Imperial College QFFF, 2018-19

Cosmology
Lecture notes

Toby Wiseman

Toby Wiseman; Huxley 507, email: `t.wiseman@imperial.ac.uk`

Books

This course is not based directly on any one book. Appropriate reading for the course is;

- Weinberg, "Cosmology"
  also an older book "Gravitation and cosmology"
- Kolb and Turner, "The early universe"
  (somewhat dated now!)
- Liddle and Lyth, "Cosmological inflation and large scale structure"
- Peacock, "Cosmological physics"
- Mukhanov, "Physical foundations of cosmology"
- Dodelson, "Modern cosmology"

For issues of GR, my favourite book is obviously Wald, "General Relativity".

Mathematica

I have written some mathematica notebooks to illustrate certain calculations. These can be downloaded from my PWP. In order to run them you will need Mathematica. Instructions for downloading this are at;
`www3.imperial.ac.uk/physics/staff/computing/software`

Conventions

I will use the following conventions; $c = 1$ and $(- + + +)$ signature.

For geometry I will use;

$$\Gamma^{c}{}_{ab}(x) \equiv \frac{1}{2} g^{cd} \left(\frac{\partial g_{db}}{\partial x^{a}} + \frac{\partial g_{da}}{\partial x^{b}} - \frac{\partial g_{ab}}{\partial x^{d}} \right)$$

so,

$$\nabla_{a} v^{b} \equiv \partial_{a} v^{b} + \Gamma^{b}{}_{ac} v^{c}$$

And,

$$[\nabla_{\alpha}, \nabla_{\beta}] v_{\mu} = R_{\alpha\beta\mu}{}^{\delta} v_{\delta}$$

so,

$$R_{\alpha\beta\mu}{}^{\delta} \quad = \quad \partial_{\beta} \Gamma^{\delta}{}_{\alpha\mu} - \partial_{\alpha} \Gamma^{\delta}{}_{\beta\mu} + \Gamma^{\nu}{}_{\alpha\mu} \Gamma^{\delta}{}_{\beta\nu} - \Gamma^{\nu}{}_{\beta\mu} \Gamma^{\delta}{}_{\alpha\nu}$$

Brief history of universe

10^{-43} s, 10^{19} GeV

Planck scale - quantum gravity - GUTs???

$\sim 10^{(-25,-40)}$ s, $\sim 10^{10-18}$ GeV

Inflation - quantum perturbations created, reheating

Hot big bang begins...

$\sim 10^{(-25,-38)}$ s, $\sim 10^{10-16}$ GeV

Baryogenesis

10^{-10} s, 10^{16} K, 100 GeV

Electroweak scale - dark matter annihilates to give relics??

10^{-4} s, 10^{13} K, 100 MeV

Quarks condense to hadrons

10 s, 10^{11} K, 1 MeV

Neutrino decoupling and e^+e^- annihilation ($e^+ + e^- \leftrightarrow \nu\bar{\nu}$)

Observation era begins...

10^2 s, 10^{10} K, 0.1 MeV

nucleosynthesis ($n + \nu \leftrightarrow p + e^-$)

10^5 yrs, 10^{4-5} K

Matter domination (Radiation-Matter equality)

3×10^5 yrs, 10^{3-4} K, 1 eV

Photon decoupling ($p + e \leftrightarrow H$), CMB formed

10^8 **yrs**

Structure formation; First objects - universe no longer close to FRW

5×10^9 **yrs**

Dark energy comes to dominate

13.8×10^9 **yrs,** 2.7 **K (effective temperature of photons)**

Now

Plan

1. **FRW**; symmetries, Friedmann equation, cosmologies for perfect fluids and scalar fields, observables, ΛCDM model

2. **Hot matter**; stat. mech. description of matter in thermal equibilirium, out of equilibrium description (Boltzmann equation), relics (dark matter)

3. **Thermal history**; hot radiation era, nucleosynthesis, recombination

4. **Inflation**; cosmological puzzles (flatness, horizon, monopole), scalar field cosmologies, slow roll inflation, generation of fluctuations during inflation

1 FRW

1.1 Cosmological principle

The universe is statistically the same at every location in it, and in every direction, ie. there is nothing special about where we live.

On suitably large scales (essentially beyond scales which are gravitationally bound - galaxy cluster scales $\sim Mpc$) and after a suitable averaging, the universe should be **homogeneous** and **isotropic** for some suitably chosen set of freely falling observers.

These observers define a foliation of spacetime into spatial slices. For any spacetime filled with observers we may (locally at least) write the metric as,

$$ds^2 = -dt^2 + h_{ij}(t, x) dx^i dx^j \tag{1}$$

This is simply a coordinate choice, where the time coordinate is the proper time as measured for a freely falling observer sitting at the constant spatial location x - (ie. x =const is a geodesic).

Technically we take **homogeneity and isotropy** to mean we may write the metric of spacetime as,

$$ds^2 = -dt^2 + d\Sigma^2 \ , \quad d\Sigma^2 = h_{ij}(t, x) dx^i dx^j \tag{2}$$

where at each constant time t, then $d\Sigma$ is the metric on a space which is homogeneous and isotropic.

1.1.1 Homogeneity and isotropy spatial geometries

Now at a fixed time t then Σ - the constant time slice - is simply a spatial geometry. Let us now suppress time in this discussion and simply consider homogeneous isotropic spatial geometries.

Consider first isotropy. Then at any point in Σ all directions must appear the same. Recall that any Riemannian geometry locally is flat. Consider a point, and then locally about that point we may write,

$$d\Sigma^2 = h_{ij} dx^i dx^j = A(\rho) d\rho^2 + B(\rho) d\Omega^2_{(2)} \tag{3}$$

where $d\Omega^2_{(2)} = d\theta^2 + \sin^2\theta d\phi^2$ is the round unit 2-sphere.

An important point is that A and B do not depend on θ and ϕ as we wish to have isotropy, and also there are no off diagonal $\rho\theta$ or $\rho\phi$ terms in the metric for the same reason.

[More precisely the metric has isometry group $SO(3)$, generated by the 3 Killing vectors of the 2-sphere.]

Let us chose a convenient coordinate $r^2 = B(\rho)$, and then,

$$\begin{aligned} d\Sigma^2 &= A(r)\left(\frac{2rdr}{B'(\rho)}\right)^2 + r^2 d\Omega^2_{(2)} \\ &= S(r)dr^2 + r^2 d\Omega^2_{(2)} \end{aligned} \tag{4}$$

Now since any space is locally flat, as $r \to 0$ we must have $\lim_{r\to 0} S(r) = 1$.

Now consider homogeneity. A necessary, but not sufficient condition for homogeneity is that the Ricci scalar is constant - ie. it doesn't depend on x^i. Now,

$$R_\Sigma = \frac{2}{r^2}\left(1 - \frac{1}{S} + \frac{rS'(r)}{S^2}\right) \tag{5}$$

where R_Σ is the Ricci scalar of the metric on Σ. Let us set this equal to a constant k. Then choose the normalization of the constant so,

$$R_\Sigma = 6k \tag{6}$$

Then one may solve the resulting equation, and with the condition that the space is locally flat at $r \sim 0$ we have,

$$S(r) = \frac{1}{1 - kr^2} \tag{7}$$

One then finds that the Ricci tensor is;

$$(R_\Sigma)_{ij} = 2kh_{ij} \tag{8}$$

and the metric is,

$$d\Sigma^2 \;\; = \;\; \frac{dr^2}{1-kr^2} + r^2 d\Omega^2_{(2)} \tag{9}$$

We see the Ricci tensor is proportional to the metric - this is a special condition called an **Einstein space**.

It is convenient to perform a coordinate transformation;

$$r \to r' = \frac{1}{a} r \tag{10}$$

for a constant a so then,

$$d\Sigma^2 \;\; = \;\; a^2 \left(\frac{dr'^2}{1 - ka^2\, r'^2} + r'^2 d\Omega^2_{(2)} \right) \tag{11}$$

By performing this scaling we may always write the metric as;

$$d\Sigma^2 \;\; = \;\; a^2 \left(\frac{dr'^2}{1 - k'\, r'^2} + r'^2 d\Omega^2_{(2)} \right) = a^2 d\Sigma'^2 \tag{12}$$

where now $k' = 0, \pm 1$, and a is a constant. Note now, $R_{\Sigma'} = 6\, k'$.

The 3 cases of k' are distinct, and a simply sets the size or 'scale' of the geometry. Let us now drop the 'primes' on r, k and consider the 3 cases $k = 0, \pm 1$.

Case $k = 0$ - flat
Then clearly $k = 0$ is the case of Euclidean 3-d flat space;

$$d\Sigma^2 = dr^2 + r^2 d\Omega^2_{(2)} = dx^2 + dy^2 + dz^2 \tag{13}$$

The homogeneity and isotropy are apparent in the cartesian and spherical coordinates; 3 translations and 3 rotations respectively.

Case $k = 1$ **- 3-sphere**
For $k = 1$ then Σ is a unit radius (round) 3-sphere. Let us see why. Consider 4-dimensional Euclidean space with coordinates $x^A = (z, x^i)$. Then we may write the metric as,

$$ds^2_{(4d)} = dz^2 + dr^2 + r^2 d\Omega^2_{(2)} = \delta_{AB} dX^A dX^B \tag{14}$$

Now consider embedding a unit 3-sphere Σ into this space,

$$z^2 + r^2 = 1 \quad \Longrightarrow \quad rdr = -zdz \tag{15}$$

Hence the induced metric is,

$$\begin{aligned} d\Sigma^2 &= \left(\frac{dz}{dr}\right)^2 dr^2 + dr^2 + r^2 d\Omega^2_{(2)} \\ &= \left(\frac{1}{1-r^2}\right) dr^2 + r^2 d\Omega^2_{(2)} \end{aligned} \tag{16}$$

using,

$$1 + \left(\frac{dz}{dr}\right)^2 = 1 + \frac{r^2}{z^2} = \frac{z^2 + r^2}{z^2} = \frac{1}{1-r^2} \tag{17}$$

Homogeneity and isotropy; We may see the isometries of the 3-sphere by noting that the full isometry group $SO(4)$ acts on the 3-sphere in the obvious way in the Cartesian coordinates X^A in the 4-d Euclidean space. One can translate this action into the z, r, θ, ϕ coordinates although it is complicated. The action of homogeneity are the 3 rotations of a point to another point, and of isotropy are the 3 rotations about an axis.

