Early Universe: Background FRW universe and thermal history

This first part of these lectures on the Early Universe is concerned with its overall content and evolution. To this end, we will consider the simplest possible model where the universe is highly symmetric. The formation and evolution of the structure of the universe is left for the second part of this course. The lectures will roughly follow these topics:

- geometric structure, metric, distance, and the age of the universe
- Einstein equations, matter and radiation in the universe, dynamics of the expansion
- the evolution of the universe, statistical mechanics in and out of equilibrium
- thermal history of the universe

The notes are fairly self-contained, but sections headed "Revision" will not be gone through in lectures: they are revising content you should mostly be familiar with from the cosmology or GR course (but with slightly different notation here in some cases). You should go through these in your own time, and ask if there are any revision topics you are not happy with. In particular Sec. IV A introduces core material on statistical mechanics, deriving the Fermi-Dirac and Bose-Einstein distributions, which may require careful revision to bring yourself up to speed.

Useful numbers for reference

$\hbar = 1.05457148 \times 10^{-34} \mathrm{m}^2\,\mathrm{kg}\,\mathrm{s}^{-1}$ $\qquad c \equiv 299792458\ \mathrm{m}\,\mathrm{s}^{-1}$

$k_B = 1.3806504 \times 10^{-23} \mathrm{m}^2\,\mathrm{kg}\,\mathrm{s}^{-2}\,\mathrm{K}^{-1}$ $\qquad G = 6.67428 \times 10^{-11} \mathrm{m}^3\,\mathrm{kg}^{-1}\,\mathrm{s}^{-2}$

$\sigma_T = 6.6524616 \times 10^{-29} \mathrm{m}^2$ $\qquad \sigma_{SB} = 5.6704 \times 10^{-8} \mathrm{J}\,\mathrm{s}^{-1}\,\mathrm{m}^{-2}\,\mathrm{K}^{-4}$

$m_p = 1.672621637 \times 10^{-27} \mathrm{kg} \approx 0.938 \mathrm{GeV}$ $\qquad m_e = 9.10938215 \times 10^{-31} \mathrm{kg} \approx 0.511 \mathrm{MeV}$

$m_n = 1.6749272 \times 10^{-27} \mathrm{kg}$ $\qquad m_n - m_p \approx 1.293 \mathrm{MeV}$

$1 \mathrm{eV} = 1.60217646 \times 10^{-19} J = 1.78266173 \times 10^{-36} \mathrm{kg}$ $\qquad 1 \mathrm{Mpc} = 3.08568025 \times 10^{22} \mathrm{m}$

$1 \mathrm{GYr} = 3.1556926 \times 10^{16} \mathrm{s}$ $\qquad H_0^{-1} \equiv [100 h \mathrm{km}\,\mathrm{s}^{-1}\,\mathrm{Mpc}^{-1}]^{-1} \approx h^{-1} 9.7781 \mathrm{Gyr}.$

$\zeta(3) \approx 1.202056903$ $\qquad \zeta(4) = \pi^4/90 \approx 1.082323234$

$$\text{Planck mass } m_P \equiv \sqrt{\hbar c/G} = 1.2209 \times 10^{19} \mathrm{GeV} = 2.17644 \times 10^{-8} \mathrm{kg}$$

$$\text{Reduced Planck mass } M_P \equiv \sqrt{\hbar c/(8\pi G)} \approx 2.4354 \times 10^{18} \mathrm{GeV} \approx 4.3414 \times 10^{-9} \mathrm{kg}$$

We will always (except where explicitly stated) use "natural" units where $k_B = c = \hbar = 1$. As an example, a temperature of 1 K corresponds to about 8.617×10^{-5} eV and the Hubble constant $H_0 = 100h$km/s/Mpc is also $h(9.7781\mathrm{Gyr})^{-1}$.
The CMB temperature today is measured to be approximately $T_\gamma^0 \approx 2.7255K$.

I. THE GEOMETRIC STRUCTURE

A. GR revision: The metric

Matter in the universe is on average very accurately charge neutral, so on large scales (long distances) gravity is the dominant force. It is therefore the theory of gravity, General Relativity, that we need to use when studying cosmology. The gravitational field equations (the *Einstein equations*) connect the curvature and the energy content of the universe, and we will spend the first two sections studying them in a cosmological context. The curvature is derived from the metric, which is the basic geometric quantity that we will use to determine the local geometric structure of spacetime.

In general, the metric is a covariant, symmetric tensor with two indices, $g_{\mu\nu}$. In General Relativity in four dimensions (3 space + 1 time) it has (a priori) ten different independent entries. We will often consider the *line element* or infinitesimal 4-distance,

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu. \tag{1}$$

Here we are going to use the $+---$ signature, which may be different to what you have used before when you learnt about General Relativity — beware!

We will try to make life easy by reducing the number of different components of the metric to the minimum possible by using symmetries. The fundamental, *cosmological principle* says that no point in the universe is preferred. Thus the universe needs to be *isotropic* (i.e. look the same in all directions, rotationally symmetric) and *homogeneous* (the same at all points, symmetric under translations). Of course we know the universe is not in fact smooth (the same at all points), because it contains galaxies, stars, planets, etc, but more generally the symmetry properties can be assumed to apply statistically (e.g. the galaxy is a priori equally likely to be *here* as *here*$+\ 10^8$ light years in the x direction). If we look on the very largest scales, the distribution of galaxies in fact looks very random, with a roughly isotropic and homogeneous distribution (e.g. the density of galaxies looks the same in all directions), so on the very largest scales the density does still look homogenous and isotropic. The fundamental assumption that we shall make is

- The large-scale universe is accurately modelled as spatially homogeneous and isotropic, with the geometry determined from the energy densities by General Relativity.

As we shall see, this assumption is self-consistent in that at early times the universe can be very smooth indeed, with gravitational growth gradually forming the structures (galaxies etc) that we see today. What is much less obvious is that this assumption is valid in the late universe, e.g. the last billions of years till today, when the universe is actually very lumpy; in fact it remains an open research question to what extent the assumption is valid. Here we shall simply assume that it is valid, and we shall see that this is sufficient to describe a wide range of cosmological phenomena. In the second half of the course we will systematically look at perturbations about a homogeneous model, and show how these can be used to accurately model the evolution of structure.

The spatially homogeneous and isotropic model is called the Friedmann-Lemaître-Robertson-Walker (FLRW or FRW) model, and making these assumptions the form of the metric can be severely restricted. A formal derivation of the general form can be found in e.g. in "Gravitation and Cosmology" by Steven Weinberg. We shall just use the fact that homogeneity implies that the spatial curvature[1] $R^{(3)}$ must be the same at every point in space, and hence only a function of time t, so $R^{(3)}(t)$. It is useful to define a *scale factor* $a(t)$ determining the overall scale of the universe, so for example $a(t) = 2$ corresponds to a universe which is twice as large as one with $a(t) = 1$. Since Riemann curvature can be defined by the commutation of two covariant derivatives, it depends on two derivatives; so if you re-scale something with curvature $R^{(3)}$ by a factor a, the curvature changes to $R^{(3)}/a^2$ (imagine blowing up a balloon - the surface locally looks less and less curved as you increase the radius). In other words we expect $K \equiv a(t)^2 R^{(3)}(t)/6$ to be constant (the factor of 6 is conventional for convenience).

Then it can be checked that in spherical coordinates

$$ds^2 = dt^2 - a(t)^2 d\Sigma_3^2 = dt^2 - a(t)^2 \left(\frac{dr^2}{1 - Kr^2} + r^2 d\Omega \right) \tag{2}$$

has the required symmetries and 3-curvature $R^{(3)}(t) = 6K/a(t)^2$. Here $d\Omega = d\theta^2 + \sin^2\theta d\phi^2$ is the two-dimensional angular volume element.

---

[1] This can be defined here as the Ricci scalar calculated from the spatial part of the metric

There are three qualitatively distinct values for the 3-curvature : open ($K < 0$), flat ($K = 0$) and closed ($K > 0$). If $K = 0$ and $a(t) = 1$ we have the Minkowski metric.

Conventionally, the scale factor can be normalised in two ways, either so that K takes the values $-1, 0$ or 1, or so that the scale factor today (t_0) is unity, $a(t_0) = a_0 = 1$. Note that we do not have the freedom to set both, either $K = -1, 0$ or 1, or else $a(t_0) = a_0 = 1$; both cannot be true. I prefer the convention where $a(t_0) = 1$, and I will generally make that choice.

It is possible and often useful to rewrite the metric (2) by changing coordinate from r to χ, where $r = S_K(\chi)$ and for general K

$$S_K(\chi) \equiv \begin{cases} \frac{1}{\sqrt{K}} \sin(\sqrt{K}\chi), & K > 0 \\ \chi & K = 0 \\ \frac{1}{\sqrt{|K|}} \sinh(\sqrt{|K|}\chi), & K < 0. \end{cases} \tag{3}$$

The resulting metric is

$$ds^2 = dt^2 - a^2(t) \left(d\chi^2 + S_K(\chi)^2 d\Omega \right). \tag{4}$$

(See the question sheet where you will derive this). The quantity $a(t)\chi$ is the proper radial distance of an object at radial coordinate r at time t. The coordinate χ is called the *comoving distance* because it is unchanged by an overall re-scaling of the universe by $a(t)$. If we set up two observers without relative motion at a comoving distance χ_0, then they will remain at this distance (hence the name), but their physical distance $a\chi_0$ will change over time.

We can also introduce the time analogue to the comoving distance, called *conformal time* η and defined by $dt \equiv a(t)d\eta$. Written in these coordinates, the line element (2) is

$$ds^2 = a(\eta)^2 \left(d\eta^2 - \left\{ \frac{dr^2}{1 - Kr^2} + r^2 d\Omega \right\} \right) \tag{5}$$

and of course the equivalent expression for (4) exists as well. The form given above shows that the metric is conformally related to the Minkowski metric for $K = 0$, i.e. that there is a transformation $g_{\mu\nu} = f\eta_{\mu\nu}$ where f is an arbitrary multiplicative factor (the scale factor in our case).

Remark 1 *The Einstein equations of general relativity link the geometry (the curvature) to the energy (matter / radiation) in the universe. We have so far only considered the former, and we will have to check in the next section if these solutions apply in the case of the matter and radiation we know actually exist.*

B. Revision: the curvature

By construction K is constant on the slices of equal time. Given the three different possibilities for K we distinguish the three following cases:

1. **closed universe:** The finite-sized spatial sections are spheres S^3 and $K > 0$.
2. **flat (critical) universe:** The unbounded spatial sections are flat and $K = 0$.
3. **open universe:** The unbounded spatial sections correspond to the hyperbolic space H^3 with $K < 0$.

In may be helpful to briefly consider simple two-dimensional analogues:

1. flat space: We can use as coordinates either x and y with $dl^2 = dx^2 + dy^2$ or else spherical coordinates r and θ with $dl^2 = dr^2 + r^2 d\theta^2$. The two are related by the coordinate transform $x = r\cos\theta$, $y = r\sin\theta$.

2. closed space: For the surface of a unit sphere we have $x^2 + y^2 + z^2 = 1$ and so $z = \pm\sqrt{1 - r^2}$ where $r^2 = x^2 + y^2$. The line element then becomes

$$\begin{aligned} dl^2 = dx^2 + dy^2 + dz^2 &= dr^2 + r^2 d\theta^2 + \frac{r^2}{1 - r^2} dr^2 \\ &= \frac{dr^2}{1 - r^2} + r^2 d\theta^2. \end{aligned} \tag{6}$$

If we change variable to χ as before, with $r = S_K(\chi)$, then χ describes the radial distance along the surface (e.g. in the closed case the distance along a great circle from the north pole). The coordinate r is the distance in the projection onto the x,y plane, which is somewhat less intuitive from the viewpoint of an observer on the surface.

3. open space: This is more tricky, we'd need to embed in a space with $++-$ signature so that $dl^2 = dx^2 + dy^2 - dz^2$ and $-x^2 - y^2 + z^2 = 1$. Then $z = \pm\sqrt{1+r^2}$. The line element then becomes

$$dl^2 = \frac{dr^2}{1+r^2} + r^2 d\theta^2 \tag{7}$$

Have a look on Wikipedia under *hyperbolic space* for some nice pictures, including classic Esher print representations. In terms of χ, the key difference with the close case is that instead of the circumference of a circle at constant χ being less than $2\pi\chi$, in the open case it is greater than $2\pi\chi$. There is also no periodicity, so like the flat case the surfaces of constant time can be infinite (unbounded, open).

