LECTURE VII: THE EARLY UNIVERSE, FLUCTUATIONS, AND THE DEVELOPMENT OF STRUCTURE

AST 204 25 Feb 2008

I. The Planck Era and the Notion of Time and Space

First thing today, we'll discuss the fact that there's an epoch, associated with a density, a temperature, a mass, and a timescale 'earlier' than which it makes no sense to even discuss the notions of time and space. This is called the *Planck time*.

The energy of a photon or other highly relativistic ($\gamma >> 1$) particle is just $E = h\nu = hc/\lambda$. The Schwarzschild radius associated with a black hole of mass m is $r_s = 2Gm/c^2$ in ordinary units, or, since the energy E associated with a mass m is mc^2 , $r_s = 2GE/c^4$. Thus an energetic particle so energetic that its own self-gravity makes a black hole in which its wavelength just fits has a wavelength which satisfies

$$\lambda = r_s = \frac{2GE}{c^4} = \frac{2Gh}{\lambda c^3},$$

so this critical wavelength, the *Planck length*, is

$$l_P = \lambda = \sqrt{\frac{2Gh}{c^3}} = 5 \times 10^{-33} cm$$

The energy, the *Planck energy*, is

$$E_P = \frac{hc}{\lambda} = \sqrt{\frac{hc^5}{2G}} = 4 \times 10^{16} erg = 2 \times 10^{19} GeV$$

And, of course, an associated temperature, the *Planck temperature*

$$T_P = E_P / k = 3 \times 10^{32} K.$$

The associated mass, the *Planck mass*, is

$$m_P = E/c^2 = 4 \times 10^{-5}g$$

There is a *density* associated with this notion, namely one Planck mass per Planck length cubed, or

$$\rho_P = m_P / l_P^3 = 3 \times 10^{92} g / cm^3$$

and a time which might be either of two expressions—the inverse of the *frequency* associated with the wavelength l_P or the expansion time of the universe associated with ρ_P . These two quantities are the same within a factor of two; the *Planck time* is usually called $l_p/c = 2 \times 10^{-43} sec$. The fact that the two notions agree is profound. Can you figure out why they agree?

We do not know whether the universe was ever this dense, hot, or rapidly expanding, but we do know that it does not make sense to discuss conditions even more extreme than this. In order to *study* a more extreme condition (smaller length scales, say), one would have to resolve finer detail, which would involve particles of even shorter wavelength. But Planck energy particles already cannot propagate because they are within their own event horizon. So the concepts of space and time as a classical arena for events breaks down here, or probably actually well before—it makes no sense to talk about the spacetime continuum on scales shorter than the Planck length and time; the universe is a fundamentally quantum gravitational entity at this epoch. The notion of even what quantum gravity is is rapidly evolving at the present time, and the whole field of fundamental physics is currently doing a grand experiment of an unprecedented kind today. There is much hope that superstring theory will somehow produce the Holy Grail of the Theory of Everything, but the motivation now is mathematical beauty rather than any real connection with physics. We will see (hopefully within your lifetimes, but I suspect not within mine) whether this crusade will succeed. Perhaps the inflation is associated with this epoch, not with the GUT epoch at all (see Aurelien's lectures next week), and with the birthplace of this and possibly/probably an infinity of other universes in a seething timeless quantum gravitational froth, some tiny fluctuation in which spawned the whole thing.

V. A Summary Table, From the Beginning Till Now

We present below a table which summarizes the physical conditions in the universe from the Planck time to the present, which we'll use in the next few lectures.

