow \hat{f}(k)$ becomes the pair $\partial_x f(x) \leftrightarrow ik \hat{f}(k)$. This means that if we Fourier transform our set of *linear* partial differential equations, it will become a set of linear ordinary differential

equations! Let's do that:

$$\frac{\partial \hat{\delta}}{\partial t} + a^{-1} i\mathbf{k} \cdot \hat{\mathbf{v}} = 0 \quad (2.17)$$

$$\frac{\partial \hat{\mathbf{v}}}{\partial t} + H\hat{\mathbf{v}} = -\frac{c_s^2}{a} i\mathbf{k} \hat{\delta} - \frac{1}{a} i\mathbf{k} \hat{\phi} \quad (2.18)$$

$$-k^2 \hat{\phi} = \frac{3}{2} H_0^2 \Omega_m a^{-1} \hat{\delta}. \quad (2.19)$$

We see that we can express everything in terms of the velocity divergence $\theta = \nabla \cdot \mathbf{v}$, $\hat{\theta} = i\mathbf{k} \cdot \hat{\mathbf{v}}$ by multiplying the second equation by $i\mathbf{k}$ to find

$$\dot{\hat{\delta}} = -a^{-1} \theta \quad (2.20)$$

$$\dot{\theta} + H\theta = \dot{\hat{\theta}} - Ha\dot{\delta} = \frac{c_s^2}{a} k^2 \delta - \frac{3}{2} H_0^2 \Omega_m a^{-2} \delta, \quad (2.21)$$

where we have omitted the hat for Fourier transforms now. Taking the second time derivative of the first equation yields

$$\ddot{\delta} = -a^{-1} \dot{\theta} + Ha^{-1} \theta = -a^{-1} \dot{\theta} - H\dot{\delta}. \quad (2.22)$$

2.2.4 Putting it all together

We see that we can combine our results into one single equation

$$\ddot{\delta} + 2H\dot{\delta} + \left(k^2 \frac{c_s^2}{a^2} - \frac{3H_0^2 \Omega_m}{2a^3} \right) \delta = 0 \quad (2.23)$$

Since this equation is an ordinary differential equation in time only, it will allow us to separate time and spatial parts as $\delta(\mathbf{x}, t) = D(t)\delta(\mathbf{x}, t_0)$, which is a feature of such linear partial differential equations. We call the function $D(t)$ the **(linear) growth factor**. Note that this equation embodies the **Jeans stability criterion**! Depending on the sign of the term in brackets in the last equation, the solution will be either growing or oscillating in time. We see that there is a critical wave number k_J which separates these two regimes given by

$$k_J = a^{-1} \frac{3H_0^2 \Omega_m}{2c_s^2}. \quad (2.24)$$

In the case of dark matter, $c_s = 0$, so that there are no oscillating solutions. One can show that in that case (and assuming one can neglect the radiation component Ω_r), the growth factor takes the form:

$$D(a) \propto \frac{H(a)}{H_0} \int_0^a [\Omega_m a'^{-1} + \Omega_\Lambda a'^2 + (1 - \Omega_m - \Omega_\Lambda)]^{-3/2} da', \quad (2.25)$$

and one usually chooses $D(t_0) = 1$ at the present time. In the presence of radiation, the full differential equation for D has to be solved.

Exercise 6 Integrate numerically the equation for the growth factor, eq. (2.25), and plot the solutions for various choices of Ω_m and Ω_Λ . Show that $D(a) = a$ for $\Omega_m = 1, \Omega_\Lambda = 0$.

Exercise 7 Re-derive similarly to what we did above, from the full nonlinear fluid equations, the linearised equations for the evolution of a fluid of dark matter and baryons, coupled through gravity. The linearised equations are

$$\frac{\partial \delta_c}{\partial t} = -a^{-1} \nabla \cdot \mathbf{v}_c \quad (2.26)$$

$$\frac{\partial \delta_b}{\partial t} = -a^{-1} \nabla \cdot \mathbf{v}_b \quad (2.27)$$

$$\frac{\partial \mathbf{v}_c}{\partial t} = -a^{-1} \nabla \phi - H \mathbf{v}_c \quad (2.28)$$

$$\frac{\partial \mathbf{v}_b}{\partial t} = -a^{-1} \nabla \phi - H \mathbf{v}_b - \frac{c_s^2}{a} \nabla \delta_b \quad (2.29)$$

$$a^{-2} \nabla^2 \phi = 4\pi G \bar{\rho} (f_b \delta_b + f_c \delta_c), \quad (2.30)$$

where $f_b = \Omega_b/\Omega_m$ and $f_c = 1 - f_b$ are the baryon fraction and cold dark matter fraction. Write a program in PYTHON that shows the transfer function $T_c(t) = \delta_c(t)/\delta_c(0)$, $T_b(t) = \delta_b(t)/\delta_b(0)$ for various times. Choose appropriate initial conditions yourself. If you want to get fully realistic, the sound speed in the gas is

$$c_s = \sqrt{\frac{\gamma k T_b}{\mu m_H}}, \quad (2.31)$$

where $\gamma = 5/3$ for an ideal monoatomic gas, $\mu = 1.22$ is the mean molecular weight of the gas assuming a Helium fraction of 24 per cent, m_H is the mass of the hydrogen atom, and T_b is the kinetic temperature of the baryons. The latter can be approximated from more involved calculations as roughly [8]

$$T_b(a) = \frac{2.726 \text{ K}}{a} \left[1 + \frac{119a}{1 + (115a)^{-3/2}} \right]^{-1}. \quad (2.32)$$

Based on numerical experiments, investigate (1) what is the effect of a higher baryon fraction on $\delta_c(k)$ and $\delta_b(k)$, (2) what is the effect of the sound speed after recombination, (3) what changes if you change the redshift of recombination, (4) describe how the *baryon bias* $\delta_b(k)/\delta_c(k)$ evolves over time before and after recombination, and (5) can you identify features seen in the CMB spectrum from Figure 2.2 in the baryon spectrum $P_b = \delta_b^2(k)$?

Note that while this exercise shows the emergence of the *baryon acoustic oscillation* (BAO) in the gravitationally coupled fluid of dark matter and baryons, many important effects are not included, most notably the relativistic effect of different growth rates of perturbations that are larger/smaller than the horizon.

Hints for this Exercise:

The skeleton of a PYTHON program that numerically integrates the coupled ODEs is given in Appendix A.2. You just have to fill in the actual equations and get it to run.

Exercise 8 Obviously we assumed that $\delta \ll 1$ when linearising the equations. What do you think will be the dominant error when δ becomes larger? Is there a limit for δ when the solution becomes unphysical?

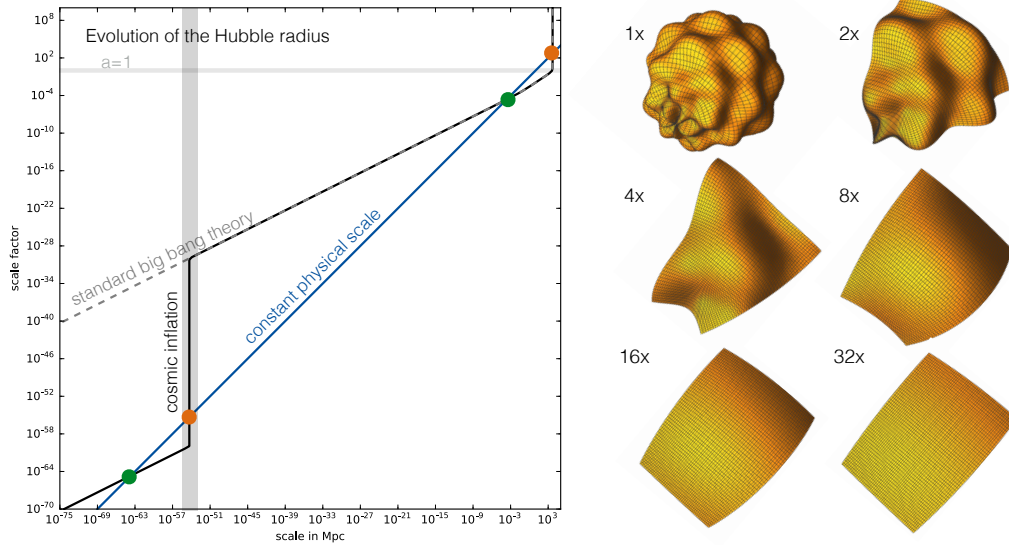


Figure 2.3: *Left*: The evolution of the Hubble radius in a standard big bang universe (dashed gray line) and in an inflationary universe (solid black line). During the inflationary phase, the Hubble radius remains constant, while the scale factor increases dramatically. A constant physical scale will only enter the horizon once in a standard universe, meaning that this scale can not have been in causal contact before then. In an inflationary universe, the scale was inside the horizon before the inflationary phase, then leaves it, before it re-enters it again at a later time. Note the identical behaviour of the inflationary phase and the dark energy phase. *Right*: A strongly curved region will appear more and more flat as it is expanded. Note that during inflation the scale factor grows by at least 10^{30} .

2.3 Origin of Inhomogeneities and Inflation

We have encountered various conceptual puzzles or even paradoxes on our journey through the paradigm of big bang cosmology as you may have noticed. Let's summarise them briefly: The big bang paradigm has no satisfying answer in itself to the following puzzling observations

1. The **horizon problem**: Why are regions that can have never been in causal contact (e.g. opposite directions on the sky) basically identical?
2. The **flatness problem**: Why is the curvature of the Universe so close to zero (i.e. $\Omega_k \simeq 0$)? There is no a priori reason why this needs to be so, and if there is some curvature it will even grow over time...
3. The **origin of inhomogeneity**: Where do the fluctuations that we observe in the CMB come from? What determines their spectrum?

In the early 1980s Alan Guth, Andrei Linde and Paul Steinhardt had the ingenious idea that an epoch of very rapid expansion (in fact exponentially fast, just like what we have seen for the cosmological constant) is able to solve all three problems. How does this work?

While we will not go into the details, inflationary theory assumes that there existed a phase in the very early Universe the vacuum energy density was much larger than today. The energy density was

dominated by a so called **inflaton field**, which behaves very similar to a cosmological constant Λ , just with a much larger value than the one that dominates the universe today. We had already seen that this means that $H(a) = \sqrt{\Lambda/3}$ and this implies that

$$a(t) \propto \exp\left(\sqrt{\frac{\Lambda}{3}}t\right). \quad (2.33)$$

After a short period of time, this vacuum decays in a phase transition into normal matter and radiation and the exponentially fast expansion is thus halted.

