Particle Cosmology

LaTeX notes by Ruihao Li

October 19, 2018

Declaration

This set of notes is based on the lectures of the 2016 Honours physics course, Particle Cosmology and Baryonic Astrophysics, given at the University of Sydney. The first two chapters are essentially the reproduction of Dr. Michael Schmidt's lecture notes, which cover basic concepts in cosmology and the thermal history of the universe. Some calculation details and additional explanations are included here. Chapters 3 and 4 are the contents covered by Dr. Jan Harmann, whose focus is on inflation and structure formation. Materials are heavily borrowed from the Cambridge lecture notes by Daniel Baumann and the book, *Modern Cosmology*, by Dodelson.

Contents

Chapter 1

The Basic Ingredients of the Universe

1.1 FRW metric

First, spatial homogeneity and isotropy mean that the universe can be described by a metric of the form

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a(t)^2 \times \mathrm{d}\ell^2 \qquad \text{with} \qquad \mathrm{d}\ell^2 = \gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j. \tag{1.1}$$

Then, depending on the geometry of space, we have different metrics for the symmetric 3-space $d\ell^2$.

- Flat space (E^3: 3-dimensional Euclidean space): the line element is simply

$$\mathrm{d}\ell^2 = \mathrm{d}\mathbf{x}^2 = \delta_{ij}\mathrm{d}x^i\mathrm{d}x^j. \tag{1.2}$$

- Positively curved space (S^3: 3-sphere): a 3-space with constant positive curvature can be represented as a 3-sphere embedded in 4-dimensional Euclidean space E^4:

$$\mathrm{d}\ell^2 = \mathrm{d}\mathbf{x}^2 + \mathrm{d}u^2, \qquad \mathbf{x}^2 + u^2 = a^2, \tag{1.3}$$

  where a is the radius of the 3-sphere.

- Negatively curved space (H^3: 3-hyperboloid): a 3-space with constant negative curvature can be represented as a 3-hyperboloid embedded in four-dimensional Lorentzian space $\mathbb{R}^{1,3}$:

$$\mathrm{d}\ell^2 = \mathrm{d}\mathbf{x}^2 - \mathrm{d}u^2, \qquad \mathbf{x}^2 - u^2 = -a^2. \tag{1.4}$$

For the last two cases, it is convenient to rescale the coordinates, $\mathbf{x} \to a\mathbf{x}$ and $u \to au$. Then the line elements of the spherical and hyperbolic cases are then

$$\mathrm{d}\ell^2 = a^2(\mathrm{d}\mathbf{x}^2 \pm \mathrm{d}u^2), \qquad \mathbf{x}^2 \pm u^2 = \pm 1. \tag{1.5}$$

The differential of the embedding condition $\mathbf{x}^2 \pm u^2 = \pm 1$ tells us

$$u\ \mathrm{d}u = \mp \mathbf{x} \cdot \mathrm{d}\mathbf{x}. \tag{1.6}$$

So

$$\mathrm{d}\ell^2 = a^2\left[\mathrm{d}\mathbf{x}^2 \pm \frac{(\mathbf{x}\cdot\mathrm{d}\mathbf{x})^2}{u^2}\right] = a^2\left[\mathrm{d}\mathbf{x}^2 \pm \frac{(\mathbf{x}\cdot\mathrm{d}\mathbf{x})^2}{1 \mp \mathbf{x}^2}\right]. \tag{1.7}$$

Then we can unify it with the Euclidean line element Eq. (1.2) by writing

$$\mathrm{d}\ell^2 = a^2\left[\mathrm{d}\mathbf{x}^2 + k\frac{(\mathbf{x}\cdot\mathrm{d}\mathbf{x})^2}{1 - k\mathbf{x}^2}\right] \equiv a^2\gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j, \tag{1.8}$$

with

$$\gamma_{ij} = \delta_{ij} + k\frac{x_i x_j}{1 - k(x_k x^k)} \tag{1.9}$$

and

$$k = \begin{cases} 0, & \text{Euclidean} \\ +1, & \text{spherical (closed universe)} \\ -1, & \text{hyperbolic (open universe)} \end{cases}. \tag{1.10}$$

Often it is more convenient to work in the spherical polar coordinates (r, θ, ϕ), where we have

$$\begin{aligned} \mathrm{d}\mathbf{x}^2 &= \mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2) \equiv \mathrm{d}r^2 + r^2\mathrm{d}\Omega^2 \\ \mathbf{x}\cdot\mathrm{d}\mathbf{x} &= r\ \mathrm{d}r. \end{aligned} \tag{1.11}$$

Then the metric in Eq. (1.8) becomes diagonal:

$$\begin{aligned} \mathrm{d}\ell^2 &= a^2\left[\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2 + k\frac{r^2\mathrm{d}r^2}{1-kr^2}\right] \\ &= a^2\left[\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2 + \frac{kr^2\mathrm{d}r^2 - \mathrm{d}r^2 + \mathrm{d}r^2}{1-kr^2}\right] \\ &= a^2\left(\frac{\mathrm{d}r^2}{1-kr^2} + r^2\mathrm{d}\Omega^2\right). \end{aligned} \tag{1.12}$$

Remember that a shows up as a scaling factor, and k is called the curvature parameter. Then the *Friedmann-Robertson-Walker* (FRW) metric can be written as

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)\left(\frac{\mathrm{d}r^2}{1-kr^2} + r^2\mathrm{d}\Omega^2\right) = -\mathrm{d}t^2 + a^2(t)\gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j. \tag{1.13}$$

The *comoving coordinates* and *physical coordinates* (remember that the comoving distance between points on an imaginary coordinate grid remains constant as the universe expands) are simply related by the time-dependent scale factor $x^i_{\text{phys}} = a(t)x^i$. Then the physical velocity of an object is then

$$v^i_{\text{phys}} = \frac{\mathrm{d}x^i_{\text{phys}}}{\mathrm{d}t} = a(t)\frac{\mathrm{d}x^i}{\mathrm{d}t} + \frac{\mathrm{d}a}{\mathrm{d}t}x^i \equiv v^i_{\text{pec}} + Hx^i_{\text{phys}}, \tag{1.14}$$

where we have defined the *Hubble parameter*

$$H \equiv \frac{\dot{a}}{a}. \tag{1.15}$$

We see that the physical velocity has two contributions: the so-called *peculiar velocity*, $v^i_{\text{pec}} \equiv a(t)\dot{x}^i$, and the *Hubble flow*, Hx^i_{phys}. The peculiar velocity of an object is the velocity measured by a comoving observer, i.e. an observer who follows the Hubble flow.

It is also noted that the complicated γ_{rr} component of Eq. (1.12) can sometimes be inconvenient. In that case, we may redefine the radial coordinate as (comoving distance infinitesimal interval)

$$\mathrm{d}\chi \equiv \frac{\mathrm{d}r}{\sqrt{1-kr^2}}. \tag{1.16}$$

Then the line element can be written as

$$\mathrm{d}\ell^2 = a^2\left[\mathrm{d}\chi^2 + S_k^2(\chi)\mathrm{d}\Omega^2\right], \tag{1.17}$$

where

$$S_k(\chi) = \begin{cases} \sinh\chi, & k=-1 \\ \chi, & k=0 \\ \sin\chi, & k=+1 \end{cases}. \tag{1.18}$$

Then the FRW metric is just

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)\left[\mathrm{d}\chi^2 + S_k^2(\chi)\mathrm{d}\Omega^2\right]. \tag{1.19}$$

This form of the metric is particularly convenient for studying the propagation of light. We can introduce *conformal time*,

$$\mathrm{d}\tau = \frac{\mathrm{d}t}{a(t)} \tag{1.20}$$

so that

$$\mathrm{d}s^2 = a^2(\tau)\big[\underbrace{-\mathrm{d}\tau^2 + \mathrm{d}\chi^2 + S_k^2(\chi)\mathrm{d}\Omega^2}_{\text{static Minkowski metric}}\big]. \tag{1.21}$$

Since light travels along the null geodesics, $\mathrm{d}s^2 = 0$, the propagation of light in FRW is the same as in Minkowski space if we first transform to conformal time. Along the path (i.e. the line of sight, defined by $\mathrm{d}\phi = 0 = \mathrm{d}\theta$), the change in conformal time equals the change in comoving distance,

$$\mathrm{d}s^2 \overset{!}{=} 0 = a^2(\tau)\left(-\mathrm{d}\tau^2 + \mathrm{d}\chi^2\right) \qquad \Longrightarrow \qquad \Delta\tau = \Delta\chi. \tag{1.22}$$

This also implies that

$$\frac{\mathrm{d}t}{a(t)} = \frac{\mathrm{d}r}{\sqrt{1-kr^2}}, \tag{1.23}$$

so the comoving distance can also be computed by

$$\chi = \int \frac{\mathrm{d}r}{\sqrt{1-kr^2}} = \int \frac{\mathrm{d}t}{a(t)} = \int \mathrm{d}a\,\frac{1}{\dot{a}(t)a(t)}. \tag{1.24}$$

Using the Friedmann equations which we will develop later, we can rewrite this integral in terms of energy densities describing today's Universe, as well as today's Hubble constant. Note that the comoving distance factors out the expansion of the Universe, giving a distance that does not change in time due to the expansion.

We will focus on a flat universe ($k = 0$) with the metric (in Cartesian coordinates)

$$g_{\mu\nu} = \begin{pmatrix} -1 & & & \\ & a(t)^2 & & \\ & & a(t)^2 & \\ & & & a(t)^2 \end{pmatrix}. \tag{1.25}$$

1.2 Kinematics

1.2.1 Geodesics

How do particles evolve in the FRW spacetime? In the absence of additional non-gravitational forces, freely-falling particles in a curved spacetime move along *geodesics*. A geodesic is a curve which extremizes the proper time $\Delta s/c$ between two points in spacetime. Consider a particle of mass m that traces out a path $X^\mu(s)$. The 4-velocity of the particle is defined by

$$U^\mu = \frac{\mathrm{d}X^\mu}{\mathrm{d}s}. \tag{1.26}$$

We show that the extremal path satisfies the *geodesic equation*

$$\frac{\mathrm{d}U^\mu}{\mathrm{d}s} + \Gamma^\mu_{\alpha\beta}U^\alpha U^\beta = 0, \tag{1.27}$$

where $\Gamma^{\mu}_{\alpha\beta}$ is the *Christoffel symbols* defined as

$$\Gamma^{\mu}_{\alpha\beta} = \frac{g^{\mu\nu}}{2}\left[\frac{\partial g_{\alpha\nu}}{\partial X^{\beta}} + \frac{\partial g_{\beta\nu}}{\partial X^{\alpha}} - \frac{\partial g_{\alpha\beta}}{\partial X^{\nu}}\right] \equiv \frac{g^{\mu\nu}}{2}[\partial_\beta g_{\alpha\nu} + \partial_\alpha g_{\beta\nu} - \partial_\nu g_{\alpha\beta}], \tag{1.28}$$

where we use the notation $\partial_\mu \equiv \partial/\partial X^\mu$.

Let us start by considering the *relativistic* action of the particle moving from point A to B along a timelike path in spacetime:

$$S = -m\int_A^B \mathrm{d}s. \tag{1.29}$$

We parameterize the path by λ[1] that increases monotonically from an initial value $\lambda(A) = 0$ to a final value $\lambda(B) = 1$. The action is a functional of the path $X^\mu(\lambda)$,

$$S[X^\mu(\lambda)] = -m\int_0^1 \sqrt{-g_{\mu\nu}(X^\mu)\frac{\mathrm{d}X^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}X^\nu}{\mathrm{d}\lambda}}\mathrm{d}\lambda \equiv \int_0^1 L(X^\mu, \dot{X}^\mu)\mathrm{d}\lambda, \tag{1.30}$$

where the Lagrangian is

$$L(X^\mu, \dot{X}^\mu) = -m\sqrt{-g_{\mu\nu}(X^\mu)\dot{X}^\mu\dot{X}^\nu} \tag{1.31}$$

and $\dot{X}^\mu \equiv \partial X^\mu/\partial\lambda$. Then the Lagrangian satisfies the Euler-Lagrange equation

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\partial L}{\partial \dot{X}^\mu}\right) - \frac{\partial L}{\partial X^\mu} = 0. \tag{1.32}$$

The derivatives are calculated to be

$$\frac{\partial L}{\partial \dot{X}^\mu} = -\frac{m^2}{L}g_{\mu\nu}\dot{X}^\nu, \qquad \frac{\partial L}{\partial X^\mu} = -\frac{m^2}{2L}\partial_\mu g_{\nu\rho}\dot{X}^\nu\dot{X}^\rho. \tag{1.33}$$

Before continuing, it is convenient to switch from the general parameterization λ to the parameterization using the proper time s. (We could not have used s from the beginning since the value of s at B is different for different curves. The range of integration would then have been different for different curves.) We have

$$\frac{\mathrm{d}s}{\mathrm{d}\lambda} = \sqrt{-g_{\mu\nu}\dot{X}^\mu\dot{X}^\nu} = -\frac{L}{m}. \tag{1.34}$$

Then we can replace $\mathrm{d}/\mathrm{d}\lambda$ by $\mathrm{d}/\mathrm{d}s$ via

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} = \frac{\mathrm{d}s}{\mathrm{d}\lambda}\frac{\mathrm{d}}{\mathrm{d}s} = -\frac{L}{m}\frac{\mathrm{d}}{\mathrm{d}s}. \tag{1.35}$$

Together with Eq. (1.33), the E-L equation (1.32) can be written as

$$\begin{aligned} &-\frac{L}{m}\frac{\mathrm{d}}{\mathrm{d}s}\left[-\frac{m^2}{L}g_{\mu\nu}\cdot\left(-\frac{L}{m}\right)\frac{\mathrm{d}X^\nu}{\mathrm{d}s}\right] + \frac{m^2}{2L}\partial_\mu g_{\nu\rho}\left(-\frac{L}{m}\frac{\mathrm{d}X^\nu}{\mathrm{d}s}\right)\left(-\frac{L}{m}\frac{\mathrm{d}X^\rho}{\mathrm{d}s}\right) = 0 \\ \Longrightarrow\quad &\frac{\partial}{\partial s}\left(g_{\mu\nu}\frac{\mathrm{d}X^\nu}{\mathrm{d}s}\right) - \frac{1}{2}\partial_\mu g_{\nu\rho}\frac{\mathrm{d}X^\rho}{\mathrm{d}s}\frac{\mathrm{d}X^\nu}{\mathrm{d}s} = 0 \\ \Longrightarrow\quad &g_{\mu\nu}\frac{\mathrm{d}^2X^\mu}{\mathrm{d}s^2} + \partial_\rho g_{\mu\nu}\frac{\mathrm{d}X^\rho}{\mathrm{d}s}\frac{\mathrm{d}X^\nu}{\mathrm{d}s} - \frac{1}{2}\partial_\mu g_{\nu\rho}\frac{\mathrm{d}X^\rho}{\mathrm{d}s}\frac{\mathrm{d}X^\nu}{\mathrm{d}s} = 0. \end{aligned} \tag{1.36}$$

In the second term we can replace $\partial_\rho g_{\mu\nu}$ with $\frac{1}{2}(\partial_\rho g_{\mu\nu} + \partial_\nu g_{\mu\rho})$ because it is contracted with an object that is symmetric in ν and ρ. Contracting the last equation in (1.36) with the inverse metric

[1] s is the proper time and we can think of λ as the coordinate time t and they are related by $\mathrm{d}s = \mathrm{d}\lambda/\gamma$.

and relabelling indices, we will obtain the geodesic equation

$$\frac{\mathrm{d}^2 X^\mu}{\mathrm{d}s^2} + \Gamma^\mu_{\alpha\beta}\frac{\mathrm{d}X^\alpha}{\mathrm{d}s}\frac{\mathrm{d}X^\beta}{\mathrm{d}s} = 0, \tag{1.37}$$

with the Christoffel symbol given by Eq. (1.28).

The derivative term in the geodesic equation (1.27) can be manipulated by using the chain rule

$$\frac{\mathrm{d}}{\mathrm{d}s}U^\mu(X^\alpha(s)) = \frac{\mathrm{d}X^\alpha}{\mathrm{d}s}\frac{\partial U^\mu}{\partial X^\alpha} = U^\alpha \frac{\partial U^\mu}{\partial X^\alpha}, \tag{1.38}$$

so we get

$$U^\alpha\left(\frac{\partial U^\mu}{\partial X^\alpha} + \Gamma^\mu_{\alpha\beta}U^\beta\right) = 0. \tag{1.39}$$

The term in the bracket is called the *covariant derivative* of U^μ

$$\nabla_\alpha U^\mu \equiv \partial_\alpha U^\mu + \Gamma^\mu_{\alpha\beta}U^\beta. \tag{1.40}$$

Then the geodesic equation can be written in a more compact way:

$$U^\alpha \nabla_\alpha U^\mu = 0. \tag{1.41}$$

In terms of the 4-momentum of the particle, $P^\mu = mU^\mu = (E, \vec{P})$, it reads

$$P^\alpha \frac{\partial P^\mu}{\partial X^\alpha} = -\Gamma^\mu_{\alpha\beta}P^\alpha P^\beta, \tag{1.42}$$

which is also valid for massless particles.

1.2.2 Geodesic motion in FRW

Using the FRW metric (1.13) in the Christoffel symbol defined in Eq. (1.28), we find that all Christoffel symbols with two time indices (0-index) vanish, i.e. $\Gamma^\mu_{00} = \Gamma^0_{0\beta} = 0$. The only non-zero components are

$$\Gamma^0_{ij} = a\dot{a}\gamma_{ij}, \qquad \Gamma^i_{0j} = \frac{\dot{a}}{a}\delta^i_j, \qquad \Gamma^i_{jk} = \frac{1}{2}\gamma^{il}(\partial_j\gamma_{kl} + \partial_k\gamma_{jl} - \partial_l\gamma_{jk}), \tag{1.43}$$

and those which are related to these by symmetry (note that $\Gamma^\mu_{\alpha\beta} = \Gamma^\mu_{\beta\alpha}$). For example, let us look at the Christoffel symbol with upper index equal to zero,

$$\begin{aligned}\Gamma^0_{\alpha\beta} &= \frac{1}{2}g^{0\nu}(\partial_\alpha g_{\beta\nu} + \partial_\beta g_{\alpha\nu} - \partial_\nu g_{\alpha\beta}) \\ &= -\frac{1}{2}(\partial_\alpha g_{\beta 0} + \partial_\beta g_{\alpha 0} - \partial_0 g_{\alpha\beta})\end{aligned} \tag{1.44}$$

because the factor $g^{0\nu}$ vanishes unless $\nu = 0$ in which case it is equal to -1. Note that the first two terms reduce to derivatives of g_{00} since $g_{i0} = 0$. The FRW metric has constant g_{00}, so these two terms vanish and we are left with

$$\Gamma^0_{\alpha\beta} = \frac{1}{2}\partial_0 g_{\alpha\beta}. \tag{1.45}$$

The derivative is non-zero only if α and β are spatial indices, $g_{ij} = a^2(t)\gamma_{ij}$. Then we find

$$\Gamma^0_{ij} = a\dot{a}\gamma_{ij}. \tag{1.46}$$

The homogeneity of the FRW universe implies that $\partial_i P^\mu = 0$, so the geodesic equation (1.42) reduces to

$$P^0 \frac{dP^\mu}{dt} = -\Gamma^\mu_{\alpha\beta} P^\alpha P^\beta = -\left(2\Gamma^\mu_{0j} P^0 + \Gamma^\mu_{ij} P^i\right) P^j, \tag{1.47}$$

where the factor of 2 arises due to the symmetry of the Christoffel symbol, $\Gamma^\mu_{0j} = \Gamma^\mu_{j0}$.

- The first thing to notice is that massive particles at rest in the comoving frame, $P^j = 0$ (remember U^μ is defined as the derivative of X^μ w.r.t. the proper time s), will stay at rest:

$$P^j = 0 \qquad \Longrightarrow \qquad \frac{dP^i}{dt} = 0. \tag{1.48}$$

- Next, we consider $\mu = 0$ in Eq. (1.47), but don't require the particles to be at rest. The first term on the RHS vanishes because $\Gamma^0_{0j} = 0$. Then we have

$$P^0 \frac{dP^0}{dt} = E\frac{dE}{dt} = -\Gamma^0_{ij} P^i P^j = -\frac{\dot{a}}{a} p^2, \tag{1.49}$$

where we define the amplitude of the physical 3-momentum as

$$p^2 \equiv g^{ij} P^i P^j = a^2 \gamma_{ij} P^i P^j. \tag{1.50}$$

The components of the 4-momentum satisfy the constraint $g_{\mu\nu} P^\mu P^\mu = m^2$, or $E^2 - p^2 = m^2$. So we have $E\,dE = p\,dp$, then Eq. (1.49) can be written as

$$\frac{\dot{p}}{p} = -\frac{\dot{a}}{a} \qquad \Longrightarrow \qquad p \propto \frac{1}{a}. \tag{1.51}$$

This indicates that the physical 3-momentum of any particle (both massive and massless) decays with the expansion of the universe. In particular, the energy of massless particles decays as $E \propto a(t)^{-1}$.

