Journeys 2018 cosmology [version 1, revised 7/20 at 1pm]

Kendrick Smith[1]

[1]*Perimeter Institute for Theoretical Physics, Waterloo, ON N2L 2Y5, Canada*
(Dated: July 20, 2018)

Note: these notes were written recently and may contain typos! If you notice any typos, please email me at `kmsmith@perimeterinstitute.ca`, and I'll correct them as soon as possible. An up-to-date version of the notes can be found at `https://github.com/kmsmith137/journeys2018_cosmology`.

I. EXPANSION HISTORY

- In special relativity, there is no preferred inertial frame, and there is no experiment that can tell whether an observer is "moving" or "at rest". In cosmology, the universe has a preferred rest frame! If we are allowed to make cosmological observations, then we can do experiments which determine our velocity relative to the cosmic rest frame.

  For example, we can measure the dipole in the cosmic microwave background (CMB). The largest source of CMB anisotropy is the CMB dipole: the CMB temperature on opposite sides of the sky differs by 3 millikelvin. The dipole arises because the Earth is moving at $10^{-3}c$ in the cosmic rest frame, and this Doppler-shifts the CMB, which is nearly uniform with temperature 2.726 Kelvin. The CMB dipole is a "cosmic speedometer" that gives the direction and magnitude of the Earth's velocity in the cosmic rest frame. In fact, the CMB dipole slowly varies over the course of a year, due to the Earth's motion around the sun.

- An observer (or worldline) is said to be "comoving" if the observer is at rest in the cosmic rest frame, for all times.

- In special relativity, there is no preferred clock. A Lorentz transformation will transform the time coordinate t to a different time coordinate t'. In cosmology, the universe has a preferred global time coordinate t: the proper time elapsed since the big bang, as measured by a comoving observer. When we say that the age of the universe is 13.8 billion years, we are referring to the value of this global time coordinate.

- The statement that the "universe is expanding" has the following precise meaning. Physical distances between comoving observers are changing with time. If $D_{\mathrm{phys}}(t)$ denotes the physically measured distance between two comoving observers, as a function of the time t when the measurement is made, then:

$$D_{\mathrm{phys}}(t) = a(t)D_{\mathrm{phys}}(t_{\mathrm{today}}) \tag{1}$$

  The function $a(t)$ defined by this equation is called the "scale factor", and is an increasing function of time. At $t = t_{\mathrm{today}}$, we have $a(t) = 1$, by definition. At the big bang ($t = 0$), we have $a(t) = 0$.

  A cartoon "plot" of $a(t)$ in the standard cosmological model (ΛCDM) is shown in Figure 1. The precise evolution of $a(t)$, including the $t^{1/2}$, $t^{2/3}$, and e^{Ht} scalings, will be explained in the next few pages!

- Given an event in spacetime, we define its *comoving coordinates* (x, y, z) to be the spatial coordinates at $t = t_{\mathrm{today}}$ of a comoving observer who passes through the event. In other words, the comoving coordinates are the spatial coordinates that would be measured if we wait until today to make the measurement.

  We have now defined global coordinates (x, y, z, t) on spacetime: the comoving spatial coordinates (x, y, z) and the proper time t since the big bang. We define the *comoving distance* D_c between two events (x, y, z, t) and (x', y', z', t') by:

$$D_c = \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2} \tag{2}$$

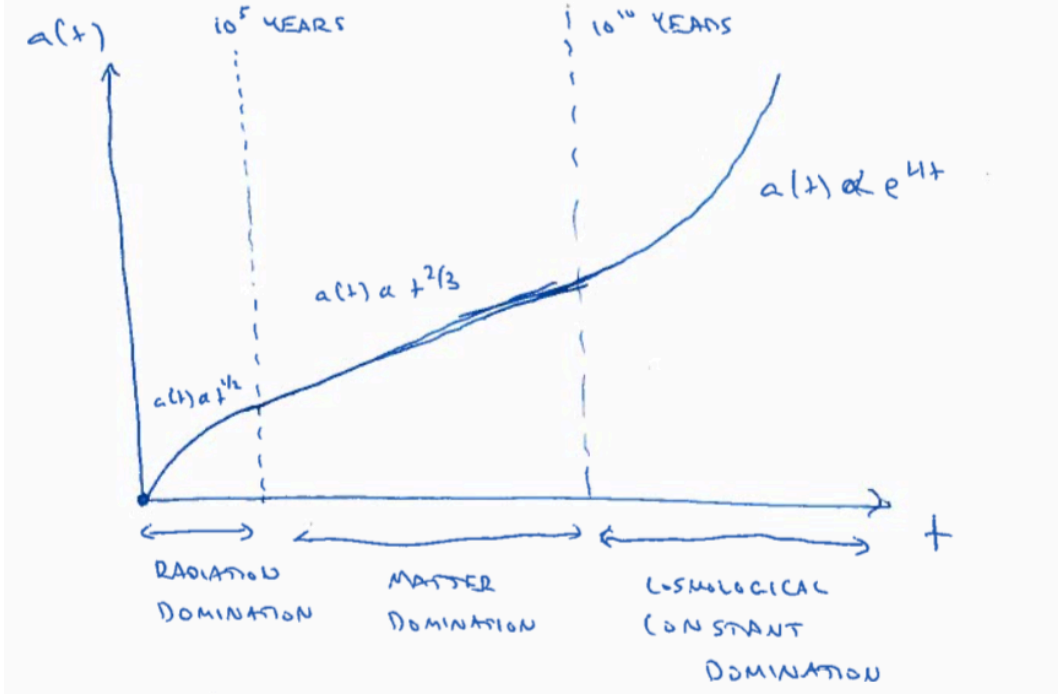 In other words, D_c is the distance measured today between two comoving observers who pass through the events.

FIG. 1: Expansion history in ΛCDM.

- If you've taken GR, it may be helpful to say that the expanding universe is described by the metric $ds^2 = -dt^2 + a(t)^2(dx^2 + dy^2 + dz^2)$. The results in this section can all be derived by doing GR calculations using this metric. (For example, Eq. (3) below for the photon redshift can be derived by specializing the geodesic equation $\ddot{x}^\mu = \Gamma^\mu_{\nu\rho}\dot{x}^\nu \dot{x}^\rho = 0$ to the case of a null geodesic.)