2.3.1 A solution to the horizon problem

Such a phase in the early universe has however profound consequences. To understand this, let us look at the so-called **Hubble radius** R_H , which is the distance at which the expansion velocity equals the speed of light, which means that

$$R_H(a) = \frac{c}{H(a)}. \quad (2.34)$$

As we have learned before, two points separated by a distance larger than the particle horizon cannot have been in causal contact. Since the particle horizon is difficult to integrate, one usually simply uses the Hubble radius to estimate the scale of causality. Since during both the inflationary and the dark energy dominated phase $H \simeq \text{const}$, also the Hubble radius is constant, while at the same time, the universe expands greatly. The evolution of the Hubble radius as a function of the scale factor of the universe is shown in the left panel of Figure 2.3. For the standard big bang cosmology (dashed gray line), the Hubble radius increases slowly as the universe expands. Also shown in the graph is a blue line that indicates a fixed physical length scale that is expanding along with the universe (chosen to be 100 Mpc in the graph). We see that this line intersects the standard big bang Hubble radius first in the right green point. This indicates the time when points separated by this distance first came into causal contact and it thus highlights that in the standard big bang model, they could never exchange information before that time: this is exactly the horizon problem. Now look at the curve for the inflationary universe. Here, during the inflationary phase, the universe grows by about 10^{30} in size, while the Hubble radius remains constant during that phase. Now the blue line indicating a constant physical scale intersects two more times. The first, left, green point shows that the scale entered the Horizon well before the inflationary phase, allowing everything on this scale to be in causal contact. During the inflationary phase, the scale gets however transported outside the horizon until it re-enters the horizon at the same time as in the standard big bang model. This solves our problem: all observables scales have been in causal contact before inflation allowing them to be in some kind of thermal equilibrium and thus becoming thermodynamically identical. Then inflation stretches them to very large scales, so that we see them at separations that are no longer in causal contact. Note that this seems like magic, just as if there is something moving faster than light, but it isn't! It's just the universe expanding, there is never any exchange between different locations that would violate the principle of relativity. So the uniformity of the CMB is a natural consequence of cosmic inflation.

2.3.2 ...and the flatness problem

However, inflation does even more. It also solves the flatness problem! If we assume there is some initial curvature along with the inflaton field, then

$$H(a) \propto \sqrt{\Lambda/3 + (1 - \Lambda/3)a^{-2}}, \quad (2.35)$$

Since during the inflationary phase a grows exponentially fast in time, any curvature will become completely negligible! Since the curvature will become so tiny (a changes by 10^{30} !), it will remain very small until today, thus explaining why we observe a flat Universe. The effect of expansion on curvature is rather intuitive: look at the right panel of Figure 2.3. We start with a heavily curved object, then we stretch it by a factor of and focus on a region of the same area as before: it is still clearly curved. However, let's repeat this process a few more times. Once we have stretched it by a factor of 32, we barely see any curvature any more, it has become virtually flat! This was just a factor of 32, inflation would stretch everything by a factor of $10^{30} - 10^{40}$!

2.3.3 ...and how inflation generates perturbations

The inflation of scales by a factor $\sim 10^{30}$ has one other important consequence. If there are any random quantum fluctuations in the inflation field, they will also get stretched to cosmological scales. Inflationary theory – and in fact there are many different flavours – predicts that after the inflationary phase, these fluctuations will follow a very simple power spectrum

$$P(k) = |\delta|^2 \propto k^{n_s}, \quad \text{where } n_s \simeq 1, \quad (2.36)$$

and n_s is called the **spectral index**. In fact, the simplest models of inflation already predict that n_s is slightly smaller than 1, which is exactly what is observed. Recent measurements of this indicate that $n_s \simeq 0.96$. Inflation also predicts that δ will follow a distribution that must be close to Gaussian (due to the origin from quantum fluctuations), which again agrees with observations. The temperature fluctuations seen in the CMB show no indication that the underlying distribution of fluctuations in the gravitational potential that produces them deviates from a Gaussian distribution.

To summarise, inflation solves all the big conceptual problems of the standard big bang cosmology by introducing a phase of exponentially fast expansion in the very early universe. In this scenario, all matter and radiation that we observe today would come from a phase transition in which the inflaton field decays into a new vacuum state at the end of inflation. It naturally explains why the CMB is so uniform, why the Universe is flat and gives an explanation for the presence and character of inhomogeneities from which all cosmic structure forms through gravitational instability.

2.4 Properties of Inhomogeneities – Gaussian random fields

As we discussed above, the density perturbations generated during inflation from quantum fluctuations in the inflaton field have a Gaussian distribution in space. We have seen that they remain small until well after the plasma phase where we can describe them by linear perturbation theory. Since linear perturbation theory does not couple scales, the perturbations will remain Gaussian until they have grown under gravity large enough that the linear theory breaks down. A field whose values follow a Gaussian distribution is called a **Gaussian random field**. Let our density fluctuations $\delta(\mathbf{x})$ be such a field. Since a Gaussian distribution is fully described by its mean $\mu = \langle \delta \rangle$ and its variance $\sigma^2 = \langle \delta^2 \rangle - \langle \delta \rangle^2$ also a Gaussian random field will be described by spatial equivalents to these two moments.

The first just means that the following integral has well defined meaning

$$\langle \delta \rangle = \frac{1}{V} \int_V \delta(\mathbf{x}) d^3x. \quad (2.37)$$

If δ is the overdensity field, then its mean must vanish, i.e. $\langle \delta \rangle = 0$, for large enough volumes V . In this case, the underlying field is called statistically **homogeneous**.

The variance generalises to the [two-point correlation function](#), which, for a field with vanishing mean, is defined as

$$\xi(\mathbf{x}_1, \mathbf{x}_2) = \langle \delta(\mathbf{x}_1) \delta(\mathbf{x}_2) \rangle. \quad (2.38)$$

If the field is statistically homogeneous (i.e. the mean vanishes in large enough volumes) one can write this as

$$\xi(\mathbf{x}) = \langle \delta(\mathbf{x}') \delta(\mathbf{x}' + \mathbf{x}) \rangle = \frac{1}{V} \int \delta(\mathbf{x}') \delta(\mathbf{x}' + \mathbf{x}) d^3x', \quad (2.39)$$

which is simply a convolution of the field with itself. If it is also [statistically isotropic](#), which means that there are no preferred directions in it, then the two-point correlation function will only be a function of the separation of the two points $r = \|\mathbf{x}_1 - \mathbf{x}_2\| = \|\mathbf{x}\|$. In this case, the function $\xi(r)$ has a very simple meaning: it describes the excess probability, with respect to random, of finding correlated values of δ at separations r .

Since eq. (2.39) is a convolution, we can turn it into a simple product in Fourier space. This gives us the [power spectrum](#) $P(\mathbf{k})$ of the random field as

$$\langle \hat{\delta}(\mathbf{k}) \hat{\delta}^*(\mathbf{k}') \rangle = (2\pi)^3 \delta_D(\mathbf{k} - \mathbf{k}') P(\mathbf{k}). \quad (2.40)$$

The right-hand-side is only nonzero when $\mathbf{k} = \mathbf{k}'$, and since we shall assume that $\delta(\mathbf{x})$ is a real valued field, then $\hat{\delta}^*(\mathbf{k}) = \hat{\delta}(-\mathbf{k})$ which implies that

$$P(\mathbf{k}) = \frac{1}{(2\pi)^3} |\hat{\delta}(\mathbf{k})|^2. \quad (2.41)$$

If the underlying field is isotropic, then there is no preferred direction for the \mathbf{k} -vector and the field will be fully described by $P(k) = (2\pi)^{-3} |\delta(k)|^2$.

This means that a statistically homogeneous and isotropic Gaussian random field is fully described by its power spectrum $P(k)$ or two-point correlation function $\xi(r)$ which depend on the norm of \mathbf{k} and \mathbf{r} only. Note also that the requirement of vanishing mean simply means that $P(k=0) = 0$.

A short calculation shows that for homogeneous isotropic fields also the following relations hold between power spectrum and two-point correlation function:

$$\xi(r) = \frac{1}{2\pi^2} \int_0^\infty k^2 P(k) \frac{\sin(kr)}{kr} dk \quad \text{and} \quad P(k) = 4\pi \int_0^\infty r^2 \xi(r) \frac{\sin(kr)}{kr} dr. \quad (2.42)$$

Finally, in the limit $r \rightarrow 0$, the two-point correlation function reduces to the variance of the field

$$\sigma^2 = \frac{1}{2\pi^2} \int_0^\infty k^2 P(k) dk. \quad (2.43)$$

While this is the point-wise variance, it is also helpful to define the variance of the field on arbitrary scale R . To this end, one simply filters the underlying field with a low-pass filter W : $\delta_R(\mathbf{x}) * W(\mathbf{x})$, e.g. a Gaussian, or more commonly, a top-hat filter $W_{TH}(r, R) = \frac{3}{4\pi R^3} \Theta(R - r)$, where $\Theta(x)$ is the Heaviside step function. The convolution simply becomes a multiplication in Fourier space, so that

$$\sigma_R^2 = \frac{1}{2\pi^2} \int_0^\infty k^2 \hat{W}^2(kR) P(k) dk. \quad (2.44)$$

For the top hat filter $\hat{W}_{TH}(kR) = 3(\sin(kR) - kR \cos(kR)) / (kR)^3$, while for a Gaussian filter $\hat{W}_G(kR) = \exp(-(kR)^2/2)$.

To summarise this section: due to the Gaussian nature of density fluctuations in the young universe, and because of the cosmological principle of homogeneity and isotropy (that was the assumption to derive the Friedmann equation), small density fluctuations *are fully characterised by their power spectrum, or, equivalently, their correlation function.*

2.5 How to generate a realisation of the early Universe - Cosmological Initial Conditions

When we put everything together, the field of density fluctuations in the young universe can be written in a very simple form in Fourier space

$$\hat{\delta}(\mathbf{k}) = A k^{n_s/2} T(k) \hat{G}(\mathbf{k}), \quad (2.45)$$

where A is a constant that sets the overall normalisation, $k^{n_s/2}$ comprises the spectrum of fluctuations generated by inflation (with power spectrum $\propto k^{n_s}$). The function $T(k)$ is the result of our linear perturbation theory and is called the **transfer function**, and $\hat{G}(\mathbf{k})$ is just the Fourier transform of a Gaussian random field of zero mean and unit variance. The field $\hat{G}(\mathbf{k})$ can be simply generated by drawing a random number from a Gaussian distribution of zero mean and unit variance for each point in space and Fourier transforming the result¹. In order to find $\delta(\mathbf{x})$ one just multiplies everything in Fourier space and transforms the result back to real space.

There is one aspect that we haven't included yet when we discussed linear perturbation theory before. We only performed a Newtonian calculation of the growth of perturbations, but in fact there are important relativistic effects that need to be taken into account. They cause perturbations on scales that are larger than the horizon at a given cosmic time to grow at a different rate than those that are smaller than the horizon. A detailed discussion of this is beyond the scope of this course. A convenient fitting formula that takes into account the relativistic effects, but not the effect of the baryon acoustic oscillations has been obtained by Bardeen, Bond, Kaiser & Szalay (BBKS) and reads

$$T_{\text{BBKS}}(k) = \frac{\log(1 + 2.34q)}{2.34q} (1 + 3.89q + (16.1q)^2 + (5.46q)^3 + (6.71q)^4)^{-1/4}, \quad (2.46)$$

where $q = k \Omega_m^{-1} \text{Mpc}^{-1}$. You will obtain an almost realistic transfer function for CDM and baryons if you multiply your result from the linear theory calculation of the coupled fluid with this function.

Finally, the amplitude, that we just called A above is typically fixed by requiring that the top-hat filtered variance σ_8^2 on a scale of $8 h^{-1} \text{Mpc}$, given by eq. (2.44), equals a given measured value. Recent measurements of cosmological parameters indicate that $\sigma_8 \simeq 0.81$.

Exercise 9 Compute the value of the top-hat filtered variance $\sigma^2(R)$ as a function of filter radius R for the BBKS power spectrum. Plot $\sigma(R)$ and normalise the spectrum correctly to a specifiable value of σ_8 .