1.2.3 Redshift

Everything we know about the universe is inferred from the light we receive from distant objects. To interpret the observations correctly, we need to take into account that the wavelength of the light gets stretched (or, equivalently, the photons lose energy) by the expansion of the universe.

We consider, quantum mechanically, the redshifting of photons. The wavelength of light is inversely proportional to the photon momentum,

$$\lambda = \frac{h}{p}. \tag{1.52}$$

Since the momentum of a photon evolves as $a(t)^{-1}$, the wavelength scales as $a(t)$. Light emitted at time t_1 with wavelength λ_1 will be observed at t_0 with wavelength

$$\lambda_0 = \frac{a(t_0)}{a(t_1)} \lambda_1. \tag{1.53}$$

Since $a(t_0) > a(t_1)$ ($t_0 > t_1$), the wavelength of light increases, i.e. light is redshifted. The *redshift parameter* is defined as the fractional shift in wavelength

$$z \equiv \frac{\lambda_0 - \lambda_1}{\lambda_1} \tag{1.54}$$

and thus

$$1 + z = \frac{a(t_0)}{a(t_1)}. \tag{1.55}$$

Using the convention $a(t_0) = 1$, the redshift is then directly related to the scale factor of the emitter,

$$1 + z = \frac{1}{a(t_1)}. \tag{1.56}$$

For nearby sources (so z is small), we can expand $a(t_1)$ in a power series,

$$a(t_1) = a(t_0)[1 + (t_1 - t_0)H_0 + \cdots], \tag{1.57}$$

where H_0 is the Hubble constant

$$H_0 \equiv \frac{\dot{a}(t_0)}{a(t_0)}. \tag{1.58}$$

Using this expansion in Eq. (1.55) gives

$$z \simeq H_0(t_0 - t_1) = H_0 d, \tag{1.59}$$

where $t_0 - t_1$ is simply the physical distance d for closed objects. Therefore, the redshift increases linearly with distance.

1.3 Dynamics

1.3.1 Einstein equations

The dynamics of the universe is determined by the Einstein equation

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu}, \tag{1.60}$$

where G is Newton's constant, $T_{\mu\nu}$ is the energy-momentum tensor, $G_{\mu\nu}$ is the Einstein tensor, $R_{\mu\nu}$ is called the Ricci tensor and R is the Ricci scalar. The Ricci tensor is obtained from the Christoffel symbols

$$R_{\mu\nu} = \partial_\alpha \Gamma^\alpha_{\mu\nu} - \partial_\nu \Gamma^\alpha_{\mu\alpha} + \Gamma^\alpha_{\beta\alpha}\Gamma^\beta_{\mu\nu} - \Gamma^\alpha_{\beta\nu}\Gamma^\beta_{\mu\alpha}, \tag{1.61}$$

and the Ricci scalar is defined as the contraction of the Ricci tensor with the metric

$$R = g^{\mu\nu} R_{\mu\nu}. \tag{1.62}$$

We first look at the 00 component of the Ricci tensor in a flat FRW metric (1.25), which is given by [cf. Eq. (1.43) for non-zero Christoffel symbols for an FRW metric]

$$\begin{aligned} R_{00} &= \partial_\alpha \Gamma^\alpha_{00} - \partial_0 \Gamma^\alpha_{0\alpha} + \Gamma^\alpha_{\beta\alpha}\Gamma^\beta_{00} - \Gamma^\alpha_{\beta 0}\Gamma^\beta_{0\alpha} \\ &= -\partial_0 \Gamma^i_{0i} - \Gamma^i_{j0}\Gamma^j_{0i} \\ &= -\frac{\partial}{\partial t}\left(\frac{\dot{a}}{a}\delta^i_i\right) - \left(\frac{\dot{a}}{a}\right)^2 \delta^i_j \delta^j_i \\ &= -\frac{\partial}{\partial t}\left(3\frac{\dot{a}}{a}\right) - 3\left(\frac{\dot{a}}{a}\right)^2 = -3\frac{\ddot{a}}{a}. \end{aligned} \tag{1.63}$$

Similarly, the spatial components of the Ricci tensor are

$$R_{ij} = \delta_{ij}\left(2\dot{a}^2 + a\ddot{a}\right). \tag{1.64}$$

Then the Ricci scalar can be evaluated as

$$R = g^{\mu\nu} R_{\mu\nu} = -R_{00} + \frac{1}{a^2} R_{ii} = 6\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2\right]. \tag{1.65}$$

Plugging everything into the Einstein equations, we obtain two independent equations describing the evolution of a flat ($k = 0$) FRW universe,

$$\begin{aligned} R_{00} - \frac{1}{2} g_{00} R &= 3\left(\frac{\dot{a}}{a}\right)^2 = 3H^2 = 8\pi G T_{00}, \\ g^{\mu\nu} G_{\mu\nu} &= -R = -6\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2\right] = 8\pi G T^{\mu}_{\mu}. \end{aligned} \tag{1.66}$$

1.3.2 Friedmann equations

Now we make the assumption that we can describe the content of the Universe by different *perfect fluids* as a leading approximation, i.e. the fluid can be described by macroscopic quantities, its energy density and pressure, while there is no stress or viscosity, in agreement with the metric being homogeneous and isotropic. In this case, the energy-momentum tensor of a perfect fluid in its rest-frame in the Minkowski space is given by

$$T^{\mu\nu} = \begin{pmatrix} \rho & & & \\ & \mathcal{P} & & \\ & & \mathcal{P} & \\ & & & \mathcal{P} \end{pmatrix}. \tag{1.67}$$

For a particle at rest, its four-velocity is

$$U^{\mu} = \frac{\mathrm{d}x^{\mu}}{\mathrm{d}\tau} = (1, 0, 0, 0), \tag{1.68}$$

and thus the energy-momentum tensor can be written in terms of this four-velocity as

$$T^{\mu\nu} = (\rho + \mathcal{P}) U^{\mu} U^{\nu} + \mathcal{P} \eta^{\mu\nu}. \tag{1.69}$$

Its generalization to general relativity is straightforward,

$$T^{\mu\nu} = (\rho + \mathcal{P}) U^{\mu} U^{\nu} + \mathcal{P} g^{\mu\nu}. \tag{1.70}$$

Thus, in the rest frame of the fluid in the FRW universe,

$$T^{\mu\nu} = \begin{pmatrix} \rho & & & \\ & a^{-2}\mathcal{P} & & \\ & & a^{-2}\mathcal{P} & \\ & & & a^{-2}\mathcal{P} \end{pmatrix} \quad \text{or} \quad T^{\mu}_{\nu} = \begin{pmatrix} -\rho & & & \\ & \mathcal{P} & & \\ & & \mathcal{P} & \\ & & & \mathcal{P} \end{pmatrix}. \tag{1.71}$$

Therefore, we have

$$T_{00} = \rho, \qquad T^{\mu}_{\mu} = -\rho + 3\mathcal{P}. \tag{1.72}$$

Plugging them into the two equations in (1.66), we obtain the *Friedmann equations*:

$$\begin{aligned} H^2 = \left(\frac{\dot{a}}{a}\right)^2 &= \frac{8\pi G}{3}\rho, \\ \frac{\ddot{a}}{a} &= -\frac{4\pi G}{3}(\rho + 3\mathcal{P}). \end{aligned} \tag{1.73}$$

A flat universe ($k = 0$) corresponds to the critical density

$$\rho_{\text{crit}} = \frac{3H^2}{8\pi G}. \tag{1.74}$$

Then the dimensionless density parameter can be expressed as

$$\Omega = \frac{\rho}{\rho_{\text{crit}}}. \tag{1.75}$$

1.3.3 Continuity equation and cosmic inventory

The next question is: how does the energy-momentum tensor of this perfect fluid evolve with time? In the absence of forces and gravity, we have the continuity equation,

$$\partial_\mu T^\mu_\nu = 0. \tag{1.76}$$

The generalization to GR is by replacing the partial derivative with a covariant derivative to ensure the continuity equation correctly transforms under a change of coordinates:

$$0 = \nabla_\mu T^\mu_\nu = \partial_\mu T^\mu_\nu + \Gamma^\mu_{\alpha\mu} T^\alpha_\nu - \Gamma^\alpha_{\nu\mu} T^\mu_\alpha. \tag{1.77}$$

For $\nu = 0$, Eq. (1.77) becomes

$$0 = \partial_\mu T^\mu_0 + \Gamma^\mu_{\alpha\mu} T^\alpha_0 - \Gamma^\alpha_{0\mu} T^\mu_\alpha = -\frac{\partial\rho}{\partial t} + \Gamma^i_{i0} T^0_0 - \Gamma^i_{0i} T^i_i, \tag{1.78}$$

and thus,

$$\frac{\partial\rho}{\partial t} + 3\frac{\dot{a}}{a}(\rho + \mathcal{P}) = 0, \quad \text{or} \quad a^{-3}\frac{\partial\left[\rho a^3\right]}{\partial t} = -\frac{\dot{a}}{a}\mathcal{P}. \tag{1.79}$$

Introducing the *equation of state parameter*

$$w \equiv \frac{\mathcal{P}}{\rho}, \tag{1.80}$$

the continuity equation (1.79) for $\nu = 0$ can be rewritten as

$$0 = \frac{\partial\rho}{\partial t} + 3(1+w)\frac{\dot{a}}{a}\rho = a^{-3(1+w)}\frac{\partial\left[\rho a^{3(1+w)}\right]}{\partial t}. \tag{1.81}$$

Therefore, we conclude that $\rho \propto a^{-3(1+w)}$. Inserting this result into the first Friedmann equation, it gives us the time dependence of the scale factor for $w \neq 1$:

$$a(t) \propto t^{\frac{2}{3(1+w)}}, \quad \text{or} \quad H(t) = \frac{2}{3(1+w)t}. \tag{1.82}$$

For $\rho = \Lambda/8\pi G$ with $w = -1$, the time dependence is

$$a(t) \propto e^{\sqrt{\Lambda/3}\,t}, \quad \text{or} \quad H = \sqrt{\Lambda/3}. \tag{1.83}$$

Finally, we give a summary table of the evolution of the three main components of the universe. Note that today, the most dominant contribution is dark energy (or a cosmological constant), which is responsible for the accelerating expansion of the universe.

Table 1.1: Evolution of different fluids.

	w	$\rho(a)$	$a(t)$	$H(t)$
matter	0	a^{-3}	$t^{2/3}$	$\frac{2}{3t}$
radiation	$\frac{1}{3}$	a^{-4}	$t^{1/2}$	$\frac{1}{2t}$
cosm. const.	-1	ρ_0	$e^{\sqrt{\Lambda/3}\,t}$	$\sqrt{\frac{\Lambda}{3}}$

1.3.4 General Friedmann equations

Here we consider the general FRW metric tensor (1.13) (written in spherical coordinates)

$$g_{\mu\nu} = \begin{pmatrix} -1 & & & \\ & \frac{a^2}{1-kr^2} & & \\ & & a^2r^2 & \\ & & & a^2r^2\sin^2\theta \end{pmatrix}. \tag{1.84}$$

So

$$g^{\mu\nu} = \begin{pmatrix} -1 & & & \\ & \frac{1-kr^2}{a^2} & & \\ & & \frac{1}{a^2r^2} & \\ & & & \frac{1}{a^2r^2\sin^2\theta} \end{pmatrix}. \tag{1.85}$$

The nonvanishing Ricci tensors are

$$\begin{aligned} R_{00} &= -3\frac{\ddot{a}}{a}; \\ R_{11} &= \frac{a\ddot{a}+2\dot{a}^2+2k}{1-kr^2}; \\ R_{22} &= r^2(a\ddot{a}+2\dot{a}+2k); \\ R_{33} &= r^2(a\ddot{a}+2\dot{a}+2k)\sin^2\theta, \end{aligned} \tag{1.86}$$

and the Ricci scalar is

$$R = g^{\mu\nu}R_{\mu\nu} = 6\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2}\right]. \tag{1.87}$$

We also consider the Einstein equation with a cosmological constant Λ:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}. \tag{1.88}$$

Again, the energy-momentum tensors can be derived via

$$\begin{aligned} T^{\mu\nu} &= (\rho+\mathcal{P})U^\mu U^\nu + \mathcal{P}g^{\mu\nu} \\ &= \begin{pmatrix} \rho & & & \\ & \frac{1-kr^2}{a^2}\mathcal{P} & & \\ & & \frac{1}{a^2r^2}\mathcal{P} & \\ & & & \frac{1}{a^2r^2\sin^2\theta}\mathcal{P} \end{pmatrix}, \end{aligned} \tag{1.89}$$

and so

$$T^\mu_\nu = T^{\mu\rho} g_{\rho\nu} = \begin{pmatrix} -\rho & & & \\ & \mathcal{P} & & \\ & & \mathcal{P} & \\ & & & \mathcal{P} \end{pmatrix},$$
$$T_{\mu\nu} = g_{\mu\rho} T^\rho_\nu = \begin{pmatrix} \rho & & & \\ & \frac{a^2}{1-kr^2}\mathcal{P} & & \\ & & a^2 r^2 \mathcal{P} & \\ & & & a^2 r^2 \sin^2\theta \mathcal{P} \end{pmatrix}. \tag{1.90}$$

If we now consider the 00 component of the Einstein equation, we get the first Friedmann equation:

$$\begin{aligned} & R_{00} - \frac{1}{2} g_{00} R + \Lambda g_{00} = 8\pi G T_{00} \\ \implies \quad & -3\frac{\ddot{a}}{a} + 3\left(\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right) - \Lambda = 8\pi G \rho \\ \implies \quad & \boxed{\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3}.} \end{aligned} \tag{1.91}$$

Similarly, if we consider the 11 (or equivalently 22/33) component of the Einstein equation, we obtain the other equation:

$$\begin{aligned} & R_{11} - \frac{1}{2} g_{11} R + \Lambda g_{11} = 8\pi G T_{11} \\ \implies \quad & \frac{a\ddot{a} + 2\dot{a}^2 + 2k}{1-kr^2} - 3\frac{a^2}{1-kr^2}\left(\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right) + \Lambda \frac{a^2}{1-kr^2} = 8\pi G \frac{a^2}{1-kr^2}\mathcal{P} \\ \implies \quad & \boxed{2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = -8\pi G \mathcal{P} + \Lambda.} \end{aligned} \tag{1.92}$$

Combining these two equations, we easily get

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3\mathcal{P}) + \frac{\Lambda}{3}. \tag{1.93}$$

Note that the cosmological constant term can be omitted if we make the following replacements:

$$\begin{aligned} \rho &\to \rho + \frac{\Lambda}{8\pi G}, \\ \mathcal{P} &\to \mathcal{P} - \frac{\Lambda}{8\pi G}. \end{aligned} \tag{1.94}$$

The first replacement can be understood by noting that the dark energy density can be written as

$$\rho_\Lambda(t) \equiv \Lambda(t) M_{\text{Pl}}^2 = \frac{\Lambda(t)}{8\pi G}. \tag{1.95}$$

So ρ is now the total energy density:

$$\rho(t) = \rho_m(t) + \rho_r(t) + \rho_\Lambda(t), \tag{1.96}$$

and the first Friedmann equation just becomes

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi G}{3}\rho \equiv \frac{8\pi G}{3}[\rho_m(t) + \rho_r(t) + \rho_\Lambda(t)]. \tag{1.97}$$

By defining dimensionless density parameters for today's Universe,

$$\Omega_{i,0} \equiv \frac{\rho_{i,0}}{\rho_{\text{crit},0}}, \quad \text{with} \quad \rho_{\text{crit},0} = \frac{3H_0^2}{8\pi G}, \tag{1.98}$$

Eq. (1.97) can be written as

$$H^2(a) = H_0^2\left[\Omega_{r,0}\left(\frac{a_0}{a}\right)^4 + \Omega_{m,0}\left(\frac{a_0}{a}\right)^3 + \Omega_{k,0}\left(\frac{a_0}{a}\right)^2 + \Omega_{\Lambda,0}\right], \tag{1.99}$$

where we have defined a "curvature" density parameter,

$$\Omega_{k,0} \equiv -\frac{k}{a_0^2 H_0^2}. \tag{1.100}$$

Observations show that

$$|\Omega_k| \leq 0.01, \quad \Omega_r = 9.4 \times 10^{-5}, \quad \Omega_m = 0.32, \quad \Omega_\Lambda = 0.68. \tag{1.101}$$

Moreover, the matter splits into 5% ordinary matter (baryons) and 27% (cold) dark matter:

$$\Omega_b = 0.05, \quad \Omega_{\text{CDM}} = 0.27. \tag{1.102}$$

Similarly, the second replacement in (1.94) can be understood by interpreting the cosmological constant as arising from a form of energy which has negative pressure, equal in magnitude to its (positive) energy density ($\mathcal{P}_\Lambda = -\rho_\Lambda$, or $w = -1$).

Chapter 2

Thermal History

In this chapter, we want to describe the first three minutes in the history of the universe, starting from the hot and dense state following inflation. At early times, the thermodynamical properties of the universe were determined by local equilibrium. However, it is the departure from thermal equilibrium that makes life interesting. As we will see, non-equilibrium dynamics allows massive particles to acquire cosmological abundances and therefore explains why there is something rather than nothing. Deviations from equilibrium are also crucial for understanding the origin of the cosmic microwave background and the formation of the light chemical elements.

2.1 Equilibrium thermodynamics

We have good observational evidence (from the perfect blackbody spectrum of the CMB) that the early universe was in *local thermal equilibrium*. The key to understanding the thermal history of the universe is the comparison between the following two quantities: the rate of expansion (Hubble rate) $H = \dot{a}/a$, which corresponds to the characteristic time scale $t_H = H^{-1}$, and the rate of interactions Γ, which corresponds to the time scale for particle interactions $t_C = \Gamma^{-1}$. If $t_C \ll t_H$, i.e. $\Gamma \gg H$, then local thermal equilibrium is reached before the effect of expansion becomes relevant. If $t_C \sim t_H$, or $\Gamma \sim H$, then particles decouple from the thermal bath ("*freeze-out*"). Different particle species may have different interaction rates and so may decouple at different times.

For example, let us look at a $2+2$ scattering process, of the form $1+2 \leftrightarrow 3+4$. Assuming that at high energies $n_1 \sim n_2 \equiv n$, the interaction rate is given by

$$\Gamma = n\sigma v, \tag{2.1}$$

where n is the number density of the particles, σ is the interaction cross section and v is their average velocity. At high temperature $T \gg m$, usually $T \gtrsim 100$ GeV, all known particles are ultra-relativistic ($v \sim 1$), we neglect the mass and by dimensional analysis,

$$n \sim T^3. \tag{2.2}$$

Interactions are mediated by gauge bosons, which are massless above the scale of electroweak symmetry breaking. The cross sections for the strong and electroweak interactions then have a similar dependence, which also can be estimated using dimensional analysis

$$\sigma \sim \frac{\alpha^2}{T^2}, \tag{2.3}$$

where $\alpha \equiv g_A^2/4\pi$ is the generalized structure constant associated with the gauge boson A. Then the interaction rate

$$\Gamma \sim T^3 \frac{\alpha^2}{T^2} = \alpha^2 T. \tag{2.4}$$

Similarly, dimensional argument gives

$$\rho \sim T^4 \tag{2.5}$$

and hence the Hubble rate

$$H^2 = \frac{8\pi G}{3}\rho = \frac{\rho}{3M_{Pl}^2} \qquad \Longrightarrow \qquad H \sim \frac{\sqrt{\rho}}{M_{Pl}} \sim \frac{T^2}{M_{Pl}}, \tag{2.6}$$

where $M_{Pl} = 1/\sqrt{8\pi G}$ is the reduced Planck mass. Then the ratio of interaction rate and expansion rate is just

$$\frac{\Gamma}{H} \sim \frac{\alpha^2 M_{Pl}}{T} \sim \frac{10^{16}\ \text{GeV}}{T} \qquad (\text{for } \alpha \sim 0.01). \tag{2.7}$$

This means that the particles in the SM are in thermal equilibrium for 100 GeV $\lesssim T \lesssim 10^{16}$ GeV, i.e. the condition $t_C \ll t_H$ is satisfied. Below $T \sim 100$ GeV, which is the scale of electroweak symmetry breaking, $W^{\pm}$ and Z^0 receive their masses. This leads to a drastic change in the strength of the weak interaction. This is known as the *electroweak phase transition.* Alternatively, one can compare the *mean free path* (see App. A) of a particle species, $1/(\sigma n)$, to the Hubble length, v/H, to define the point of decoupling implicitly as:

$$\frac{1}{\sigma(T_{\text{dec}})n(T_{\text{dec}})} \overset{!}{\sim} \frac{v}{H(T_{\text{dec}})} \qquad \Longrightarrow \qquad \frac{n(T_{\text{dec}})\sigma(T_{\text{dec}})v}{H(T_{\text{dec}})} \overset{!}{\sim} 1. \tag{2.8}$$

This is exactly the same condition as $\Gamma(T_{\text{dec}})/H(T_{\text{dec}}) \sim 1$.