au	$R(\tau)$	1 + z	Т	kT	$ ho_T$	$ ho_m$	$R_0 u_h$	r_h	m_h	description
(sec)			(K)	(eV)	g/cm^3	g/cm^3	(cm)	(cm)	M_{\odot}	
4(17)		1	2.7	2.3(-4)	9(-30)	3(-30)	4(28)	4(28)	3.8(23)	present
8(12)	$ au^{2/3}$	1500	4000	0.3	1.2(-20)	1(-20)	1.1(27)	7(23)	(18)	combination
4(11)		7000	2(4)	1.9	2.5(-18)	1(-18)	1.7(26)	2.4(22)	3(16)	equal m&r
10		5(9)	1(10)	1M	5(5)	0.2	4(21)	(11)	130	nuclear reac
3(-7)	$ au^{1/2}$	8(12)	1(13)	1G	1.6(18)	5(8)	1.4(17)	1.8(4)	5(-12)	quark-gluon
2(-11)		1(15)	1(15)	100G	4(26)	1(15)	1.2(5)	1.2	3(-18)	baryogenesis
2(-34)		1(28)	1(28)	1(15)G	3(77)	1(54)*	1.5(3)	1.5(-25)	7(-54)	end inflation
1(-34)	$e^{H\tau}$	2(45)	??	??	3(77)	(106)*	3.5(20)	1.5(-25)	0.1	mid-inflation
2(-36)		5(62)	1(28)	1(15)G	3(77)	2(158)*	8(37)	1.5(-25)	1(51)	beg. inflation
1(-43)	??	??	2(32)	2(19)G	3(92)	??	??	5(-33)	??	Planck era

Cosmological Conditions as Functions of Time and Redshift

The numbers in parentheses are powers of 10, so that, for example, $1(-43) = 1 \times 10^{-43}$. The columns in the table are mostly self-explanatory, but briefly are as follows: τ is the cosmic time at the epoch in question, Rlaw the relation between the scale factor and τ . 1+z the redshift factor, T the temperature in K, kT the typical thermal energy in eV–later MeV and GeV; ρ_T the total mass-energy density in g cm⁻³.

The column headed ρ_m is the density of rest mass Actually, this is a swindle for early times, because what it is is simply the present rest mass density multiplied by $(1 + z)^3$, and as such is just a 'tracer' for the mass today. Clearly early on at energies where protons and neutrons do not exist, we cannot easily calculate the rest mass density, and during inflation, if the identities of protons and neutrons were carried along, the densities would become ridiculous. These entries are noted by the presence of a (*). Note that this simple bookkeeping calculation assumes that the baryons, were they conserved, continue to carry one proton mass, which is probably wrong–in fact, they may carry no mass at all, because the mechanism which gives particles mass may do so only at relatively low energies. But this calculation illustrates graphically that baryon number cannot be conserved in inflation.

 R_0u_h is the comoving size of the particle horizon computed as if there were no inflation referred to the present universe. At and prior to the end of the inflation era, we just follow an incoming light ray which is at that horizon radius at the end of inflation, which illustrates how immense the effective horizon becomes during inflation. Note that the *physical* size of the horizon remains at the *event horizon* during this time, but the redshift factors become so large that the comoving distances become huge. About 55 efolds are required for the inflated particle horizon to encompass the entire observable universe today, and after about 80, roughly the supposed number, the horizon is 7 orders of magnitude larger than the present observable universe. Thus any point in the universe at the end of inflation had been able to receive information and send information to a volume which is now immensely larger than the present observable universe.

The adjacent column, the physical size of the horizon at the relevant period, is just the comoving one divided by (1 + z).

The column m_h is the rest mass, again referred to the present universe, within the horizonjust the present rest mass density within the comoving radius $R_0 u_h$.

We have talked about most of the phenomena identified in the table: The present, recombination, the equality of radiation and mass density, the era of nuclear reactions. Above a temperature corresponding to something like 1 GeV (though it is somewhat uncertain) baryons and mesons no longer exist and are replaced by a plasma of free quarks and gluons; at somewhat higher energies, above about 100 Gev, the electroweak symmetry is restored, and at something like these energies the processes which lead to baryogenesis probably freeze out and create the baryon number we see today.

Before this, at least according to our present hazy understanding, not very much interesting happens over a very large range in temperature and expansion until the energies of the GUT era are reached. There are only the electroweak and strong forces, only the fundamental leptons, quarks, any decaying heavy remnants of the GUT era, the gauge bosons responsible for the forces, including photons, Ws, Zs, and gluons. Then we hit the GUT era and inflation, prior to which there may (or may not) have been a more-or-less ordinary expansion era from the Planck time.

The numbers in this table are mostly, especially in the columns pertaining to the earliest times, very uncertain. The assumptions made in calculating the entries are that H_0 is about 70km/s/Mpc, that Ω_0 is unity but is made up currently of $\Omega_m = 0.3$ and $\Omega_{\Lambda} = 0.7$; that baryogenesis is associated with the decay or interaction of weakly interacting particles of mass about 100 GeV, that inflation is associated with the GUT energy scale at about 10^{15} GeV, and that the universe undergoes about 80 efolds of inflation.

