MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Physics Department

Physics 8.286: The Early Universe May 1, 2000
Prof. Alan Guth

QUIZ 3 SOLUTIONS

PROBLEM 1: DID YOU DO THE READING? *(25 points)*

Solution to Problem 1 written by Edward Keyes.

(a) Birkhoff's theorem

Birkhoff's theorem states that "the gravitational effect of a uniform medium external to a spherical cavity is zero." This is a theorem from general relativity, and necessary to know in order to extrapolate our Newtonian cosmology results to the whole universe: it might have been the case that the global curvature of space would have interfered with our Newtonian results. The other choices in the question were generally true statements from other areas of cosmology.

(b) Special-case cosmological models

The Einstein-de Sitter model is not, as some answered, Einstein's original, static universe with a cosmological constant. Instead, this model describes a flat ($k = 0$) universe with a critical density of ordinary matter ($\rho = \rho_c$). As we showed earlier in the class, this means that its scale factor grows as $R(t) \propto t^{2/3}$.

The Milne model describes an empty universe: it is open ($k = -1$) and has no matter or radiation in it ($\rho = 0$). Its scale factor grows linearly with time, since there's no matter to slow down the Hubble expansion. (One normally includes "test" particles in the description of the Milne universe, so that we can talk about their motion. But the mass of these test particles is taken to be arbitrarily small, so we completely ignore any gravitational field that they might produce.)

As an interesting aside, we might ask why the Milne model has $k = -1$. Since there is no matter, there shouldn't be any general relativity effects, and so we would ordinarily expect that the metric should be the normal, flat, Minkowski special relativity metric. Why is this space hyperbolic instead?

The answer is an illustration of the subtleties that can arise in changing coordinate systems. In fact, the metric of the Milne universe can be viewed as ***either*** a flat, Minkowski metric, ***or*** as the negatively curved metric of an open universe, depending on what coordinate system one uses. If one uses coordinates for time and space as they would be measured by a single inertial observer, then one finds a Minkowski metric; in this way of describing the model, it is clear that special relativity is sufficient, and general relativity plays no role. In this coordinate system all the test particles start at the origin at time $t = 0$, and they move outward from the origin at speeds ranging from zero, up to (but not including) the speed of light.

On the other hand, we can describe the same universe in a way that treats all the test particles on an equal footing. In this description we define time not as it would be measured by a single observer, but instead we define the time at each location as the time that would be measured by observers riding with the test particles at that location. This definition is what we have been calling "cosmic time" in our description of cosmology. One can also introduce a comoving spatial coordinate system that expands with the motion of the particles. With a particular definition of these spatial coordinates, one can show that the metric is precisely that of an open Robertson-Walker universe with $R(t) = t$.

The derivation is left as an exercise for the curious student. You should find that the normal special-relativistic time dilation and Lorentz contraction formulas, when applied to the velocities of a Hubble expansion to construct the comoving coordinate system, introduce the negative curvature to the metric.

(c) Neutron-proton ratio

In the early universe, neutrons and protons first formed when the temperature dropped far enough to keep them from being torn apart into their constituent quarks. This happened around a microsecond after the Big Bang, and at this time there were roughly equal numbers of neutrons and protons.

In fact, neutrons could be converted into protons and vice-versa in several weak reactions with electrons, positrons, and neutrinos:

$$n + \nu_e \longleftrightarrow p + e^- \quad , \quad n + e^+ \longleftrightarrow p + \bar{\nu}_e \quad , \quad n \longleftrightarrow p + e^- + \bar{\nu}_e$$

Since these are weak-force reactions, though, their rates are strongly dependent on the temperature. Once T drops below 10^{10} K, the neutrinos stop interacting with matter, and these reactions freeze, except for the forward direction of the third reaction, which describes free neutron decay (this process has a half-life of 15 minutes, so it doesn't affect things very much).

Before the freeze-out, which essentially fixes the neutron/proton ratio, the reaction rates shift as the temperature changes. The neutron is 1.3 MeV heavier than the proton, while the mass/energy of an electron is only 0.5 MeV. This means that the conversion of a neutron to a proton and electron is energetically favorable, while the reverse process costs energy. As the temperature drops so that kT is of the order of 1 MeV, these energy differences become significant compared to the available free thermal energy, and the reaction rates shift so that thermal equilibrium favors protons over neutron by an increasing margin.

When the weak reactions freeze out, this unequal ratio of neutrons and protons is preserved. Since essentially all of the neutrons end up in helium after nucleosynthesis, this also fixes the ratio of hydrogen to helium formed by the Big Bang.

