MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Physics Department

Physics 8.286: The Early Universe — November 21, 2011
Prof. Alan Guth

PROBLEM SET 8
Revised Version*

DUE DATE: Tuesday, November 29, 2011. The extra credit problems can be handed in any time before 5 pm on Friday, December 9. (There will also be a Problem Set 9, due Tuesday, December 6, 2011.)

READING ASSIGNMENT: Barbara Ryden, *Introduction to Cosmology*, Chapter 9 *(The Cosmic Microwave Background)*, except that you can skip Section 9.3. Also read "Inflation and the New Era of High-Precision Cosmology," by Alan Guth, written for the MIT Physics Department annual newsletter, 2002. It is available at

http://web.mit.edu/physics/news/physicsatmit/physicsatmit_02_cosmology.pdf

The data quoted in the article about the nonuniformities of the cosmic microwave background radiation has since been superceded by much better data; the conclusions have only gotten stronger.

UPCOMING QUIZ: Thursday, December 8, 2011.

PROBLEM 1: THE HORIZON PROBLEM *(8 points)*

The success of the big bang predictions for the abundances of the light elements suggests that the universe was already in thermal equilibrium at one second after the big bang. At this time, the region which later evolves to become the observed universe was many horizon distances across. Try to estimate how many. You may assume that the universe is flat.

PROBLEM 2: THE FLATNESS PROBLEM *(7 points)*

Although we now know that $\Omega_0 = 1$ to an accuracy of a few percent, let us pretend that the value of Ω today is 0.1. It nonetheless follows that at 10^{-37} second after the big bang (about the time of the grand unified theory phase transition), Ω must have been extraordinarily close to one. Writing $\Omega = 1 - \delta$, estimate the value of δ at $t = 10^{-37}$ sec (using the standard cosmological model).

* Revised December 3, 2011. Minor changes to Problem 3 to improve clarity. The preamble to Problem 5 was revised and enlarged slightly, and the formula for the probability that a system is in a particular microstate was corrected to read $\exp(-E/kT)\exp(\sum_i \mu_i Q_i/kT)$. There were minor changes to the remainder of Problem 5.

PROBLEM 3: THE GREISEN-ZATSEPIN-KUZMIN (GZK) CUTOFF, *(10 points)*

Very shortly after the CMB was discovered, it was pointed out* that the existence of the radiation would impose a cutoff on very high energy cosmic rays. Protons with an energy above about 6×10^{19} eV would have a high cross section for scattering off the photons of the CMB, limiting the range that they could travel to something like 50 Mpc. Since there are no known sources within this distance, there is a prediction that we should not see cosmic rays higher than this energy.

(a) Using the formulas for the energy density and number density of black-body radiation, calculate the average energy of a photon $\bar{E}_\gamma$ for radiation with an arbitrary temperature T. Your answer should be in the form of a dimensionless number times kT. For $T = 2.725$ K, the temperature of the CMB, what is this energy, in MeV?

(b) The cross section for proton-photon scattering has a strong enhancement when the particles have just enough energy to create a very short-lived particle called the $\Delta(1232)$, which has a rest energy of 1232 MeV. The Δ then decays immediately (in about 10^{-23} second) to a proton and π^0 particle, or a neutron and a π^+ particle:

$$p + \gamma \longrightarrow \Delta \begin{cases} p + \pi^0 \\ n + \pi^+ \end{cases}$$

Suppose that photons with an energy E_γ of 3 times the mean are plentiful enough to scatter the cosmic ray protons. What energy E_p must the proton have so that it is possible to create a $\Delta(1232)$ when it collides head-on with a photon of energy $E_\gamma = 3\bar{E}_\gamma$? The mass of the proton is given by $m_p c^2 = 938.27$ MeV.