- Now that we have defined global coordinates (x, y, z, t), we can draw spacetime diagrams. In the left panel of Figure 2, we show the past and future lightcones, using a comoving spatial coordinate x and proper time t on the axes. (The right panel uses "conformal time" as the time coordinate, and will be defined later!)

  In this spacetime diagram, lightcones are not at $45°$ angles. Instead, the lightcone $x(t)$ obeys the differential equation $dx/dt = a(t)^{-1}$. This is equivalent to the statement that the physically measured speed of light is always equal to 1. Note that dx/dt is the *comoving* distance per proper time, so the physically measured speed (physical distance per proper time) is $a(dx/dt)$.

  The size of the past lightcone turns out to be bounded as $a \to 0$. That is, the lightcone comoving coordinate $x(t)$ approaches a finite limit as $t \to 0$. (If we assume that $a(t)$ behaves as $a \propto t^{1/2}$ as $t \to 0$ as shown in Figure 1, then this follows since the integral $\Delta x \propto \int a(t)^{-1}$ converges as $t \to 0$.) This means that our past causal horizon is finite: there are parts of the universe that have never been in causal contact with us. This may be counterintuitive, since you would think that near the big bang ($a = 0$) all points in the universe become aribtrarily close to each other, and come into causal contact.

  The size of the future lightcone also turns out to be bounded as $a \to \infty$. That is, there is a finite patch of the universe that we can ever reach, even if we travel at the speed of light. (If we assume that $a(t)$ behaves as $a \propto e^{Ht}$ as $t \to \infty$ as shown in Figure 1, then this follows since the integral $\Delta x \propto \int a(t)^{-1}$ converges as $t \to \infty$.)

- In an expanding universe, photons are "redshifted": their energy decreases with time. The physically measured energy E_{phys} measured by a comoving observer scales as:

$$E_{\mathrm{phys}} \propto a(t)^{-1} \tag{3}$$

(In lecture, I gave a plausibility argument for Eq. (3) involving wavefronts of the electric field, but I didn't reproduce it in the notes!)

- We define $\rho(t)$ to be the energy density of the universe at time t. This is a physically measured density, that is, energy per physical (not comoving) volume. In the standard cosmological model, $\rho(t)$ is the sum of contributions from baryonic matter, cold dark matter, dark energy, and radiation:

$$
\begin{aligned}
\rho(t) &= \text{energy density of universe at time } t \\
&= \rho_b(t) + \rho_c(t) + \rho_\Lambda(t) + \rho_{\text{rad}}(t)
\end{aligned} \tag{4}
$$

Today the densities are:

$$
\begin{aligned}
\rho(t_{\text{today}}) &= 4.83 \text{ keV/cm}^3 \\
\rho_b(t_{\text{today}}) &= 0.05\rho(t_{\text{today}}) \\
\rho_c(t_{\text{today}}) &= 0.27\rho(t_{\text{today}}) \\
\rho_\Lambda(t_{\text{today}}) &= 0.68\rho(t_{\text{today}}) \\
\rho_{\text{rad}}(t_{\text{today}}) &= (4.15 \times 10^{-5})\rho(t_{\text{today}})
\end{aligned}
$$

$$\tag{5}$$
$$\tag{6}$$
$$\tag{7}$$
$$\tag{8}$$
$$\tag{9}$$

When talking about expansion history, cosmologists sometimes use the term "matter" to mean "nonrelativistic particles", and "radiation" to mean "relativistic particles". In the standard cosmological model, radiation consists of photons and neutrinos. (We are assuming for simplicity that the neutrino masses are zero, so that neutrinos are guaranteed to be relativistic.) The densities evolve with time as:

$$
\begin{aligned}
\rho_b(t) &= a(t)^{-3}\rho_b(t_{\text{today}}) \\
\rho_c(t) &= a(t)^{-3}\rho_c(t_{\text{today}}) \\
\rho_\Lambda(t) &= \rho_\Lambda(t_{\text{today}}) \\
\rho_{\text{rad}}(t) &= a(t)^{-4}\rho_{\text{rad}}(t_{\text{today}})
\end{aligned}
$$

$$\tag{10}$$
$$\tag{11}$$
$$\tag{12}$$
$$\tag{13}$$

The time dependence of each component can be understood as follows. The dark energy is assumed to be a cosmological constant, which by definition means that ρ_Λ is constant in time. For nonrelativistic particles (baryons and cold dark matter), the energy density arises from particle rest masses. As the universe expands, the number of particles per comoving volume is constant, which implies that the number density (or energy density ρ) per physical volume is proportional to a^{-3}. For relativistic particles (radiation), there is an additional factor a^{-1} since the energy of each particle redshifts with time (Eq. (3)), leading to an overall time dependence of the form $\rho_{\text{rad}} \propto a^{-4}$.

Taken together, Eqs. (4)–(13) specify the energy density of each component ρ_i, and the total energy density ρ, as a function of the scale factor a. At very early times (roughly $a \lesssim 1/3000$), radiation is the dominant component of the universe, since its energy density scales as a^{-4}. Next, the universe is matter dominated until $a \sim 0.7$, when it starts to become cosmological constant dominated.

- So far, we have not explained how the time dependence of $a(t)$ shown in Figure 1 arises. The evolution of $a(t)$ is determined by *Friedmann's equation*, which relates the expansion rate to the energy density:

$$
\left(\frac{1}{a}\frac{da}{dt}\right)^2 = \frac{8\pi G}{3}\rho(t) \tag{14}
$$

As a bit of notation, the *Hubble expansion rate* $H(t)$ is defined to be the fractional expansion rate per unit time:

$$
H(t) = \frac{1}{a}\frac{da}{dt} \tag{15}
$$

and another way of writing Friedmann's equation is:

$$
H^2 = \frac{8\pi G}{3}\rho \tag{16}
$$

- The "plot" of $a(t)$ in Figure 1 can be obtained by integrating Friedmann's equation as follows. First, we write $\rho(a)$ as follows.

$$\rho(a) = \rho_0 \left(\Omega_\Lambda + \Omega_b a^{-3} + \Omega_c a^{-3} + \Omega_{\text{rad}} a^{-4} \right) \tag{17}$$

where we have introduced some standard cosmology notation: $\rho_0 = \rho(t_{\text{today}})$ is the energy density today, and $\Omega_i = \rho_i(t_{\text{today}})/\rho(t_{\text{today}})$ is the fractional density of component i today (where $i \in \{b, c, \Lambda, \text{rad}\}$).