Hint: The upper limit of the integral in eq. (2.44) being infinity, simply choose a large value for it (depending on R of course) and make sure that your result does not depend on it.

¹In fact, the Fourier transform of a real Gaussian random field is also a Gaussian random field in Fourier space. It is however complex, with both real part and imaginary part being Gaussian random fields and the additional symmetry of real fields of $\hat{\delta}^*(-\mathbf{k}) = \hat{\delta}(\mathbf{k})$, which follows directly from the definition of the Fourier transform.

Exercise 10 Create a three-dimensional realisation of the young Universe! Start by filling a three-dimensional array of $N \times N \times N$ cells (use $N = 32$ or 64) with Gaussian random numbers of zero mean and unit variance. Fourier transform the array and multiply the result with $k^{n_s/2} T_{\text{BBKS}}(k)$. You will have to choose a size of the volume that this represents since the wave numbers that you will be able to represent using such an array are $k_0, 2k_0, 3k_0 \dots \frac{N}{2}k_0$ in each linear dimension with $k_0 = \frac{2\pi}{L}$. Transform the result back to real space. Visualize a slice of size $N \times N$ from the array. Make a histogram of all values of $\delta(\mathbf{x})$ – is it still Gaussian?

Calculate the power spectrum of the actual realisation. You will have to average values $|\hat{\delta}(\mathbf{k})|^2$ over shells in modulus of k , i.e. between k and $k + dk$ for some suitable k . Compare your result to the spectrum that you wanted to sample. What do you observe for small values of k and for large ones?

Hint: You find more tips and instructions how to do this in the Appendix A.3.

2.6 Lagrangian Perturbation Theory and Zel'dovich's approximation

The equations for evolution of perturbations that we have derived in Section 2.2 were at fixed spatial coordinates, i.e. we chose a coordinate frame which is at rest in comoving coordinates. It is however always possible to use a coordinate frame in which we are freely moving along with the fluid. Such a frame that is comoving with the fluid is called a **Lagrangian frame**. It can be easily shown that the **convective derivative** in the Eulerian description becomes a simple partial derivative

$$\frac{D}{Dt}X := \frac{\partial X}{\partial t} + \mathbf{v} \cdot \nabla X \quad [\text{Eulerian}] \quad \rightarrow \quad \frac{\partial X}{\partial t} \quad [\text{Lagrangian}]. \quad (2.47)$$

It is clear that this will significantly simplify the full nonlinear Euler equation (2.6), particularly if we also *neglect pressure forces*, to become

$$\frac{\partial \mathbf{v}}{\partial t} + H\mathbf{v} = -a^{-1} \nabla \phi. \quad (2.48)$$

where of course still Poisson's equation holds for the gravitational potential

$$\nabla^2 \phi = \frac{3}{2} H_0^2 \Omega_m a^{-1} \delta. \quad (2.49)$$

Since the Lagrangian evolution equation for \mathbf{v} only depends on the gravitational potential, the idea of **Lagrangian perturbation theory (LPT)** is to derive a perturbative expansion for it. For us it will suffice to only go to first order which will give us 1LPT, or also called **the Zeldovich approximation**. Since we know that at first order in Eulerian perturbation theory $\delta(\mathbf{x}, t) = D(t) \delta(\mathbf{x}, t = t_0)$, we can simply get from Poisson's equation that then $\phi(\mathbf{x}, t) = (D(t)/a(t)) \phi_i(\mathbf{x})$ – with the initial potential $\phi_i(\mathbf{x}) = \phi(\mathbf{x}, t = t_0)$. We see immediately that in an Einstein-de Sitter cosmology ($\Omega_m = 1$ and hence $D = a$) the potential is constant. Even in the general case, we can rewrite eq. (2.48) as

$$\frac{\partial \mathbf{v}}{\partial t} + H\mathbf{v} = \frac{1}{a} \frac{\partial(a\mathbf{v})}{\partial t} = -\frac{D}{a^2} \nabla \phi_i \quad (2.50)$$

$$\Rightarrow \quad \mathbf{v} = -\frac{\nabla \phi_i}{a} \int \frac{D}{a} dt. \quad (2.51)$$

In order to simplify this integral, we next recall that D is a solution of the second order differential equation, eq. (2.23), $\ddot{D} + 2H\dot{D} = (3/2)H_0^2\Omega_m a^{-3}D$. We can rewrite this using $a^2\ddot{D} + 2a^2H\dot{D} = \frac{\partial}{\partial t}(a^2\dot{D})$ to see that simply

$$\int \frac{D}{a} dt = \frac{a^2\dot{D}}{\frac{3}{2}H_0^2\Omega_m}. \quad (2.52)$$

We thus find finally

$$\mathbf{v} = -\frac{a\dot{D}}{\frac{3}{2}H_0^2\Omega_m} \nabla\phi_i. \quad (2.53)$$

Since the position of the fluid parcel moves as $\mathbf{v} = a\dot{\mathbf{x}}$, we immediately see that the position must be given by

$$\mathbf{x} = \mathbf{q} - \frac{D}{\frac{3}{2}H_0^2\Omega_m} \nabla\phi_i \quad (2.54)$$

relative to an initial position \mathbf{q} for $D \rightarrow 0$ so that $\phi_i = \phi(\mathbf{q})$. This initial position \mathbf{q} is also called the [Lagrangian coordinate](#) of the fluid element and one can think of it simply as a label that one attaches to the fluid element and that stays with the fluid element for all times as it moves through the space of [Eulerian coordinates](#) \mathbf{x} . If we define a new velocity potential $\psi := -\left(\frac{3}{2}H_0^2\Omega_m\right)^{-1}\phi_i$, then these equations take the particularly simple form

$$\mathbf{x} = \mathbf{q} + D\nabla\psi \quad (2.55)$$

$$\mathbf{v} = a\dot{D}\nabla\psi \quad (= a\dot{\mathbf{x}}). \quad (2.56)$$

A final interesting aspect of this analysis is that the formula for the density at the location of a fluid element can be derived by assuming mass conservation when going from Lagrangian to Eulerian frame. Then, for any arbitrary volume V

$$\int_V \bar{\rho} d^3q = \int_V \bar{\rho}(1 + \delta(\mathbf{x})) d^3x \quad (2.57)$$

$$= \int_V \bar{\rho}(1 + \delta) \left| \frac{\partial \mathbf{x}}{\partial \mathbf{q}} \right| d^3q, \quad (2.58)$$

where we have just used the Jacobian determinant when transforming integration variables in the second equality. Since the volume V is arbitrary, the integrands have to be identical, i.e.

$$\Rightarrow 1 + \delta(\mathbf{q}) = \left| \frac{\partial \mathbf{x}}{\partial \mathbf{q}} \right|^{-1} = \left| \delta_{ij} + D \frac{\partial^2 \psi}{\partial q_i \partial q_j} \right|^{-1} \quad (2.59)$$

$$= \frac{1}{(1 + D\lambda_1)(1 + D\lambda_2)(1 + D\lambda_3)}, \quad (2.60)$$

where λ_i are the *real* eigenvalues of the second derivative tensor $\partial_{ij}\psi$ (the eigenvalues are real because the tensor is symmetric). Note that this is the overdensity at the Lagrangian coordinates. In order to obtain the Eulerian density, one would need to invert the coordinates, which is not always possible.

We see that the overdensity δ will become infinite if $D\lambda = -1$ for one of the λ . Since D is the growth factor, and thus monotonously increasing in time, this will always happen at some point in time if the initial $\lambda < 0$. If $\lambda > 0$, it will never happen of course. This singularity of δ is called [shell crossing](#): it occurs when several fluid elements starting from different initial \mathbf{q} (i.e. with different Lagrangian coordinates) end up at the same \mathbf{x} (i.e. at the same Eulerian coordinate). Since there are three dimensions, there are four different combinations of positive and negative eigenvalues in each point, and they correspond to different features of the cosmic web (see Figure 2.4).

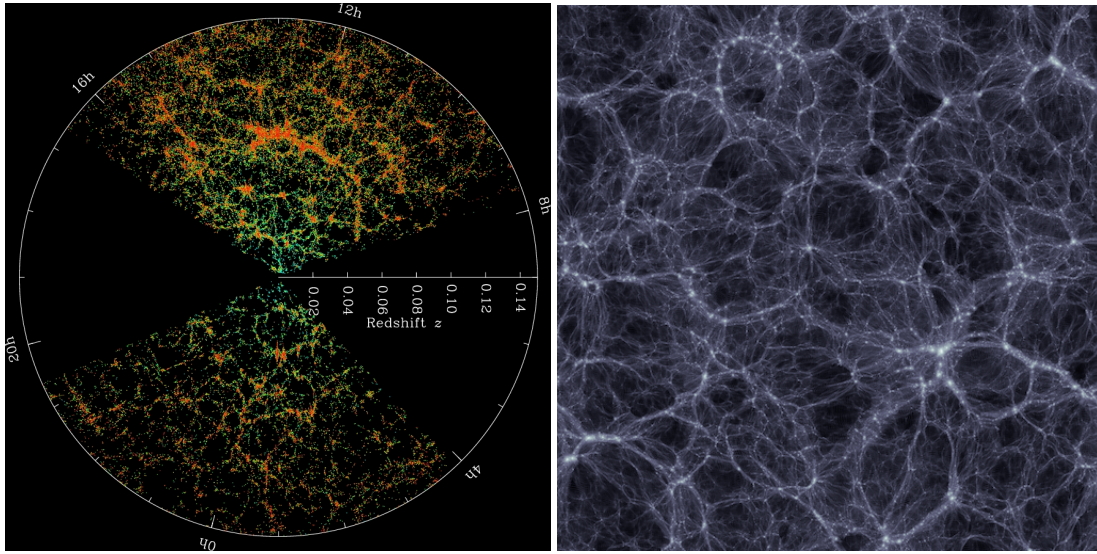


Figure 2.4: The cosmic web as it emerges in redshift space in galaxy redshift surveys such as the SDSS survey (left panel) and in cosmological simulations (right panel). The different structures, such as nodes, filaments and voids are directly related to the eigenvalues of the tensor $\partial^2\psi/\partial q_i\partial q_j$ computed on the initial Gaussian random field. [left panel: Copyright by the SDSS Collaboration]

Exercise 11 Calculate the particle velocities \mathbf{u} in the Zel'dovich approximation on a grid for the power spectrum realisation that you have already calculated in Exercise 10. Note that assuming you have $\delta(\mathbf{x})$ on a grid, it is trivial to compute $\hat{\mathbf{v}}$ in Fourier space. Visualise a slice of the velocity field for one component, e.g. v_x .

Hint: Remember that under Fourier transform $\nabla \rightarrow i\mathbf{k}$, so that $\frac{\nabla}{\nabla^2} \rightarrow -i\frac{\mathbf{k}}{k^2}$.

Exercise 12 Create an initial distribution of points located at the centres of the cells where you calculated δ and \mathbf{u} from the realisation of the random field in Exercise 11. Now use the Zel'dovich approximation to move these points. Plot the points in a thin slice and describe what you see as time progresses.

Hint: See Appendix A.4 for help how to plot slices of points as well as 3D projections of points.