[Hint: one cannot expect that $E_p + E_\gamma = 1232$ MeV, since the conservation of momentum implies that the final Δ must have nonzero momentum, and hence nonzero kinetic energy. One could solve the conservation of energy and momentum equations simultaneously, but it is easiest to remember that the square of the energy-momentum four-vector is Lorentz-invariant:

$$p^\mu = \left(\frac{E}{c}, \vec{p}\right) \quad \Longrightarrow \quad p^2 = |\vec{p}|^2 - \frac{E^2}{c^2} = \text{Lorentz-invariant quantity.}$$

By applying this Lorentz-invariance to the total energy-momentum vector, you can deduce that

$$|\vec{p}_{\rm tot}|^2 - \frac{E_{\rm tot}^2}{c^2} = -\frac{E_{\rm rest}^2}{c^2} \ ,$$

* K. Greisen, *Phys. Rev. Lett.* **16,** 748 (1966); G.T. Zatsepin and V.A. Kuzmin, *Pis'ma Zh. Eksp. Teor. Fiz.* **4,** 114 (1966) [*JETP Lett.* **4,** 78 (1966)].

where $E_{\rm rest}$ is the total energy in the rest frame of the system. When there is just enough energy to produce a Δ particle, the energy in the rest frame must be 1232 MeV. In doing the calculation, you may use the fact that $m_p c^2 \gg E_\gamma$, and that $E_p \gg m_p c^2$.]

PROBLEM 4: BIG BANG NUCLEOSYNTHESIS *(8 points)*

The calculations of big bang nucleosynthesis depend on a large number of measured parameters. Below you are asked to qualitatively describe the effects of changing some of these parameters. Include a sentence or two to explain each of your answers.

(a) Suppose an extra neutrino species is added to the calculation. Would the predicted helium abundance go up or down?

(b) Suppose the weak interactions were stronger than they actually are, so that the thermal equilibrium distribution between neutrons and protons were maintained until $kT \approx 0.25$ MeV. Would the predicted helium abundance be larger or smaller than in the standard model?

(c) Suppose the proton-neutron mass difference were larger than the actual value of 1.29 MeV/c^2. Would the predicted helium abundance be larger or smaller than in the standard calculation?

(d) The standard theory of big bang nucleosynthesis assumes that the matter in the universe was distributed homogeneously during the era of nucleosynthesis, but the alternative possibility of inhomogeneous big-bang nucleosynthesis has been discussed since the 1980s. Inhomogeneous nucleosynthesis hinges on the hypothesis that baryons became clumped during a phase transition at $t \approx 10^{-6}$ second, when the hot quark soup converted to a gas of mainly protons, neutrons, and in the early stages, pions. The baryons would then be concentrated in small nuggets, with a comparatively low density outside of these nuggets. After the phase transition but before nucleosynthesis, the neutrons would have the opportunity to diffuse away from these nuggets, becoming more or less uniformly distributed in space. The protons, however, since they are charged, interact electromagnetically with the plasma that fills the universe, and therefore have a much shorter mean free path than the neutrons. Most of the protons, therefore, remain concentrated in the nuggets. Does this scenario result in an increase or a decrease in the expected helium abundance?

PROBLEM 5: THE DEUTERIUM BOTTLENECK *(10 points)*

The "deuterium bottleneck" plays a major role in the description of big bang nucleosynthesis: all of the nuclear reactions involved in nucleosynthesis depend on deuterium forming at the start, but deuterium does not become stable until the temperature reaches a rather low value. In this problem we will explore the statistical mechanics of the deuterium bottleneck.

A dilute ideal gas of classical nonrelativistic particles of type X, in thermal equilibrium, has a number density given by

$$n_X = g_X \left(\frac{m_X kT}{2\pi\hbar^2}\right)^{3/2} \exp\left(-\frac{m_X c^2}{kT}\right) \exp\left(\frac{\mu_X}{kT}\right) , \qquad \text{(P5.1)}$$

where $\hbar = h/2\pi$, c, and k have their usual meanings: Planck's constant, the speed of light, and the Boltzmann constant. Here g_X is the number of spin degrees of freedom associated with the particle (like the factor $g = 2$ that we encountered with photons), m_X is the mass of the particle, T is the temperature, and μ_X is the chemical potential of the particle. Because the gas is dilute there are no corrections associated with fermions and the Pauli exclusion principle; the diluteness implies that the probability for any two particles to occupy the same quantum state is negligible in any case, so it does not matter whether such states are allowed. The dilute gas approximation is valid provided that