(d) The dipole anisotropy

When we look at the temperature of the cosmic microwave background radiation, to first order it appears uniform across the sky. When we look closer, though, we see that it is hotter in one direction and smoothly shades into cooler in the opposite direction, at a level of about one part in 1000. This is the dipole anisotropy.

The explanation is quite simple: the Earth is not at rest with respect to the cosmic background radiation. The motion of our Sun around the center of the Galaxy, and the motion of our Galaxy towards the Virgo Cluster, etc., all give us a net velocity of around 600 km/sec, which causes us to see blueshifted CMB photons in one direction, and redshifted ones in the opposite direction. As we learned earlier in the class, a redshifted blackbody spectrum just shifts its temperature, so we see the effects of this motion as a smooth temperature variation across the sky.

(e) Events in the early universe

The correct order is:

(iii) Quark confinement, at $t \sim 10^{-6}$ sec.

(v) Muon annihilation, at $t \sim 10^{-4}$ sec.

(ii) Decoupling of electron neutrinos, at $t \sim 1$ sec.

(i) Primordial nucleosynthesis, at $t \sim 10^{2}$ sec.

(iv) Recombination, at $t \sim 10^{5}$ years.

A surprising number of students did not realize that recombination is the final stage of the early universe. After this event takes place, the universe is transparent to photons and the temperature has dropped to just a few thousand K. Nothing interesting happens after this until the processes of structure formation begins.

PROBLEM 2: FREEZE-OUT OF MUONS *(35 points)*

(a) The factors contributing to g from the muons are the following:

$$2 \quad \text{since there are two particles, the muon and the antimuon}$$

$$2 \quad \text{since there are two spin states for each particle}$$

$$\frac{7}{8} \quad \text{since the } \mu^- \text{ and the } \mu^+ \text{ are fermions}$$

Thus

$$\boxed{g_{\mu^+\mu^-} = 2 \times 2 \times \frac{7}{8} = \frac{7}{2}\ .}$$

(b) Besides the muons, the particles in thermal equilibrium when kT is just above 106 MeV are photons, neutrinos, and electron-positron pairs. As found in class

$$g_\gamma = 2 \qquad \text{(bosons, 2 spin states)}$$

$$g_\nu = \underbrace{3}_{\text{No. of species}} \times \underbrace{2}_{\text{Particle/antiparticle}} \times \underbrace{\frac{7}{8}}_{\text{Fermion factor}} = \frac{21}{4} \ .$$

$$g_{e^+e^-} = \underbrace{2}_{\text{Particle/antiparticle}} \times \underbrace{2}_{\text{Spin states}} \times \underbrace{\frac{7}{8}}_{\text{Fermion factor}} = \frac{7}{2} \ .$$

So, for kT just above 106 MeV, g is the sum of all of these contributions:

$$\boxed{g = g_{\mu^+\mu^-} + g_\gamma + g_\nu + g_{e^+e^-} = \frac{57}{4} = 14.25 \ .}$$

(c) We know that entropy is to a high degree of accuracy conserved as the universe expands, although of course it is thinned by the expansion. The conservation of entropy means that the entropy contained within any comoving volume does not change. The entropy per comoving volume S is therefore constant, and can be written as

$$S = R^3(t)\, s(t) \ ,$$

where $s(t)$ is the entropy density. The expression for the entropy density of black body radiation is given on the cover of the exam:

$$s = g\frac{2\pi^2}{45}\frac{k^4T^3}{(\hbar c)^3} \quad .$$

This formula describes the radiation of effectively massless particles, so it is not be valid when $kT \approx 106$ MeV, since for such temperatures the mass of the muons cannot be neglected. We can apply this formula, however, when $kT \gg 106$ MeV, when the muons are effectively massless, and we can also apply it when $kT \ll 106$ MeV, when the muons are essentially nonexistent. So let t_1 denote a time when kT is well ***above*** 106 MeV (but below the threshold for producing other particles), and let t_2 denote a time when kT is well ***below*** 106 MeV (but well above 0.5 MeV, when the electron-positron pairs will disappear). (Realistically the freeze-out of the muons will overlap the freezing out of the pions, with mass/energies of 135-140 MeV, but for the purpose of this problem we are ignoring the pions.)