Case $k < 0$ **- 3-hyperboloid**
The $k = -1$ case is similar to the sphere. Now instead of embedding a surface in Euclidean space, we instead embed in 4d Minkowski spacetime,

$$ds^2 = -dz^2 + dr^2 + r^2 d\Omega^2_{(2)} = -dz^2 + \delta_{ij} dX^i dX^j \tag{18}$$

where now the extra coordinate z is a 'time' coordinate. Now a hyperboloid, Σ, with unit radius is embedded as,

$$z^2 - r^2 = \quad \Longrightarrow \quad rdr = zdz \tag{19}$$

and its induced metric is,

$$
\begin{aligned}
d\Sigma^2 &= -\left(\frac{dz}{dr}\right)^2 dr^2 + dr^2 + r^2 d\Omega^2_{(2)} \\
&= \left(\frac{1}{1+r^2}\right) dr^2 + r^2 d\Omega^2_{(2)}
\end{aligned}
\tag{20}
$$

now using,

$$
1 - \left(\frac{dz}{dr}\right)^2 = 1 - \frac{r^2}{z^2} = \frac{z^2 - r^2}{z^2} = \frac{1}{1+r^2}
\tag{21}
$$

Homogeneity and isotropy; We see the embedding and metric of the hyperboloid are invariant under a Lorentz transformation of the 4-d embedding Minkowski spacetime, ie. under $SO(1,3)$. Hence the isometry group of the 3-hyperboloid Σ is $SO(1,3)$. The action on the Minkowski coordinates z, X^i is straightforward, but is complicated in the coordinates r, θ, ϕ. The action of homogeneity and isotropy are generated by combinations of the 3 boosts and 3 rotations.

1.1.2 The FRW spacetime

Now let us return to the full spacetime, rather than a constant t slice, and consider the time dependence.

We have seen that the geometry of a constant t slice can be written as,

$$
d\Sigma^2 = a^2 \left(\frac{1}{1-kr^2} dr^2 + r^2 d\Omega^2_{(2)} \right)
\tag{22}
$$

for a constant a setting the scale of the space, and $k = 0, \pm 1$ setting its character.

However we must recall that a was a constant of integration, and may be different depending on the spatial slice we pick. We can use the above coordinates on each time slice provided we let $a = a(t)$. Now the full spacetime metric takes the form,

$$
ds^2 = -dt^2 + a(t)^2 d\Sigma^2_{(k)} \ , \quad d\Sigma^2_{(k)} = \frac{1}{1-kr^2} dr^2 + r^2 d\Omega^2_{(2)} \ , \quad k = 0, \pm 1
\tag{23}
$$

and the function $a(t)$ is the 'scale factor' controlling how the size of the homogeneous isotropic spatial slices changes in time.

[Note that one might wonder why $k \neq k(t)$? Such topology change is not possible in a smooth manner.]

We use the terminology;

1. $k = 0$ is a **flat** universe
2. $k = 1$ is a **closed** universe (sphere spatial sections)
3. $k = -1$ is an **open** universe (hyperbolic spatial sections)

Note that in the flat and open cases the geometry of the spatial slices is infinite - the spatial volume is infinite. However, for the closed case the universe has finite volume.

1.1.3 Conformal time

We may choose a new time coordinate, τ called conformal time such that;

$$ds^2 = a(\tau)^2 \left(-d\tau^2 + d\Sigma^2_{(k)}\right) \tag{24}$$

defined by $dt = a(\tau)d\tau$ so that $d\tau = dt/a(t)$ and hence,

$$\tau(t) = \int_{t_0}^{t} \frac{dt'}{a(t')} \tag{25}$$

One simplification of using conformal time is that null geodesics in this spacetime follow the same curves as in the spacetime $ds^2 = -d\tau^2 + d\Sigma^2_{(k)}$. However τ does not have an interpretation as a proper time.

1.1.4 Another coordinate chart

We have seen that we may write,

$$d\Sigma^2_{(k)} = \frac{1}{1-kr^2}dr^2 + r^2 d\Omega^2_{(2)} \tag{26}$$

in isotropic coordinates. There is another convenient chart for computing the Christoffel symbol ${\Gamma^{(h)i}}_{jk}$,

$$d\Sigma^2_{(k)} \;=\; h_{ij}dx^i dx^j = \left(\delta_{ij} + \frac{k\,x^i x^j}{1-k|x|^2}\right) dx^i dx^j \;, \quad |x|^2 = x^n x^m \delta_{nm} \tag{27}$$

and in this chart one can conveniently write,

$${\Gamma^{(h)i}}_{jk} \;=\; k\,x^i\,h_{jk} \tag{28}$$

1.2 Properties of FRW

Writing FRW in the general form earlier,

$$ds^2 = -dt^2 + a(t)^2 \Sigma^2_{(0,\pm1)} = -dt^2 + a(t)^2 h_{ij}(x) dx^i dx^j \tag{29}$$

then one finds the metric and inverse metric in coordinates $x^\mu = (t, x^i)$ is,

$$g_{\mu\nu} = \begin{pmatrix} -1 & 0 \\ 0 & a^2(t) h_{ij} \end{pmatrix} \; , \quad g^{\mu\nu} = \begin{pmatrix} -1 & 0 \\ 0 & \frac{1}{a^2(t)} h^{ij} \end{pmatrix} \tag{30}$$

where h^{ij} is the inverse metric to h_{ij}.

The non-vanishing Christoffel components are then,

$$\begin{aligned} \Gamma^t{}_{ij} &= a\,\dot{a} h_{ij} \\ \Gamma^i{}_{tj} &= \frac{\dot{a}}{a}\delta_{ij} \\ \Gamma^i{}_{jk} &= \Gamma^{(h)i}{}_{jk} \end{aligned} \tag{31}$$

where $\Gamma^{(h)i}{}_{jk}$ is the Christoffel symbol of the 3-d spatial geometry $h_{ij}(x)$.

The components, $\Gamma^t{}_{ti} = \Gamma^i{}_{tt} = 0$ by isotropy, and the components $\Gamma^t{}_{tt} = 0$ happen to vanish using this proper time coordinate t.

Then one finds the Ricci tensor components;

$$\begin{aligned} R_{tt} &= -3\frac{\ddot{a}}{a} \\ R_{ti} &= 0 \\ R_{ij} &= R^{(h)}_{ij} + \left(2\dot{a}^2 + a\ddot{a}\right) h_{ij} \end{aligned} \tag{32}$$

where $R^{(h)}_{ij}$ is the Ricci tensor of h_{ij} and we recall that $R^{(h)}_{ij} = 2k h_{ij}$ and hence,

$$R_{ij} = \left(2k + 2\dot{a}^2 + a\ddot{a}\right) h_{ij} \tag{33}$$

Again $R_{ti} = 0$ due to isotropy.

This yields an Einstein tensor $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R$ as,

$$\begin{aligned} G_{tt} &= \frac{3k}{a^2} + \frac{3\dot{a}^2}{a^2} \\ G_{ti} &= 0 \\ G_{ij} &= \left(-k - \dot{a}^2 - 2a\ddot{a}\right) h_{ij} \end{aligned} \tag{34}$$

1.2.1 Special geodesics of FRW

By symmetry, x^i =constant is a timelike geodesic - with proper time t. These are the worldlines our preferred observers. These observers are called **comoving observers** as they free fall with the homogeneous, isotropic frame. They have 4-velocity $v^\mu = dx^\mu/d\tau = dx^\mu/dt = (1, 0, 0, 0)$.

A simple set of geodesics are the null curves passing through $r = 0$ is isotropic coordinates,

$$ds^2 = -dt^2 + a^2(t)\left(\frac{1}{1-kr^2}dr^2 + r^2 d\Omega^2_{(2)}\right) \tag{35}$$

By symmetry θ, ϕ =constant for these null geodesics, so the metric on their world line is,

$$ds^2_{curve} = -dt^2 + \frac{a^2(t)}{1-kr^2}dr^2 = 0 \tag{36}$$

which must vanish for a null curve so,

$$\frac{1}{a}dt = \pm\frac{1}{\sqrt{1-kr^2}}dr \tag{37}$$

for the curve. Consider a null geodesic arriving at an observer at $r = 0$ at time $t = t_o$, having previously passed through radius R at time T (with $t_o > T$). Thus r is decreasing with t and so we require the negative sign above. Then integrating (and flipping the limits in $\int dr$) we obtain,

$$\int_T^{t_o} \frac{1}{a(t)}dt = \int_0^R \frac{1}{\sqrt{1-kr^2}}dr = \begin{cases} R & k = 0 \\ \sin^{-1}(R) & k = 1 \\ \sinh^{-1}(R) & k = -1 \end{cases} \equiv f_k(R) \tag{38}$$

1.2.2 General geodesics of FRW

Consider a geodesic $x^\mu = (T(\lambda), R(\lambda), \Theta(\lambda), \Phi(\lambda))$ with affine parameter λ. In order to obtain the geodesic equations we may vary the action,

$$S = \int d\lambda L \ , \quad L = g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = -\dot{T}^2 + a^2(T)\left(\frac{\dot{R}^2}{1-kR^2} + R^2\left(\dot{\Theta}^2 + \sin^2\Theta\dot{\Phi}^2\right)\right) \tag{39}$$

where $\dot{=} d/d\lambda$ (not d/dt).

Consider a geodesic passing through $r = 0$. Then by symmetry again it should have Θ, Φ =constant. More generally symmetry implies there should be a class of geodesics with only radial motion, i.e. with Θ, Φ =constant.

Check: Φ and Θ equations are (respectively;

$$a^2 R^2 \sin^2 \Theta \dot{\Phi} = J \ , \quad \frac{d}{d\lambda}\left(a^2 R^2 \dot{\Theta}\right) = a^2 R^2 \sin\Theta \cos\Theta \dot{\Phi}^2 \tag{40}$$

for constant of integration J. These are indeed satisfied for Θ, Φ =constant, and $J = 0$.

Note: We might also consider geodesic motion which is not radial. For this we would have to solve the above complicated system. However, we know from homogeneity that we can consider any point as $r = 0$, so other rays not going through $r = 0$ are always related to ones through the origin by homogeneity.

