Oxford Physics

M. Phys. Major Option C1

Astrophysics

Cosmology Lecture Notes (2022-2023)

Julien Devriendt

DENYS WILKINSON BUILDING,
KEBLE ROAD, OXFORD OX1 3RH

julien.devriendt@physics.ox.ac.uk

JD—NOVEMBER 15, 2022

1 Disclaimer

Special thanks to Pedro Ferreira who kindly provided his Cosmology lectures notes and Adrianne Slyz who graciously shared the material of a chapter she wrote on the relation between fine-tuning and large scale structure (OUP, J. Silk et al editors). My lectures notes heavily draw from these two sources, and in particular cover much the same material and adopt a very similar structure. I have also endeavoured to use the same notation as that of Steve Balbus' lectures on GR & Cosmology in the 3rd year course. Thus, although I present a brief overview of the homogeneous Universe model at the beginning of these lecture notes to make them more self-contained and refresh the memory of the reader concerning this notation, I assume that the material related to this topic is known. Readers not familiar with GR will benefit greatly from reading Steve's notes on the Newtonian expanding Universe. This latter is all that is required to follow my course, provided one is not interested in the nitty-gritty details of how GR perturbation theory underpins the validity of the Newtonian approach in tackling large scale structure.

These notes are, by and large, a direct transcription of what I write on the blackboard (with the exception of GR perturbation theory where I provide more intermediate steps in this manuscript). Thus they should be considered as an introduction to the subject rather than a definitive treatment. I perused many text books to write these notes, and here are my main sources of inspiration:

[L]: A. Liddle, *An Introduction to Modern Cosmology*, Wiley

[P]: J. Peacock, *Cosmological Physics*, CUP

[TP] T. Padmanabhan, *Large Scale Structure*, CUP

[D]: S. Dodelson, *Modern Cosmology*, Academic Press

[KT]: E. Kolb & M. Turner, *The Early Universe*, Addison Wesley.

[W]: S. Weinberg, *Gravitation and Cosmology*, Wiley

In each section of these notes, I have suggested which of these books I believe is best suited to study the topic at hand in more depth.

Finally, I apologise in advance for any typo/error in these notes: please do not hesitate to contact me when you find one (including a departure of notation from the 3rd year course) so I can get the chance to fix them quickly. This will be much appreciated.

2 The Homogeneous Universe

As was established in the 3rd year GR & Cosmology course, the most general spacetime metric in an homogeneous and isotropic Universe, where one can define a uniform time is given by (see e.g. [W], chapter 13 for proof):

$$\begin{aligned} ds^2 &= g_{\mu\nu} dx^\mu dx^\nu \\ &= \begin{pmatrix} cdt & dr & d\theta & d\phi \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & a^2[t]/(1-Kr^2) & 0 & 0 \\ 0 & 0 & a^2[t]r^2 & 0 \\ 0 & 0 & 0 & a^2[t]r^2\sin^2\theta \end{pmatrix} \begin{pmatrix} cdt \\ dr \\ d\theta \\ d\phi \end{pmatrix} \\ &= -c^2 dt^2 + a^2[t]\left(\frac{dr^2}{1-Kr^2} + r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2\right) \end{aligned} \quad (1)$$

where $g_{\mu\nu}$ is the (diagonal) *covariant* metric tensor, c is the speed of light in vacuum, K is a constant indicative of the spatial curvature of the Universe and $a[t]$ is the time dependent expansion factor*. r, θ, ϕ define a *comoving* (spherical) coordinate system — i.e. a system in which an observer at rest moves along with the expansion of the Universe. This metric is called the Friedmann–Lemaître–Robertson–Walker (FLRW) metric, and we have adopted the space-like sign convention (−,+,+,+) for the metric signature, as in the 3rd year course.

The same considerations about homogeneity and isotropy also imply that if we sit in the comoving frame in which the fluid† which fills the Universe is at rest‡, its energy–momentum tensor can be written in a simple diagonal form as:

$$T_\mu^\nu = \begin{pmatrix} -\rho[t]c^2 & 0 & 0 & 0 \\ 0 & P[t] & 0 & 0 \\ 0 & 0 & P[t] & 0 \\ 0 & 0 & 0 & P[t] \end{pmatrix} \quad (2)$$

where $\rho[t]$ and $P[t]$ are the (time dependent) *average* density and pressure of of the fluid. Indeed, starting from the general form of the energy-momentum tensor

$$T_{\mu\alpha} = (\rho[t] + P[t]/c^2)U_\mu U_\alpha + P[t]g_{\mu\alpha}$$

where $U_\mu = (-c, 0, 0, 0)$ is the *covariant* 4-velocity of the fluid at rest, we can multiply it by the *contravariant* (or inverse) metric tensor,

$$g^{\alpha\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & (1-Kr^2)/a^2[t] & 0 & 0 \\ 0 & 0 & 1/(a^2[t]r^2) & 0 \\ 0 & 0 & 0 & 1/(a^2[t]r^2\sin^2\theta) \end{pmatrix}$$

---

*$K = 0$ for a flat (Euclidean) universe, $K > 0$ for a closed (spherical) universe and $K < 0$ for an open (hyperbolic) one. In what follows, contrary to what was done in the 3rd year course, we will choose that $K^{-1/2}$ (= a in 3rd year course parlance), like r, has the dimensions of a length, so as to be able to normalise the *dimensionless* expansion factor (called $a[t]$ here instead of $R[t]$ in the 3rd year course) to unity at present, i.e. set $a_0 \equiv a[t_0] = 1$. This arbitrary choice takes advantage of our freedom to simultaneously rescale a, K and r without changing the geometry of spacetime. Unless otherwise mentioned, in these notes the subscript '0' indicates time dependent quantities evaluated at the present time t_0.

†We only consider *ideal* fluids in these lectures. For a discussion of cases when this approximation breaks down, see e.g. [D].

‡Note that this choice is even possible for a fluid of photons, in spite of these latter travelling at the speed of light in any reference frame, as the velocity of the co-moving frame is simply the bulk velocity of the fluid.

to obtain

$$T_{\mu}^{\nu} = (\rho[t] + P[t]/c^2)U_{\mu}U^{\nu} + P[t]\delta_{\mu}^{\nu}\,. \tag{3}$$

Using the normalisation $g^{\mu\nu}U_{\mu}U_{\nu} = U_{\mu}U^{\mu} = U_0U^0 = -c^2$ then yields equation (2). The nature of the fluid then needs to be specified by an equation of state of the form $P = P[\rho]$.

Finally, Einstein's field equations derived in the 3rd year course — with the same convention (+,−,−) for the signs in front of the metric, Riemann tensor and energy-momentum tensor respectively — relate metric and energy momentum tensor, thus describing the interaction between space-time and matter:

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

where $R_{\mu\nu}$ and R, the Ricci tensor and scalar, are functions of $g_{\mu\nu}$ and its first and second derivatives with respect to coordinates, G is the universal gravitational constant and Λ is the so-called cosmological constant. As you have seen in the 3rd year course, using the explicit forms of the FLRW metric and ideal fluid energy-momentum tensor given by equations (1) and (2), these field equations (4) yield the following pair of simple independent equations (R_{00} and $R_{ij}\delta_{ij}$ ($= R_{rr}$) terms respectively):

$$\left(\frac{da}{dt}\right)^2 = \frac{8\pi G a^2[t]}{3}\rho[t] + \frac{\Lambda a^2[t]c^2}{3} - Kc^2 \tag{5}$$

$$\frac{d^2a}{dt^2} = -\frac{4\pi G a[t]}{3}\left(\rho[t] + 3\frac{P[t]}{c^2}\right) + \frac{\Lambda a[t]c^2}{3} \tag{6}$$

called the Friedmann–Lemaître (FL) equations. Together with the equation of state of the fluid, these two equations constitute a closed system, i.e. they completely describe the dynamical evolution of an homogeneous and isotropic Universe.

It is convenient to define the *Hubble parameter* (a.k.a. the expansion rate of the Universe), $H[t] = \dot{a}[t]/a[t]$, where the dot over $a[t]$ stands for the derivative w.r.t. time, and a *critical density*

$$\rho_c[t] = \frac{3H^2[t]}{8\pi G}$$

obtained by setting $\Lambda = 0$ and $K = 0$ in the first FL equation (i.e. equation (5)). Finally, rather than the expansion factor, observers will talk about the cosmological *redshift*, $z \equiv a_0/a[t] - 1$, as it is the quantity which is directly observable as a wavelength shift, $z = \lambda_0/\lambda[t] - 1$. Given these definitions, one can rewrite equation (5) as:

$$\rho_c[z] = \rho[z] + \frac{\Lambda c^2}{8\pi G} - \frac{3Kc^2(1+z)^2}{8\pi G}$$

We can also compare all energy densities of the form $\rho_X[z]c^2$ — where the subscript $X = \{\gamma, M, V, K\}$ indicates contributions from relativistic matter, non-relativistic matter, vacuum energy and curvature respectively — to the critical energy density value $\rho_c[z]c^2$, and write the ratios as dimensionless functions of redshift, $\Omega_X[z] = \rho_X[z]/\rho_c[z]$. We already have $\rho_V[z] = \Lambda c^2/(8\pi G)$ and $\rho_K[z] = -3Kc^2(1+z)^2/(8\pi G)$ in the previous equation, so we just need to separate $\rho[z]$ into a non-relativistic contribution $\rho_M[z]$ and a relativistic one $\rho_\gamma[z]$ to obtain:

$$\Omega_M[z] + \Omega_\gamma[z] + \Omega_V[z] + \Omega_K[z] = 1 \tag{7}$$

Note that only $\Omega_K[z]$ can be negative in this equation. Making the scalings with redshift explicit, i.e. $\rho_M[z] \propto (1+z)^3$, $\rho_\gamma[z] \propto (1+z)^4$, $\rho_V[z] = \text{constant}$, $\rho_K[z] \propto (1+z)^2$ and choosing as a reference the present day values, we can rewrite the equation governing the expansion rate of the Universe in the following convenient form:

$$H^2[z] = H_0^2\left(\Omega_{M,0}(1+z)^3 + \Omega_{\gamma,0}(1+z)^4 + \Omega_{V,0} + \Omega_{K,0}(1+z)^2\right) \tag{8}$$

Current best estimates for these are $\Omega_{M,0} = 0.317$, $\Omega_{V,0} = 0.683$, $\Omega_{K,0} = 0.0$, $\Omega_{\gamma,0} = 9 \times 10^{-5}$, and $H_0 = 67.2$ km/s/Mpc (c.f. Planck collaboration results 2018)[§].

Going back to equation (1), one notices that by defining a new time variable η, called the *conformal* time, as $d\eta \equiv dt/a[t]$, it is possible to factor the expansion of the Universe out of the spacetime interval ds^2, which then reads:

$$ds^2 = a^2[\eta]\left(-c^2 d\eta^2 + \frac{dr^2}{1-Kr^2} + r^2 d\theta^2 + r^2\sin^2\theta\, d\phi^2\right). \tag{9}$$

η is the amount of time it would take a photon to travel from where an observer is located to the largest distance it can observe, provided the Universe stopped expanding. Which leads us to define the important notion of *comoving*, or *particle horizon*, $R_h[t]$, as the maximum distance a massless particle can have travelled since the Big Bang:

$$R_h[t] = c\eta = c\int_0^t \frac{dt'}{a[t']} = R_h[z] = c\int_z^\infty \frac{dz'}{H[z']} \tag{10}$$

Note that since the *proper* distance is equal to the comoving distance times the expansion factor (i.e. these two distances are equal today, at $z = 0$, because by convention $a_0 = 1$), the proper distance to the particle horizon at any redshift is simply $R_h[z]/(1+z)$.

3 Newtonian Evolution of Cosmological Perturbations

So far, we have built our understanding of the expanding Universe on an extraordinary simplifying assumption — that at any given time, it looks the same everywhere and any direction. This assumption is underpinned by observations of the cosmic microwave background radiation which we find to be

[§]In the remainder of these notes we will drop the '0' subscript and write $\Omega_i = \Omega_{i,0}$ i.e. assume these parameters have their present day values, unless the situation is ambiguous. Note that the Ω_V contribution in equation (8) can also be generalised to a more complex form of 'dark energy', with a different equation of state.

uniform to within one part in a hundred thousand. Yet we also know that the Universe cannot be perfectly smooth. The latest observational surveys have identified large scale structures in abundance: galaxies grouped together in clusters, filaments and walls or separated by large empty voids that can span hundreds of millions of light years. Indeed, the mere existence of galaxies, stars and planets indicate that the Universe looks more and more anisotropic as scales become smaller and smaller.

Therefore, if we are to understand the formation of large scale structure, we must go one step beyond the homogeneous Universe and derive the differential equations which govern the growth of small inhomogeneities, i.e. allow for spatial variations in the evolution of energy density, pressure, and gravity. In this section, we will use Newtonian gravity when doing so. However, before we go any further, it is important to make the following disclaimer, to clarify the situation: *it is impossible to study cosmology using Newtonian gravity in a mathematically rigorous and self-consistent way, for all type of fluids and at all perturbation wavelengths*. Having said that, as we will see in the following section, Newtonian gravity does provide us with the qualitative and quantitative behaviour of perturbations that we would find in a proper, general relativistic treatment, at least in the case of massive, non-relativistic particles.

More specifically, let us consider the case where matter is an ideal, non-relativistic self-gravitating fluid. At the epoch where it is the dominant component* the equations governing its dynamics are:

$$\frac{\partial \rho}{\partial t}[\vec{r},t] + \frac{\partial}{\partial \vec{r}}\left(\rho \vec{u}\right)[\vec{r},t] = 0\,, \tag{11}$$

where $\vec{r}$ is the position of a fluid element in physical (proper) coordinates and $\vec{u}$ its corresponding velocity. Equation (11) represents mass (energy) conservation. We also have conservation of momentum (Euler's equations)

$$\frac{\partial \vec{u}}{\partial t}[\vec{r},t] + \left(\vec{u}\frac{\partial}{\partial \vec{r}}\right)\vec{u}[\vec{r},t] = -\frac{\partial}{\partial \vec{r}}\Psi[\vec{r},t] - \frac{1}{\rho}\frac{\partial}{\partial \vec{r}}P[\vec{r},t]\;, \tag{12}$$

and finally, Poisson's equation for the gravitational potential Ψ which writes†:

$$\left(\frac{\partial}{\partial \vec{r}}\right)^2 \Psi[\vec{r},t] = 4\pi G \rho[\vec{r},t]\,. \tag{13}$$

Rigorously speaking, these equations are only valid on scales which are small compared to the particle horizon previously defined, and when the fluid is far away enough from a black hole.

One can more conveniently re-write these equations in a comoving coordinate system defined by $\vec{x} = \vec{r}/a[t]$. In such a system, we have $\vec{u} = \dot{a}\vec{x} + \vec{v}[\vec{x},t]$ where $\dot{a}\vec{x}$ is called the *Hubble flow* as it is the

*This is less restrictive than it may seem, given the values of the cosmological parameters quoted in the previous section. According to these our Universe becomes matter dominated when $\Omega_M[z] = \Omega_\gamma[z]$, i.e at $z \sim 3400$ and stops being so when $\Omega_V[z] = \Omega_M[z]$, i.e. at $z \sim 0.3$. This spans a time window of ~ 10.2 Gyr, to be compared to the age of the Universe, ~ 13.7 Gyr.

†If one wanted to include the cosmological constant Λ it would appear on the RHS of this equation with a negative sign.

velocity contributed by the expansion of the Universe and $\vec{v}[\vec{x},t]$ is the *peculiar* velocity of the fluid element. This leads to[‡]:

$$\frac{\partial\rho}{\partial t}[\vec{x},t] + \frac{3\dot{a}}{a}\rho[\vec{x},t] + \frac{1}{a}\frac{\partial}{\partial\vec{x}}(\rho\vec{v})[\vec{x},t] = 0 \tag{14}$$

$$\frac{\partial\vec{v}}{\partial t}[\vec{x},t] + \frac{1}{a}\left(\vec{v}\frac{\partial}{\partial\vec{x}}\right)\vec{v}[\vec{x},t] + \frac{\dot{a}}{a}\vec{v}[\vec{x},t] = -\frac{1}{a}\frac{\partial}{\partial\vec{x}}\Phi[\vec{x},t] - \frac{1}{a\rho}\frac{\partial}{\partial\vec{x}}P[\vec{x},t] \tag{15}$$

$$\left(\frac{1}{a}\frac{\partial}{\partial\vec{x}}\right)^2\Phi[\vec{x},t] = 4\pi G\rho[\vec{x},t] + 3\frac{\ddot{a}}{a} \tag{16}$$

where we have defined a new gravitational potential $\Phi = \Psi + \frac{1}{2}a\ddot{a}\vec{x}^2$ to cast equation (15) in a simpler form.

In the limit where the fluid is only slightly perturbed compared to the homogeneous and isotropic background (which we know is a valid assumption at the time when the cosmic microwave background is observed, around $z \sim 1100$) we can write $\rho[\vec{x},t] = \bar{\rho}[t](1+\delta[\vec{x},t])$ with the density contrast $\delta[\vec{x},t] \equiv (\rho[\vec{x},t]-\bar{\rho}[t])/\bar{\rho}[t] \ll 1$, $\vec{v}[\vec{x},t] \ll \dot{a}\vec{x}$, $P[\vec{x},t] = \bar{P}[t] + \delta P[\vec{x},t] = \bar{P}[t] + (\partial P/\partial\rho)_S\,\delta\rho[\vec{x},t] = \bar{P}[t] + c_s^2\bar{\rho}\,\delta[\vec{x},t]$ for adiabatic perturbations with associated sound speed c_s, $\Phi[\vec{x},t] \ll \frac{1}{2}a\ddot{a}\vec{x}^2$ and the bar above a variable indicates the homogeneous background average.

Linearising equations (14) (15) and (16), i.e. using the FL equations to cancel the 0^{th} order terms and keeping only the first order terms in all the perturbed variables (which we leave as an exercise), one obtains:

$$\frac{\partial\delta}{\partial t}[\vec{x},t] + \frac{1}{a}\frac{\partial\vec{v}}{\partial\vec{x}}[\vec{x},t] = 0 \tag{17}$$

$$\frac{\partial\vec{v}}{\partial t}[\vec{x},t] + \frac{\dot{a}}{a}\vec{v}[\vec{x},t] = -\frac{1}{a}\frac{\partial}{\partial\vec{x}}\Phi[\vec{x},t] - \frac{c_s^2}{a}\frac{\partial}{\partial\vec{x}}\delta[\vec{x},t] \tag{18}$$

$$\nabla^2_{\vec{x}}\,\Phi[\vec{x},t] = 4\pi G\bar{\rho}a^2\delta[\vec{x},t] \tag{19}$$

Finally, taking the gradient of the perturbed Euler equation (18) and using equations (17) and (19) to get rid of $\nabla_{\vec{x}}\,\vec{v}$ and $\nabla^2_{\vec{x}}\,\Phi$ respectively, yields the following 2^{nd} order linear partial differential equation for δ:

$$\ddot{\delta}[\vec{x},t] + 2H\dot{\delta}[\vec{x},t] - \frac{c_s^2}{a^2}\nabla^2_{\vec{x}}\,\delta[\vec{x},t] = 4\pi G\bar{\rho}\,\delta[\vec{x},t]\,. \tag{20}$$

This equation is more easily studied and solved in Fourier space. Taking its Fourier transform[§],

[‡]Being careful about the variable changes:

$$\left(\frac{\partial}{\partial t}\right)_{\vec{r}} = \left(\frac{\partial}{\partial t}\right)_{\vec{x}} - \frac{\dot{a}}{a}\left(\vec{x}\cdot\frac{\partial}{\partial\vec{x}}\right)_t \qquad \& \qquad \left(\frac{\partial}{\partial\vec{r}}\right)_t = \frac{1}{a}\left(\frac{\partial}{\partial\vec{x}}\right)_t$$

[§]The Fourier transform of δ is defined as

$$\delta[\vec{x},t] = \frac{1}{(2\pi)^3}\int d^3k\,\delta_{\mathbf{k}}\exp(-i\vec{k}\cdot\vec{x})$$

equation (20) becomes

$$\ddot{\delta}_{\mathbf{k}} + 2H\dot{\delta}_{\mathbf{k}} + \left(\frac{c_s^2}{a^2}k^2 - 4\pi G\bar{\rho}\right)\delta_{\mathbf{k}} = 0\,, \tag{21}$$

or

$$\boxed{\delta''_{\mathbf{k}} + \mathcal{H}\delta'_{\mathbf{k}} + \left(c_s^2 k^2 - 4\pi G a^2\bar{\rho}\right)\delta_{\mathbf{k}} = 0} \tag{22}$$

using the conformal variables defined in equation (9), and replacing the partial derivative w.r.t. conformal time, $\partial/\partial\eta \to {}'$. The *conformal* Hubble parameter is thus $\mathcal{H} \equiv a'/a = aH$. This is arguably the most important equation of these Cosmology lectures, as we will come back to it time and again to understand the growth of structures in the Universe, even in the non-linear regime relevant to galaxy formation!