$$n_X \ll \left(\frac{m_X kT}{2\pi\hbar^2}\right)^{3/2} . \qquad \text{(P5.2)}$$

You may or may not be familiar with chemical potential, but it will suffice for you to know that it is a concept introduced to treat quantities that are conserved or at least effectively conserved over the time scales of interest. Examples of such conserved quantities include electric charge, baryon number, electron number, muon number, and tau number. Such conserved quantities can have any value in thermal equilibrium, since the value is determined by the initial conditions and cannot be changed. The chemical potential is related to the value of such conserved quantities in a way that is closely analogous to the way that temperature is related to the energy of the system. (Energy is of course also a conserved quantity.) Temperature does not determine the energy of a system, but a high temperature indicates a high propensity for high energy states. Similarly the chemical potential for a conserved quantity Q indicates a high propensity for the value of the quantity to be large.

For each such conserved quantity Q_i one introduces a chemical potential μ_i. The chemical potential of particle X is just the sum of the chemical potentials of whatever conserved quantities the particle might possess:

$$\mu_X = \sum_i \mu_i q_i^X , \qquad \text{(P5.3)}$$

where q_i^X is the amount of quantity Q_i contained in one particle of type X. In the grand canonical ensemble, which gives the probability distribution that leads to Eq. (P5.1), each possible state for the system as a whole is assigned a probability proportional to $\exp(-E/kT)\exp(\sum_i \mu_i Q_i/kT)$, where E is the energy of the state and Q_i is the amount of conserved quantity i in the state. In this description neither the energy nor the values of the conserved charges Q_i are fixed, but for a large system the probability distributions for the energy and the conserved quantities Q_i become narrowly peaked about a central value. Note that Eq. (P5.3) implies that for any allowed reaction, such as

$$A + B \longleftrightarrow C \ , \tag{P5.4}$$

we are guaranteed that

$$\mu_A + \mu_B = \mu_C \ , \tag{P5.5}$$

since the conserved quantities must balance on the two sides of the equation.

(a) I mentioned in lecture that our textbook writes Eq. (P5.1) incorrectly, omitting the chemical potential factor. See for example Eqs. (10.11) and (10.12). The author does, however, have a footnote about this (p. 156), which concludes that "in most cosmological contexts, as it turns out, the chemical potential is small enough to be safely neglected." We can check this statement by using the author's formula to calculate the proton density at 3 minutes into the big bang, at the time of Steven Weinberg's Fifth Frame, from chapter 5 of *The First Three Minutes*. At that time the temperature was $T = 10^9$ K. To compare with the right answer, we make use of the fact that the ratio of the number density n_b of baryons to the number density n_γ of photons is estimated from WMAP data* as

$$\eta \equiv \frac{n_b}{n_\gamma} = (6.1 \pm 0.2) \times 10^{-10} \ . \tag{P5.6}$$

According to Weinberg, at that time 14% of the baryons were neutrons, with 86% protons. At the risk of appearing impertinent toward the author (but physicists are known for their impertinence), I will phrase the question this way: By how many kilo-orders of magnitude is the author's formula for n_p in error?† (Be prepared to have your calculators overflow — if they do, calculate the logarithm of the answer.)

(b) For deuterium production, the relevant reaction is

$$n + p \longleftrightarrow D \ , \tag{P5.7}$$

* D.N. Spergel et al., "Wilkinson Microwave Anisotropy Probe (WMAP) three year results: implications for cosmology," *Astrophys. J. Suppl.* **170,** 377 (2007), also available at http://arxiv.org/abs/astro-ph/0603449. They actually write it as $6.116^{+0.197}_{-0.249} \times 10^{-10}$, but I don't think that we have any need for the extra digits.