Now we are left with varying the line element along the curve,

$$L = -\dot{T}^2 + \frac{a^2(T)}{1 - kR^2}\dot{R}^2 \tag{41}$$

We have already solved this in the null case above. Consider now the timelike case. Using,

$$\frac{\partial L}{\partial \dot{R}} = \frac{2a^2 \dot{R}}{1 - kR^2} \ , \quad \frac{\partial L}{\partial R} = \frac{2ka^2 R\dot{R}^2}{(1 - kR^2)^2} \tag{42}$$

this yields the Euler-Lagrange equation, $d/d\lambda(\partial L/\partial \dot{R}) = \partial L/\partial R$;

$$\frac{d}{d\lambda}\left(\frac{a^2 \dot{R}}{1 - kR^2}\right) = \frac{ka^2 R\dot{R}^2}{(1 - kR^2)^2} \tag{43}$$

Now dividing by $a^2 \dot{R}/(1 - kR^2)$ gives,

$$\frac{d}{d\lambda}\left(\ln\left(\frac{a^2 \dot{R}}{1 - kR^2}\right)\right) = \frac{kR\dot{R}}{1 - kR^2} = -\frac{1}{2}\frac{d}{d\lambda}\left(\ln\left(1 - kR^2\right)\right) \tag{44}$$

So then,

$$\dot{R} = \frac{A}{a^2}\sqrt{1 - kR^2} \tag{45}$$

where A is a constant of integration.

Now using $L = \mu \equiv 0, \pm 1$ we have,

$$\dot{T}^2 = \frac{a^2(T)}{1 - kR^2}\dot{R}^2 - \mu = \frac{A^2 - \mu a^2}{a^2} \tag{46}$$

so that,

$$\dot{T} = \frac{\sqrt{A^2 - \mu a^2}}{a} \tag{47}$$

Hence the curve obeys,

$$\frac{dR}{dT} = \frac{\dot{R}}{\dot{T}} = \frac{A\sqrt{1 - kR^2}}{a\sqrt{A^2 - \mu a^2}} \tag{48}$$

Consider a ray through $r = 0$ at $t = t_o$ and previously through $r = R$ at $t = T$ with $t_o > T$. Hence during the interval $T < t < t_0$ then $dR/dT < 0$ and so $A < 0$. Then (flipping the limits in R);

$$\int_T^{t_o} \frac{|A|dT}{a(T)\sqrt{A^2 - \mu\, a^2(T)}} = \int_0^R \frac{dR}{\sqrt{1 - kR^2}} = \begin{cases} R & k = 0 \\ \sin^{-1}(R) & k = 1 \\ \sinh^{-1}(R) & k = -1 \end{cases} \tag{49}$$

and so we may calculate $R = R(T)$. This obviously agrees with the case $\mu = 0$ from before.

Aside on photons/wave

Consider a high frequency photon - so its wavelength is much shorter than the local curvature scales. Then in an LIF we may write that,

$$A_\mu = e^{ik_\mu x^\mu} e_\mu \tag{50}$$

for constant polarisation e_μ (obeying constraints) and constant wave vector k_μ, obeying $k^2 = 0$. The 4-momentum of the photon is,

$$p^\mu = \hbar k^\mu \tag{51}$$

The fact that k^μ is constant may be written covariantly as,

$$k^\mu \nabla_\mu k^\nu = 0 \tag{52}$$

ie. the wave crests follow null curves. Now more generally in a curved spacetime, a high frequency photon may be written as,

$$A_\mu = e^{i\phi(t,x)} e_\mu(t,x) \tag{53}$$

where the phase ϕ is rapidly varying, and now the field $k_\mu = \partial_\mu \phi$ again in the short wavelength limit, obeys the geodesic equation,

$$k^\mu \nabla_\mu k^\nu = 0 \tag{54}$$

and the 4-momentum is again $p^\mu = \hbar k^\mu$. This is what is known as the *geometric optics approximation.*

1.2.3 Gravitational redshift

Consider a photon travelling along a null geodesic ($\mu = 0$). Then its 4-momentum is $p^\mu = \hbar k^\mu$ with k^μ obeying the null geodesic condition $k^\mu \nabla_\mu k^\nu = 0$ with $k^\mu k_\mu = 0$. Hence we may write,

$$k^\mu = \frac{dx^\mu}{d\bar{\lambda}} \tag{55}$$

for some affine parameter $\bar{\lambda}$. Now (for $\hbar = 1$, $c = 1$)

$$p^\mu = \frac{dx^\mu}{d\lambda} = \left(\dot{T}, \dot{R}, 0, 0\right) = \left(\frac{A}{a(T)}, \frac{A}{a^2(T)}\sqrt{1 - kR^2}, 0, 0\right) \tag{56}$$

Consider a comoving observer with 4-velocity $v^\mu = (1, 0, 0, 0)$. They measure the photon to have an energy $E_{comove} = -p^\mu v_\mu$, so,

$$E_{comove} = +p^t = \frac{A}{a(T)} \tag{57}$$

Hence the energy of the photon measured by observers in the cosmological frame redshifts as $E \sim 1/a(T)$.

A unit norm spatial vector (tangent to constant t slices) for such an observer, in the direction of the photon, is $n^\mu = (0, 1/\sqrt{g_{rr}}, 0, 0)$, so that $n^\mu n_\mu = 1$ and $n^\mu v_\mu = 0$. Then the magnitude of momentum $|p|_{comove}$ that the observer measures the photon to have in the direction n^μ is,

$$\begin{aligned} |p|_{comove} &= |p^\mu n_\mu| = |g_{rr} p^r n^r| = |\sqrt{g_{rr}} p^r| = \frac{a}{\sqrt{1-kR^2}} \frac{|A|}{a(T)^2} \sqrt{1-kR^2} \\ &= \frac{|A|}{a(T)} \end{aligned} \tag{58}$$

Thus we have, $a(T)|p|_{comove}$ =constant. In the case $k = 0$, this result is seen as a conservation law using Cartesian spatial coordinates due to translation invariance in space. More generally it is a conservation law due to the isometry of homogeneity. The fact that $E_{comove} \sim 1/a$ is simply a result of the fact that in any LIF a photon has $E = |p|$.

We have derived the famous gravitational redshift result, namely that for a photon emitted at time t_e with comoving energy E_e (and momentum $|p|_e = E_e$), and received at t_o, comoving energy E_o (and momentum $|p|_o = E_o$), then,

$$\frac{E_o}{E_e} = \frac{a(t_e)}{a(t_o)} \tag{59}$$

Hence a photon is redshifted by an expanding universe - in a sense it is tunnelling out of a gravitational potential well. Since frequency of a photon $E = \hbar\nu$, one also has,

$$\frac{\nu_o}{\nu_e} = \frac{a(t_e)}{a(t_o)} \tag{60}$$

It is conventional to define the **redshift** Z as,

$$1 + Z \equiv \frac{\nu_e}{\nu_o} = \frac{a(t_o)}{a(t_e)} \tag{61}$$

where now t_o should be taken as the time **today**. Thus redshift $Z(t_e)$ vanishes for photons emitted today (ie. very close by) and is $Z > 0$ for photons emitted in our past. It is important as it provides a **very physical** measure of time, $Z = Z(t)$ by,

$$Z(t) = \frac{a(t_o)}{a(t)} - 1 \tag{62}$$

where we think of $t < t_o$ as the time in the past when a photon which we receive today was originally emitted. Inverting this relation, $t = t(Z)$ gives an elegant relation between comoving proper time and redshift which is **directly observed**. However, this relation clearly depends on the detail of $a(t)$, and hence the cosmological model.

A more pedestrian derivation:
Consider a null ray emitted at $r = R$ at time $t = t_e$ and propagating to the origin at time $t = t_o$. Then we have previously found,

$$\int_{t_e}^{t_o} \frac{dt}{a(t)} = \int_0^R \frac{dr}{\sqrt{1-kr^2}} = \begin{cases} R & k=0 \\ \sin^{-1}(R) & k=1 \\ \sinh^{-1}(R) & k=-1 \end{cases} \tag{63}$$

Now consider a second ray emitted slightly later at $t = t_e + \delta t_e$ and received at $t = t_o + \delta t_o$. Then,

$$\int_{t_e+\delta t_e}^{t_o+\delta t_0} \frac{dt}{a(t)} = \int_0^R \frac{dr}{\sqrt{1-kr^2}} \tag{64}$$

Hence, differencing these relations we obtain,

$$\begin{aligned} 0 &= \int_{t_e+\delta t_e}^{t_o+\delta t_0} \frac{dt}{a(t)} - \int_{t_e}^{t_o} \frac{dt}{a(t)} \\ &= \frac{\delta t_e}{a(t_e)} - \frac{\delta t_o}{a(t_o)} \end{aligned} \tag{65}$$

Now since the relation between the emitted and observed comoving frequencies is $\frac{\nu_o}{\nu_e} = \frac{\delta t_e}{\delta t_o}$ then we obtain as before;

$$1 + Z \equiv \frac{\nu_e}{\nu_o} = \frac{a(t_o)}{a(t_e)} \tag{66}$$

Massive geodesics
Let us continue our discussion of geodesics and conclude with a massive particle ($\mu = -1$ and $\lambda = \tau \neq t$). Then it has 4-momentum $p^\mu = mv^\mu$,

$$p^\mu = m\frac{dx^\mu}{d\lambda} = m\left(\dot{T}, \dot{R}, 0, 0\right) = \left(\frac{m\sqrt{A^2 + a^2(T)}}{a(T)}, \frac{mA}{a^2(T)}\sqrt{1-kR^2}, 0, 0\right) \tag{67}$$

Now a comoving observer measures E_{comove} and $|p|_{comove}$ as;

$$E_{comove} = p^t = m\sqrt{1 + \frac{A^2}{a^2(T)}} = m\gamma \tag{68}$$

where γ is the usual Lorentz factor $\gamma = 1/\sqrt{1-v^2}$ for observed velocity v (so $v^2 = 1 - 1/\gamma^2$), and so,

$$v^2 = \frac{A^2}{a^2 + A^2} \tag{69}$$

Then note that,

$$\gamma v = \frac{|A|}{\sqrt{a^2 + A^2}}\sqrt{1 + \frac{A^2}{a^2(T)}} = \frac{|A|}{a(T)} \tag{70}$$

And for the momentum,

$$\begin{aligned} |p|_{comove} &= |p^\mu n_\mu| = \frac{a}{\sqrt{1-kR^2}}\frac{m|A|}{a(T)^2}\sqrt{1-kR^2} \\ &= \frac{m|A|}{a(T)} = m\gamma v \end{aligned} \tag{71}$$

with observed velocity v. Thus $E^2_{comove} = |p|^2_{comove} + m^2$ as usual, but with a red shifting momentum, $|p|_{comove} \sim 1/a(T)$.

Note that as for the photon, we have $a|p|_{comove}$ =constant, which follows from the conservation law due to spatial homogeneity. Recall this is easily seen computing geodesics for the flat $k = 0$ case in Cartesian coordinates where homogeneity is manifest.