III. Small Perturbations: The Linear Regime

We will see that fluctuations which arose in the inflation era reappear across the 'particle horizon' after inflation ends; fluctuations which enter the horizon before the end of radiation dominance at $z \approx 2500$ do not grow appreciably until radiation dominance ends, though fluctuations in the baryonic fluid are coupled to the radiation and participate in acoustic wave motion both before the era of matter domination and after until recombination. (We will follow common practice and call the time and redshift of the switchover from radiation to matter dominance τ_{eq} and z_{eq} , though the epoch has no good name that one does not stumble over.) We believe, however, that most of the matter is dark matter which does not couple to the radiation at all, and so fluctuations in it can begin to grow immediately after τ_{eq} . The initial flicker-noise power spectrum is modified on large scales, though it is essentially preserved in the dark matter on smaller scales. The table above shows that the horizon at the turnover is on comoving scales of about $2 \times 10^{26} cm$, about 70 Mpc, and that the horizon encloses a total mass or order 3×10^{16} solar masses. The largest bound objects we know in the universe are the great clusters, with masses of a few times 10^{15} solar masses; galaxies have masses three orders of magnitude smaller. So we are interested in small enough regions that we can proceed with purely Newtonian physics. We believe that the fluctuations grow to become the clusters of galaxies and galaxies we see in the universe today, and in this lecture we will see how they grow and how they can be characterized.

Since perturbations in the matter cannot grow before matter dominance, we can assume as the initial conditions, which we take at the beginning of matter dominance, that the power spectrum is whatever emerged from inflation, and that the Hubble constant is uniform, since the radiation, which up until the initial time provided essentially all the energy density, had time to smooth itself out on the length scales we are considering.

Consider for the sake of argument a 'tophat' perturbation in which the matter density is different from the *critical* density in the universe by some factor $(1 + \delta^+)$ in some spherical region of initial radius r_i , and is the average density in the universe outside this region. Now if Ω_m is different from unity, whether or not the total Ω is different from unity, the *mean* density is different from the critical density as well. Let $(1 + \delta_e)$ be the ratio of the mean matter density $\bar{\rho}$ to the critical density ρ_c (*e* stands for 'external'); note that δ_e is probably negative. Now when we were developing the description for the expanding universe, we wrote a Newtonian equation which was the direct analog of the Friedman equation for the radius of a shell in an expanding universe, just balancing the kinetic and potential energies:

$$\dot{r}^2 - \frac{8\pi G\rho r^2}{3} = 2\epsilon \tag{1}$$

We begin, we argue, with the same value of $H = \dot{r}/r$ for all the shells in our model, inside the perturbation and out. Let us think about the behavior of a perturbation with the density just such as to make $\epsilon = 0$. It is clear that in this case $\rho = \rho_c$ and $r = C\tau^{2/3}$, because this is just part of a critical-density universe. If the external density is less than ρ_c , it will expand faster and leave the little tophat alone.

Now consider a density in the tophat greater than the critical density. Then ϵ is negative, and the material in the tophat will reach some maximum radius at some point in the future, turn around, and collapse again. How does the density excess behave early in the expansion? We can perturb Equation 1 to follow the small difference in behavior from the critical case, which we can solve exactly, but it is messy and we do not need to; we know the answer already. The cosmological relations you derived in your problem sets say that at early times

$$1 - \Omega_m \approx \frac{1 - \Omega_0}{\Omega_m (1 + z)} \propto (1 + z)^{-1}$$

and furthermore that any contribution to the density in the present universe in the form of a cosmological constant is completely negligible early in the universe. The perturbation is a little piece of a higher-density universe–it does not know about the exterior, remember, and

$$1 - \Omega_m = 1 - \frac{\rho}{\rho_c} = \frac{\rho - \rho_c}{\rho_c} = \delta^+$$

The shell in question is not expanding, to be sure, exactly the way the universe is, but when the perturbation is still small, it is certainly expanding approximately the way the universe is, so we can neglect for a while the difference in the factor (1+z) inside and out; so we get, for no work,

$$\delta^+(z) \propto (1+z)^{-1} \propto R$$

Of course, the same proportionality applies to δ_e , so whether we refer the density perturbation to the critical density or the mean density, the *contrast* $\delta \propto R$. This is a very important result.

