M1 Cosmology - 4 - The Hot Big Bang and Inflation

Nick Kaiser

Contents

1 The Hot-Big-Bang model

1.1 The radiation dominated era

- In addition to the dark and luminous matter in the universe there is the *cosmic microwave background* (CMB) radiation
- It was discovered in 1965 by Penzias and Wilson, is coming to us isotropically from all directions and has a very accurately thermal spectrum with $T \simeq 3$K.
- this currently has a very small energy density
  - expressed as a fraction of critical density $\Omega_r \equiv \rho_r/\rho_c \simeq 6 \times 10^{-5}$
  - or about 10^{-3} times the density of baryonic matter today
  - or about 2×10^{-4} times the total matter density today
- but the CMB photons are red-shifting, and are therefore losing energy as time goes by
  - so their energy density is falling as $1/a^4$ (the number density of photons is decreasing as $n \propto 1/a^3$, while the energy per photon is $h\nu \propto 1/a$)
  - hence beyond a redshift of about 5000 – the redshift of matter-radiation equality – the radiation dominated over the matter (dark plus baryonic) density
  - and radiation has a significant pressure
    - for isotropic radiation like the CMB the pressure is $P = \mathcal{E}/3 = \rho c^2/3$

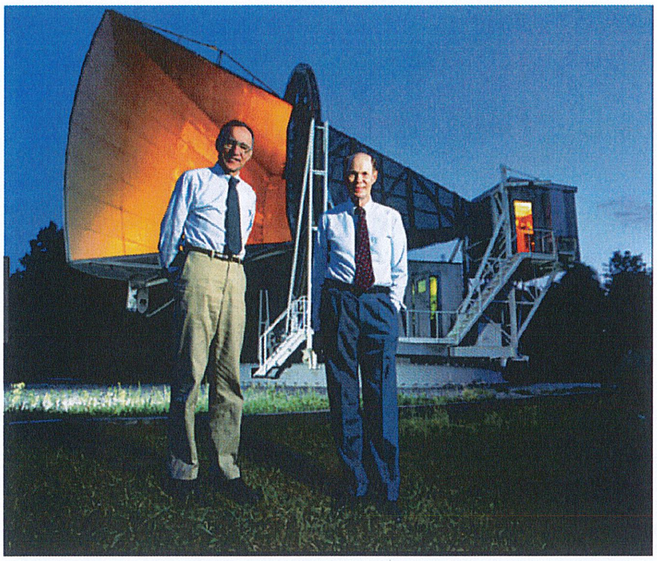


Figure 1: Penzias and Wilson in front of the radio telescope with which they discovered the cosmic microwave background radiation. It appeared to be coming uniformly from all directions.
It was later found to have a nearly precisely thermal spectrum with a current temperature of about 3K.
It is interpreted as a 'fossil' remnant of an earlier hot phase of the universe when the matter was ionised and the plasma was locked in thermal equilibrium with the radiation.
According to *Saha's equation*, the matter and radiation would have decoupled when the temperature was a few thousand Kelvin – and thus at a redshift of about 1000. So this radiation provides us with a 'snapshot' of the universe as it was a few thousand years after the big bang.