4 Relativistic Cosmological Perturbation Theory

The Newtonian analysis developed in the previous section clearly has limitations. In particular, it fails for perturbations with sizes comparable or larger than the particle horizon, and for fully relativistic fluids at both large and short wavelengths, as the fluid pressure also contributes significantly to the active gravitational mass in this case, unlike in the Newtonian case where only the inertial mass matters.

Unfortunately, the physical interpretation of the results obtained is less transparent in General Relativity (GR) than in the Newtonian theory of gravitation. The main problem arises from the freedom in the choice of coordinates that GR allows to describe perturbations. Indeed, in contrast to the homogeneous and isotropic Universe where the preferable coordinate system is fixed by the symmetry properties there is no obvious choice as soon as we introduce perturbations. Freedom in the coordinate choice, called *gauge freedom*, generally leads to the appearance of fictitious perturbation modes which only reflect the properties of the coordinate system used.

To illustrate this point, let us consider the unperturbed homogeneous and isotropic Universe, where $\rho[\vec{x},t] = \bar{\rho}[t]$, and define a new time coordinate, t' related to t via $t' = t + \delta t[\vec{x},t]$. The mass (energy) density in this new coordinate system, $\rho'[\vec{x},t'] \equiv \bar{\rho}[t[\vec{x},t']]$, when evaluated on the hypersurface $t' = C^{\rm st}$, will in general depend on the spatial coordinate $\vec{x}$. Assuming $\delta t \ll t$ we can then write $\bar{\rho}[t] = \bar{\rho}[t' - \delta t[\vec{x},t]] \simeq \bar{\rho}[t'] - \dot{\bar{\rho}}\,\delta t[\vec{x},t] \equiv \bar{\rho}[t'] + \delta\bar{\rho}[\vec{x},t']$, where the (fictitious) linear perturbation term $\delta\bar{\rho}[\vec{x},t']$ is entirely due to our choice of a new "disturbed" time coordinate!

In the same way that fictitious terms can appear, real perturbation terms may disappear depending on our choice of coordinate system. There exist two (related) ways to overcome this problem. The first one is to calculate only combinations of metric perturbations which are gauge invariant, i.e. that are identical under any change of coordinate system. These combinations are called the *Bardeen* *variables*. Alternatively, we can fix the gauge, and keep track of *all* perturbations, metric *and* matter. It is this approach that we will follow here, adopting the convenient Newtonian gauge, which is uniquely defined for perturbations that decay at spatial infinity[*] and commonly used to study the formation and evolution of large scale structure and cosmic microwave background anisotropies.

We will further restrict ourselves to *scalar* perturbations of the metric and associated energy density inhomogeneities, as these are the most relevant for structure formation, which is the main topic of these lectures[†]. Finally to avoid overly complicated calculations, but also because the measured values of the cosmological parameters given in the previous section indicate that our Universe satisfies this constraint to a high degree of accuracy, we will restrict ourselves to a flat FLRW spacetime.

More specifically, this means that our starting point will be the conformal metric of equation (9), with $K = 0$, cartesian spatial coordinates instead of spherical coordinates as spherical symmetry is broken anyways when perturbations are present. We then perturb this metric using a single scalar component Φ, such that[‡]:

$$ds^2 = a^2[\eta] \left(-\left(1 + \frac{2\Phi}{c^2}\right) c^2 d\eta^2 + \left(1 - \frac{2\Phi}{c^2}\right) \delta_{ij} dx^i dx^j \right). \tag{23}$$

You should be familiar with the Newtonian limit where you set $a[\eta] = 1$ and Φ is the standard Newtonian gravitational potential with the same definition than in the previous section (hence the name of (conformal) Newtonian gauge). As is customary, wherever we use latin indices, they run over the spatial part only, whereas greek indices span the whole of spacetime. We can view this metric as a 0^{th} order part (which is just the usual FLRW metric) and a linearly perturbed part, i.e. as $g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu}$ with $\bar{g}_{\mu\nu}$ the usual metric of the homogeneous background Universe.

From there, we need to work out the linearly perturbed Einstein field equations which, if we ignore the cosmological constant term on the RHS[§], will formally look like:

$$\delta G_{\mu\nu} = -\frac{8\pi G}{c^4} \delta T_{\mu\nu}, \tag{24}$$

---

[*] see [D] for a discussion of what specific gauge choices entail.

[†] Once again we refer to [D] for a discussion of the scalar-vector-tensor decomposition of metric perturbations, but note that, in any case, in the linear regime which we will explore in these lectures, these different perturbation modes evolve independently of one another, so that our results for scalar perturbations will hold regardless of whether vector or tensor modes are present.

[‡] A couple of remarks. Firstly, we use this form because it makes direct contact with the gauge invariant Bardeen variables previously discussed. Second, the scalar perturbations of the time part and spatial part of the metric have a priori no reason to be the same, which is why if you look at Cosmology text books, you will see two different variables Ψ and Φ used. In writing the metric as we do here, we are ignoring any source of anisotropic stress, in which case $\Psi = \Phi$. Since anisotropic stress generally is a second order effect, ignoring it is a reasonable approximation in linear perturbation theory.

[§] It is easy to add it back in at the very end, so we leave it as an exercise, especially since we ignored its presence in the Newtonian theory as well in the previous section and we wish to make the results of the two sections as easily comparable as possible.

where we have written the Einstein tensor $G_{\mu\nu}$ as the usual linear combination of the Ricci tensor $R_{\mu\nu}$ and the metric tensor times the Ricci scalar R:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R\,.$$

Let us start with evaluating the LHS of equation (24). To do so, the first step is to calculate the connection coefficients:

$$\Gamma^{\mu}_{\nu\rho} = \frac{1}{2} g^{\mu\alpha} (\partial_\nu g_{\alpha\rho} + \partial_\rho g_{\alpha\nu} - \partial_\alpha g_{\nu\rho}). \tag{25}$$

We thus write the co- and contra-variant conformal metric of equation (23):

$$g_{\alpha\beta} = a^2[\eta] \begin{pmatrix} -(1+2\Phi/c^2) & 0 \\ 0 & (1-2\Phi/c^2)\delta_{ij} \end{pmatrix}$$

and

$$g^{\alpha\beta} = \frac{1}{a^2[\eta]} \begin{pmatrix} -(1-2\Phi/c^2) & 0 \\ 0 & (1+2\Phi/c^2)\delta^{ij} \end{pmatrix}$$

respectively, where we used the fact that for $\epsilon \ll 1$, $(1+\epsilon)^{-1} \simeq 1-\epsilon$. We then substitute these in the definition of the connection coefficients of equation (25) to obtain[¶]:

$$\begin{aligned}
\Gamma^0_{00} &= \frac{1}{2} g^{00}\, \partial_0 g_{00} = \frac{-1+2\Phi/c^2}{2a^2}\, \partial_0 \left(-a^2(1+2\Phi/c^2)\right) &&= \frac{\mathcal{H}}{c} + \frac{\Phi'}{c^3} + \cancel{\mathcal{O}\left(\frac{\Phi^2}{c^4}\right)} \\
\Gamma^0_{0i} &= \frac{1}{2} g^{00}\, \partial_i g_{00} = \frac{-1+2\Phi/c^2}{2a^2}\, \partial_i \left(-a^2(1+2\Phi/c^2)\right) &&= \frac{\partial_i \Phi}{c^2} \\
\Gamma^i_{00} &= \frac{1}{2} g^{ij} (-\partial_j g_{00}) = \frac{(1+2\Phi/c^2)\,\delta^{ij}}{2a^2}\, \partial_j \left(a^2(1+2\Phi/c^2)\right) &&= \frac{\partial^i \Phi}{c^2} \\
\Gamma^0_{ij} &= \frac{1}{2} g^{00} (-\partial_0 g_{ij}) = \frac{1-2\Phi/c^2}{2a^2}\, \partial_0 \left(a^2(1-2\Phi/c^2)\delta_{ij}\right) &&= \left(\frac{\mathcal{H}}{c} - 4\mathcal{H}\frac{\Phi}{c^3} - \frac{\Phi'}{c^3}\right) \delta_{ij} \\
\Gamma^i_{j0} &= \frac{1}{2} g^{ik} (\partial_0 g_{kj}) = \frac{(1+2\Phi/c^2)\,\delta^{ik}}{2a^2}\, \partial_0 \left(a^2(1-2\Phi/c^2)\delta_{kj}\right) &&= \left(\frac{\mathcal{H}}{c} - \frac{\Phi'}{c^3}\right) \delta^i_j \\
\Gamma^i_{jk} &= \frac{1}{2} g^{il} (\partial_j g_{lk} + \partial_k g_{lj} - \partial_l g_{jk}) &&= -\frac{\partial_j \Phi}{c^2}\delta^i_k - \frac{\partial_k \Phi}{c^2}\delta^i_j + \frac{\partial_l \Phi}{c^2}\delta_{jk}\delta^{il}\,.
\end{aligned}$$

where we have, by analogy with equation (22), used the following notation for the partial derivative w.r.t. conformal time: $c\,\partial_0 \to \partial/\partial\eta \to {}'$. Notice that the expressions for Γ^0_{00}, Γ^0_{ij} and Γ^i_{j0} have a 0^{th} order bit (the $\mathcal{H}$ terms) and a first order bit – the 0^{th} order bit you have already derived when working out the FL equations (5) from scratch in the 3rd year GR & Cosmology course.

[¶] Remember we neglect all quadratic or higher order terms in small quantities – which we denote generically by $\mathcal{O}(\Phi^2/c^4)$ in the first calculation, but drop altogether in the remainder of these lecture notes.

Now we turn to combining the connection coefficients together to obtain the Ricci tensor:

$$R_{\mu\nu} = \partial_\nu \Gamma^\alpha_{\mu\alpha} - \partial_\alpha \Gamma^\alpha_{\mu\nu} + \Gamma^\beta_{\mu\alpha}\Gamma^\alpha_{\nu\beta} - \Gamma^\beta_{\mu\nu}\Gamma^\alpha_{\alpha\beta}\,.$$

Starting with the time component of the tensor

$$R_{00} = \partial_0 \Gamma^\alpha_{0\alpha} - \partial_\alpha \Gamma^\alpha_{00} + \Gamma^\beta_{0\alpha}\Gamma^\alpha_{0\beta} - \Gamma^\beta_{00}\Gamma^\alpha_{\alpha\beta}\,,$$

we notice that if $\alpha = 0$, each pair of terms cancels, so we only need to sum over $\alpha = i$. We are then left with four terms:

$$\begin{aligned}
\partial_0 \Gamma^i_{0i} &= \partial_0(\mathcal{H}/c - \Phi'/c^3)\,\delta^i_i && = 3\frac{\mathcal{H}'}{c^2} - 3\frac{\Phi''}{c^4} \\
\partial_i \Gamma^i_{00} &= \partial_i\partial^i\Phi/c^2 && = \frac{\nabla^2\Phi}{c^2} \\
\Gamma^\beta_{0i}\Gamma^i_{0\beta} &= \cancel{\partial_i\Phi\partial^i\Phi/c^4} + (\mathcal{H}/c - \Phi'/c^3)\delta^j_i(\mathcal{H}/c - \Phi'/c^3)\delta^i_j && = 3\frac{\mathcal{H}^2}{c^2} - 6\mathcal{H}\frac{\Phi'}{c^4} \\
\Gamma^\beta_{00}\Gamma^i_{i\beta} &= (\mathcal{H}/c + \Phi'/c^3)(\mathcal{H}/c - \Phi'/c^3)\,\delta^i_i && = 3\frac{\mathcal{H}^2}{c^2}\,.
\end{aligned}$$

So we find, keeping only up to first order terms in the perturbed potential, Φ:

$$R_{00} = 3\frac{\mathcal{H}'}{c^2} - 3\frac{\Phi''}{c^4} - \frac{\nabla^2\Phi}{c^2} - 6\mathcal{H}\frac{\Phi'}{c^4}\,.$$

Note that, once again, it splits into a 0^{th} order bit (first term) and a linear bit (last 3 terms). Repeating this procedure for the mixed components of the Ricci tensor, one obtains:

$$R_{0i} = \partial_i \Gamma^\alpha_{0\alpha} - \partial_\alpha \Gamma^\alpha_{0i} + \Gamma^\beta_{0\alpha}\Gamma^\alpha_{i\beta} - \Gamma^\beta_{0i}\Gamma^\alpha_{\alpha\beta}\,,$$

and splitting the index α into time and spatial indices, we now have 8 terms to calculate:

$$\begin{aligned}
\partial_i \Gamma^0_{00} &= \partial_i(\mathcal{H}/c + \Phi'/c^3) && = \frac{\partial_i\Phi'}{c^3} \\
\partial_0 \Gamma^0_{0i} &= \partial_0\partial_i\Phi/c^2 && = \frac{\partial_i\Phi'}{c^3} \\
\Gamma^\beta_{00}\Gamma^0_{i\beta} &= (\mathcal{H}/c + \Phi'/c^3)\partial_i\Phi/c^2 + \partial^j\Phi(\mathcal{H}/c - 4\mathcal{H}\Phi/c^3 - \Phi'/c^3)\delta_{ij}/c^2 && = 2\mathcal{H}\frac{\partial_i\Phi}{c^3} \\
\Gamma^\beta_{0i}\Gamma^0_{0\beta} &= \partial_i\Phi(\mathcal{H}/c + \Phi'/c^3)/c^2 + (\mathcal{H}/c - \Phi'/c^3)\partial_i\Phi/c^2 && = 2\mathcal{H}\frac{\partial_i\Phi}{c^3} \\
\partial_i \Gamma^k_{0k} &= \partial_i(\mathcal{H}/c - \Phi'/c^3)\delta^k_k && = -\,3\frac{\partial_i\Phi'}{c^3} \\
\partial_k \Gamma^k_{0i} &= \partial_k(\mathcal{H}/c - \Phi'/c^3)\delta^k_i && = -\,\frac{\partial_i\Phi'}{c^3} \\
\Gamma^\beta_{0k}\Gamma^k_{i\beta} &= \partial_i\Phi(\mathcal{H}/c - \Phi'/c^3)/c^2 + (\mathcal{H}/c - \Phi'/c^3)(-\delta^k_k\,\partial_i\Phi - \delta^j_i\,\partial_j\Phi + \delta_{ik}\delta^{kl}\,\partial_l\Phi)/c^2 && = -\,2\mathcal{H}\frac{\partial_i\Phi}{c^3} \\
\Gamma^\beta_{0i}\Gamma^k_{k\beta} &= \partial_i\Phi(\mathcal{H}/c - \Phi'/c^3)\delta^k_k/c^2 + (\mathcal{H}/c - \Phi'/c^3)\delta^j_i(-\delta^k_j\partial_k\Phi - \delta^k_k\,\partial_j\Phi + \delta^l_j\,\partial_l\Phi)/c^2 && = 0\,.
\end{aligned}$$

So we simply have:

$$R_{0i} = -2\mathcal{H}\frac{\partial_i\Phi}{c^3} - 2\frac{\partial_i\Phi'}{c^3}\,.$$

Finally, for the spatial components of the Ricci tensor:

$$R_{ij} = \partial_j\Gamma^\alpha_{i\alpha} - \partial_\alpha\Gamma^\alpha_{ij} + \Gamma^\beta_{i\alpha}\Gamma^\alpha_{j\beta} - \Gamma^\beta_{ij}\Gamma^\alpha_{\alpha\beta}\,,$$

the same splitting of the index α yields the following 8 terms:

$$\begin{aligned}
\partial_j\Gamma^0_{i0} &= \partial_j\partial_i\Phi/c^2 &&= \frac{\partial_i\partial_j\Phi}{c^2}\\
\partial_0\Gamma^0_{ij} &= \partial_0(\mathcal{H}/c - 4\mathcal{H}\Phi/c^3 - \Phi'/c^3)\,\delta_{ij} &&= \left(\frac{\mathcal{H}'}{c^2} - 4\mathcal{H}'\frac{\Phi}{c^4} - 4\mathcal{H}\frac{\Phi'}{c^4} - \frac{\Phi''}{c^4}\right)\delta_{ij}\\
\Gamma^\beta_{i0}\Gamma^0_{j\beta} &= (\mathcal{H}/c - \Phi'/c^3)(\mathcal{H}/c - 4\mathcal{H}\Phi/c^3 - \Phi'/c^3)\,\delta_{ij} &&= \left(\frac{\mathcal{H}^2}{c^2} - 4\mathcal{H}^2\frac{\Phi}{c^4} - 2\mathcal{H}\frac{\Phi'}{c^4}\right)\delta_{ij}\\
\Gamma^\beta_{ij}\Gamma^0_{0\beta} &= (\mathcal{H}/c - 4\mathcal{H}\Phi/c^3 - \Phi'/c^3)(\mathcal{H}/c + \Phi'/c^3)\,\delta_{ij} &&= \left(\frac{\mathcal{H}^2}{c^2} - 4\mathcal{H}^2\frac{\Phi}{c^4}\right)\delta_{ij}\\
\partial_j\Gamma^k_{ik} &= \partial_j(-\delta^k_k\,\partial_i\Phi - \delta^k_i\,\partial_k\Phi + \delta_{ik}\delta^{kl}\,\partial_l\Phi)/c^2 &&= -\,3\frac{\partial_i\partial_j\Phi}{c^2}\\
\partial_k\Gamma^k_{ij} &= \partial_k(-\delta^k_i\,\partial_j\Phi - \delta^k_j\,\partial_i\Phi + \delta_{ji}\delta^{kl}\,\partial_l\Phi)/c^2 &&= -\,2\frac{\partial_i\partial_j\Phi}{c^2} + \frac{\nabla^2\Phi}{c^2}\delta_{ij}\\
\Gamma^\beta_{ik}\Gamma^k_{j\beta} &= (\mathcal{H}/c - 4\mathcal{H}\Phi/c^3 - \Phi'/c^3)(\mathcal{H}/c - \Phi'/c^3)\,\delta_{ij} &&= \left(\frac{\mathcal{H}^2}{c^2} - 4\mathcal{H}^2\frac{\Phi}{c^4} - 2\mathcal{H}\frac{\Phi'}{c^4}\right)\delta_{ij}\\
\Gamma^\beta_{ij}\Gamma^k_{k\beta} &= (\mathcal{H}/c - 4\mathcal{H}\Phi/c^3 - \Phi'/c^3)(\mathcal{H}/c - \Phi'/c^3)\,\delta^k_k\,\delta_{ij} &&= 3\left(\frac{\mathcal{H}^2}{c^2} - 4\mathcal{H}^2\frac{\Phi}{c^4} - 2\mathcal{H}\frac{\Phi'}{c^4}\right)\delta_{ij}
\end{aligned}$$

which we collect to write:

$$R_{ij} = \left(-\frac{\mathcal{H}'}{c^2} - 2\frac{\mathcal{H}^2}{c^2} + 8\mathcal{H}^2\frac{\Phi}{c^4} + 4\mathcal{H}'\frac{\Phi}{c^4} + 6\mathcal{H}\frac{\Phi'}{c^4} + \frac{\Phi''}{c^4} - \frac{\nabla^2\Phi}{c^2}\right)\delta_{ij}\,.$$