The entropy per comoving volume is conserved, and by combining the two formulas above it can be written

$$S = C \times g(T) R^3 T^3 \quad \text{where} \quad C = \text{constant.}$$

Since $S(t_1) = S(t_2)$, we find

$$C \times g(t_1) \big[R(t_1) T(t_1) \big]^3 = C \times g(t_2) \big[R(t_2) T(t_2) \big]^3 ,$$

which implies that

$$\frac{(RT)|_{t_2}}{(RT)|_{t_1}} = \left[\frac{g(t_1)}{g(t_2)} \right]^{1/3} .$$

We found $g(t_1)$ in part (b): $g(t_1) = 14.25$. After the muons disappear from the black body radiation they no longer contribute to the g in the expression for the entropy, so

$$g(t_2) = g_\gamma + g_\nu + g_{e^+e^-} = 2 + \frac{21}{4} + \frac{7}{2} = \frac{43}{4} = 10.75 .$$

Using these values in the expression above we obtain the increase in RT,

$$(RT)|_{t_2} = \left(\frac{14.25}{10.75} \right)^{1/3} (RT)|_{t_1} = \boxed{ \left(\frac{57}{43} \right)^{1/3} (RT)|_{t_1} . }$$

Numerically,

$$(RT)|_{t_2} \approx (1.10)\, (RT)|_{t_1} .$$

PROBLEM 3: GEODESICS IN A CLOSED UNIVERSE (*40 points + 5 points extra credit*)

(a) *(7 points)* For purely radial motion, $d\theta = d\phi = 0$, so the line element reduces do

$$-c^2\, d\boldsymbol{\tau}^2 = -c^2\, dt^2 + R^2(t) \left\{ \frac{dr^2}{1 - r^2} \right\} .$$

Dividing by dt^2,

$$-c^2 \left(\frac{d\boldsymbol{\tau}}{dt} \right)^2 = -c^2 + \frac{R^2(t)}{1 - r^2} \left(\frac{dr}{dt} \right)^2 .$$

Rearranging,

$$\boxed{\frac{d\boldsymbol{\tau}}{dt} = \sqrt{1 - \frac{R^2(t)}{c^2(1-r^2)}\left(\frac{dr}{dt}\right)^2}\ .}$$

(b) *(3 points)*

$$\frac{dt}{d\boldsymbol{\tau}} = \frac{1}{\frac{d\boldsymbol{\tau}}{dt}} = \boxed{\frac{1}{\sqrt{1 - \frac{R^2(t)}{c^2(1-r^2)}\left(\frac{dr}{dt}\right)^2}}\ .}$$

(c) *(10 points)* During any interval of clock time dt, the proper time that would be measured by a clock moving with the object is given by $d\boldsymbol{\tau}$, as given by the metric. Using the answer from part (a),

$$d\boldsymbol{\tau} = \frac{d\boldsymbol{\tau}}{dt}\,dt = \sqrt{1 - \frac{R^2(t)}{c^2(1-r_p^2)}\left(\frac{dr_p}{dt}\right)^2}\,dt\ .$$

Integrating to find the total proper time,

$$\boxed{\boldsymbol{\tau} = \int_{t_1}^{t_2} \sqrt{1 - \frac{R^2(t)}{c^2(1-r_p^2)}\left(\frac{dr_p}{dt}\right)^2}\,dt\ .}$$

(d) *(10 points)* The physical distance $d\ell$ that the object moves during a given time interval is related to the coordinate distance dr by the spatial part of the metric:

$$d\ell^2 = ds^2 = R^2(t)\left\{\frac{dr^2}{1-r^2}\right\} \quad \Longrightarrow \quad d\ell = \frac{R(t)}{\sqrt{1-r^2}}\,dr\ .$$

Thus

$$\boxed{v_{\rm phys} = \frac{d\ell}{dt} = \frac{R(t)}{\sqrt{1-r^2}}\,\frac{dr}{dt}\ .}$$

Discussion: A common mistake was to include $-c^2\,dt^2$ in the expression for $d\ell^2$. To understand why this is not correct, we should think about how an observer would measure $d\ell$, the distance to be used in calculating the velocity of a passing object. The observer would place a meter stick along the path of the object, and she would mark off the position of the object at the beginning and end of a time interval dt_{meas}. Then she would read the distance by subtracting the two readings on the meter stick. This subtraction is equal to the physical distance between the two marks, measured at the ***same*** time t. Thus, when we compute the distance between the two marks, we set $dt = 0$. To compute the speed she would then divide the distance by dt_{meas}, which is nonzero.