Thus more generally the redshift is really related to observed and emitted comoving 3-momentum;

$$1 + Z = \frac{a(t_e)}{a(t_o)} = \frac{|p|_o}{|p|_e} \tag{72}$$

for any free falling object - null or timelike. However, for photons $E = |p|$. Note in terms of observed energy for a massive particle we have,

$$1 + Z = \frac{|p|_o}{|p|_e} = \sqrt{\frac{E_o^2 - m^2}{E_e^2 - m^2}} \tag{73}$$

1.3 Observables

Let us now consider some important quantities which are related to observables in cosmology.

1.3.1 Comoving distance

If we are at $r = 0$ and a distant object is a radial coordinate position R at time t then we say it has comoving distance R. Note that if we and the object are comoving, then this distance remains constant in time.

1.3.2 Proper distance

The foliation of spacetime by the family of cosmological observers allows us to define the **proper distance** between two objects at a given time t. Suppose we consider an observer at time t at the origin $r = 0$, and an object at that time at radius $r = R$. Then the proper distance $d(t, R)$ is the geodesic distance within the spatial section Σ. The spatial slice Σ at constant t is;

$$d\Sigma^2 = a(t)^2 \left(\frac{dr^2}{1 - kr^2} + r^2 d\Omega^2_{(2)} \right) \tag{74}$$

and by symmetry the geodesic in Σ between the observer and the object travels on the radial line θ, ϕ =constant. Hence its proper distance is,

$$d(t, R) = a(t) \int_0^R \frac{dr}{\sqrt{1 - kr^2}} = a(t) \begin{cases} R & k = 0 \\ \sin^{-1}(R) & k = 1 \\ \sinh^{-1}(R) & k = -1 \end{cases} \tag{75}$$

Note that this is not a geodesic in the full spacetime, only within the geometry of the constant spatial slice defined by our cosmological observers.

Note also that for a null geodesic emitted at time $t = t_e$ from radius $r = R$ and reaching the origin $r = 0$ at time $t = t_o$ we have,

$$\int_{t_e}^{t_o} \frac{dt}{a(t)} = \int_0^R \frac{dr}{\sqrt{1 - kr^2}} \tag{76}$$

Hence, we have for a null ray,

$$\frac{d(t_o, R)}{a(t_o)} = \frac{d(t_e, R)}{a(t_e)} = \int_{t_e}^{t_o} \frac{dt}{a(t)} \tag{77}$$

1.3.3 Hubble function

A very important quantity is the **Hubble function**,

$$H(t) = \frac{a'(t)}{a(t)} \tag{78}$$

and its value today, ie. $t = t_0$, is called the Hubble constant $H_0 = H(t_0)$.

Consider a comoving observer at $r = 0$, and a comoving object at $r = R$. Then the Hubble function determines the rate of change of the proper distance between these objects.

$$\begin{aligned} \frac{d}{dt}d(t,R) &= a'(t) \begin{cases} R & k = 0 \\ \sin^{-1}(R) & k = 1 \\ \sinh^{-1}(R) & k = -1 \end{cases} \\ &= H(t)d(t,R) \end{aligned} \tag{79}$$

Note that $\dot{d}/d$ is independent of R.

This can be interpreted as a version of Hubble's law, namely that objects in an expanding universe recede from us with a velocity proportional to their distance. This movement of objects away from us is called the **Hubble flow**.

Note however that the 'velocity' $\dot{d}$ is not a very physical one. Let us consider now a more physical version of Hubble's law.

1.3.4 Nearby sources

Consider a source nearby to us at $r = 0$, $t = t_o$, emitting at time t_e and at radius $r = R$, so $\Delta t = t_o - t_e$ is small.

Now using our previous relation for the proper distance travelled by a null ray,

$$\frac{d(t_o, R)}{a(t_o)} = \int_{t_e}^{t_o} \frac{dt}{a(t)} \sim \frac{\Delta t}{a(t_o)} \tag{80}$$

then we see,

$$d(t_o, R) \sim \Delta t \tag{81}$$

This simply says the proper distance is approximately the light travel time.

Then we may also expand,

$$
\begin{aligned}
a(t_e) &= a(t_o) + a'(t_o)(t_e - t_o) + \dots \\
&= a(t_o)\,(1 - \Delta t \cdot H_o + \dots)
\end{aligned} \tag{82}
$$

Hence,

$$
1 + Z = \frac{a(t_o)}{a(t_e)} = 1 + \Delta t H_o + \dots \quad \Longrightarrow \quad Z \simeq \Delta t H_o \tag{83}
$$

Thus the Hubble constant determines the redshift of local sources.
Putting these results together we obtain Hubble's experimental result for nearby sources;

$$
Z \simeq H_o \cdot d(t_o, R) \tag{84}
$$

1.3.5 Value of Hubble constant

The Hubble constant H_o has a value, usually quoted in peculiar units;

$$
H \simeq 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}} \tag{85}
$$

What funny units! Recall a parsec; $1pc$ =distance at which 1 AU subtends an angle of 1 arc second! So $pc = 3.2\,\mathrm{light}\,yr = 3.1 \times 10^{16} m$. In terms of astronomical scales;

- Nearest star (Proxima Centauri) $\sim$ 1 pc
- Milky way (our galaxy) $\sim$ 1 Kpc
- Galaxy cluster $\sim$ 1-10 Mpc
- Hubble horizon (approx size of observable universe) $\sim$ 3000 Mpc

These units are useful in the sense of $\dot{d} = Hd$ so that an object $1Mpc$ away looks as though it is receding at a velocity $70\,\mathrm{km\,s^{-1}}$.

For reference, note that our peculiar motion, ie. the motion relative to the cosmological frame, is $\sim 400\,\mathrm{km\,s^{-1}}$ (which is typical value). Thus on large scales ($>$ cluster size) peculiar motion become negligible compared to the Hubble flow. On small scales it is an important effect, dominating the motion due to cosmological expansion.

1.3.6 Angular diameter and Luminosity distance

So far we have the relation $R(t)$ and $Z(t)$ for a null ray provided we know the cosmological expansion $a(t)$. Suppose we observe light from an object and measure its redshift, then we know the time of emission t and the radius R the light was emitted from.

Now since we also know the proper distance relation $d(t, R)$ to the object then if we have an independent measure of distance then we can check this agrees. This will be a check of our cosmological model.

Turned around, if we know the distance d and redshift Z of many observed objects, we can hope to reconstruct the function $a(t)$ and deduce the correct cosmological expansion rate (and then using Einstein's equations understand the matter driving the expansion).

However while Z is something directly and accurately observed, the proper distance d is not something that is easy to deduce simply looking at a distance source. Hence it is useful to have two more practical distance measures; **angular diameter distance** d_A and **luminosity distance** d_L.

Angular diameter distance d_A
Let us be comoving observers at $r = 0$. Consider a comoving object at radius $r = R$ with proper size Δ (for example a distant galaxy). Assume we know this size - for example we have a good model for galaxies. Suppose it emits light at time $t = t_e$, and we observe this at time $t = t_o$. Let us see the object to have an angular diameter Θ on the sky.

Then the angular diameter distance is defined to be **the distance the object *appears* to be from us** ie. assuming a simple Euclidean spatial geometry and infinite speed of light. So,

$$d_A = \frac{\Delta}{\Theta} \tag{86}$$

Now from the spatial geometry at time t_e,

$$ds^2_{space} = a(t_e)^2 \left(\frac{dr^2}{1 - kr^2} + r^2 \left(d\theta^2 + \sin^2\theta d\phi^2 \right) \right) \tag{87}$$

we have,

$$\Delta = a(t_e)\, R\, \Theta \tag{88}$$

and thus we see,

$$d_A = a(t_e)\, R = \frac{1}{1+Z} a(t_0)\, R \tag{89}$$

Recall that,

$$d(t,R) = a(t) \begin{cases} R & k=0 \\ \sin^{-1}(R) & k=1 \\ \sinh^{-1}(R) & k=-1 \end{cases} \quad , \quad \frac{d(t_o,R)}{a(t_o)} = \frac{d(t_e,R)}{a(t_e)} \tag{90}$$

Hence we can relate d_A to the proper distance today $d(t_o, R)$. For example if $k = 0$,

$$\begin{aligned} d_A &= a(t_e)\, R = d(t_e, R) = \frac{a(t_e)}{a(t_o)} d(t_o, R) \\ &= \frac{1}{(1+Z)}\, d(t_o, R) \end{aligned} \tag{91}$$

Thus if we know the size of an object Δ we can measure d_A and its redshift Z directly, and then deduce its proper distance d from these.

Luminosity distance d_L
Suppose we have a similar source, but now instead of knowing the size Δ of the source, we instead know its **intrinsic luminosity** L = energy emitted (isotropically) per time. When we observe the light we see an **apparent luminosity** ℓ =energy at receiver per area per time.

We define the luminosity distance which is **the distance the object *appears* to be from us**;

$$\ell = \frac{L}{4\pi d_L^2} \tag{92}$$

From the spatial geometry at time t_o, there light from the source is spread over a 2-sphere with area,

$$A(S_2) = 4\pi \left(a(t_o) R\right)^2 \tag{93}$$

The apparent luminosity is given by;

$$\ell = \frac{1}{(1+Z)^2} \frac{L}{A(S_2)} \tag{94}$$

where the two factors of $1/(1+Z)$ are due to;

- The energy of each photon is redshifted by $\nu_o/\nu_e = 1/(1+Z)$
- The rate of receiving photons is reduced relative to emission frequency by the same factor. Recall our previous calculation of gravitational redshift precisely showed in section 1.2.3 (see equation 65) that $\delta t_o = (1+Z)\delta_e$.

Hence we see, that;

$$\ell = \frac{L}{4\pi d_L^2} = \frac{1}{(1+Z)^2}\frac{L}{A(S_2)} \quad \Longrightarrow \quad d_L = (1+Z)a(t_o)\,R \qquad (95)$$

Recalling that $d_A = a(t_e)\,R$, we see,

$$d_L = (1+Z)^2 d_A \qquad (96)$$

Again we can relate d_L to the proper distance today $d(t_o, R)$. For example if $k = 0$,

$$d_L = (1+Z)a(t_o)\,R = (1+Z)d(t_o, R) \qquad (97)$$

Experiment
We measure a distant source, and determine its redshift Z. Either knowing its size Δ, or its luminosity L we then determine its distance d_A or d_L.

Recall,

$$d_A = a(t_e)\,R = \frac{1}{(1+Z)}a(t_o)\,R\ , \quad d_L = (1+Z)a(t_o)\,R \qquad (98)$$

and so our measurements give us directly Z and $a(t_o)R$.