In local thermal equilibrium, the main object of interest is the *distribution function*

$$f(\mathbf{x}, \mathbf{p}, t) \simeq \frac{\text{number of particles}}{\text{volume in space and momentum}} = \frac{\text{number of particles}}{\mathrm{d}^3x\mathrm{d}^3p/(2\pi\hbar)^3}. \tag{2.9}$$

In an homogeneous and isotropic universe, the distribution function does not depend on the position $\mathbf{x}$ and the direction of the momentum, but only on the absolute magnitude of the momentum. For example, the number density of species i is given as

$$n_i = g\int \frac{\mathrm{d}^3x\mathrm{d}^3p}{(2\pi\hbar)^3} f(\mathbf{x}, \mathbf{p}) = g\int \frac{\mathrm{d}^3p}{(2\pi)^3} f(p) \tag{2.10}$$

where $\hbar = 1$, and g is the internal degrees of freedom.

A system of particles is said to be in *kinetic equilibrium* if the particles exchange energy and momentum efficiently. This leads to a state of maximum entropy in which the distribution functions are given by the usual Bose-Einstein and Fermi-Dirac distributions:

$$f(p) = \frac{1}{e^{(E(p)-\mu)/T} \mp 1} \tag{2.11}$$

where the $-$ sign is for bosons, the $+$ sign is for fermions and μ is the chemical potential (characterizing the response of a system to a change in particle number). At low temperatures, $T < E - \mu$, both distributions reduce to the Maxwell-Boltzmann distribution

$$f(p) \approx e^{(E(p)-\mu)/T}. \tag{2.12}$$

At early times, the chemical potentials of all particles are so small that they can be neglected. Then the number density can be evaluated:

$$n_{\text{eq}} = g\int \frac{\mathrm{d}^3p}{(2\pi)^3} f(p) = 4\pi g \int_0^\infty \frac{E\mathrm{d}E}{(2\pi)^3} \frac{\sqrt{|E^2 - m^2|}}{\mathrm{e}^{E/T} \mp 1} \xrightarrow{T \gg m,\mu} \begin{cases} \frac{\zeta(3)}{\pi^2} gT^3 & \text{bosons} \\ \frac{3}{4}\frac{\zeta(3)}{\pi^2} gT^3 & \text{fermions} \end{cases}, \tag{2.13}$$

where the Riemann zeta function has the value $\zeta(3) = 1.2$. The distribution functions can also be used to obtain the energy density

$$\rho_{\rm eq} = g \int \frac{\mathrm{d}^3 p}{(2\pi)^3} f(p) E(p) \xrightarrow{T \gg m,\mu} \begin{cases} \frac{\pi^2}{30} g T^4 & \text{bosons} \\ \frac{7}{8}\frac{\pi^2}{30} g T^4 & \text{fermions} \end{cases}. \tag{2.14}$$

Similarly the pressure can be written as

$$\mathcal{P}_{\rm eq} = g \int \frac{\mathrm{d}^3 p}{(2\pi)^3} f(p) \frac{p^2}{3E(p)} \xrightarrow{T \gg m,\mu} \frac{1}{3}\rho. \tag{2.15}$$

In the non-relativistic limit (low temperatures $T \ll m$), the results are

$$n_{\rm eq} = g\left(\frac{mT}{2\pi}\right)^{3/2} e^{-m/T}, \tag{2.16}$$

$$\rho_{\rm eq} = mg\left(\frac{mT}{2\pi}\right)^{3/2} e^{-m/T} = m n_{\rm eq}. \tag{2.17}$$

Note that the photon fluid does not have a chemical potential because the number of photons is not conserved.

Next, we show that the entropy is conserved in thermal equilibrium[1] (= kinetic equilibrium + chemical equilibrium). Since there are far more photons than baryons in the universe, the entropy of the universe is dominated by the entropy of the photon bath (at least as long as the universe is sufficiently uniform). For negligible chemical potentials due to the number-changing processes of photons, e.g. $e^+e^- \to \gamma\gamma, \gamma\gamma\gamma$, the first law of thermodynamics tells us

$$\mathrm{d}U = T\mathrm{d}S - \mathcal{P}\mathrm{d}V. \tag{2.18}$$

Then we can write

$$\mathrm{d}(\rho(T)V) = T\,\mathrm{d}(s(T)V) - \mathcal{P}(T)\mathrm{d}V, \tag{2.19}$$

where $s(T)$ is the entropy density. For a constant temperature T, using product rule and dropping terms proportional to $\mathrm{d}T$, we have

$$\rho(T)\mathrm{d}V = Ts(T)\mathrm{d}V - \mathcal{P}(T)\mathrm{d}V. \tag{2.20}$$

Thus the entropy density can be written as

$$s(T) = \frac{\rho(T) + \mathcal{P}(T)}{T}. \tag{2.21}$$

Similarly using one of the Maxwell relations and considering the coefficient in front of the differential $T\mathrm{d}V$ in Eq. (2.19),

$$s(T) = \frac{\partial \mathcal{P}}{\partial T}. \tag{2.22}$$

Then we have

$$\frac{\partial \mathcal{P}}{\partial T} = \frac{\rho(T) + \mathcal{P}(T)}{T}. \tag{2.23}$$

This relation can be used to show that the entropy density in the universe scales as a^{-3}. To see

[1] Thermal equilibrium is achieved for species which are both in kinetic and chemical equilibrium. These species then share a common temperature, $T_i = T$. This common temperature is often identified with the photon temperature T_γ — the "temperature of the Universe".

this, let us first rewrite the conservation law (continuity equation) Eq. (1.79) as

$$a^{-3}\frac{\partial\left[(\rho+\mathcal{P})a^3\right]}{\partial t}-\frac{\partial \mathcal{P}}{\partial t}=0. \tag{2.24}$$

The derivative of the pressure w.r.t. time can be written as

$$\frac{\partial \mathcal{P}}{\partial t}=\frac{\mathrm{d}T}{\mathrm{d}t}\frac{\partial \mathcal{P}}{\partial T}, \tag{2.25}$$

so now

$$\begin{aligned}
&a^{-3}\frac{\partial\left[(\rho+\mathcal{P})a^3\right]}{\partial t}-\frac{\mathrm{d}T}{\mathrm{d}t}\frac{\rho+\mathcal{P}}{T}\\
=&a^{-3}T\frac{\partial}{\partial t}\left[\frac{(\rho+\mathcal{P})}{T}a^3\right]\\
=&a^{-3}T\frac{\partial}{\partial t}\left[s(T)a^3\right]=0,
\end{aligned} \tag{2.26}$$

where we have used Eq. (2.23). Therefore, the condition of thermal equilibrium tells us that the entropy in a comoving volume is fixed[2],

$$s(T)a^3=\text{const.} \tag{2.27}$$

In the radiation-dominated era, it is convenient to define the *effective relativistic degrees of freedom* $g_*^{\rho,s}(T)$[3] to calculate the total (radiation) density and entropy:

$$\rho=\sum_i \rho_i=\frac{\pi^2}{30}g_*^{\rho}(T)T^4, \qquad s=\sum_i\frac{\rho_i+\mathcal{P}_i}{T_i}=\frac{2\pi^2}{45}g_*^{s}(T)T^3, \tag{2.28}$$

where T is the temperature of the photon gas. The sum over particle species may receive two types of contributions, $g_*^{\rho}(T)=g_{*\rho}^{\mathrm{th}}(T)+g_{*\rho}^{\mathrm{dec}}(T)$:

- Relativistic species in thermal equilibrium with the photons, $T_i=T\gg m_i$,

  $$g_{*\rho}^{\mathrm{th}}(T)=\sum_{i=b}g_i+\frac{7}{8}\sum_{i=f}g_i. \tag{2.29}$$

  Notice that for these particle species, their thermal contribution is independent of temperature (when the condition $T\gg m_i$ is satisfied). When the temperature drops below the mass m_i of a particle species, it becomes non-relativistic and is removed from the sum.

- Relativistic species that are not in thermal equilibrium with the photons, $T_i\neq T\gg m_i$,

  $$g_{*\rho}^{\mathrm{dec}}(T)=\sum_{i=b}g_i\left(\frac{T_i}{T}\right)^4+\frac{7}{8}\sum_{i=f}g_i\left(\frac{T_i}{T}\right)^4. \tag{2.30}$$

  We have allowed the decoupled species to have different temperatures T_i. For example, this will be relevant for neutrinos after e^+e^- annihilation.

Similarly, for the effective relativistic degrees of freedom in entropy, there are also these two contributions, $g_*^{s}(T)=g_{*s}^{\mathrm{th}}(T)+g_{*s}^{\mathrm{dec}}(T)$:

[2] To a good approximation, expansion of the Universe is adiabatic, so the total entropy stays constant beyond equilibrium.

[3] Since the energy density of relativistic species is much greater than that of non-relativistic species (in the radiation-dominated era or the early Universe), it suffices to include the relativistic species only.

- Relativistic species in thermal equilibrium with the photons, $T_i = T \gg m_i$,

$$g_{*s}^{\text{th}}(T) = \sum_{i=b} g_i + \frac{7}{8}\sum_{i=f} g_i = g_{*\rho}^{\text{th}}(T). \tag{2.31}$$

- Relativistic species that are not in thermal equilibrium with the photons, $T_i \neq T \gg m_i$,

$$g_{*s}^{\text{dec}}(T) = \sum_{i=b} g_i \left(\frac{T_i}{T}\right)^3 + \frac{7}{8}\sum_{i=f} g_i \left(\frac{T_i}{T}\right)^3. \tag{2.32}$$

  Note that $g_{*s}^{\text{dec}} \neq g_{*\rho}^{\text{dec}}$ since $s_i \propto T_i^3$.

Whenever particles decouple from the thermal plasma (the temperature of the universe drops below the mass of a particle species and it becomes non-relativistic), $g_*^{\rho,s}$ decreases. $g_*^{\rho}(T) = g_*^{s}(T)$ only when *all* the relativistic species are in equilibrium at the same temperature. In the real Universe, this is the case until $t \approx 1$ s, i.e. before neutrino decoupling [cf. Fig. 2.3].

2.2 Relativistic decoupling of neutrinos

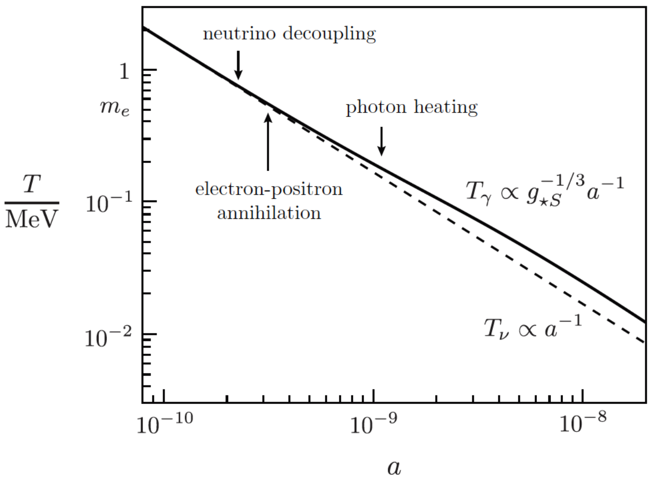


Figure 2.1: Thermal history through electron-positron annihilation.

Neutrinos are coupled to the thermal bath via weak interaction processes like

$$\begin{aligned} \nu_e + \overline{\nu}_e &\leftrightarrow e^+ + e^-, \\ e^- + \overline{\nu}_e &\leftrightarrow e^- + \overline{\nu}_e. \end{aligned} \tag{2.33}$$

They decouple at $T \sim 1$ MeV (when $\Gamma/H \sim 1$), before $e^\pm$ become non-relativistic. After decoupling, the neutrinos move freely along geodesics and preserve to an excellent approximation the relativistic Fermi-Dirac distribution. Shortly after the neutrinos decouple, the temperature drops below the electron mass, $e^\pm$ then become non-relativistic and annihilate,

$$e^+ + e^- \leftrightarrow \gamma + \gamma. \tag{2.34}$$

The energy density and entropy of the electrons and positrons are transferred to the photons, but not to the decoupled neutrinos. In such a way they heat up the photons but not the decoupled neutrinos. Thus neutrinos are effectively colder than the cosmic plasma [cf. Fig. 2.1].

To relate the neutrino temperature to the photon temperature today, we can use entropy conservation. Entropy before neutrino decoupling at scale factor a_1 is

$$s(a_1) = \frac{2\pi^2}{45} T_1^3 \left[2(\gamma) + \frac{7}{8}(2(e^-) + 2(e^+) + 3 \times 2(\nu)) \right] = \frac{43\pi^2}{90} T_1^3, \tag{2.35}$$

where the numbers with brackets indicate the effective *relativistic* d.o.f. of particles in the corresponding brackets. After $e^\pm$ become non-relativistic ($m_e = 0.5$ MeV vs. decoupling temperature for neutrinos $T \sim 1$ MeV), they transfer their entropy to the cosmic plasma and effectively reheat the cosmic plasma. Thus the entropy after $e^\pm$ annihilation at a late-enough redshift a_2 is

$$s(a_2) = \frac{2\pi^2}{45} \left[2T_\gamma^3 + \frac{7}{8} \cdot 3 \cdot 2T_\nu^3 \right], \tag{2.36}$$

where photons have a temperature T_γ and neutrinos have a temperature T_ν. Entropy conservation $s(a_1)a_1^3 = s(a_2)a_2^3$ tells us

$$\frac{43}{2}(a_1 T_1)^3 = 4 \left[\left(\frac{T_\gamma}{T_\nu} \right)^3 + \frac{21}{8} \right] (T_\nu(a_2)a_2)^3. \tag{2.37}$$

Finally we need to relate the temperature T_1 to the temperature at a later time. After neutrinos are decoupled, they still preserve the shape of the Fermi-Dirac distribution and the temperature is inversely proportional to the scale factor, $T \propto a^{-1}$, because its energy scales as a^{-1} as discussed in Sec. 1.2.2. Thus the temperature of neutrinos T_ν satisfy

$$a_1 T_1 = T_\nu(a_2)a_2. \tag{2.38}$$

Solving Eq. (2.37) gives

$$\frac{T_\nu}{T_\gamma} = \left(\frac{4}{11} \right)^{1/3}. \tag{2.39}$$

We can thus conclude that the temperature of neutrino background today is lower than the cosmic microwave background (CMB). We find that the neutrinos of neutrinos today is (1 eV $= 1.16 \times 10^4$ K)

$$T_\nu^0 = T_\gamma^0 \left(\frac{4}{11} \right)^{1/3} = 2.73 \text{ K} \times \left(\frac{4}{11} \right)^{1/3} = 1.95 \text{ K} = 1.68 \times 10^{-4} \text{ eV}. \tag{2.40}$$

The temperature of neutrinos today T_ν^0 is smaller than the square root of the solar mass square difference, $\sqrt{\Delta m_\odot^2} = 8.66 \times 10^{-3}$ eV, and thus at least two neutrinos are massive today.

For $T \ll m_e$, the effective relativistic d.o.f. in energy density and entropy are

$$g_*^\rho = 2 + \frac{7}{8} \times 2 \times \left(\frac{4}{11} \right)^{4/3} N_{\text{eff}} = 3.36, \tag{2.41}$$

$$g_*^s = 2 + \frac{7}{8} \times 2 \times \left(\frac{4}{11} \right) N_{\text{eff}} = 3.94, \tag{2.42}$$

where we have introduced the parameter N_{eff} as the effective number of neutrino species in the Universe. If neutrinos decoupling was instantaneous, then $N_{\text{eff}} = 3$. However, neutrino decoupling was not quite complete when e^+e^- annihilation began, so some of the energy and entropy did leak to the neutrinos. Taking this into account raises the effective number of neutrinos to $N_{\text{eff}} = 3.046$[4].

[4]The Planck constraint on N_{eff} is 3.36 ± 0.34. This still leaves room for discovering that $N_{\text{eff}} \neq 3.046$.

Moreover, we used to believe that neutrinos were massless ($m_\nu = 0$), in which case we would have

$$\rho_\nu = \frac{7}{8} N_{\rm eff} \left(\frac{4}{11} \right)^{4/3} \rho_\gamma \qquad \Longrightarrow \qquad \Omega_\nu h^2 \approx 1.7 \times 10^{-5}. \tag{2.43}$$

But neutrino oscillation experiments have since shown that neutrinos do have mass and they gave a lower bound on the total neutrino mass, $\sum_i m_{\nu,i} > 60$ meV. Massive neutrinos behave as radiation-like particles in the early Universe, and as matter-like particles in the late Universe. Taking this into account, one can show that the energy density of massive neutrinos, $\rho_\nu = \sum_i m_{\nu,i} n_{\nu,i}$, corresponds to

$$\Omega_\nu h^2 \approx \frac{\sum_i m_{\nu,i}}{94 \text{ eV}}. \tag{2.44}$$

By demanding that the neutrinos don't over close the Universe, that is, $\Omega_\nu < 1$, one sets a cosmological upper bound on the sum of neutrino masses, $\sum_i m_{\nu,i} < 15$ eV. In fact, measurements of tritium β-decay find that $\sum_i m_{\nu,i} < 6$ eV. Observations of the cosmic microwave background, galaxy clustering and type Ia supernovae together put an even strong bound, $\sum_i m_{\nu,i} < 1$ eV.

2.3 The Boltzmann equation

What happens when we go out of equilibrium? The out-of-equilibrium phenomena played a role in: (i) the formation of the light elements during Big Bang Nucleosynthesis (BBN); (ii) recombination of electrons and protons into neutral hydrogen when the temperature was of order 1/4 eV; and quite possibly in (iii) production of dark matter in the early universe. The formal tool to describe the evolution beyond equilibrium is the Boltzmann equation.

The Boltzmann equation in the compact form is

$$\frac{{\rm d} f(\mathbf{x}, \mathbf{p}, t)}{{\rm d}\lambda} = C'[f], \tag{2.45}$$

where LHS is the Liouville term (the change of the distribution function w.r.t. the affine parameter λ) and RHS is the collision term taking into account all interactions. We use

$$\frac{{\rm d}x^\mu}{{\rm d}\lambda} = P^\mu \qquad \Longrightarrow \qquad \frac{{\rm d}t}{{\rm d}\lambda} = E \tag{2.46}$$

to write

$$\frac{{\rm d}f}{{\rm d}t} = \frac{{\rm d}\lambda}{{\rm d}t}\frac{{\rm d}f}{{\rm d}\lambda} = \frac{1}{E} C'[f] \equiv C[f]. \tag{2.47}$$

In an FRW universe (homogeneous and isotropic), the distribution function is a function of only the magnitude of 3-momentum $p \equiv |\mathbf{p}|$ (and time), $f = f(p, t)$.

We will look at Boltzmann equation for number density. Recall that

$$n = g \int \frac{{\rm d}^3 p}{(2\pi)^3} f(p). \tag{2.48}$$

We use the chain rule

$$\frac{{\rm d}f(p,t)}{{\rm d}t} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial p}\frac{{\rm d}p}{{\rm d}t} = \frac{\partial f}{\partial t} - pH \frac{\partial f}{\partial p}, \tag{2.49}$$

where in the last equal sign we have used $\dot{p}/p = -\dot{a}/a = -H$ (cf. Eq. (1.51)). Integrating this equation over the three-momentum we obtain

$$\begin{aligned}
g\int\frac{\mathrm{d}^3p}{(2\pi)^3}\frac{\mathrm{d}f}{\mathrm{d}t} &= g\int\frac{\mathrm{d}p}{(2\pi)^3}4\pi p^2\left(\frac{\partial f}{\partial t}-pH\frac{\partial f}{\partial p}\right)\\
&= \frac{\partial}{\partial t}\left[g\int\frac{\mathrm{d}p}{(2\pi)^3}4\pi p^2 f(p,t)\right]-H\Bigg\{\underbrace{g\int\frac{\mathrm{d}p}{(2\pi)^3}\frac{\partial}{\partial p}\left[4\pi p^3 f(p,t)\right]}_{\propto\, p^3 f(p,t)\,|_0^\infty=0}-3g\int\frac{\mathrm{d}p}{(2\pi)^3}4\pi p^2 f(p,t)\Bigg\}\\
&= \frac{\partial}{\partial t}\left[g\int\frac{\mathrm{d}^3p}{(2\pi)^3}f(p,t)\right]+3Hg\int\frac{\mathrm{d}^3p}{(2\pi)^3}f(p,t)\\
&= \frac{\partial n}{\partial t}+3Hn = a^{-3}\frac{\mathrm{d}}{\mathrm{d}t}(na^3),
\end{aligned} \tag{2.50}$$

where the boundary term $p^3 f(p,t)\,|_0^\infty$ vanishes because a finite energy density requires that $f(p)$ decreases faster than $1/p^4$ for large p. Then in the absence of interactions (collisions), the the number density evolves as

$$\frac{\partial n}{\partial t}+3\frac{\dot a}{a}n = a^{-3}\frac{\mathrm{d}}{\mathrm{d}t}(na^3) = 0. \tag{2.51}$$

This means that in the absence of a collision term, the number density n scales as a^{-3}, i.e. the density dilutes with the expanding volume, as expected. This is simply a reflection of the fact that the number of particles in a fixed physical volume is conserved.