C. Revision: The cosmological redshift

In their rest frames, many substances emit or absorb radiation at characteristic frequencies, for example the Lyman-α transition in hydrogen. By observing known radiation emission or absorbtion at distant sources we can often tell by how much the frequency has changed between emission and when we received it. Consider a static source at a fixed radial comoving distance χ. Since light travels on geodesics with $ds^2 = 0$, from the metric we have for incoming light $a(t)d\chi = -dt$ and hence

$$\chi = \int_0^\chi d\chi' = -\int_{t_2}^{t_1} \frac{dt}{a(t)} = \int_{t_1}^{t_2} \frac{dt}{a(t)}. \tag{8}$$

We can think of t_1 as being the time a wave crest is emitted, and t_2 the time when it is received. We can also consider the next wave crest emitted a time interval $\delta t_1 = 1/\nu_1$ later in the source rest frame, which will reach $\chi = 0$ at a time $t_2 + \delta t_2$. Since the comoving distance of the source is assumed to be fixed we also have

$$\chi = \int_{t_1+\delta t_1}^{t_2+\delta t_2} \frac{dt}{a(t)} = \int_{t_1}^{t_2} \frac{dt}{a(t)}. \tag{9}$$

The only what that this can be true is if

$$\frac{\delta t_2}{a(t_2)} - \frac{\delta t_1}{a(t_1)} = 0 \quad \Longrightarrow \quad \frac{\nu_1}{\nu_2} = \frac{\delta t_2}{\delta t_1} = \frac{a(t_2)}{a(t_1)}. \tag{10}$$

Since we can observe the frequency of the radiation ν_2 when observed, and we know the source frequency ν_1, we can define the very useful observable called the *redshift* between emission and observation, given by

$$1 + z \equiv \frac{\nu_1}{\nu_2} = \frac{a(t_2)}{a(t_1)}. \tag{11}$$

If $t_2 \geq t_1$ and the universe is expanding, as is the case for observation today of objects in the past, then $1 + z \geq 1$ and so the redshift is positive with $0 \leq z < \infty$: frequencies are observed to be redder than when they were emitted, and the ratio simply tells us how much smaller the universe was when the light was emitted. This makes sense, as one can think of the wavelength of the light stretching with the universe as it expands, so the total amount of stretching just gives the overall expansion ratio.

D. Cosmological redshift for massive particles

We can also see the redshifting effect more generally from the geodesic equation for a freely propagating particle. If choose to normalize the affine parameter λ so that the particle's four momentum

p^μ is given by

$$p^\mu = \frac{dx^\mu}{d\lambda}, \tag{12}$$

the geodesic equation becomes

$$\frac{dp^\mu}{d\lambda} + \Gamma^\mu{}_{\alpha\beta}p^\alpha p^\beta = 0. \tag{13}$$

We can align our FRW metric coordinates so that the particle is moving along a purely radial direction, so that $p^\phi = p^\theta = 0$. The $\mu = 0$ component of the geodesic equation for the evolution of $p^0 = E$ then gives

$$\begin{aligned}\frac{dE}{d\lambda} + \Gamma^0{}_{rr}p^r p^r &= \frac{dE}{dt}\frac{dt}{d\lambda} + \frac{a\dot{a}}{1-Kr^2}p^r p^r \\ &= \frac{dE}{dt}E - \frac{\dot{a}}{a}g_{rr}p^r p^r \\ &= \frac{dE}{dt}E + \frac{\dot{a}}{a}\mathbf{p}^2 = 0\end{aligned} \tag{14}$$

where $p_\mu p^\mu = E^2 - \mathbf{p}^2 \equiv E^2 - p^2 = m^2$. Differentiating $E^2 - p^2 = m^2$ implies $EdE = pdp$, so the geodesic equation can be written in terms of the momentum as

$$\frac{1}{p}\frac{dp}{dt} = -\frac{\dot{a}}{a} \implies p \propto 1/a. \tag{15}$$

Hence the momentum of a freely propagating particles decays $\propto 1/a$ as the universe expands, in agreement with our previous result in the case of light (where $p = E = h\nu$, so the energy and frequency also redshift $\propto 1/a$). For a non-relativistic particle with $p \approx mv$, the velocity decays $\propto 1/a$.

II. COSMOLOGICAL DISTANCES

In order to test different cosmological models, distance measures are very useful. We have already discussed redshift, which can be used as a kind of distance measure (if we know the source is static), though to relate redshifts to distance we need to know $a(t)$. There are two basic other ways in which distances can be measured: we can either consider the apparent angular size of an object and compare it to its known diameter, or we can measure the apparent observed luminosity and compare to a known (or modelled) source luminosity of an object.

To think about different distances we will use the metric of Eq. (4), recapping results that were derived/motivated in the Cosmology course without using any GR.

1. *Comoving distance*

For purely radial distances, we have $d\Omega = 0$. For a fixed time ($dt = 0$) the element of proper radial distance is given by $ds = a(t)d\chi$, so at fixed time $a(t)\chi$ is the proper radial distance. The coordinate χ is the comoving distance, which has the overall expansion scale factored out.

For observations of objects in the universe, the relevant line-of-sight comoving distance is the distance travelled by the light since it was emitted. Light travels on geodesics with $ds = 0$, and hence for radial rays the comoving distance travelled by light is determined by

$$ds^2 = dt^2 - a(t)^2 d\chi^2 = 0 \tag{16}$$

$$\implies \chi = \int d\chi = \int \frac{dt}{a(t)} = \int d\eta = \Delta\eta, \tag{17}$$

where $\Delta\eta$ is the change in conformal time between emission and observation. So we shall often be interested in radial comoving distances given by the conformal time, $\chi = \Delta\eta$ (as we were in Sec. I C when we were thinking about the observed frequency of light from a distance source)

2. Transverse comoving distance (comoving angular diameter distance)

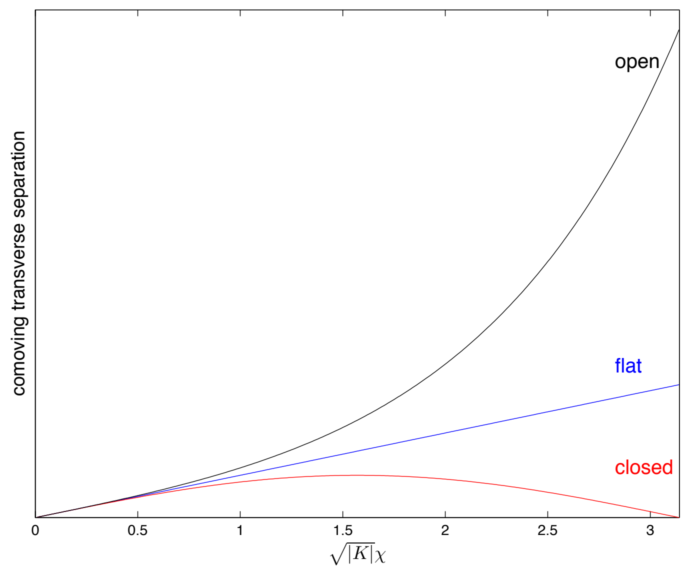

FIG. 1: Imagine sending out radial light rays with some small separation. This figure shows the comoving separation of the rays as a function of the radial comoving distance (in curvature radius units). In the closed universe the rays re-focus at the antipodal distance $\chi = \pi/\sqrt{K}$. In an open universe the rays get forever further apart.

This distance is useful for considering objects of small angular size transverse to the line of sight, i.e. with $d\chi = 0$ at fixed time. If we choose $d\phi = 0$ the proper transverse distance element is $ds = a(t)S_K(\chi)d\theta$, which has comoving length $ds/a(t)$. We define the transverse comoving distance d_m so that

$$\text{transverse comoving distance element} = \frac{ds}{a(t)} = d_m d\theta \tag{18}$$

hence

$$d_m = S_K(\chi). \tag{19}$$

If we imagine sending out two light beams separated by a small angle θ, $d_m\theta = S_K(\chi)\theta$ would be the comoving separation of the beams once they have reached a comoving distance χ. In a flat geometry this is just the standard Euclidean result, $\chi\theta$. However in a closed universe $S_K(\chi) < \chi$, and hence the beams are closer than they would be in a flat universe: the light rays are converging. So looking at objects in a closed universe is a bit like looking through a magnifying lens. Indeed at $\chi = \pi/\sqrt{K}$ the rays focus as the separation goes to zero, and at $\chi = 2\pi/\sqrt{K}$ they converge again at the point of emission having gone all the way round the universe[2]! Conversely in an open universe $S_K(\chi) > \chi$, and the light rays are diverging and become for ever (exponentially) further apart than they would be in a flat universe.

---

[2] Assuming of course the universe actually lasts long enough for light to have the time to go all the way round.

3. *The (physical) angular diameter distance*

For an object of physical length ds (transverse to the line of sight, $d\chi = 0$), observed with small angular size $d\theta$, the (physical) angular diameter distance d_A is defined by

$$ds = d_A d\theta. \tag{20}$$

Since $ds = a(t)S_K(\chi)d\theta$, this is just the same as the comoving angular diameter distance but keeping the expansion factor, hence

$$d_A = a(t)S_K(\chi) = a(t)d_m = \frac{d_m}{1+z}. \tag{21}$$

See Figure 2.

4. *Revision: The luminosity distance*

We assume that we know the intrinsic, absolute luminosity L of a certain object, called a *standard candle*, that radiates isotropically. Type Ia supernovae are believed to be approximate standard candles, and they are used to map out cosmological distances. Let us assume that we observe a flux S from the standard candle with absolute luminosity L at a fixed comoving distance. The definition of the luminosity distance d_L is then

$$S \equiv \frac{L}{4\pi d_L^2} \tag{22}$$

This is shown in Figure 3.

Imagine a pulse of a number of photons reaching us from the source. In the source rest frame the pulse lasts time δt_1 and contains energy $L\delta t_1$. The total number of photons reaching us per comoving area δA will be diluted as the photons move out over a spherical wavefront, so we only receive a fraction $\delta A/(4\pi d_m^2)$ of the photons. The energy of each photon ($h\nu$) is redshifted by a factor $1/(1+z)$ as the frequency is redshifted, so the total energy we will receive is $\delta E = \delta A \delta t_1 L/[(1+z)4\pi d_m^2]$. The flux (energy per unit time per unit area) at observation time t_2 is then

$$S = \frac{\delta E}{\delta A \delta t_2} = \frac{L}{(1+z)4\pi d_m^2}\frac{\delta t_1}{\delta t_2} = \frac{L}{(1+z)^2 4\pi d_m^2} \tag{23}$$

from Eq. (10). Hence the luminosity distance is given by

$$d_L = (1+z)d_m. \tag{24}$$

We see that there is a relation between the luminosity distance and the angular diameter distance, $d_L = (1+z)^2 d_A$. This actually also holds very generally (e.g. in general metrics) using a theorem called the *reciprocity relation* as long as no photons are lost along the line of sight (e.g. by scattering).

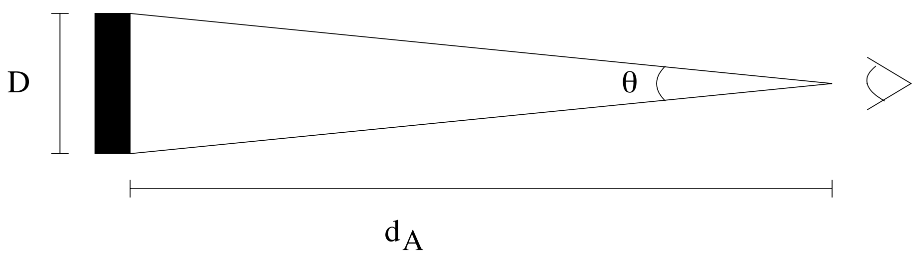

FIG. 2: The angular diameter distance, where an object with known physical size D is observed to have angular size θ.

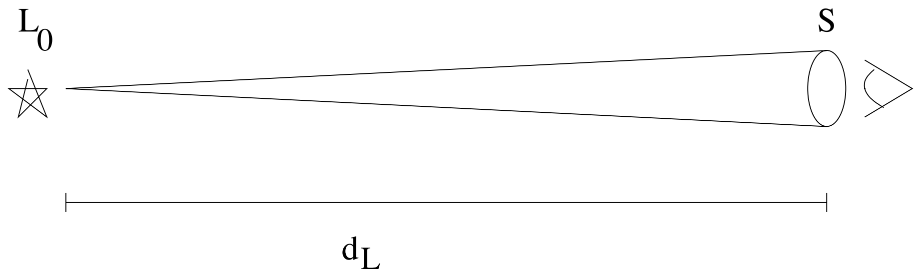

FIG. 3: The luminosity distance, where a source with absolute luminosity L_0 is observed on earth with Flux S.