Suppose we understand $a(t)$ and wish to test our cosmological model. Then our measurement for a source gives both Z and R. But the photon relations,

$$1+Z = \frac{a(t_o)}{a(t_e)}\ , \quad \int_{t_e}^{t_o} \frac{dt}{a(t)} = \begin{cases} R & k=0 \\ \sin^{-1}(R) & k=1 \\ \sinh^{-1}(R) & k=-1 \end{cases} \qquad (99)$$

relates these as we know $R = R(t_e)$ and $Z = Z(t_e)$ so we know $R = R(Z)$. Thus we can check the agreement.

In practice we use lots of sources to build up the relation between Z and R, and then constrain parameters in our model to fit this.

1.4 Horizons and the big bang

The inverse of the Hubble function, H^{-1}, defines a distance scale, denoted the **Hubble scale**. This is sometimes loosely referred to as the 'horizon size'. Generally it doesn't refer to any horizon, but is an important physical scale that essentially determines the scale beyond which the spacetime doesn't look flat for causal physics. However in certain situations there are two types of horizon - **particle** and **event** horizons.

Suppose we have an expanding universe which started at $t = t_{BB}$ at a Big Bang, so that $a(t_{BB}) = 0$. Using the null relation for a ray reaching $r = 0$ at time t starting at a radius $r = R_H$ at the Big Bang;

$$\int_{t_{BB}}^{t} \frac{dt'}{a(t')} = \begin{cases} R_H & k = 0 \\ \sin^{-1}(R_H) & k = 1 \\ \sinh^{-1}(R_H) & k = -1 \end{cases} \tag{100}$$

Thus this gives a relation $R_H = R_H(t)$. The **Particle horizon size** $d_H(t)$ at time t is the proper radius of the comoving volume with this radius $R_H(t)$. Recall the proper distance,

$$d(t, R) = a(t) \begin{cases} R & k = 0 \\ \sin^{-1}(R) & k = 1 \\ \sinh^{-1}(R) & k = -1 \end{cases} \tag{101}$$

and hence we see,

$$d_H(t) = d(t, R_H(t)) = a(t) \int_{t_{BB}}^{t} \frac{dt'}{a(t')} \tag{102}$$

Existence of a particle horizon
Note that this integral **may or may not converge** depending on the behaviour of $a(t)$ at the big bang. We shall discuss this later. For usual matter and radiation this does converge (with $a \sim (t - t_{BB})^{1/2}$ for hot radiation). However, for exotic inflaton matter, this does not.

Event horizon
If the integral

$$\int_{t}^{\infty} \frac{dt'}{a(t')} = \begin{cases} R_E & k = 0 \\ \sin^{-1}(R_E) & k = 1 \\ \sinh^{-1}(R_E) & k = -1 \end{cases} \tag{103}$$

converges, it means that an event at $r = 0$ and time t can only influence comoving observers within a radius $r \leq R_E(t)$. An observer outside the radius R_E will never see the event, however long they wait. This comoving size is known as the event horizon for the event at $r = 0$, time t, and its proper size is,

$$d_E(t) = d(t, R_E(t)) = a(t) \int_t^{\infty} \frac{dt'}{a(t')} \tag{104}$$

Note that if the universe recollapses then t may have a maximum value t_{max}. In this case the event horizon is defined in the obvious way as,

$$\int_t^{t_{max}} \frac{dt'}{a(t')} = \begin{cases} R_E & k = 0 \\ \sin^{-1}(R_E) & k = 1 \\ \sinh^{-1}(R_E) & k = -1 \end{cases} \tag{105}$$

Comment: The existence of dark energy leads to an event horizon, where $d_E \sim H_0^{-1}$. Hence as the universe expands, objects which are not gravitationally bound to us move further away ($\dot{d} = Hd$). Once they move beyond a proper distance d_E they can never influence us again.

Note that while the terminology **horizon** is used, these are not proper causal horizons in the sense of black holes. Their definition is dependent on the choice of an observer. The region associated to one observers horizon will not coincide with that of another.

1.5 Perfect fluids

Consider the stress tensor for a single perfect fluid;

$$T_{\mu\nu} = (\rho + P)\, u_\mu u_\nu + P g_{\mu\nu} \tag{106}$$

where ρ, P are the energy density and pressure, and u^μ is the local fluid 4-velocity (so $u^2 = -1$). The relation $P = P(\rho)$ is the **equation of state** of the fluid. Recall that for such a fluid the dynamics is simply determined by the conservation of stress-energy,

$$\nabla_\mu T^{\mu\nu} = 0 \tag{107}$$

which yields **relativistic hydrodynamics**.

Important: We may have several such fluids in a spacetime and provided they don't exchange energy-momentum with other matter, **each one will individually obey the above equations**.

The equations of hydro are best described by projecting into the u^μ and orthogonal directions. Firstly project the conservation equation into the direction u^μ;

$$u_\mu \nabla_\nu T^{\mu\nu} = 0 \quad \Longrightarrow \quad u^\nu \partial_\nu \rho + (\rho + P) \nabla_\nu u^\nu = 0 \tag{108}$$

recalling $u^2 = -1$. This gives an evolution equation for the density.

Let us now project onto an orthogonal direction n^μ to the motion u^μ, so that $u^\mu n_\mu = 0$. Note that n^μ must be space like (since v^μ is timelike). Then,

$$n_\mu \nabla_\nu T^{\mu\nu} = 0 \quad \Longrightarrow \quad n_\mu \left((\rho + P) u^\nu \nabla_\nu u^\mu + g^{\mu\nu} \partial_\nu P \right) = 0 \tag{109}$$

and this is equivalent to,

$$(\rho + P) u^\nu \nabla_\nu u_\mu + \left(\delta_\mu^\nu + u_\mu u^\nu \right) \partial_\nu P = 0 \tag{110}$$

again recalling that $u^2 = -1$. The quantity $\perp_\mu^\nu = \delta_\mu^\nu + u_\mu u^\nu$ is a projector into the directions orthogonal to u^μ. It has the property that,

$$\perp_\mu^\nu u^\mu = 0 \; , \quad \perp_\mu^\nu n^\mu = n^\nu \tag{111}$$

for any n^μ orthogonal to u^μ. Then,

$$(\rho + P) u^\nu \nabla_\nu u_\mu = - \perp_\mu^\nu \partial_\nu P \tag{112}$$

gives an evolution equation for the velocity.

We will derive this later, but taking $P = w\rho$ for a constant w, then,

- **Dust / cold matter**: $w = 0$, so $P \simeq 0$
- **Radiation / hot matter**: $w = \frac{1}{3}$, so $P \simeq \frac{1}{3}\rho$
- **Vacuum energy**: $w = -1$

Note in the case of vacuum energy it isn't really a fluid;

$$T_{\mu\nu} = -\rho g_{\mu\nu} \tag{113}$$

However, the condition $\nabla_\mu T^{\mu\nu} = 0$ then implies $\partial_\mu \rho = 0$, and hence the energy density is not dynamical, but constant in time and space. It is natural to define the **cosmological constant**

$$\Lambda \equiv 8\pi G_N \rho \tag{114}$$

so we may write,

$$T_{\mu\nu} = -\frac{1}{8\pi G_N} \Lambda g_{\mu\nu} \tag{115}$$

and then we move this term to the LHS of the Einstein equations,

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = 8\pi G_N \tilde{T}_{\mu\nu} \tag{116}$$

where $\tilde{T}_{\mu\nu}$ is the stress tensor due to other matter - not a cosmological term. Now Λ is an inverse length squared, where the length gives the radius of curvature associated to the cosmological constant.

1.6 The cosmological stress tensor

Now consider cosmological matter in FRW,

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = -dt^2 + a(t)^2 h_{ij}(x) dx^i dx^j \tag{117}$$

so that the stress tensor is homogeneous and isotropic. Then since there is no preferred direction we must have $T_{ti} = 0$, and $T_{ij} \propto h_{ij}$. Due to homogeneity we must have T_{tt} is a function of time only, and $T_{ij} = f(t) h_{ij}$ for a function of time f. Then we **define** the total density and total pressure of cosmological matter in FRW to be,

$$T_{tt} \equiv \rho_{tot}(t) \ , \quad T_{ij} \equiv a^2(t) P_{tot}(t) h_{ij}(x) \tag{118}$$

where $T_{\mu\nu}$ is the sum of all matter components. Why? Consider a perfect fluid which is homogeneous and isotropic. Then $u^\mu = (1, 0, 0, 0)$ and $\rho = \rho(t)$, $P = P(t)$. Thus,

$$\begin{aligned} T_{tt} &= (\rho + P) u_t u_t + P g_{tt} = \rho \\ T_{ti} &= 0 \\ T_{ij} &= P g_{ij} = a^2 P h_{ij} \end{aligned} \tag{119}$$

and thus the density and pressure agree.

Note however, that while any matter with cosmological symmetry has a stress tensor with the form above, it does **not** mean it behaves like a single perfect fluid with a simple local equation of state $P_{tot}(t) = F(\rho_{tot}(t))$. In general matter with have a complicated non-local (in time) equation of state.

Let us consider the conservation equation $\nabla^\mu T_{\mu\nu} = 0$ for a stress tensor sharing the cosmological symmetry. Recall that this applies to the total stress tensor (as a result of the Bianchi identities), but also applies separately to the stress tensor of any matter component that doesn't interact with other matter. The time component of the conservation equation implies;

$$\begin{aligned} 0 &= \nabla^\mu T_{\mu t} = \partial^t T_{tt} - \Gamma^\mu{}_{\mu\alpha} T^\alpha{}_t - \Gamma^\mu{}_{t\alpha} T^\alpha{}_\mu \\ &= \dot{\rho} - \Gamma^i{}_{it} T^t{}_t - \Gamma^j{}_{ti} T^i{}_j \\ &= \dot{\rho} - 3\frac{\dot{a}}{a}(-\rho) + \frac{\dot{a}}{a}\delta^j_i P \delta^i_j \\ &= \dot{\rho} + 3\frac{\dot{a}}{a}(\rho + P) \end{aligned} \tag{120}$$

This applies to both the total stress tensor, but also separately to any component that doesn't interact with other components. The only assumption is that the matter has the cosmological symmetry.

1.7 Cosmological perfect fluids

Consider now the behaviour of a single perfect fluid with cosmological symmetry that doesn't interact with other matter. Let it have equation of state $P = w\rho$ for constant w. Then conservation implies,

$$\dot{\rho} + 3\frac{\dot{a}}{a}\rho(1+w) = 0 \quad \Longrightarrow \quad \rho = \frac{k}{a^{3(1+w)}} \tag{121}$$

for a constant of integration k. We usually write this constant in terms of the scale factor today, a_o, and the value of the fluid density today, ρ_o, as,

$$\rho = \rho_0 \left(\frac{a_0}{a}\right)^{3(1+w)} \tag{122}$$

Consider the important cases;

- **Dust / cold matter**: $w = 0$; $\rho \sim \frac{1}{a^3}$, so the matter simply dilutes.