The Boltzmann equation formalizes the statement that the rate of change in the abundance of a given particle is the difference between the rates for producing and eliminating that species. Consider a $2\to 2$ scattering, $1+2\leftrightarrow 3+4$. The Boltzmann equation for this system in an expanding universe is given by

$$a^{-3}\frac{\mathrm{d}(n_1a^3)}{\mathrm{d}t} = \int\prod_{i=1}^{4}\mathrm{d}\Pi_i(2\pi)^4\delta^{(4)}(P_1+P_2-P_3-P_4)|\mathcal{M}|^2\times[f_3f_4(1\pm f_1)(1\pm f_2)-f_1f_2(1\pm f_3)(1\pm f_4)], \tag{2.52}$$

where

$$\int\mathrm{d}\Pi_i = g_i\int\frac{\mathrm{d}^3p_i}{(2\pi)^3 2E_i(p_i)} \tag{2.53}$$

is the *Lorentz invariant phase space integral* for particle i with g_i internal degrees of freedom. The factor of $2E_i$ in the denominator arises because, relativistically, the phase space integrals should really be 4 dimensional, over the three components of momentum and one of energy. However, these are constrained to lie on a 3-sphere fixed by the on-shell condition, $E^2 = p^2+m^2$:

$$\begin{aligned}
\int\mathrm{d}^4p\;\delta(E^2-p^2-m^2)\theta(E) &= \int\mathrm{d}^3p\int_0^\infty\mathrm{d}E\;\delta(E^2-p^2-m^2)\\
&= \int\mathrm{d}^3p\int_0^\infty\mathrm{d}E\;\frac{\delta\left(E-\sqrt{p^2+m^2}\right)}{2E},
\end{aligned} \tag{2.54}$$

where we have used the identity

$$\delta[f(x)] = \left|\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x_0}^{-1}\delta(x-x_0). \tag{2.55}$$

On the RHS of Eq. (2.52), the 4D delta function $\delta^{(4)}(P_1+P_2-P_3-P_4)$ enforces energy-momentum conservation. The matrix element squared $|\mathcal{M}|^2$, which governs the strength of the interaction, is averaged over initial and final states. The term $f_3f_4(1\pm f_1)(1\pm f_2)$ is called the *source term* and describes $3 + 4 \to 1 + 2$. It is proportional to the distribution functions in the initial state because we have to average over the thermal distribution. The factors $1 \pm f_i$ are called the Bose-enhancement (+) and Pauli-blocking (−) factors. These factors encapsulate the fact that it is easier (harder) for a boson (fermion) to transition to a state that already contains a boson (fermion). Similarly, $f_1 f_2 (1 \pm f_3)(1 \pm f_4)$ is called the *loss term*. Note that we have assumed that the process is reversible so that $|\mathcal{M}_{34\to12}|^2 = |\mathcal{M}_{12\to34}|^2 = |\mathcal{M}|^2$.

We will assume that all particles are in kinetic equilibrium (all species share the same temperature) so then we can apply Bose-Einstein or Fermi-Dirac distribution. But they are not necessarily in chemical equilibrium (equilibrium in number densities). For systems with $T \ll E - \mu$, we can neglect ± 1 in the denominators of the Fermi-Dirac and Bose-Einstein distributions, and so we just use the Maxwell-Boltzmann distribution, which is given by

$$f^{MB}(E) = e^{-(E-\mu)/T} = e^{\mu/T} e^{-E/T}. \tag{2.56}$$

Similarly, approximating $1 \pm f \simeq 1$ for the Pauli-blocking and Bose-enhancement factors, we can get the expression

$$\begin{aligned} f_3 f_4 (1 \pm f_1)(1 \pm f_2) - f_1 f_2 (1 \pm f_3)(1 \pm f_4) \simeq & f_3^{MB} f_4^{MB} - f_1^{MB} f_2^{MB} \\ = & e^{-(E_1+E_2)/T} \left[e^{(\mu_3+\mu_4)/T} - e^{(\mu_1+\mu_2)/T} \right] \end{aligned} \tag{2.57}$$

where we have used energy conservation $E_1 + E_2 = E_3 + E_4$. The number density of species i is given by

$$n_i = n_i^{(0)} e^{\mu_i/T}, \tag{2.58}$$

where $n_i^{(0)}$ is called the *equilibrium number density* and is given as

$$n_i^{(0)} = g_i \int \frac{\mathrm{d}^3 p}{(2\pi)^3} e^{-E_i/T} = \begin{cases} g_i \left(\frac{m_i T}{2\pi}\right)^{3/2} e^{-m_i/T} & m_i \gg T \\ g_i \zeta(3) \frac{T^3}{\pi^2} & m_i \ll T \end{cases}. \tag{2.59}$$

Using this, we can rewrite Eq. (2.57) in terms of the number densities

$$f_3 f_4 (1 \pm f_1)(1 \pm f_2) - f_1 f_2 (1 \pm f_3)(1 \pm f_4) = e^{-(E_1+E_2)/T} \left\{ \frac{n_3 n_4}{n_3^{(0)} n_4^{(0)}} - \frac{n_1 n_2}{n_1^{(0)} n_2^{(0)}} \right\}. \tag{2.60}$$

So the Boltzmann equation (2.52) can just be written as

$$\boxed{a^{-3} \frac{\mathrm{d}(n_1 a^3)}{\mathrm{d}t} = n_1^{(0)} n_2^{(0)} \langle \sigma v \rangle \left\{ \frac{n_3 n_4}{n_3^{(0)} n_4^{(0)}} - \frac{n_1 n_2}{n_1^{(0)} n_2^{(0)}} \right\}}, \tag{2.61}$$

with the *thermally averaged cross section*

$$\langle \sigma v \rangle = \frac{1}{n_1^{(0)} n_2^{(0)}} \prod_{i=1}^{4} \int \mathrm{d}\Pi_i e^{-(E_1+E_2)/T} (2\pi)^4 \delta^{(4)}(P_1 + P_2 - P_3 - P_4) |\mathcal{M}|^2. \tag{2.62}$$

Before moving on, a few comments are in order. First, note that we could equally well use $E_3 + E_4$ and the equilibrium number densities of the particles 3 and 4. It is also straightforward to generalize the expression of the Boltzmann equation to other processes such as decays or scatterings with more than 2 particles in the final state. Next, if the interaction rate is much larger than the Hubble rate, $\langle \sigma v \rangle n_2^{(0)} \gg H$, the chemical equilibrium is also achieved. This can be seen by noting that the LHS of Eq. (2.61) is of order n_1/t, or, since the typical cosmological time is H^{-1}, $n_1 H$;

whereas the RHS is of order $n_1 n_2 \langle \sigma v \rangle$. Therefore, if the interaction rate is much larger than the Hubble rate, the terms on the RHS will be much larger than the one on the LHS. The only way to maintain the equality is then for the individual terms on the right to cancel. So the number densities satisfy

$$\frac{n_3 n_4}{n_3^{(0)} n_4^{(0)}} = \frac{n_1 n_2}{n_1^{(0)} n_2^{(0)}}. \tag{2.63}$$

Consequently, the chemical potentials are related by [cf. Eq. (2.58)]

$$\mu_3 + \mu_4 = \mu_1 + \mu_2. \tag{2.64}$$

2.4 Big Bang Nucleosynthesis (BBN)

BBN is the formation of light elements. It occurs at temperatures $T \lesssim 1$ MeV since nuclear binding energies are of order MeV. After the decoupling of neutrinos, the universe consisted of relativistic particles such as photons, electrons, positrons and neutrinos (decoupled), and non-relativistic particles such as baryons. Baryons did not completely annihilate due to a small initial baryon-antibaryon asymmetry, $(n_b - n_{\bar{b}})/n_b \simeq 3 \times 10^{-8}$. Rather than the freeze-out mechanism for e.g. neutrino decoupling, the baryons decoupled when all antibaryons annihilated away. This mechanism then explains the large excess of baryons over antibaryons in today's Universe, which is given by

$$\eta_b \equiv \frac{n_b - n_{\bar{b}}}{n_\gamma} \simeq \frac{n_b}{n_\gamma} \simeq 6.05 \times 10^{-10}. \tag{2.65}$$

In principle, BBN is a very complicated process involving many coupled Boltzmann equations to track all the nuclear abundances. However, in practice, we can make two simplifications. First, essentially no elements heavier than helium are produced at appreciable levels. So the only neuclei we need to track are hydrogen and helium, and their isotopes. Second, above $T \simeq 0.1$ MeV, no light nuclei are formed but only free protons and neutrons exist. Then deuterium and other light nuclei formed at $T \leq 0.1$ MeV.

2.4.1 Deuterium

Deuterium (an isotope of hydrogen with one proton and one neutron) is produced by

$$n + p \leftrightarrow D + \gamma. \tag{2.66}$$

When the interaction rate $\Gamma \gg H$, the chemical equilibrium is reached and thus Eq. (2.63) applies to the deuterium production reaction (with $n_\gamma = n_\gamma^{(0)}$)

$$\left.\frac{n_D}{n_n n_p}\right|_{\rm eq} = \frac{n_D^{(0)}}{n_n^{(0)} n_p^{(0)}}. \tag{2.67}$$

Using the equilibrium number density given in Eq. (2.59) (with $g_D = 3$ and $g_n = g_p = 2$), we get

$$\left.\frac{n_D}{n_n n_p}\right|_{\rm eq} = \frac{3}{4}\left(\frac{2\pi m_D}{m_n m_p T}\right)^{3/2} e^{(m_n + m_p - m_D)/T} \simeq \frac{3}{4}\left(\frac{4\pi}{m_p T}\right)^{3/2} e^{B_D/T}, \tag{2.68}$$

where $B_D \equiv m_n + m_p - m_D$ is the binding energy and the approximation is made for $m_p \simeq m_n \simeq m_D/2$. Therefore, as long as the chemical equilibrium holds, the deuterium-to-proton ratio is

$$\left.\frac{n_D}{n_p}\right|_{\text{eq}} = \frac{3}{4} n_n^{\text{eq}} \left(\frac{4\pi}{m_p T}\right)^{3/2} e^{B_D/T}. \tag{2.69}$$

To get an order of magnitude estimate, we note that the number densities of both proton and neutron are proportional to the baryon density n_b. Thus, using the expression

$$n_n \sim n_b \simeq \eta_b n_\gamma = \eta_b \frac{2\zeta(3)}{\pi^2} T^3, \tag{2.70}$$

and dropping the numerical factors, Eq. (2.69) can be written as

$$\left.\frac{n_D}{n_b}\right|_{\text{eq}} \sim \eta_b \left(\frac{T}{m_p}\right)^{3/2} e^{B_D/T}. \tag{2.71}$$

The smallness of the baryon-to-photon ratio (as in Eq. (2.65)) inhibits the production of deuterium until the temperature drops well beneath the binding energy B_D. When $\Gamma \sim H$, the deuterium abundance "freeze-out" and there is a non-zero deuterium abundance.

2.4.2 Neutron to proton ratio

The primordial ratio of neutrons to protons is of particular importance to the outcome of BBN, since essentially all the neutrons become incorporated into ^{4}He. Neutrons and protons are converted into each other via weak interactions

$$p + \bar{\nu} \leftrightarrow n + e^+, \qquad p + e^- \leftrightarrow n + \nu, \qquad n \leftrightarrow p + e^- + \bar{\nu}. \tag{2.72}$$

At high temperatures the dominant processes are the first two scattering processes. Protons and neutrons are in equilibrium down to $T \sim 1$ MeV. Their number density ratio evolves as

$$\frac{n_p^{(0)}}{n_n^{(0)}} = \frac{e^{-m_p/T} \int \mathrm{d}p\, p^2 e^{-p^2/2m_p T}}{e^{-m_n/T} \int \mathrm{d}p\, p^2 e^{-p^2/2m_n T}} \simeq e^{Q/T}, \tag{2.73}$$

where $Q \equiv m_n - m_p \simeq 1.3$ MeV.

The neutron fraction is defined as

$$X_n \equiv \frac{n_n}{n_n + n_p} \qquad \Longrightarrow \qquad X_n^{\text{eq}}(T) = \frac{1}{1 + n_p^{(0)}/n_n^{(0)}} \simeq \frac{1}{1 + e^{Q/T}}. \tag{2.74}$$

The Boltzmann equation (2.61) for the reactions (2.72) $(1 = n;\ 3 = p;\ 2, 4 = \ell)$ can be written as

$$a^{-3} \frac{\mathrm{d}(n_n a^3)}{\mathrm{d}t} = n_\ell^{(0)} \langle \sigma v \rangle \left[\frac{n_p n_n^{(0)}}{n_p^{(0)}} - n_n \right], \tag{2.75}$$

with the number density of leptons $n_\ell = n_\ell^{(0)}$. If we define the interaction rate $\Gamma_{np} \equiv n_\ell^{(0)} \langle \sigma v \rangle$ and use the definition of the neutron fraction, we can rewrite the Boltzmann equation as

$$\frac{\mathrm{d}X_n}{\mathrm{d}t} = \Gamma_{np} \Big[(1 - X_n) e^{-Q/T} - X_n \Big]. \tag{2.76}$$

We now define a dimensionless quantity $x \equiv Q/T$. And we use the first Friedmann equation for radiation domination to change variables to time t to x:

$$\left(\frac{1}{2t}\right)^2 = H^2 = \frac{8\pi G}{3}\rho = \frac{8\pi G}{3}\frac{\pi^2}{30}g_* T^4. \tag{2.77}$$

The effective relativistic degrees of freedom is $g_* = 10.75$ because $g_\gamma = 2,\ g_{e^+} = g_{e^-} = 2$ and $g_\nu = 6$. Then we find

$$\frac{\mathrm{d}X_n}{\mathrm{d}x} = \frac{x\Gamma_{np}}{H(Q)}\left[e^{-x} - X_n(1+e^{-x})\right] \tag{2.78}$$

with $H(Q) \simeq 1.3\ s^{-1}$. The interaction rate is given by

$$\Gamma_{np} = \frac{255}{\tau_n x^5}(12 + 6x + x^2). \tag{2.79}$$

with the neutron lifetime $\tau_n = 886.7\ s$. Thus, $\Gamma_{np}(T = Q) \simeq 5.5\ s^{-1}$.

2.5 Freeze-out

The prime example for freeze-out is the dark matter production via freeze-out. For concreteness, we will focus on the hypothesis that the dark matter is a weakly interacting massive particle (WIMP). We consider a massive Dirac fermion X (the dark matter particle) with mass m_X, which is initially in thermal equilibrium with the cosmic plasma, but later decouples from it (freeze-out). We consider a process where a dark matter particle and its antiparticle annihilate to produce two light particles:

$$X + \overline{X} \leftrightarrow l + \bar{l}. \tag{2.80}$$

We assume that the light particles are in chemical as well as kinetic equilibrium with the cosmic plasma (in other words, they are tightly coupled to the cosmic plasma), so $n_l = n_l^{(0)}$. Then the Boltzmann equation for $n_X = n_{\overline{X}}$ is [cf. Eq. (2.61)]

$$a^{-3}\frac{\mathrm{d}(n_X a^3)}{\mathrm{d}t} = \langle \sigma v \rangle \left\{ (n_X^{(0)})^2 - n_X^2 \right\}. \tag{2.81}$$

We will assume g_* is constant, which is a good approximation for temperatures well above the QCD phase transition. In this case, entropy conservation tells us that the temperature scales like a^{-1}, i.e. $aT = \text{const}$.[5] We also define

$$Y \equiv \frac{n_X}{T^3}, \qquad Y_{(0)} \equiv \frac{n_X^{(0)}}{T^3}. \tag{2.82}$$

Eq. (2.81) can then be written as

$$\begin{aligned} \frac{1}{a^3}\frac{\mathrm{d}}{\mathrm{d}t}(n_X a^3) &= \langle \sigma v \rangle \left\{ (n_X^{(0)})^2 - n_X^2 \right\} \\ \Longrightarrow\quad T^3 \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{n_X}{T^3}\right) &= \langle \sigma v \rangle \left\{ (n_X^{(0)})^2 - n_X^2 \right\} \\ \Longrightarrow\quad \frac{\mathrm{d}Y}{\mathrm{d}t} &= T^3 \langle \sigma v \rangle \left(Y_{(0)}^2 - Y^2 \right). \end{aligned} \tag{2.83}$$

The freeze-out process is characterized by the mass m_X of the particle X. Thus, it is convenient to define a new dimensionless quantity

$$x \equiv \frac{m_X}{T}. \tag{2.84}$$

[5] More generally, we have $\left(\frac{a(T_1)T_1}{a(T_0)T_0}\right)^3 = \frac{g_*(T_0)}{g_*(T_1)}$, as a consequence of conservation of entropy.

Very high temperature corresponds to $x \ll 1$, in which case reactions proceed rapidly so $Y \simeq Y_{(0)}$. Since the dark matter particles are relativistic at these epochs, the $m \ll T$ limit of Eq. (2.59) implies that $Y \simeq 1$. For low temperatures, or high x, the equilibrium abundance $Y_{(0)}$ becomes exponentially suppressed (e^{-x}). Ultimately, the dark matter particles will become so rare because of this suppression that they will not be able to find each other fast enough to maintain the equilibrium abundance. This is the onset of freeze-out (see Fig. 2.2).

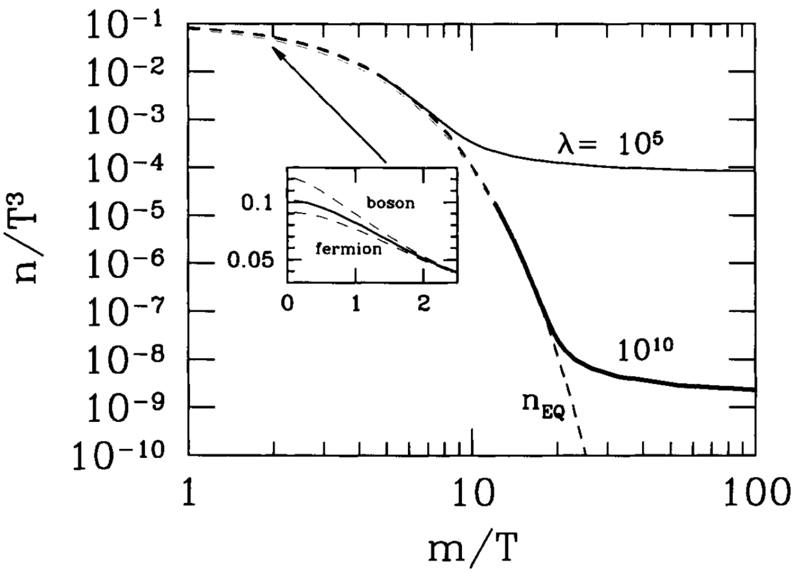


Figure 2.2: Abundance of heavy stable (dark matter) particle as the temperature drops beneath its mass. Dashed line is equilibrium abundance. Two different solid curves show heavy particle abundance for two different values of λ, the ratio of the annihilation rate to the Hubble rate.