III. THE DYNAMICS OF THE EXPANSION

A. Revision: Introducing matter and energy

As mentioned in the last section, we are missing a central part of our theory. The metric that we have been using so far,

$$ds^2 = dt^2 - a(t)^2 \left(\frac{dr^2}{1 - Kr^2} + r^2 d\Omega \right) \tag{25}$$

contains a free function $a(t)$. This should not surprise us. So far we have just been using differential geometry and the cosmological principle. Now we need to use General Relativity in order to determine this function from the contents of the universe. To this end, we need to solve the Einstein equations,

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu} + \Lambda g_{\mu\nu}. \tag{26}$$

$G_{\mu\nu}$ is the Einstein tensor, given by a combination the Ricci tensor $R_{\mu\nu} \equiv R^{\alpha}_{\mu\alpha\nu}$ and the Ricci scalar $R \equiv g^{\mu\nu} R_{\mu\nu}$ which are related to the curvature of space-time, and are determined by the metric:

$$\Gamma^{\alpha}_{\mu\nu} = \frac{1}{2} g^{\alpha\beta} \left(g_{\beta\mu,\nu} + g_{\beta\nu,\mu} - g_{\mu\nu,\beta} \right) \tag{27}$$

$$R^{\alpha}_{\beta\mu\nu} = \Gamma^{\alpha}_{\nu\beta,\mu} - \Gamma^{\alpha}_{\mu\beta,\nu} + \Gamma^{\delta}_{\nu\beta} \Gamma^{\alpha}_{\mu\delta} - \Gamma^{\delta}_{\mu\beta} \Gamma^{\alpha}_{\nu\delta}. \tag{28}$$

$T_{\mu\nu}$ is the stress-energy tensor or energy-momentum tensor (EM tensor). It is determined by the matter and energy in the universe. The term Λ is the cosmological constant, which is often considered to be a component of the stress-energy tensor with $T^{(\Lambda)}_{\mu\nu} = \Lambda g_{\mu\nu}/(8\pi G)$.

Clearly, if the left hand side of the Einstein equations have to obey the cosmological principle (isotropy and homogeneity), then the same is required for the right hand side. If the matter in the universe is not homogeneous, then we cannot describe the structure of space-time with a metric that is. Thus, in the inertial frame moving with the matter, the energy-momentum tensor needs to have the following form:

$$T^{\mu}_{0} = 0, \quad T^{1}_{1} = T^{2}_{2} = T^{3}_{3}, \tag{29}$$

and none of the elements can depend on space, only on time. This specific form of the EM tensor arises for ideal or perfect fluids. It is then given by

$$T^{\nu}_{\mu} = \mathrm{diag} \left(\rho(t), -P(t), -P(t), -P(t) \right), \tag{30}$$

where $\rho(t)$ is the energy density of the perfect fluid, and $P(t)$ its pressure. In general, if the observer is moving with respect to the rest-frame of the fluid, the EM tensor can be written in explicitly covariant form as

$$T^{\mu\nu} = -P g^{\mu\nu} + (\rho + P) U^{\mu} U^{\nu} \tag{31}$$

and we recover the previous case when the fluid velocity four-vector is $U^0 = 1$, $U^i = 0$.

The relation between the pressure and the energy density, $P = P(\rho)$ is called the *equation of state*. For (ideal) matter, radiation and vacuum energy it takes a very simple form,

$$P = w\rho, \quad \begin{cases} w = 0 & \text{"pressureless" matter} \\ w = 1/3 & \text{radiation} \\ w = -1 & \text{vacuum energy or cosmological constant} \end{cases} \tag{32}$$

For the third case, we have moved the cosmological constant term $\sim \Lambda g_{\mu\nu}$ from the geometric side of the Einstein equation to the matter side and interpreted it as $T^{(\Lambda)}_{\mu\nu}$ as discussed previously. Comparing it with eq. (30) leads immediately to the equation of state given above.

B. Revision: Energy conservation

Now we need to solve the Einstein equations. As a first step, we remember that the EM tensor is conserved,

$$T^{\mu\nu}_{\ ;\mu} = 0 = G^{\mu\nu}_{\ ;\mu}. \tag{33}$$

The right hand side of this equation follows not only from the Einstein equations, but are also a geometric property in differential geometry and known as *Bianchi identities*. The time-component is

$$T^{\nu}_{0;\nu} = \dot{\rho} + \Gamma^i_{i0}(\rho + P) = \dot{\rho} + 3\left(\frac{\dot{a}}{a}\right)(\rho + P) = 0. \tag{34}$$

It may also be helpful to rewrite this equation as

$$d(\rho a^3) = -P da^3, \tag{35}$$

which can be seen as the first law of thermodynamics, as ρa^3 is proportional to the total energy, and a^3 to the volume. For the simple equation of state $P = w\rho$ with w constant we can also easily solve for $a(t)$ by writing

$$\frac{\dot{\rho}}{\rho} = -3(1+w)\frac{\dot{a}}{a} \quad \Longrightarrow \quad d\ln\rho = -3(1+w)d\ln a. \tag{36}$$

Integrating we find immediately that

$$\rho \propto a(t)^{-3(1+w)} \propto \begin{cases} a(t)^{-3} & \text{for } w = 0 & \text{(pressureless matter)} \\ a(t)^{-4} & \text{for } w = 1/3 & \text{(radiation)} \\ \text{const.} & \text{for } w = -1 & \text{(vacuum energy)} \end{cases} \tag{37}$$

For pressureless matter all the energy is in the mass, and the mass density simply dilutes with the volume scale $a(t)^3$. The energy density in radiation decreases more rapidly because as the universe expands the wavelength and hence energy is also redshifted $\propto 1/a(t)$. In an expanding universe, radiation will dominate at early times, while matter will become more important later on. If there is a non-vanishing contribution from vacuum energy, it will always start to dominate eventually (as shown in figure 4).

C. Revision: The Friedmann equations

The non-zero components of the Ricci tensor and scalar are:

$$R_{00} = -3\frac{\ddot{a}}{a} \tag{38}$$

$$R_{ij} = -\left[\frac{\ddot{a}}{a} + 2\left(\frac{\dot{a}}{a}\right)^2 + 2\frac{K}{a^2}\right] g_{ij} \tag{39}$$

$$R = -6\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2}\right] \tag{40}$$

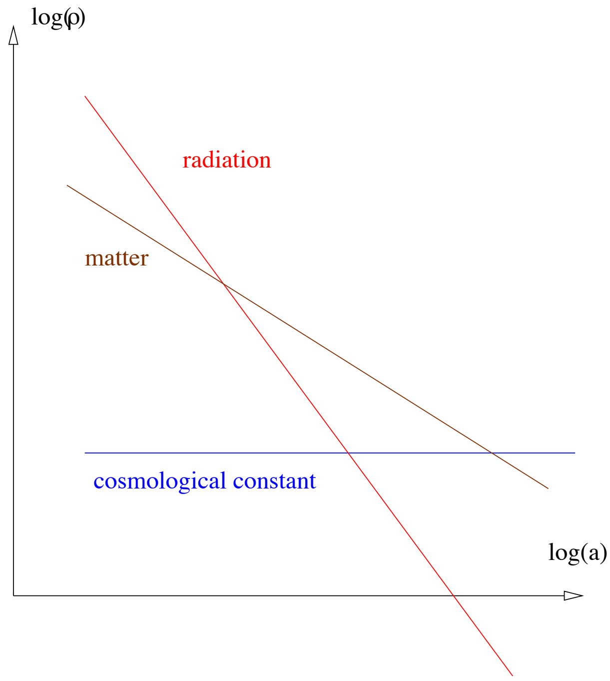

FIG. 4: The densities of the different ingredients of the universe (radiation, matter and a cosmological constant) as a function of scale factor.

The $0-0$ component of the Einstein equations is then

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2} = \frac{8\pi G}{3}\rho. \tag{41}$$

The ρ on the right hand side is the sum of all contributions to the energy density in the universe. This Friedmann equation can also be written in terms of $H = \dot{a}/a$ and $\kappa \equiv 8\pi G$ as

$$H^2 + \frac{K}{a^2} = \frac{\kappa}{3}\rho, \tag{42}$$

which tells us that the expansion rate is directly related to the curvature and density. It may seem unintuitive that universes with *higher* densities expand *faster* (H is bigger), for the same curvature; this is basically because if it didn't the curvature would instead have to be very different, and observations constrain the curvature to be small.

The $i-i$ component is

$$2\left(\frac{\ddot{a}}{a}\right) + \left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2} = -8\pi G P, \tag{43}$$

and if we replace the second and third term on the left hand side by $8\pi G\rho/3$ by using the first Friedmann equation, (41), we find the second Friedmann equation,

$$\left(\frac{\ddot{a}}{a}\right) = -\frac{4\pi G}{3}\left(\rho + 3P\right). \tag{44}$$

The equations (35), (41) and (44) are not independent. Any one of them can be derived from the other two. Written in terms of H and $\kappa = 8\pi G$ we can summarize these important results as

$$H^2 + \frac{K}{a^2} = \frac{\kappa}{3}\rho, \tag{45}$$

$$\dot{H} + H^2 = -\frac{\kappa}{6}(\rho + 3P), \tag{46}$$

$$\dot{\rho} + 3H(\rho + P) = 0. \tag{47}$$

IV. THERMAL HISTORY

A. Revision: equilibrium distributions

If particles interact often and exchange energy, they will rapidly reach an equilibrium state where at any given time each possible micro-configuration with the same total energy is equally likely. For example, for a gas in a box any give atom is just as likely to be in one position in the box as any another. Many of the micro-configurations are however macroscopically indistinguishable, i.e. can be described by a macrostate with a particular temperature T, total number of particles N, etc. The most likely macroscopic state is the one which has the largest fraction of the possible microstate configurations for the system (subject to any constraints, like conservation of energy or particle number), and this is the state the system will almost always be in when it has reached equilibrium. For example, the macroscopic state corresponding to having 50 coins heads and 50 coins tails is vastly more likely in a distribution of 100 coin tosses than a macroscopic state corresponding to 100 heads, because there are vastly more sequences of tosses that give a 50/50 result than an unlikely sequence of 100 heads. Likewise gas particles in a box could all be in one corner, but there are vastly more ways of arranging them dispersed throughout the box, and hence the latter is what we expect to see if the particles are free to move around and randomly interact.

In the case of particles, the form of the most likely distribution depends on whether the particles are bosons (integer spin, like the photon), or fermions (half-integer spin, like neutrinos, which obey the Pauli exclusion principle and hence cannot have more than one particle in each distinct quantum state).

Consider a set of energy levels, each having energy ϵ_i and occupied by n_i particles. The levels may be degenerate, in that there are g_i distinct quantum states (sub-levels) all of the same energy. For example in the case of a gas of particles the energy is determined by the particle momentum, and there are many states of the same energy because the particle could be in many possible different locations, and the momentum could be in many different directions. Statistical mechanics applies to large systems, so that the occupation probability of any quantum is independent of the number of states that exist. We can therefore consider large g_i and calculate what fraction of these states are occupied in order to calculate the average occupation number.

Consider having n_i identical fermions, so that each distinct quantum state can only have either 0 or 1 fermion in it. We now want to find the most likely distribution of the fermions amongst the available energy levels, which depends on the number of different ways that you can arrange the particles in the levels[3]. The number of ways of putting n_i identical particles into g_i distinct states is given by the binomial coefficient

$$w_f(n_i, g_i) = \frac{g_i!}{n_i!(g_i - n_i)!}. \tag{48}$$

For bosons each state can have more than one particle in, and calculating the number of ways the states can be more populated is more tricky. Consider writing down a list of the particles and grouping them into the different states, with a boundary separating particles in different states (and consecutive boundary lines if there are no particles in a given state). There are $g_i - 1$ boundaries between each group of particles in this list (because there are g_i sub-levels), and the total number of particles is n_i. The number of ways this could happen is the number of ways of choosing $g_i - 1$ boundaries and n_i particles from a set of $n_i + g_i - 1$ boundaries and particles; this gives[4]

$$w_b(n_i, g_i) = \frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!}. \tag{49}$$

---

[3] This corresponds to maximizing the entropy. In statistical mechanics the entropy is $S \equiv -k_B \sum_i P_i \ln P_i$, where P_i is the probability of being in the ith micro-configuration of the system (microstate). We assume each microstate is equally likely to be occupied, so that $P_i = 1/W$ where W is the number of microstates (W is the number of distinct ways of arranging things). In this case $S = k_B \ln W$.

[4] If you are confused, see `http://en.wikipedia.org/wiki/Bose-Einstein_distribution` first notes section.