- **Radiation / hot matter**: $w = \frac{1}{3}$; $\rho \sim \frac{1}{a^4}$, so the radiation both dilutes and its constituents redshift.

- **Vacuum energy**: $w = -1$, as we saw before we have $\rho =$constant.

1.8 Dynamics and the Friedmann equation

We have previously states the components of the Einstein tensor. Recall,

$$\begin{aligned} ds^2 &= -dt^2 + a(t)^2 h_{ij}(x) dx^i dx^j \ , \quad R^{(h)}_{ij} = 2k h_{ij} \\ G_{tt} &= \frac{3k}{a^2} + \frac{3\dot{a}^2}{a^2} \ , \quad G_{ij} = \left(-k - \dot{a}^2 - 2a\ddot{a}\right) h_{ij} \end{aligned} \tag{123}$$

Then the Einstein equations are $G_{\mu\nu} = 8\pi G_N T_{\mu\nu}$ where we emphasise that $T_{\mu\nu}$ is the **total** stress tensor, and hence is conserved and in our FRW approximation must have cosmological symmetry. Using equations (119) we then arrive at two important relations. Firstly the tt component directly gives;

$$\frac{\dot{a}^2}{a^2} + \frac{k}{a^2} = \frac{8\pi G_N}{3} \rho_{tot} \qquad \Longrightarrow \qquad H^2 = \frac{8\pi G_N}{3} \rho_{tot} - \frac{k}{a^2} \tag{124}$$

This is the **Friedmann equation** and directly determines the Hubble function $H = \dot{a}/a$. Note that we also have the matter conservation equation;

$$\dot{\rho} + 3\frac{\dot{a}}{a} \left(\rho_{tot} + P_{tot}\right) = 0 \tag{125}$$

Suppose we have a single fluid for matter. Recall the fluid equations of motion are simply given by stress energy conservation. Then given the universe at a time t_i with scale factor $a(t_i)$ and density $\rho(t_i)$, with known equation of state $P = P(\rho)$ then the Friedmann and conservation equations allow us to integrate forward in time to determine $a(t)$ and $\rho(t)$.

For a combination of perfect fluids (with cosmological symmetry), then the total energy and pressure will be the sum of the various components

$$\rho_{tot} = \sum_i \rho_i \ , \quad P_{tot} = \sum_i P_i \tag{126}$$

and each will be conserved separately;

$$\dot{\rho}_i + 3\frac{\dot{a}}{a} \left(\rho_i + P_i\right) = 0 \tag{127}$$

(Obviously this implies the total stress tensor is conserved). Again the Friedmann equation, the individual conservation equations and equations of state $P_i = P_i(\rho_i)$, allow one to integrate $a(t)$ and $\rho_i(t)$ forward in time.

More generally, for any matter with cosmological symmetry, the Friedmann equation together with the matter equations of motion allow integration forward in time. For example, for an (interacting) scalar field ϕ the equation of motion is,

$$\nabla^2 \phi = \nabla^\mu \nabla_\mu \phi = V'(\phi) \tag{128}$$

and the stress tensor is,

$$T_{\mu\nu} = \partial_\mu \phi \partial_\nu \phi - g_{\mu\nu} \left(\frac{1}{2} (\partial_\alpha \phi)^2 + V(\phi) \right) \tag{129}$$

and it is easy to show the matter equation of motion implies the stress tensor is conserved. For a mass, then $V(\phi) = \frac{m^2}{2}\phi^2$. Then for cosmological symmetry the scalar equation becomes;

$$\ddot{\phi} + 3\frac{\dot{a}}{a}\dot{\phi} = -V'(\phi) \tag{130}$$

and so knowing a, ϕ and $\dot{\phi}$ at a time t_i then the Friedmann equation and scalar equation allow integration forward in time. In this case,

$$\rho_{tot} = \frac{1}{2}\dot{\phi}^2 + V(\phi) \ , \quad P_{tot} = \frac{1}{2}\dot{\phi}^2 - V(\phi) \tag{131}$$

Note that the case $V(\phi) = 0$ is $P = \rho$ (so $w = 1$ - a 'stiff' equation of state). However, for general $V(\phi)$ this is not of a simple $P = P(\rho)$ form.

What about the spatial components of the Einstein equations? The spatial ij-components may be combined with the tt component to eliminate terms involving H^2 to give another equation which is useful;

$$\frac{\ddot{a}}{a} = -\frac{4\pi G_N}{3} (\rho + 3P) \tag{132}$$

Note that this equation is independent of k. An important point is that (due to the Bianchi identities) this equation is equivalent to the Friedmann equation and total matter conservation equation. Thus it contains no new information, but nicely summarises how $\ddot{a}$ behaves.

Acceleration and expansion
Cosmological expansion is characterised by $\dot{a} > 0$. However, acceleration is characterised by $\ddot{a} > 0$. We see for this to hold we require;

$$\rho + 3P < 0 \tag{133}$$

where we note this is the total energy density and pressure. Suppose the stress tensor is dominated by a perfect fluid with equation of state $P = w\rho$. In this case acceleration implies,

$$w < -\frac{1}{3} \tag{134}$$

Hence we see cold or hot matter ($w = 0, \frac{1}{3}$) give a decelerated expansion. However dark energy gives acceleration.

Note that if $\rho > 0$ then if $\dot{a} > 0$ at some time, then the universe remains expanding unless $k = 1$ (the closed case). In the case $k = 0, -1$ this is decelerated or accelerated expansion depending on the sign of $\ddot{a}$. In the closed case $k = 1$, if we have $\ddot{a} < 0$ then if the scale factor reaches a radius such that $\dot{a} = 0$ so,

$$\frac{8\pi G}{3}\rho = \frac{1}{a^2} \tag{135}$$

then after this point $\dot{a} < 0$, and the universe will recollapse to a big crunch.

1.9 Simple cosmological solutions

Consider the flat case $k = 0$ with a single perfect fluid $P = w\rho$ for constant w with $w > -1$. Then from conservation we already saw;

$$\rho = \rho_o \left(\frac{a_o}{a}\right)^{3(1+w)} \tag{136}$$

Putting this into the Friedmann equation;

$$\frac{\dot{a}^2}{a^2} = \frac{8\pi G_N \rho_o}{3}\left(\frac{a_o}{a}\right)^{3(1+w)} = H_o^2 \left(\frac{a_o}{a}\right)^{3(1+w)} \tag{137}$$

Hence (taking the positive root for an expanding universe);

$$a^{\frac{1+3w}{2}}\dot{a} = H_o a_o^{\frac{3(1+w)}{2}} \tag{138}$$

and integrating,

$$\frac{2}{3(1+w)} a^{\frac{3(1+w)}{2}} = H_o a_o^{\frac{3(1+w)}{2}} (t-k) \tag{139}$$

for a constant of integration $k = 0$. We now see why we restricted to $w > -1$. We conventionally fix this by taking the big bang $a = 0$ to be at time $t = 0$, so,

$$a = a_o \left(\frac{3(1+w)}{2} H_o t \right)^{\frac{2}{3(1+w)}} \tag{140}$$

Then the density goes as,

$$\begin{aligned} \frac{\rho}{\rho_o} &= \left(\frac{a}{a_o} \right)^{-3(1+w)} = \left(\left(\frac{3(1+w)}{2} H_o t \right)^{\frac{2}{3(1+w)}} \right)^{-3(1+w)} \\ &= \left(\frac{2}{3(1+w)} \frac{1}{H_o t} \right)^2 \end{aligned} \tag{141}$$

Note then that for all values of w we have $\rho \sim t^{-2}$.

Hubble function
The Hubble function is;

$$H = \frac{\dot{a}}{a} = \frac{2}{3(1+w)} \frac{1}{t} \tag{142}$$

Null geodesics
Null geodesics are governed by the equation (recall $k = 0$);

$$R = \int_{t_e}^{t_o} \frac{dt}{a(t)} = \frac{1}{a_o} \left(\frac{3(1+w)}{2} H_o \right)^{-\frac{2}{3(1+w)}} \int_{t_e}^{t_o} dt\, t^{-\frac{2}{3(1+w)}} \tag{143}$$

where,

$$\int_{t_e}^{t_o} dt\, t^{-\frac{2}{3(1+w)}} = \frac{3(1+w)}{1+3w} \left[t^{\frac{1+3w}{3(1+w)}} \right]_{t_e}^{t_o} \tag{144}$$

Lower limit: For the lower limit to converge for $t_e \to 0$ we require;

$$0 < \frac{1+3w}{3(1+w)} \quad \Longrightarrow \quad w > -\frac{1}{3} \quad \text{or} \quad w < -1 \tag{145}$$

Thus for non-accelerating matter $-\frac{1}{3} < w$ precisely gives a finite lower limit. For accelerating matter in the range $-1 \leq w < -\frac{1}{3}$ it does not. (The range $w < -1$ is unlikely to be physically relevant).
Upper limit: For the upper limit to converge for $t_o \to \infty$ we require the opposite;

$$\frac{1+3w}{3(1+w)} < 0 \quad \Longrightarrow \quad -1 < w < -\frac{1}{3} \tag{146}$$

Thus this requires accelerated expansion.

Particle horizon
In such a model the particle horizon size (recall $k = 0$) at time $t = t_o$ is;

$$d_H(t_o) \;=\; a(t_o) \int_0^{t_o} \frac{dt}{a(t)} = \left(\frac{3(1+w)}{2} H_o \right)^{-\frac{2}{3(1+w)}} \int_0^{t_o} dt \, t^{-\frac{2}{3(1+w)}} \tag{147}$$

Hence it is finite for non-accelerating matter with $-\frac{1}{3} < w$ such as dust $w = 0$ and radiation $w = 1/3$. Then the lower limit gives zero, and so,

$$\begin{aligned} d_H(t_o) &= \left(\frac{3(1+w)}{2} H_o \right)^{-\frac{2}{3(1+w)}} \frac{3(1+w)}{1+3w} (t_o)^{\frac{1+3w}{3(1+w)}} \\ &= \frac{3(1+w)}{1+3w} \left(\frac{3(1+w)}{2} H_o \right)^{-1} \\ &= \frac{2}{1+3w} \frac{1}{H_o} \end{aligned} \tag{148}$$

recalling that $H_o = \frac{2}{3(1+w)} \frac{1}{t_o}$.