To change the variable from t to x, we need the Jacobian $\mathrm{d}x/\mathrm{d}t$. In the radiation-dominated era, the first Friedmann equation can be written as

$$H(T) \overset{\text{Tab. 1.1}}{=} \frac{1}{2t} \overset{(1.73)}{=} \sqrt{\frac{8\pi G}{3}\rho_r(T)} \overset{(2.28)}{=} \sqrt{\frac{8\pi G}{3}\frac{\pi^2}{30}g_*^\rho(T)\,T^2} = \sqrt{\frac{8\pi^3 G}{90}g_*^\rho(T)}\,\frac{m_X^2}{x^2} = \frac{H(m_X)}{x^2}$$
$$\implies \quad t = \frac{x^2}{2H(m_X)}. \tag{2.85}$$

So the evolution equation (2.83) becomes

$$\frac{\mathrm{d}Y}{\mathrm{d}x} = -\frac{\lambda}{x^2}\left(Y^2 - Y_{(0)}^2\right), \tag{2.86}$$

where

$$\lambda \equiv \frac{m_X^3 \langle\sigma v\rangle}{H(m_X)}, \tag{2.87}$$

which parametrizes the interaction strength. The cross section might depend on temperature, but in many theories it is constant or its temperature dependence can be neglected. We will assume it to be constant here. Note that we can cast this evolution equation in a more suggestive way:

$$\frac{x}{Y_{(0)}}\frac{\mathrm{d}Y}{\mathrm{d}x} = -\frac{\lambda Y_{(0)}}{x}\left[\left(\frac{Y}{Y_{(0)}}\right)^2 - 1\right]. \tag{2.88}$$

Using $H(m_X) = x^2 H(T)$, $Y = n_X/T^3$, and $x = m_X/T$, we can rewrite the factor on the RHS as

$$-\frac{\lambda Y_{(0)}}{x} = -\frac{m_X^3 \langle\sigma v\rangle n_{(0)}}{x^3 H(T) T^3} = -\frac{-n_{(0)} \langle\sigma v\rangle}{H(T)} = -\frac{\Gamma_{(0)}}{H(T)}. \tag{2.89}$$

Then Eq. (2.88) becomes

$$\frac{x}{Y_{(0)}} \frac{\mathrm{d}Y}{\mathrm{d}x} = -\frac{\Gamma_{(0)}}{H(T)} \left[\left(\frac{Y}{Y_{(0)}} \right)^2 - 1 \right]. \tag{2.90}$$

In this form, we see that the evolution of the number density of X is controlled by the effectiveness of annihilations, the usual Γ/H factor, times a measure of the deviation from equilibrium.

Eq. (2.86) is a form of the Riccati equation[6] and there is no general analytic solution to it. However, we can obtain an approximate analytic solution. Since the constant λ is generally large, the abundance Y of the particle X will track its equilibrium value. But at late times well after freeze-out, $T \ll m_X$, that is, $x \gg 1$, when the equilibrium abundance is exponentially suppressed, we can neglect $Y_{(0)} \ll Y$. Then Eq. (2.86) simplifies to

$$\frac{\mathrm{d}Y}{\mathrm{d}x} \simeq -\frac{\lambda}{x^2} Y^2 \tag{2.91}$$

at late times ($x \gg 1$). Integrating this equation from the epoch of freeze-out x_f to very late times $x = \infty$, we obtain

$$\frac{1}{Y_\infty} - \frac{1}{Y_f} = \frac{\lambda}{x_f}. \tag{2.92}$$

Typically Y_f at freeze-out is significantly larger than Y_∞ at late times, so a simple analytic approximation is

$$Y_\infty = Y(x = \infty) \simeq \frac{x_f}{\lambda}. \tag{2.93}$$

An analytic estimate for the freeze-out temperature $T_f = m_X/x_f$ can be obtained from considering the size of the coefficient at freeze-out

$$\frac{\lambda Y_{(0)}(x_f)}{x_f} \simeq 1 \tag{2.94}$$

in the rescaled evolution equation (2.88)

$$\frac{x}{Y_{(0)}} \frac{\mathrm{d}Y}{\mathrm{d}x} = -\frac{\lambda Y_{(0)}}{x} \left[\left(\frac{Y}{Y_{(0)}} \right)^2 - 1 \right]. \tag{2.95}$$

From this, we can obtain an implicit equation for the freeze-out temperature,

$$\begin{aligned} \frac{\lambda Y_{(0)}(x_f)}{x_f} &\overset{(2.87)}{=} \frac{m_X^3 \langle\sigma v\rangle}{H(m_X)} \frac{Y_{(0)}(x_f)}{x_f} \simeq 1 \\ \implies H(m_X) \simeq x_f^2 \langle\sigma v\rangle n^{(0)}(x_f) &\overset{(2.59)}{=} \frac{g_X m_X \langle\sigma v\rangle}{(2\pi)^{3/2}} x_f^{1/2} e^{-x_f}, \end{aligned} \tag{2.96}$$

where we have assumed that the dark matter particles are non-relativistic at freeze-out ($m_X \gg T_f$). Typical values for x_f are a few times 10.

Finally, we want to determine the present-day abundance of these heavy particle relics. At temperature T_1 when Y has reached its asymptotic value Y_∞, the number density of the particle X is given by $Y_\infty T_1^3$. For later times, the particles are non-relativistic and the number density scales

[6] In mathematics, a Riccati equation in the narrowest sense is any first-order ordinary differential equation taht is quadratic in the unknown function.

as a^{-3}. Using entropy conservation

$$g_*^s(T_0)(a_0T_0)^3 = g_*^s(T_1)(a_1T_1)^3, \tag{2.97}$$

similar to the case of CMB photons and neutrinos, we can relate the two temperatures and find the energy density today:

$$\rho_X(T_0) = m_X \frac{a_1^3}{a_0^3} n_X(T_1) = m_X \left(\frac{a_1T_1}{a_0T_0}\right)^3 \frac{T_0^3}{T_1^3} n_X(T_1) = m_X Y_\infty T_0^3 \frac{g_*^s(T_0)}{g_*^s(T_1)}, \tag{2.98}$$

with $g_*^s(T_0)/g_*^s(T_1) \sim 1/30$. Using this result we can compute the dimensionless dark matter density today

$$\begin{aligned}\Omega_X h^2 = \frac{\rho_X(T_0)}{\rho_{\text{crit},0}} h^2 = m_X Y_\infty T_0^3 h^2 \frac{g_*^s(T_0)}{g_*^s(T_1)} \cdot \frac{8\pi G}{3H_0^2} &\overset{(2.93)}{=} m_X \frac{x_f}{\lambda} T_0^3 h^2 \frac{g_*^s(T_0)}{g_*^s(T_1)} \cdot \frac{8\pi G}{3H_0^2} \\ &= 0.3 \frac{x_f}{10} \sqrt{\frac{g_*^\rho(m_X)}{100}} \frac{10^{-39}\ \text{cm}^2}{\langle \sigma v \rangle}.\end{aligned} \tag{2.99}$$

This is a remarkable result that nicely ties in with particle physics, because the cross section needed to obtain the correct dark matter relic abundance today ($\Omega_X h^2 \simeq 0.12$) for a particle X with masses $\sim$100 GeV is of order of the weak-interaction cross section G_F^2, where G_F is the Fermi's constant. This coincidence is often called the *WIMP miracle*, because a weakly interacting massive particle (WIMP) automatically obtains the correct relic abundance via freeze-out to explain dark matter.

2.6 Summary tables

Event	time t	redshift z	temperature T
Inflation	10^{-34} s (?)	–	–
Baryogenesis	?	?	?
EW phase transition	20 ps	10^{15}	100 GeV
QCD phase transition	20 μs	10^{12}	150 MeV
Dark matter freeze-out	?	?	?
Neutrino decoupling	1 s	6×10^9	1 MeV
Electron-positron annihilation	6 s	2×10^9	500 keV
Big Bang nucleosynthesis	3 min	4×10^8	100 keV
Matter-radiation equality	60 kyr	3400	0.75 eV
Recombination	260–380 kyr	1100–1400	0.26–0.33 eV
Photon decoupling	380 kyr	1000–1200	0.23–0.28 eV
Reionization	100–400 Myr	11–30	2.6–7.0 meV
Dark energy-matter equality	9 Gyr	0.4	0.33 meV
Present	13.8 Gyr	0	0.24 meV

Figure 2.3: Summary of key events in the thermal history of the universe.

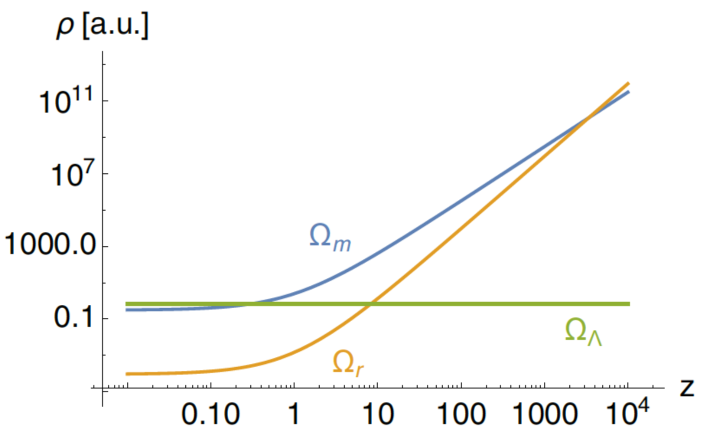


Figure 2.4: Evolution of the energy density ρ of radiation, matter and dark energy. A larger redshift means an earlier time.

Combining Figs. 2.3 and 2.4, we should be able to figure out what component (radiation, matter, or dark energy) dominated at a particular event in the thermal history of the universe by matching the redshift z.

	Flavors	Particle + Antiparticle	Colors	Spins	Total
Quarks (u, d, c, s, t, b)	6	2	3	2	72
Charged leptons (e, μ, τ)	3	2	1	2	12
Neutrinos (ν_e, ν_μ, ν_τ)	3	2	1	1	6
Gluons (g)	1	1	8	2	16
Photon (γ)	1	1	1	2	2
Massive gauge bosons ($W^\pm$, Z^0)	2	2, 1	1	3	9
Higgs bosons (H^0)	1	1	1	1	1
All elementary particles	17				118

Figure 2.5: The SM particles and their degrees of freedom - 1.[7]

type		mass	spin	g
quarks	$t, \bar{t}$	173 GeV	$\frac{1}{2}$	$2 \cdot 2 \cdot 3 = 12$
	$b, \bar{b}$	4 GeV		
	$c, \bar{c}$	1 GeV		
	$s, \bar{s}$	100 MeV		
	$d, \bar{s}$	5 MeV		
	$u, \bar{u}$	2 MeV		
gluons	g_i	0	1	$8 \cdot 2 = 16$
leptons	$\tau^\pm$	1777 MeV	$\frac{1}{2}$	$2 \cdot 2 = 4$
	$\mu^\pm$	106 MeV		
	$e^\pm$	511 keV		
	$\nu_\tau, \bar{\nu}_\tau$	< 0.6 eV	$\frac{1}{2}$	$2 \cdot 1 = 2$
	$\nu_\mu, \bar{\nu}_\mu$	< 0.6 eV		
	$\nu_e, \bar{\nu}_e$	< 0.6 eV		
gauge bosons	W^+	80 GeV	1	3
	W^-	80 GeV		
	Z^0	91 GeV		
	γ	0		2
Higgs boson	H^0	125 GeV	0	1

Figure 2.6: The SM particles and their degrees of freedom - 2.

[7]The neutrinos are assumed to have only one helicity state: left-handed neutrinos and right-handed antineutrinos. For the other fermions, there are LH & RH particles, as well as LH & RH antiparticles. Moreover, massive spin-1 particles ($W^\pm, Z^0$) have three polarizations, while massless spin-1 particles (gluons and photons) have two polarizations. The Higgs boson is a scalar, so it has only one polarization.

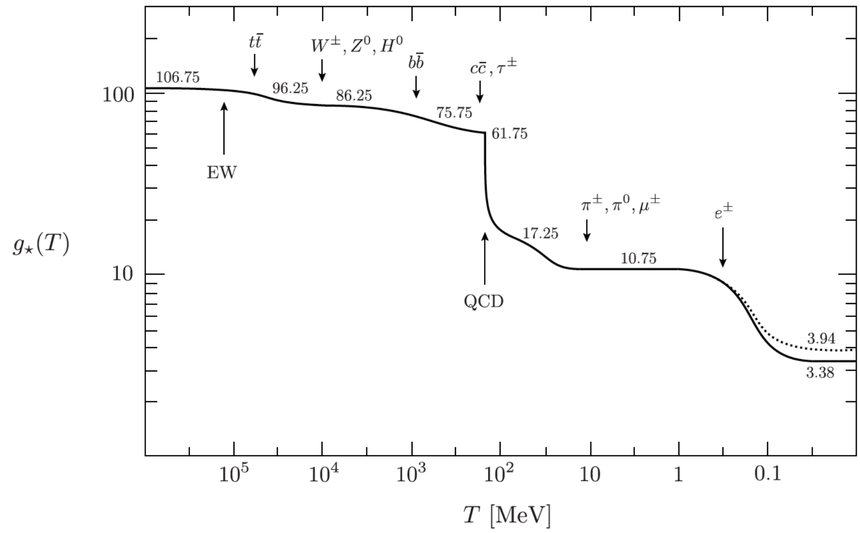


Figure 2.7: Evolution of the effective number of relativistic degrees of freedom $g_*(T) \equiv g_*^\rho(T)$. The dotted line corresponds to the number of effective degrees of freedom in entropy $g_*^s(T)$.

Chapter 3

Inflation

For a homogeneous and isotropic universe, then it is described by the FLRW metric

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a(t)^2\left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2\mathrm{d}\Omega^2\right], \tag{3.1}$$

where $a(t)$ is the scale factor and $k \in \{-1, 0, 1\}$ is the curvature factor. Plugging this into the Einstein field equations

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{3.2}$$

with

$$T^\mu_\nu = \begin{pmatrix} -\rho & 0 & 0 & 0 \\ 0 & \rho & 0 & 0 \\ 0 & 0 & \rho & 0 \\ 0 & 0 & 0 & \rho \end{pmatrix}, \tag{3.3}$$

we obtain the Friedmann equations

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \tag{3.4a}$$

$$\dot{H}^2 + H \equiv \frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p). \tag{3.4b}$$

However, there are problems with this.

3.1 Shortcomings of naive FLRW cosmology

3.1.1 Horizon problem

The size of a causal patch of space it determined by how far light can travel in a certain amount of time. In an expanding spacetime the propagation of light (photons) is best studied using conformal time. We can now define two different types of cosmological horizons, one which limits the distances at which past events can be observed and one which limits the distances at which it will ever be possible to observe future events.

The **particle horizon** is the maximum distance from which particles could have traveled to the observer in the age of the universe. It represents the boundary between the observable and the unobservable regions of the universe, so its distance at the present epoch defines the size of the observable universe. Since nothing can travel faster than light, the particle horizon is determined by the distance that a photon could have traveled since the bang till time t. Photons travel along null geodesics characterized by $\mathrm{d}r = \mathrm{d}t/a(t)$, so the particle horizon is given by

$$R_H(t) = a(t)\int_{t_i}^{t} \frac{\mathrm{d}t'}{a(t')}. \tag{3.5}$$

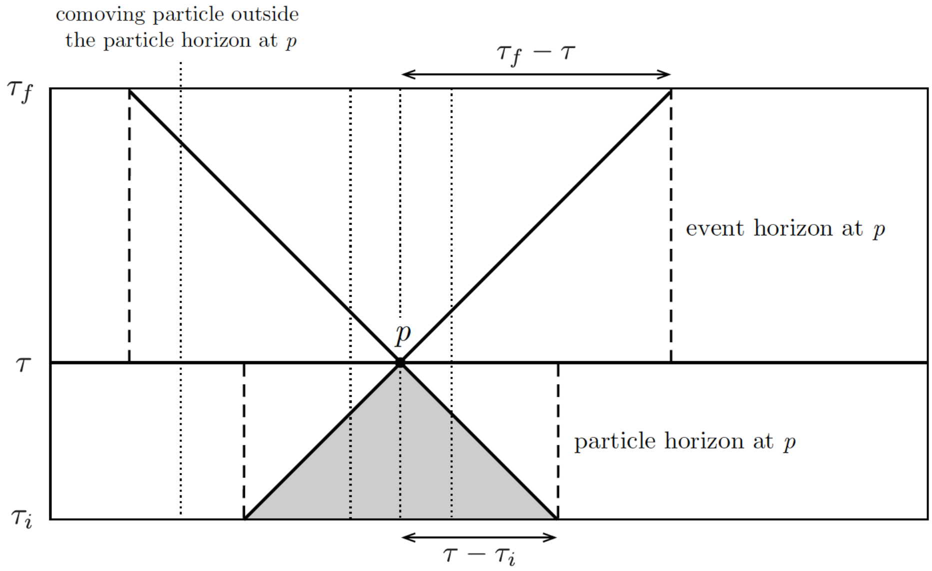


Figure 3.1: Spacetime diagrams illustrating the concept of horizons. Dotted lines show the worldlines of comoving objects. The event horizon is the maximal distance to which we can send signal. The particle horizon is the maximal distance from which we can receive signals.

Using the conformal time, the particle horizon becomes

$$R_H(t) = a(\tau) \int_{\tau_i}^{\tau} \mathrm{d}\tau, \tag{3.6}$$

where τ_i is the conformal time corresponding to t_i. The size of the particle horizon at time τ may be visualized by the intersection of the past light cone of an observer p with the spacelike surface $\tau = \tau_i$, as shown in Fig. 3.1. Causal influences have to come from within this region. Only comoving particles whose worldlines intersect the past light cone of p can send a signal to an observer at p. Notice that every observer has his or her own particle horizon.

Just as there are past events that we cannot see now, there may be future events that we will never be able to see (and distant regions that we will never be able to influence). The cosmic **event horizon** is the largest distance from which light emitted now can ever reach an observer in the future. In general, it is given by

$$R_e(t) = a(t) \int_{t}^{t_f} \frac{\mathrm{d}t}{a(t)} = a(\tau) \int_{\tau}^{\tau_f} \mathrm{d}\tau. \tag{3.7}$$

Notice that this may be finite even if the physical time is infinite, $t_f = +\infty$.

For a fluid with equation of state $p = w\rho$, the particle horizon is

$$R_H = \frac{2}{1+3w} H^{-1}. \tag{3.8}$$

Then if $w > -1/3$, the horizon exists and is finite. Also we see that in the standard cosmology, the particle horizon is, up to numerical factors, equal to the age of the universe or the Hubble radius, H^{-1}. For this reason, we can use horizon and Hubble radius interchangeably. However, as we shall see, in inflational models the horizon and Hubble radius are not roughly equal as the horizon grows exponentially relative to the Hubble radius.

About 380,000 years after the Big Bang, the universe had cooled enough to allow the formation of hydrogen atoms and the decoupling of photons from the primordial plasma. We observe this

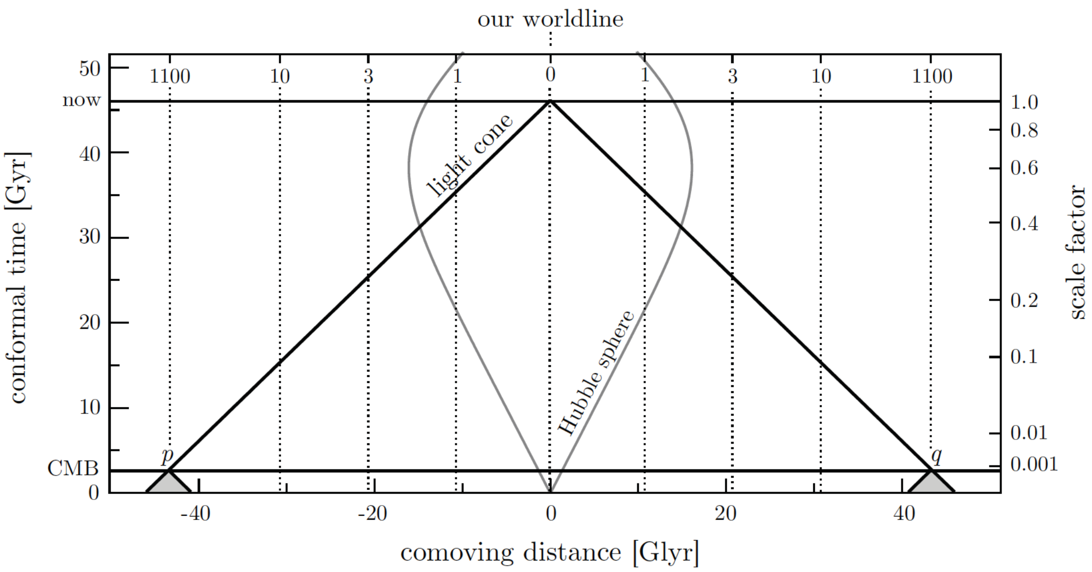


Figure 3.2: The horizon problem in the conventional Big Bang model. The intersection of our past cone with the spacelike CMB surface corresponds to two opposite points in the observed CMB. Their past light cones do not overlap before they hit the singularity, $a = 0$, so the points appear never to have been in causal contact.

event in the form of the cosmic microwave background (CMB), an afterglow of the hot Big Bang. Remarkably, this radiation is almost perfectly isotropic, with anisotropies in the CMB temperature being smaller than one part in ten thousand. However, the finiteness of the conformal time elapsed between $t_i = 0$ and the time of the formation of the CMB, $t_{\rm rec}$, implies a serious problem: it means that most spots in the CMB have non-overlapping past light cones and hence never were in causal contact (see Fig. 3.2). We see that the photons were emitted sufficiently close to the Big Bang singularity that the past light cones of p and q do not overlap. This implies that these two points are not causally connected. Then the puzzle is: how do the photons coming from p and q "know" that they should be at almost exactly the same temperature? In fact, in the standard cosmology the CMB is made of about 10^4 disconnected patches of space. If there wasn't enough time for these regions to communicate, why do they look so similar? This is the **horizon problem**. The quantitative analysis of this is the following.