† I have exchanged email with Ryden about this, and she said she would fix it in the next edition.

so Eq. (P5.5) tells us that $\mu_n + \mu_p = \mu_D$. This equality implies that if we form the ratio

$$\frac{n_D}{n_p \, n_n} \ , \tag{P5.8}$$

expressing each number density as in Eq. (P5.1), then the chemical potential factors will cancel out. (This is how the formula is normally used, and this is how Ryden uses it on p. 180. From here on her treatment is correct, but we will proceed with slightly more detail.) To describe the bookkeeping for the reaction of Eq. (P5.7), we need to define our variables. I am using n_n, n_p, and n_D to mean the number densities of free neutrons, free protons, and deuterium nuclei. n_b denotes the total baryon number density, so

$$n_b = n_n + n_p + 2n_D \ . \tag{P5.9}$$

Finally, I will use $n_n^{\rm TOT}$ and $n_p^{\rm TOT}$ to denote the total number densities of neutrons and protons respectively, whether free or bound inside deuterium. We assume that deuterium production happens fast enough so that there is no further change in the neutron-proton balance while deuterium is forming, so the ratio

$$f \equiv \frac{n_n^{\rm TOT}}{n_b} \tag{P5.10}$$

is fixed. We will describe the extent to which the reaction has proceeded by specifying the fraction x of neutrons that remain free,

$$x \equiv \frac{n_n}{n_n^{\rm TOT}} \ . \tag{P5.11}$$

Using these definitions, write the equation that equates the ratio $n_D/(n_p \, n_n)$ to a function of temperature, using Eq. (P5.1) for each of the number densities. The deuteron is spin-1, with $g = 3$, and the proton and neutron are each spin-$\frac{1}{2}$, with $g = 2$. Except in the exponential factor, you may approximate $m_n = m_p = m_D/2$. Manipulate this formula so that it has the form

$$F(\eta, f, x) = G(T) \ , \tag{P5.12}$$

where F and G are functions that you must determine. You will need the binding energy of deuterium,

$$B = (m_p + m_n - m_D)c^2 \approx 2.22 \text{ MeV}. \tag{P5.13}$$

Eq. (P5.12) determines x as a function of T, or vice versa, but we will not try to write the function explicitly in either case.

(c) Using your result in part (b), and taking $f = 0.14$ from Weinberg's book, find the value of x, the fraction of neutrons that have been bound in deuterium, at the time of the Fifth Frame, when $T = 10^9\,\mathrm{K}$. You will probably want to solve the equation numerically. Two significant figures will be sufficient.

(d) Again using your result from part (b), and assuming that $f = 0.14$ is still accurate, find the temperature at which $x = \frac{1}{2}$, i.e., the temperature for which half of the neutrons have become combined into deuterium. Again you will presumably find the answer numerically, and 2 significant figures will be sufficient. What is the value of kT at this temperature? Qualitatively, what feature of the calculation causes this number to be small compared to B?

PROBLEM 6: A ZERO MASS DENSITY UNIVERSE— GENERAL RELATIVITY DESCRIPTION

(This problem is not required, but can be done for 7 points extra credit.)

In this problem and the next we will explore the connections between special relativity and the standard cosmological model which we have been discussing. Although we have not studied general relativity in detail, the description of the cosmological model that we have been using is precisely that of general relativity. In the limit of zero mass density the effects of gravity will become negligible, and the formulas must then be compatible with the special relativity which we discussed at the beginning of the course. The goal of these two problems is to see exactly how this happens.

These two problems will emphasize the notion that a coordinate system is nothing more than an arbitrary system of designating points in spacetime. A physical object might therefore look very different in two different coordinate systems, but the answer to any well-defined physical question must turn out the same regardless of which coordinate system is used in the calculation.

From the general relativity point of view, the model universe is described by the Robertson-Walker spacetime metric:

$$ds^2_{\mathrm{ST}} = -c^2 dt^2 + a^2(t)\left\{\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta\, d\phi^2\right)\right\} .$$

I have included the subscript "ST" to remind us that this formula gives the full spacetime metric, as opposed to the purely spatial metric which we discussed earlier. This formula describes the analogue of the "invariant interval" of special relativity, measured between the spacetime points (t, r, θ, ϕ) and $(t+dt, r+dr, \theta+d\theta, \phi+d\phi)$.

The evolution of the model universe is governed by the general relation

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3}G\rho - \frac{kc^2}{a^2} ,$$

except in this case the mass density term is to be set equal to zero.