Exercise 13 Show mathematically that shell crossing occurs where the flow converges and that it never occurs where it diverges. Show that there are four possible combinations of converging and diverging flow along different directions in three dimensions. Sketch the flow patterns for each of these four possible cases. Can you relate this analysis to what you saw visually in Exercise 12. This analysis explains the emergence of the cosmic web.

Chapter 3

The Nonlinear Universe

3.1 Modelling Dark Matter

For evidence for the existence of dark matter, its properties and efforts to detect it see slides... (Will be summarised here in text form later)

Dark matter in its simplest form has the following macroscopically defining properties: it is

1. **nonbaryonic** : it does not emit or absorb light, nor does it interact with ordinary matter other than through gravity,
2. **collisionless** : it does not appreciably interact with itself except through gravity,
3. **cold** : it has a negligible temperature, which means that it has negligible intrinsic velocity dispersion.

These properties make dark matter different from ordinary matter. The fact that it does not interact with itself means also that there will be no collisions between dark matter particles that would drive it towards a thermodynamic equilibrium state (as e.g. a Maxwell Boltzmann distribution with isotropic velocity dispersion). In most general form, it would thus follow on macroscopic scales the **collisionless Boltzmann equation** which is just a conservation equation for the **phase space density** $f(\mathbf{x}, \mathbf{v}, t)$:

$$\frac{\partial f}{\partial t} + \frac{\tilde{\mathbf{v}}}{a^2} \cdot \nabla_x f - \nabla_x \phi \cdot \nabla_{\tilde{\mathbf{v}}} f = 0, \quad (3.1)$$

where $\tilde{\mathbf{v}} = a\mathbf{v}$ and \mathbf{v} is our usual comoving velocity¹, and ϕ follows Poisson's equation

$$\nabla_x^2 \phi = 4\pi G \bar{\rho} \delta = \frac{3}{2} H_0^2 \Omega_m \delta / a, \quad (3.2)$$

where we have used that $\bar{\rho} = \rho_{\text{crit}} \Omega_m a^{-3}$.

Since f gives the density of dark matter particles in phase space, the density in configuration (or normal) space is given by marginalising over the velocity dimensions

$$n(\mathbf{x}) = \int f(\mathbf{x}, \mathbf{v}, t) d^3v =: \rho(\mathbf{x})/m_p, \quad (3.3)$$

¹Note that this is a general feature that we have encountered before. Conservation laws in comoving quantities include a Hubble drag, e.g. $\partial_t v + H v = \dots$ which can be eliminated by using $\tilde{v} = av$ so that $a^{-1} \partial_t \tilde{v} = \partial_t v + H v$. This is usually called a super-comoving coordinate.

where m_p is the (microscopic) particle mass. The system of collisionless Boltzmann and Poisson's equation is usually called the **Vlasov-Poisson** system of equation. Basically, it simply states that the phase space density is conserved. It is advected in configuration space by velocities, and in velocity space by gravitational forces. In particular, the equation implies that the phase space density is conserved along so-called **characteristics**, which are the trajectories

$$\frac{d\mathbf{x}_i}{dt} = a^{-2}\tilde{\mathbf{v}}, \quad \frac{d\tilde{\mathbf{v}}_i}{dt} = -\nabla\phi. \quad (3.4)$$

This can be easily seen by writing $f(\mathbf{x}_i(t), \tilde{\mathbf{v}}_i(t), t)$ and inserting it into the collisionless Boltzmann equation. The equations 3.4 are however simply describing the motion of points in the gravitational field given by the potential ϕ , given by the (separable) Hamiltonian

$$H = \sum_i \frac{p_i^2}{2ma^2} + \phi(\{x_i\}), \quad \text{with} \quad \tilde{\mathbf{v}}_i = \mathbf{p}_i/m. \quad (3.5)$$

The last stumbling block to obtain a solution of the system is thus given by constructing a gravitational potential from the points moving on characteristics. This is usually done by assuming that if the points are fairly sampling f , then they can also be used to estimate a density, from which then the gravitational force can be calculated through Poisson's equation. This point is a little bit subtle however. Assume we have a large number N of points $(\mathbf{x}_i, \mathbf{v}_i)$ that are uniformly sampling f . Then we can attach a mass m_b to them so that $\bar{\rho} = Nm_p/V$ gives just the mean density. A simple way of estimating the density will then be given by

$$\rho(\mathbf{x}, t) = m_b \sum_{i=1}^N \delta_D(\mathbf{x} - \mathbf{x}_i(t)). \quad (3.6)$$

This density estimate will however not be smooth, and in fact, it can lead to arbitrarily high accelerations if two points come arbitrarily close. Such accelerations would however be a pure artefact of the sampling. In the limit $N \rightarrow \infty$ they should never appear. To this end, a better estimate is given by

$$\rho(\mathbf{x}, t) = m_b \sum_{i=1}^N \delta_D(\mathbf{x} - \mathbf{x}_i(t)) * W_h(\mathbf{x}), \quad (3.7)$$

where W_h is some **softening kernel**, e.g. a Gaussian kernel, that 'smears' out each of the points on a scale h so that nearby points overlap and the density field becomes smooth. This procedure is called **force softening** and is needed to suppress spurious two-body effects. Choosing the width h is a bit like black magic, and we will encounter this aspect in more detail later on again.

In summary, we have seen that dark matter can be described by solving for the characteristic curves of the collisionless Boltzmann equation, along which the phase space density f is conserved. These are simply the equations of motion of N meta-particles that are just place-holders for sampling the local flow in phase space. They can be used to estimate a density field, which, in turn, can be used to calculate the gravitational forces. These forces are in turn used to advance the particles.

3.2 Collisionless simulations in one dimension

The simplest case, in which one can best see the peculiar properties of dark matter, is given by a 1+1 dimensional phase space, meaning we only have a one dimensional configuration space and a one dimensional velocity space. The beauty of it is that we can visualise all of phase space easily, since it is only two-dimensional (for 2+2 it is already 4 dimensional and impossible to plot without making projections).

3.2.1 Solving Poisson's equation in 1D

Another significant simplification in 1+1 dimensions is that we can trivially solve Poisson's equation, since now

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{3}{2} H_0^2 \Omega_m a^{-1} \delta, \quad (3.8)$$

which can be integrated once to yield the gravitational force field F

$$-F(x) = \frac{\partial \phi}{\partial x} = \frac{3}{2} H_0^2 \Omega_m a^{-1} \int_0^x \delta(x') dx', \quad (3.9)$$

Recall that $\delta(x) = (\rho(x) - \bar{\rho})/\bar{\rho}$, so that

$$\int_0^x \delta(x') dx' = \int_0^x \frac{\rho(x')}{\bar{\rho}} dx' - \int_0^x dx' = \frac{M(< x) - \bar{\rho}x}{\bar{\rho}}. \quad (3.10)$$

The last equality just says that the force is given by the mass to the left of point x minus the mass one would have at mean density at that point. Since the whole is divided by $\bar{\rho}$, it is simply a statement about distances. If we use the density estimator from eq. (3.6), one can easily see that if the particles are contained in a finite 'volume' L

$$\frac{M(< x_i)}{\bar{\rho}} - x_i = (\# \text{ particles to the left of } x_i) \frac{L}{N} - x_i. \quad (3.11)$$

The number of particles to the left of a given particle can be trivially found if one sorts all particles by increasing x_i , then the number of particles to the left is simply the rank of any given particle in the sorted list minus one.

Note that this does not include any force softening yet, we will ignore it here for the moment, not least because we can afford very large particle numbers in 1D.

3.2.2 Time Stepping

Since we know now how to calculate the gravitational force given the positions $\{x_i\}_{i=1\dots N}$ of N particles, we only have to find a way to integrate the equations of motion for the characteristics.

One can derive an explicit time integration scheme by remembering the definition of derivatives as the limit of a difference operation and then letting the difference take on a finite value

$$\dot{x}(t) = \lim_{\Delta t \rightarrow 0} \frac{x(t + \Delta t) - x(t)}{\Delta t} \rightarrow x(t + \Delta t) = x(t) + \Delta t \dot{x}(t) \quad (3.12)$$

where Δt should be small. We can then insert the right-hand-side of the differential equation for $\dot{x}(t)$ and we have a forward integration scheme. In total we would have

$$x_i(t + \Delta t) = x_i(t) + \Delta t a(t)^{-2} v_i(t) \quad (3.13)$$

$$v_i(t + \Delta t) = v_i(t) + \Delta t F(x_i(t)), \quad (3.14)$$

where $F(x_i(t)) = \nabla \phi|_{x_i(t)}$ is the gravitational force at position x_i evaluated at time t . This integration scheme is called the **forward Euler** scheme. It works, but it is not particularly stable since it does not conserve energy. This implies that in order to conserve energy, one needs to choose a very small time step, which makes the integration costly.

For this reason, one usually resorts to so-called [symplectic time integrators](#). They have the nice property that they are time-symmetric (i.e. they give back the same result if one integrates one step forward and then one backward in time). It is a consequence of [Noether's theorem](#) that such symmetries always lead to conservation properties. The simplest symplectic integrator is the so-called [leap-frog integrator](#). It has the form

$$x_i(t + \Delta t/2) = x_i(t) + \frac{\Delta t}{2} a(t)^{-2} v_i(t) \quad (3.15)$$

$$v_i(t + \Delta t) = v_i(t) + \Delta t F(x_i(t + \Delta t/2)) \quad (3.16)$$

$$x_i(t + \Delta t) = x_i(t + \Delta t/2) + \frac{\Delta t}{2} a(t + \Delta t)^{-2} v_i(t + \Delta t) \quad (3.17)$$

One can easily show that this integrator is indeed symmetric now. It uses always positions already advanced by half a time-step in order to compute the gravitational forces and then update the velocities and similarly for the positions. Obviously, for a whole simulation, strict reversibility and thus strict energy conservation would require that the time step Δt is constant. This criterion is however usually relaxed since if the variability is not too large, energy is still well conserved. The advantage of the leap-frog scheme is also that it does not require storing intermediate data for the time integration. This would however be the case for higher order symplectic methods.

Exercise 14 Show explicitly that the forward Euler method is not time-reversible, while the leap frog scheme is for constant time steps.

Exercise 15 Integrate a Kepler two-body problem with the forward Euler method and with the leap frog integrator. How does the lack of energy conservation manifest itself? What is the effect of the leading error of the leap-frog integrator in this case? Do you know which quantity is not conserved?