    * this can be seen from the stress-energy tensor $T^{\mu\nu} = \int (d^3p/p^0) f(\mathbf{p}) p^\mu p^\nu$ in which, since photons have $|\mathbf{p}| = p^0$, the only difference between $P = T^{zz}$ (the flux of z-momentum in the z direction) and $\mathcal{E} = T^{00}$ is the presence of a factor μ^2, with $\mu = p^z/|\mathbf{p}| = \cos\theta$, and the average of μ^2 over directions, for an isotropic distribution function $f(\mathbf{p}) = f(|\mathbf{p}|)$, is $\langle \mu^2 \rangle = 1/3$.
    * an alternative way to understand this is to think of a photon as a particle moving at $v = c$ with momentum $|\mathbf{p}| = E/c$ and bouncing around in a reflecting box and transferring momentum to the walls when it bounces. If the length of the side of the box is L then the time between collisions with the $+z$-wall is $T = 2L/\mu c$ (as the speed in the z-direction is μc) and the momentum transferred is $2p_z = 2\mu|\mathbf{p}| = 2\mu E/c$. So the momentum transfer per unit time per unit area is $P = (2\mu E/c)/(T = 2L/\mu c)/L^2 = \mu^2 E/L^3$, and averaging over directions then agains gives $P = \mathcal{E}/3$.
    * Both of the above are considering photons as particles of light. But the same conclusions are reached from the electromagnetic stress energy tensor for an isotropic bath of EM waves with $\mathcal{E} = \epsilon_0 |\mathbf{E}|^2 + \mu_0^{-1} |\mathbf{B}|^2$ and using the *Maxwell stress tensor* $T_{ij} = -\epsilon_0 (E_i E_j - \frac{1}{2}\delta_{ij}|\mathbf{E}|^2) - mu_0^{-1}(B_i B_j - \frac{1}{2}\delta_{ij}|\mathbf{B}|^2)$.
  - thus the equation of state parameter for radiation is $w = w_{\mathrm{r}} = 1/3$

- The next question is: how does this change the expansion history of the universe?
  - You might think that this would speed up the expansion.
  - This, after all, is what happens in the hot plasma in, say, an H-bomb explosion.
  - But there the acceleration of the expansion of the matter is caused by *pressure gradient forces*.
  - Here the pressure is spatially homogeneous, so there is no ∇P force density.
  - And, as we have seen, as it is $\rho + 3P/c^2$ that appears in the acceleration equation, the pressure actually enhances the *deceleration* of the universe.

1.2 The effect of the cosmic radiation on the dynamical evolution of the universe

1.2.1 The continuity equation and energy non-conservation

- The continuity equation is modified from $\dot{\rho} = -3H\rho$ to
  - $\boxed{\dot{\rho} = -3H(\rho + P/c^2)}$
  - the extra term coming from the fact that any volume element will in the process of expanding do PdV work and since $E = Mc^2$ this implies a decrease in the density over and above the $\rho \propto 1/V$ for pressure-less matter

- this is simply the first law of thermodynamics
- and, applied to the radiation itself, which has $P = \rho c^2/3$ gives $\dot{\rho}/\rho = -4H = -4\dot{a}/a \Rightarrow \rho \propto 1/a^4$
- so consistent with the idea that there is a constant comoving density of photons (physical number density $n_\gamma \propto 1/a^3$) and each photon losing energy – as measured by fundamental observers it is passing – as $h\nu \propto 1/a$.

- on the other hand, this may seem surprising
  - each and every volume element is losing energy by doing PdV work
  - what has happened to conservation of energy?
  - for a bounded sphere the total energy is conserved; there is energy flowing out through spherical shells and this appears in the increased kinetic energy of expansion of the matter being blown off
  - but for an unbounded universe the concept of total energy is not useful and it is better to simply come to terms with the fact that every volume element is losing energy (if $P > 0$), with the flux through any closed boundary being the integral of what's being lost in the enclosed volume, and, in a very real sense, energy in the expanding universe is not conserved
- it is perhaps worth re-emphasizing that the density ρ here is the total energy density
  - for a gas with kinetic pressure, for instance, $P \sim \rho\sigma_v^2$ so one might imagine that the pressure correction to the density on the RHS of the continuity equation $\rho \times (1 + \mathcal{O}(\sigma_v^2/c^2))$ is just some special relativistic correction to the rest-mass energy density
  - that is entirely wrong – any extra energy density associated with the random motion of particles has already been incorporated in ρ, as this is the *total* energy, not the rest-mass, density

1.2.2 The energy equation and the spatial curvature

- The Friedmann energy equation is
- $\boxed{\dot{a}^2 = (8\pi/3)G\rho a^2 - kc^2}$
- At the time of the CMB discovery, the density and expansion rate $H = \dot{a}/a$ were quite uncertain.
- but it was known that the kinetic term and the potential energy term are quite similar today
- But, as kc^2 is constant, and the potential term varies as $\rho a^2 \propto 1/a$, the kinetic and potential terms must have been equal to each other not so long ago (on a logarithmic scale)
- and during the radiation era these would have been enormous compared to the 'curvature term' kc^2 and would have continued to increase indefinitely going back in time.
- So, during the radiation era, it is a very good approximation to ignore the curvature term and that the density and expansion rate are directly related.