Second, we use Friedmann's equation to write $H(a)$ as

$$H(a) = H_0 \left(\Omega_b a^{-3} + \Omega_c a^{-3} + \Omega_\Lambda + \Omega_{\text{rad}} a^{-4} \right)^{1/2} \tag{18}$$

where $H_0 = (8\pi G \rho_0/3)^{1/2}$ is the Hubble expansion rate today.

Third, we integrate to get t as a function of a.

$$\frac{dt}{da} = \frac{1}{aH(a)} \tag{19}$$

$$t = \int_0^a \frac{da'}{a' H(a')}$$

$$= H_0^{-1} \int_0^a da' \, (a')^{-1} \left(\Omega_b a^{-3} + \Omega_c a^{-3} + \Omega_\Lambda + \Omega_{\text{rad}} a^{-4} \right)^{-1/2} \tag{20}$$

This integral cannot be done analytically, but it is straightforward to integrate numerically. For example, if we integrate from $a = 0$ to $a = 1$, we get the age of the universe, which turns out to be $t = 1.38 \times 10^{10}$ years.

- Let's look at an example where Friedmann's equation can be integrated analytically. Suppose the density ρ is a power law in the scale factor a:

$$\rho(a) = \rho_0 a^{-3(1+w)} \tag{21}$$

We have written the exponent as $3(1 + w)$ to conform to some standard notation in cosmology. The parameter w defined in Eq. (21) can be interpreted as the ratio $w = p/\rho$, where p is the total pressure of the universe, and w is called the "equation of state".

The power-law cosmology we are considering is called a "constant-w cosmology". It includes the following special cases:

  - A "matter dominated" universe, where the energy density is dominated by nonrelativistic particles ($w = 0$).
  - A "radiation dominated" universe, where the energy density is dominated by relativistic particles ($w = 1/3$).
  - A cosmological constant dominated universe ($w = -1$).

Note that the ΛCDM universe is radiation dominated at early times, and cosmological constant dominated at late times.

- In the constant-w universe defined by Eq. (21), Friedmann's equation can be integrated as follows:

$$H(a) = \left(\frac{8\pi G}{3} \rho(a) \right)^{1/2}$$

$$= H_0 a^{-3(1+w)/2} \qquad \text{where } H_0 = \left(\frac{8\pi G}{3} \rho_0 \right)^{1/2} \tag{22}$$

$$\frac{dt}{da} = \frac{1}{aH} \tag{23}$$

$$t = \int \frac{da}{aH(a)}$$

$$= H_0^{-1} \int da \, a^{(1+3w)/2} \tag{24}$$

Depending on the value of w, this integral may or may not converge, in the limits $a \to 0$ and $a \to \infty$.

$$\text{(Integral in Eq. (24) converges as } a \to 0) \Leftrightarrow \text{(Universe begins in a big bang)}$$
$$\Leftrightarrow w > -1 \qquad (25)$$
$$\text{(Integral in Eq. (24) converges as } a \to \infty) \Leftrightarrow \text{(Universe ends in a "big rip")}$$
$$\Leftrightarrow w < -1 \qquad (26)$$

For ΛCDM, we have $w = 1/3$ at early times and $w = -1$ at late times, so the universe begins in a big bang singularity, but does not end in a big rip singularity (barely!)

- Continuing to assume the constant-w universe defined by Eq. (21), let's solve for the trajectory $x(t)$ of a light ray. Here, x is a comoving spatial coordinate and t is proper time. Previously, we argued that the trajectory was determined by the differential equation $dx/dt = 1/a(t)$. Therefore,

$$
\begin{aligned}
x &= \int \frac{dt}{a} \\
&= \int \frac{da}{a^2 H(a)} \\
&= H_0^{-1} \int da \, a^{(-1+3w)/2} \qquad (27)
\end{aligned}
$$

Depending on the value of w, this integral may or may not converge, in the limits $a \to 0$ and $a \to \infty$.

$$\text{(Integral in Eq. (27) converges as } a \to 0) \Leftrightarrow \text{(Past causal horizon is finite)}$$
$$\Leftrightarrow w > -\frac{1}{3} \qquad (28)$$
$$\text{(Integral in Eq. (27) converges as } a \to \infty) \Leftrightarrow \text{(Future causal horizon is finite)}$$
$$\Leftrightarrow w < -\frac{1}{3} \qquad (29)$$

For ΛCDM, we have $w = 1/3$ at early times and $w = -1$ at late times, so the past and future causal horizons are both finite (as previously stated).

- As you can see from the calculations in the last few bullet points, it is useful to switch between time coordinates t and a. There are a few more time coordinates which are useful in cosmology. First, the redshift z, which is related to a by the change of variables:

$$a = \frac{1}{1+z} \qquad (30)$$

By convention, physicists usually use a, whereas astronomers use z, but there is no real difference.

- Another useful time coordinate is *conformal time* τ, which is defined by the differential equation

$$\frac{d\tau}{dt} = \frac{1}{a(t)} \qquad (31)$$

with initial condition $\tau = 0$ at the big bang ($a = 0$). Some equivalent definitions:

$$
\begin{aligned}
\tau &= \int_0^t \frac{dt'}{a(t')} \\
&= \int_0^a \frac{da'}{a'^2 \, H(a')} \qquad (32)
\end{aligned}
$$

This time coordinate has the property that "lightcones are at 45° angles", that is, the trajectory of a light ray is $dx/d\tau = 1$.

- There are multiple notions of time and distance in an expanding universe. For example, the age of the universe (Δt) is 14 billion years old, but the size of the observable universe (Δx) is 46 billion comoving light-years. That is, if we look back to the big bang along our past lightcone, the resulting point in space would be 46 billion light-years away today. (Figure 2, left panel.)