3.2.3 Initial Conditions for Plane Wave Collapse

Now that we have a way to calculate gravitational forces and a time integration scheme, we are almost done. All we need is initial particle positions $x_i(t_0)$ and velocities $v_i(t_0)$ at some initial time t_0 . Since we are dealing with particles that are following the fluid flow, we want to use Lagrangian perturbation theory in this case to set the initial conditions. This means that we can simply use the Zeldovich approximation applied to initial overdensity perturbation $\delta(x, t_0)$. The simplest initial condition that one can imagine here is that of an initial sinusoidal density perturbation of the form

$$\delta(x, t_0) = -A D(t_0) \cos(k_0 x) \quad \text{with } k_0 = \frac{2\pi}{L}, \quad (3.18)$$

where L is the length of the simulation volume that we want to simulate and A is some normalisation constant. This problem of a single collapsing wave is traditionally called a [Zeldovich pancake](#) or [Zeldovich plane wave collapse](#), and we will later see why. From this we can calculate the initial gravitational potential and velocity potential to be at the initial positions q

$$\phi_i(q) = A \frac{3H_0^2 \Omega_m}{2k_0^2} \frac{D(t_0)}{a(t_0)} \cos(k_0 q) \quad \psi(q) = -\frac{A}{k_0^2} \frac{D(t_0)}{a(t_0)} \cos(k_0 q) \quad (3.19)$$

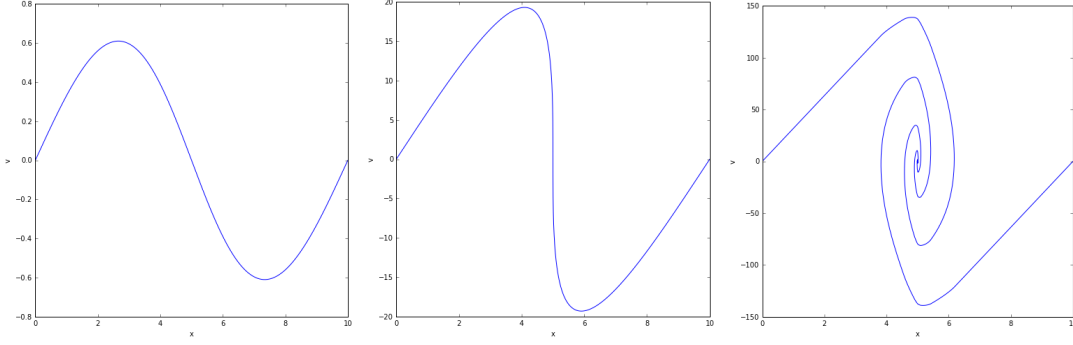


Figure 3.1: Phase space portrait of the plane wave collapse from an initial single sinusoidal perturbation. Shell-crossing occurs at some a_c . The left panel shows the distribution function at $a = a_c/10$, the central one at shell-crossing $a = a_c$ and the right panel at a much later time $a = 10a_c$.

The initial particle positions and velocities are then given by

$$x_i = q_i + D\nabla\psi = q_i + \frac{A}{k_0} \frac{D(t_0)^2}{a(t_0)} \sin(k_0 q_i) \quad (3.20)$$

$$v_i = a\dot{D}\nabla\psi = \frac{A}{k_0} D(t_0)\dot{D}(t_0) \sin(k_0 q_i) \Rightarrow \tilde{v}_i = av_i \quad (3.21)$$

We note that we can for simplicity often just assume that $D(t) = a(t)$, which is true in an Einstein-de Sitter universe (i.e. for matter domination).

Let us remember that the Zeldovich approximation predicted density singularities that were related to certain properties of the eigenvalues of the second derivative tensor $\partial^2\psi/\partial q_i/\partial q_j$. We can evaluate the density given in eq. (2.59) for our plane wave. We find

$$1 + \delta(q) = \left| 1 + Ak_0 \frac{D(t)^2}{a(t)} \cos(k_0 q) \right|^{-1} \quad (3.22)$$

which will become infinite if

$$Ak_0 \cos(k_0 q) = -\frac{a(t)}{D(t)^2} \sim -a(t)^{-1}. \quad (3.23)$$

Clearly this equation can have multiple solution in general. It will first happen at a time t_c for the point $q = \pi/k_0 = L/2$ where the cosine is most negative at -1 . For that point the condition then becomes $A = 1/(k_0 a(t_c))$ which allows us to choose the time when first shell-crossing appears conveniently in terms of the amplitude.

Figure 3.1 shows the result of a simulation of the plane wave collapse. The graphs show the distribution function f in 1+1 dimensional phase space for three times $a = a_c/10$, $a = a_c$ and $a = 10a_c$, before, at, and after shell crossing. At shell crossing, the distribution function has a vertical tangent, which implies that the density is formally infinite. After shell crossing, gravity leads to so called **filamentation** of the distribution function, which means that it starts to roll up in ever finer structures. Note that this is a genuine feature of the collisionless nature of dark matter. The particles coming from the left and right just pass through each other. No Maxwell-Boltzmann distribution (which would be a Gaussian in the velocity dimension at every point in configuration space) ever appears. Instead, in the region for

which shell-crossing has happened, the distribution function folds back and forth over the same region of space leading to a [multi-stream region](#).

Exercise 16 *Simulate the plane wave collapse yourself. You find a PYTHON skeleton to run such a simulation in Appendix A.5. All you need to do is add the leap-frog integrator yourself and specify the initial conditions. Plot the solution for different particle numbers at different times. How does the time step influence the solution at late times? Overplot the result of the Zeldovich approximation. Where does it work and where does it fail? Variation: instead of the single mode wave, set up random initial conditions with an initial power spectrum $P(k) \sim k^{-1}$.*

Hint: *It might be more straightforward to integrate everything in terms of a instead of t . Note that $dt = da/\dot{a}$ and you can use the Friedmann equation to compute \dot{a} .*

3.3 Collisionless simulations in three dimensions

For full three dimensional simulations both the computational effort and the algorithmic complexity increases substantially. Typically, one will run three dimensional simulations on big parallel supercomputers. It is beyond the scope of this introduction to teach parallel programming, or even the details of the algorithms involved, so we will just explain the underlying techniques and see what existing software can be used to run such simulations.

3.3.1 Force calculation in three dimensions – the PM method

The first challenge we are facing is that in three-dimensions, a simple integral of Poisson’s equation like eq. (3.10) will no longer be possible. There are different possibilities forward.

Fourier-space solution to Poisson’s equation

We had seen before that Fourier transform can be used very efficiently to solve Poisson’s equation, since in Fourier space the force field \mathbf{F} is just given by

$$\hat{\mathbf{F}}(\mathbf{k}) = -\frac{i\mathbf{k}}{|\mathbf{k}|^2} \hat{\delta}(\mathbf{k}). \quad (3.24)$$

This means that if we find a method to compute an overdensity field $\delta(\mathbf{x})$ from the positions of the N particles, and after solving eq. (3.24), are able to interpolate the force field back to the particle positions, we should have a workable method. So let us look at this aspect next.

Mass assignment schemes

The simplest method to assign particle positions to grid cells is to associate a particle with the nearest cell centre. This method is called [Nearest-Grid-Point \(NGP\)](#). It would be implemented as follows. For a grid spacing $\Delta x = L/N_g$, if L is the length of the computational domain and N_g the number of grid cells per dimension, then assign the mass of a particle with coordinates (x, y, z) entirely to the cell

$$(i, j, k) = \lfloor x/\Delta x, y/\Delta x, z/\Delta x \rfloor. \quad (3.25)$$

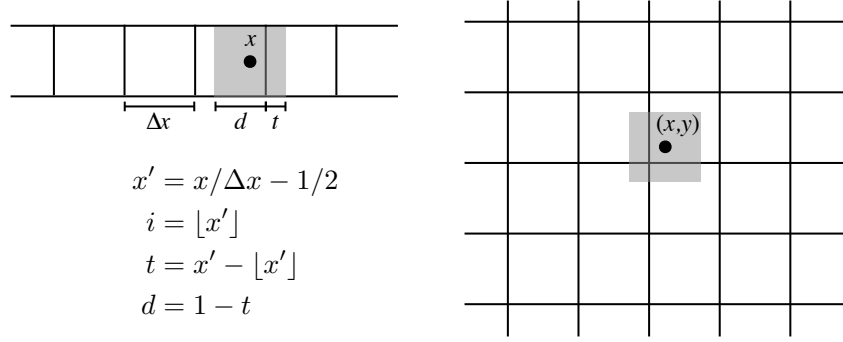


Figure 3.2: The Cloud-in-Cell (CIC) mass assignment scheme in one (left panel) and two (right panel) dimensions. For a particle at location x , one determines the left grid cell i as well as the overlap of the particle assuming a size Δx with cell i and the neighbour $i + 1$. In more than one dimension, one has to repeat this procedure for each dimension ending up with 2^d intersections.

Repeating this assignment for all particles yields an estimate of the density field which allows evaluation of eq. 3.24. **Important:** In order for a mass assignment scheme to be consistent, one has to use the same scheme for mass assignment and back-interpolation of the force \mathbf{F} , i.e. once we have \mathbf{F} , we simply look up the value from cell (i, j, k) for the particle at (x, y, z) according to eq. 3.25. Note that the force changes discontinuously for NGP assignment as a particle moves through the grid.

At the next higher order, one can use multi-linear interpolation, which is called **Cloud-in-Cell (CIC)** interpolation in the context of mass/charge interpolation methods. Instead of assigning all mass to a single cell, one calculates the overlap of a particle of hypothetical size Δx in each dimension with the 2^d cells it can intersect in d dimensions. The procedure is illustrated in Figure 3.2. The left intersection with cell i shall be called d , the one with the right neighbour $i + 1$ shall be called t . A fraction of the mass $m_p d$ is assigned to cell i , and the remainder $m_p t = m_p(1 - d)$ to cell $i + 1$. In three dimensions, the respective assignments become e.g.

$$\begin{aligned}
(i, j, k) &\leftarrow m_p d_x d_y d_z \\
(i + 1, j, k) &\leftarrow m_p t_x d_y d_z \\
(i, j + 1, k) &\leftarrow m_p d_x t_y d_z \\
&\dots
\end{aligned} \tag{3.26}$$

where the factor for the left cell is always the d and the one for the right the t , in total contributing to eight cells. In the reverse interpolation, the values of the force in the eight cells are summed with weights given by the products of t 's and d 's on the right hand side. For the CIC assignment, the force changes linearly as the particle moves through the grid.

In principle one can go to arbitrarily higher order interpolation. The next higher order is called **Triangular-shaped-clouds (TSC)** and it corresponds to quadratic multi-dimensional interpolation. It is rare that schemes with higher order than TSC are used in computational cosmology.

These three most important assignment functions can be explicitly written out as (following the book

by Hockney & Eastwood 1981)

$$W_{\text{NGP}}(x) = \begin{cases} 1 & |x| \leq \frac{H}{2} \\ 0 & \text{otherwise} \end{cases} \quad (3.27)$$

$$W_{\text{CIC}}(x) = \begin{cases} 1 - \frac{|x|}{H} & |x| < H \\ 0 & \text{otherwise} \end{cases} \quad (3.28)$$

$$W_{\text{TSC}}(x) = \begin{cases} \frac{3}{4} - \left(\frac{x}{H}\right)^2 & |x| \leq \frac{H}{2} \\ \frac{1}{2} \left(\frac{3}{2} - \frac{|x|}{H}\right)^2 & \frac{H}{2} \leq |x| \leq \frac{3H}{2} \\ 0 & \text{otherwise} \end{cases} \quad (3.29)$$

The 3D assignment function is then just the product $W_{3d}(x, y, z) = W(x) W(y) W(z)$.