Today's particle horizon corresponds roughly to the radius of the surface of last scattering[1], $R_H(t_0) \simeq H_0^{-1} \simeq \lambda_{LSS}(t_0)$. The physical size of today's last scattering surface at last scattering is

$$\lambda_{LSS}(t_{LS}) = R_H(t_0)\frac{a_{LS}}{a_0} \simeq \frac{1}{1000} R_H(t_0) \simeq \frac{1}{1000}\lambda_{LSS}(t_0). \tag{3.9}$$

Particle horizon at t_{LS}

$$\begin{aligned} &R_H(t_{LS}) \simeq H_{LS}^{-1} \\ &H^2 \propto \rho_m \propto a^{-3} \quad \text{(matter domination)} \\ &R_H(t_{LS}) \simeq R_H(t_0)\left(\frac{a_{LS}}{a_0}\right)^{3/2}. \end{aligned} \tag{3.10}$$

[1]According to the standard cosmology, photons decoupled from the thermal bath at a temperature of the order of 0.3 eV. This corresponds to the so-called surface of "last scattering" at a redshift of about 1100 and an age of about $180{,}000(\Omega_0 h^2)^{-1/2}$ yrs. From the epoch of last-scattering onwards, photons free-stream and reach us basically untouched.

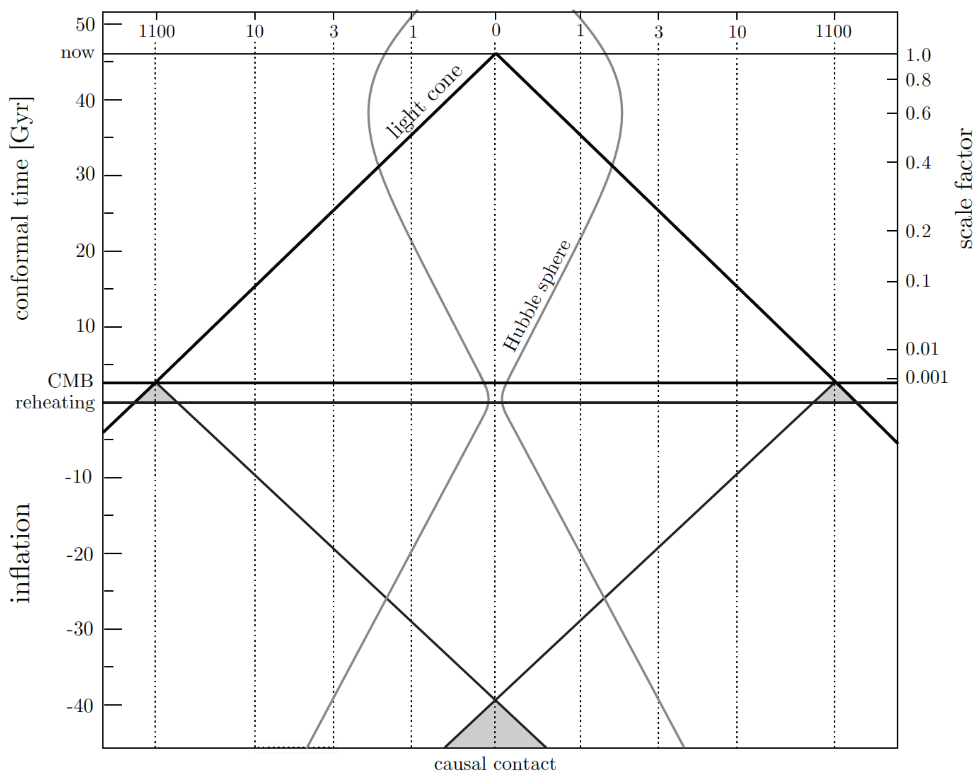


Figure 3.3: Inflationary solution to the horizon problem. The comoving Hubble sphere shrinks during inflation and expands during the conventional Big Bang evolution. The spacelike singularity of the standard Big Bang is replaced by the reheating surface, i.e. rather than marking the beginning of time it now corresponds simply to the transition from inflation to the standard Big Bang evolution. All points in the CMB have overlapping past light cones and therefore originated from a causally connected region of space.

At the time of last scattering, today's Hubble volume consisted of

$$\frac{\lambda^3_{LSS}(t_{LS})}{H^{-3}_{LS}} = \left(\frac{a_{LS}}{a_0}\right)^{-3/2} \simeq 3 \times 10^4 \tag{3.11}$$

causally disconnected patches!

So how do solve the horizon problem? Our description of the horizon problem has highlighted the fundamental role played by the growing comoving Hubble sphere of the standard Big Bang cosmology. Specifically, the comoving Hubble radius $(aH)^{-1}$ scales as

$$(aH)^{-1} = H_0^{-1} a^{\frac{1}{2}(1+3w)}. \tag{3.12}$$

So $w > -\frac{1}{3}$ leads to a growing Hubble sphere as the universe expands. Then a simple solution to the horizon problem therefore suggests itself: let us conjecture a phase of decreasing comoving Hubble radius in the early universe, that is, with

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(aH^{-1}\right) < 0. \tag{3.13}$$

Another way of saying it is that the universe has to go through a phase during which the physical scales $\lambda \sim a$ evolve faster than the horizon scale H^{-1}. Physically, the shrinking Hubble sphere requires a fluid with $w < -1/3$, i.e. no horizon exists. If this lasts long enough, the horizon problem can be avoided. Using the second Friedmann equation, this implies $\ddot{a} > 0$. So this phase corresponds to an accelerated expansion. This is called the **inflation**. See Fig. 3.3.

3.1.2 Flatness problem

Rewrite the first Friedmann equation using the critical density

$$\rho_{\text{crit}} = \frac{3H^2}{8\pi G} \tag{3.14}$$

and

$$\Omega_i \equiv \frac{\rho_i}{\rho_{\text{crit}}}, \tag{3.15}$$

then we have

$$\sum_i \Omega_i - 1 = \frac{k}{a^2H^2} \equiv \Omega_k, \tag{3.16}$$

which is related to the curvature of the universe. We know from (CMB) observations that today $|\Omega_k| \lesssim 10^{-2}$, implying that the space is close to flat. At earlier times,

$$\begin{aligned} \text{Matter domination}: H^2 &\propto \rho_m \propto a^{-3} \\ \implies |\Omega_k| &\propto \frac{1}{a^2a^{-3}} = a \\ \text{Radiation domination}: H^2 &\propto \rho_R \propto a^{-4} \\ \implies |\Omega_k| &\propto \frac{1}{a^2a^{-4}} = a^2. \end{aligned} \tag{3.17}$$

In both cases $\Omega_k = \Omega - 1$ decreases going backwards with time. Let us make a tremendous extrapolation and assume that Einstein equations are valid until the Plank era, when the temperature of the universe is $T_{Pl} \sim m_{Pl} \sim 10^{19}$ GeV. Then we can deduce the value of Ω_k at the Planck time (corresponding to T_{Pl}):

$$\frac{|\Omega_k|_{T_{Pl}}}{|\Omega_k|_{T_0}} \simeq \frac{a_{Pl}^2}{a_0^2} \simeq \frac{T_0^2}{T_{Pl}^2} \simeq \mathcal{O}(10^{-62}). \tag{3.18}$$

The flatness problem is actually a fine-tuning problem. In order to get the correct value of $\Omega_k \sim 1$ at present, the value of Ω_k at early times have to be fine-tuned to values extremely close to zero but without being exactly zero. It turns out that inflation can solve this problem too. If we add a new component ϕ at early times, with $p_\phi = w\rho_\phi$ and $\rho_\phi \propto a^{-3(1+w)}$, then the first Friedmann equation can be written as

$$\frac{3H^2}{8\pi G} = \frac{\rho_{\phi,i}}{a^{3(1+w)}} - \frac{3}{8\pi G}\frac{k}{a^2} + \frac{\rho_{m,i}}{a^3} + \frac{\rho_{r,i}}{a^4}. \tag{3.19}$$

Therefore, as long as $w < -1/3$, all initial curvature (as well as radiation and matter) will be diluted away.

3.2 Inflaton

To implement inflation, we introduce a scalar field ϕ called the **inflaton**, with the action

$$S = \int \mathrm{d}^4x\sqrt{-g}\,\mathscr{L} = \int \mathrm{d}^4x\sqrt{-g}\left[\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi)\right], \tag{3.20}$$

where $\sqrt{-g} = \sqrt{-\det(g_{\mu\nu})} = a^3(t)$ for the FRW metric. We apply the Euler-Lagrange equation

$$\partial_\mu \frac{\partial(\sqrt{-g}\mathscr{L})}{\partial(\partial_\mu\phi)} - \frac{\partial(\sqrt{-g}\mathscr{L})}{\partial\phi} = 0. \tag{3.21}$$

Using the metric signature $(-,+,+,+)$, we get

$$\begin{aligned}
&\partial_\mu(\sqrt{-g}\partial^\mu\phi)-\sqrt{-g}\frac{\mathrm{d}V}{\mathrm{d}\phi}=0\\
\Longrightarrow\quad &\partial_\mu a^3\partial^\mu\phi+a^3\partial_\mu\partial^\mu\phi-a^3\frac{\mathrm{d}V}{\mathrm{d}\phi}=0\\
\Longrightarrow\quad &-\frac{\partial a^3}{\partial t}\frac{\partial\phi}{\partial t}+a^3\left(-\frac{\partial^2\phi}{\partial t^2}+\nabla^2\phi\right)-a^3\frac{\mathrm{d}V}{\mathrm{d}\phi}=0\\
\Longrightarrow\quad &3a^2\dot a\dot\phi+a^3\ddot\phi-a^3\nabla^2\phi-a^3\frac{\mathrm{d}V}{\mathrm{d}\phi}=0.
\end{aligned}\tag{3.22}$$

So we obtain

$$\ddot\phi+3H\dot\phi-\frac{\nabla^2}{a^2}\phi+\frac{\mathrm{d}V}{\mathrm{d}\phi}=0.\tag{3.23}$$

This is the Klein-Gordon equation. The energy-momentum tensor of the scalar field is given by

$$T_{\mu\nu}=\partial_\mu\phi\partial_\nu\phi-g_{\mu\nu}\mathscr{L}.\tag{3.24}$$

and then the corresponding energy density and pressure density are

$$\begin{aligned}
\rho_\phi=T_{00}&=\frac{1}{2}\dot\phi^2+V(\phi)+\frac{(\nabla\phi)^2}{2a^2}\\
p_\phi=T_{ii}&=\frac{1}{2}\dot\phi^2-V(\phi)+\frac{(\nabla\phi)^2}{6a^2}.
\end{aligned}\tag{3.25}$$

The three terms are the kinetic energy, the potential energy, and the gradient energy, respectively. Notice that if the gradient energy dominates, we would have $w=p_\phi/\rho_\phi=-1/3$, not enough to drive inflation. On the other hand, if the potential energy dominates (as in the slow-roll scenario), $V(\phi)\gg 1/2\dot\phi^2$ and $V(\phi)\gg(\nabla\phi)^2/a^2$, then

$$p\simeq-\rho\quad\Longrightarrow\quad w\simeq-1\quad\overset{(3.4b)}{\Longrightarrow}\quad\frac{\ddot a}{a}\simeq+\frac{8\pi G}{3}V(\phi).\tag{3.26}$$

Therefore, we realize that a scalar field whose energy is dominant in the universe and whose potential energy dominates over the kinetic energy gives inflation. Therefore, neglecting the kinetic term totally (**de Sitter approximation**), we have

$$H^2\overset{(3.4a)}{=}\frac{8\pi G}{3}V(\phi)=\text{const.}\equiv H_{\text{inf}}^2\tag{3.27}$$

and

$$a(t)\propto\exp(H_{\text{inf}}\cdot t).\tag{3.28}$$

Therefore, during the inflationary (de Sitter) epoch, the scale factor grows exponentially and the horizon scale H^{-1} is constant. If inflation lasts long enough, all the physical scales that have left the horizon during the radiation-dominated or matter-dominated phase can re-enter the horizon in the past: this is because such scales are exponentially reduced. This then explains both the problem of the homogeneity of CMB and the initial condition problem of small cosmological perturbations. Once the physical length is within the horizon, microphysics can act, the universe can be made approximately homogeneous and the primeval inhomogeneities can be created.

How much inflation do we need to solve the horizon problem? We define the number of **e-foldings** as

$$N_e\equiv\ln\frac{a(t_e)}{a(t_i)}.\tag{3.29}$$

In the de Sitter approximation, it is given by

$$N_e = H_{\text{inf}}(t_e - t_i). \tag{3.30}$$

Today's horizon scale must have been smaller than the Hubble radius at the time of inflation H_{inf}^{-1},

$$\lambda_{H_0}(t_i) = H_0^{-1}\frac{a(t_i)}{a(t_0)} = H_0^{-1}\frac{a(t_i)}{a(t_e)}\frac{a(t_e)}{a(t_0)} = H_0^{-1}e^{-N_e}\frac{T_0}{T_e} \lesssim H_{\text{inf}}^{-1}. \tag{3.31}$$

This gives

$$N_e \gtrsim \ln\left(\frac{T_0}{H_0}\right) + \ln\left(\frac{T_e}{H_{\text{inf}}}\right) \simeq 67 + \underbrace{\ln\left(\frac{T_e}{H_{\text{inf}}}\right)}_{<0}, \tag{3.32}$$

where the value of the last term depends on the physics of reheating. Typically, $N_e \geq 50 \sim 60$ is needed to solve the horizon problem. We can calculate the inflation time from Eq. (3.30)

$$\begin{aligned} t_e - t_i &\gtrsim \frac{N_e}{H_{\text{inf}}} \simeq \frac{50}{10^{15}\ \text{GeV}} \qquad (H_{\text{inf}} \lesssim 10^{15}\ \text{GeV}) \\ t_e - t_i &\gtrsim 50 \cdot 6.6 \times 10^{-41}\ \text{s} \simeq 3 \times 10^{-39}\ \text{s}. \end{aligned} \tag{3.33}$$

3.3 Slow-roll inflation

The slow-roll approximation is

- Potential energy dominates over the kinetic energy: $V(\phi) \gg 1/2\dot{\phi}^2$;
- Friction term dominates over the kinematics of the field: $3H\dot{\phi} \gg \ddot{\phi}$.

This will give an attractor solution.

We use these approximations to solve the Friedmann equations and the Klein-Gordon equation for the inflaton field:

$$(1):\ H^2 = \frac{8\pi G}{3}\rho = \frac{8\pi G}{3}\left(\frac{1}{2}\dot{\phi}^2 + V\right), \tag{3.34a}$$

$$(2):\ \dot{H} + H^2 = -\frac{4\pi G}{3}(\rho + 3p) = \frac{8\pi G}{3}\left(V - \dot{\phi}^2\right), \tag{3.34b}$$

$$(3):\ \ddot{\phi} + 3H\dot{\phi} + \frac{\mathrm{d}V}{\mathrm{d}\phi} = 0. \tag{3.34c}$$

We can express these two conditions in terms of the (Hubble) **slow-roll parameters**. First if we do the subtraction (2) − (1)

$$\dot{H} = -4\pi G\dot{\phi}^2, \tag{3.35}$$

then (1) becomes

$$H^2 = \frac{8\pi G}{3}V - \frac{1}{3}\dot{H}, \tag{3.36}$$

or

$$1 = \frac{8\pi G}{3H^2}V - \frac{1}{3}\frac{\dot{H}}{H^2}. \tag{3.37}$$

So we can define the first slow-roll parameter [cf. Eq. (3.34b) and the first slow-roll condition]:

$$\epsilon_H \equiv -\frac{\dot{H}}{H^2} \ll 1, \tag{3.38}$$

and the second slow-roll parameter [cf. the second slow-roll condition]:

$$\eta_H \equiv \frac{\ddot{\phi}}{H\dot{\phi}} \quad \text{with} \quad |\eta_H| \ll 1. \tag{3.39}$$

Note that by manipulating the above equations [use Eqs. (3.35) and (3.34a) to rewrite $\dot{\phi}^2$ and V in terms of H and $\dot{H}$ and then substitute them into Eq. (3.34b)], we can obtain

$$\dot{H} + H^2 = H^2(1 - \epsilon_H) = \frac{\ddot{a}}{a}. \tag{3.40}$$

This means that when $\epsilon_H = 1$, the inflation ends. We can also define the second set of slow-roll parameters that are sometimes called the potential slow-roll parameters:

$$\begin{aligned} \epsilon_V &\equiv \frac{1}{16\pi G}\left(\frac{V'}{V}\right)^2 \\ \eta_V &\equiv \frac{1}{8\pi G}\frac{V''}{V}, \end{aligned} \tag{3.41}$$

where $V' \equiv \mathrm{d}V/\mathrm{d}\phi$. It can be shown that to first order, these two sets of slow-roll parameters are related by

$$\epsilon_V \simeq \epsilon_H, \qquad \eta_V \simeq \eta_H + \epsilon_H. \tag{3.42}$$

Note that ϵ_V and η_V only depend on the potential and its derivatives, so they do not require solving a differential equation, and can be used to identify the point where inflation ends. If $V(\phi)$ supports $\epsilon_V \ll 1, |\eta_V| \ll 1$, then $V(\phi)$ can potentially describe inflation (necessary but not sufficient condition). Additionally, we need to make sure that $V(\phi)$ can support enough e-foldings of inflation:

$$\begin{aligned} N_e = \int_{t_i}^{t_e} \mathrm{d}t\; H = \int_{\phi_i}^{\phi_e} \mathrm{d}\phi\; \frac{H}{\dot{\phi}} &\overset{(3.35)}{=} \sqrt{4\pi G}\int_{\phi_i}^{\phi_e} \mathrm{d}\phi\; \frac{H}{\sqrt{-\dot{H}}} \\ &= \sqrt{4\pi G}\int_{\phi_i}^{\phi_e} \mathrm{d}\phi\; \frac{1}{\sqrt{\epsilon_H}} \\ &\simeq \sqrt{4\pi G}\int_{\phi_i}^{\phi_e} \mathrm{d}\phi\; \frac{1}{\sqrt{\epsilon_V(\phi)}} \\ &= 8\pi G\int_{\phi_i}^{\phi_e} \mathrm{d}\phi\; \frac{V(\phi)}{V'(\phi)}. \end{aligned} \tag{3.43}$$

Some simple examples of potentials (with two parameters) that can do the job include

- Monomial inflation: $V(\phi) = \lambda M_p^4(\phi/M_p)^n$ (this needs something to stop it falling into negative infinity).
- Hilltop inflation: $V(\phi) = \Lambda^4[1 - (\phi/\mu)^p + \cdots]$.
- Natural inflation: $V(\phi) = \Lambda^4[1 + \cos(\phi/f)]$ (periodic potential that can be seen in complex scalar fields).

3.4 Perturbations from inflation

Our current understanding of the origin of structure in the universe is that it originated from small ‘seed’ perturbations, which over time grew to become all of the structure we observe. Once the universe becomes matter dominated (around 1000 yrs after the bang), primeval density inhomogeneities are amplified by gravity and grow into the structure we see today. The fact that a fluid of self-gravitating particles is unstable to the growth of small inhomogeneities was first pointed out by Jeans and is known as the Jeans instability. In order for structure formation to occur via gravitational instability, there must have been small preexisting fluctuations on physical length scales when they crossed the Hubble radius in the radiation-dominated and matter-dominated eras. Our best guess for the origin of these perturbations is quantum fluctuations during an inflationary era in the early universe.

General idea: quantize the inflaton field and look at the behavior of quantum fluctuations around the mean ("background") value.