(a) Since the mass density is zero, it is certainly less than the critical mass density, so the universe is open. We can then choose $k = -1$. Derive an explicit expression for the scale factor $a(t)$.

(b) Suppose that a light pulse is emitted by a comoving source at time t_e, and is received by a comoving observer at time t_o. Find the Doppler shift ratio z.

(c) Consider a light pulse that leaves the origin at time t_e. In an infinitesimal time interval dt the pulse will travel a physical distance $ds = c dt$. Since the pulse is traveling in the radial direction (i.e., with $d\theta = d\phi = 0$), one has

$$c dt = a(t)\frac{dr}{\sqrt{1-kr^2}} .$$

Note that this is a slight generalization of Eq. (3.8), which applies for the case of a Euclidean geometry ($k = 0$). Derive a formula for the trajectory $r(t)$ of the light pulse. You may find the following integral useful:

$$\int \frac{dr}{\sqrt{1+r^2}} = \sinh^{-1} r .$$

(d) Use these results to express the redshift z in terms of the coordinate r of the observer. If you have done it right, your answer will be independent of t_e. (In the special relativity description that will follow, it will be obvious why the redshift must be independent of t_e. Can you see the reason now?)

PROBLEM 7: A ZERO MASS DENSITY UNIVERSE— SPECIAL RELATIVITY DESCRIPTION

(This problem is also not required, but can be done for 8 points extra credit.)

In this problem we will describe the same model universe as in the previous problem, but we will use the standard formulation of special relativity. We will therefore use an inertial coordinate system, rather than the comoving system of the previous problem. Please note, however, that in the usual case in which gravity is significant, there is no inertial coordinate system. Only when gravity is absent does such a coordinate system exist.

To distinguish the two systems, we will use primes to denote the inertial coordinates: (t', x', y', z'). Since the problem is spherically symmetric, we will also introduce "polar inertial coordinates" (r', θ', ϕ') which are related to the Cartesian inertial coordinates by the usual relations:

$$x' = r'\sin\theta'\cos\phi'$$
$$y' = r'\sin\theta'\sin\phi'$$
$$z' = r'\cos\theta' .$$

In terms of these polar inertial coordinates, the invariant spacetime interval of special relativity can be written as

$$ds^2_{\text{ST}} = -c^2 dt'^2 + dr'^2 + r'^2 \left(d\theta'^2 + \sin^2\theta' d\phi'^2\right) \ .$$

For purposes of discussion we will introduce a set of comoving observers which travel along with the matter in the universe, following the Hubble expansion pattern. (Although the matter has a negligible mass density, I will assume that enough of it exists to define a velocity at any point in space.) These trajectories must all meet at some spacetime point corresponding to the instant of the big bang, and we will take that spacetime point to be the origin of the coordinate system. Since there are no forces acting in this model universe, the comoving observers travel on lines of constant velocity (all emanating from the origin). The model universe is then confined to the future light-cone of the origin.

(a) The cosmic time variable t used in the previous problem can be defined as the time measured on the clocks of the comoving observers, starting at the instant of the big bang. Using this definition and your knowledge of special relativity, find the value of the cosmic time t for given values of the inertial coordinates— i.e., find $t(t', r')$. [Hint: first find the velocity of a comoving observer who starts at the origin and reaches the spacetime point (t', r', θ', ϕ'). Note that the rotational symmetry makes θ' and ϕ' irrelevant, so one can examine motion along a single axis.]

(b) Let us assume that angular coordinates have the same meaning in the two coordinate systems, so that $\theta = \theta'$ and $\phi = \phi'$. We will verify in part (d) below that this assumption is correct. Using this assumption, find the value of the comoving radial coordinate r in terms of the inertial coordinates— i.e., find $r(t', r')$. [Hint: consider an infinitesimal line segment which extends in the θ-direction, with constant values of t, r, and ϕ. Use the fact that this line segment must have the same physical length, regardless of which coordinate system is used to describe it.] Draw a graph of the t'-r' plane, and sketch in lines of constant t and lines of constant r.