The Ricci scalar is by definition:

$$R \;\equiv\; g^{\mu\nu}R_{\mu\nu} = g^{00}R_{00} + \cancel{2g^{0i}R_{0i}} + g^{ij}R_{ij} = g^{00}R_{00} + g^{ij}R_{ij}\,,$$

as $g^{0i} = 0$. Using the expressions previously derived for R_{00} and R_{ij}, we get:

$$\begin{aligned}
Ra^2 &= -(1 - 2\Phi/c^2)(3\mathcal{H}'/c^2 - 3\Phi''/c^4 - \nabla^2\Phi/c^2 - 6\mathcal{H}\Phi'/c^4) + (1 + 2\Phi/c^2)\delta^{ij}\\
&\quad \times\left(-\mathcal{H}'/c^2 - 2\mathcal{H}^2/c^2 + 8\mathcal{H}^2\Phi/c^4 + 4\mathcal{H}'\Phi/c^4 + 6\mathcal{H}\Phi'/c^4 + \Phi''/c^4 - \nabla^2\Phi/c^2\right)\delta_{ij}\\
&= -6\frac{\mathcal{H}'}{c^2} - 6\frac{\mathcal{H}^2}{c^2} + 6\frac{\Phi''}{c^4} - 2\frac{\nabla^2\Phi}{c^2} + 24\mathcal{H}\frac{\Phi'}{c^4} + 12\mathcal{H}'\frac{\Phi}{c^4} + 12\mathcal{H}^2\frac{\Phi}{c^4}\,.
\end{aligned}$$

We now have all the ingredients to work out the different components of the Einstein tensor. They are, starting with the time component:

$$\begin{aligned}
G_{00} &\equiv R_{00} - \frac{1}{2}g_{00}R \\
&= 3\mathcal{H}'/c^2 - 3\Phi''/c^4 - \nabla^2\Phi/c^2 - 6\mathcal{H}\Phi'/c^4 + \frac{1}{2}(1+2\Phi/c^2)(-6\mathcal{H}'/c^2 - 6\mathcal{H}^2/c^2 + 6\Phi''/c^4 \\
&\qquad -2\nabla^2\Phi/c^2 + 24\mathcal{H}\Phi'/c^4 + 12\mathcal{H}'\Phi/c^4 + 12\mathcal{H}^2\Phi/c^4) \\
&= -3\frac{\mathcal{H}^2}{c^2} - 2\frac{\nabla^2\Phi}{c^2} + 6\mathcal{H}\frac{\Phi'}{c^4}\,.
\end{aligned} \tag{26}$$

Once again, the first term on the RHS is the 0^{th} order term you have derived in the 3^{rd} year GR & Cosmology course, which yields (simply by dividing by -3 and remembering that we assumed a flat Universe, i.e. $K = 0$, without a cosmological constant, i.e. $\Lambda = 0$) the LHS term of the FL equation (5). The mixed space-time component is simply R_{0i}, since $g_{0i} = 0$, so:

$$G_{0i} = -2\mathcal{H}\frac{\partial_i\Phi}{c^3} - 2\frac{\partial_i\Phi'}{c^3}\,, \tag{27}$$

with no 0^{th} order term, as expected from the FL equations for the homogeneous Universe! Finally, the spatial component reads:

$$\begin{aligned}
G_{ij} &\equiv R_{ij} - \frac{1}{2}g_{ij}R \\
&= (-\mathcal{H}'/c^2 + 4\mathcal{H}'\Phi/c^4 + 6\mathcal{H}\Phi'/c^4 + \Phi''/c^4 - \nabla^2\Phi/c^2 - 2\mathcal{H}^2/c^2 + 8\mathcal{H}^2\Phi/c^4)\,\delta_{ij} \\
&\quad -\frac{1}{2}(1-2\Phi/c^2)(-6\mathcal{H}'/c^2 - 6\mathcal{H}^2/c^2 + 6\Phi''/c^4 - 2\nabla^2\Phi/c^2 + 24\mathcal{H}\Phi'/c^4 + 12\mathcal{H}'\Phi/c^4 \\
&\qquad +12\mathcal{H}^2\Phi/c^4)\,\delta_{ij} \\
&= \left(2\frac{\mathcal{H}'}{c^2} + \frac{\mathcal{H}^2}{c^2} - 8\mathcal{H}'\frac{\Phi}{c^4} - 6\mathcal{H}\frac{\Phi'}{c^4} - 2\frac{\Phi''}{c^4} - 4\mathcal{H}^2\frac{\Phi}{c^4}\right)\delta_{ij}\,.
\end{aligned} \tag{28}$$

with the 0^{th} order term providing two of the terms in the FL equation (6).

We now turn to the RHS of the Einstein field equation (24). From equation (3), we have:

$$\bar{T}^\mu_\nu = (\bar{\rho} + \bar{P}/c^2)\bar{U}^\mu\bar{U}_\nu + \bar{P}\delta^\mu_\nu$$

where the bars, as usual, denote averaged, homogeneous background Universe quantities. Therefore, perturbing the stress-energy tensor w.r.t. the homogeneous background yields:

$$\begin{aligned}
T^\mu_\nu &= \bar{T}^\mu_\nu + \delta T^\mu_\nu \\
&= \left(\bar{\rho} + \delta\rho + (\bar{P} + \delta P)/c^2\right)\bar{U}^\mu\bar{U}_\nu + (\bar{\rho} + \bar{P}/c^2)(\delta U^\mu\bar{U}_\nu + \bar{U}^\mu\delta U_\nu) + (\bar{P} + \delta P)\delta^\mu_\nu
\end{aligned} \tag{29}$$

where we have neglected any anisotropic stress perturbation, in keeping with our approximation for the metric tensor perturbation. We can then write:

$$g_{\mu\nu}U^\mu U^\nu = (\bar{g}_{\mu\nu} + \delta g_{\mu\nu})(\bar{U}^\mu + \delta U^\mu)(\bar{U}^\nu + \delta U^\nu) = -c^2$$

and since $\bar{g}_{\mu\nu}\bar{U}^\mu\bar{U}^\nu = -c^2$ as well, this previous equation simplifies to (keeping only linear terms):

$$\delta g_{\mu\nu}\bar{U}^\mu\bar{U}^\nu + 2\bar{g}_{\mu\nu}\delta U^\mu\bar{U}^\nu = 0\,.$$

As $\bar{U}_\mu = -ac\,\delta^0_\mu$ and $\bar{U}^\mu = c/a\,\delta^\mu_0$, this equation then yields:

$$-2\Phi - 2a^2\delta U^0 \times \frac{c}{a} = 0\,,\ \text{i.e. } \delta U^0 = -\frac{\Phi}{ac}\,.$$

Defining $\delta U^i \equiv v^i/a$, with $v^i \equiv dx^i/d\eta$ the *coordinate velocity*, we can write:

$$U^\mu = \frac{c}{a}(1 - \Phi/c^2, v^i/c)\,,$$

and since $U_\mu = g_{\mu\nu}U^\nu$, we obtain $U_0 = g_{00}U^0 + g_{0i}U^i = -a^2(1 + 2\Phi/c^2) \times c/a(1 - \Phi/c^2)$, i.e. $U_0 = -ac(1 + \Phi/c^2)$ and $U_i = g_{i0}U^0 + g_{ij}U^j = a^2(1 - 2\Phi/c^2)\delta_{ij} \times v^j/a = av_i$, so that

$$U_\mu = -ac(1 + \Phi/c^2, -v_i/c)\,.$$

Injecting these expressions for the 4-velocities into equation (29), we find

$$\begin{aligned}
T^0_0 &= -(\bar{\rho}+\delta\rho)c^2 - \cancel{(\bar{P}+\delta P)} + \cancel{(\bar{\rho}+\bar{P}/c^2)(\Phi-\Phi)} + \cancel{(\bar{P}+\delta P)\delta^0_0} \\
&= -(\bar{\rho}+\delta\rho)c^2 \\
T^i_0 &= \cancel{(\bar{\rho}+\delta\rho+(\bar{P}+\bar{P})/c^2)\times 0} + (\bar{\rho}+\bar{P}/c^2)(-v^i c + 0) + \cancel{(\bar{P}+\delta P)\delta^i_0} \\
&= -(\bar{\rho}+\bar{P}/c^2)\,cv^i \\
T^0_j &= \cancel{(\bar{\rho}+\delta\rho+(\bar{P}+\delta P)/c^2)\times 0} + (\bar{\rho}+\bar{P}/c^2)(0 + v_j c) + \cancel{(\bar{P}+\delta P)\delta^0_j} \\
&= (\bar{\rho}+\bar{P}/c^2)\,cv_j \\
T^i_j &= \cancel{(\bar{\rho}+\delta\rho+(\bar{P}+\delta P)/c^2)\times 0} + \cancel{(\bar{\rho}+\bar{P}/c^2)(0+0)} + (\bar{P}+\delta P)\delta^i_j \\
&= (\bar{P}+\delta P)\delta^i_j\,.
\end{aligned} \tag{30}$$

An important thing to note is that if we define the *momentum density* as $q^i \equiv (\bar{\rho} + \bar{P}/c^2)v^i$, and consider the stress energy tensor as the sum of its various component, that is to say we write

$$T_{\mu\nu} = \sum_X T_{\mu\nu\,X}\,,$$

where X stands for photons, baryons, dark matter, etc ..., then the perturbations in density, pressure and momentum density of these components simply add, i.e.:

$$\delta\rho = \sum_X \delta\rho_X \ ; \ \ \delta P = \sum_X \delta P_X \ ; \ \ q^i = \sum_X q^i_X \, .$$

We can now plug all these expressions for the perturbed stress energy tensor components into equations ((26), (27) and (28)) to obtain the linearised Einstein field equations:

$$\begin{aligned} G_{00} &= -\frac{8\pi G}{c^4} g_{0\mu} T^\mu_0 \\ -3\frac{\mathcal{H}^2}{c^2} - 2\frac{\nabla^2\Phi}{c^2} + 6\mathcal{H}\frac{\Phi'}{c^4} &= -\frac{8\pi G a^2}{c^2}\bar{\rho}\left(1 + \delta + 2\frac{\Phi}{c^2}\right) \end{aligned}$$

where $\delta \equiv \delta\rho/\bar{\rho}$ as usual, the 0^{th} order terms on both sides yield the FL equation (5) in conformal coordinates for a flat spacetime and without a cosmological constant term:

$$\mathcal{H}^2 = \frac{8\pi G a^2}{3}\bar{\rho}\, , \tag{31}$$

and the first order ones give the 00 component of the perturbed equations (24) we are looking for:

$$\nabla^2\Phi = 4\pi G a^2 \bar{\rho}\delta + \frac{8\pi G a^2}{c^2}\bar{\rho}\Phi + 3\mathcal{H}\frac{\Phi'}{c^2}\, .$$

Upon substitution of equation (31) in the second RHS term, this first order equation simplifies to:

$$\boxed{\nabla^2\Phi = 4\pi G a^2 \bar{\rho}\delta + 3\frac{\mathcal{H}}{c^2}(\mathcal{H}\Phi + \Phi')} \tag{32}$$

which, for $c \to \infty$ gives us the Newton-Poisson equation (19). We can also see that if we assume $\Phi' \sim \mathcal{H}\Phi$ and recall that $\mathcal{H} \sim 1/\eta$, the relativistic correction term on the RHS is small in front of the LHS term when $kc\eta \gg 1$[‖], that is when the wavelengths of the perturbations are much smaller than the particle horizon.

The "relativistic" Newton-Poisson equation (32) can be written in another form if we consider the mixed term linearised Einstein equation:

$$\begin{aligned} G_{0i} &= -\frac{8\pi G}{c^4} g_{0\mu} T^\mu_i \\ -\frac{2}{c^3}\partial_i(\mathcal{H}\Phi + \Phi') &= \frac{8\pi G a^2}{c^3}(\bar{\rho} + \bar{P}/c^2)v_i \end{aligned}$$

‖Obtained by taking the Fourier transform of the equation.

and integrate it assuming perturbations vanish at spatial infinity and that $v_i = \partial_i \Theta$ i.e. only the scalar part of the perturbation matters**, which gives

$$\mathcal{H}\Phi + \Phi' = -4\pi G a^2 (\bar{\rho} + \bar{P}/c^2)\Theta\,.$$

Injecting this expression in the second term on the RHS of equation (32) then leads to the desired form:

$$\nabla^2 \Phi = 4\pi G a^2 \bar{\rho} \left(\delta - 3\frac{\mathcal{H}}{c^2} \left(1 + \frac{\bar{P}}{\bar{\rho} c^2} \right) \Theta \right) \equiv 4\pi G a^2 \bar{\rho} \Delta \tag{33}$$

which also is a Poisson equation but one where the gravitational field is sourced by the gauge invariant scalar quantity Δ rather than the energy density contrast δ only. Note that, once again, when $c \to \infty$, we do recover the Newtonian limit given by equation (19) as $\Delta \to \delta$.

Finally the last equation comes from the spatial component of the linearised Einstein equation:

$$\begin{aligned} G_{ij} &= -\frac{8\pi G}{c^4} g_{i\mu} T^{\mu}_{j} \\ 2\frac{\mathcal{H}'}{c^2} + \frac{\mathcal{H}^2}{c^2} - 8\mathcal{H}'\frac{\Phi}{c^4} - 6\mathcal{H}\frac{\Phi'}{c^4} - 2\frac{\Phi''}{c^4} - 4\mathcal{H}^2\frac{\Phi}{c^4} &= -\frac{8\pi G a^2}{c^4} \left(\bar{P} + \delta P - 2\frac{\Phi}{c^2}\bar{P} \right)\,. \end{aligned}$$

In a similar way to the time component, this equation yields a 0^{th} order part:

$$\mathcal{H}' = -\frac{4\pi G a^2}{3} \left(\bar{\rho} + 3\frac{\bar{P}}{c^2} \right)\,, \tag{34}$$

where we have used equation (31) to replace the $\mathcal{H}^2$ term on the LHS. This simply is the second FL equation (6) in conformal coordinates and without the cosmological constant term on the RHS. The linear part is thus the only new equation we have derived:

$$8\mathcal{H}'\frac{\Phi}{c^4} + 6\mathcal{H}\frac{\Phi'}{c^4} + 2\frac{\Phi''}{c^4} + 4\mathcal{H}^2\frac{\Phi}{c^4} = \frac{8\pi G a^2}{c^4} \left(\delta P - 2\frac{\Phi}{c^2}\bar{P} \right)\,.$$

Once again, we can use the 0^{th} order equation to eliminate the $\bar{P}$ term on the RHS, and introduce the adiabatic sound speed of the fluid $c_s^2 = \delta P/\delta\rho$, to simplify this equation a bit:

$$\boxed{\Phi'' + 3\mathcal{H}\Phi' + (2\mathcal{H}' + \mathcal{H}^2)\Phi = 4\pi G a^2 c_s^2\, \bar{\rho}\, \delta} \tag{35}$$

**More specifically, this means we apply the scalar-vector-tensor decomposition to the perturbed velocity (analogous to a Helmoltz decomposition of the vector into a curl free and a divergence free part, and writing the curl free component as the gradient of a scalar – sometimes called the velocity potential), thus writing $v_i = \partial_i \Theta + \hat{v}_i$, and neglecting the vector part, $\hat{v}_i$, in keeping with the assumptions we made for the metric perturbations.

Although in principle the Einstein field equations (along with the equation of state of the fluid) provide a complete description of the dynamics of the system, to better make the link with the Newtonian perturbation theory developed in the previous section, we now derive the relativistic equivalent of energy conservation and the Euler equations. These two sets of equations are related through the Bianchi identities as you have seen in the third year course. We thus write the conservation of the energy-momentum tensor, $\nabla_\mu T^\mu_\nu = 0$, i.e.:

$$\partial_\mu T^\mu_\nu + \Gamma^\mu_{\mu\alpha} T^\alpha_\nu - \Gamma^\alpha_{\mu\nu} T^\mu_\alpha = 0\,. \tag{36}$$

Let's consider the $\nu = 0$ component of this equation first. It writes:

$$\partial_0 T^0_0 + \partial_i T^i_0 + \cancel{\Gamma^0_{00} T^0_0} + \Gamma^i_{i0} T^0_0 + \cancel{\Gamma^0_{0i} T^i_0} + + \Gamma^j_{ji} T^i_0 - \cancel{\Gamma^0_{00} T^0_0} - \Gamma^i_{00} T^0_i - \cancel{\Gamma^0_{i0} T^i_0} - \Gamma^j_{i0} T^i_j = 0\,,$$

and injecting the expressions previously obtained for the linear connection coefficients and for the perturbed stress energy tensor, we get:

$$\begin{aligned} & -\partial_0(\bar\rho + \delta\rho)c^2 - (\bar\rho + \bar P/c^2)c\,\partial_i v^i - (\mathcal{H}/c - \Phi'/c^3)\,\delta^i_i(\bar\rho + \delta\rho)c^2 \\ & +3\,\partial_i\Phi(\bar\rho + \bar P/c^2)\,v^i/c - \partial^i\Phi(\bar\rho + \bar P/c^2)\,v_i/c - (\mathcal{H}/c - \Phi'/c^3)\,\delta^j_i(\bar P + \delta P)\delta^i_j \;=\; 0\,, \end{aligned}$$

and keeping only up to the linear terms, this simplifies to:

$$\bar\rho\,' + \bar\rho\,'\delta + \bar\rho\,\delta' + (\bar\rho + \bar P/c^2)\frac{\partial\vec v}{\partial\vec x} + 3(\mathcal{H} - \Phi'/c^2)(\bar\rho + \bar P/c^2) + 3\mathcal{H}(\delta\rho + \delta P/c^2) = 0\,.$$

Splitting this equation in a 0^{th} order and a first order one, as usual, yields:

$$\bar\rho\,' + 3\mathcal{H}\bar\rho\left(1 + \frac{\bar P}{\bar\rho c^2}\right) = 0\,,$$

which is simply the conservation of energy in the homogenous background[††], and

$$\boxed{\delta' + \left(1 + \frac{\bar P}{\bar\rho c^2}\right)\left(\frac{\partial\vec v}{\partial\vec x} - 3\frac{\Phi'}{c^2}\right) + 3\mathcal{H}\left(\frac{c_s^2}{c^2} - \frac{\bar P}{\bar\rho c^2}\right)\delta = 0} \tag{37}$$

where we have used the 0^{th} order equation to replace $\bar\rho\,'$. Note that this gives equation (17)[‡‡], in the Newtonian limit where $c \to \infty$. Once again general relativistic correction terms are small for $kc\eta \gg 1$, i.e. sub-horizon scale perturbations.