1.2.3 The acceleration equation

- As we have seen, the acceleration equation is
  - $\boxed{\ddot{a} = -\frac{4\pi}{3}G(\rho + 3P/c^2)a}$
  - This can be considered to be a consequence of special relativity and thermodynamics (in the continuity equation) and the idea that pressure should not appear in the energy equation – as it would seem unreasonable for $\dot{a}$ to change discontinuously if we 'switch on' pressure
  - as we saw, it follows fairly directly from the geodesic deviation equation, using the fact that the Ricci tensor appearing there is sourced by the 'trace-reversed' stress-energy tensor
  - is it often characterised as saying 'pressure gravitates in GR'.
  - What does this mean more physically?

- First, this is a real and, in principle, locally measurable effect
  - as mentioned, there are no pressure gradient forces; the plasma fluid elements are in free-fall
  - so if we have a pair of massive test particles these must expand from each other in the same way as elements of the plasma
  - so if we simply watch two free but massive test particles, we will see that their separation obeys the modified acceleration equation
  - one can similarly measure the local density of pressure-less invisible non-interacting dark matter by the pressure in a spring that would be needed to keep them from moving together
  - if the DM is relativistic neutrinos that have $P = \rho c^2/3$, the pressure you will measure will be twice as large
- But, on the other hand, it does *not* mean that, if the pressure inside a spherical star were to change because of the ignition of nuclear reactions say this would be measurable in the *external* gravitational field
  - the external field can be expressed as in integral of an effective stress energy that keeps the external field that of a constant mass
  - 'what happens in Vegas stays in Vegas'
  - and a balloon containing gas with pressure does not create an excess gravitational field in its environment; the tension in the membrane cancelling the additional attraction term from the pressure

1.2.4 Solutions for the scale-factor in the radiation era

- The consequences for cosmology is an increase the deceleration of the universe as compared to the case for a matter dominated cosmology
  - with $\rho \propto 1/a^4$ in the energy equation, and ignoring there the curvature term as it is negligible, one finds that $H \propto a^{-2}$ and since $H \sim 1/t$ this implies
    * $\boxed{a \propto t^{1/2}}$

    rather than $t^{2/3}$ as in the pressure-less case
  - so this decreases the age of the universe – but not by all that much since the universe was matter dominated for the great majority of its life.

1.3 The thermal history of the universe and big-bang nucleosynthesis

- the current temperature of the CMB is $T \simeq 3K$
  - corresponding to an energy $kT \sim 3 \times 10^{-4}$eV
- so at the epoch of matter and radiation equality ($z_{\rm eq} \sim 3 \times 10^3$) the temperature was $kT \sim 1$eV
- and the expansion rate was $H^2 \simeq 2 \times (1 + z_{\rm eq})^3 H_0^2$ so $H \simeq 2 \times 10^5 H_0$
- and the age of the universe was $t_{\rm eq} \sim 1/H_{\rm eq} \sim 3 \times 10^4$yrs $\simeq 10^{12}$s
  - this is somewhat earlier than decoupling at $t \sim 3 \times 10^5$yrs.
  - note that in doing these calculations more precisely one needs to take into account the massless – or nearly massless – neutrinos
  - these are believed to have been in thermal equilibrium with the photons at early times and would have had similar energy density (not exactly equal as they have a Fermi-Dirac rather than Bose-Einstein distribution)
- prior to this, $\rho \propto T^4 \propto 1/t^2$ so $T \propto 1/t^{1/2}$