3.4.1 Massless scalar field

We start with a toy example. We consider a free scalar field in the de Sitter background, with the Lagrangian

$$\mathscr{L} = \partial_\mu \chi \partial^\mu \chi + V_0 \qquad (V(\chi) = 0). \tag{3.44}$$

For the field, we separate into the classical background part and the perturbation part,

$$\chi(\vec{x}, t) = \overline{\chi}(t) + \delta\chi(\vec{x}, t). \tag{3.45}$$

Given the Lagrangian, the equation of motion is

$$\ddot{\chi} + 3H\dot{\chi} - \frac{\nabla^2}{a^2}\chi = 0. \tag{3.46}$$

Let us go to the momentum space by doing a Fourier transformation

$$\delta\chi(\vec{x}, t) = \int \frac{\mathrm{d}^3 k}{(2\pi)^{3/2}} e^{i\vec{k}\cdot\vec{x}} \delta\chi_k(t), \tag{3.47}$$

where $\delta\chi_k(t)$ are the Fourier components. Then we have an equation of motion just for the perturbations

$$\delta\ddot{\chi}_k + \underbrace{3H\delta\dot{\chi}_k}_{\text{damping term}} + \frac{k^2}{a^2}\delta\chi_k = 0. \tag{3.48}$$

We can study the evolution of the fluctuation in a more quantitative way. Let us define $\delta\sigma_k \equiv a \cdot \delta\chi_k$ and switch to conformal time $\mathrm{d}\tau = \mathrm{d}t/a$. Then Eq. (3.48) becomes

$$\delta\sigma_k'' + \left(k^2 - \frac{a''}{a}\right)\delta\sigma_k = 0, \tag{3.49}$$

where we use prime to indicate the derivatives w.r.t. the conformal time (we used dot to indicate derivatives w.r.t. the coordinate time). For the time being, we solve the problem for a pure de Sitter expansion and we take the scale factor exponentially growing as $a \sim e^{Ht}$ (H is constant); the corresponding conformal factor reads

$$a(\tau) = -\frac{1}{H\tau} \qquad (\tau < 0). \tag{3.50}$$

Eq. (3.49) is the equation of motion for a harmonic oscillator with action

$$\delta S_k = \frac{1}{2}\int \mathrm{d}\tau \left[(\delta\sigma_k')^2 + \left(k^2 - \frac{a''}{a}\right)\delta\sigma_k^2\right], \tag{3.51}$$

where the negative time-dependent mass term $-a''/a = -2/\tau^2$. If we promote $\delta\sigma_k$ to an operator, then the canonical commutation relations tell us

$$\langle 0|\delta\sigma^{\dagger}_{\vec{k}_1}\delta\sigma_{\vec{k}_2}|0\rangle = \delta^{(3)}(\vec{k}_1 - \vec{k}_2)\delta\sigma^{*}_{\vec{k}_1}\delta\sigma_{\vec{k}_2} = |\delta\sigma_k|^2 \tag{3.52}$$

Then the real-space variance of $\delta\sigma$ is

$$\begin{aligned}\mathrm{Var}(\delta\sigma) &= \left\langle |\delta\sigma(\vec{x},t)|^2 \right\rangle \\ &= \int \frac{\mathrm{d}^3 k}{(2\pi)^3}|\delta\sigma_k|^2 \\ &= \int \frac{\mathrm{d}k}{k}\underbrace{\frac{k^2}{2\pi^2}|\delta\sigma_k|^2}_{\equiv P_{\delta\sigma}(k)}\end{aligned} \tag{3.53}$$

where $P_{\delta\sigma}(k)$ is the **power spectrum** that tells us the correlation between two points in space.

Let's return to the equation of motion. First, notice in de Sitter, $a''/a = 2a^2H^2$ (remember that the primes indicate the derivatives w.r.t. the conformal time τ), then Eq. (3.49) becomes

$$\delta\sigma_k'' + \left(k^2 - 2a^2H^2\right)\delta\sigma_k = 0. \tag{3.54}$$

If $k^2/a^2 \gg H^2$ (**sub-Hubble or sub-horizon scale**: wavelengths shorter than the Hubble radius), $\delta\sigma_k$ oscillates around the origin. This implies $\delta\chi_k \sim e^{-Ht}e^{i\omega t}$ (since $a \sim e^{Ht}$) oscillates with exponentially damping amplitude. On the other hand, if $k^2/a^2 \ll H^2$ (**super-Hubble or super-horizon scale**: wavelengths longer than the Hubble radius), $\delta\sigma_k \sim e^{Ht}$ grows exponentially, while $\delta\chi_k \sim$ const. (supercritical damping). The above equation has an exact solution

$$\delta\sigma_k(\tau) = \frac{e^{-ik\tau}}{\sqrt{2k}}\left(1 + \frac{i}{k\tau}\right). \tag{3.55}$$

In the limit $|k\tau| \gg 1$, i.e. $k/a \gg H$ (sub-Hubble scale), the solution reduces to the simple harmonic oscillator plane-wave solution, as it should.

Plugging the solution (3.55) into the expression for the power spectrum, we have

$$\begin{aligned}P_{\delta\chi}(k) = \frac{k^3}{(2\pi)^2}|\delta\chi_k|^2 &= \frac{k^3}{(2\pi)^2a^2}|\delta\sigma_k|^2 \\ &= \frac{k^3}{(2\pi)^2a^2}\left|\frac{e^{-ik\tau}}{\sqrt{2k}}\left(1+\frac{i}{k\tau}\right)\right|^2 \\ &= \frac{k^2}{(2\pi)^2a^2}\left|1+\frac{i}{k\tau}\right|^2.\end{aligned} \tag{3.56}$$

We evaluate this expression at $k = aH$, where $a(\tau) = -1/H\tau$, then it becomes

$$P_{\delta\chi}(k) = \frac{H^2}{(2\pi)^2}. \tag{3.57}$$

Since H is constant in de Sitter, the resulting power spectrum is scale invariant ("white noise"). If we allow a small time dependence of H (**quasi de Sitter expansion**), which is parametrized by the slow-roll parameter $\epsilon_H = -\dot{H}/H^2$, the the power spectrum picks up a weak scale dependence,

$$\frac{\mathrm{d}\ln P_{\delta\chi}(k)}{\mathrm{d}\ln k} = \underbrace{\frac{\mathrm{d}\ln P_{\delta\chi}(k)}{\mathrm{d}t}}_{\sim 2\dot{H}/H}\underbrace{\frac{\mathrm{d}t}{\mathrm{d}\ln k}}_{\sim 1/H} \simeq -2\epsilon_H < 0. \tag{3.58}$$

So we can write it as a power law:

$$P_{\delta\chi}(k) = A \cdot \left(\frac{k}{k_*}\right)^{-2\epsilon_H}, \tag{3.59}$$

where k_* is called the pivot scale and the slow-roll parameter $\epsilon_H \simeq \epsilon_V$.

3.4.2 Massive scalar field

If we now consider the actual inflaton field ϕ, then

$$\delta\ddot{\phi}_k + 3H\delta\dot{\phi}_k + \left(\frac{k^2}{a^2} - \frac{\mathrm{d}^2 V}{\mathrm{d}\phi^2}\right)\delta\phi_k = 0, \tag{3.60}$$

Again, introducing the conformal time and scaling gives

$$\delta\sigma_k'' + \left[k^2 - \frac{1}{\tau^2}\left(\nu_\phi^2 - \frac{1}{4}\right)\right]\delta\sigma_k = 0, \tag{3.61}$$

with

$$\nu_\phi^2 = \frac{9}{4} - \frac{\mathrm{d}^2 V}{\mathrm{d}\phi^2}\frac{1}{H^2}. \tag{3.62}$$

If we further assume that $\mathrm{d}^2V/\mathrm{d}\phi^2 \equiv m_\phi^2 = \text{const.}$, an analytical solution exists in the de Sitter limit:

$$\delta\sigma_k = \frac{\sqrt{\pi}}{2} e^{i(\nu_\phi + \frac{1}{2})\frac{\pi}{2}} \sqrt{-\tau} H_{\nu_\phi}^{(1)}(-k\tau), \tag{3.63}$$

where $H_{\nu_\phi}^{(1)}(-k\tau)$ is the Hankel function of the first kind:

$$\begin{aligned} H_{\nu_\phi}^{(1)}(x \gg 1) &\sim \sqrt{\frac{2}{\pi x}} e^{i\left(x - \frac{\pi}{2}\nu_\phi - \frac{\pi}{2}\right)} \quad \text{(for sub-Hubble scales)} \\ H_{\nu_\phi}^{(1)}(x \ll 1) &\sim \sqrt{\frac{2}{\pi}} e^{-i\frac{\pi}{2}} 2^{\nu_\phi - \frac{3}{2}} \frac{\Gamma(\nu_\phi)}{\Gamma(\frac{3}{2})} x^{\nu_\phi} \quad \text{(for super-Hubble scales).} \end{aligned} \tag{3.64}$$

After some algebra, we find that the power spectrum is given by

$$P_{\delta\phi}(k) = \frac{H^2}{(2\pi)^2}\left(\frac{k}{aH}\right)^{3-2\nu_\phi}. \tag{3.65}$$

Then we can calculate the scale-dependence

$$\begin{aligned} \frac{\mathrm{d}\ln P_{\delta\phi}(k)}{\mathrm{d}\ln k} = 3 - 2\nu_\phi &= 3 - 2\sqrt{\frac{9}{4} - \frac{V''}{H^2}} \\ &\simeq 3 - 3\left(1 - \frac{2}{9}\frac{V''}{H^2}\right) \\ &= \frac{2}{3}\frac{V''}{H^2} \\ &\overset{(3.34a)}{\simeq} \frac{2}{3}\frac{V''}{\frac{8\pi G}{3}V} = 2\eta_V. \end{aligned} \tag{3.66}$$

Going to the quasi de Sitter expansion, we again pick up an extra $-\epsilon_H \simeq -2\epsilon_V$, then

$$\frac{\mathrm{d}\ln P_{\delta\phi}(k)}{\mathrm{d}\ln k} \simeq 2\eta_V - 2\epsilon_V. \tag{3.67}$$

This can be positive or negative, but is always close to zero. So it is an almost scale-invariant spectrum.

3.5 Metric perturbations

We looked at the field perturbation and worked in the framework of GR. But any perturbation in ϕ will induce perturbation in the metric $g_{\mu\nu}$ according to the Einstein's equations, $\delta\phi \leftrightarrow \delta g_{\mu\nu}$. Then the metric tensor can be decomposed into

$$g_{\mu\nu}(\vec{x}, t) = \overline{g}_{\mu\nu}(t) + \delta g_{\mu\nu}(\vec{x}, t), \tag{3.68}$$

where $\overline{g}_{\mu\nu}(t)$ is a flat FLRW metric. The metric perturbations can be decomposed according to their spin with respect to a local rotation of the spatial coordinates on hypersurfaces of constant time. This leads to scalar perturbations, vector perturbations, and tensor perturbations. Assuming spatial flatness,

$$(\overline{g}_{\mu\nu} + \delta g_{\mu\nu})\mathrm{d}x^\mu \mathrm{d}x^\nu = a^2(t)\big[-(1+2A)\mathrm{d}\tau^2 + 2B_i \mathrm{d}x^i \mathrm{d}\tau + (\delta_{ij} + h_{ij})\mathrm{d}x^i \mathrm{d}x^j\big]. \tag{3.69}$$

In total, we have 10 degrees of freedom.

Tensor perturbations or gravitational waves have spin 2 and are the "true" degrees of freedom of the gravitational fields in the sense that they can exist even in the vacuum. Vector perturbations are spin 1 modes arising from rotational velocity fields and are also called vorticity modes. Finally, scalar perturbations have spin 0.

3.5.1 Scalar-vector-tensor decomposition of $\delta g_{\mu\nu}$

We can first decompose B_i as

$$B_i = \partial_i B + B_i, \quad \text{with} \quad \partial^i B_i = 0, \tag{3.70}$$

and h_{ij} as

$$h_{ij} = -2C\delta_{ij} + 2\underbrace{\left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\nabla^2\right)}_{D_{ij}} E + \partial_i E_j + \partial_j E_i + 2E_{ij}. \tag{3.71}$$

Then we have 4 scalar + 4vector + 2 tensor degrees of freedom. (scalar perturbations: density perturbations, electric fields; vector perturbations: magnetic fields, vorticity field; tensor perturbations: gravitational waves.) At linear order in perturbation theory, different types of perturbations are independent of each other.

Let us focus on the scalar perturbations for now. The general scalar perturbation of the metric can be written as

$$\delta g^S_{\mu\nu} = \begin{pmatrix} -2A & \partial_i B \\ \partial_i B & -2C\delta_{ij} + 2D_{ij}E \end{pmatrix}. \tag{3.72}$$

Given the perturbed metric and energy-momentum tensor, one can derive the perturbed Einstein and Klein-Gordon equations.

3.5.2 The gauge issue

GR is symmetric under general coordinate transformations:

$$g_{\mu\nu}(x) \to \tilde{g}_{\mu\nu}(\tilde{x}) = \frac{\mathrm{d}x^\alpha}{\mathrm{d}\tilde{x}^\mu}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tilde{x}^\nu} g_{\alpha\beta}(x). \tag{3.73}$$

We want to make sure that we calculate quantities that are not dependent on the choice of coordinates. The issue is that the inflation perturbation $\delta\phi$ is not gauge invariant! We need to identify a set of quantities that are gauge invariant (invariant under general coordinate transformations). Altogether, there are $2+1+1$ independent gauge-invariant quantities. The set that we choose is called the **Bardeen variables**. If we define

$$\mathcal{H} \equiv \frac{a'}{a} = aH, \tag{3.74}$$

then the Bardeen variables can be written as

$$\begin{aligned} \Psi &= A + \mathcal{H}(B - E') + (B - E')' \\ \Phi &= -C + \mathcal{H}(B - E') - \frac{1}{3}\nabla^2 E \\ \Phi_i &= E_i' - B_i \\ E_{ij}&. \end{aligned} \tag{3.75}$$

After identifying appropriate gauge-invariant quantities, one can remove 2 scalar and 2 vector degrees of freedom by fixing the gauge. For example, in the Newtonian gauge, $B = E = 0$, then

$$g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = a^2(\tau)\big[-(1+2\Psi)\mathrm{d}\tau^2 + (1+2\Phi)\delta_{ij}\mathrm{d}x^i\mathrm{d}x^j\big]. \tag{3.76}$$

We note that Ψ is the gravitational potential. As another example, for a scalar field, we can show the spatial part of δT^μ_ν is diagonal,

$$\delta T^i_j = \delta^i_j. \tag{3.77}$$

This means that there is no anisotropic stress in a scalar field (the anisotropic stress Π^i_j is traceless, $\delta^j_i \Pi^i_j = 0$, and transverse, $\partial^i \Pi_{ij} = 0$). What follows is

$$\Psi + \Phi = 0. \tag{3.78}$$

Therefore, in this case we only need to consider one scalar degree of freedom. A particular useful way to choose this d.o.f. is the **comoving curavature perturbation** $\mathcal{R}$, which is gauge invariant and conserved on super-Hubble scales. In the Newtonian gauge, it is given by

$$\mathcal{R} = \Phi + H\frac{\delta\phi}{\dot{\phi}} = \Phi + \mathcal{H}\frac{\delta\phi}{\phi'}. \tag{3.79}$$

Solving the perturbed Klein-Gordon equation and Einstein's equations, we can construct the power spectrum of $\mathcal{R}$:

$$P_\mathcal{R}(k) = \frac{k^3}{4\pi^2 m_{pl}^2}\frac{1}{\epsilon_H}|\delta\phi_k|^2, \tag{3.80}$$

where $m_{pl} = 1/\sqrt{G}$ is the Planck mass. For a given k, we evaluate this at $k = aH$,

$$|\delta\phi_k|^2 \sim \frac{H^2}{2k^3}\left(\frac{k}{aH}\right)^{\frac{3}{2}-\tilde{\nu}_\phi}, \tag{3.81}$$

where
$$\tilde{\nu}_\phi = \frac{9}{4} + 9\epsilon_H - 3\eta_H, \tag{3.82}$$
then Eq. (3.80) gives
$$\begin{aligned} P_\mathcal{R}(k) &= \frac{1}{8\pi^2\epsilon_H}\frac{H^2}{m_{pl}^2}\bigg|_{k=aH} \\ \Longrightarrow \quad \frac{\mathrm{d}\ln P_\mathcal{R}(k)}{\mathrm{d}\ln k} &\equiv n_S - 1 \simeq 2\eta_H - 4\epsilon_H \simeq 2\eta_V - 6\epsilon_V, \end{aligned} \tag{3.83}$$
where n_S is called the spectral index. Then the phenomenological parametrization is
$$P_\mathcal{R}(k) \simeq A_S\left(\frac{k}{k_*}\right)^{n_S-1}. \tag{3.84}$$
The Planck CMB anisotropy data implies $A_S \simeq (2.142 \pm 0.049) \times 10^{-9}$, which gives $n_S \simeq 0.9667 \pm 0.0004$.

During inflation the curvature perturbation is generated on super-Hubble scales with a spectrum which is nearly scale invariant, i.e. nearly independent from the wavelength $\lambda = \pi/k$; the amplitude of the fluctuation on superhorizon scales does not (almost) depend upon the time at which the fluctuations crosses the horizon and becomes frozen in. The small tilt of the power spectrum arises from the fact that the inflaton field is massive, giving rise to a non-vanishing η_V and because during inflation the Hubble rate is not exactly constant, but nearly constant, where 'nearly' is quantified by the slow-roll parameters ϵ_V.

3.5.3 Tensor perturbations

The tensor partial of $\delta g_{\mu\nu}$ is simply given by
$$\delta g^T_{\mu\nu} = a^2(\tau)h_{ij}\mathrm{d}x^i\mathrm{d}x^j, \tag{3.85}$$
where h_{ij} is a traceless and transverse tensor with 2 d.o.f. (polarizations):
$$h_{ij} = h_+e^+_{ij} + h_\times e^\times_{ij}, \tag{3.86}$$
where e^+_{ij} and $e^\times_{ij}$ are the polarization tensors that satisfy
$$\begin{aligned} &e_{ij} = e_{ji}, \quad k^ie_{ij} = 0, \quad e_{ii} = 0 \\ &e_{ij}(-\vec{k},\lambda') = e^*_{ij}(\vec{k},\lambda) \qquad (\lambda \in \{+,\times\}) \\ &\sum_\lambda e^*_{ij}(\vec{k},\lambda)e^{ij}(\vec{k},\lambda) = 4. \end{aligned} \tag{3.87}$$
Luckily, h_{ij} is already gauge-invariant. Its action reads
$$S = \frac{m^2_{Pl}}{2}\int \mathrm{d}^4x\sqrt{-g}\partial_\mu h^{ij}\partial^\mu h_{ij}. \tag{3.88}$$

Notice that this is the action of a massless scalar field, so we can use the result obtained previously. After properly accounting for d.o.f., the power spectrum is given by

$$\begin{aligned} P_T(k) &= \frac{2}{\pi^2}\frac{H^2}{m_{pl}^2}\bigg|_{k=aH} \\ \frac{\mathrm{d}\ln P_T(k)}{\mathrm{d}\ln k} &= -2\epsilon_H. \end{aligned} \tag{3.89}$$

The power law is then written as

$$P_T(k) = A_T\left(\frac{k}{k_*}\right)^{n_T}. \tag{3.90}$$

Defining the tensor-to-scalar ratio

$$r \equiv \frac{A_T}{A_S} = 16\epsilon_H = -8n_T, \tag{3.91}$$

we obtain the slow-roll consistency relation:

$$n_T = -\frac{1}{8}r. \tag{3.92}$$

3.5.4 Summary of predictions of inflation

- Flat universe: $\Omega_k = 0$ ($\sim 10^{-5}$).
- Scalar perturbations: $P_{\mathcal{R}} \simeq A_S(k/k_*)^{n_S-1}$.
- Tensor perturbations: $P_T \simeq A_T(k/k_*)^{n_T}$, with $r = A_T/A_S = -8n_T$. It turns out that $r < 0.1$ at 90% C.L.

Chapter 4

Structure Formation

We have seen how inflation can generate initial inhomogeneities. Now let us look at how these perturbations evolve at later times (post-inflationary evolution).

4.1 Newtonian perturbation theory

As long as $k/a \gg H$, i.e. scales are within the Hubble radius, Newtonian gravity is a reasonable approximation.