Event horizon
In this model the event horizon size (recall $k = 0$) at time $t = t_o$ is;

$$d_E(t) = a(t_o) \int_{t_o}^{\infty} \frac{dt}{a(t)} = \left(\frac{3(1+w)}{2} H_o \right)^{-\frac{2}{3(1+w)}} \int_{t_o}^{\infty} dt \, t^{-\frac{2}{3(1+w)}} \tag{149}$$

Thus we see for $-1 < w < -\frac{1}{3}$ that the upper limit gives zero and so up to a sign we get the same as for the particle horizon above,

$$\begin{aligned} d_E(t) &= \left(\frac{3(1+w)}{2}H_o\right)^{-\frac{2}{3(1+w)}} \frac{3(1+w)}{1+3w}\left[t^{\frac{1+3w}{3(1+w)}}\right]_{t_o}^{\infty} \\ &= -\frac{3(1+w)}{1+3w}\left(\frac{3(1+w)}{2}H_o\right)^{-1} \\ &= -\frac{2}{1+3w}\frac{1}{H_o} \end{aligned} \quad (150)$$

where the quantity is positive due to the range of w.

Cold matter and $k = 0$
Consider our special case of cold matter. Note that this case is called the **Einstein-de Sitter model** (although it has nothing to do with de Sitter spacetime). Then,

$$\rho = \rho_o \left(\frac{a_o}{a}\right)^3 \quad (151)$$

Then we have decelerated expansion with;

$$\frac{a}{a_o} = \left(\frac{3}{2}H_o t\right)^{\frac{2}{3}} \ , \quad H = \frac{2}{3t} \quad (152)$$

Recall (as for all w) then $\rho \sim t^{-2}$. The particle horizon size is;

$$d_H(t_o) = \frac{2}{H_o} \quad (153)$$

but there is no event horizon.

Hot matter/Radiation $w = 1/3$ and $k = 0$
For radiation we have,

$$\rho = \rho_o \left(\frac{a_o}{a}\right)^4 \quad (154)$$

Again this yields $\rho \sim t^{-2}$. As for dust we have decelerated expansion;

$$\frac{a}{a_o} = (2H_o t)^{\frac{1}{2}} \ , \quad H = \frac{1}{2t} \quad (155)$$

and the particle horizon is;

$$d_H(t_o) \quad = \quad \frac{1}{H_o} \tag{156}$$

and again there is no event horizon without acceleration.

Special case: Cosmological constant $w = -1$ **and** $k = 0$
This is an example of **de Sitter** spacetime. Then as we saw earlier the energy density is constant,

$$\rho = \rho_o \equiv \frac{\Lambda}{8\pi G_N} \tag{157}$$

with Λ a constant inverse length squared. Let us assume that $\Lambda > 0$ - a **positive cosmological constant**, so that the energy density $\rho = -P > 0$. This negative pressure leads to accelerated expansion of the spacetime. Then the Friedmann equation gives;

$$H^2 = \frac{\dot{a}^2}{a^2} = \frac{8\pi G_N \rho_o}{3} = \frac{\Lambda}{3} \tag{158}$$

so (taking the positive root);

$$a = a_0 \, e^{\sqrt{\frac{\Lambda}{3}}(t-t_o)} \ , \quad H = H_o = \sqrt{\frac{\Lambda}{3}} \tag{159}$$

where the Hubble function is constant. An important point is that now $a \to 0$ as $t \to -\infty$. So the singularity or 'big bang' is not at $t = 0$, but is an **infinite proper time in the past**. For this reason there is no particle horizon;

$$\int_{-\infty}^{t_o} \frac{dt}{a(t)} = \int_{-\infty}^{t_o} dt e^{-\sqrt{\frac{\Lambda}{3}}(t-t_o)} = \frac{1}{-\sqrt{\frac{\Lambda}{3}}} \left[e^{-\sqrt{\frac{\Lambda}{3}}(t-t_o)} \right]_{-\infty}^{t_o} \to \infty \tag{160}$$

Another important feature is that we see **exponential expansion**. So the rate of acceleration is very quick. This acceleration leads to an event horizon;

$$\begin{aligned} d_E(t) &= a(t_o) \int_{t_o}^{\infty} \frac{dt}{a(t)} = \int_{t_o}^{\infty} dt\, e^{-\sqrt{\frac{\Lambda}{3}}(t-t_o)} \\ &= \frac{1}{-\sqrt{\frac{\Lambda}{3}}} \left[e^{-\sqrt{\frac{\Lambda}{3}}(t-t_o)} \right]_{t_o}^{\infty} \\ &= \sqrt{\frac{3}{\Lambda}} \end{aligned} \tag{161}$$

with the upper limit now giving zero.

1.10 Some other cosmological solutions

More generally the de Sitter solutions are simple to write down for $k = \pm 1, 0$ and $\Lambda > 0$. They solve the equation,

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = 0 \quad \Longrightarrow \quad R_{\mu\nu} = \Lambda g_{\mu\nu} \tag{162}$$

Let us define;

$$H_o = \sqrt{\frac{\Lambda}{3}} \tag{163}$$

as above.

Flat slicing of de Sitter, $k = 0$: we have already seen this;

$$ds^2 = -dt^2 + e^{2H_o t} \delta_{ij} dx^i dx^j \tag{164}$$

where $a \to 0$ as $t \to -\infty$.

Global de Sitter, $k = +1$:

$$ds^2 = -dt^2 + \frac{1}{H_o^2} \cosh H_o t\, d\Sigma_{k=+1}^2 \tag{165}$$

Note that in this case we have no $a = 0$ 'big bang', and for $t < 0$ we have a **contracting** universe!

Hyperbolic slicing of de Sitter, $k = -1$**:**

$$ds^2 = -dt^2 + \frac{1}{H_o^2} \sinh H_o t \, d\Sigma^2_{k=-1} \tag{166}$$

where $a \to 0$ as $t \to 0$.

In fact all these are coordinates on the same spacetime, with the global $k = +1$ case covering the **whole** of de Sitter, and the others only covering portions of it. Thus in fact for $k = 0$ and $k = -1$, then $a \to 0$ isn't really the big bang at all but rather just a coordinate singularity.

Matter plus radiation ($k = 0$**):** Consider now again the case $k = 0$ with both a matter and radiation perfect fluid. Now,

$$\frac{\dot{a}^2}{a^2} = \frac{8\pi G_N}{3} \left(\rho_m \left(\frac{a_0}{a} \right)^3 + \rho_r \left(\frac{a_0}{a} \right)^4 \right) \tag{167}$$

and thus,

$$\dot{a}^2 = \frac{A}{a} + \frac{B}{a^2} \,, \quad A = \frac{8\pi G_N}{3} \rho_m a_o^3 \,, \quad B = \frac{8\pi G_N}{3} \rho_r a_o^4 \tag{168}$$

Then,

$$\frac{a}{\sqrt{Aa + B}} da = dt \quad \Longrightarrow \quad t - t_k = \frac{2}{3\sqrt{A}} \left(a - 2\frac{B}{A} \right) \sqrt{a + \frac{B}{A}} \tag{169}$$

so,

$$(t - t_k) = \sqrt{\frac{\rho_m a_o^3}{6\pi G}} \left(a(t) - 2\frac{\rho_r a_0}{\rho_m} \right) \sqrt{a(t) + \frac{\rho_r a_0}{\rho_m}} \tag{170}$$

Note we may set the big bang to $t = 0$ by choosing an appropriate t_k;

$$\sqrt{6\pi G} t = \sqrt{\rho_m a_o^3} \left(a(t) - 2\frac{\rho_r a_0}{\rho_m} \right) \sqrt{a(t) + \frac{\rho_r a_0}{\rho_m}} + 2\frac{\rho_r^{3/2} a_o^3}{\rho_m} \tag{171}$$

Note that this is not straightforward to invert to obtain $a = a(t)$. Note that setting $\rho_r = 0$ we recover $t \sim a^{3/2}$ (i.e.. $a \sim t^{2/3}$). Note that taking $\rho_m \to 0$ is possible - a divergence must be absorbed into t_k - and then one is left with

$t \sim a^2$ (ie. $a \sim t^{1/2}$).

Matter plus Λ ($k = 0$): One can also solve for matter plus a cosmological constant. Then,

$$\frac{\dot{a}^2}{a^2} = \frac{8\pi G_N}{3}\left(\rho_\Lambda + \rho_m \left(\frac{a_0}{a}\right)^3\right) \tag{172}$$

and so,

$$\begin{aligned} \frac{8\pi G_N}{3} dt &= \sqrt{\frac{a}{\rho_\Lambda a^3 + \rho_m a_o^3}} da \\ &= \frac{2}{3\sqrt{\rho_\Lambda}} \ln\left[a^{\frac{3}{2}}\rho_\Lambda + \sqrt{\rho_\Lambda}\sqrt{a^3\rho_\Lambda + \rho_m a_o^3}\right] \end{aligned} \tag{173}$$

1.11 General cosmologies

Let us use the notation that t_o is the time today, with H_o and ρ_o being the values of $H(t_o)$ and $\rho_{tot}(t_o)$ today. We then define the **critical density**;

$$\rho_{crit} = \frac{3}{8\pi G_N} H_o^2 \tag{174}$$

This is the matter density a flat ($k = 0$) would have in order to reproduce the expansion rate H_o. Experimentally we find,

$$\rho_{crit} \simeq 10^{-26}\, h^2\, kg\, m^{-3} \sim 1\, \text{proton}/m^{-3} \tag{175}$$

We may write the Friedmann equation today as;

$$\rho_{crit} = \rho_o - \frac{3}{8\pi G_N}\frac{k}{a^2} \tag{176}$$

Hence we see;

- By construction for a flat universe $\rho_o = \rho_{crit}$
- A closed universe is **over dense**: $\rho_o > \rho_{crit}$
- An open universe is **under dense**: $\rho_o < \rho_{crit}$

Now consider an FRW universe with matter given by a cosmological constant Λ, cold matter ρ_M and hot matter/radiation ρ_R, all of which are not interacting with the other components. Then we write,

$$\begin{aligned}\rho_\Lambda &= \frac{\Lambda}{8\pi G_N} \\ \rho_M &= \rho_{M,o}\left(\frac{a_0}{a}\right)^3 \\ \rho_R &= \rho_{R,o}\left(\frac{a_0}{a}\right)^4\end{aligned} \tag{177}$$

for matter and radiation densities $\rho_{M,o}$, $\rho_{R,o}$ today. Then we have,

$$H^2 + \frac{k}{a^2} = \frac{8\pi G_N}{3}\left(\rho_\Lambda + \rho_{M,o}\left(\frac{a_0}{a}\right)^3 + \rho_{R,o}\left(\frac{a_0}{a}\right)^4\right) \tag{178}$$

Now let us use the critical density to rewrite this expression by defining;

$$\Omega_i = \frac{\rho_{i,o}}{\rho_{crit}} \quad , \qquad \frac{8\pi G_N}{3}\rho_{crit} = H_o^2 \tag{179}$$

and we also define,

$$\frac{8\pi G_N}{3}\rho_{crit}\Omega_k = H_o^2\Omega_k \equiv -\frac{k}{a_o^2} \tag{180}$$

These Ω_i for $i = \Lambda, M, R, k$ give the fraction of the energy density at $t = t_o$ in a component compared to the energy density ρ_{crit}. Note that for $k = 0$ then Ω_i give the fraction in a component compared to the total energy density. For $k \neq 0$ we should think of ρ_{crit} as the effective total energy density including thinking of the spatial curvature as a matter component.