(c) Show that the radial coordinate r of the comoving system is related to the magnitude of the velocity in the inertial system by

$$r = \frac{v/c}{\sqrt{1 - v^2/c^2}} \ .$$

Suppose that a light pulse is emitted at the spatial origin ($r' = 0$, $t' =$ *anything*) and is received by another comoving observer who is traveling at speed v. With what redshift z is the pulse received? Express z as a function of r, and compare your answer to part (d) of the previous problem.

(d) In this part we will show that the metric of the comoving coordinate system can be derived from the metric of special relativity, a fact which completely establishes the consistency of the two descriptions. To do this, first write out the equations of transformation in the form:

$$\begin{aligned} t' &= ? \\ r' &= ? \\ \theta' &= ? \\ \phi' &= ? \ , \end{aligned}$$

where the question marks denote expressions in t, r, θ, and ϕ. Now consider an infinitesimal spacetime line segment described in the comoving system by its two endpoints: (t, r, θ, ϕ) and $(t + dt, r + dr, \theta + d\theta, \phi + d\phi)$. Calculating to first order in the infinitesimal quantities, find the separation between the coordinates of the two endpoints in the inertial coordinate system— i.e., find dt', dr', $d\theta'$, and $d\phi'$. Now insert these expressions into the special relativity expression for the invariant interval ds^2_{ST} , and if you have made no mistakes you will recover the Robertson-Walker metric used in the previous problem.

DISCUSSION OF THE ZERO MASS DENSITY UNIVERSE:

The two problems above demonstrate how the general relativistic description of cosmology can reduce to special relativity when gravity is unimportant, but it provides a misleading picture of the big-bang singularity which I would like to clear up.

First, let me point out that the mass density of the universe increases as one looks backward in time. If the mass density parameter $\Omega \equiv \rho/\rho_c$ for our universe has a value of 0.2, at the low end of the empirically allowed range, then the universe today can be approximately modeled by the zero mass density universe. However, provided that Ω is greater than zero today, the zero mass density model cannot be taken as a valid model for the early history of the universe.

In the zero mass density model, the big-bang "singularity" is a single spacetime point which is in fact not singular at all. In the comoving description the scale factor $a(t)$ equals zero at this time, but in the inertial system one sees that the spacetime metric is really just the usual smooth metric of special relativity, expressed in a peculiar set of coordinates. In this model it is unnatural to think of $t = 0$ as really defining the beginning of anything, since the the future light-cone of the origin connects smoothly to the rest of the spacetime.

In the standard model of the universe with a nonzero mass density, the behavior of the singularity is very different. First of all, it really is singular— one can mathematically prove that there is no coordinate system in which the singularity disappears. Thus, the spacetime cannot be joined smoothly onto anything that may have happened earlier.

The differences between the singularities in the two models can also be seen by looking at the horizon distance. We learned in Lecture Notes 4 that light can travel only a finite distance from the time of the big bang to some arbitrary time t, and that this "horizon distance" is given by

$$\ell_p(t) = a(t) \int_0^t \frac{c}{a(t')} dt' \ .$$

For the scale factor of the zero mass density universe as found in the problem, one can see that this distance is infinite for any t— for the zero mass density model there is ***no*** horizon. For a radiation-dominated model, however, there is a finite horizon distance given by $2ct$.

Finally, in the zero mass density model the big bang occurs at a single point in spacetime, but for a nonzero mass density model it seems better to think of the big bang as occurring everywhere at once. In terms of the Robertson-Walker coordinates, the singularity occurs at $t = 0$, for all values of r, θ, and ϕ. There is a subtle issue, however, because with $a(t = 0) = 0$, all of these points have zero distance from each other. Mathematically the locus $t = 0$ in a nonzero mass density model is too singular to even be considered part of the space, which consists of all values of $t > 0$. Thus, the question of whether the singularity is a single point is not well defined. For any $t > 0$ the issue is of course clear— the space is homogeneous and infinite (for the case of the open universe). If one wishes to ignore the mathematical subtleties and call the singularity at $t = 0$ a single point, then one certainly must remember that the singularity makes it a very unusual point. Objects emanating from this "point" can achieve an infinite separation in an arbitrarily short length of time.

Total points for Problem Set 8: 43, plus up to 15 points extra credit.