These mass assignment schemes also have a softening effect on the mass distribution and facilitate to reach the *collisionless limit* when the ratio of particles/cell is sufficiently high. For the “naked” particle distribution that we had above

$$n(\mathbf{x}) = \sum_{i=1}^N \delta_D(\mathbf{x} - \mathbf{x}_i), \quad (3.30)$$

the density distribution after assigning the mass to the grid using an assignment function $W(\mathbf{x})$ is given at grid point locations \mathbf{x}_g by the convolution

$$\rho(\mathbf{x}_g) = \frac{m}{(\Delta x)^3} \int d^3x' n(\mathbf{x}') W(\mathbf{x}_g - \mathbf{x}') \quad (3.31)$$

where the assignment kernel for the various schemes can be shown to be in Fourier space given by the simple formula

$$\hat{W}(\mathbf{k}) = (\Delta x)^3 \text{sinc}^n(k/k_g), \quad (3.32)$$

with $n = 1$ for NGP, $n = 2$ for CIC and $n = 3$ for TSC. Since the back interpolation of forces to particles utilizes the same scheme for consistency, the effective smoothing is \hat{W}^2 so that the higher order assignment schemes are increasingly smoother and thus effectively low-pass filtered.

3.3.2 High force resolution and force splitting

One obvious problem is that the PM method does not scale well if high resolution is needed in certain regions of space. This scenario is often the case in astrophysics due to the nature of gravity as an attractive force. One possibility is due to introduce higher resolution meshes, deposit particles on these higher resolution meshes and then solve using an adaptive “relaxation method” such as the adaptive multigrid method. This is algorithmically a bit more involve, but is e.g. what is done in the RAMSES code.

A simpler alternative is to decompose the problem into a long-range force and a short range force. For the long-range force, one can then use the PM method on a relatively coarse mesh, and for the short range force, one can consider the direct force between particles only in their neighbourhood – since the particles further away contribute through the PM part. Instead of the direct force, one often uses here a tree-method (see next section) for the short range force (this is e.g. what the GADGET-2 code is doing).

In order to achieve a consistent cut, the force is split into a long-range and short-range part using a sigmoidal cut function in Fourier space, e.g. a Gaussian, cutting at scale r_s

$$\begin{aligned}\hat{R}_{\text{long}}(k; r_s) &= \exp(-k^2 r_s^2) \\ \hat{R}_{\text{short}}(k; r_s) &= 1 - \exp(-k^2 r_s^2).\end{aligned}\tag{3.33}$$

The appropriately modified Green's functions in Fourier and real space obtained by multiplication or convolution respectively, i.e.

$$\hat{G}(k) = -\frac{1}{k^2} \hat{R}(k; r_s)\tag{3.34}$$

$$G(r) = -\frac{1}{r} R(r; r_s)\tag{3.35}$$

for which we find in real space

$$\begin{aligned}G_{\text{long}}(r; r_s) &= -\frac{1}{r} \text{erf}\left(\frac{r}{2r_s}\right) \\ G_{\text{short}}(r; r_s) &= -\frac{1}{r} \text{erfc}\left(\frac{r}{2r_s}\right).\end{aligned}\tag{3.36}$$

Instead of the normal Green's functions, one thus simply uses these truncated functions instead and obtains a hybrid code. In order to use this approach, one simply replaces the PM Green's function with G_{long} and instead of the particle-particle interaction in the direct summation or tree force (see below), one uses G_{short} for the potential, and ∇G_{short} for the force.

3.3.3 Tree-PM codes

As we have said before, while easy and relatively fast, the problem with solving Poisson's equation on a uniform grid is that this approach does not lend itself well to situations that require a large dynamic range. At the same time, direct summation of individual particle-particle forces is quadratic in the number of particles and thus becomes quickly computationally prohibitive. In such cases, e.g. in the context of galactic dynamics, or the dynamics of star clusters, tree methods have however proven their worth. The idea of the Barnes & Hut tree algorithm is to circumvent the N^2 scaling problem in direct summation of the mutual interactions between N bodies by a hierarchical approach and 'meta'-particles as illustrated in Figure 3.3.

At each node in the tree, one inserts a 'meta' particle that carries the mass and sits at the centre of mass of the branch of the tree. The total interaction

$$\phi(\mathbf{x}) \propto \sum_j \frac{1}{|\mathbf{x} - \mathbf{x}_j|} = \sum_j \frac{1}{|(\mathbf{x} - \boldsymbol{\lambda}) - (\mathbf{x}_j - \boldsymbol{\lambda})|}\tag{3.37}$$

sum can be multipole expanded using

$$\frac{1}{|\mathbf{y} + \boldsymbol{\lambda} - \mathbf{x}_j|} \simeq \underbrace{\frac{1}{|\mathbf{y}|}}_{\text{monopole}} - \underbrace{\mathbf{y} \cdot \frac{\boldsymbol{\lambda} - \mathbf{x}_j}{|\mathbf{y}|^3}}_{\text{dipole}} + \underbrace{\frac{1}{2} \mathbf{y}^T \frac{3(\boldsymbol{\lambda} - \mathbf{x}_j)(\boldsymbol{\lambda} - \mathbf{x}_j)^T - \mathbf{I}(\boldsymbol{\lambda} - \mathbf{x}_j)^2}{|\mathbf{y}|^5} \mathbf{y}}_{\text{quadrupole}} + \dots\tag{3.38}$$

If we set $\boldsymbol{\lambda}$ to be the centre of mass, then the dipole vanishes. The trick is now that since we know that $|\mathbf{x}_j - \boldsymbol{\lambda}|$ is bounded by the space partitioning cell size, we can directly control the accuracy by accepting

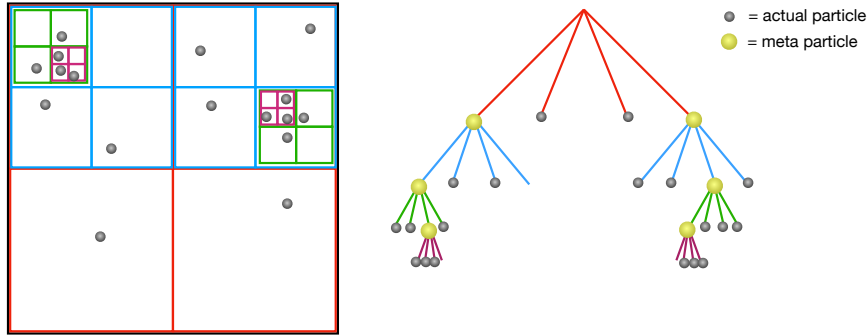


Figure 3.3: The Barnes & Hut tree algorithm. Instead of a direct summation, one organises particles in a hierarchical tree with ‘meta’ particles located at the centre of mass of the respective sub-tree at whose root they are located and carrying the total mass of the branch. Depending on the distance of the particle which wishes to compute the interaction, one can replace the entire tree branch with the meta-particle to good accuracy. For even better accuracy, one can carry out a multi-pole expansion of the branch so that quadrupoles, etc. can also be carried along. The effective algorithm ends up being $N \log N$ in the number of particles instead of N^2 .

a meta particle as a valid approximation for the entire branch if

$$\theta = \frac{\ell}{|\mathbf{y}|} < \theta_c, \quad (3.39)$$

where ℓ is the cell size of the node, and θ_c the threshold opening angle that we set. In the limit of small θ_c , one simply recovers direct summation since all nodes are opened and only the leafs are used.

The state-of-the-art codes are exploiting all existing techniques. A very promising one for very large particle count simulations is the Fast-Multipole-Method, used e.g. in the PKDGRAV3-code, or the geometric multigrid method, used e.g. in the RAMSES code. The newest codes also readily utilise thousands of GPUs to generate simulated Universes for the upcoming generations of cosmological observations.

3.3.4 Running a simulation with MUSIC and GADGET-2

There is no need to implement all of the techniques, you are very welcome to try out the existing and freely available software MUSIC (<https://www-n.oca.eu/ohahn/MUSIC/>) to generate cosmological initial conditions, and GADGET-2 (<http://wwwmpa.mpa-garching.mpg.de/gadget/>) to run cosmological N -body simulations.

We’ll discuss in detail in the course (interactively) how to set up and start a simulation.

3.4 Halo Finding and Merger Trees

3.4.1 Extracting gravitationally bound structures from N-body simulations

One of the most important tasks in the post-processing of cosmological simulations is the identification of gravitationally bound structures: the haloes. The existence of a local minimum of the gravitational potential is a necessary condition for the existence of a bound structure. A local minimum of the potential however corresponds also to local maximum of the matter density field. Thus, most halo finders focus on determining the region of gravitational boundness around a density maximum. Many methods exist, below the most commonly used for isolated haloes are explained, and we discuss a few existing software packages after. Calculations using the model spherical collapse of a single top hat perturbation in EdS cosmology indicate that regions that are in virial equilibrium should have an overdensity ρ_{vir} of

$$\rho_{\text{vir}} = \Delta_{\text{vir}} \rho_c \quad \text{with} \quad \Delta_{\text{vir}} = 18\pi^2 \simeq 178 \quad (3.40)$$

where ρ_c is the critical density of the universe.

The Friends-of-Friends (FoF) Algorithm

Press & Davis (1982) and Davis et al. (1985) introduced the so-called friends-of-friends algorithm that is particularly suited for N-body simulations. The algorithm builds equivalence classes among all particle pairs in the N-body simulation whose distance falls below a given threshold. The resulting equivalence classes thus comprise regions bounded by an isodensity surface. It is common practice to relate the distance threshold to the mean inter-particle distance b and adopt a value of $b = 0.2$. This choice can be roughly motivated by considering the density of an isothermal sphere at the virial radius. Expressed in units of the mean inter-particle separation, $b \sim (\Delta_{\text{vir}}/3)^{-1/3} \approx 0.25$ in this case. One main problem of the friends-of-friends algorithm is that it tends to spuriously connect nearby haloes that are linked by a bridge which may purely be a result of the stochastic nature of N-body simulations. This problem is most severe in regions of high environmental density.

However, friends-of-friends haloes are not constrained in their shape by the algorithm and mass functions obtained with this algorithm show a nearly universal behaviour across redshift.

The Spherical Overdensity (SO) Algorithm

The second most commonly employed halo finding method is the spherical overdensity algorithm introduced by Lacey & Cole (1994). Here, particles are ranked according to their local density. Then a sphere is grown around the first particle until the mean density within that sphere falls below the virial overdensity, i.e. the virial radius R_{vir} is defined as $M_{\text{vir}} = M(< R_{\text{vir}})$ (where $M(< r)$ is the mass enclosed in a sphere of radius r) so that

$$\frac{M(< R_{\text{vir}})}{\frac{4\pi}{3} R_{\text{vir}}^3} = \Delta_{\text{vir}} \rho_c. \quad (3.41)$$

All particles within the sphere make up the first halo and are removed from the list of density ranked particles. The procedure is then repeated with the next highest density particle left in the list and stopped, when no more groups are found that contain a given minimum number of particles. The resulting haloes do not suffer from spurious linking of distinct haloes. They are however spherical by definition which

might not reflect the true shape of the gravitationally bound volume. A further advantage of the spherical overdensity method is that it can be readily used also in hydrodynamic simulations where particles and mesh cells come with basically arbitrary masses so that the simple friends-of-friends algorithm cannot be applied with a fixed distance threshold. Instead of ranking particles by their density, any method to identify peaks of the density field is sufficient to serve as the starting point before growing spheres around the peaks.