  Because lightcones are 45° in conformal time, the conformal age of the universe ($\Delta \tau$) is the same as the size of the observable universe (Δx), i.e. 46 billion comoving light-years. (Figure 2, right panel.) In general, you can think of conformal time as "comoving distance along the lightcone".
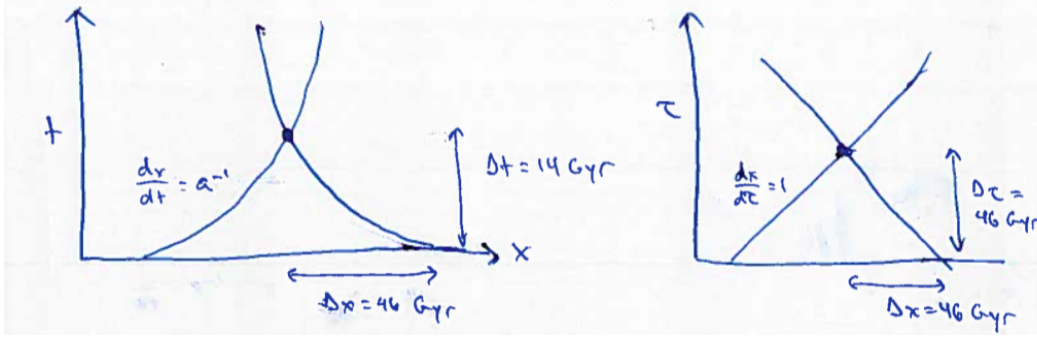

FIG. 2: FRW lightcone geometry in proper time t (left) and conformal time τ (right), and comoving spatial coordinate x.

This concludes our main discussion of the expanding universe. In the following bullet points, we mention a few additional points for completeness, in less detail (we just sketch the key equations). These points will either not be used at all in the next section, or will only be used briefly.

- So far we have analyzed the dynamics of massless particles (photons), which are fully described by the statements that the trajectory satisfies $dx/dt = a(t)^{-1}$, and the photon energy redshifts as $E \propto a^{-1}$. Let us now consider the case of massive particles. We will state the main results without derivation, since the derivation requires GR. These results will be used briefly at the end of the next section, in the context of massive neutrinos.

  A massive particle has a 4-momentum (E_{phys}, p_{phys}), where

$$E_{\mathrm{phys}} = \frac{m_0}{\sqrt{1 - v^2}} \qquad p_{\mathrm{phys}} = \frac{m_0 v}{\sqrt{1 - v^2}} \tag{33}$$

  If the particle is moving in the expanding universe with no applied force, then the momentum redshifts as $p_{\mathrm{phys}} \propto a^{-1}$. This statement is true for both massive and massless particles, and generalizes our previous statement that photon energies redshift as a^{-1}. Note that for a massive particle, the velocity eventually v becomes arbitrarily small as the universe expands.

  If an external force F acts on a massive particle, then

$$\frac{dp_{\mathrm{phys}}}{dt} = -Hp_{\mathrm{phys}} + F \tag{34}$$

  The term ($-Hp_{\mathrm{phys}}$) is a gravitational force (sometimes called "Hubble friction" or "Hubble drag") which slows down particles in the cosmic rest frame.

- A few more definitions. Let $p(t)$ be the total pressure of the universe, defined as follows.

  Operational definition: consider an ensemble of particles with physical number density n_{phys} and 4-momentum ($E_{\mathrm{phys}}, p_{\mathrm{phys}}$). (We assume that the magnitude of each particle's momentum is the same, but the directions of the momenta are uniformly distributed.) Then:

$$\rho(t) = n_{\mathrm{phys}} E_{\mathrm{phys}} \qquad p(t) = n_{\mathrm{phys}} \left(\frac{p_{\mathrm{phys}}^2}{3E_{\mathrm{phys}}} \right) \tag{35}$$

  Note that we have $p(t) \to 0$ and $p(t) \to \rho(t)/3$ in the nonrelativistic and relativistic limits respectively.

Conceptual definition: suppose we place a 3D "box" anywhere in the universe, and assume that particles bounce elastically from the sides the box. Then $p(t)$ is the force per unit area on the sides of the box.

Here are two important equations, which we haven't mentioned so far:

$$\frac{d\rho}{d\log a} = -3(\rho + p) \qquad \text{("Continuity equation")} \tag{36}$$

$$\frac{dH}{dt} = -4\pi G(\rho + p) \qquad \text{("Friedmann's second equation")} \tag{37}$$

The continuity equation generalizes the law of conservation of energy to the expanding universe, and can be integrated to give the evolution of the energy density ρ. We haven't needed the continuity equation yet, because we have been considering particle species in either the nonrelativistic or relativistic limits, where we argued directly that the energy density evolves as a^{-3} or a^{-4} respectively. However, in an intermediate case (say $v = 0.5c$), we would need the continuity equation.

Friedmann's second equation is equivalent to the continuity equation (if Friedmann's first equation is assumed), but is sometimes convenient on its own.

- So far, we have assumed that the universe is spatially flat. There is a generalization where the universe is allowed to have spatial curvature, which we now sketch. (These results won't be used in the rest of the notes, and weren't included in the lecture, except in response to a question!)

First, we consider the case of positive curvature, that is a sphere of comoving radius R_c. (The comoving radius is fixed throughout the expansion.) Then Friedmann's equation becomes:

$$H^2 + \frac{K}{a^2} = \frac{8\pi G}{3}\rho \qquad \text{where } K = 1/R_c^2 \tag{38}$$

(Derivation omitted, since it requires GR!)

The effect of positive curvature is the same as a *negative* contribution to the energy density, of the form

$$\rho_K = -\frac{3}{8\pi G}\frac{K}{a^2} \tag{39}$$

Through analysis of the CMB and other datasets, curvature is known to be small. The current observational bound is roughly:

$$|\Omega_K| = \left|\frac{\rho_K(a=1)}{\rho_{\text{tot}}(a=1)}\right| \lesssim 10^{-2} \tag{40}$$

Since $\rho_K \propto a^{-2}$, it follows that curvature is also small at all times in the expansion history, since at early times it is dominated by matter (a^{-3}) or radiation (a^{-4}), and at late times it is dominated by the cosmological constant.

The curvature K can also be negative. In this case, the spatial geometry is a hyperboloid rather than a sphere. The corresponding contribution to the energy density (Eq. (39)) is positive.

II.   THERMAL HISTORY

- Today, the cosmic microwave background is a blackbody with temperature $T = 2.726$ Kelvin. As the universe expands, the temperature decreases as $T \propto a^{-1}$. (This is not obvious, but starting from the standard formula for the blackbody spectrum, one can show that if every photon redshifts as $1/a$, the result is another blackbody spectrum with temperature T/a.)

Therefore, if we imagine looking back to the early universe, the universe will get hotter and denser at earlier times. At a time when the temperature is ≈ 3000 K, corresponding to scale factor $a = 10^{-3}$, something interesting happens. Above this temperature, there are enough high-energy CMB photons to ionize neutral hydrogen atoms into free protons and electrons. Therefore, as the universe expands and cools, it transitions from an opaque p^+e^- plasma to a gas of transparent neutral hydrogen. This is the time when the CMB is formed.