Given a set of occupation numbers $\{n_i\}$ for each level, the total number of ways the levels and sub-levels can be populated is

$$W = \prod_i w(n_i, g_i). \tag{50}$$

We now want to find the most likely distribution (that with the largest W), subject to the constraint of fixed energy $E = \sum_i n_i \epsilon_i$ and number of particles $N = \sum_i n_i$ in a fixed large volume. We can do this with Lagrange multipliers[5], i.e. maximize

$$f \equiv \ln(W) + \alpha(N - \sum_i n_i) + \beta(E - \sum_i n_i \epsilon_i). \tag{51}$$

In general maximizing using $\frac{\partial f}{\partial n_i} = 0$ gives

$$\frac{\partial \ln W}{\partial n_i} = \alpha + \beta \epsilon_i. \tag{52}$$

For the cases in hand, in the large g limit we can use Stirling's approximation $\ln n! \approx n \ln n - n$. For Fermions this gives

$$\ln W \approx \sum_i \left[-n_i \ln n_i - (g_i - n_i) \ln(g_i - n_i) + g_i \ln g_i \right] \tag{53}$$

and hence the maximum is for $\hat{n}_i$ where

$$\frac{\partial \ln W}{\partial n_i} = -\ln \hat{n}_i + \ln(g_i - \hat{n}_i) = \alpha + \beta \epsilon_i \tag{54}$$

which rearranges to give

$$\hat{n}_i = \frac{g_i}{e^{\alpha + \beta \epsilon_i} + 1} \qquad \text{(fermions)}. \tag{55}$$

For bosons with $g_i \gg 1$ so $g - 1 \approx g$ we have

$$\ln W \approx \sum_i \left[(n_i + g_i) \ln(n_i + g_i) - (n_i + g_i) - n_i \ln n_i + n_i - g_i \ln g_i + g_i \right] \tag{56}$$

which gives the maximum at

$$\hat{n}_i = \frac{g_i}{e^{\alpha + \beta \epsilon_i} - 1} \qquad \text{(bosons)}. \tag{57}$$

For large g_i we expect the distribution to be symmetric, so the maximum is also the mean: the average occupation of each sub-level is given by the large-g_i fraction $\hat{n}_i / g_i$:

$$\mathcal{N}_i \equiv \frac{\langle n_i \rangle}{g_i} \approx \frac{1}{e^{\alpha + \beta \epsilon_i} \pm 1}, \tag{58}$$

with $+$ for fermions and $-$ for bosons. This is called the *occupation number* for a particular state[6]. For fermions is it just the probability that any given state is occupied; for bosons, each state can have more than one occupant, so the number can go larger than one (e.g. in Bose-Einstein condensation).

The distribution is determined by two constant α and β, which can be used as the thermodynamic variables labelling the distinct macrostates. As such they must be related to the usual thermodynamic state variables (temperature and chemical potential). Also, the number of ways of arranging things

---

[5] For an introduction see `http://en.wikipedia.org/wiki/Lagrange_multiplier`

[6] Note that sometimes the occupation number is alternatively defined including the g_i degeneracy factor.

is conventionally measured by the *entropy*, defined with a convenient constant so that $S = k_B \ln W$. The mostly likely state, the equilibrium state, is therefore the state of maximum entropy (subject to the constraints).

Note that since the constraints must be satisfied $S = k_B \ln(W) = k_B f$. From Eq. 51 we then see that

$$\frac{\partial S}{\partial E} = k_B \beta, \qquad \frac{\partial S}{\partial N} = k_B \alpha. \tag{59}$$

So α describes how the entropy responds to changes in the total number of particles, and β describes the response to changes in total energy. Eq. (59) can then be used give a statistical mechanical *definition* of the temperature and chemical potential in terms of β and α, and hence the response of the entropy.

We can also relate α and β to familiar classical thermodynamic quantities by comparing with the relation

$$\mathrm{d}E = T\mathrm{d}S - P\mathrm{d}V + \mu\mathrm{d}N \qquad \Longrightarrow \qquad \mathrm{d}S = \frac{1}{T}\left(\mathrm{d}E + P\mathrm{d}V - \mu\mathrm{d}N\right), \tag{60}$$

so that for a fixed volume using Eqs. (59) we see that

$$\left.\frac{\partial S}{\partial E}\right|_{V,N} = k_B\beta = \frac{1}{T}, \qquad \left.\frac{\partial S}{\partial N}\right|_{V,E} = k_B\alpha = -\frac{\mu}{T}. \tag{61}$$

Hence β is related to the temperature, and α to the chemical potential (and temperature):

$$\beta = \frac{1}{k_B T}, \qquad \alpha = \frac{-\mu}{k_B T}. \tag{62}$$

Another way to see the relation to the classical quantities is using Eq. (52): assuming the energy levels do not change we then have[7]

$$\mathrm{d}S = k_B \mathrm{d}\ln W = k_B \sum_i \frac{\partial \ln W}{\partial n_i}\mathrm{d}n_i = k_B \sum_i (\alpha + \beta\epsilon_i)\mathrm{d}n_i = k_B(\alpha dN + \beta dE). \tag{63}$$

Hence comparing coefficients with Eq. (60) at fixed volume gives Eq. (62).

Finally we can now go back and write the equilibrium occupation number of Eq. (58) directly in terms of the temperature and chemical potential as

$$\mathcal{N}_i = \frac{1}{e^{(\epsilon_i - \mu)/k_B T} \pm 1}. \tag{64}$$

For brevity we will often use units with $k_B = 1$; remember to put k_B back in where required to make the dimensions correct when calculating numerical answers.

For a particle in a large volume we can calculate the relevant density of states using the fact that in a box of side L the wavelengths in each direction are quantized so that $\lambda_i = 2L/n_i$, where n_i is integer. In terms of momentum $|\mathbf{p}| = h\nu = \hbar 2\pi c/\lambda = 2\pi/\lambda = 2\pi\sqrt{n_x^2 + n_y^2 + n_z^2}/2L$ in natural units. In n-space, the states are spaced in a cubic grid with unit grid point separation. However the n-space points are only in the range where $n_x, n_y, n_z \geq 0$, but $\mathbf{p}$ can point in any direction, so there are eight points in momentum space for every point in the positive octant of n-space. So the number of states per unit momentum volume is $\frac{1}{8} \times \left(\frac{2L}{2\pi}\right)^3$. Over a particular range of momenta and positions the number of states is therefore

$$dg_{\mathbf{p}} = \frac{1}{8}\frac{(2L)^3 d^3\mathbf{p}}{(2\pi)^3} = \frac{d^3\mathbf{p}d^3\mathbf{x}}{(2\pi)^3}. \tag{65}$$

For a particle A in addition there may be a spin-degeneracy factor g_A.

---

[7] If the energy levels can change $\mathrm{d}E = \sum_i (n_i \mathrm{d}\epsilon_i + \epsilon_i \mathrm{d}n_i)$ so that $\sum_i \beta\epsilon_i \mathrm{d}n_i = \beta\mathrm{d}E - \beta\sum_i n_i \mathrm{d}\epsilon_i$, where $\sum_i n_i \mathrm{d}\epsilon_i =$ dWork $= -PdV$ is the work done on the system, consistent with the usual thermodynamical result when the volume is not fixed.

B. Distribution function

For a particle species A (with mass m) in statistical equilibrium, the number density n, energy density ρ and pressure P are given as integrals over the distribution function $f_A(\mathbf{x}, \mathbf{p}, t)$. This is defined so that in a 3-momentum element $d^3\mathbf{p}$ and spatial volume element $d^3\mathbf{x}$ there are $f_A(\mathbf{x}, \mathbf{p}, t)d^3\mathbf{p}d^3\mathbf{x}$ particles, where $\mathbf{p}$ is the 3-momentum. We only consider the homogeneous case here, so $f_A(\mathbf{x}, \mathbf{p}, t) = f_A(\mathbf{p}, t)$, and statistical isotropy implies $f_A(\mathbf{p}, t) = f_A(|\mathbf{p}|, t) \equiv f_A(p, t)$. We will mainly leave the time dependence implicit, which will be manifest via the temperature dependence of the equilibrium distribution function. Particles of different species will be interacting constantly, exchanging energy and momentum. If the rate of these reactions $\Gamma(t) = n\langle\sigma v\rangle$ (where σ is the cross-section and v is the rms velocity) is much higher than the rate of expansion $H(t)$, then these interactions can produce and maintain thermodynamic equilibrium with some temperature T. Therefore, particles may be treated as an ideal (Bose or Fermi) gas, with the equilibrium distribution function $f_A d^3\mathbf{p}d^3\mathbf{x} = \mathcal{N}_\mathbf{p} g_A dg_\mathbf{p}$ so that

$$f_A(p) = \frac{g_A}{(2\pi)^3}\frac{1}{e^{(E_A-\mu_A)/T_A} \pm 1} \tag{66}$$

where g_A is the spin degeneracy factor, μ_A is the chemical potential, T_A is the temperature of this species and $E(p) = \sqrt{p^2 + m^2}$, where $p = |\mathbf{p}|$. The "+" sign corresponds to fermions, and the "-" sign to bosons, and we are using units with Boltzmann constant $k_B = 1$.

C. Chemical potential

The chemical potential may not be very familiar, and for a given system in general is unknown; however we know some things about it.

If the number of particles is not constrained (so that chemical as well as kinetic equilibrium is obtained), we do not need the $\sum_i n_i = N$ Lagrange multiplier, i.e. $\alpha = \mu = 0$ and the chemical potential is zero. For example in the very early universe photons are not conserved (double Compton scattering $e^- + \gamma \leftrightarrow e^- + \gamma + \gamma$ happens in equilibrium at high temperatures), so the number of photons can change to maximize the entropy. The maximum is where $\partial S/\partial N = 0$, and hence from Eq. (61) $\mu_\gamma = 0$.

For any interaction between particles that takes place frequently in the equilibrium (where $dS = 0$), we must also have

$$dS = \sum_i \frac{\partial S}{\partial N_i}\mathrm{d}N_i = -\sum_i \frac{\mu_i}{T}\mathrm{d}N_i = 0, \tag{67}$$

and hence $\sum_i \mu_i dN_i = 0$ (since the temperatures must also be the same in equilibrium). If this were not the case the particles could convert into each other to increase the entropy further.

If different species are in chemical equilibrium through the reactions $A + B \rightleftharpoons C + D$, then the chemical potentials satisfy $\sum_i \mu_i \mathrm{d}N_i = -\mu_A - \mu_B + \mu_C + \mu_D = 0 \implies \mu_A + \mu_B = \mu_C + \mu_D$[8]. This can be used to relate unknown chemical potentials to each other. For example as we mentioned photons are not conserved at high temperature, so we know[9] $\mu_\gamma = 0$, and if pair production and annihilation takes place, eg. $\mathrm{e}^- + \mathrm{e}^+ \leftrightarrow \gamma + \gamma$, then the particle and antiparticle have equal and opposite chemical potentials, e.g. $\mu_{\mathrm{e}^-} = -\mu_{\mathrm{e}^+}$.

[8] You can also derive this by considering the constraints (and hence Lagrange multipliers) $A + B \rightleftharpoons C + D$ imposes on the individual numbers and energy of the joint system; see e.g. the Mukhanov book, Sec 3.3.

[9] When the universe cools enough ($z \lesssim 2 \times 10^6$) double Compton scattering is inefficient, and it's possible for a "μ-distortion" to develop where μ_γ becomes non-zero if new energy is injected; see e.g. arXiv:1201.5375 and refs therein.