We now turn to the $\nu = i$ component of the stress-energy tensor conservation equation (36). This reads:

$$\partial_0 T^0_i + \partial_j T^j_i + \Gamma^0_{00} T^0_i + \Gamma^j_{j0} T^0_i + \Gamma^0_{0j} T^j_i + \Gamma^j_{jk} T^k_i - \Gamma^0_{0i} T^0_0 - \Gamma^j_{0i} T^0_j - \Gamma^0_{ji} T^j_0 - \Gamma^j_{ki} T^k_j = 0\,,$$

[††]Obtained by differentiating equation (5) w.r.t. time (setting $\Lambda = 0$ first), and using equation (6) to get rid of the second order derivative of the expansion factor w.r.t. time.

[‡‡]In conformal coordinates.

which yields, upon injecting the linear expressions previously calculated for all these terms,

$$\begin{aligned} \partial_0\left((\bar\rho+\bar P/c^2)\,cv_i\right)+\partial_j(\bar P+\delta P)\delta_i^j+(\mathcal{H}/c+\Phi'/c^3)(\bar\rho+\bar P/c^2)\,cv_i \\ +3(\mathcal{H}/c-\Phi'/c^3)(\bar\rho+\bar P/c^2)\,cv_i+\partial_j\Phi(\bar P+\delta P)\delta_i^j/c^2-3\,\partial_k\Phi(\bar P+\delta P)\delta_i^k/c^2 \\ +\partial_i\Phi(\bar\rho+\delta\rho)-(\mathcal{H}/c-\Phi'/c^3)\delta_i^j(\bar\rho+\bar P/c^2)\,cv_j \\ +(\mathcal{H}/c-4\mathcal{H}\Phi/c^3-\Phi'/c^3)\delta_{ji}(\bar\rho+\bar P/c^2)\,cv^j \\ +(\delta_i^j\,\partial_k\Phi+\delta_k^j\,\partial_i\Phi-\delta_{ki}\delta^{jl}\,\partial_l\Phi)(\bar P+\delta P)\delta_j^k/c^2 &= 0\,. \end{aligned}$$

Keeping only the linear terms in small quantities, we can simplify the previous expression as:

$$(\bar\rho+\bar P/c^2)\,v_i'+\partial_i(\delta P)+\left(\bar\rho\,'+\bar P\,'/c^2+4\mathcal{H}(\bar\rho+\bar P/c^2)\right)v_i+(\bar\rho+\bar P/c^2)\,\partial_i\Phi = 0\,,$$

and using once more the 0^{th} order conservation of energy equation to re-arrange the $\bar\rho\,'$ terms, we arrive at the following, more familiar form:

$$\boxed{\vec v\,'+\mathcal{H}\vec v-3\mathcal{H}\frac{c_s^2}{c^2}\vec v=-\frac{\partial\Phi}{\partial\vec x}-\frac{c_s^2}{1+\bar P/(\bar\rho c^2)}\frac{\partial\delta}{\partial\vec x}} \tag{38}$$

where we have taken advantage that for adiabatic fluctuations $\bar P\,'/\bar\rho\,'=\delta P/\delta\rho=c_s^2$. When $c\to\infty$ this gives the Newtonian equation (18), but more specifically, the GR correction terms are small when the sound speed of the fluid is small compared to the speed of light.

As in the Newtonian case, taking the spatial gradient of the perturbed "relativistic" Euler equation (38) and using equations (37) and (33) to get rid of $\nabla\vec v$ and $\nabla^2\,\Phi$ respectively, yields a (somewhat complicated and hence not generally useful) 2^{nd} order linear partial differential equation for δ. It does, however, simplify to yield equation (22) in the Newtonian limit, i.e. when the fluid considered is non-relativistic and the scale of the perturbation is small compared to the particle horizon, approximations generally accurate enough to describe the evolution of large scale structure. For instance, the largest super clusters of galaxies that we observe today on the sky, like the local Laniakea or the more distant Saraswati super cluster, have sizes $\sim\,150$ Mpc, which are much smaller than the current particle horizon ($R_h(z=0)\sim 14.4$ Gpc).

5 The Evolution of Large Scale Structure

Let us therefore go back and have a look at equation (22). We can identify a number of features in the evolution of $\delta_{\mathbf{k}}$ without actually solving this equation. For a start, it is quite clearly the equation of a damped harmonic oscillator with time dependent damping coefficient and spring constant. The damping (second term on the LHS) is due to the expansion of the Universe and will tend to suppress growth. The spring constant (third term in between brackets on the LHS) will change sign depending on whether k is large or small. If the positive part of the spring constant, $c_s^2k^2$, dominates then we should expect oscillatory behaviour in the form of acoustic waves in the fluid. If the negative term, $4\pi G a^2 \bar{\rho}$ dominates, then the evolution will be unstable and we should expect $\delta_{\mathbf{k}}$ to grow. The *physical* wavelength λ_J*, that defines the transition between these two behaviours is given by:

$$\lambda_J = \frac{2\pi}{k_J} = c_s \left(\frac{\pi}{G\bar{\rho}} \right)^{\frac{1}{2}}$$

and is known as the *Jeans wavelength*. For $\lambda > \lambda_J$ gravitational collapse dominates and perturbations grow. For $\lambda < \lambda_J$ pressure wins and perturbations do not grow. We can have a rough idea of how a given system of particles will behave if we note that $c_s^2 \sim k_B T/m$ where k_B is the Boltzmann constant, T is the temperature and m is the mass of the individual particles. We can then rewrite the Jean length as $\lambda_J = \sqrt{\pi k_B T/(Gm\bar{\rho})}$. It is clear that a system which is hot and/or made up of light particles will have a large λ_J; a cold system with heavy particles will have a small λ_J. It is now useful to find solutions for specific scenarios.

5.1 Pressureless Fluid in the Matter Dominated Era

In this situation we have $c_s^2 \simeq 0$ and hence $\lambda \gg \lambda_J$. We can therefore discard the term which depends on pressure in equation (22). As we further restrict ourselves to the matter dominated era, we then neglect all contributions to the energy densities of the Universe in front of that of matter, so that $\bar{\rho} \simeq \bar{\rho}_M$. Working in conformal coordinates, equation (7) thus reduces to $\Omega_M[\eta] = 1$, i.e. $\bar{\rho}_M[\eta] = \rho_c[\eta] = 3\mathcal{H}^2[\eta]/(8\pi a^2 G)$. In that case, the fluctuations in the pressureless fluid are also those of the component driving the expansion of the background, and equation (22) simplifies to:

$$\delta_{\mathbf{k}}'' + \mathcal{H}\delta_{\mathbf{k}}' - \frac{3}{2}\mathcal{H}^2\delta_{\mathbf{k}} = 0\,.$$

Finally, equation (8) yields† $a \propto \eta^2$, so that $\mathcal{H} = 2/\eta$ and trying power law solutions of the form $\delta_{\mathbf{k}} \propto \eta^\alpha$, one easily finds:

$$\boxed{\delta_{\mathbf{k}} = C_1\eta^2 + C_2\eta^{-3}} \tag{39}$$

where C_1 and C_2 are constants of integration. Rewriting both terms of this solution as a function of the expansion factor a, one gets $\delta_{\mathbf{k}} \propto a$ and $\delta_{\mathbf{k}} \propto a^{-3/2}$ respectively. As a increases with time when the Universe expands, one readily deduces that the second term decays and becomes sub-dominant very fast. We are then left with the first term which continues to grow under the effect of gravity as the Universe expands: for this reason it is generally called the *growing mode*. Note that we could have performed the same calculation using the 'normal' time t rather than the conformal time η, i.e. we could have solved equation (21) instead of (22). We would then have obtained $\delta_{\mathbf{k}} \propto t^{2/3}$ and $\delta_{\mathbf{k}} \propto t^{-1}$ for the growing and decaying mode respectively.

*as opposed to the *comoving* wavelength, $\lambda_J^c = \lambda_J/a$.

†when recast in conformal coordinates, and after separating variables a and η and integrating.

5.2 Pressureless Matter in the Λ Dominated Era

Proceeding the same way as previously, we still have $c_s^2 \simeq 0$ and $\lambda \gg \lambda_J$ and can therefore discard the term which depends on pressure in equation (22). However, since we are in the Λ dominated era, we now neglect all contributions to the energy densities of the Universe in front of that of the cosmological constant, so that $\bar{\rho} \simeq \bar{\rho}_V$. Working in conformal coordinates once again, equation (7) reduces this time to $\Omega_V[\eta] = 1$, i.e. $\bar{\rho}_V[\eta] = \rho_c[\eta] = 3\mathcal{H}^2[\eta]/(8\pi a^2 G)$. As there exists by definition no fluctuation in the cosmological constant which could source perturbations in the gravitational potential[‡], equation (22) simplifies to:

$$\delta''_{\mathbf{k}} + \mathcal{H}\delta'_{\mathbf{k}} - 4\pi G a^2 \bar{\rho}_M \delta_{\mathbf{k}} = \delta''_{\mathbf{k}} + \mathcal{H}\delta'_{\mathbf{k}} \cancel{- \frac{3}{2}\mathcal{H}^2 \frac{\bar{\rho}_M}{\bar{\rho}_V}} \delta_{\mathbf{k}} = 0\,.$$

where the last term has been crossed out because it is a second order effect: the pressureless fluid is *not* driving the expansion of the background Universe and as such its average density $\bar{\rho}_M \ll \bar{\rho}_V \simeq \bar{\rho}$. Going back to equation (8) yields $a \propto 1/\eta$, so that $\mathcal{H} = -1/\eta$ and trying once again power law solutions of the form $\delta_{\mathbf{k}} \propto \eta^\alpha$, one easily finds:

$$\boxed{\delta_{\mathbf{k}} = C_1 + C_2 \eta^2} \tag{40}$$

where C_1 and C_2 are constants of integration as usual. Rewriting both terms of this solution as a function of the expansion factor a, one gets $\delta_{\mathbf{k}} = C^{\text{st}}$ and $\delta_{\mathbf{k}} \propto a^{-2}$ for each of these. Again, since a increases with time when the Universe expands, one readily deduces that the second term decays and becomes sub-dominant very fast. The first term, however, does *not* grow under the effect of gravity as was the case in the matter dominated era: it stays constant. Note that had we performed the same calculation using the 'normal' time t rather than the conformal time η, we would then have obtained $\delta_{\mathbf{k}} = C^{\text{st}}$ and $\delta_{\mathbf{k}} \propto \exp(-2\sqrt{\Lambda/3}\,t)$ for the (non) growing and decaying mode respectively. Clearly a faster rate of expansion[§] inhibits the growth of fluctuations.

5.3 Relativistic Fluid in the Radiation Dominated Era

The characteristic properties of the fluid will also affect how it evolves under gravity in an expanding universe. Consider the growth of perturbations of a relativistic fluid with pressure. An example of this scenario is of radiation interacting strongly with baryons before recombination. During this epoch, baryons are dissociated into protons and electrons which interact with photons through Thomson scattering. The net result is that radiation behaves as a gravitating fluid with pressure $\bar{P} = 1/3\,\bar{\rho}c^2$, i.e. the adiabatic sound speed is $c_s = 1/\sqrt{3}\,c$. Although we could extend equations (17), (18) and

[‡] i.e. a term of the form $4\pi G \bar{\rho}_V \delta_V$ with $\delta_V \neq 0$ which would then appear on the RHS of the perturbed Newton-Poisson equation (19).

[§] It is easy to show from equation (8) that the expansion factor, a, grows exponentially with time as $\exp(\sqrt{\Lambda/3}\,t)$ in the Λ dominated era, rather than as a power law.

(19) so that they describe a relativistic fluid[¶], and use these modified equations to derive the equivalent of equation (22), it is much more satisfactory to go to the full general relativistic equations that we have worked quite hard to derive in the previous section. We will thus follow this route and leave the exploration of the 'modified Newtonian path' as an exercise. In the radiation dominated era, we have $\bar\rho \simeq \bar\rho_\gamma$, and switching to conformal coordinates, equation (7) reduces to $\Omega_\gamma[\eta] = 1$, i.e. $\bar\rho_\gamma[\eta] = \rho_c[\eta] = 3\mathcal{H}^2[\eta]/(8\pi a^2 G)$. Equation (8) further yields $a \propto \eta$, so that $\mathcal{H} = 1/\eta$ and finally equations (31) and (34) give $\mathcal{H}^2 = -\mathcal{H}' = 1/\eta^2$ for the background Universe in that case.

Our starting point for the perturbations will be Einstein's field equation (35):

$$\Phi'' + 3\mathcal{H}\Phi' + (2\mathcal{H}' + \mathcal{H}^2)\Phi - 4\pi G a^2 \bar\rho\, \delta\, c_s^2 = 0\,.$$

Re-writing the relativistic Poisson equation (32) as:

$$4\pi G a^2 \bar\rho\, \delta = \nabla^2\Phi - 3\frac{\mathcal{H}}{c^2}(\mathcal{H}\Phi + \Phi')\,, \tag{41}$$

we can use it to get rid of the δ term which appears in the previous field equation to get, upon taking the Fourier transform and using $\mathcal{H}^2 = -\mathcal{H}'$,

$$\Phi''_{\mathbf{k}} + 4\mathcal{H}\Phi'_{\mathbf{k}} + c_s^2 k^2 \Phi_{\mathbf{k}} = 0\,.$$

Finally, using the expansion factor scaling with conformal time we can simplify this equation a bit further as:

$$\Phi''_{\mathbf{k}} + \frac{4}{\eta}\Phi'_{\mathbf{k}} + \frac{c^2 k^2}{3}\Phi_{\mathbf{k}} = 0\,.$$

The solutions of this differential equation are:

$$\Phi_{\mathbf{k}} = A_1\,\frac{\sqrt{3}}{kc\eta}\, j_1\left[\frac{kc\eta}{\sqrt{3}}\right] + A_2\,\frac{\sqrt{3}}{kc\eta}\, y_1\left[\frac{kc\eta}{\sqrt{3}}\right],$$

where A_1 and A_2 are constants, $j_1[x] = \sin[x]/x^2 - \cos[x]/x$ and $y_1[x] = -\cos[x]/x^2 - \sin[x]/x$ are spherical Bessel functions of the first and second kind respectively. Considering the 'early times' limit, i.e. $\eta \to 0$ we deduce that the constant A_2 must vanish, as otherwise the gravitational potential would blow up since $\lim_{x\to 0} y_1[x]/x = -\infty$. Note that, contrary to what would have happened had we followed the modified Newtonian road, this solution is valid for *any* wavenumber k, even those

[¶]Basically by treating the fluid as a special relativistic fluid which sources the Newtonian gravitational potential. The perturbed equations then become:

$$\begin{aligned}\dot\delta + 4\nabla\vec{v}/(3a) &= 0\\ \dot{\vec{v}} + H\vec{v} &= -\nabla\Phi/a - c^2\nabla\delta/(3a)\\ \nabla^2\Phi &= 8\pi G\bar\rho a^2\delta\end{aligned}$$

where both the extra third in the continuity equation and the factor 2 in the Poisson equation come from the contribution of the relativistic fluid pressure (respectively $\bar P$ and $3\bar P$ for these two equations) to the energy density.

corresponding to super-horizon wavelengths[‖]. However, in keeping with the previous sub-sections, if we only consider sub-horizon modes, i.e. $kc\eta \gg 1$, we get the solution

$$\Phi_{\mathbf{k}} \simeq -\frac{3A_1}{(kc\eta)^2} \cos\left[\frac{kc\eta}{\sqrt{3}}\right],$$

which means that the gravitational potential on small scales oscillates at a frequency $kc/\sqrt{3}$ with an amplitude that decays with (conformal) time as η^{-2}, or in terms of expansion factor as a^{-2}.

Turning back to the energy density fluctuations in the relativistic fluid, equation (41) can be conveniently re-arranged into:

$$\delta_{\mathbf{k}} = \frac{2}{3\mathcal{H}^2}\left(-k^2\Phi_{\mathbf{k}} - 3\frac{\mathcal{H}}{c^2}(\mathcal{H}\Phi_{\mathbf{k}} + \Phi'_{\mathbf{k}})\right) = -\left(\frac{2}{3}k^2\eta^2 + \frac{2}{c^2}\right)\Phi_{\mathbf{k}} - \frac{2\eta}{c^2}\Phi'_{\mathbf{k}}\,. \tag{42}$$

For sub-horizon energy density perturbations, i.e. such that $kc\eta \gg 1$, the first RHS term dominates over the other two. Plugging in the solution previously obtained for the gravitational potential, we then easily derive that:

$$\boxed{\delta_{\mathbf{k}} \simeq -\frac{2}{3}k^2\eta^2\Phi_{\mathbf{k}} = \frac{2A_1}{c^2}\cos\left[\frac{kc\eta}{\sqrt{3}}\right]} \tag{43}$$

In other words, sub-horizon energy density perturbations in the relativistic fluid oscillate around $\delta_{\mathbf{k}} = 0$ at a frequency $kc/\sqrt{3}$ with constant amplitude. Note that we could also have plugged in the full solution for the gravitational potential in terms of Bessel functions and considered all the terms in equation (42), which would allow us to conclude that for super-horizon perturbations with $kc\eta \ll 1$, both $\Phi_{\mathbf{k}}$ and $\delta_{\mathbf{k}}$ are constant[**]. Finally let us stress that this way of deriving the evolution of perturbations can also be applied to the matter dominated Universe that we studied earlier on. We leave it as an exercise to show that in that case, one recovers the evolution given by equation (39) for $\delta_{\mathbf{k}}$ and that the gravitational potential $\Phi_{\mathbf{k}} = A_1 + A_2\eta^{-5}$ stays $\sim$ constant.

5.4 Pressureless Matter in the Radiation Dominated Era

Now that we have obtained the evolution of energy density perturbations in the dominant relativistic fluid, we can look at what happens to the sub-dominant pressureless matter perturbations ($c_s^2 \simeq 0$) in

[‖]Even though, unlike for the sub-horizon modes, the results are not easy to interpret physically, as we discuss further below.

[**]This is counter-intuitive, as one expects that for large enough wavelengths, fluid pressure will not be able to counter-balance gravity. Our intuition is correct here, super-horizon wavelength fluctuations always grow, and in particular, they grow $\propto \eta^2$ in the radiation dominated era as can be shown (exercise) by, most simply, going the 'modified Newtonian path' to derive an equation equivalent to (22) and neglecting the pressure of the relativistic fluid in front of the gravitational term. This is a warning: as we previously mentioned the Newtonian gauge is only uniquely defined for perturbations which decay at spatial infinity, so it is best to restrict its use to sub-horizon perturbations for which there is no gauge ambiguity.

the radiation dominated era, just as we did in the Λ dominated era. A notable example is cold dark matter density fluctuations during the radiation dominated era: they are decoupled from the baryon + radiation fluid. In this situation, the pressureless matter will not only play a sub-dominant role in the expansion of the Universe, i.e. $\bar{\rho}_\gamma[\eta] \simeq \bar{\rho} = 3\mathcal{H}^2[\eta]/(8\pi a^2 G)$ will still hold, but it will also make a negligible contribution to the general relativistic version of the Newton-Poisson equation (32). In other words, on sub-horizon scales ($kc\eta \gg 1$) we will have:

$$\nabla^2\Phi = 4\pi G(\bar{\rho}_\gamma\delta_\gamma + \bar{\rho}_M\delta_M) = 4\pi G\bar{\rho}_\gamma\left(\delta_\gamma + \cancel{\frac{\bar{\rho}_M}{\bar{\rho}_\gamma}\delta_M}\right) \simeq 4\pi G\bar{\rho}_\gamma\delta_\gamma\,,$$

where we have crossed out the second order term. We have already calculated the evolution of δ_γ in the previous sub-section and found that on sub-horizon scales these perturbations do not grow, but oscillate rapidly around zero, so that their time averaged amplitude, $\langle\delta_{\mathbf{k},\gamma}\rangle_\eta$, is nil. Thus, in this case, equation (22) can be simplified to:

$$\delta''_{\mathbf{k}} + \mathcal{H}\delta'_{\mathbf{k}} = \delta''_{\mathbf{k}} + \frac{1}{\eta}\delta'_{\mathbf{k}} = 0\,.$$

The solution to this equation is:

$$\boxed{\delta_{\mathbf{k}} = C_1 \ln\eta + C_2} \tag{44}$$

or if we recast things in terms of the expansion factor, $\delta_{\mathbf{k}} = C_1 \ln a + C_2$. In other words, the interesting point here is that even though perturbations in the dominant tightly coupled relativistic baryon+radiation fluid do *not* grow, perturbations in the sub-dominant pressureless matter component do. As it turns out, this logarithmically slow growth is of key importance as it means that once recombination is finished and baryons have decoupled from photons, baryonic density fluctuation growth will be sped up as these baryons will fall in deeper dark matter potential wells. It is this very acceleration of the baryonic collapse which allows high density contrast structures such as the galaxies we observe today to form rapidly enough.