- Continuing to go back in time, another interesting event in the thermal history occurs at $T = 0.5$ MeV, when the temperature is equal to the electron mass. Above this temperature, there are enough high-energy photons to produce e^+e^- pairs via the process $2\gamma \to e^+e^-$, and the universe is a plasma of photons, positrons and electrons (plus a few protons and neutrons). Below this temperature, the positrons and electrons have mostly annihilated to photons, and their abundance is small.

  (As an aside, you might wonder why there are still electrons in the universe today. This is because there is a small asymmetry between the abundance of positrons and electrons in the initial conditions, so that some electrons remain after all the positrons have annihilated. The reason for this initial asymmetry is still an unsolved problem in cosmology!)

- Going slightly further back in the thermal history, another interesting event occurs at $T = 1$ MeV. Above this temperature, neutrinos are in thermal equilibrium with the $e^+e^-\gamma$ plasma, due to scattering processes like $\nu e^- \to \nu e^-$ or $e^+e^- \to \nu\bar{\nu}$. Below this temperature, these scattering events become infrequent enough that the neutrinos decouple from the thermal plasma and can be treated as noninteracting. This event is called "neutrino freezeout".

- The preceding events in the thermal history are shown in Figure 3. Note that there are many interesting events not shown! For example, at 100 keV, the electroweak interactions which convert protons to neutrons freeze out, and the cosmological abundance of neutrons (and other atomic nuclei, such as lithium) is determined. This stage in the thermal history is called "Big Bang Nucleosynethesis", but we won't cover it in these notes. Our goal in this section is to analyze the "neutrino freezeout" and "e^+e^- annihilation" stages, and show how they *predict a cosmic neutrino background with temperature* $T_\nu = (4/11)^{1/3}T_{\rm CMB} = 1.95$ Kelvin.

FIG. 3: Thermal history of the universe. (With some interesting steps omitted, for example Big Bang Nucleosynthesis at $T = 100$ keV!)

- As a starting point for analyzing the thermal history, consider a box of volume V, containing a multiparticle plasma at temperature T. The plasma can consist of multiple particle species. Each species can be a boson or fermion, massive or massless, and can have arbitrary spin. We will assume that all multiparticle states are in thermal equilibrium, so that the probability of every multiparticle state is given by the Boltzmann distribution $\propto e^{-E/T}$, where E is the total energy of the state (obtained by summming the energies of all particles).

- Consider a single momentum p, a single particle species with rest mass m, and a single spin state. This defines a mode, whose occupation number N will be a random number whose probability distribution is given by the Boltzmann distribution:

$$p(N) \propto e^{-NE_p/T} \qquad \text{where } E_p = \sqrt{p^2 + m^2} \tag{41}$$

We will be interested in calculating the expectation value $f_p = \langle N \rangle$. If the particle is a fermion, then N can be

either 0 or 1, and:

$$p(N) = \frac{e^{-NE_p/T}}{1 + e^{-E_p/T}} \qquad \text{(fermion assumed)} \qquad (42)$$

$$f_p = \sum_{N=0}^{1} N p(N)$$

$$= \frac{e^{-E_p/T}}{1 + e^{-E_p/T}} \qquad \text{(fermion assumed)} \qquad (43)$$

If the particle is a boson, then N can be any integer, and:

$$p(N) = \frac{e^{-NE_p/T}}{\sum_{N=0}^{\infty} e^{-NE_p/T}} \qquad \text{(boson assumed)} \qquad (44)$$

$$f_p = \sum_{N=0}^{\infty} N p(N)$$

$$= \frac{\sum_{N=0}^{\infty} N e^{-NE_p/T}}{\sum_{N=0}^{\infty} e^{-NE_p/T}}$$

$$= \frac{e^{-E_p/T}(1 - e^{-E_p/T})^{-2}}{(1 - e^{-E_p/T})}$$

$$= \frac{e^{-E_p/T}}{1 - e^{-E_p/T}} \qquad \text{(boson assumed)} \qquad (45)$$

where we have used Eqs. (A3), (A4). We see that Eqs. (43), (45) may be combined into the single expression:

$$f_p = \frac{e^{-E_p/T}}{1 \mp e^{-E_p/T}} \qquad (46)$$

where *in Eq. (46) and throughout this section, the upper sign means "boson", and the lower sign means "fermion".*

- Now we would like to compute the total energy density, by summing over all momenta, spin states, and particle species. When we sum over momenta in a box with volume V, we will replace the sum by the following "mode-counting" integral:

$$\sum_p \to V \int \frac{d^3 p}{(2\pi)^3} \qquad (47)$$

To derive this, suppose that the side lengths of the box are (L_x, L_y, L_z), with $V = L_x L_y L_z$, and the box has periodic boundary conditions. In such a box, the momentum $p = (p_x, p_y, p_z)$ of a particle must be quantized as:

$$(p_x, p_y, p_z) = \left(\frac{2\pi}{L_x}, \frac{2\pi}{L_y}, \frac{2\pi}{L_z} \right) \qquad ((n_x, n_y, n_z) \text{ integers}) \qquad (48)$$

in order for the wavefunction $\psi(x) = e^{ip \cdot x}$ to satisfy the periodic boundary conditions. Therefore, when we sum over momenta, we are summing over a discrete lattice of momenta defined by Eq. (48). A little thought shows that the number of lattice elements per volume $(d^3 p)$ is $(L_x L_y L_z)/(2\pi)^3 = V/(2\pi)^3$. Therefore, the sum over momenta can be approximated by the integral $\int d^3 p \, V/(2\pi)^3$. This completes the derivation of Eq. (47).

- Now we can compute the energy density of the plasma:

$$\rho = \frac{1}{V} \sum_{\text{species}} \sum_{\text{spins}} \sum_p E_p f_p$$

$$= \frac{1}{V} \sum_{\text{species}} g V \int \frac{d^3 p}{(2\pi)^3} E_p \frac{e^{-E_p/T}}{1 \mp e^{-E_p/T}}$$

$$= \sum_{\text{species}} \frac{g}{2\pi^2} \int dp \, p^2 E_p \frac{e^{-E_p/T}}{1 \mp e^{-E_p/T}} \qquad (49)$$

Here, g denotes the number of spin states (which can be different for each particle species). The integral in Eq. (49) gives the exact energy density ρ, for a particle with arbitrary mass and spin. For arbitrary m the integral cannot be done analytically, but we will be able to determine the leading behavior in the relativistic limit $T \gg m$ and nonrelativistic limit $T \ll m$, in the next two bullet points.