Think again about the energetics. Let us write the energy equation (1) in the following way for the initial conditions:

$$\dot{r}^2 - \frac{8\pi G\rho_c r^2}{3} - \frac{8\pi G(\rho - \rho_c)r^2}{3} = 2\epsilon.$$

Here we have just added and subtracted the term $8\pi G\rho_c r^2/3$. But the first two terms just cancel (this is the definition of the critical density), and leave us with (dividing by 2):

$$-\frac{4\pi G(\rho-\rho_c)r^2}{3} = \epsilon.$$

We can rephrase this in the suggestive way

$$-\frac{G\delta m}{r} = \epsilon$$

where δm is just the *excess mass* in the perturbation interior to the shell at radius r compared to the mass that would be contained if the shell contained just the critical density. Clearly

$$\delta m = m\delta^+$$

where m is the total mass inside the shell. Then δm goes like $\tau^{2/3}$ as well, and this can be thought of profitably and interchangeably in two very different ways. The excess mass is growing because the density contrast is growing because the shell is expanding somewhat less rapidly than the external universe. This is correct and precise. The excess mass is growing because the density contrast is growing because matter is slowly moving inward with respect to a shell which is expanding at the same rate as the external universe is expanding. This is also correct, but expresses a rather different point of view.

We have restricted ourselves here to spherical perturbations; clearly one does not need to restrict oneself to uniform ones. Everything we have said applies to a shell in a general spherical perturbation, with all the densities and density contrasts replaced with the *mean* density inside the shell–what matters, after all, is just how much *mass* there is within the shell. Rather more surprising is that these results are not confined to *spherical* perturbations while the perturbations are very small–if one has some general density excess field $\delta(\mathbf{u}) = (\rho(\mathbf{u}) - \bar{\rho})/\bar{\rho}$, then

$\delta(\mathbf{u}) \propto R$

This result is true as long as $\delta \ll 1$ but is crudely correct as long as $\delta \ll 1-i.e.$, as long as the perturbation in the density is not much larger than the density itself. This approximation is the *linear* approximation and the realm in which it is valid is called the *linear regime*.

IV. Spherical Perturbations in the Nonlinear Regime

Now if we are going to make a galaxy or a cluster or any other kind of bound structure, we are clearly interested in perturbations with negative ϵ . We are interested in structures with enormous density contrast, and furthermore structures which *are bound*. Remember that most of the matter is dark matter, which does not radiate or otherwise interact with either itself or other matter except gravitationally, which means that it cannot lose energy. The galaxy is bound *now* and must have been bound for all of its existence if most of its mass cannot get rid of any energy.

So consider a shell in a perturbation which has positive δ^+ and hence negative ϵ . Its energy, we have seen, is $-G\delta m/r_i$ initially, and this is conserved. At some point in its history, it will reach some maximum radius, and then collapse. We can easily calculate how big it will get. At its maximum radius, its energy is all potential, and is Gm/r_{max} . So

$$G\delta m/r_i = Gm/r_{max},$$

and we obtain the exquisitely simple result

$$\frac{r_{max}}{r_i} = \frac{m}{\delta m} = \frac{1}{\delta^+}.$$

if a shell at the initial time has interior to it a 1 percent density excess, it will expand a factor of 100 before turning around. If we do a little work, we can also calculate how long it takes to reach maximum expansion; twice this is the collapse time, the time it takes the structure it is forming to reach high density, and, well—form. We will solve the equation of motion, which we can write

$$\left(\frac{dr}{d\tau}\right)^2 = 2Gm\left(\frac{1}{r} - \frac{1}{r_{max}}\right);$$

clearly, as we saw above, $\epsilon = -Gm/r_{max}$. Then we write

$$\left(\frac{2Gm}{r_{max}^3}\right)^{1/2} d\tau = \frac{dr/r_{max}}{\sqrt{r_{max}/r - 1}}$$
$$= \frac{dq}{\sqrt{1/q - 1}}$$
$$= \frac{\sqrt{q}dq}{\sqrt{1-q}},$$

where we have here let $q = r/r_{max}$. If we let $q = \sin^2(\theta/2)$, (don't worry about where the /2 comes from-it is convenient later but is not necessary) $\sqrt{(1-q)}$ is $\cos(\theta/2)$ and $dq = \sin(\theta/2)\cos(\theta/2)d\theta$, so

$$\left(\frac{2Gm}{r_{max}^3}\right)^{1/2} d\tau = \sin^2(\theta/2)d\theta$$
$$= \frac{1 - \cos\theta}{2}d\theta.$$