For simplicity, often 200 is used instead of 178 and haloes that are defined using this overdensity criterion are often denoted with an index $200c$, so that the usual mass definition for a halo is M_{200c} . In the sub-field of cosmology that uses galaxy clusters as cosmological probes, usually a higher density threshold of $500c$ is used that better correlates with the hot X-ray emitting region of the cluster halo. Sometimes also the mean density is used instead of the critical density so that $200m$ refers to $\rho_{\text{vir}} = 200\rho_m = 200\Omega_m\rho_c$. In the presence of a cosmological constant, in principle the overdensity criterion becomes time dependent and a convenient parameterisation has been given by Bryan&Norman (1998) as

$$\Delta_{\text{vir}}(a) = 18\pi^2 + 82x - 39x^2, \quad \text{with} \quad x := \frac{\Omega_m a^{-3}}{\Omega_m a^{-3} + \Omega_\Lambda} - 1. \quad (3.42)$$

3.4.2 Running a Halo Finder

For practical purposes, we shall use the ROCKSTAR software package (<https://bitbucket.org/gfcanford/rockstar>).

3.5 Quantifying Large-Scale Structure

3.5.1 The Halo Mass Function

3.5.2 The Halo Correlation Function and Halo Bias

Chapter 4

The Baryonic Universe

Appendix A

Hints for the computational exercises

A.1 Numerically integrating the Friedmann equation in PYTHON

This section provides hints for Exercise 2, where a numerical solution of the Friedmann equation is sought. For this exercise, we will need the following PYTHON packages

```
1 import numpy as np
2 import matplotlib.pyplot as plt
3 from scipy.integrate import quad
4 %matplotlib inline
```

We will use the `scipy` routine `quad` to numerically integrate a function. The main problem we have to circumvent is that the limits of integration that are passed to `quad` cannot be vectors. This means that we will have to write a function around `quad` that loops over an array of values, passes them one by one to `quad` and stores the result in a new array. The skeleton to integrate the Friedmann equation in order to find $t(a)$, i.e. the cosmic time as a function of the scale factor, then might look like this:

```
1 def Hubble_a(a):
2     return ...
3
4 def t_of_a(a):
5     res = np.zeros_like(a)
6     for i, ai in enumerate(a):
7         t, err = quad(lambda ap : 1.0/(ap*Hubble_a(ap)), 0, ai)
8
9         res[i] = t
10    return res
11
12 a = np.logspace(-8, 1, 100)
13
14 plt.loglog(t_of_a(a)*H0, a)
```

```

15 plt.xlabel('t * H0')
16 plt.ylabel('a(t)')

```

We have not provided the implementation for the function $\text{Hubble_a} = H(a) = \frac{\dot{a}}{a}$, but leave this as the main purpose of the exercise besides learning how to integrate such functions numerically. There is one more exotic bit in this listing. We are making use of PYTHON lambda expressions. This is just a short hand notation for a function definition, i.e. we define in place a nameless function that maps $a_p \mapsto (a_p H(a_p))^{-1}$, which is the function that we want to integrate over. Finally, we create an array of 100 values of logarithmically spaced scale factors $10^{-8} \leq a \leq 10$, calculate the cosmic times (in units of the Hubble time after multiplying with H_0), and plot the result as time against a .

A.2 A linear dark matter + baryon model in PYTHON

This section provides a skeleton for the time integration of the gravitationally coupled dark matter and baryonic fluid through recombination from Exercise 7. We will use the following PYTHON packages, most of them as usual

```

1 import numpy as np
2 import matplotlib.pyplot as plt
3 from scipy.integrate import quad
4 from scipy.integrate import odeint
5 from scipy.integrate import ode
6 from scipy.special import erf,erfc
7 %matplotlib inline

```

Next we define the usual constants, in addition the baryon fraction $f_b = \Omega_b/\Omega_m$ and the cold dark matter fraction $f_c = 1 - f_b$.

```

1 H0      = 70.0
2 Omegab  = 0.045
3 Omegam  = 0.3
4 Omegal  = 0.7
5 Omegar  = 8e-5
6 fb      = Omegab/Omegam # baryon fraction
7 fc      = 1.0-fb        # CDM fraction

```

The more involved bit is that we need to model the recombination of the plasma and the evolution of the baryon sound speed in some crude way. We assume that prior to recombination, the baryons are tightly coupled to the photons and that the sound speed is $c_s = c/3$. After recombination, we will use the fitting formula given in the exercise to estimate the baryon temperature, and calculate the sound speed from it (we will make this value artificially $20\times$ larger in order to have the effect on larger scales). Then, we can either assume instantaneous recombination at $a = 10^{-3}$, or spread it out artificially thus mocking that it is not instantaneous. The two functions for slow and fast recombination are just that. Altogether we thus have something like

```

1 def T_baryon(a):
2     # parameterisation of baryon temperature from Tseliakhovich & Hirata 2010
3     return 2.726/a/(1+119.*a/(1.0+(115.*a)**-1.5))
4
5 def slow_recombination( a, c1, c2 ):
6     f = 0.5*(1.0-erf((np.log10(a)+3)*16))
7     return f*c1 + (1.0-f)*c2
8
9 def fast_recombination( a, c1, c2 ):
10    if a<1e-3:
11        return c1
12    else:
13        return c2
14
15 def sound_speed(a):
16     c1 = clight/3.0
17     c0_squared = 0.0112764987
18     c2 = np.sqrt(c0_squared*T_baryon(a)) * 20.0 # we make it 20 times larger!!
19     cs = slow_recombination(a,c1,c2)
20     #cs = fast_recombination(a,c1,c2)
21     return cs

```

We will use the `scipy` function `odeint` to integrate the ordinary differential equations. It will take an array of variables `u[...]` describing the current state vector. We will store in it all fields that we want to evolve in time, i.e. we set $\mathbf{u} = (\delta_c, \delta_b, \theta_c, \theta_b, a)$. The set of coupled ordinary differential equations then takes the form $\dot{\mathbf{u}} = F(\mathbf{u}, t)$ and we will also have to provide the function F taking the right hand side of the differential equations, which is why we will call this function `RHS`. We will need only one additional parameter (all constants we have as global variables), the wave number k that we want to evolve. The function `RHS` then has the form

```

1 #####
2 # the RHS of the coupled equations of baryons and dark matter
3 # u[0] = delta_c
4 # u[1] = delta_b
5 # u[2] = theta_c
6 # u[3] = theta_b
7 # u[4] = a
8 def RHS(u,t,param):
9     k = param[0]
10    du = np.zeros_like(u)    # make du the same shape as u
11    a = u[4]
12    H = Hubble_a(a)          # we will need H(a), so calculate it
13    cs = sound_speed(a)      # we will also need the sound speed at a, so get it
14
15    du[0] = ... # RHS for d delta_c / dt
16    du[1] = ... # RHS for d delta_b / dt
17    du[2] = ... # RHS for d theta_c / dt
18    du[3] = ... # RHS for d theta_b / dt
19    du[4] = ... # RHS for da / dt
20    return du

```

and returns the vector **du** containing the right-hand-sides of the ODEs. The last thing we will need is to set up initial conditions, and write the actual loop over the range of wave numbers k for which we want to evolve **u**. For the initial conditions, we will simply set $\delta_c(t_0) = 1, \delta_b(t_0) = \theta_c(t_0) = \theta_b(t_0) = 0$, we choose to start at $a = 10^{-10}$ and evolve to some $a = \text{astop}$ that we can adjust as we like. For the wave numbers, we will choose 200 values logarithmically spaced $10^{-3} \leq k \cdot \text{Mpc} \leq 10$. We get something like this

```

1  ainit=1e-10 # the starting scale factor
2  astop=1e-4  # we want to finish at astop
3  tstart= time_of_a(ainit) # compute the cosmic time for ainit
4  tstop = time_of_a(astop) # and for astop
5
6  ## setup the initial conditions for CDM density and theta, also for baryon
7  deltac0 = 1.0
8  thetac0 = 0.0
9  deltab0 = 0.0
10 thetab0 = 0.0
11
12 ## package all initial conditions into one vector
13 uinit = np.array([deltac0, deltab0, thetac0, thetab0, ainit])
14
15 ## these will be the wave numbers k for which we compute the result:
16 kvals = np.logspace(-3,1,200)
17
18 ## setup vectors for the results deltab(k), deltac(k), ...
19 deltab = np.zeros_like(kvals)
20 deltac = np.zeros_like(kvals)
21 thetab = np.zeros_like(kvals)
22 thetac = np.zeros_like(kvals)
23 afinal = np.zeros_like(kvals)

```

Finally, we need the main loop with the integration. We choose a fixed number of 1000 integration steps in logarithmic time. If we increase to larger values of k this will not be enough. In fact, the correct time step should be chosen based on the sound speed. For simplicity we will simply choose a fixed number and make sure that the result is converged. The main loop then looks like this

```

1  ## set the number of time steps in log time in a rough way
2  nsteps = 1000 # normally this should depend on k, so check convergence!
3
4  ## these will be the times at which we evaluate
5  time = np.logspace(np.log10(tstart), np.log10(tstop), nsteps)
6
7  ## loop over all wave numbers k and evolve them
8  for i,k in enumerate(kvals):
9      ## integrate the system for all times in 'time',
10     u = odeint(RHS, uinit, time, args=([k],))
11
12     ## store last result (in time) in the respective results
13     deltac[i] = u[-1,0]
14     deltab[i] = u[-1,1]
15     thetac[i] = u[-1,2]

```

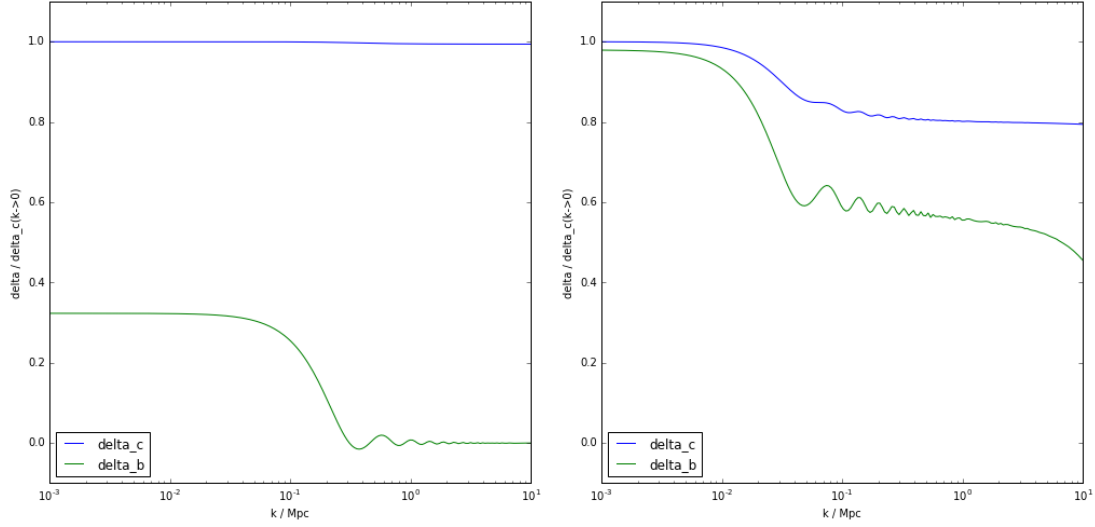


Figure A.1: The amplitude of dark matter and baryon density fluctuations prior to recombination at $a = 10^{-4}$ (left) and some time after recombination at $a = 10^{-2}$ calculated with a simple linear perturbation approximation for the coupled fluid. The amplitudes are normalised so that $\delta_c = 1$ for the lowest k that we calculated.

```

16     thetab[i] = u[-1,3]
17     afinal[i] = u[-1,4]
18
19 print 'Finished integration to a=',afinal[0]

```

This completes the program. We can plot the results and find a k -dependent amplitude for δ_c and δ_b as shown in Figure A.1 at $a = 10^{-4}$.