5.5 Damping of Cosmological Perturbations

There exists situations which do not fit well into the formalism we have been using so far. These occur when there is imperfect coupling between different fluid components or when the system cannot be described purely in terms of a fluid and one must revert to a more microscopic description in terms of the particle distribution function.

5.5.1 The Boltzmann Equation

How to treat the matter distribution without assuming that it is a fluid? Consider the phase space distribution function $f(\vec{x}, \vec{p})$ where $\vec{x}$ is the position and $\vec{p}$ is the momentum of individual particles.

The evolution of f is described by the Boltzmann equation:

$$\left(\frac{\partial}{\partial t} + \frac{\partial \vec{x}}{\partial t} \cdot \nabla_{\vec{x}} + \frac{\partial \vec{p}}{\partial t} \cdot \nabla_{\vec{p}}\right) f = \mathcal{Q}(f, f)$$

where the collision term on the RHS is generally a complicated multiple integral over momentum space involving differential cross-sections. Its general relativistic generalisation writes:

$$\left(p^\alpha \frac{\partial}{\partial x^\alpha} - \Gamma^\alpha_{\mu\nu} p^\mu p^\nu \frac{\partial}{\partial p^\alpha}\right) f = \left(\frac{\partial f}{\partial t}\right)_{\text{coll}} \tag{45}$$

5.5.2 Damping During Recombination

Before recombination, photons and charged particles are tightly coupled together, i.e. the mean free path of the photons is negligible and this radiation and baryon combination can be treated as a single relativistic fluid as we did in section 5.3. However as the Universe expands, the density and temperature of such a fluid drop, and photons end up not possessing enough energy to prevent protons from capturing free electrons and form neutral hydrogen. Their mean free path, given by $\lambda_{\text{mfp}} \approx 1/(\sigma_T n_e)$ where n_e is the number density of free electrons and σ_T the Thompson scattering cross-section, thus progressively approaches infinity and the Universe becomes transparent to radiation. This entails that during this recombination epoch, for the brief period of time when it transitions from negligible to infinite, the mean free path of the photons becomes finite.

In other words, as the Universe recombines (around $z_\star \simeq 1100$), n_e plummets over a redshift change of $\Delta z \simeq 80$ and because photons and electrons are not perfectly coupled at that time, the photons will be able to random walk out of overdensities as they scatter off free electrons. In doing so they will drag matter from over dense to under dense regions and therefore damp out perturbations on scales smaller than the characteristic random walk scale. In order to derive the damping scale (known as the *Silk damping* scale, named after former Oxford Savilian professor Joe Silk) rigorously, one needs to use the Boltzmann equation previously given (eq (45) see e.g. [P]), but it can be derived approximately using the following route.

Consider a photon that suffers $N = c\Delta t/\lambda_{\text{mfp}}[t]$ collisions during recombination. Between two collisions it travels a comoving distance $\lambda_{\text{mfp}}/a[t]$. Since this is a random walk, this photon acquires a mean square comoving displacement $(\Delta x)^2 = N(\lambda_{\text{mfp}}/a)^2 = c\lambda_{\text{mfp}}[t]\Delta t/a^2[t]$, so that the total mean square comoving distance it travels until the time of decoupling is:

$$x^2 \equiv \int_0^{t_\star} \frac{c\,dt}{\sigma_T n_e[t] a^2[t]} = \int_0^{\eta_\star} \frac{c\,d\eta}{\sigma_T n_e[\eta] a[\eta]} = \frac{c\eta_\star}{5\sigma_T n_e[\eta_\star] a[\eta_\star]}$$

where we have used the fact that recombination occurs during the matter dominated epoch, so that $a[\eta] = a[\eta_\star](\eta/\eta_\star)^2$ and $n_e[\eta] = n_e[\eta_\star](a[\eta]/a[\eta_\star])^{-3} = n_e[\eta_\star](\eta/\eta_\star)^{-6}$. This yields a comoving Silk damping scale,

$$\boxed{\lambda_S^c \equiv x = \sqrt{\frac{c\eta_\star}{5a[\eta_\star]\sigma_T n_e[\eta_\star]}}} \tag{46}$$

using the conformal time at decoupling, $\eta_\star$. Plugging in numbers for $c\eta_\star = 262\,\mathrm{Mpc}$, $a[\eta_\star] \simeq 1/1101 = 9.08 \times 10^{-4}$, $\sigma_T = 6.65 \times 10^{-25}\,\mathrm{cm}^2$ and $n_e[\eta_\star] = \Omega_{b,0}\,\rho_{c,0}/m_p \times 1101^3 \simeq 332\,\mathrm{cm}^{-3}$ (where the subscript b indicates, 'normal' matter, i.e. baryons) yields $\lambda_S^c \simeq 9.2\,\mathrm{Mpc}$. Fluctuations below this scale will be washed out.

5.5.3 Massive Neutrinos During Any Era

Massive neutrinos cannot be described as a fluid, as they interact very weakly with each other and their evolution must be studied using the Boltzmann equation (45). However, as for the Silk damping, it is possible to derive an approximation of their effect. On very large scales ($> R_h(t_{eq})$, the size of the particle horizon at matter-radiation equality) they will tend to cluster just like matter and radiation but on small scales, they will *free-stream* i.e. move relativistically from one region of space to another. Neutrinos decouple when the temperature of the Universe drops below $k_B T_D \sim 1$ MeV, around $a[t_D] \simeq 3 \times 10^{-10}$. This energy is very much larger than their rest mass energy, which means they still are ultra-relativistic when they decouple. As their 3-momentum decays linearly with redshift, we can define a time, t_{nr}, when they transition from being relativistic particles to being non-relativistic, i.e. $k_B T_{nr} \sim m_\nu c^2$. The exact time thus depends on the neutrino mass, m_ν, but it will happen before matter-radiation equality if neutrinos are more massive than ~ 1 eV, so that the proper length travelled by the massive neutrinos, can be written as:

$$\begin{aligned}
\lambda_{FS}[t_0] &= a[t_0]\int_0^{t_0} \frac{v[t]}{a[t]}dt \simeq a[\eta_0]c\left(\int_0^{\eta_{nr}} d\eta + \int_{\eta_{nr}}^{\eta_{eq}} \frac{\eta_{nr}}{\eta}d\eta + \int_{\eta_{eq}}^{\eta_0} \frac{\eta_{nr}\eta_{eq}}{\eta^2}d\eta\right) \\
&= c\eta_{nr}\left(1 + \ln[\eta_{eq}/\eta_{nr}] + \eta_{eq}(1/\eta_{eq} - 1/\eta_0)\right) \\
&\simeq c\eta_{nr}\left(2 + \ln[\eta_{eq}/\eta_{nr}]\right)
\end{aligned}$$

where we have used the fact that $a[\eta] = a[\eta_{nr}](\eta/\eta_{nr})$ during the radiation dominated epoch, and $a[\eta] = a[\eta_{eq}](\eta/\eta_{eq})^2$ during the matter dominated epoch. This will lead an overall damping effect, wiping out structure on small scales as the neutrino velocity is larger than the escape velocity of the gravitational potential fluctuations on these scales. Plugging in numbers for $c\eta_{nr} \sim c\eta_{eq} \simeq 10\,\mathrm{Mpc}$ (i.e. assuming 1 eV neutrinos which become non-relativistic exactly at matter-equality, even though truly speaking they only do so a bit later) yields a free-streaming length roughly given by:

$$\boxed{\lambda_{FS} \equiv \lambda_{FS}[t_0] \simeq 20\left(\frac{1\,\mathrm{eV}}{m_\nu}\right)\mathrm{Mpc}} \tag{47}$$

Given the current upper limits on the neutrino mass (of the order of a fraction of an eV; Planck collaboration) this corresponds to scales much larger than galaxy clusters.

6 Building models of large scale structure

Thus far we have studied the evolution of structure in a variety of scenarios and we should, by now, have a qualitative understanding of how cosmological perturbations may evolve. We now need complete this analysis by defining the initial conditions, i.e. the seeds of structure, then characterizing how perturbations of different length scales evolve and and finally identifying how we should ultimately characterize large scale structure today.

Over the decades there have been a plethora of proposals for the initial conditions of structure formation. One set of possibilities is that the Universe started off in a quasi-chaotic initial state and that the thermal initial state smoothed out the large inhomogeneities leaving a residue of fluctuations when then evolved to form structure. Clearly this is not a viable proposal unless we severely modify the nature of the Universe at those early time- as we saw in the previous section, structure on very large scales (larger than the Jeans wavelength) will tend to grow under the force of gravity. Furthermore, there is a limit to how much the Universe could homogenize set by causality so it is in fact physically impossible to implement such a simple idea.

We do, however, have a proposal that tends to smooth out the Universe and that changes the causal structure of space-time. Inflation will take a microscopic patch of the Universe which is in thermal equilibrium and is well within the Jeans wavelength at that time, and expands it to macroscopic, cosmological proportions. In doing so, Inflation solves the problem of how to homogenize but also provides a mechanism for seeding structure. We expect that, due to the quantum nature of space-time and matter, that the Universe will be riven by quantum fluctuations on microscopic scales. A period of Inflationary expansion will amplify and stretch these quantum fluctuations to macroscopic scales which will be many times larger than the cosmological horizon by the time Inflation ends. As the Universe resumes its normal expansion in the radiation era, the fluctuations will seed structure in the cosmological fluid which will then evolve in the manner described in the previous sections.

The form of the intial conditions arising from Inflation have a deeply appealing feature: they will correspond to a *random field* which has a *scale invariant* gravitational potential. In this context, a random field is a three-dimensional function which can be generated through some random process; this should not come as a surprise given that the source of the fluctuations is a quantum process. And if you think about what we are trying to do, and look at the structure of the sky, you will realized that there must be an element of randomness. Our theory won't tell us if a cluster of galaxies, or a filament of galaxies or more generally an overdensity or underdensity is going to be at an exact position in space. All we can talk about is how much more probable structures of a given size are going to be relative to others. For example, we may expect to see more structure of 1 Mpc than of 100 Mpc, but we don't know exactly where they will be. Hence we talk about our density contrast, δ, or gravitational potentials being a random fields for which we can calculate their statistical properties.

We characterize a random field in much the same way we would characterize any other random process. For example we will have that the density contrast, δ satisfies

$$\langle\delta(\vec{x})\rangle = 0$$

and that we can characterize its variance in terms of a *correlation function*, $\xi(r)$ through

$$\xi(|\vec{x} - \vec{x}'|) \equiv \langle \delta(\vec{x})\delta(\vec{x}')\rangle$$

or alternatively in terms of its *power spectrum*

$$\boxed{P(k) \equiv \langle |\delta_{\vec{k}}|^2 \rangle} \tag{48}$$

By defining $\xi(r)$ or $P(k)$ we can characterize the statistical properties of the random field* It is often useful to consider the dimensionless version of the power spectrum, the *mass variance* which is given by

$$\Delta^2(k) = \frac{k^3 P(k)}{4\pi}$$

We now need to understand what is meant by scale invariance. Let us define the average gravitational potential in a ball of radius R to be

$$\Phi(R) = \frac{1}{V_R} \int_{V_R} d^3x \Phi(\vec{x})$$

where V_R is the volume of the ball. We can define the variance of Φ on that scale to simply be

$$\sigma_R^2(\Phi) = \langle \Phi^2(R) \rangle$$

where $\langle \cdots \rangle$ is an ensemble average, i.e. an average over many possible configurations of Φ. A scale invariant spectrum corresponds to a variance which is independent of R, i.e. $\sigma_R^2(\Phi) \propto \text{constant}$. It turns out that we can relate $\sigma_R^2(\Phi)$, to $\delta(t, \vec{k})$ through the Newton-Poisson equation. Indeed we have that

$$\sigma_R^2(\Phi) \simeq \frac{k^3}{4\pi} \langle |\Phi(t, \vec{k})|^2 \rangle \propto \frac{k^3}{k^4} \langle |\delta(t, \vec{k})|^2 \rangle \ \text{ with } \ k = \frac{2\pi}{R}$$

If it is scale invariant we then have that the power spectrum of the density fluctuations at initial time t_i, $P_i(\vec{k}) \equiv \langle |\delta(t_i, \vec{k})|^2 \rangle$ has the form

$$P_i(k) \propto k$$

In practice, choosing scale invariant initial conditions for the density field corresponds to picking the amplitude of the density field† to be given by $|\delta(t, \vec{k})| \propto k^{1/2}$.

Having chosen a set of initial conditions we can predict what the large scale structure of the Universe for different sets of assumptions. We will do so for three model universes: a universe where matter is solely of baryons and known as the Baryonic Dark Matter (BDM) universe ; a universe which also contains pressureless dark matter and is known as the Cold Dark Matter (CDM) universe; and a universe in which matter is dominated by massive neutrinos and is known as the Hot Dark Matter (HDM) universe.

*This is only strictly true of the random process is Gaussian. For non-Gaussian processes one has to go further and characterize such quantities as $\langle \delta(\vec{x}_1)\delta(\vec{x}_1)\delta(\vec{x}_1)\rangle$ and higher order products. It turns out the Inflation predicts that the random fields are, to a very good approximation, Gaussian.

†This is not strictly true, otherwise we would have $\langle \delta \rangle \neq 0$ but given that we are not interested in $\delta_{\vec{k}}$ today, but in $P(k)$, this prescription will suit us

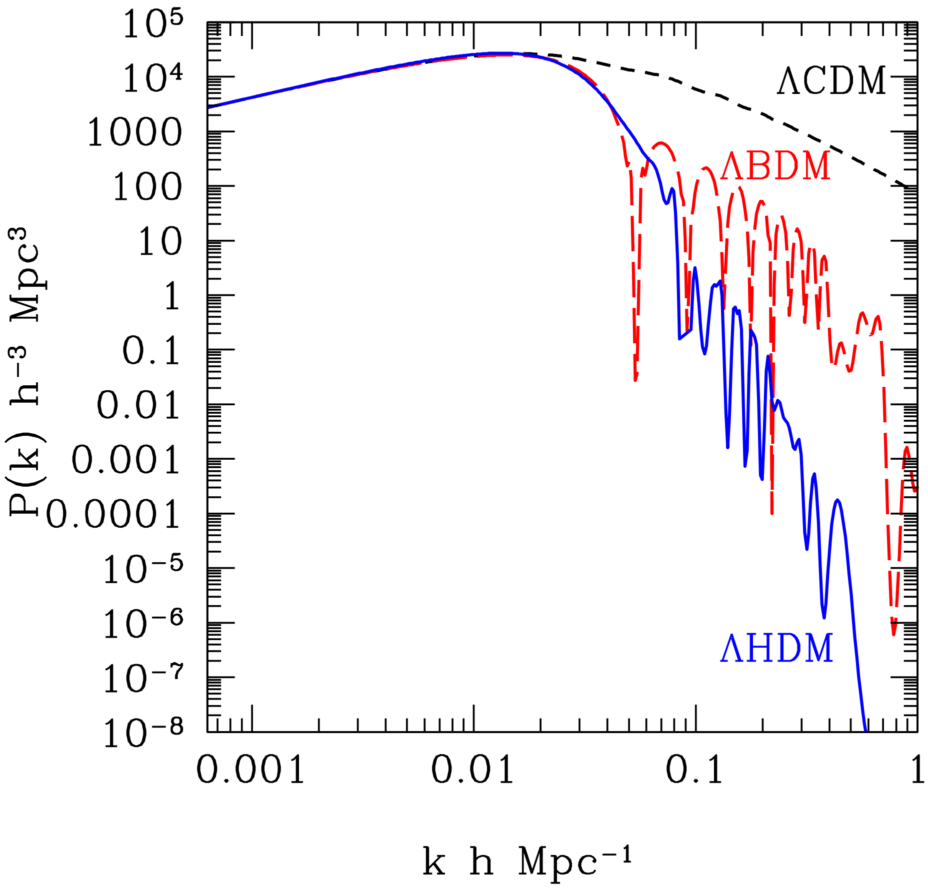

Figure 1: The power spectrum for three different models of the Universe: ΛHDM, i.e. a universe with baryons, massive neutrinos and a cosmological constant; ΛBDM, i.e. a universe with only baryons and a cosmological constant; ΛCDM, i.e. a universe with baryons, cold dark matter and a cosmological constant. Note that the first two models have much less power on small scales (large k) than the last one.

6.1 A Baryonic Dark Matter Universe

Consider the simplest scenario we can imagine with the tools we have been give: a flat Universe consisting solely of photons, atoms (i.e. Baryons) and a cosmological constant. There is a limit on how much of the Universe can be made of baryons: the abundance of light elements restricts $\Omega_B h^2 \simeq 0.024$. With our current constraints on the Hubble constant, this means that fractional energy density in baryons must be around 5% and given that we are considering a Universe with $\Omega = 1$ we must have $\Omega_\Lambda = 0.95$. On very large scales (i.e. on scales larger than the sound horizon at equality between matter and radiation, i.e. $> c\eta_{eq}/\sqrt{3}$) perturbations will grow until they reach the Λ dominated era, after which they will be constant. On scales below the sound horizon at matter-radiation equality, perturbations will oscillate acoustically and once the Universe recombines, they will grow again until the freeze in during the Λ era. Hence we expect a series of peaks and troughs on intermediate to small scales. On very small scales, i.e. on scales which are smaller than the Silk length at recombination, perturbations will experience severe exponential damping.