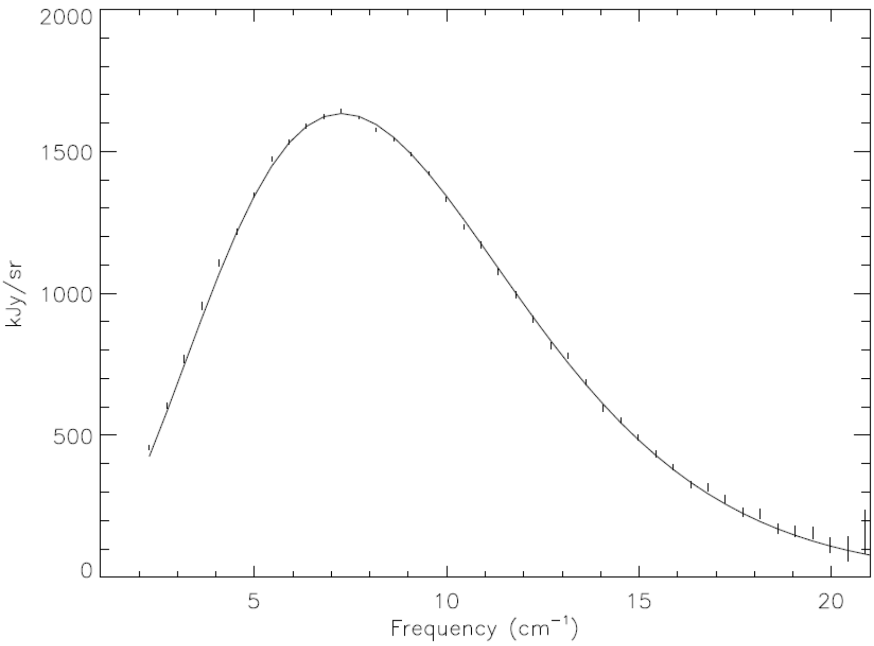

FIG. 5: Measurements of the CMB intensity as a function of frequency by COBE's Firas experiment. The data are in good agreement with a blackbody spectrum (solid line) with $T = 2.726K$. *From* arXiv:astro-ph/9605054

D. Revision: Blackbody spectrum

Of particular importance are the photons in the early universe and observed today in the CMB. The Firas experiment on the COBE satellite measured the energy spectrum of the CMB with high accuracy. If the photons originated from a thermal equilibrium distribution in the early universe what would we expect to see? Since photons are bosons with $\mu = 0$, and including the two photon spins (polarizations) we have the equilibrium distribution function

$$f_\gamma(E) = \frac{2}{(2\pi)^3}\frac{1}{e^{E/k_BT} - 1} \tag{68}$$

where in terms of the photon frequency ν we know $E = p = h\nu$. Each photon has an energy E so the total energy received over an area dA with photon direction (momentum) in solid angle $d\Omega$ is $Ef_\gamma d^3\mathbf{p}d^3\mathbf{x} = Ef_\gamma p^2 dp d\Omega_\mathbf{p} d^3\mathbf{x} = E^3 f_\gamma dE d\Omega_\mathbf{p} dA c dt$. Hence the intensity (power per unit area per unit solid angle per unit frequency) observed is (with constants put back in)

$$B(\nu) = \frac{2h\nu^3}{c^2}\frac{1}{e^{\frac{h\nu}{k_BT}} - 1}. \tag{69}$$

This is the blackbody spectrum, and observations by Firas fit this spectrum to high accuracy with a temperature of $T = 2.726K$. So we know that photons were very close to equilibrium at some point in the early universe.

For small frequencies (large wavelengths) a series expansion in ν gives the *Rayleigh-Jeans* approximation

$$B(\nu) \approx \frac{2\nu^2 k_B T}{c^2} \qquad (h\nu \ll k_B T). \tag{70}$$

In the limit of high frequencies (the *Wien tail*) there is instead the Wien approximation

$$B(\nu) \approx \frac{2h\nu^3}{c^2} e^{-\frac{h\nu}{k_BT}} \qquad (h\nu \gg k_B T). \tag{71}$$

Sometimes people use $B(\lambda)$ rather than $B(\nu)$, remember to account for the Jacobian if you are switching between the two.

E. Number densities, energy densities and pressure

Given the distribution function, we can then calculate the number density n, energy density ρ and pressure P:

$$n = \int f(\mathbf{p}) d^3\mathbf{p} = \frac{g}{2\pi^2} \int_0^\infty \frac{p^2 dp}{e^{(E-\mu)/T} \pm 1} = \frac{g}{2\pi^2} \int_m^\infty \frac{(E^2 - m^2)^{1/2}}{e^{(E-\mu)/T} \pm 1} E dE \,, \tag{72}$$

$$\rho = \int E(p) f(\mathbf{p}) d^3\mathbf{p} = \frac{g}{2\pi^2} \int_0^\infty \frac{E}{e^{(E-\mu)/T} \pm 1} p^2 dp = \frac{g}{2\pi^2} \int_m^\infty \frac{(E^2 - m^2)^{1/2}}{e^{(E-\mu)/T} \pm 1} E^2 dE \,, \tag{73}$$

$$P = \int \frac{p^2}{3E(p)} f(\mathbf{p}) d^3\mathbf{p} = \frac{g}{2\pi^2} \frac{1}{3} \int_m^\infty \frac{(E^2 - m^2)^{3/2}}{e^{(E-\mu)/T} \pm 1} dE \,. \tag{74}$$

Don't confuse P for pressure with p for momentum[10]. We now give the number density, pressure, temperature, etc, for some useful limits:

- **Relativistic species**: $m \ll T$, $\mu \ll T$:

$$n = T^3 \frac{g}{2\pi^2} \int_0^\infty \frac{x^2 dx}{e^x \pm 1} \propto T^3 \,, \tag{75}$$

where $x = E/T$. Using the Riemann zeta function

$$\zeta(n) = \frac{1}{\Gamma(n)} \int_0^\infty du \frac{u^{n-1}}{e^u - 1}$$

for bosons we have

$$n_B = T^3 \frac{g\zeta(3)}{\pi^2} \,, \tag{76}$$

where $\zeta(3) \approx 1.202056903$, and for fermions, we can use a cunning trick of writing:

$$\frac{1}{e^x + 1} = \frac{1}{e^x - 1} - \frac{2}{e^{2x} - 1} \,, \tag{77}$$

and then:

$$n_F = T^3 \frac{g\zeta(3)}{\pi^2} (1 - \frac{1}{4}) = \frac{3}{4} n_B \,. \tag{78}$$

---

[10] Why $p^2/3E$ factor? Pressure is the force per unit area, so momentum change per unit time per unit area; momentum in the x direction is p_x per particle, hence a change in momentum of $\Delta p_x = 2|p_x|$ if it hits the perpendicular area dA. The volume swept out in time dt is $d^3\mathbf{x} = |v_x| dt dA = |p_x| dt dA/E$, so the number that hit is $dN = f d^3\mathbf{p} d^3\mathbf{x} = (|p_x| dt dA/E) f d^3\mathbf{p}$. Hence the contribution to the pressure (force per unit area), is $dP = \Delta p_x/(dt dA) = (2p_x^2/E) f d^3\mathbf{p}$ for particles moving in the right direction ($p_x > 0$). For given p_x^2 half are moving in the wrong direction, so the total pressure is $P = \int (p_x^2/E) f d^3\mathbf{p}$. For an isotropic distribution $\int d^3\mathbf{p} p_x^2 = \int d^3\mathbf{p} p_y^2 = \int d^3\mathbf{p} p_z^2$, hence $P = \int \frac{p^2}{3E} f d^3\mathbf{p}$. If the distribution is *not* isotropic, the pressure is defined to be the angle-averaged quantity given by the first part of Eq. (74).
Once can also write a general four-vector definition for the stress-energy tensor $T^{\mu\nu} = \int d^4p\, f(p) p^\mu p^\nu\, \delta(p^\alpha p_\alpha - m^2)$, which gives the same results from $T^{\mu\nu} = \int \frac{d^3\mathbf{p}}{E} f(\mathbf{p}) p^\mu p^\nu$ in the isotropic limit (changing variables to do the energy integral and accounting for the two poles in the delta function).

For the energy density, using $\zeta(4) = \pi^4/90$ and $\Gamma(n) = (n-1)!$ we get:

$$\rho_B = T^4 \frac{g}{2\pi^2} \int_0^\infty \frac{x^3 dx}{e^x - 1} = \frac{g}{30}\pi^2 T^4 \,, \tag{79}$$

$$\rho_F = T^4 \frac{g}{30}\pi^2 (1 - \frac{1}{8}) = \frac{7}{8}\rho_B \,, \tag{80}$$

and pressure

$$P = \frac{\rho}{3} \,. \tag{81}$$

There are several points here worthy of note: Firstly, we have derived now from statistical mechanics the equation of state for radiation, and indeed we find that $w = P/\rho = 1/3$. Secondly, in eq. (79) we derived the Stefan-Boltzmann relation for photons ($g = 2$), $\rho_\gamma = a_R T^4 = 4\sigma_{SB} T^4/c$, where the Stefan-Boltzmann constant is given by

$$\sigma_{SB} \equiv \frac{k_B^4 \pi^2}{60\hbar^3 c^2} = 5.6704 \times 10^{-8} \mathrm{Js^{-1}m^{-2}K^{-4}}. \tag{82}$$

We also showed in the last chapter that the energy density in radiation scales like $\rho_\gamma \propto a^{-4}$. Combined with the above, it follows immediately that the temperature of radiation (and of any relativistic species) scales like

$$T_\gamma \propto 1/a. \tag{83}$$

The important consequence is that the universe was much *hotter* when it was smaller. Here we use the temperature of the radiation as the "temperature of the universe", both because it is well defined as the radiation has a thermal spectrum (and thus a unique, well-defined temperature) and because at early times the other particle species interact with the radiation and so share its temperature. We will later discuss what happens when this is no longer true.

- **Relativistic species, small chemical potential**: $m \ll T$, $|\mu| \ll T$ (but μ non-negligible) Consider the number density of Fermions, and doing a series expansion in μ:

$$\begin{aligned} n &= \frac{g}{2\pi^2} \int_0^\infty \frac{p^2 dp}{e^{(p-\mu)/T} + 1} \\ &\approx n(\mu = 0) + \frac{\mu}{T} \frac{g}{2\pi^2} \int_0^\infty \frac{e^{p/T} p^2 dp}{[e^{p/T} + 1]^2} + \ldots \\ &= n(\mu = 0) + \frac{g\mu T^2}{12} + \mathcal{O}(\mu^2/T^2). \end{aligned} \tag{84}$$

In particular for a particle X and anti-particle $\bar{X}$ with $\mu_X = -\mu_{\bar{X}}$ (as expected if e.g. $X + \bar{X} \leftrightarrow \gamma + \gamma$), then

$$n_X - n_{\bar{X}} = \frac{g\mu_X T^2}{6} + \mathcal{O}(\mu_X^3/T^3). \tag{85}$$

In particular, if $n_X \approx n_{\bar{X}}$ (e.g. because the universe is charge neutral), then $\mu_X/T \approx 0$. The chemical potential is closely related to conserved charges and particle-antiparticle asymmetries.

- **Non-relativistic (massive) species**, $m \gg T$:

$$\begin{aligned} n &\approx \frac{g}{2\pi^2} e^{\mu/T} \int_0^\infty e^{-E/T} p^2 dp \approx \frac{g}{2\pi^2} e^{\mu/T} \int_0^\infty e^{-(m+p^2/2m)/T} p^2 dp \\ &= g \left(\frac{mT}{2\pi}\right)^{3/2} e^{-(m-\mu)/T} \,, \end{aligned} \tag{86}$$

$$\rho = mn \,, \tag{87}$$

$$P \approx \frac{g}{2\pi^2} e^{\mu/T} \int_0^\infty e^{-(m+p^2/2m)/T} \frac{p^2}{3m} p^2 dp \simeq nT \ll \rho \quad (P \simeq 0) \,. \tag{88}$$

This recovers the ideal gas law, $P = nT$. Again we can look at the relative numbers of particles and anti-particles in the case that they are massive, and now get

$$n_X - n_{\bar{X}} = 2g \left(\frac{mT}{2\pi}\right)^{3/2} e^{-m/T} \sinh(\mu_X/T). \tag{89}$$

Again this is zero if $\mu_X = 0$, and for small chemical potentials the numbers are all very small because of the exponential $e^{-m/T}$ suppression.

Therefore, in general for relativistic species the number densities go as T^3 and the energy density behaves as T^4, while for massive species they are suppressed by the Boltzmann factor $\exp(-m/T)$. This exponential suppression of the number density with the mass means that non-relativistic particles (with $m \gg T$) either become exponentially rare, or drop below the density where they interact sufficiently often to stay in equilibrium and hence go out of equilibrium. Both cases will be of interest. For example, electrons and positrons interact very strongly via the Colomb force and so maintain equilibrium until there eare almost no positrons left at $T \ll m_{e+} \sim 0.5$MeV (they all annihilate with electrons). Other particles that interact weakly, like weakly-interactive dark matter, would rapidly go out of equilibrium.