- First consider the relativistic limit $(T \gg m)$. In this case we can approximate $E_p \approx p$ in the integral:

$$
\begin{aligned}
\rho &= \frac{g}{2\pi^2} \int dp\, p^2 E_p \frac{e^{-E_p/T}}{1 \mp e^{-E_p/T}} \\
&\approx \frac{g}{2\pi^2} \int dp\, p^3 \frac{e^{-p/T}}{1 \mp e^{-p/T}} \\
&= \frac{g}{2\pi^2} \int dp\, p^3 \left(e^{-p/T} \pm e^{-2p/T} + e^{-3p/T} \pm e^{-4p/T} + \cdots \right) \\
&= \frac{g}{2\pi^3}(6T^4) \times \begin{cases} (\pi^4/90) & \text{[boson]} \\ (7/8)(\pi^4/90) & \text{[fermion]} \end{cases}
\end{aligned} \tag{50}
$$

In the last line we have used Eqs. (A1), (A6), and (A7). In the cosmology literature, the result is usually written this way:

$$
\rho \approx \frac{\pi^2}{30} g_* T^4 \qquad (T \gg m \text{ assumed}) \tag{51}
$$

where we have defined

$$
g_* = \begin{cases} g & \text{[boson]} \\ (7/8)g & \text{[fermion]} \end{cases} \tag{52}
$$

This result generalizes the usual expression $\rho = (\pi^2/15)T^4$ for the energy density of blackbody radiation (for photons, $g = g_* = 2$).

- Now consider the nonrelativistic limit $(T \ll m)$. In this case:

$$
\begin{aligned}
\rho &= \frac{g}{2\pi^2} \int dp\, p^2 E_p \frac{e^{-E_p/T}}{1 \mp e^{-E_p/T}} \\
&\approx \frac{g}{2\pi^2} \int dp\, p^2 E_p e^{-E_p/T} \\
&= \frac{g}{2\pi^2} \int_m^\infty dE\, E^2 \sqrt{E^2 - m^2}\, e^{-E/T} && \text{change variable } E = \sqrt{p^2 + m^2} \\
&= \frac{g}{2\pi^2} e^{-m/T} \int_0^\infty d\epsilon\, (m+\epsilon)^2 \epsilon^{1/2}(2m+\epsilon)^{1/2} e^{-\epsilon/T} && \text{change variable } \epsilon = E - m \\
&\approx \frac{g}{2\pi^2} e^{-m/T} \int_0^\infty d\epsilon\, m^2 \epsilon^{1/2}(2m)^{1/2} e^{-\epsilon/T} \\
&= \frac{g}{2\pi^2} e^{-m/T} m^2 (2m)^{1/2} \frac{\sqrt{\pi}}{2} T^{3/2} \\
&= gm \left(\frac{mT}{2\pi} \right)^{3/2} e^{-m/T}
\end{aligned} \tag{53}
$$

where we have used Eq. (A1) to do the integral (note that $(1/2)! = \sqrt{\pi}/2$). We see that the abundance of a massive particle is exponentially suppressed when the temperature of the plasma falls below the particle mass.

- Summarizing the previous bullet points, for a single particle species, the energy density ρ is given by:

$$
\begin{aligned}
\rho(T) &= \frac{g}{2\pi^2} \int dp\, p^2 E_p \frac{e^{-E_p/T}}{1 \mp e^{-E_p/T}} \\
\end{aligned} \tag{54}
$$

$$
= \begin{cases} \frac{\pi^2}{30} g_* T^4 & (T \gg m) \\ gm \left(\frac{mT}{2\pi} \right)^{3/2} e^{-m/T} & (T \ll m) \end{cases} \tag{55}
$$

- We will also need formulas for the pressure p, which was defined briefly near the previous section. According to Eq. (35), a particle with 4-momentum (E, p) contributes E to the total energy and $(p^2/3E)$ to the pressure. Therefore, the pressure is given by the following integral, which was obtained by replacing E_p in Eq. (54) by $p^2/(3E_p)$:

$$p(T) = \frac{g}{2\pi^2} \int dp\, p^2 \left(\frac{p^2}{3E_p} \right) \frac{e^{-E_p/T}}{1 \mp e^{-E_p/T}} \tag{56}$$

As in the case of the density ρ, the leading behavior of this integral in the relativistic and nonrelativistic limits can be calculated analytically. We omit the details and quote the final result:

$$p(T) = \begin{cases} (1/3)\rho(T) & (T \gg m) \\ (T/m)\rho(T) & (T \ll m) \end{cases} \tag{57}$$

- So far, we have not considered the expansion of the universe. We would like to know how the temperature T of the plasma evolves with the scale factor a. In principle, this can be determined by using the continuity equation, which was mentioned near the end of the last section:

$$\frac{d\rho}{d\log a} = -3(\rho + p) \tag{58}$$

Since ρ and p are functions of T, given explicitly by Eqs. (54), (56) above, the continuity equation (a single first-order differential equation) suffices to determine $T(a)$ (since there is only one dynamical degree of freedom, the temperature T).

For example, if we assume $p = \rho/3$, which is a good approximation if all species are relativistic (Eq. (57)), then the continuity equation reads $d\rho/d\log a = -4\rho$, which implies $\rho \propto a^{-4}$. If all species are relativistic, then we also know that $\rho \propto T^4$ (Eq. (55)), so this implies $T \propto a^{-1}$.

Note that most of the time, all species in the plasma are relativistic, and T evolves as $T \propto a^{-1}$. This is because when a species becomes nonrelativistic, its abundance becomes exponentially suppressed, and it can be neglected. For example, if the temperature is significantly higher than 0.5 MeV (say 2 MeV), then the plasma consists of positrons, electrons, and photons, and all species are relativistic. If the temperature is significantly lower than 0.5 MeV (say 0.1 MeV), then the plasma consists only of photons (which are always relativistic), and the e^+e^- abundance is small enough to be negligible. In both of these regimes, the temperature evolves as $T \propto a^{-1}$. During the e^+e^- annihilation event when $T \sim 0.5$ MeV, the evolution of the temperature is more complicated, but in principle it can be obtained from the continuity equation (Eq. (58)).