As we have discussed in the previous section, the initial fluctuation power spectrum given by inflation is $P(k) \propto k$. Therefore we expect fluctuations on very large scales ($kc\eta_{eq}\sqrt{3} < 1$), to retain this power spectrum shape long after recombination ($z_\star \simeq 1100$) and almost to the present day, $z = 0$, as they will enter the sound horizon at a later time, during the matter dominated era, where they can continue to grow independently of size. Only the very largest of them, which enter the horizon during the late Λ dominated era will have their growth suppressed. On the other hand, smaller scale fluctuations (with $kc\eta_{eq}\sqrt{3} > 1$, corresponding to galaxies and galaxy clusters) entered the horizon *before* matter-radiation equality (as we saw for galaxies, around $z \sim 10^6 \gg z_{eq} \simeq 3300$). From the time η at which they entered and until η_{eq}, they will have stopped growing. As their size was $k = \sqrt{3}/(c\eta)$ when they entered and the fluctuations outside the horizon continue to grow like $\delta_{\mathbf{k}} \propto \eta^2$ until the end of the radiation dominated era, the amount of growth suppression will be $\propto k^{-2}$, so that the shape of the power spectrum on these scales after recombination will become $P(k) \propto k^{-3}$. Putting all these pieces together, we find that the mass variance for this theory is roughly described[‡] by:

$$\begin{array}{ll} \Delta^2(k) \propto k^4 & \text{if } kc\eta_{eq}/\sqrt{3} < 1 \\ \Delta^2(k) \propto \cos^2[kc\eta_{eq}/\sqrt{3}] \times \exp[-2(k/k_S)^2] & \text{if } kc\eta_{eq}/\sqrt{3} > 1 \end{array} \tag{49}$$

We can also consider a Universe in which, for some reason, we have underestimated the density of baryons. If we assume that the majority of baryons are dark (in the form of dark nuggets of matters, brown dwarfs or even black holes) we can choose $\Omega_B = 0.25$ and $\Omega_\Lambda = 0.75$. In Figure 1 we plot the power spectrum of this theory which we can call the Baryonic Dark Matter scenario (or BDM for short) and we can see the gross qualitative features we identified above. In particular, you will note the transition from an increasing power spectrum, $P(k) \propto k$, on large scales, to a decreasing one, $P(k) \propto k^{-3}$, on small scales, with power peaking around $k \sim k_{eq} = \sqrt{3}/(c\eta_{eq})$.

6.2 A Cold Dark Matter Universe

An interesting scenario arises if we consider a Universe in which, once again, we have radiation, baryons and Λ but now add a component of pressureless, non-relativistic matter that does not interact with the radiation. We shall call this the Cold Dark Matter scenario (or CDM) and has a qualitatively different behaviour to the BDM case. There are effectively two regimes that will define the shape of the power spectrum. On very large scales, i.e. scales such that $k\eta \ll 1$ for all times before equality between radiation and matter, the density contrast will grow as η^2 until it reaches the Λ dominated era. For scales that cross this threshold, i.e. such that $k\eta_{eq} < 1$, the density contrast will have its evolution suppressed to logarithmic growth; this suppression in growth will last between the time the

[‡]We have made simplifying assumptions ...

wavelength of the perturbation has the same size as the sound horizon and the transition from radiation to matter domination. During the matter era, perturbations will grow again until Λ domination sets in.

A rough estimate of the mass variance, $\Delta^2(k)$ gives us

$$\boxed{\begin{array}{ll}\Delta^2(k) \propto k^4 & \text{if } k\eta_{eq} < 1 \\ \Delta^2(k) \propto (\ln(k))^2 & \text{if } k\eta_{eq} > 1\end{array}} \tag{50}$$

The overall shape can be clearly seen in Figure 1.

6.3 A Hot Dark Matter Universe

There is yet another simple model we can consider. If we now replace the pressureless matter in the CDM model by light massive neutrinos, we will have an altogether different cosmology. The motivation is clear: we know that neutrinos exist and there is even tentative evidence that they may have a mass. As we saw in the previous section, neutrinos will not evolve as a fluid and will free stream while they are relativistic, exponentially damping all perturbations on small scales. The neutrinos are weakly interacting, dark (i.e. they don't interact strongly with light) and move relativistically so can be considered a "hot" component of the Universe. For these reasons, a Universe in which neutrinos make up the bulk of the energy density today is called the Hot Dark Matter scenario (or HDM). The mass variance can be roughly approximated to

$$\boxed{\Delta^2(k) \propto k^4 \exp(-k\lambda_{FS})} \tag{51}$$

and is plotted in Figure 1.

7 The non-linear regime: towards a theory of galaxy formation

In the previous section on the comparison of large scale structure to observations, we have modelled galaxies as an ensemble of mathematical point masses which trace the underlying continuous density field. However, using the perturbative framework we have developed in the previous lectures, it is possible to go one step further and derive the one point statistics (a.k.a. the *mass function* and its time evolution) of the collapsed dark matter structures which host galaxies, if not of the galaxies themselves. This is the purpose of this section.

7.1 Growth of an isolated 'top hat' dark matter density perturbation

Although our reasoning holds in the most general of cases, we will only consider the case of an Einstein – de Sitter universe (i.e $\Omega_M[z] = 1$, $\Omega_V[z] = 0$, $\Omega_\gamma[z] = 0$, $\Omega_K[z] = 0$). The reason for such a choice is purely practical: all the integrals we need to calculate have an explicit closed analytic form in this specific universe. Also, as previously mentioned, this is not as restrictive as it seems because we know that our Universe has been in the matter dominated era for most of its lifetime.

Let us consider a small spherical[*] perturbation of radius r_i in an otherwise homogeneous dark matter (collisionless fluid) density field at some early time t_i (equivalent to high redshift z_i). Let us assume that its peculiar velocity $\vec{v}_i = \vec{0}$, i.e. this perturbation exactly follows the expansion of the Universe at t_i[†]. Its density contrast is $\delta_i \equiv (\rho[t_i] - \bar{\rho}_M[t_i])/\bar{\rho}_M[t_i] \ll 1$ where $\bar{\rho}_M[t_i]$ is the density of the homogeneous Universe at time t_i. From section 5.1, we know that the *growing mode* of this density fluctuation in the linear regime scales as $\delta \propto a$. Expressing the density contrast of our perturbation in terms of redshift rather than expansion factor, we write[‡]:

$$\delta[z] = \frac{3}{5}\frac{1+z_i}{1+z}\,\delta_i \tag{52}$$

and use it to define the *linearly extrapolated density contrast* $\delta_0 \equiv \delta[0] = 3/5(1+z_i)\delta_i$ as the density contrast that our perturbation would have today ($z = 0$) if it had never ceased to grow linearly. This will allow us to use the present time as a common time origin for all perturbations, rather than having to deal with a collection of different initial times when we discuss different initial perturbations.

Moving on to the non-linear regime, Birkhoff's theorem (GR equivalent of Newton's second theorem) tells us that our spherically symmetric homogeneous a.k.a. *top hat* density perturbation evolves as an independent Universe with a slightly different density. We therefore use the (Newtonian) equations of motion (equation 6):

$$\begin{aligned}\frac{d^2 r_b[t]}{dt^2} &= -\frac{4\pi G\bar{\rho}_M[t_i]r_i^3}{3r_b^2[t]}\\ \frac{d^2 r[t]}{dt^2} &= -\frac{4\pi G\bar{\rho}_M[t_i](1+\delta_i)r_i^3}{3r^2[t]}\end{aligned} \tag{53}$$

for the background Universe and the perturbation respectively, where $r_i = r[t_i] = r_{b,i} = r_b[t_i]$. Note that in considering only an EdS Universe we have effectively dropped a term $\Lambda/3 \times r_b[t]$ from the first equation of group (53) and a similar term $\Lambda/3 \times r[t]$ from the second equation, as we did in the linear regime. Multiplying these equations by $\dot{r}_b$ and $\dot{r}$ respectively, and integrating them with respect to time, one gets:

$$\begin{aligned}\frac{\dot{r_b}^2[t]}{2} - \frac{4\pi G\bar{\rho}_M[t_i]r_i^3}{3r_b[t]} &= E_b\\ \frac{\dot{r}^2[t]}{2} - \frac{4\pi G\bar{\rho}_M[t_i](1+\delta_i)r_i^3}{3r[t]} &= E_p\end{aligned} \tag{54}$$

[*]This assumption is only necessary to explore the non-linear regime: our linear calculations never assumed anything about the shape of the perturbations.

[†]This assumption does not reduce the generality of the calculation since a simple change in t_i always allows one to start from such initial conditions.

[‡]You will no doubt wonder where the factor 3/5 comes from, because when $z = z_i$, we should have $\delta[z_i] = \delta_i$. And indeed we do, but at early times we cannot neglect the decaying mode of the perturbation (see section 5.1) in front of its growing mode, for the good reason that it has not yet decayed away! I leave it to you as an exercise to show that given our chosen initial conditions, this decaying mode accounts for 2/5 of the initial density contrast of the perturbation.

Without loss of generality, one can set $E_b = 0$[§] and derive $E_p = -4\pi G\bar{\rho}_M[t_i]\delta_i r_i^2/3$ using the initial conditions $r[t_i] = r_b[t_i]$ and $\dot{r}[t_i] = \dot{r_b}[t_i]$. We can then rewrite the second equation of group (54) as:

$$\dot{r}[t] = \sqrt{\Omega_M[z_i]H^2[z_i]r_i^2\left[(1+\delta_i)\frac{r_i}{r[t]} - \delta_i\right]}$$

Note that although we elect to only define Ω_M and H for the homogeneous Universe, $\dot{r_i}/r_i = H[z_i] = H_0(1+z_i)^{3/2}$ with the last equality being valid in an EdS Universe only. In such a Universe, any overdensity is bound because the background density already is at the critical density, i.e. that which divides open and closed Universes. This means that the expansion of our initially slightly overdense perturbation will become slower and slower with respect to the homogeneous background Universe, until time t_{max} when it reaches its radius of maximum expansion r_{max}, also called *turn around* radius, where $\dot{r}[t_{max}] = 0$. Plugging these conditions in the previous equation we get $r_{max} = (1+\delta_i)r_i/\delta_i$. Further separating variables t and r, and integrating after setting $u = \sqrt{r/r_i}$ and $\Omega_M[z_i] = \Omega_M[0] = 1$, yields:

$$\int_{t_i}^{t_{max}} dt = \frac{1}{H_0(1+z_i)^{3/2}} \int_1^{\sqrt{r_{max}/r_i}} \frac{2u^2 du}{\sqrt{1+\delta_i - \delta_i u^2}}$$

i.e.

$$t_{max} = t_i + \frac{1}{H_0(1+z_i)^{3/2}} \left\{\frac{r_{max}}{r_i} - 1 + \frac{r_{max}}{r_i}\sqrt{\frac{r_{max}}{r_i} - 1}\left(\frac{\pi}{2} - \arcsin\sqrt{\frac{r_{max}}{r_i}}\right)\right\}$$

Since $r_{max}/r_i \gg 1$ and $t_{max} \gg t_i$, replacing r_{max} by its expression as a function of δ_i and r_i and using the definition of δ_0, we can recast the time of turn around as:

$$t_{max} \simeq \frac{\pi}{2H_0}\left(\frac{3}{5\delta_0}\right)^{3/2}$$

Now our perturbation will collapse to a point in a time $t_{coll} = 2t_{max} - t_i \simeq 2t_{max}$ given the symmetry of the problem. In the homogeneous Universe, this time of collapse will correspond to the redshift of collapse z_{coll} defined by:

$$t_{coll} = \frac{2}{3H_0}\frac{1}{(1+z_{coll})^{3/2}}$$

Equating the two expressions obtained for t_{coll}, we can define $\delta_{0,c}$ as the *linearly extrapolated critical density contrast* that a perturbation must have in order to collapse **exactly** at redshift z_{coll}:

$$\boxed{\delta_{0,c}[1+z_{coll}] = \frac{3(12\pi)^{2/3}}{20}(1+z_{coll})} \qquad (55)$$

[§]We can pick the zero point we want for the perturbation, because we can always write $E_p = E_b + E'_p$ and consider the constant E'_p instead of E_p

In other words, if the linearly extrapolated density contrast δ_0 of a perturbation with density contrast $\delta_i \ll 1$ at redshift z_i is equal to $3(12\pi)^{2/3}/20 \approx 1.686$, then this density perturbation will collapse at $z = 0$. If δ_0 is twice this value, it will collapse at $z = 1$, if it is four times greater, collapse will occur at $z = 3$ and so on and so forth.

Note that for an exactly homogeneous perturbation as we have assumed, the density becomes infinite at the center at z_{coll}, since all the mass arrives there at the same time. What it means is that the model is not a very good description of the reality of dark matter halo collapse: in the real Universe, none of the density fluctuations are exactly homogeneous (at least on galactic dark matter halo scales) and 'sub-halos' will collapse first and merge together later to form a bigger halo in a process that is called *violent relaxation* and happens under the only constraint of the integrals of motion. To cut a long story short, in that case, dark matter particles end up populating the available phase sphace equiprobably and the end (stationary) state for the halo is the most probable one. In the case where only mass and total energy are conserved, the final density profile is proportional to r^{-2} and is called *isothermal sphere* profile, and one can show that the *virial radius* of the halo is $r_{vir} \approx 0.5\, r_{max}$. The final density contrast of the halo at z_{coll} is therefore $\delta_h \approx 18\pi^2 \simeq 178$. To a large extent this scenario and numbers are corroborated by direct numerical modelling of the gravitational collapse process using $N-$body simulations.

7.2 The mass function of virialised dark matter halos

Now that we have solved the problem of non-linear growth for an isolated perturbation, we can wonder how these dark matter halos are distributed in mass, and how this mass function evolves in time[¶].

To tackle this problem, we come back to the initial density fluctuations which constitute a Gaussian random field, and consider this field in Fourier space where it is also a Gaussian random field $\delta[\vec{k}]$. Rewriting it as $\delta[V]$ in terms of its variance $V = \sigma_0^2[M]$ at mass scale M, since the Universe is assumed to be smooth on large scales, we have $\lim_{M\to\infty} V = 0$ and $\lim_{M\to\infty} \delta[V] = 0$. As we reduce the smoothing scale (increase $\vec{k}_V$) $\delta[V]$ becomes different from zero in such a way that its average value $< |\delta[V]|^2 >= V$ by definition of the variance[‖]. The net result is that for a smoothing function that has the shape of a 'top hat' in Fourier space:

$$W_k = \begin{cases} 1 & \text{if } k \leq k_V \,, \\ 0 & \text{if } k > k_V \end{cases}$$

$\delta[V]$ performs a true randow walk: each increment to $\delta[V]$ when $V[k_V]$ increases comes from a different shell of Fourier modes and thus is not correlated with the previous increment, as the random field is Gaussian. This means that the trajectories $\delta[V]$ follow a simple diffusion equation:

$$\frac{\partial D}{\partial V} = \frac{1}{2}\frac{\partial^2 D}{\partial \delta^2} \tag{56}$$

[¶]This subsection contains advanced material which will not be used to set either problem or exam questions
[‖]See Lacey and Cole (1993, MNRAS, 262, 627) for details

where $D[\delta, V]$ is the number of trajectories in $[\delta,\ \delta + \partial\delta]$ at "time" V. To compute the number of virialized halos of a given mass at time t (or redshift z) one just needs to place a barrier which absorbs trajectories at $\delta_{0,c}[t]$. Equation (56) then has a unique solution which was first computed by Chandrasekhar (1943):

$$D[\delta, V, \delta_{0,c}[t]]d\delta = \frac{1}{\sqrt{2\pi V}}\left(\exp\left[-\frac{\delta^2}{2V}\right] - \exp\left[-\frac{(\delta - 2\delta_{0,c}[t])^2}{2V}\right]\right) d\delta$$

Now the probability PdV that a trajectory will be absorbed between V and $V + dV$ must equal the difference between the total number of trajectories before and after crossing the barrier, i.e.:

$$\begin{aligned} P[V, \delta_{0,c}[t]]dV &= -\frac{\partial}{\partial V}\left(\int_{-\infty}^{\infty} D[\delta, V, \delta_{0,c}[t]]d\delta\right) dV \\ &= -\left[\frac{1}{2}\frac{\partial D}{\partial \delta}\right]_{-\infty}^{\infty} dV, \end{aligned}$$

Therefore, inserting the analytic expression of D in the previous equation yields:

$$P[V, \delta_{0,c}[t]]dV = \frac{\delta_{0,c}[t]}{\sqrt{2\pi V^3}} \exp\left[-\frac{\delta_{0,c}^2[t]}{2V}\right] dV$$

which is the fraction of mass of the Universe which is enclosed in virialized halos whose masses are comprised in $[M, M + dM]$ (corresponding to $[V, V + dV]$). We can then easily calculate the number density of virialized halos of mass M as a function of time t:

$$\boxed{\begin{aligned} \frac{dn}{dM}[M, t]dM &= \frac{\bar{\rho}_M[0]}{M} P(V, \delta_{0,c}[t]) \left|\frac{dV}{dM}\right| dM \\ &= \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \frac{\bar{\rho}_M[0]}{M^2} \frac{\delta_{0,c}[t]}{\sigma_0(M)} \left|\frac{d\ln\sigma_0}{d\ln M}\right| \exp\left[-\frac{\delta_{0,c}^2[t]}{2\sigma_0[M]^2}\right] dM \end{aligned}} \quad (57)$$

which is the Press-Schechter (1974) formula.

Note that it is possible to extend this approach to take into account mergers with a two barrier, conditional probability calculation (see e.g. Lacey and Cole 1993) but that in any case the method lacks any spatial or dynamical information about such dramatic events. This can only be remedied by running dark matter $N-$body simulations, against which the analytic results presented here have been checked (see Figure 2). Finally, let us emphasise again that to form galaxies, one cannot be content with describing the evolution of dark matter haloes alone but one must also account for the relevant baryonic (normal gas) physics, and that is **much** more complicated, as for example we do not yet fully understand how a star forms!

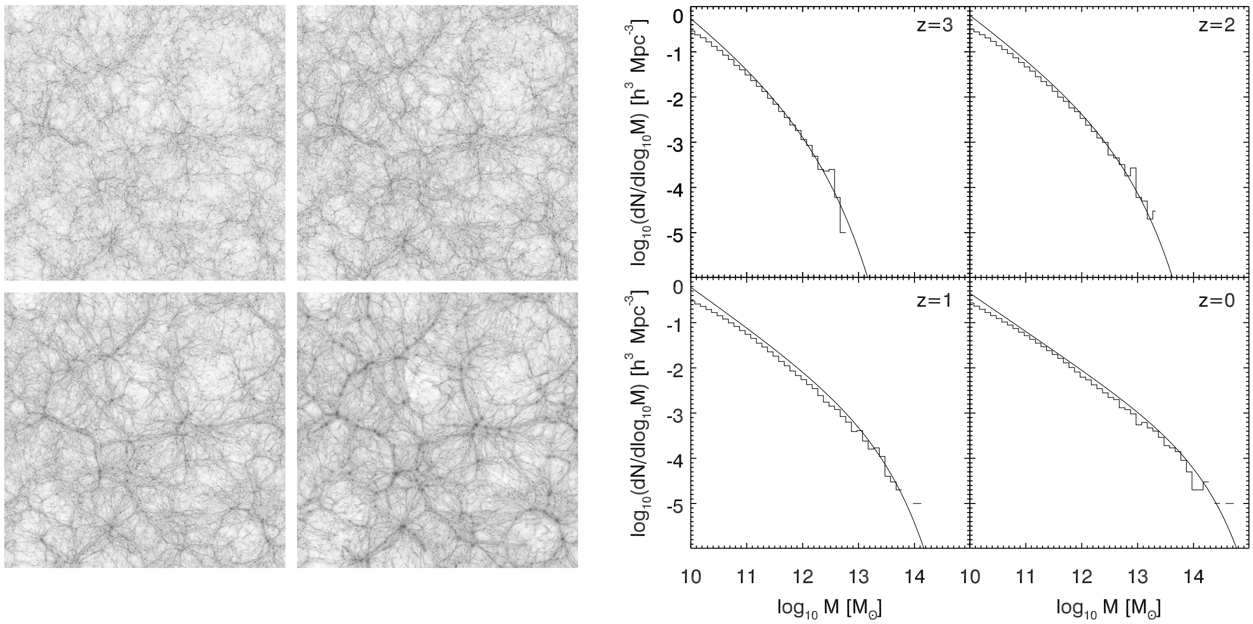

Figure 2: Evolution of a slice of the gas density field in a cosmological simulation (left panels, Horizon-AGN simulation: see `https://www.horizon-simulation.org` for detail) along with the corresponding evolution of the virialized dark matter halo mass function directly measured in the simulation (right panels, histograms) and the Press-Schechter analytic description discussed in the text (Eq. 57, right panels, solid lines). In both cases, redshift decreases from $z = 3$ (top left panels) to $z = 0$ (bottom right panels) as indicated on the halo mass function plot.