Notice that the term $(1 - \cos \theta)/2 = \sin^2(\theta/2)$ is just r/r_{max} again. We can integrate this trivially, and get the *parametric solution*

$$\tau = \left(\frac{r_{max}^3}{8Gm}\right)^{1/2} \left(\theta - \sin\theta\right)$$
$$\frac{r}{r_{max}} = \frac{1 - \cos\theta}{2}.$$

For the geometry afficionados among you, this is the pair of equations which describe the development of a *cycloid*, the locus of a point on a circular hoop of unit radius as it rolls along the x-axis, starting at x = 0, y = 0, and ending at $x = 2\pi, y = 0$; θ is the angle through which the hoop has rolled. In this identification, x is $\tau \sqrt{8Gm/r_{max}^3}$; y is $2r/r_{max}$.

Look at the behavior of r/r_{max} . It is zero for $\theta = 0$ and 1 for $\theta = \pi$, so $\theta = \pi$ corresponds to maximum expansion. Then it is zero again for $\theta = 2\pi$, which is fully collapsed. If we let

$$t_c = 2\pi \sqrt{r_{max}^3 / (8Gm)} = \pi \sqrt{r_{max}^3 / (2Gm)}$$

then the expression for τ is

$$\tau = \left(\frac{t_c}{2\pi}\right)(\theta - \sin\theta);$$

when θ is π , we are at maximum expansion, and $\tau = t_c/2$; when θ is 2π , we have collapsed and $\tau = t_c$. The quantity t_c is the *collapse time*. With a little manipulation you can also relate the collapse time to the initial conditions:

$$t_c = \frac{\pi}{H_i} (\delta^+)^{-3/2}$$

where H_i is the Hubble constant at the initial conditions. You will be asked to derive this in the next problem set.

We have done a lot of manipulation; where has it gotten us? We see that a positive perturbation is like a little piece of a *closed universe*, which reaches some maximum radius and then collapses. If we draw a spacetime diagram, the positive part of the perturbation separates out from the expanding universe and collapses to form a bound object.

We might well ask what the *interface* looks like—do we leave a vacuum at the (admittedly unrealistic) sharp edge of the tophat? The answer is no. Remember that the maximum radius and collapse time depend on the *average* density inside a shell. A shell just outside the tophat feels the same average density as the last shell *in* the tophat, so it collapses in the same time. As we go away from the edge of the tophat the average density clearly tends to the mean density in the universe, and the collapse time gets longer and longer; if the mean density of the universe is less than the critical density, there is a *last bound shell*. Let us see how this goes:

The average density inside a shell is

$$\rho_{av} = \frac{3m}{4\pi r_i^3},$$

but m is just the mass belonging to the mean density in the universe, $4\pi\bar{\rho}r^3/3$, plus the excess mass in the tophat, which we can call δM :

$$\delta M = 4\pi R_i^3 \rho_c (\delta^+ - \delta_e)/3$$

Here everything refers to conditions at the initial epoch, and capital R and M to the *whole* tophat, *i.e.* at its edge. Then clearly, outside the tophat,

$$\rho_{av} = \bar{\rho} + \frac{3\delta M}{4\pi r_i^3}, \quad r_i > R_i$$

If we subtract the critical density and divide by it, we can write the much more intuitive expression:

$$\delta^+(r_i) = \delta_e + (\delta^+ - \delta_e) \left(\frac{R_i^3}{r_i^3}\right), \quad r_i > R_i,$$
$$= \delta^+, \quad r_i < R_i.$$

here δ^+ with no argument is the density excess with respect to the critical density *inside* the tophat. Thus we see that if δ_e is negative, there will be a radius $r_{i,last}$,

$$r_{i,last} = R_i \left(\frac{-\delta_e}{\delta^+ - \delta_e}\right)^{1/3}$$

which has energy zero. All shells inside of this eventually fall into the object being formed, and can represent an amount of mass large compared to the mass in the original tophat, especially if the original tophat is a large-amplitude perturbation.

In the next lecture we will investigate the energetics of the object being formed, and see how the present properties of galaxies and clusters connect to the initial conditions.