A.3 Generating a realisation of the young Universe

Here, we seek to make a realisation of a Gaussian random field with a prescribed power spectrum $P(k)$, i.e. we want to create a density field $\delta(\mathbf{x})$ with

$$\hat{\delta}(\mathbf{k}) = \sqrt{P(k)} \hat{G}(\mathbf{k}), \quad \text{so that} \quad P(k) \propto \hat{\delta}(\mathbf{k}) \hat{\delta}^*(\mathbf{k}), \quad (\text{A.1})$$

where $\hat{G}(\mathbf{k})$ is a Gaussian random field of zero mean and unit variance. Let's first begin by creating the Gaussian random field. We will need the following imports in this exercise.

```

1 import matplotlib.pyplot as plt
2 import numpy as np
3 import numpy.fft
4 import numpy.random
5 import matplotlib.mlab as mlab
6 %matplotlib inline

```

A.3.1 Generating the realisation

First, let us create the Gaussian random field of size N^3 . This is achieved using the PYTHON function `np.random.normal` for which one can specify the size of the field, the mean and the variance. By default the mean is zero and the variance is unity. Next we will plot an image of a slice of the field and also visualise its distribution:

```
1  # create a Gaussian random field of size N
2  N = 128
3  delta0 = np.random.normal(size=[N,N,N])
4
5  # plot a slice of the field (center of box)
6  fig = plt.figure()
7  cax = plt.pcolor(delta0[:, :, N/2], vmin=-3, vmax=3)
8  plt.axis('image')
9  cbar = fig.colorbar(cax)
10
11 # plot a histogram of the values, they should follow
12 # a Gaussian distribution
13 fig2 = plt.figure()
14 n = plt.hist( np.reshape(delta0, [N*N*N]), 100, normed=1 )
15
16 # overplot a Gaussian distribution of zero mean, unit variance
17 x = np.linspace(-5,5,100)
18 y = mlab.normpdf( x, 0.0, 1.0)
19 plt.plot(x, y, 'r--', linewidth=2)
```

The next part is a little bit more involved. We will need to compute the three-dimensional Fourier transformation of the field, and then multiply it with the square root of the power spectrum. Let's assume we have defined a function that returns the square root of $P(k)$ called `delta_amplitude(k)`. Since $\delta(\mathbf{x})$ is a real valued field, we can use the numpy function `rfftn` to compute the 3D Fourier transformation by means of an FFT (Fast Fourier Transform):

```
1  # compute the Fourier transform
2  fdelta0 = np.fft.rfftn( delta0 )
```

Now `fdelta0` holds the transformed field, but be aware that the modes \mathbf{k} are not ordered in a trivial way. We can use the following piece of code to get both the vector components of \mathbf{k} as well as the norm $k = \|\mathbf{k}\|$:

```
1  Lbox = 1000          # the box size, sets the fundamental mode
2  kmin = 2.0*np.pi/Lbox # the fundamental mode of the box
3  kmax = kmin * N/2     # the Nyquist mode of each linear dimension
4
5  d=1.0/kmin/N
6  k_full = np.fft.fftfreq(N,d=d)
7  k_half = np.fft.rfftfreq(N,d=d)
8  kx,ky,kz = np.meshgrid(k_full,k_full,k_half) # get vec. components
9  k = np.sqrt(kx**2+ky**2+kz**2)              # get norm k
```

If you want to learn more about the data layout in FFT transformed real fields, see [here](#).

Now, since you have k , you can simply compute $\delta(k) = \sqrt{P(k)} \hat{G}(\mathbf{k})$, since \hat{G} is already stored in `fdelta0`. The final result can be obtained by an inverse Fourier transformation, and we'll once again display a slice of the field as an image:

```

1 fdelta = ...
2
3 # transform back
4 delta = np.fft.irfftn(fdelta)
5
6 # plot a slice of the image
7 plt.pcolor(delta[:, :, N/2])
8 plt.axis('image')

```

A.3.2 Measuring the sampled power spectrum

Now we should have a realisation of the field which follows the power spectrum we wanted. Let's verify that this is indeed the case. For this we can estimate a power spectrum from the field like this:

```

1 # decide number of bins and set up bin boundaries
2 nbins = 20
3 kbinlimits = np.logspace(np.log10(kmin*2), np.log10(kmax), nbins+1)
4
5 # initialise fields for results
6 kbin = np.zeros(nbins)
7 Pk = np.zeros(nbins)
8 ePk = np.zeros(nbins)
9
10 # a normalisation factor we need for the FFT
11 normfac = 1.0/np.sqrt(N*N*N)
12
13 for i in range(nbins):
14     kleft = kbinlimits[i] # left bin boundary
15     kright = kbinlimits[i+1] # right bin boundary
16
17     # get indices of all values of k that fall into bin
18     ii = np.logical_and( k>kleft, k<=kright)
19
20     # get values of fdelta that belong to these k
21     data = fdelta[ ii ] * normfac
22     kbin[i] = np.mean( k[ii] )
23     Pk[i] = ...
24     ePk[i] = ...

```

Estimate $P(k)$ from the mean of $\hat{\delta}\hat{\delta}^*$ that fall into the bin, and the error on that mean from the standard deviation divided by the square root of the number of values in that bin. The number of values in the bin can be obtained from `len(data)`. Then plot the measured power spectrum against $P(k)$. Did it work? The result should look like what is shown in Figure [A.2](#).

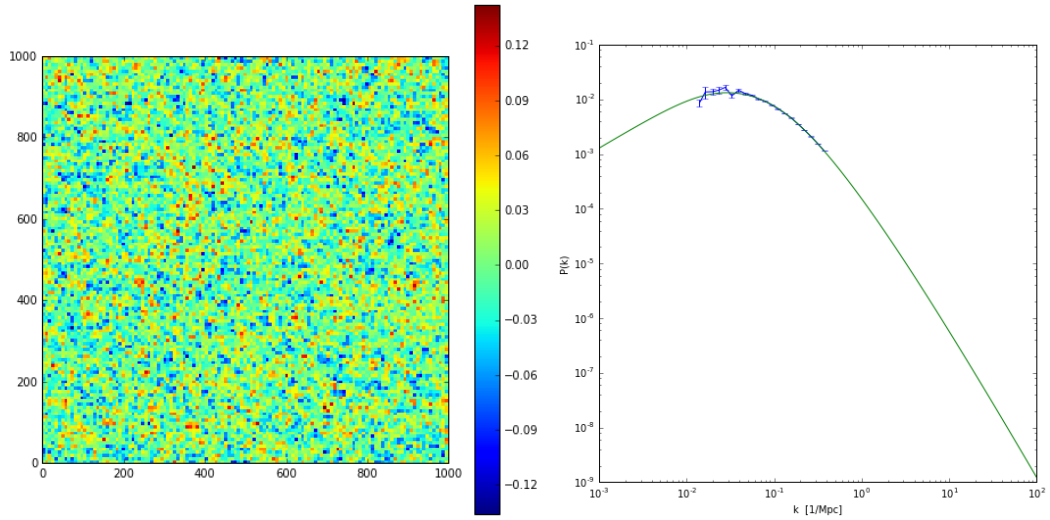


Figure A.2: Left: A realisation of a Gaussian random field with a prescribed power spectrum $P(k)$. Right: sampled power spectrum of the field realisation compared to the $P(k)$ from which we sampled. The box size was 1 Gpc with $N = 128$ for a BBKS transfer function and $n_s = 0.961$.

A.4 Plotting point distributions in projection and in 3D

A.5 Simulating Plane Wave Collapse

This is the code to compute the collapse of dark matter in 1+1 dimensional phase space. You need to fill in the initial conditions in function `compute_Zeldovich` yourself and write the leap-frog integrator yourself. You can calculate the acceleration that you use to update the velocities by calling `data = compute_acc(data)`, you can then access the accelerations via `data['a']`.

```

1  import matplotlib.pyplot as plt
2  import numpy as np
3  %matplotlib inline
4
5  Lbox = 1.0      # length of the simulation volume
6  Np = 512       # number of particles in the simulation
7
8  Omegam = 1.0
9  Omegal = 0.0
10 Omegak = 1.0 - Omegam - Omegal
11 H0 = 70.0
12
13 astart = 0.01  # starting time
14 across = 0.1   # time when shell crossing shall appear
15 astop = 0.1    # time when to stop the simulation
16
17 # calculation of the amplitude
18 amp = Lbox / (2*np.pi*across)

```

```

1 dtype = [('q', 'float'), ('x', 'float'), ('v', 'float'), ('a', 'float')]
2
3 def adot(a):
4     return H0 * a * np.sqrt( Omegal + Omegak/a**2 + Omegam/a**3 )
5
6 # compute initial conditions using the Zeldovich approx.
7 def compute_Zeldovich(q,a):
8     data = np.zeros(np.shape(q), dtype=dtype)
9     data['q'] = q
10    data['x'] = Lbox*q + ...
11    data['v'] = ...
12    return data
13
14 # set up the Lagrangian coordinates
15 q = np.linspace(0.0,1.0,Np)
16
17 # set up the data vector using Zeldovich approx.
18 data = compute_Zeldovich(q,astart)

```

The snippet above sets the global variables and the initial conditions. The fields for q , x and v should be initialised for all particles now. Next is the main time stepping loop, you will need to write the integrator yourself.

```

1 # computes the accelerations of each particle
2 xundisp = q*Lbox
3 def compute_acc( dat ):
4     meanpos = np.mean(dat['x']) - 0.5*Lbox
5     dat = np.sort( dat, order='x' )
6     dat['a'] = -(xundisp - (dat['x']-meanpos))
7     return dat
8
9 # the main time stepping loop
10 a = astart
11 da = 1e-4
12 while a<astop:
13     # drift
14     data['x'] = data['x'] + ...
15     a = a + 0.5*da
16
17     # kick
18     data = compute_acc(data)
19     data['v'] = data['v'] + ...
20     a = a + 0.5*da
21
22     #drift
23     data['x'] = data['x'] + ...
24
25     istep = istep+1
26
27 print 'arrived at a=',a, ' in ',istep, ' time steps'

```

In order to see the phase space spirals, you need to plot `data['x']` against `data['v']`.

Bibliography

- [1] S. Burles, K. M. Nollett, and M. S. Turner. Big-Bang Nucleosynthesis: Linking Inner Space and Outer Space. *ArXiv Astrophysics e-prints*, March 1999.
- [2] S. Dodelson. *Modern cosmology*. Academic Press, 2003.
- [3] A. Einstein. Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften (Berlin)*, Seite 142-152., 1917.
- [4] A. Friedmann. Über die Krümmung des Raumes. *Zeitschrift für Physik*, 10:377–386, 1922.
- [5] E. Hubble. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, 15:168–173, March 1929.
- [6] G. Lemaître. Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques. *Annales de la Société Scientifique de Bruxelles*, 47:49–59, 1927.
- [7] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck 2015 results. XIII. Cosmological parameters. *A&A*, 594:A13, September 2016.
- [8] D. Tseliakhovich and C. Hirata. Relative velocity of dark matter and baryonic fluids and the formation of the first structures. *Phys. Rev. D*, 82(8):083520, October 2010.