8 Comparing to observations

8.1 The spatial distribution of galaxies

A typical survey of galaxies will be like the image in Figure 3

Furthermore, we have been talking about δ, a continuous field defined over all of space. But what we actually see are galaxies, bright dots in the sky. We must relate our theory, the theory of how δ arises and evolves with our data, a catalogue of positions of galaxies in the sky. Once again let us take as our starting point, the mean density, ρ_0 and expand it in terms of the density contrast:

$$\rho = \rho_0(1 + \frac{\delta\rho}{\rho}) = \rho_0(1 + \delta)$$

The galaxies must trace this density somehow. For example, if the density is high (i.e. where there is an overdensity) we expect to see more galaxies, more bright dots. If the density is low, we expect to see less galaxies, a void. We can make this comparison more quantitative.Take the distribution of galaxies in the Universe and lay down balls of radius λ all over. Calculate the mass contained in each ball (i.e. add up the mass in all the galaxies contained in each ball). We have that

$$M(\lambda) = \int_{S_\lambda} d^3x \rho(\vec{x})$$

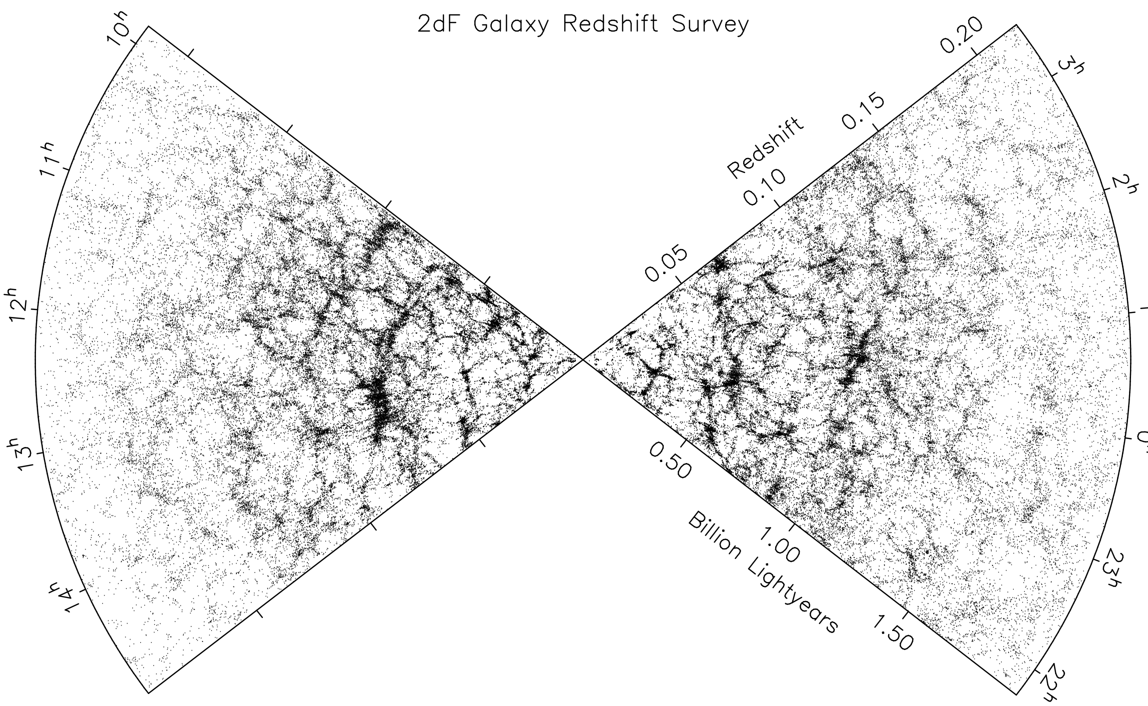

Figure 3: A survey of galaxies over the sky illustrates there are inhomogeneities spanning a wide range of scales (2dF Galaxy Redshift Survey -based at Oxford)

where S_λ is a sphere of radius λ. We can find the average over all balls to get

$$\bar{M}(\lambda) \equiv \langle M(\lambda) \rangle = \frac{4\pi}{3}\lambda^3\rho_0$$

Suppose we now calculate the variance

$$[\delta M(\lambda)]^2 \equiv \langle [M(\lambda) - \bar{M}(\lambda)]^2 \rangle$$

Some balls will be heavier and others will be lighter. There should be a scatter. One can show that the mass variance is given by

$$\Delta^2(k) \equiv \left[\frac{\delta M(\lambda)}{\bar{M}(\lambda)}\right]^2 \simeq \frac{k^3}{2\pi^2}P(k) \quad \text{with } k = \frac{2\pi}{\lambda}$$

So $P(k)$ gives you the fluctuations in mass of balls with a given radius. Different types of clumping will lead to different $P(k)$ and these can be compared to the clumping that we see in the distribution of galaxies.

If we were to consider a universe with baryons and radiation, as described above, we would find a power spectrum as in Figure 1.

Clearly the problem is that perturbations in the baryons are prevented from growing because of the way they link up to radiation. One solution is to have a different form of non-relativistic matter

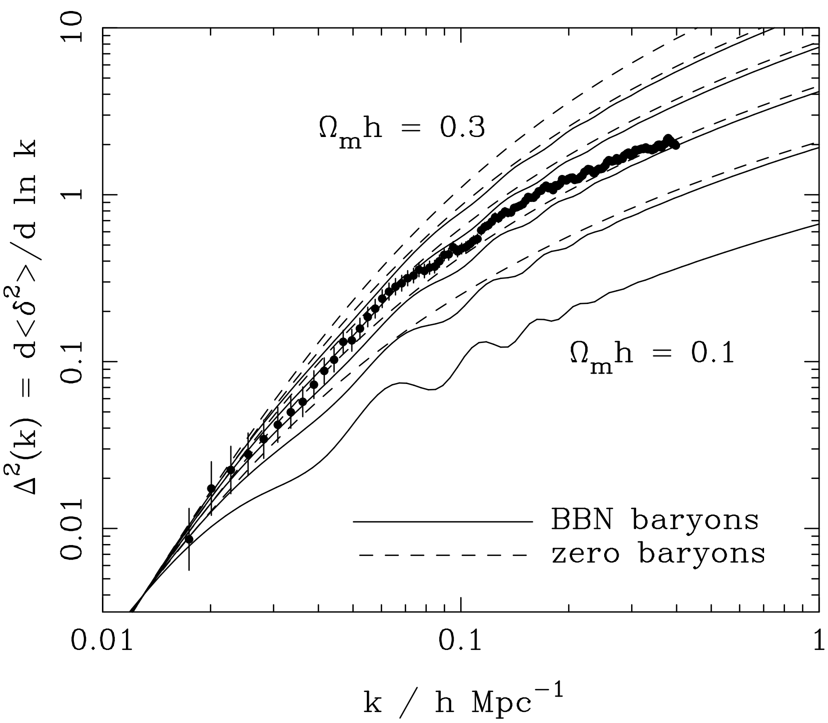

Figure 4: The mass variance, $\Delta^2(k)$ of inhomogeneities with cold dark matter (either on its own (dashed lines) or including baryons (solid lines)) compared to that of the real Universe as measured from the galaxy distribution (solid points with error bars).

that doesn't interact with radiation. It won't feel that baryon pressure and will have its Jean's length equal to zero. Because this matter does not interact with light, it can be called "dark matter". An example of the mass variance $\Delta(k)^2$ compared to data is shown in Figure 4.

8.2 Anisotropies in the Cosmic Microwave Background

We have focused on the evolution of perturbations in the density field and how we can connect them to the distribution of galaxies. There is, of course, another very important component to the Universe, which plays a significant role at early times: radiation.

Let us briefly recall the history of radiation in the Universe. At very early times the universe is highly ionized, Hydrogen is dissociated into free electrons and protons and the mean free path of photons is effectively 0. The universe is opaque. At late times, the universe is neutral, protons and electrons are in Hydrogen atoms and photons are free to propagate. The Universe is transparent. The transition from one state to another is naturally related to the binding energy of hydrogen and the evolution of the ionization fraction

$$X \equiv \frac{n_p}{n_p + n_H} \tag{58}$$

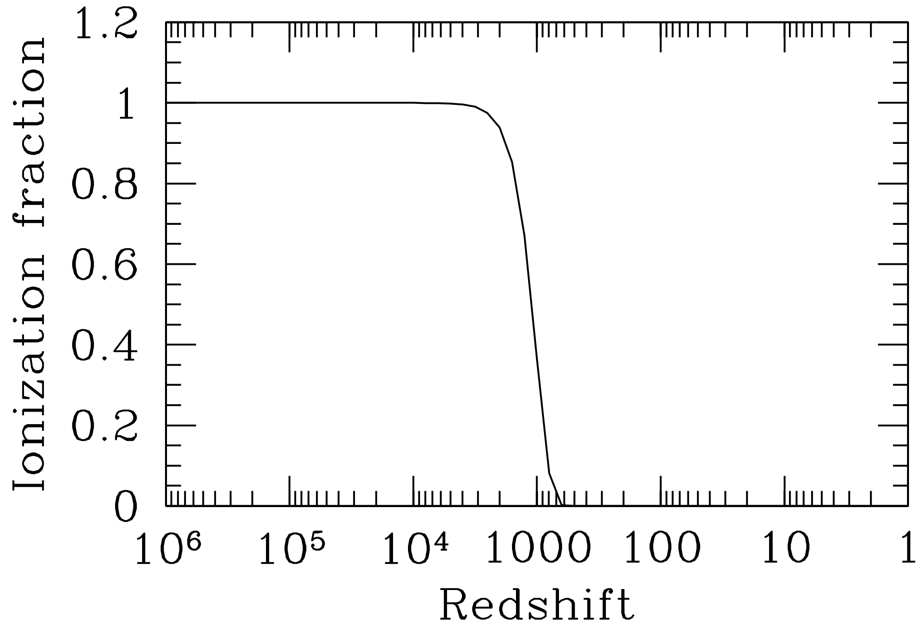

Figure 5: The evolution of the ionisation fraction as a function of redshift. Note that. even though the transition is sharp at recombination, $z \sim 1100$, it is not instantaneous (width $\Delta z \sim 80$). This gives rise to the so-called Silk damping effect, where CMB photons continue to scatter off free e^- for a while before freely propagating to us.

At sufficiently early times we will find that $X = 1$, i.e. the Universe is completely ionized. As it crosses a certain threshold, electrons and protons combine to form Hydrogen. This happens when the temperature of the Universe is $T \simeq 3570K$ or $0.308eV$, i.e. when it was approximately $380,000$ years old, at a redshift of $z \simeq 1100$. We would naively expect this to happen at $13.6eV$. One way to think about why this isn't so is that, at a given temperature there will always be a few photons with energies larger than the average temperature. This energetic photons only become unimportant at sufficiently low temperatures.

We can now reconstruct the history of a photon left over from recombination. We know that, post recombination, a photon has been travelling along a straight path from then until now. It has travelled a distance $d_* \simeq \eta_0 - \eta_*$ where η_0 (η_*) is the conformal time today (at recombination). Before recombination, the mean free path of the photon was negligable, it was effectively standing still. So we can think of recombination as the time when these photons were released to travel through space. When we look from a fixed point in the sky, we will received photons that have travelled straight towards us since η_*, the point at which they are released will map out a two dimensional sphere, of radius d_* which is known as the surface of last scatter. We can think of this light as an image of the the surface of last scatter, a photograph of a spherical slice of the universe at η_*.

Recall that the Universe was in thermal equilibrium and hence the radiation should have a black body spectrum. And given that it is approximately homogeneous on large scales, the black body will be the same in all directions, with the peak at the same temperature in all directions. But we also know that it perturbed and hence we should see deviations. These deviations will have various contributions.

For a start, we have the Stefan-Boltzman law: $\rho_\gamma = \sigma T^4$ where σ is the Stefan Boltzman constant. But we then have

$$\frac{\delta T}{T} = \frac{1}{4}\frac{\delta\rho_\gamma}{\rho_\gamma}.$$

This seems quite easy to understand: the more compressed the radiation is, the hotter it is. So points on the surface of last scattering which have denser radiation will look hotter. This is known as the *intrinsic term*

When the photons are released at the surface of last scattering, the will collide with one last electron or proton before they propagate towards us. That proton or electron may have a peculiar velocity, $\vec{v}_B$. This will impart a Doppler shift on the photon, i.e. the observed temperature will be $T' = T(1 - \vec{v}_B \cdot \vec{n})$ so that

$$\frac{\delta T}{T} = -\vec{v}_B \cdot \vec{n}$$

This is known as the *Doppler* term.

There will be gravitational effects too. If the photon is caught in a gravitational well at the surface of last scatter, it will be held back, i.e. gravitationally redshifted. The bigger the well, the colder the photon becomes:

$$\frac{\delta T}{T} = -\Phi$$

This is known as the *Sachs-Wolfe* term. Finally, as the photon propagates towards us through empty space, space time is changing and warping as it evolves. The photon will be redshifted or blue shiftet according to:

$$\frac{\delta T}{T} = -2\int_{\eta_*}^{\eta_0} d\eta\,\dot{\Phi} \tag{59}$$

This is known as the *Integrated Sachs-Wolfe* term.

Let us put all this together now. Recall: the surface of last scatter is a sphere with radius d_*. Suppose we look in a give direction $\hat{n}$. We will see

$$\frac{\delta T}{T}(\hat{n}) \;=\; \frac{1}{4}\delta(\eta_*, d_*\hat{n}) - [\vec{v}_B \cdot \hat{n}](\eta_*, d_*\hat{n}) - \Phi(\eta_*, d_*\hat{n}) - 2\int_{\eta_*}^{\eta_0} d\eta'\,\dot{\Phi}[\eta', (\eta_0 - \eta')\hat{n}] \tag{60}$$

In summary, a measurement of $\frac{\delta T}{T}$ is a snapshot of the universe at t_*, it is related to quantities that we know from studying large scale structure and we can use it to do "archeology"of the Universe.

What do we expect to see? First of all, recall the we looked at how perturbations in the radiation and baryons evolve in the radiation era, before recombination. On small scales, there should be a

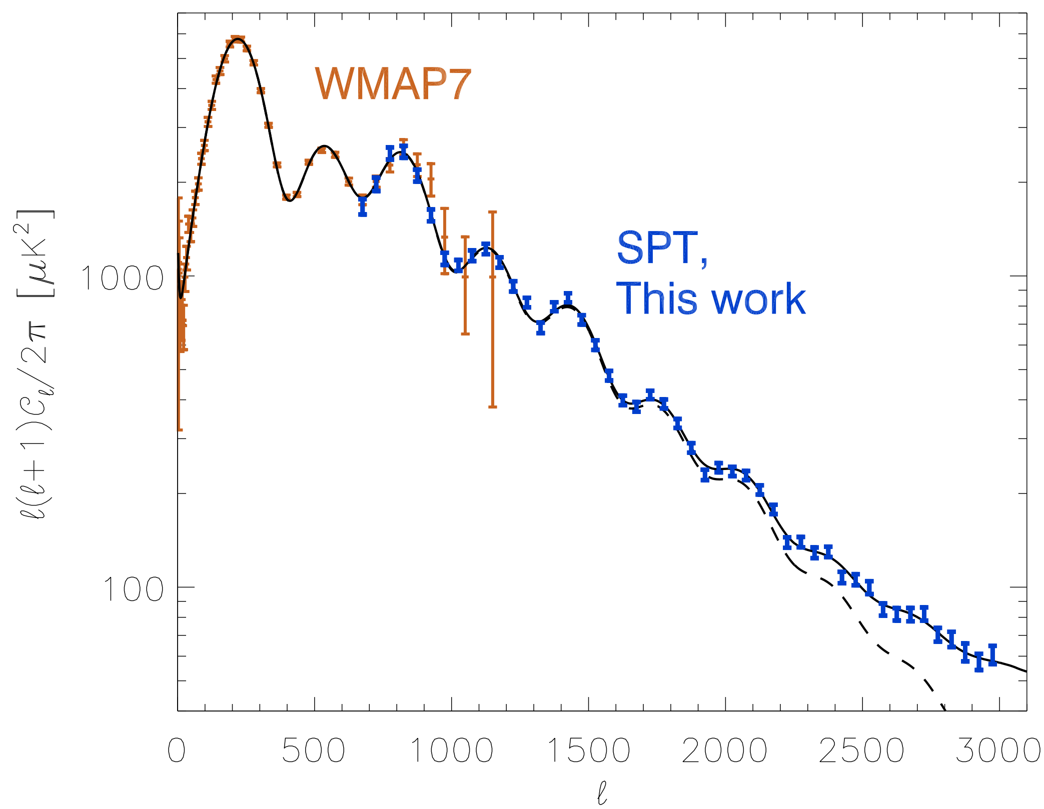

Figure 6: The best fit power spectrum of the CMB as compared to the WMAP and SPT data (Keisler 2011) Note the exquisite accuracy of the measurement and the damped oscillations we discuss in the text.

series of acoustic oscillations. We found a solution, $\delta_\gamma \propto J_2(k\eta/\sqrt{3})$. We also know that $\vec{\nabla}\cdot\vec{v} = -\dot{\delta}$ So if $\delta_\gamma \propto \cos(k\eta/\sqrt{3})$ then $\vec{v}_B\cdot\hat{n} \propto \sin(k\eta/\sqrt{3})$. We should see an oscillatory pattern in $\frac{\delta T}{T}$ with the spacing between peaks and troughs set by the angular projection of $k\eta_*/\sqrt{3}$. On very large scales we expect it to be relatively featureless.

On small scales, an altogether different phenomenon kicks in. During recombination, as the photons decouple from the baryons, they will slowly start to propagate. They will move around a little and in doing so, they will leak energy from high density regions to low density regions. The net effect is to reduce high density regions, fill in low density regions and effectively smooth out perturbations. features on small scales will be smoothed out.

How do we analyse a map of the cosmic microwave background? We need to take the equivalent of the Fourier transform except now it is on the surface of the sphere. You will recall from mathematical methods and quantum mechanics that there is a useful basis to this in, the Spherical Harmonic functions. So we can take

$$\frac{\delta T}{T}(\hat{n}) = \Sigma_{\ell m} a_{\ell m} Y_{\ell m}(\hat{n}). \tag{61}$$

and plot the power spectrum

$$C_\ell = \frac{1}{2\ell+1}\Sigma_m |a_{\ell m}|^2 \tag{62}$$

In Figure 6 you can see the structure of the power spectrum. There is clear evidence of the oscillatory structure as well as of the damping on very small scales and you can appreciate how remarkably well this power spectrum has been measured.

9 Gravitational Lensing

Gravity bends the path of photons. This makes gravitational lensing an extremely powerful tool to study the mass distribution of matter in the Universe, since this effect is independent of whether or not structures (galaxies, groups or clusters) have reached an equilibrium or are still growing and changing.