(e) *(10 points)* We start with the standard formula for a geodesic, as written on the front of the exam:

$$\frac{d}{d\boldsymbol{\tau}}\left\{g_{\mu\nu}\frac{dx^\nu}{d\boldsymbol{\tau}}\right\} = \frac{1}{2}\left(\partial_\mu g_{\lambda\sigma}\right)\frac{dx^\lambda}{d\boldsymbol{\tau}}\frac{dx^\sigma}{d\boldsymbol{\tau}} .$$

This formula is true for each possible value of μ, while the Einstein summation convention implies that the indices ν, λ, and σ are summed. We are trying to derive the equation for r, so we set $\mu = r$. Since the metric is diagonal, the only contribution on the left-hand side will be $\nu = r$. On the right-hand side, the diagonal nature of the metric implies that nonzero contributions arise only when $\lambda = \sigma$. The term will vanish unless $dx^\lambda/d\boldsymbol{\tau}$ is nonzero, so λ must be either r or t (i.e., there is no motion in the θ or ϕ directions). However, the right-hand side is proportional to

$$\frac{\partial g_{\lambda\sigma}}{\partial r} .$$

Since $g_{tt} = -c^2$, the derivative with respect to r will vanish. Thus, the only nonzero contribution on the right-hand side arises from $\lambda = \sigma = r$. Using

$$g_{rr} = \frac{R^2(t)}{1-r^2} ,$$

the geodesic equation becomes

$$\frac{d}{d\boldsymbol{\tau}}\left\{g_{rr}\frac{dr}{d\boldsymbol{\tau}}\right\} = \frac{1}{2}\left(\partial_r g_{rr}\right)\frac{dr}{d\boldsymbol{\tau}}\frac{dr}{d\boldsymbol{\tau}} ,$$

or

$$\frac{d}{d\boldsymbol{\tau}}\left\{\frac{R^2}{1-r^2}\frac{dr}{d\boldsymbol{\tau}}\right\} = \frac{1}{2}\left[\partial_r\left(\frac{R^2}{1-r^2}\right)\right]\frac{dr}{d\boldsymbol{\tau}}\frac{dr}{d\boldsymbol{\tau}} ,$$

or finally

$$\boxed{\frac{d}{d\boldsymbol{\tau}}\left\{\frac{R^2}{1-r^2}\frac{dr}{d\boldsymbol{\tau}}\right\} = R^2\frac{r}{(1-r^2)^2}\left(\frac{dr}{d\boldsymbol{\tau}}\right)^2 .}$$

This matches the form shown in the question, with

$$A = \frac{R^2}{1-r^2} \ , \ \text{and } C = R^2 \frac{r}{(1-r^2)^2} \ ,$$

with $B = D = E = 0$.

(f) *(5 points EXTRA CREDIT)* The algebra here can get messy, but it is not too bad if one does the calculation in an efficient way. One good way to start is to simplify the expression for p. Using the answer from (d),

$$p = \frac{m v_{\text{phys}}}{\sqrt{1 - \frac{v_{\text{phys}}^2}{c^2}}} = \frac{m \frac{R(t)}{\sqrt{1-r^2}} \frac{dr}{dt}}{\sqrt{1 - \frac{R^2}{c^2(1-r^2)} \left(\frac{dr}{dt}\right)^2}} \ .$$

Using the answer from (b), this simplifies to

$$p = m \frac{R(t)}{\sqrt{1-r^2}} \frac{dr}{dt} \frac{dt}{d\boldsymbol{\tau}} = m \frac{R(t)}{\sqrt{1-r^2}} \frac{dr}{d\boldsymbol{\tau}} \ .$$

Multiply the geodesic equation by m, and then use the above result to rewrite it as

$$\frac{d}{d\boldsymbol{\tau}} \left\{ \frac{Rp}{\sqrt{1-r^2}} \right\} = mR^2 \frac{r}{(1-r^2)^2} \left(\frac{dr}{d\boldsymbol{\tau}} \right)^2 \ .$$

Expanding the left-hand side,

$$\begin{aligned} LHS = \frac{d}{d\boldsymbol{\tau}} \left\{ \frac{Rp}{\sqrt{1-r^2}} \right\} &= \frac{1}{\sqrt{1-r^2}} \frac{d}{d\boldsymbol{\tau}} \{Rp\} + Rp \frac{r}{(1-r^2)^{3/2}} \frac{dr}{d\boldsymbol{\tau}} \\ &= \frac{1}{\sqrt{1-r^2}} \frac{d}{d\boldsymbol{\tau}} \{Rp\} + mR^2 \frac{r}{(1-r^2)^2} \left(\frac{dr}{d\boldsymbol{\tau}} \right)^2 \ . \end{aligned}$$

Inserting this expression back into left-hand side of the original equation, one sees that the second term cancels the expression on the right-hand side, leaving

$$\frac{1}{\sqrt{1-r^2}} \frac{d}{d\boldsymbol{\tau}} \{Rp\} = 0 \ .$$

Multiplying by $\sqrt{1-r^2}$, one has the desired result:

$$\frac{d}{d\boldsymbol{\tau}} \{Rp\} = 0 \quad \Longrightarrow \quad \boxed{p \propto \frac{1}{R(t)} \ .}$$