1. Several relativistic species: number of degrees of freedom

If we have a collection of relativistic species, each of them in equilibrium at different temperatures T_i, we can write the total energy density ρ_R, summing over all the contributions and neglecting the chemical potentials ($\mu_i \ll T$), as:

$$\rho_R = \sum_i \rho_i = \frac{T_\gamma^4}{30} \pi^2 g_* \tag{90}$$

where where g_* is the "effective" number of degrees of freedom, given by:

$$g_* = \sum_{\text{bosons}} g_i \left(\frac{T_i}{T_\gamma}\right)^4 + \frac{7}{8} \sum_{\text{fermions}} g_i \left(\frac{T_i}{T_\gamma}\right)^4 . \tag{91}$$

As the temperature decreases, the effective number of degrees of freedom in radiation will decrease, as massive particles start behaving non-relativistically when their mass become larger then T:

$T \ll 1$ MeV: the only relativistic particles would be the 3 neutrino species (fermions with 2 degrees of freedom each) and the photon (boson, 2 polarisation states). Neutrinos at this temperature are decoupled from the thermal bath and they are slightly colder than the photons, with a temperature $T_\nu = (4/11)^{1/3} T_\gamma$ [we'll derive this later on]. Therefore:

$$g_* = 2 + \frac{7}{8} \times 3 \times 2 \times \left(\frac{4}{11}\right)^{4/3} \simeq 3.36 \,. \tag{92}$$

1 MeV $< T <$ 100 MeV: Electrons and positrons have a mass of about 0.5 MeV, and so they are now also relativistic. As the difference between neutrino and photon temperature is due to the electron-positron annihilation (see discussion in the section IV I), we have $T_\nu = T_\gamma$ and so

$$g_* = 2 + \frac{7}{8} (3 \times 2 + 2 \times 2) = 10.75 \tag{93}$$

$T <$ 300 GeV: this is above the electroweak unification scale, and for the particles in the Standard Model we have $g_* \simeq 106.75$

⋮

F. Entropy

The universe has far more photons than baryons, so the entropy of a uniform universe is dominated by that of the relativistic particles. The fundamental relation of thermodynamics for a system in equilibrium with negligible chemical potential (or no change in particle number) is

$$dE = TdS - PdV. \tag{94}$$

The change in energy is the work done changing the volume plus the temperature times the change in entropy (the temperature is effectively the energy per internal state, each with energy $\sim k_B T$). In a cosmological volume V we have $E = \rho V$ so

$$V d\rho + \rho dV = TdS - PdV. \tag{95}$$

We know about $d\rho/dt$ from the energy conservation equation (Eq. 34); using $V \propto a^3$ this gives

$$\frac{d\rho}{dt} = -3H(\rho + P) = -\frac{1}{V}\frac{dV}{dt}(\rho + P), \tag{96}$$

so substituting in Eq. (95) we have

$$-\frac{dV}{dt}(\rho + P) + \rho\frac{dV}{dt} = T\frac{dS}{dt} - P\frac{dV}{dt} \tag{97}$$

$$\implies \frac{dS}{dt} = 0. \tag{98}$$

So the total entropy is in a comoving volume is conserved, which is what we might expect for a closed system (there is nowhere for heat to flow from or to). It is also useful to consider the entropy density $s = S/V$, where substituting for $dS = d(sV)$ in Eq. (95)

$$T(sdV + V ds) = V d\rho + \rho dV + PdV. \tag{99}$$

$$\implies d\rho - T ds = (Ts - \rho - P)\frac{dV}{V}. \tag{100}$$

For a system at equilibrium the entropy density, energy density and pressure are intensive quantities that can be written as functions only of the temperature, $\rho = \rho(T)$, $s = s(T)$, $P = P(T)$, so that $d\rho - Tds \propto dT$. The coefficients of the dT and dV terms must separately be zero because one is intensive (independent of volume) and the other is extensive (depends on the size of the system). For example you could consider a volume change at constant temperature $dT = 0$, which implies then the dV term must be zero. Hence for the dV coefficient to be zero we get an expression for the entropy density

$$s = \frac{1}{T}(\rho + P). \tag{101}$$

The dT coefficient is then zero by the $\dot{\rho}$ energy conservation equation.

We therefore expect $S \propto a^3 s$ to be conserved between different times when a homogeneous universe is in thermodynamic equilibrium. This also applies to separate decoupled (uninteracting) components if they are each separately in a thermal distribution with their own temperature.

For a relativistic species A in thermal equilibrium at a temperature T_A, we have seen that:

$$\rho_A = g_A^{\text{eff}}\frac{\pi^2}{30}T_A^4 = 3P_A\,, \tag{102}$$

where $g_A^{\text{eff}} = g_A$ for bosons, and $g_A^{\text{eff}} = 7g_A/8$ for fermions. The total entropy density s, summing over all possible contributions is given by:

$$\begin{aligned} s &= \sum_i s_i = \frac{\pi^2}{30}(1+\frac{1}{3})T_\gamma^3\left[\sum_{\text{bosons}} g_i\left(\frac{T_i}{T_\gamma}\right)^3 + \frac{7}{8}\sum_{\text{fermions}} g_i\left(\frac{T_i}{T_\gamma}\right)^3\right] \\ &= \frac{2\pi^2}{45}g_{*S}T_\gamma^3\,, \end{aligned} \tag{103}$$

where T_γ is the temperature of the photons, and we have defined the effective number of degrees of freedom in entropy:

$$g_{*S} = \sum_{\text{bosons}} g_i \left(\frac{T_i}{T_\gamma}\right)^3 + \frac{7}{8} \sum_{\text{fermions}} g_i \left(\frac{T_i}{T_\gamma}\right)^3 . \quad (104)$$

Notice that g_{*S} is equal to g_* only when *all* the relativistic species are in equilibrium at the *same* temperature.

We can also consider independent thermal components that are at different temperatures separately. Using the entropy conservation law for a thermal system i at temperature T_i we have then:

$$d(g_{*S,i} a^3 T_i^3) = 0 \;\;\Rightarrow\;\; T_i \propto g_{*S,i}^{-1/3} a^{-1} . \quad (105)$$

If $g_{*S,i}$ is a constant, then the T_i of the system decreases as the inverse of the scale factor, $T_i \propto a^{-1}$, and $s_i \propto T_i^3 \propto a^{-3}$. If there is one or more independent thermal systems, as we shall see is the case after decoupling, then this law will apply separately to each one, where the temperatures of each system can in general be different.

G. Decoupling

At any time, the Universe will contain a blackbody distribution of photons with some temperature T_γ. If a species interacts (directly or indirectly) with the photons with a rate $\Gamma_{A\gamma}$ high enough ($\Gamma_{A\gamma} \gg H$), then these particles will have the same temperature as the photons: $T_A = T_\gamma$. Any set of particles species which are interacting among themselves at a high enough rate will share the same temperature. But in general this may not be the case and the Universe could be populated by different species each with its own temperature (uninteracting species behave like independent thermal systems). As a rule of thumb, a species A will maintain an equilibrium distribution when there are many photon interactions in the time it takes the universe to expand significantly, i.e. $\Gamma_A \gg H$, and it will decouple from the thermal bath when the interaction rate drops below the rate of expansion, $\Gamma_A \ll H$.

The rate of interaction can be expressed as:

$$\Gamma_A \equiv n_T \langle \sigma_X v \rangle , \quad (106)$$

where n_T is the number density of the target particles, v is the relative velocity, and σ_X is the interaction cross section. $\langle \sigma_X v \rangle$ denotes an average value of this combination (σ_X usually depends on the energy). We should note that the interaction rates between particles A and B are not symmetric, e.g. if $n_A \gg n_B$ then also $\Gamma_B \gg \Gamma_A$. It is then possible that particle B is still in thermal equilibrium with A, while A has already decoupled from B.

As long as $\Gamma_A \gg H$ and the interactions can maintain equilibrium, the distribution function f_A maintain the form of the equilibrium distribution. But once the species A is completely decoupled ($\Gamma_A \ll H$) the particles will be just travelling freely. The distribution function is then frozen in to the form it had at decoupling, though particle momenta will gradually redshift as the universe expands.

Notice that if a massive particle decouples when it is relativistic $T_D \gg m$, then the distribution function is "frozen" in the form of the distribution function $f_{(\mathrm{eq})}$ of massless particles. These particles will become non-relativistic when the temperature of the thermal bath drops below their mass, such that their energy will now be $E \simeq m$. The distribution function and number density of the particles will still be given by the frozen-in form corresponding to relativistic particles, but the energy density will be that of non-relativistic particles $\rho \simeq nm$. This is exactly what happens for massive neutrinos.

H. Neutrino decoupling

At temperature below $T \simeq 10^{12}\mathrm{K} \simeq O(100)$ MeV, the energy density of the universe is essentially given by that of the relativistic particles $e^\pm$, ν, $\bar{\nu}$ and photons. Since they are in equilibrium with the same temperature, the effective number of degrees of freedom is $g_* = 10.75$, and the rate of expansion in this radiation-dominated epoch is given by:

$$H(T) = \frac{\rho_R^{1/2}}{\sqrt{3}M_P} \sim \frac{T^2}{M_P}\,. \tag{107}$$

Here we are going to give a very rough argument, not worrying about $\mathcal{O}(1)$ constants. Neutrinos are kept in equilibrium via weak interaction processes (for example $\bar{\nu}\nu \leftrightarrow e^+e^-$ via Z, elastic scattering of ν and e^- via Z exchange, or $e^-\bar{\nu} \leftrightarrow e^-\bar{\nu}$ via W^-, etc.), with a cross section given by:

$$\sigma_F \simeq G_F^2 E^2 \sim G_F^2 T^2\,, \tag{108}$$

where G_F is the Fermi constant ($G_F = \pi\alpha_W/(\sqrt{2}m_W^2) = 1.1664 \times 10^{-5}\,\mathrm{GeV}^{-2}$). The interaction rate per (massless, so $v = 1$) neutrino is:

$$\Gamma_F = n\langle\sigma_F v\rangle \sim G_F^2 T^5\,, \tag{109}$$

and hence

$$\frac{\Gamma_F}{H(T)} \sim T^3 G_F^2 M_P \sim \left(\frac{T}{1\ \mathrm{MeV}}\right)^3\,. \tag{110}$$

Therefore neutrinos decouple from the rest of the matter at a temperature around $T_D \sim 1$ MeV. Below 1 MeV, the neutrino temperature scales as a^{-1}.

I. Electron-positron annihilation

Shortly after neutrino decoupling, the temperature drops below the mass of the electron ($T <$ 0.5 MeV), and the electron-positron pairs all annihilate into photons[11]. Due to the strong electromagnetic forces the photon-electron-positron gas remains in thermal equilibrium, but the neutrinos are decoupled. We can therefore treat the photon-electron-positron gas as a separate thermal system from the neutrinos. Let's call them thermal system one and thermal system two, where being decoupled means there are no interactions so they behave independently. For the photon-electron-positron gas before and after e^-e^+ annihilation we have respectively:

$$g_{*S,1}(T_D > T \gg m_e) = 2 + \frac{7}{8}4 = \frac{11}{2}\,, \qquad g_{*S,1}(T \ll m_e) = 2\,. \tag{111}$$

These results are of course only approximate near $T \sim m_e$, but due to the $e^{-m/T}$ factor in the equilibrium abundance for massive particles, the number of electrons and positrons is rapidly exponentially suppressed for $T \ll m_e$.

The conservation of the entropy $S_1 = g_{*S,1}(aT_\gamma)^3$ for the particles which are in equilibrium with radiation shows that $g_{*S,1}(T_\gamma a)^3$ remains constant during expansion. Because $g_{*S,1}$ decreases after $T < m_e$, the value of $(aT_\gamma)^3$ will be larger after e^-e^+ annihilation than its value before:

$$\frac{(aT_\gamma)^3_{\mathrm{after}}}{(aT_\gamma)^3_{\mathrm{before}}} = \frac{(g_{*S,1})_{\mathrm{before}}}{(g_{*S,1})_{\mathrm{after}}} = \frac{11}{4}\,. \tag{112}$$

Neutrinos do not participate in this process and their entropy is separately conserved, with $(aT_\nu)_{\mathrm{before}} = (aT_\nu)_{\mathrm{after}}$ since the neutrino degrees of freedom do not change. But before e^-e^+ annihilation began, photons and neutrinos had the same temperature, $(aT_\gamma)_{\mathrm{before}} = (aT_\nu)_{\mathrm{before}}$ since they were in equilibrium prior to neutrino decoupling. Therefore:

$$(aT_\gamma)_{\mathrm{after}} = \left(\frac{11}{4}\right)^{1/3}(aT_\nu)_{\mathrm{after}}\,. \tag{113}$$

---

[11] Not *all* the electrons annihilate since there is an excess of electrons to have charge neutrality with the protons, $n_e = n_p$. However we will show shortly that this excess is tiny compared to the number of photons, and hence the number of electrons and positrons produced in equilibrium above 0.5 MeV.

The temperature of the photons is larger than that of the neutrinos today by a factor $(11/4)^{1/3} \sim 1.4$ (so $T_{\nu 0} \sim 1.92K$ for $T_{\gamma 0} = 2.726K$). And because $T_\gamma \neq T_\nu$, today $g_* \neq g_{*S}$, with $g_* \simeq 3.36$ and $g_{*S} \simeq 3.91$.