9.1 Lensing by a compact object

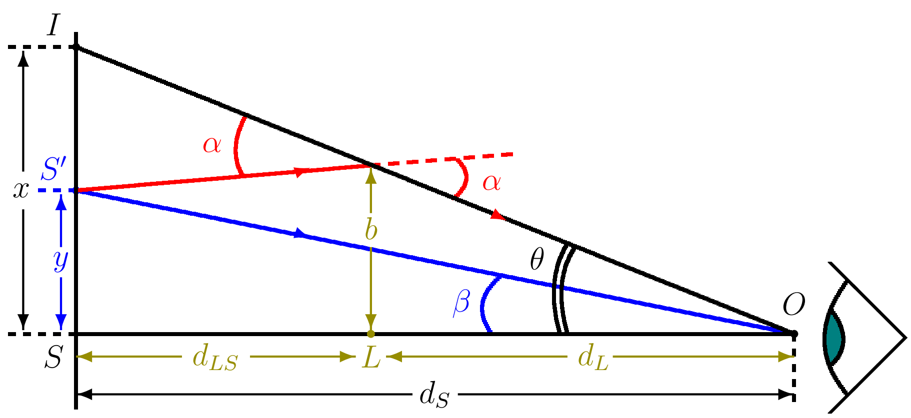

Figure 7: A point mass L is gravitationally deflecting the light emitted by a source S'.

The Sun bending of light rays was the first test of Einstein's theory of general relativity, carried out by Eddington and collaborators during a solar eclipse in May 1919. If the gravity of a mass M, located at point L, bends the light of a source located at point S' in the image plane — this plane is perpendicular to the line $O - L$ it defines with and observer located at point O —, then this source will appear to be located at point I in the image plane instead of S' (see Figure 7). As seen in the 3rd year GR course, Einstein predicted that the light passing at a distance b from L in the lens plane is bent by an angle α approximately given by:

$$\boxed{\alpha \approx \frac{4GM}{bc^2} = \frac{2R_s}{b}} \tag{63}$$

where G is the gravitational constant, c the speed of light and $R_s = 2GM/c^2$ is the Schwarschild radius beyond which the light cannot escape from the gravitational pull of the mass. Recall that this is *twice* the value of the predicted Newtonian deflection and that the approximation holds as long as the bending is small, i.e. $\alpha \ll 1$.

Using equation (63) we can now calculate where the image of a distant source will appear if a point mass is placed in front of it. If the mass lens had been absent, we would have seen the source at S' at an angle $\beta = LOS'$ from the direction $O - L$. Writing d_S as the distance from the observer to the image plane (i.e. the intersection S of $O - L$ with the plane containing S') and y the distance $S - S'$ in the plane, we have $\beta \approx y/d_S$ as long as $d_S \gg y$. Because the light is bent by an amount α, the source appears at an angle $\theta = LOI$ instead of β. If we call x the distance $S - I$ in the image plane, we have $\theta \approx x/d_S$ as long as $d_S \gg x$. For a small bending, the displacement of the point mass in the image plane is then given by $x - y = \alpha d_{LS}$ where d_{LS} is the distance between the lens at L and the image plane, i.e. $L - S$. Finally, the impact parameter b, which measures the distance between the intersection L' of a light ray emitted by the source and the lens plane (i.e. the plane perpendicular to $O - L$ and containing L), simply is $b = \theta d_L$ if $d_L \gg b$ and d_L is the distance between the observer and the lens $O - L$. Using Eq.63 and dividing it by d_S we find:

$$\theta - \beta = \alpha \frac{d_{LS}}{d_S} = \frac{1}{\theta} \frac{4GM}{c^2} \frac{d_{LS}}{d_L d_S} \equiv \frac{1}{\theta} \theta_E^2 \tag{64}$$

where the angle θ_E is called the *Einstein radius*. We then have to solve a quadratic equation to find the angular distance θ between the lens L and the image of the point source L' (or I since we are only interested in the angular separation here):

$$\theta^2 - \beta\theta - \theta_E^2 = 0$$

i.e.

$$\boxed{\theta = \frac{\beta \pm \sqrt{\beta^2 + 4\theta_E^2}}{2}} \tag{65}$$

We can see immediately from Eq.65 that a point source located exactly behind the lens (i.e. at point S) will be seen as a circle of light on the sky with radius θ_E since $\beta = 0$ in this case. When $\beta > 0$, the image θ_+ is further from the lens since $\theta_+ > \beta$ and lies outside the Einstein radius since $\theta_+ > \theta_E$. These exterior images were the ones seen around the eclipsed Sun by Eddington. The image at θ_- is inverted since $\theta_- < 0$ and lies within θ_E since $\theta_- > -\theta_E$ on the opposite side of the lens.

Quite regularly, one star of the Milky Way's bulge is gravitationally lensed by another one in the disk. However, images θ_+ and θ_- are too close from one another to distinguish individually in that case, but one can still tell the star is being lensed as it appears brighter on the sky (see Figure 8). Because of the small size of the Einstein ring, gravitational lensing by compact objects in the halo is called *microlensing*.

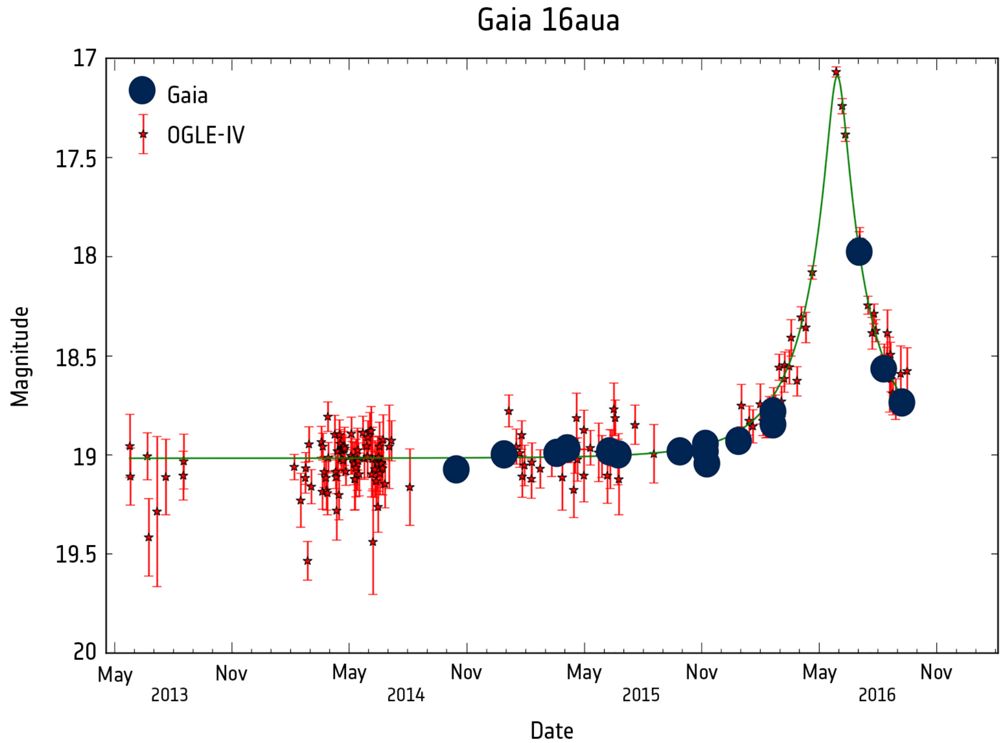

Figure 8: Example of a microlensing event (Gaia 16aua) detected by Gaia (solid circles) and Ogle (stars with error bars). The solid green line represents the best fit model that can be obtained using the theory we have described in the text (©ESA).

9.2 Lensing by extended sources

When the lens is an entire galaxy or a cluster of galaxies, we can first think of it as a collection of point masses. We then rewrite equation (63) as:

$$\vec{\alpha}(\vec{b}) = \sum_i \frac{4GM_i}{c^2} \frac{\vec{b} - \vec{b_i}}{|\vec{b} - \vec{b_i}|^2} \tag{66}$$

where M_i are the individual point masses constituting the lens, and $\vec{b_i}$ their position vectors w.r.t. the lens centre. Note that α and b become vectors, as the spherical symmetry of a single point-like lens is now broken. We can calculate the bending of the light emitted by a background source by summing up the effects of all the point masses within the lens.

If the lens is compact as compared with both its distance d_L to the observer **and** its distance d_{LS} to the source plane, then the deflection of the light only depends on the *continuous* surface density $\Sigma(\vec{b})$ of the lens. In this continuous limit, one must then specify the light ray's closest approach to the galaxy/cluster center by a vector $\vec{b}$ and integrate over the galaxy/cluster to calculate the deflection vector $\vec{\alpha}$:

$$\vec{\alpha}(\vec{b}) \equiv \nabla \Psi_L(\vec{b}) \text{ with } \Psi_L(\vec{b}) = \frac{4G}{c^2} \int_S \Sigma(\vec{b}') \ln |\vec{b} - \vec{b}'| d^2 b' \tag{67}$$

where the *lensing potential*, Ψ_L, has a similar form to the gravitational potential $\phi(\vec{b})$ but the integral is two dimensional instead of three dimensional, and the term $1/|\vec{b} - \vec{b}'|$ is replaced by the $\ln |\vec{b} - \vec{b}'|$ term. In the most general of cases, we must calculate $\Psi_L(\vec{b})$ numerically with a computer from a distribution of matter given by $\Sigma(\vec{b})$, but suppose that the lens is axisymmetric, so that Σ only depends on the projected distance R of the source to the center of the galaxy/cluster. We can then show that the bending of a ray of light passing at radius b from the centre of the lens only depends on the mass $M_{<b}$ **projected** within that circle. Equation (67) then simplifies to:

$$\alpha(b) = \frac{4G}{bc^2} \int_0^b \Sigma(R) 2\pi R dR = \frac{4G}{c^2} \frac{M_{<b}}{b}. \tag{68}$$

To prove this, we can just adapt the arguments that one uses to prove Newton's second theorem**, so that the light is bent just as if all the material projected within radius b had been replaced by a point of the same total mass located at the centre.

Let us now use equation (68) to figure out how an axisymmetric galaxy/cluster bends the light emitted by a distant galaxy located behind it. The geometry is the same as for micro lensing, except that the lens is now not a compact object! If the lensing cluster had been absent, we would have seen the background galaxy at point S', i.e. at an angle β from $O - L$, the line which joins the observer to the cluster center. Instead, we see this background galaxy's image at point I which makes an angle θ with $O - L$. Thus we only have to modify equation (64) which was valid for micro-lensing in the following way:

$$\theta - \beta = \alpha \frac{d_{LS}}{d_S} = \frac{1}{\theta} \frac{4GM_{<b}}{c^2} \frac{d_{LS}}{d_L d_S}, \tag{69}$$

Remembering that $b = \theta d_L$, we can divide both sides of equation (69) by θ, to rewrite it in terms of the *critical density for lensing* Σ_{crit}:

$$\boxed{\beta = 1 - \frac{1}{\Sigma_{crit}} \frac{M_{<b}}{\pi b^2} \quad \text{where} \quad \Sigma_{crit} \equiv \frac{c^2}{4\pi G} \frac{d_S}{d_L d_{LS}}.} \tag{70}$$

The quantity $M_{<b}/(\pi b^2)$ is simply the average surface density within radius b. Usually, $\Sigma(R)$ declines from a peak at the centre of the lens, so that this average will fall as well. It naturally follows that if the central density of the lens is greater than Σ_{crit}, then the image of a source located at $\beta = 0$ exactly in line with the observer and the cluster centre will be a thin circular *Einstein ring*

**We can also prove in the same way the equivalent of Newton's first theorem: a light ray passing through a uniform circular ring is not bent.

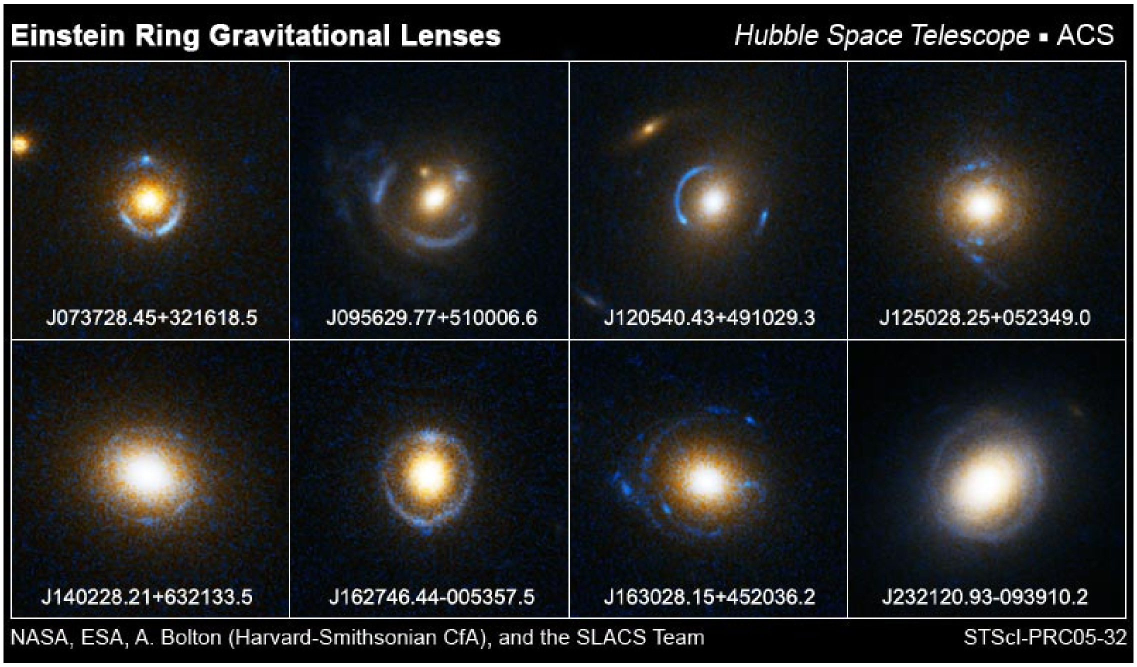

Figure 9: Examples of strong lensing events (Einstein rings and arcs) by extended lenses (galaxies) with the Hubble Space Telescope.

(see Figure 9) of angular size $\theta_E = b_E/d_L$ where b_E is the radius where the average density equals the critical density value, i.e. $M_{<b_E}/(\pi b_E^2) = \Sigma_{crit}$. On the other hand, if the central surface density of the lens is smaller than Σ_{crit}, then the cluster cannot produce multiple images of any background galaxy and no ring is observed. Also note that the other important difference with micro-lensing is that distances are now *cosmological*, so we cannot ignore the expansion of the Universe when calculating them. As we only are concerned with angles when studying gravitational lensing, we will use *angular distances* for d_L, d_{LS} and d_S. Indeed, at time $t_{\rm em}$ when it emits a photon, a galaxy of size D located at a comoving distance $r_{\rm em}$ (i.e. either the lens at point L or the source S' in Figure 7) from the observer (point O on the same Figure) will subtend an angle φ on the sky given by:

$$\varphi = \frac{D}{a_{\rm em} r_{\rm em}} = \frac{D(1+z)}{a_0 r_{\rm em}} = \frac{D(1+z)}{r_{\rm em}} \equiv \frac{D}{d_A}$$

where we have naturally defined the angular distance as $d_A \equiv r_{\rm em}/(1+z)$. To calculate $r_{\rm em}$, one then uses the metric (Eq. 1) for a photon ($ds^2 = 0$), separates variables and plugs in the expansion rate of the Universe (Eq. 8). In the most general case, the resulting integral will not have an analytic form, but for an *Einstein-de Sitter* Universe with $\Omega_M[z] = 1$ and all other contributions to the energy density nil, one gets:

$$r_{\rm em} = \frac{2c}{H_0}\left(1 - \frac{1}{\sqrt{1+z}}\right) \quad \text{and thus} \quad d_A = \frac{2c}{H_0(1+z)}\left(1 - \frac{1}{\sqrt{1+z}}\right).$$

9.3 Weak gravitational lensing

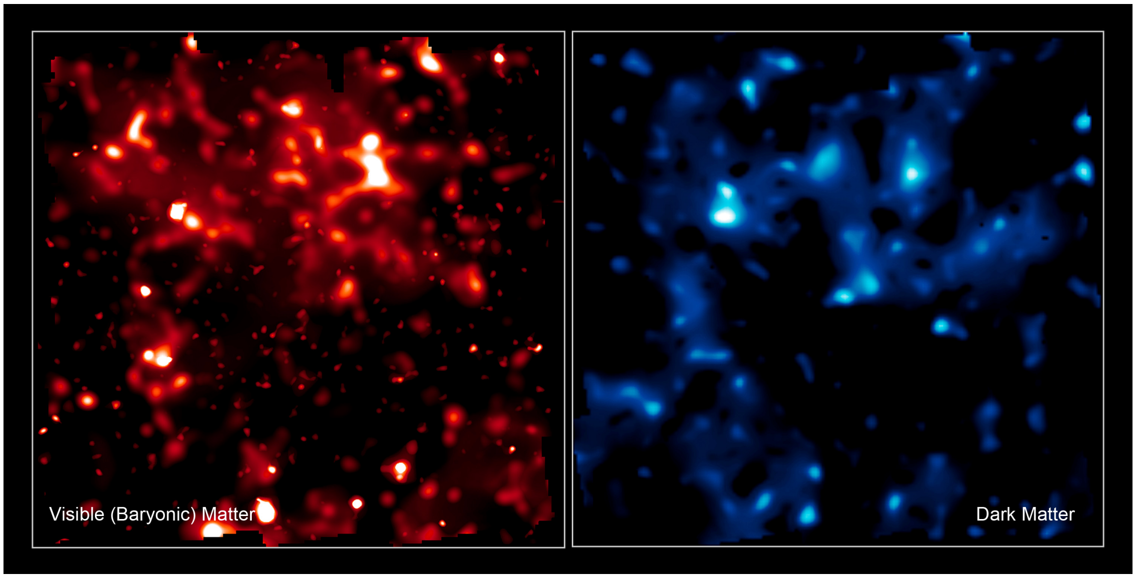

Figure 10: Reconstruction of the dark matter density distribution (right hand panel, in false blue colours) using weak lensing. For comparison the distribution of normal baryonic matter is given in false red colours on the left panel (©ESA). The map (COSMOS field) covers an area of sky nine times the angular diameter of the full Moon.

When galaxies lie behind a lensing cluster, but are located well outside of its Einstein radius, their images are only weakly magnified and slightly stretched in the tangential direction, i.e. galaxies which would otherwise appear as perfect circular discs become ellipses with tangential and radial axes having a ratio $\frac{x}{y}/\frac{\Delta x}{\Delta y}$ or $|\frac{d\beta}{d\theta}|/|\frac{\beta}{\theta}|$ when using the same notation as in the microlensing case.

The *shear* γ is therefore used in weak lensing studies, which measures the difference in the amount of compression the lens exerts on the source in the tangential and the radial directions. For an image located at a distance $b \gg \theta_E d_L$ from the lensing cluster's center, we have:

$$\boxed{\gamma \equiv \frac{1}{2}\left[\frac{d\beta}{d\theta} - \frac{\beta}{\theta}\right] = \frac{\bar{\Sigma}_{<b} - \Sigma(b)}{\Sigma_{crit}}} \tag{71}$$

where $\bar{\Sigma}_{<b} = M_{<b}/(\pi b^2)$ is the average surface density of matter projected within radius b and $\Sigma(b)$ is the surface density at radius b.

Measuring the average shape of many background galaxies that have been weakly distorted allows one to estimate the shear and hence probe the (projected) mass distribution in the outer parts of galaxy clusters[††] . Note that this technique also applies to distortions caused by the large scale structure of

[††]This is possible even though one does not a priori know the shape of the lensed galaxies because they are all compressed by the lens in a systematic way, regardless of their intrinsic shape!

the Universe (see Figure 10), but in this case the light from background galaxies is bent multiple times by each of the structure present on its way to the observer, so one cannot assume the bending occurs at a single distance d_L. The only way to calculate the amplitude of the weak lensing shear signal originating from large scale structure lensing is therefore to perform ray-tracing numerical simulations by post-processing the matter density distribution provided by a cosmological simulation.