- In fact, the continuity equation can be integrated exactly, to give the evolution of T for all values of the scale factor a, including the behavior during annihilation events. The exact solution can be written as the following conservation law:

$$a^3 \left(\frac{\rho(T) + p(T)}{T} \right) = \text{const.} \tag{59}$$

This conservation law is nontrivial to derive, and there is more than one way of deriving it. In the next two bullet points, we will use a brute-force approach based on the continuity approach. This has the advantage of being completely rigorous, but the disadvantage that the conservation law in Eq. (59) looks like an algebraic accident.

There is also a thermodynamic derivation based on identifying the quantity $(\rho + p)/T$ as the entropy per unit physcal volume, and arguing that the expansion of the universe should be an adiabatic process which conserves entropy per comoving volume. In these notes, we omit the thermodynamic derivation!

- As a first step toward proving the conservation law in Eq. (59), we prove the following lemma:

$$\frac{\partial p}{\partial T} = \frac{\rho + p}{T} \tag{60}$$

The proof is as follows. First, we take the integral expressions for ρ and p in Eqs. (54), (56), and change variables to $E = \sqrt{p^2 + m^2}$, obtaining:

$$\rho = \frac{g}{2\pi^2} \int dE\, E^2 \sqrt{E^2 - m^2}\, f(E/T) \tag{61}$$

$$p = \frac{g}{2\pi^2} \int dE\, \left(\frac{1}{3}(E^2 - m^2)^{3/2} \right) f(E/T) \tag{62}$$

where $f(E/T) = e^{-E/T}/(1 \mp e^{-E/T})$. In this equation and others below, integrals over E run from $E = m$ to $E = \infty$. Now, starting from Eq. (62), we calculate $(\partial p/\partial T)$ as follows:

$$\begin{aligned}
\frac{\partial p}{\partial T} &= \frac{g}{2\pi^2} \int dE\, \left(\frac{1}{3}(E^2 - m^2)^{3/2} \right) \left(-\frac{E}{T^2} \right) f'(E/T) \\
&= \frac{g}{2\pi^2} \int dE\, \left(\frac{1}{3}(E^2 - m^2)^{3/2} \right) \left(-\frac{E}{T} \right) \frac{\partial}{\partial E} f(E/T)
\end{aligned} \tag{63}$$

In the last line, we have expressed the result as a partial derivative with respect to E. This allows us to integrate by parts as follows:

$$\begin{aligned}
\frac{\partial p}{\partial T} &= \frac{g}{2\pi^2} \int dE\, f(E/T) \left(-\frac{\partial}{\partial E} \right) \left(-\frac{E}{3T}(E^2 - m^2)^{3/2} \right) \\
&= \frac{g}{2\pi^2} \int dE\, f(E/T) \left(\frac{1}{3T}(E^2 - m^2)^{3/2} + \frac{E^2}{T}(E^2 - m^2)^{1/2} \right)
\end{aligned} \tag{64}$$

Comparing with Eqs. (61), (62), the result can be written:

$$\frac{\partial p}{\partial T} = \frac{p}{T} + \frac{\rho}{T} \tag{65}$$

completing the proof of the lemma (Eq. (60)).

- Now we can prove the conservation law in Eq. (59). We will show that the derivative of the left-hand side $(a^3(\rho + p)/T)$ with respect to $(\log a)$ is zero. First we write:

$$\frac{d}{d\log a} \left(a^3 \frac{(\rho + p)}{T} \right) = 3a^3 \frac{(\rho + p)}{T} + \frac{a^3}{T} \frac{d(\rho + p)}{d\log a} - a^3 \frac{\rho + p}{T^2} \frac{dT}{d\log a} \tag{66}$$

We now compute the factor $d(\rho + p)/d(\log a)$ which appears in the second term:

$$\begin{aligned}
\frac{d(\rho + p)}{d\log a} &= \frac{d\rho}{d\log a} + \frac{\partial p}{\partial T} \frac{dT}{d\log a} \\
&= -3(\rho + p) + \frac{(\rho + p)}{T} \frac{dT}{d\log a}
\end{aligned} \tag{67}$$

where we have used the continuity equation (Eq. (58)) and the lemma (Eq. (60)) in the last line. Now we plug Eq. (67) into Eq. (66). All terms cancel on the RHS and the result is:

$$\frac{d}{d\log a} \left(a^3 \frac{(\rho + p)}{T} \right) = 0 \tag{68}$$

This completes the proof of the conservation law (Eq. (59)).

- Here is alternative way of writing the conservation law in Eq. (59). We define g_{tot} (a function of time) by:

$$\rho + p = \frac{2\pi^2}{45} g_{\text{tot}} T^4 \tag{69}$$

Then the conservation law can be written as:

$$T \propto a^{-1} g_{\text{tot}}^{-1/3} \tag{70}$$

The constant $(2\pi^2/45)$ on the RHS of Eq. (69) has been chosen so that a relativistic species contributes g_* to g_{tot}. The contribution of a nonrelativistic species is exponentially suppressed. (These statements follow from Eq. (55).)

- Throughout most of the thermal history, g_{tot} is roughly constant, and Eq. (70) implies that $T \propto a^{-1}$. However, during an annihilation event, the value of g_{tot} changes.

  Consider the case of e^+e^- annihilation. Before annihilation starts, the thermal plasma consists of photons, electrons, and positrons, and all three species are relativistic. The photons contribute $g_* = 2$, since the photon is a boson with two spin states. The electrons and positrons are fermions with two spin states, and therefore each contribute $g_* = 2(7/8)$. Therefore:

$$g_{\text{tot}} = 2 + 2\left(\frac{7}{8}\right) + 2\left(\frac{7}{8}\right) = \frac{11}{2} \qquad \text{(before } e^+e^- \text{ annihilation)} \qquad (71)$$

  After e^+e^- annihilation, the photons contribute $g_* = 2$, and the contribution from electrons and positrons is exponentially suppressed. Therefore:

$$g_{\text{tot}} = 2 \qquad \text{(after } e^+e^- \text{ annihilation)} \qquad (72)$$

  Comparing the previous two equations, we see that g_{tot} decreases by a factor $(4/11)$ during e^+e^- annihilation. In order to satisfy the conservation law in Eq. (70), the photon temperature must increase by a factor $(11/4)^{1/3}$. That is, let a_1 be the value of the scale factor at a time before e^+e^- annihilation, and let a_2 be the scale factor at a time after e^+e^- annihilation. Then the photon temperatures $T_{\gamma 1}$ and $T_{\gamma 2}$ are related by:

$$T_{\gamma 2} = \left(\frac{11}{4}\right)^{1/3} \frac{a_1}{a_2} T_{\gamma 1} \qquad (73)$$

- Now that we have studied the evolution of the photon temperature, let's consider the evolution of the neutrino temperature, which is much simpler! Let a_1 be the value of the scale factor just before neutrino decoupling. At this time, the neutrinos are in thermal equilibrium with the photons, so the neutrino occupation numbers $f_1(p)$ at scale factor a_1 are given by:

$$f_1(p) = \frac{e^{-p/T_{\gamma 1}}}{1 + e^{-p/T_{\gamma 1}}} \qquad (74)$$

  where $T_{\gamma 1}$ is the photon temperature just before neutrino decoupling.