4.1.1 Static universe

We ignore expansion for now and assume a non-relativistic fluid with a mass density ρ, pressure $p \ll \rho$ and velocity $\vec{u}$. What do we need to describe the dynamics of it? First, we have mass conservation, which leads to the continuity equation

$$\dot{\rho} + \boldsymbol{\nabla} \cdot (\rho \vec{u}) = 0. \tag{4.1}$$

Then we have momentum conservation, leading to the Euler equation

$$\dot{\vec{u}} + (\vec{u} \cdot \boldsymbol{\nabla})\vec{u} + \frac{1}{\rho}\boldsymbol{\nabla} p + \boldsymbol{\nabla}\Psi = 0, \tag{4.2}$$

where Ψ is the gravitational potential. Finally, the Newtonian gravity leads to the Poisson equation

$$\nabla^2 \Psi = 4\pi G \rho. \tag{4.3}$$

Next, we split everything into background and perturbations:

$$\begin{aligned} \rho(x,t) &= \overline{\rho}(t) + \delta\rho(x,t) \\ p(x,t) &= \overline{p}(t) + \delta p(x,t) \\ \vec{u}(x,t) &= \cancel{\overline{\vec{u}}(t)} + \delta\vec{u}(x,t) \\ \Psi(x,t) &= \overline{\Psi}(t) + \delta\Psi(x,t). \end{aligned} \tag{4.4}$$

These lead to the following three equations (ignoring all terms of higher than linear order in the perturbations):

$$\dot{\delta\rho} + \overline{\rho}(\boldsymbol{\nabla} \cdot \vec{u}) = 0, \tag{4.5a}$$

$$\dot{\vec{u}} + \frac{1}{\overline{\rho}}\boldsymbol{\nabla}\delta p + \boldsymbol{\nabla}\delta\Psi = 0, \tag{4.5b}$$

$$\nabla^2 \delta\Psi = 4\pi G \delta\rho. \tag{4.5c}$$

For adiabatic perturbation, we have $\delta p = c_s^2 \delta\rho$, where c_s is the speed of sound. Then Eq. (4.5b) becomes

$$\dot{\vec{u}} + \frac{c_s^2}{\overline{\rho}} \boldsymbol{\nabla} \delta\rho + \boldsymbol{\nabla} \delta\Psi = 0. \tag{4.6}$$

Differentiating Eq. (4.5a) w.r.t. time gives

$$\delta\ddot{\rho} + \cancel{\dot{\overline{\rho}}} \boldsymbol{\nabla} \cdot \vec{u} + \overline{\rho} \boldsymbol{\nabla} \cdot \dot{\vec{u}} = 0. \tag{4.7}$$

Substituting $\dot{\vec{u}}$ with Eq. (4.5b) into Eq. (4.7) yields

$$\delta\ddot{\rho} + \overline{\rho}\, \boldsymbol{\nabla} \cdot \left(-\frac{c_s^2}{\overline{\rho}} \boldsymbol{\nabla} \delta\rho - \boldsymbol{\nabla} \delta\Psi \right) = 0, \tag{4.8}$$

or

$$\delta\ddot{\rho} - c_s^2 \nabla^2 \delta\rho - \overline{\rho} \nabla^2 \delta\Psi = 0. \tag{4.9}$$

Using Eq. (4.5c) in Eq. (4.9) leads to

$$\delta\ddot{\rho} - \big(c_s \nabla^2 + 4\pi G \overline{\rho} \big) \delta\rho = 0. \tag{4.10}$$

This is just a wave equation admitting plane-wave solutions

$$\delta\rho = A_{\pm} e^{-i\vec{k}\cdot\vec{x} \pm i\omega t}, \tag{4.11}$$

with

$$\omega^2 = c_s^2 k^2 - 4\pi G \overline{\rho}. \tag{4.12}$$

For $\omega^2 > 0$, $\delta\rho$ oscillates around the origin. In this case, the pressure counteracts gravity and there is no collapse. These are called the acoustic oscillations. On the other hand, for $\omega^2 < 0$, ignoring the damped solution, $\delta\rho$ grows exponentially. This leads to the gravitational collapse. Finally, for the dividing line $\omega^2 = 0$, it leads to $k_J = (4\pi G \overline{\rho}/c_s^2)^{1/2}$, which is called the **Jeans wavenumber**. Therefore, the large-scale perturbations ($k < k_J$) grow, while the small-scale ones ($k > k_J$) don't.

4.1.2 Expanding space in Newtonian gravity

We can "simulate" the expanding space in the Newtonian picture by going to the comoving coordinates and replacing $\vec{x}$ with $a(t)\vec{x}$, $\vec{u}$ with $H(t)\vec{x} + \vec{u}$. Let us also define a density contrast, $\delta \equiv \delta\rho/\overline{\rho}$. Making these substitutions, Eq. (4.10) becomes

$$\ddot{\delta} + 2H\dot{\delta} - \left(\frac{\nabla^2}{a^2} c_s^2 + 4\pi G \overline{\rho} \right) \delta = 0, \tag{4.13}$$

or in Fourier space,

$$\ddot{\delta}_k + 2H\dot{\delta}_k - \left(\frac{k^2}{a^2} c_s^2 - 4\pi G \overline{\rho} \right) \delta_k = 0. \tag{4.14}$$

Solving this differential equation, we see that the Jeans scale is the same as in the static case but picks up a time dependence via $\overline{\rho}(t)$ and $c_s^2(t)$. The effect of the damping term is then

- $k > k_J$: acoustic oscillation with decreasing amplitude;
- $k < k_J$: growth, but no longer exponential.

4.1.3 Dark matter perturbations on sub-Hubble scales

Let us take a closer look at the dark matter (DM) fluctuations in different eras of the universe's history.

We start with matter domination, where $a(t) \propto t^{2/3}$ and thus $H = 2/3t$. So Eq. (4.13) becomes

$$\begin{aligned} &\ddot{\delta}_m + 2H\dot{\delta}_m + \left(\frac{c_s^2}{a^2}\nabla^2 - 4\pi G\overline{\rho}_m\right)\delta_m = 0 \\ \implies \quad &\ddot{\delta}_m + \frac{4}{3t}\dot{\delta}_m - \frac{2}{3t^2}\delta_m = 0, \end{aligned} \tag{4.15}$$

where we have used one of the Friedmann equations: $8\pi G\overline{\rho}_m = 3H^2$, and $c_s = 0$ for DM. Then if we make an Anzatz: $\delta_m \propto t^\beta$, we get

$$\delta_m \propto \begin{cases} t^{-1} \propto a^{-3/2} & \text{decaying mode} \\ t^{3/2} \propto a & \text{growing mode} \end{cases}. \tag{4.16}$$

So in the matter-dominated era, the DM perturbations grow proportional to the scale factor. At matter-radiation equality ($z \simeq 3500$), $\delta_m \sim \mathcal{O}(10^{-5})$. If the universe contained only DM, we would still have $\delta_m < 1$ today.

Next, during radiation domination, $a(t) \propto t^{1/2}$, and we have

$$\begin{aligned} &\ddot{\delta}_m + 2H\dot{\delta}_m - 4\pi G(\overline{\rho}_r\delta_r + \overline{\rho}_m\delta_m) = 0 \\ \implies \quad &\ddot{\delta}_m + \frac{1}{t}\dot{\delta}_m - 4\pi G(\overline{\rho}_r\delta_r + \overline{\rho}_m\delta_m) = 0. \end{aligned} \tag{4.17}$$

The radiation part has $c_s^2 \simeq 1/3$, so the Jeans scale for radiation perturbations is

$$\begin{aligned} k_{J,r} = \left(\frac{4\pi G\overline{\rho}a^2}{c_s^2}\right)^{\frac{1}{2}} &\simeq \left(\frac{3H^2a^2}{2c_s^2}\right)^{\frac{1}{2}} \\ &\simeq \left(\frac{9}{2}H^2a^2\right)^{\frac{1}{2}} \sim aH. \end{aligned} \tag{4.18}$$

Therefore, radiation perturbations oscillate on all sub-Hubble scales, and so $\langle\delta_r\rangle_t = 0$. Eq. (4.17) simplifies to

$$\ddot{\delta}_m + \frac{1}{t}\dot{\delta}_m - 4\pi G\overline{\rho}_m\delta_m = 0. \tag{4.19}$$

Furthermore, being pressure-less, δ_m can only evolve on cosmological time scales, so

$$\ddot{\delta}_m \sim H^2\delta_m \sim \frac{8\pi G}{3}\overline{\rho}_r\delta_m \gg 4\pi G\overline{\rho}_m\delta_m. \tag{4.20}$$

Then we have

$$\ddot{\delta}_m + \frac{1}{t}\dot{\delta}_m = 0. \tag{4.21}$$

Solving this gives

$$\delta_m \propto \begin{cases} \text{const.} \\ \ln t \propto \ln a \end{cases}. \tag{4.22}$$

This implies that in the radiation-dominated era, the DM perturbations grow logarithmically with the scale factor.

4.1.4 Adiabatic vs isocurvature initial conditions

If there are two different perfect fluids, we can express the total fluctuation as

$$\delta\rho_{\text{tot}}(x) = \delta\rho_1(x) + \delta\rho_2(x). \tag{4.23}$$

We can also express the total fluctuation in terms of the **adiabatic perturbation** and the **isocurvature perturbation** (in a different basis):

$$\delta\rho_{\text{tot}}(x) = \delta A + \delta J, \tag{4.24}$$

where

$$\delta A : \ \delta S = 0 \quad \Longrightarrow \quad \frac{n_1}{n_2} = \text{const.} \tag{4.25}$$

and

$$\delta J : \ \delta\Psi = 0 \quad \Longrightarrow \quad \delta\rho_1 = -\delta\rho_2. \tag{4.26}$$

Note that if inflation was single-field, $\delta S = 0$. Therefore, the perturbations must be adiabatic.

Chapter 5

Baryonic Astrophysics

Here are some sample exam questions.

1. Big Bang & cosmic fog

(i) Hydrogen was mostly formed during the recombination, and helium was mostly formed during the Big Bang nucleosynthesis (BBN).

(ii) The spectra of quasars could be used to date the final stages of the reionization epoch. If the quasar is so distant that the light we observe from it escaped during the 'dark ages', its UV light will have been absorbed by the neutral hydrogen present at the time; if the quasar is closer and the light we observe was emitted only after the reionization, there will have been no neutral hydrogen to impede it. (Note that while the neutral hydrogen atoms absorb all wavelengths of light, most wavelengths are released again. UV light, in contrast, ionizes the atoms and is completely absorbed.)

(iii) (While the electrons of neutral hydrogen can absorb photons of some wavelengths by rising to an excited state, a universe full of neutral hydrogen will be relatively opaque only at those absorbed wavelengths, but transparent throughout most of the spectrum.) Once objects started to condense in the early universe (gravitational effect), they were energetic enough to re-ionize neutral hydrogen. Then the universe reverted from being neutral, to once again being an ionized plasma. So in short, reionization caused the fog to lift. (At this time, however, matter had been diffused by the expansion of the universe, and the scattering interactions of photons and electrons were much less frequent than before electron-proton recombination. Thus, a universe full of low density ionized hydrogen will remain transparent, as is the case today.)

2. Gas across the Universe

(i)

- Intergalactic medium (IGM): The IGM is the hot, X-ray emitting gas that permeates the space between galaxies. It consists mostly of ionized hydrogen, that is, a plasma consisting of equal numbers of electrons and protons. As gas falls into the IGM from the voids, it heats up to temperatures of 10^5 K to 10^7 K, which is high enough so that the collisions between atoms have enough energy to cause the bound electrons to escape from the hydrogen nuclei. At these temperatures, it is called the warm-hot intergalactic medium (WHIM).
- Intracluster medium (ICM): The ICM is the superheated plasma present at the center of a galaxy cluster. This gas is composed mainly of ionized hydrogen and helium and strongly emits X-ray radiation. It is also enriched with heavy elements, such as iron.
- Interstellar medium (ISM): The ISM is the matter that exsists in the interstellar space, the physical space within a galaxy beyond the influence of each star on the plasma. Approximately 70% of the mass of the interstellar medium consists of lone hydrogen atoms; most of the remainder consists of helium atoms. This is enriched with trace amounts of heavier atoms

formed through stellar nucleosynthesis. A number of molecules, dust and cosmic rays also exist in ISM. Stars form within the densest regions of the ISM, molecular clouds, and replenish the ISM with matter and energy through planetary nebulae, stellar winds, and supernovae.

(ii) Both the IGM and ICM are at high temperatures, they mostly emit X-ray radiation by the bremsstrahlung process and X-ray emission lines from the heavy elements. These X-rays can be observed using an X-ray telescope. On the other hand, one of the simplest ways to detect cold interstellar gas clouds is by taking spectra of distant stars. These spectra will show the absorption lines indicative of the spectral type of the star, but may also contain additional absorption lines if the light has passed through gas along the way to our detectors. These interstellar absorption lines are created when cold ISM absorbs some of the radiation emitted by the distant star. They tend to be much sharper and more narrow than the absorption lines created in the atmosphere of the star, and indicate the temperature, density and chemical composition of the interstellar gas through which the light has passed.

3. Stellar chemistry

(i) The s-process or slow-neutron-capture-process is a nucleosynthesis process that occurs at relatively low neutron density and intermediate temperature conditions in stars. Under these conditions heavier nuclei are created by neutron capture, increasing the atomic mass of the nucleus by one. A neutron in the new nucleus decays by β^- decay to a proton, creating a nucleus of higher atomic number. The rate of neutron capture by atomic nuclei is slow relative to the rate of radioactive β^- decay. The s-process (or any neutron capture process) is secondary, meaning that it requires preexisting heavy isotopes (e.g. ^{56}Fe) as seed nuclei to be converted into other heavy nuclei. It produces approximately half of the isotopes of the elements heavier than iron, and therefore plays an important role in the galactic chemical evolution. Some isotopes are only produced via the s-process, and they include ^{86}Sr, ^{96}Mo, ^{104}Pd, ^{116}Sn.

(ii) The r-process or rapid-neutron-capture-process is a nucleosynthesis process that occurs in core-collapse supernovae (supernova nucleosynthesis at high temperature and high neutron density) and is responsible for the creation of approximately half of the neutron-rich atomic nuclei heavier than iron. The process entails a succession of rapid neutron captures by heavy seed nuclei, typically ^{56}Fe or other more neutron-rich heavy isotopes. An example of r-process isotope is ^{96}Zr.

(iii) Three processes which affect the process of climbing the neutron drip line are; a notable decrease in the neutron-capture cross section at nuclei with closed neutron shells, the inhibiting process of photodisintegration, and the degree of nuclear stability in the heavy-isotope region. This last phenomenon terminates the r-process when its heaviest nuclei become unstable to spontaneous fission.

4. Stellar physics

(i) A black body is an idealized physical body that absorbs all incident electromagnetic radiation, regardless of frequency or angle of incidence. A black body in thermal equilibrium (that is, at a constant temperature) emits electromagnetic radiation called black-body radiation. The radiation is emitted according to Planck's law, meaning that it has a spectrum that is determined by the temperature alone.

(ii) For a black body, its luminosity is given by

$$L = 4\pi R^2 \sigma T^4 = 4\pi \left(\frac{D}{2}\right)^2 \sigma T^4, \tag{5.1}$$

where σ is the Stefan-Boltzmann constant. This implies

$$\frac{L}{L_\odot} = \frac{D^2}{D_\odot^2} \cdot \frac{T^4}{T_\odot^4}. \tag{5.2}$$

Therefore,

$$\frac{D}{D_\odot} = \frac{T_\odot^2}{T^2}\sqrt{\frac{L}{L_\odot}}. \tag{5.3}$$

(iii) The luminosity of the star can be measured experimentally. The temperature of a star can be most accrately inferred from its spectral features. Then we can use Eq. (5.3) to determine the star's diameter since the size, temperature and luminosity of the Sun are all known parameters.

(iv) More massive stars are brighter, which also suggests that they consume their fuel much faster. So a star's lifetime is determined by its mass (fuel supply) and luminosity (fuel consumption):

$$t \propto \frac{M}{L}. \tag{5.4}$$

The mass-luminosity relation $L \propto M^4$ then gives

$$t \propto M^{-3}. \tag{5.5}$$

So more massive stars live shorter lives.

5. Galaxy luminosity functions

(i) The Schechter luminosity function (LF) can be written as

$$\Phi(L)\mathrm{d}L = n_* \left(\frac{L}{L_*}\right)^{\alpha} \exp\left(-\frac{L}{L_*}\right)\frac{\mathrm{d}L}{L_*}. \tag{5.6}$$

$\Phi(L)$ gives the number of galaxies in a sample per unit luminosity per unit volume. n_* is a normalization factor which defines the overall density of galaxies. L_* is a characteristic galaxy luminosity. α defines the "faint-end slope" (power law slope at low L) of the LF. The LF gives the number galaxies per luminosity interval. Given a luminosity as input, the luminosity function essentially returns the abundance of objects with that luminosity (specifically, number density per luminosity interval). They are used to study the properties of large groups or classes of objects, such as the stars in clusters or the galaxies in the Local Group.

(ii) Given the LF $\Phi(L)$, the galaxy luminosity density (of galaxies with luminosity greater than L) can be computed as

$$\int_L^\infty \Phi(L)L\,\mathrm{d}L. \tag{5.7}$$

6. Galaxy evolution

(i) Assuming all galaxies are fading as they age, then the time axis is in the opposite direction to the luminosity axis. Therefore, the value of a given luminosity function increases as the galaxies age. This is as expected, as luminous galaxies become exponentially rarer at high enough luminosities, similar to a general result in cosmology – the number density of massive objects (e.g. clusters) drops off exponentially at high masses.

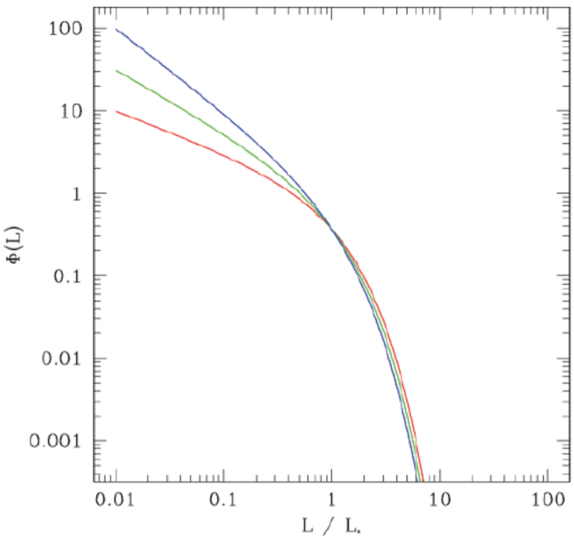


Figure 5.1: The evolution of the Schechter luminosity function with luminosity, with different faint-end slopes.

(ii) A magnitude-limited survey means that one can only probe the most luminous/massive galaxies at higher redshifts (because at high z, we cannot observe faint galaxies). Then we lose the low-mass objects in this process.

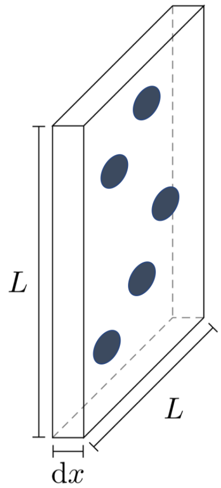
(iii) Jeans instability: In stellar physics, the Jeans instability causes the collapse of interstellar gas clouds and subsequent star formation. It occurs when the internal gas pressure is not strong enough to prevent gravitational collapse of a region filled with matter.

Taylor expansion:

$$\begin{aligned} & f(x+\Delta x) \\ = & f(x) + \frac{f'(x)}{1!}\Delta x + \frac{f''(x)}{2!}(\Delta x)^2 + \cdots \\ = & \sum_{k=0}^{\infty} \frac{f^{(k)}(x)}{k!}(\Delta x)^k \end{aligned} \tag{5.8}$$

Appendix A

Mean free path



The mean free path is the average distance traveled by a moving particle between successive collisions (typically in a medium). It is given by

$$\ell = \frac{1}{\sigma n}, \tag{A.1}$$

where n is the number density of target particles (in the medium) and σ is the effective cross-sectional area for collision (e.g. the scattering cross-section of one atom in the medium).

To derive this formula, let us imagine a beam of particles being shot through a target, and consider an infinitesimally thin slab of the target (see figure). The atoms in the target that might stop a beam particle are shown as blobs in the slab. The area of the slab is L^2 and its volume is $L^2\mathrm{d}x$. A typical number of atoms in the slab is therefore $nL^2\mathrm{d}x$, where n is the number density of the atoms. So the probability that a beam particle will be stopped in this slab is the net area of the stopping atoms divided by the total area of the slab (assuming that a beam particle will be stopped completely once it collides with one of the atoms in the slab):

$$P(\text{stopping within } \mathrm{d}x) = \frac{A_{\text{atoms}}}{A_{\text{slab}}} = \frac{(nL^2\mathrm{d}x)\sigma}{L^2} = n\sigma\mathrm{d}x, \tag{A.2}$$

where σ is the area (or the scattering cross-section) of one atom. Then the drop in beam intensity equals the incoming beam intensity multiplied by the probability of a particle being stopped within the slab:

$$\mathrm{d}I = -In\sigma\mathrm{d}x, \quad \text{or} \quad \frac{\mathrm{d}I}{\mathrm{d}x} = -In\sigma \equiv \frac{I}{\ell}. \tag{A.3}$$

The solution to the ODE is known as the Beer-Lambert law and has the form

$$I(x) = I_0 e^{-x/\ell}. \tag{A.4}$$

Now, note that the probability that a particle is absorbed between x and $x + \mathrm{d}x$ is given by

$$\mathrm{d}P(x) = \frac{I(x) - I(x + \mathrm{d}x)}{I_0} = \frac{1}{\ell} = e^{-x/\ell}\left(1 - e^{-\mathrm{d}x/\ell}\right) \simeq \frac{1}{\ell}e^{-x/\ell}\mathrm{d}x. \tag{A.5}$$

Then the mean distance traveled by a beam particle before being stopped (i.e. mean free path) is given by the expectation value of x:

$$\langle x \rangle = \int_0^\infty x \, \mathrm{d}P(x) = \int_0^\infty \frac{x}{\ell}e^{-x/\ell}\mathrm{d}x = \ell. \tag{A.6}$$

So $\ell = 1/(\sigma n)$ is the mean free path.