Then we may write,

$$H^2 = \frac{8\pi G_N}{3}\rho_{crit}\left(\Omega_\Lambda + \Omega_k\left(\frac{a_0}{a}\right)^2 + \Omega_M\left(\frac{a_0}{a}\right)^3 + \Omega_R\left(\frac{a_0}{a}\right)^4\right) \tag{181}$$

and so the Friedmann equation (note we have also used all the fluid equations to get here) takes the simple form;

$$\left(\frac{H}{H_o}\right)^2 = \Omega_\Lambda + \Omega_k\left(\frac{a_0}{a}\right)^2 + \Omega_M\left(\frac{a_0}{a}\right)^3 + \Omega_R\left(\frac{a_0}{a}\right)^4 \tag{182}$$

and this relation completely determines the dynamics of the scale factor. Note that for $t = t_o$ we obtain the important relation;

$$1 = \sum_i \Omega_i = \Omega_\Lambda + \Omega_k + \Omega_M + \Omega_R \tag{183}$$

An important comment; clearly for very small a, radiation (if present) always dominates. At late times if $a \to \infty$ then Λ (if present) dominates. In that case then;

$$H \to \sqrt{\Omega_\Lambda} H_0 \tag{184}$$

Note that for our universe we believe $\Omega_\Lambda = 0.7$. Hence H, which today is H_o, is very close to its final value, so $H \to 0.8 H_0$.

To obtain the dynamics we take the positive root (for an expanding universe) and integrate;

$$t = \frac{1}{H_o} \int_0^{a(t)} \frac{da}{a\sqrt{\Omega_\Lambda + \ldots + \Omega_R \left(\frac{a_0}{a}\right)^4}} \tag{185}$$

where we assume that we have a big bang $a = 0$ at $t = 0$. A more physical parameterization is in terms of redshift. Define;

$$x(t) \equiv \frac{a(t)}{a_o} = \frac{1}{1 + Z(t)} \tag{186}$$

Note then that;

$$\begin{aligned} t = 0\,,\ a = 0 \quad &\Longrightarrow \quad x = 0 \\ t = t_o\,,\ a = a_o \quad &\Longrightarrow \quad x = 1 \end{aligned} \tag{187}$$

Then,

$$t = \frac{1}{H_o} \int_0^{\frac{1}{1+Z(t)}} \frac{dx}{x\sqrt{\Omega_\Lambda + \Omega_k x^{-2} + \Omega_M x^{-3} + \Omega_R x^{-4}}} \tag{188}$$

and the age of the universe today is;

$$t_o = \frac{1}{H_o} \int_0^{1} \frac{dx}{x\sqrt{\Omega_\Lambda + \Omega_k x^{-2} + \Omega_M x^{-3} + \Omega_R x^{-4}}} \tag{189}$$

Consider a null geodesic received at $t = t_o$ and $r = 0$, emitted at time t and radius R. Recall our usual relation;

$$\int_t^{t_o} \frac{dt'}{a(t')} = F_k(R) , \qquad F_k(R) = \begin{cases} R & k = 0 \\ \sin^{-1}(R) & k = 1 \\ \sinh^{-1}(R) & k = -1 \end{cases} \tag{190}$$

Now,

$$\begin{aligned} F_k(R) &= \int_{a(t)}^{a_o} \frac{1}{a} \frac{dt}{da} da = \int_{a(t)}^{a_o} \frac{da}{a^2 H(a)} \\ &= \frac{1}{a_o H_o} \int_{\frac{1}{1+Z}}^{1} \frac{dx}{x^2 \sqrt{\Omega_\Lambda + \Omega_k x^{-2} + \Omega_M x^{-3} + \Omega_R x^{-4}}} \end{aligned} \tag{191}$$

A nice way to write this is;

$$a_o R(Z) = \frac{1}{H_o \sqrt{\Omega_k}} \sinh \left[\sqrt{\Omega_k} \int_{\frac{1}{1+Z}}^{1} \frac{dx}{x^2 \sqrt{\Omega_\Lambda + \Omega_k x^{-2} + \Omega_M x^{-3} + \Omega_R x^{-4}}} \right] \tag{192}$$

where we recall $\Omega_k = -k/(a_o^2 H_o^2)$.

1.12 The ΛCDM model

We now consider the ΛCDM universe which is our 'standard model' of cosmology. It is a universe that is started by inflation, goes through a period of hot big bang, and at late times is dominated by Cold Dark Matter (CDM) together with conventional matter, and a cosmological constant.

At late times in the ΛCDM universe we assume that the universe has expanded to the point where radiation is irrelevant, so $\Omega_r \ll 1$. In fact as we discuss later for our universe $\Omega_r \sim 10^{-5}$. Thus we neglect it.

We also assume that $|\Omega_k| \ll 1$. Note that this does not strictly require $k = 0$, but just that the size of the spatial sections are sufficiently large that any spatial curvature is irrelevant. This is justified by the theory of inflation we will discuss later.

Now recalling that we have the constraint (setting $\Omega_k = \Omega_r = 0$);

$$1 = \Omega_\Lambda + \Omega_m \tag{193}$$

then at late times in the universe (ie. now) there are just two parameters to measure. Firstly H_o, and secondly say Ω_m, the fraction of energy density in matter relative to the cosmological constant.

Recall the angular diameter and luminosity distances for an object at redshift Z are;

$$d_A = \frac{1}{(1+Z)} a_o R(Z) \ , \quad d_L = (1+Z) a_o R(Z) \tag{194}$$

The most elegant determination of the ΛCDM parameters is from supernovae. Their redshift is directly measured. Certain (Type IIA) supernova are believed to be 'standard candles'. Once their light curves have been measured, we believe we can infer their total luminosity L. Then we may determine their luminosity distance from observation of apparent luminosity ℓ, as $\ell = L/(4\pi d_L^2)$. Then,

$$\begin{aligned} d_L(Z) &= (1+Z) a_o R(Z) \\ &= \frac{(1+Z)}{H_o} \int_{\frac{1}{1+Z}}^{1} \frac{dx}{x^2 \sqrt{\Omega_\Lambda + \Omega_M x^{-3}}} \end{aligned} \tag{195}$$

give the relation between measured luminosity distance and redshift.

Note that this integral can be computed in closed form - it is given in terms of an elliptic integral. However the analytic expression is not terribly useful.

After observing lots of supernovae, one fits the data to this relation to determine the best fit of the parameters H_o and Ω_m. One famously finds,

$$\begin{aligned} h &\simeq 0.7 \ , \quad H_o = 100\, h\, kms^{-1} Mpc^{-1} \\ \Omega_m &\simeq 0.3 \quad \Longrightarrow \quad \Omega_\Lambda \simeq 0.7 \end{aligned} \tag{196}$$

Hence there is a **positive** cosmological constant (dark energy).

Current state of the art measurements are from Planck satellite from a full CMB analysis

$$\begin{aligned} h &\simeq 0.67 \pm 0.015 \\ \Omega_\Lambda &\simeq 0.686 \pm 0.02 \end{aligned} \tag{197}$$

As you can see these agree nicely!

It is a great mystery why the cosmological constant is so small (according to any QFT calculation it should be enormous!), and yet is of order the energy density in matter today. No one knows the answers to these questions. The idea that dark energy may be dynamic (so called quintessence) rather than a cosmological constant may try to explain this, but the real issues I believe should be addressed by quantum gravity.

There are two important epochs;

Acceleration: Recall that,

$$\begin{aligned} \frac{\ddot{a}}{a} &= -\frac{4\pi G}{3}(\rho + 3P) \\ &= -\frac{4\pi G}{3}\left(-2\rho_\Lambda + \rho_{m,0}(1+Z)^3\right) \end{aligned} \tag{198}$$

Hence the redshift Z_{acc} at which acceleration began in the late universe was,

$$2\Omega_\Lambda = \Omega_m (1 + Z_{acc})^3 \quad \Longrightarrow \quad Z_{acc} = 0.67 \tag{199}$$

which is around $6.1\, Gyr$ ago.

Λ-matter equality: Defined as the time or redshift $Z_{\Lambda-m}$ when the energy densities in the two components (which drive the Friedmann equation) are equal. Hence.

$$\Omega_\Lambda = \Omega_m (1 + Z_{\Lambda-m\,eq})^3 \quad \Longrightarrow \quad Z_{\Lambda-m\,eq} = 0.33 \tag{200}$$

which is around $3.6\, Gyr$ ago.

1.13 Dark matter

The $\Omega_m = 0.3$ matter component is the total cold matter. This is the usual baryonic matter - stars and gas - but also dark matter. There are various techniques in astronomy to measure the amount of baryonic matter in the universe. However there are also methods to measure the total mass that gravitates in a system - for example weak lensing. The mass of galaxies assessed by directly accounting for the baryonic matter, or looking at the total gravitating mass is considerably different. Likewise the galactic dynamics that would follow only form the observed baryonic matter doesn't reproduce observed dynamics at all. The implication is that cold matter is actually composed not only of baryons but also a non-luminous **dark matter**.

If we split the cold matter into a dark matter and baryonic component, so,

$$\rho_m = \rho_{DM} + \rho_B \tag{201}$$

then different methods agree that,

$$\frac{\Omega_B}{\Omega_m} \sim 0.15 \tag{202}$$

ie. around only 15% of cold matter is baryonic. The rest is dark matter. Since $\Omega_m = 0.32$ (Planck) then,

$$\Omega_B \sim 0.05 \ , \quad \Omega_{DM} \sim 0.27 \tag{203}$$