J. Matter-radiation equality (revision)

The total matter density and radiation today is (for CMB temperature today $T_{\gamma 0} = 2.726K$, which is observed to be very close to blackbody)

$$\rho_{m0} = \frac{3H_0^2\Omega_m}{8\pi G} = 1.88 \times 10^{-32}\Omega_m h^2 \text{ kg cm}^{-3} \tag{114}$$

$$\rho_{r0} = g_* \frac{\pi^2}{30}\frac{(k_B T_{\gamma 0})^4}{c^5\hbar^3} = 7.8 \times 10^{-37} \text{ kg cm}^{-3} \tag{115}$$

Here we used the definition that the Hubble parameter today is $H_0 = h 100\text{km}\,\text{s}^{-1}\text{Mpc}^{-1}$. Using the fact that $\rho_r/\rho_m \propto a_0/a = 1+z$, it follows that the redshift of equal matter and radiation energy densities (*matter-radiation equality*) is given by:

$$1 + z_{(\text{eq})} \;=\; 2.4 \times 10^4 \Omega_m h^2. \tag{116}$$

Evidence suggests $\Omega_m h^2 \sim 0.133$, so $z_{(\text{eq})} \sim 3200$ ($T_{eq} \sim 0.75$eV). Prior to this redshift the universe was *radiation dominated*, afterwards it was *matter dominated* until dark energy became important at low redshift ($z \sim 1$).

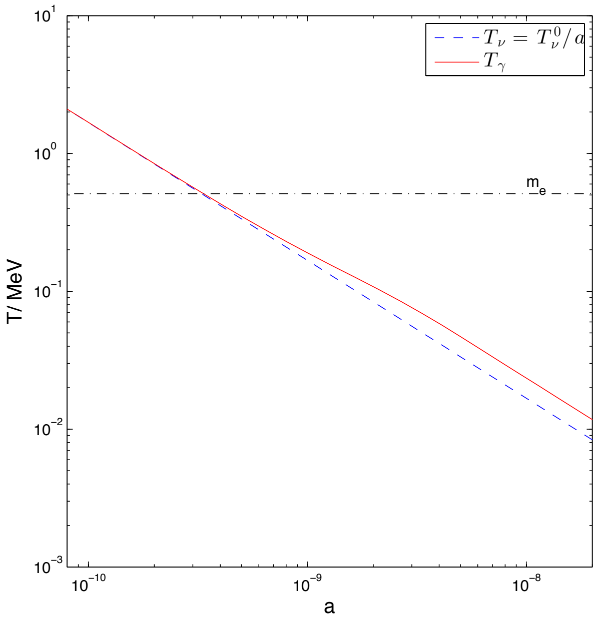

FIG. 6: Thermal history through electron-positron annihilation. Neutrinos are decoupled and their temperature redshifts $\propto 1/a$. When $T \lesssim m_e$ the electrons and positrons (which are in equilibrium with the photons) annihilate, injecting energy and reducing the cooling due to redshifting. Note that the main effect is after $T \sim m_e$, because the contributions to the energy density are skewed towards $E > T$ because of the large phase space and higher weighting for larger momenta (integrand is $\propto f(E)p^2 E dp$).
Ex: See if you can write a Matlab or python script to reproduce this plot by numerically integrating the Fermi-Dirac distribution.

K. Photon decoupling and recombination

This was covered in the cosmology course.

L. Decoupling and freeze out

As discussed in Sec. (IV G), as long as $\Gamma_A \gg H$ and the interactions can maintain equilibrium, the distribution function f_A maintain the form of the equilibrium distribution. But once the species A is completely decoupled ($\Gamma_A \ll H$) the particles will be travelling along geodesics. Consider for example dN (freely propagating) particles in a proper volume $d^3\mathbf{x} \propto a^3$ with momentum in the interval $d^3\mathbf{p} \propto a^{-3}$ (in Sec. I C we showed that the momentum of particles moving on geodesics redshifts $\propto 1/a$) . Between a time t_D and a later time, assuming the particles are completely decoupled and don't decay, the number of particles is conserved and the momentum redshifts so

$$dN \propto f(\mathbf{x}, \mathbf{p}, t) d^3\mathbf{x} d^3\mathbf{p} = f(\mathbf{x}_D, \mathbf{p}_D, t_D) d^3\mathbf{x}_D d^3\mathbf{p}_D = f(\mathbf{x}_D, \mathbf{p}a/a_D, t_D) d^3\mathbf{x} d^3\mathbf{p} \tag{117}$$

since the factors in the phase space volume element cancel. Approximating decoupling as instantaneous at t_D we can then calculate the distribution function at any later time t using the equilibrium distribution at decoupling

$$f_{t>t_D}(\mathbf{x}, \mathbf{p}, t) = f_{(\mathrm{eq})}(\mathbf{x}_D, a\mathbf{p}/a_D, t_D). \tag{118}$$

If decoupling occurs when the species is ultrarelativistic ($T_D \gg m$), then:

$$f_{t>t_D}(\mathbf{p}, t) = \frac{g}{(2\pi)^3} \left[\exp\left(\frac{pa}{T_D a_D} \right) \pm 1 \right]^{-1}. \tag{119}$$

This has the same form as the distribution function $f_{(\mathrm{eq})}$ for a relativistic species with temperature

$$T(t) = T_D \frac{a_D}{a(t)}. \tag{120}$$

The "temperature" in the distribution function falls strictly as a^{-1}; the entropy of these particles, $S_A = s_A a^3$ is conserved separately. However for the species which are still in thermal equilibrium, $T \propto g_{*S}^{-1/3} a^{-1}$ falls more slowly because the number of degrees of freedom in equilibrium has now decreased.

The number density of these decoupled particles is given by:

$$n_A = g_A^{\mathrm{eff}} \left(\frac{\zeta(3) T_D^3}{\pi^2} \right) \left(\frac{a_D}{a} \right)^3. \tag{121}$$

This number density will be comparable to the number density of photons at any given time. That is, any such decoupled species will continue to exist in our universe today as a relic background with number density comparable to the number density of photons.

Notice that if a massive particle decouples when it is relativistic $T_D \gg m$, then the distribution function is "frozen" in the form of the distribution function $f_{(\mathrm{eq})}$ of massless particles. These particles will become non-relativistic when the temperature of the thermal bath drops below their mass, such that their energy will now be $E \simeq m$. The distribution function and number density of the particles will still be given by the frozen-in form corresponding to relativistic particles, but the energy density will be that of non-relativistic particles $\rho \simeq nm$. This is exactly what happens for massive neutrinos.

If the particles decouple when they are already non-relativistic, then

$$f_{t>t_D}(\mathbf{p}) = f_{(\mathrm{eq})}(\mathbf{p}\frac{a}{a_D}, T_D) \simeq \frac{g}{(2\pi)^3} \exp\left[-\frac{(m-\mu)}{T_D} \right] \exp\left[-\frac{p^2}{2mT_D} \left(\frac{a}{a_D} \right)^2 \right]. \tag{122}$$

The distribution function has the same form as that of a non-relativistic Maxwell-Boltzmann gas with a temperature $T = T_D(a_D/a)^2$, and chemical potential $\mu(t) = m + (\mu_D - m)(T/T_D)$.

V. NON-EQUILIBRIUM PROCESSES: THE BOLTZMANN EQUATION

We have seen that, qualitatively, when the interaction rate of a particle species A, Γ_A, is much larger than H, then they are kept in thermal equilibrium; when Γ_A drops below H, they are decoupled. However, we have so far always used the *equilibrium* form of the one-particle distribution function $f(x, p, t)$. A more accurate description of the decoupling process needs to follow the *evolution* of the distribution function. This is given by the Boltzmann equation, which can be written in a very compact form as:

$$\hat{L}[f_A] = \hat{C}_A[f_A]\,, \tag{123}$$

where $\hat{L}$ is the Liouville operator, and $\hat{C}_A$ the collision operator.

A. The Non-relativistic Boltzmann equation

To help clarify what is going on, we start with a brief discussion of the non-relativistic Boltzmann equation. The Liouville operator is just the total time derivative. Recall that the number of particles a volume $d^3\mathbf{x}$ with momentum in the range $d^3\mathbf{p}$ is $dN = f(\mathbf{x},\mathbf{p},t)d^3\mathbf{x}d^3\mathbf{p}$. If particles in dN have mass m and are subject to a force F, at time dt later the particles have moved because of their momentum, and have changed momentum because of the force, hence [12].

$$f(\mathbf{x}+\frac{\mathbf{p}}{m}dt, \mathbf{p}+\mathbf{F}dt, t+dt)d^3\mathbf{x}d^3\mathbf{p} = f(\mathbf{x},\mathbf{p},t)d^3\mathbf{x}d^3\mathbf{p} + C(\mathbf{x},\mathbf{p},t)d^3\mathbf{x}d^3\mathbf{p}dt, \tag{124}$$

where the 'collision' term C gives the rate of particles being created and destroyed (annihilations, decays), and scattering events that transfer momentum from $\mathbf{p}'$ into $\mathbf{p}$, or from $\mathbf{p}$ into $\mathbf{p}''$, which all lead to a change in $f(\mathbf{x},\mathbf{p},t)$. Dividing through we therefore have

$$\begin{aligned}\frac{f(\mathbf{x}+\frac{\mathbf{p}}{m}dt, \mathbf{p}+\mathbf{F}dt, t+dt) - f(\mathbf{x},\mathbf{p},t)}{dt} &= \frac{\partial}{\partial t}f(\mathbf{x},\mathbf{p},t) + \frac{p^i}{m}\frac{\partial}{\partial x^i}f(\mathbf{x},\mathbf{p},t) + F^i\frac{\partial}{\partial p^i}f(\mathbf{x},\mathbf{p},t)\\ &= C(\mathbf{x},\mathbf{p},t).\end{aligned} \tag{125}$$

The Boltzmann equation can therefore be written

$$L_{NR}f = C, \tag{126}$$

where since $\mathbf{F} = d\mathbf{p}/dt$ and $\mathbf{p}/m = d\mathbf{x}/dt$

$$L_{NR} = \frac{\partial}{\partial t} + \frac{dx^i}{dt}\frac{\partial}{\partial x^i} + \frac{dp^i}{dt}\frac{\partial}{\partial p^i} = \frac{d}{dt}\,. \tag{127}$$

This is of course just what you'd expect when writing out the total derivative in terms of partial derivatives: the Boltzmann equation just gives the total change in the distribution function with time accounting for particles moving under a force and collisions.

---

[12] Note that we have use the fact that phase space volume elements are conserved so that $d^3\mathbf{x}d^3\mathbf{p} = d^3\mathbf{x}'d^3\mathbf{p}'$: this is one form of Liouville's theorem. Hamilton's equations are

$$\dot{x}^i = \frac{\partial \mathrm{H}}{\partial p^i} \qquad \dot{p}^i = -\frac{\partial \mathrm{H}}{\partial x^i}$$

where H is the Hamiltonian. If we define a 6-dimensional phase-space coordinate $\eta = (\mathbf{x},\mathbf{p})$, it therefore follows that $\boldsymbol{\nabla}\cdot\dot{\eta} = 0$: the flow in phase space is divergenceless and hence volumes are conserved. Mathematically you can get this by showing that the Jacobian from η to η' is constant, see e.g. `http://www.nyu.edu/classes/tuckerman/stat.mech/lectures/lecture_2/node2.html`. Clearly if dN is constant (no collision), and $d^3\mathbf{x}d^3\mathbf{p}$ is constant, then f must be constant, which is just the collisionless Boltzmann equation.

B. Relativistic Boltzmann equation

The relativistic generalisation is the total derivative with respect to an affine parameter λ along some world line (e.g. could be taken to be proportional to the proper time):

$$\hat{L} = \frac{d}{d\lambda} = \frac{dx^\mu}{d\lambda}\frac{\partial}{\partial x^\mu} + \frac{dp^\mu}{d\lambda}\frac{\partial}{\partial p^\mu}. \tag{128}$$

The second term encodes the change in the distribution function just from changes in momentum as particles move along geodesics (or in principle also the effect of external forces). As in Sec. I C we can choose to normalize the affine parameter so that

$$p^\mu = \frac{dx^\mu}{d\lambda}, \tag{129}$$

so the geodesic equation becomes

$$\frac{dp^\mu}{d\lambda} + \Gamma^\mu{}_{\alpha\beta}p^\alpha p^\beta = 0. \tag{130}$$

Assuming there are no external (non-gravitational) forces, the particles will follow geodesics, so the relativistic generalisation of $\hat{L}$ in Eq. (128) can then be written

$$\hat{L} = p^\mu \frac{\partial}{\partial x^\mu} - \Gamma^\mu_{\alpha\beta}p^\alpha p^\beta \frac{\partial}{\partial p^\mu}\,. \tag{131}$$

Here $p_\mu p^\mu = E^2 - \mathbf{p}^2 = E^2 - p^2 = m^2$.