  At all subsequent times, the neutrinos are non-interacting, i.e. not in thermal equilibrium with the plasma. Instead, their momenta "freestream" as follows. Suppose that at a time after neutrino decoupling, the scale factor is a_2. Then, the momentum of each individual neutrino redshifts as $p_2 = (a_1/a_2)p_1$. Therefore, the neutrino occupation numbers $f_2(p)$ at time a_2 are given by:

$$\begin{aligned} f_2(p) &= f_1\left(\frac{a_2}{a_1}p\right) \\ &= \frac{e^{-a_2 p/(a_1 T_{\gamma 1})}}{1 - e^{-a_2 p/(a_1 T_{\gamma 1})}} \\ &= \frac{e^{-p/T_{\nu 2}}}{1 + e^{-p/T_{\nu 2}}} \end{aligned} \qquad (75)$$

  where we have defined:

$$T_{\nu 2} = \frac{a_1}{a_2} T_{\gamma 1} \qquad (76)$$

  This calculation shows that the neutrino temperature T_ν is initially equal to the photon temperature $T_{\gamma 1}$, then redshifts as $T_\nu \propto a^{-1}$ for all subsequent times. In contrast, the photon temperature redshifts as $T_\gamma \propto a^{-1} g_{\text{tot}}^{-1/3}$. The photons and neutrinos behave differently because neutrinos freestream, whereas photons are in thermal equilibrium with the plasma, so their temperature is boosted during e^+e^- annihilation.

  Now comparing Eqs. (73), (76), we see that for all times after e^+e^- annihilation, the photon and neutrino temperatures are related by:

$$T_\nu = \left(\frac{4}{11}\right)^{1/3} T_\gamma \qquad (77)$$

In particular, today the CMB temperature is $T_\gamma = 2.726$ Kelvin, so we *predict the existence of a cosmic neutrino background with temperature* $T_\nu = (4/11)^{1/3} T_{\mathrm{CMB}} = 1.95$ Kelvin.

- (This comment is included just for fun, and derivations are omitted!) So far, have assumed that neutrinos are massless. If neutrinos have a small mass, then m_ν will be negligible at neutrino freezeout, when the number density of relic neutrinos in the universe is determined. However, at later times, neutrinos will become nonrelativistic, and their total contribution to the energy density may be non-negligible. A calculation (omitted in these notes!) shows that the fraction of the total energy density due to neutrinos today is:

$$\Omega_\nu = \frac{\rho_\nu(t_{\mathrm{today}})}{\rho(t_{\mathrm{today}})} = \frac{\sum_\nu m_\nu}{43 \text{ eV}} \tag{78}$$

where $\sum_\nu m_\nu$ denotes the sum of the neutrino masses over all 3 species.

Thus, the gravitational effect of neutrinos is surprisingly large: even a mass as small as 43 eV would have an order-one effect on cosmology! Since cosmological observables (such as the CMB power spectrum) are currently measured to percent level, one might expect the current observational limits on neutrino mass to be two orders of magnitude smaller than this. This expectation is correct: the current upper limit is

$$\sum_\nu m_\nu \lesssim 120 \text{ meV} \qquad (95\% \text{ CL}) \tag{79}$$

In fact, neutrino oscillation experiments have established a *lower bound* on the sum of the neutrino masses:

$$\sum_\nu m_\nu \gtrsim 46 \text{ meV} \tag{80}$$

The cosmology upper bound (Eq. (79)) is not too far away from the particle physics lower bound (Eq. (80)), and will continue to improve. A measurement of neutrino mass may not be far away!

Appendix A: Some integrals and sums

Here are some integrals and sums which arise in the "Thermal History" section. First, the following integral:

$$\int_0^\infty x^n e^{-x} = n! \tag{A1}$$

If n is a half-integer, then $n!$ can be computed by remembering that $n! = n(n-1)!$ and $(-1/2)! = \sqrt{\pi}$. For example:

$$\begin{aligned}
\int_0^\infty x^{3/2} e^{-x} &= (3/2)! \\
&= (3/2)(1/2)(-1/2)! \\
&= \frac{3}{4}\sqrt{\pi} \tag{A2}
\end{aligned}$$

Recall the standard formula for the geometric series:

$$\sum_{n=0}^\infty x^n = \frac{1}{1-x} \tag{A3}$$

Related sums can sometimes be done by taking derivatives on both sides. For example:

$$\begin{aligned}
\sum_{n=0}^\infty n x^n &= x\frac{d}{dx}\left(\sum_{n=0}^\infty x^n\right) \\
&= x\frac{d}{dx}\left((1-x)^{-1}\right) \\
&= \frac{x}{(1-x)^2} \tag{A4}
\end{aligned}$$

The Riemann zeta function $\zeta(s) = \sum_{n=0}^{\infty} n^{-s}$ can be evaluated analytically if n is an even integer. The cases $s = 2$ and $s = 4$ are worth remembering:

$$\zeta(2) = \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots = \frac{\pi^2}{6} \tag{A5}$$

$$\zeta(4) = \frac{1}{1^4} + \frac{1}{2^4} + \frac{1}{3^4} + \cdots = \frac{\pi^4}{90} \tag{A6}$$

An alternating sum of the form $\sum_{n=0}^{\infty} (-1)^{n+1} n^{-s}$ can be done using the following trick:

$$\frac{1}{1^4} - \frac{1}{2^4} + \frac{1}{3^4} - \frac{1}{4^4} + \cdots = \left(1 - 2\frac{1}{2^4}\right)\left(\frac{1}{1^4} + \frac{1}{2^4} + \frac{1}{3^4} + \cdots\right)$$

$$= \frac{7}{8}\left(\frac{\pi^4}{90}\right) \tag{A7}$$