For the *isotropic* and *homogenenous* FRW model, the distribution function f_A can be taken to only depend on p^0 (the energy), such that $f_A = f_A(E,t)$ (or equivalently on $|\mathbf{p}|$ via $m^2 = E^2 - p^2$). So only the $\mu = 0$ components do not vanish, and we find using the FRW metric

$$\hat{L}[f_A] = \left[E\frac{\partial}{\partial t} - \frac{\dot{a}}{a}p^2\frac{\partial}{\partial E}\right] f_A(E,t)\,. \tag{132}$$

We can also write the Boltzmann equation in terms of the number density

$$n_A = 4\pi \int dp p^2 f_A(E,t)\,, \tag{133}$$

by dividing Eq. (132) by the energy and then integrating over the momentum,

$$4\pi \int dp p^2 \frac{\hat{L}[f_A]}{E} = \frac{dn_A}{dt} - H4\pi \int dp \frac{p^4}{E}\frac{\partial f_A}{\partial E} = \frac{dn_A}{dt} + H4\pi\int dp \frac{\partial(p^3)}{\partial p} f_A \tag{134}$$

$$= \dot{n}_A + 3Hn_A\,, \tag{135}$$

where we used $m^2 = E^2 - p^2$ so $pdp = EdE$ and integrated by parts. Notice that in the absence of interactions, the Boltzmann equation would reduce to:

$$\dot{n}_A + 3Hn_A = 0\,, \tag{136}$$

just the conservation of particles per comoving volume, $d(a^3 n_A)/dt = 0$, when there are no interactions or decays. For very massive particles (so $\rho_A = m_A n_A$, $P_A = 0$) it is just the energy conservation equation.

C. The Collision operator

The collision operator roughly gives the change in the number of particles A due to interactions or spontaneous decays. That is, given a space volume $d^3\mathbf{p}d^3\mathbf{x}$, it would give the number of particles "*in*" minus the number "*out*". For a process of the kind $A + X \leftrightarrow Y$, where X and Y denote collectively any set of other particles, in the case of a homogeneous and isotropic universe we have

$$\begin{aligned} \dot{n}_A + 3Hn_A &= -R_{(AX\to Y)} n_A n_X + R_{(Y\to AX)} n_Y \quad (137) \\ &= -\Gamma_{(AX\to Y)} n_A + \Gamma_{(Y\to AX)} n_Y \quad (138) \end{aligned}$$

for some forward and backward rate coefficients, $R = \langle \sigma v \rangle$, $\Gamma_{(AX\to Y)} = R_{(AX\to Y)} n_X$, and $\Gamma_{(Y\to AX)}$ is the spontaneous decay rate. Here we've assumed occupation numbers are low (e.g. all massive particles), so there are no fermi blocking or bose enhancement effects: the rate of producing Y is independent of the existing Y. The equation immediately shows that for $\Gamma \ll H$ the collision term is small compared to the Hubble expansion, so the system will go out of equilibrium.

For annihilation and inverse annihilation $A + \bar{A} \leftrightarrow Y + \bar{Y}$ we have

$$\dot{n}_A + 3Hn_A = -R_{(\bar{A}A\to\bar{Y}Y)} n_A n_{\bar{A}} + R_{(\bar{Y}Y\to\bar{A}A)} n_Y n_{\bar{Y}} \,. \quad (139)$$

In equilibrium the RHS must be zero to have no net change in particle numbers so the rates are related by

$$R_{(\bar{Y}Y\to\bar{A}A)} n_Y^{(\mathrm{eq})} n_{\bar{Y}}^{(\mathrm{eq})} = R_{(\bar{A}A\to\bar{Y}Y)} n_A^{(\mathrm{eq})} n_{\bar{A}}^{(\mathrm{eq})}. \quad (140)$$

This relation is called *detailed balance*, and can be very useful for relating the backward and forward rates.

D. Relic abundance of massive stable particles

If the species A is stable, then the dominant process which can change the number of particles in a comoving volume are the annihilation and inverse annihilation $A + \bar{A} \leftrightarrow X + \bar{X}$. We will also assume that there is no asymmetry between particles and antiparticles, and that X and $\bar{X}$ are in thermal equilibrium, i. e.,

$$\begin{aligned} n_A &= n_{\bar{A}} \,, \quad (141) \\ n_X &= n_{\bar{X}} = n_X^{(\mathrm{eq})} \quad (142) \end{aligned}$$

The Boltzmann equation, Eq. (139), with the detailed balance result gives

$$\begin{aligned} \dot{n}_A + 3Hn_A &= -\langle \sigma_A v \rangle n_A n_{\bar{A}} + \langle \sigma_X v \rangle n_X^{(\mathrm{eq})} n_{\bar{X}}^{(\mathrm{eq})} \quad (143) \\ &= -\langle \sigma_A v \rangle [n_A^2 - (n_A^{(\mathrm{eq})})^2] \,. \quad (144) \end{aligned}$$

For any species of particles that is not being created or destroyed $n_A \propto a^{-3}$ so that $a^3 n_A$ is constant. More generally we can take out the expansion dilution by defining a comoving number density $\propto n_A a^3$, for example $Y_A \equiv n_A/s \propto a^3 n_A$ since we know the total entropy density scales as $s \propto a^{-3}$ if s is dominated by thermal radiation. Then

$$\dot{Y}_A = -n_A \langle \sigma_A v \rangle Y_A \left[1 - \left(\frac{Y_A^{(\mathrm{eq})}}{Y_A} \right)^2 \right] = -\Gamma_A Y_A \left[1 - \left(\frac{Y_A^{(\mathrm{eq})}}{Y_A} \right)^2 \right] . \quad (145)$$

We are interested in the relative change in comoving number as the universe expands, so it's useful to rewrite this as

$$\frac{d \ln Y_A}{d \ln a} = -\frac{\Gamma_A}{H} \left[1 - \left(\frac{Y_A^{(\mathrm{eq})}}{Y_A} \right)^2 \right] . \quad (146)$$

Here is our "rule of thumb" in the above equation: Γ_A/H describes the "efficiency" of the annihilations. When $\Gamma_A \gg H$ there are many collisions on the timescale of the expansion, so interactions are fast enough for the species A to remain thermalized, and $Y_A \to Y_A^{(\rm eq)}$ (the term in brackets on the RHS must be very small if $\Gamma \gg H$). When $\Gamma_A \ll H$ the right hand side turns itself off so $Y_A = \text{const}$, and the abundance Y_A "freezes in" to the value $\sim Y_A^{(\rm eq)}|_{T_f}$, where T_f is the "freeze out" temperature.

The equilibrium form of Y is given in the relativistic and non-relativistic limits by

$$Y_A^{(\rm eq)} = \frac{45}{2\pi^4}\zeta(3)\frac{g_A^{\rm n}}{g_{*S}} \qquad \text{if } T \gg m_A \tag{147}$$

$$Y_A^{(\rm eq)} = \frac{45}{4\sqrt{2\pi^7}}\frac{g_A}{g_{*S}}\left(\frac{m_A}{T}\right)^{3/2} e^{-m_A/T} \qquad \text{if } T \ll m_A \tag{148}$$

where for bosons $g_A^{\rm n} = g_A$ and for fermions $g_A^{\rm n} = 3g_A/4$ and we took $\mu_A = 0$ assuming no particle-antiparticle asymmetry.

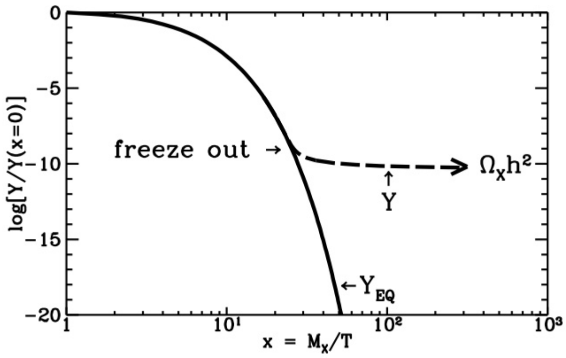

FIG. 7: $Y(x)$ for massive stable particles. For $T \gg M(x \ll 1)$, Y is at its equilibrium value, but decays exponentially as the temperature becomes lower. At freeze-out the particle stops interacting and Y asymptotes to a constant value.

1. Hot relics

If the species decouples when still relativistic, then at that time $Y_A^{(\rm eq)}$ is constant and the final, asymptotic value of Y_A is quite insensitive to the details of freeze-out:

$$Y_{A,0} \approx Y_A^{(\rm eq)}(T_f) = 0.278\frac{g_A^{\rm n}}{g_{*S}(T_f)} \,. \tag{149}$$

Assuming that there is no entropy production since A decouples, its abundance, number density and relic mass density today (assuming $m \gg T_0$) are just:

$$n_{A,0} = s_0 Y_{A,0} \qquad \Omega_{A,0} = \frac{8\pi G}{3H_0^2} s_0 Y_{A,0} m_A \,. \tag{150}$$

2. Cold relics

If the species decouples when $m \gg T$, then the abundance in equilibrium is suppressed by the Boltzmann factor $Y_A^{(\rm eq)} \simeq \exp(-m_A/T)$, and hence changes rapidly as the universe expands. Calculating the correct freeze-out value is therefore more tricky, and generally requires numerical integration of the Boltzmann equation.

We shall simply parametrise the temperature dependence of the average cross-section by

$$\langle \sigma_A v \rangle \equiv \sigma_0 \left(\frac{T}{m_A}\right)^n \tag{151}$$

such that:

$$\frac{\Gamma_A}{H} = \frac{a^2 a^{-n}}{a^3} Y_A \frac{(a^3 s)\sigma_0}{(a^2 H)} \left(\frac{aT}{m_A}\right)^n \equiv k Y_A a^{-(n+1)} \tag{152}$$

where the bracketed terms in the middle expression are constant in the radiation-dominated thermal state where we assume freeze-out happens. As soon as $T \sim m$, the equilibrium density becomes very small, and to get a rough estimate we can drop $Y_A^{(\text{eq})}/Y_A \ll 1$ giving

$$\frac{d \ln Y_A}{d \ln a} \approx -\frac{\Gamma_A}{H} = -k Y_A a^{-(n+1)}. \tag{153}$$

This can then be integrated to solve for $Y_{A,\infty}$ a long time after freeze out

$$\int_{Y_A^{(\text{eq})}}^{Y_{A,\infty}} \frac{dY_A}{Y_A^2} = -k \int_{a_f}^{a_\infty} a^{-(n+2)} da \approx -\frac{k}{n+1} a_f^{-(n+1)} \tag{154}$$

when we integrate to a time much later than freeze-out where $a_\infty \gg a_f$. The LHS is dominated by the late-time limit where $Y \sim Y_{A,\infty}$ (where Y_A is much smaller than $Y_A^{(\text{eq})}$) hence

$$Y_{A,\infty} \approx \frac{n+1}{k} a_f^{n+1} = (n+1) \left. \frac{HT}{s\sigma_0 m_A} \left(\frac{m_A}{T}\right)^{n+1} \right|_f . \tag{155}$$

Since freeze-out happens when $T \sim m_A$, we can define $x_f \equiv m_A/T$ which is order unity, and the prefactor is constant giving

$$Y_{A,\infty} \simeq \frac{HT}{s\sigma_0 m_A}(n+1) x_f^{n+1} \propto \frac{g_*^{1/2}}{g_{*S}} \frac{(n+1) x_f^{n+1}}{m_P m_A \sigma_0} , \tag{156}$$

The important point to notice is that the relic abundance is inversely proportional to the annihilation cross-section σ_0, and the mass: the larger the cross-section, the more will annihilate in a given time; the larger the mass, the earlier freeze-out will happen, and the higher the interaction rate (because of the higher density and energy), and hence the more annihilate during freeze-out. Therefore lighter, weakly interacting particles survive with a larger abundance than the heavier, strongly interacting ones. Moreover, the relic density today would be independent of the mass:

$$\rho_A = m_A n_{A,0} = m_A s_0 Y_{A,\infty} \propto \sigma_0^{-1} , \tag{157}$$

$$\Omega_{A,0} = \frac{s_0 Y_{A,\infty} m_A}{3 H_0^2 m_P^2} \propto \sigma_0^{-1} . \tag{158}$$

Weakly interacting, stable particles can survive as a cold relic and become a candidate for "dark matter".

Check, using Eq. (158), that indeed if we demand $\Omega_A \sim O(1)$ then $\sigma_0 \sim 10^{-37}\text{cm}^2 \sim G_F^2$, i. e., of the order of the cross-section for weak interactions.

E. Big-Bang Nucleosynthesis (BBN)

This is covered in the cosmology course.