- *so the age and the temperature of the universe are directly related* in the radiation era
- provided one takes into account the number of order unity from *'effective number of relativistic degrees of freedom*

- thus when the universe was a few minutes old, i.e. at a time $t \sim 100$s, the temperature was $kT \sim 0.1$MeV
- this is interestingly close – within about an order of magnitude – of the energy scale of nuclear reactions
  - the mass difference between a neutron and a proton being $\simeq 1.5$MeV
  - the binding energy of a helium nucleus is about 28MeV
- so when the universe was substantially younger than a few minutes the temperature was greater than the neutron-proton mass difference so Boltzmann's formula would predict roughly equal numbers of protons and neutrons
- and the equilibrium abundance of helium – which can be estimated from Saha's equation – would have been tiny
- but by the time the universe cooled and aged to become a few minutes old
  - the Boltzmann formula tilted the balance towards protons
  - and the neutrons were 'locked up' in helium nuclei (with trace amounts of other light elements like lithium)
  - and shortly after any free neutrons would have decayed to protons (the half life being about 11 minutes)

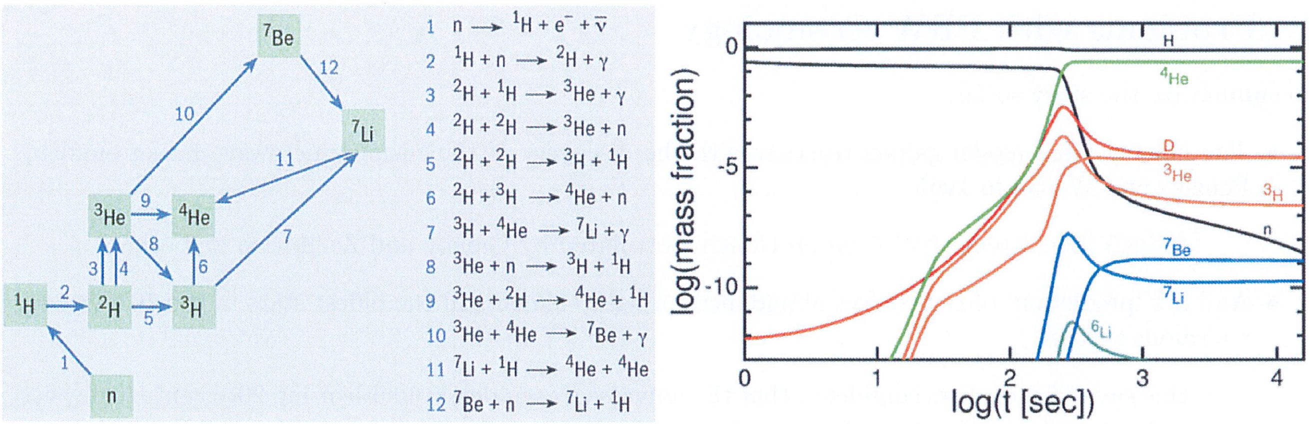


Figure 2: The network of nuclear reactions that are evolved numerically in BBN calculations is shown on the left. On the right are the results of numerical calculations of the abundances as a function of time. One can see that the number of neutrons and protons start off equal, but as the temperature drops, thermal equilibrium starts to favour the lighter protons. Around 200s, it becomes favourable for the some of the neutrons to become locked up in, primarily, helium nuclei.

- detailed calculations (figure 2) predict a final proton to neutron abundance ration $n/p \simeq 1/7$ corresponding to a universe containing about 25% helium by weight
  - this was a huge success; it nicely explained what astronomers had inferred from studying the chemical composition of surfaces of the oldest (very metal poor) stars
- Moreover, as shown in figure 3, the results are sensitive to the value of the *'entropy per baryon'* $n_\gamma/n_{\rm B}$ with agreement with measured abundances requiring a current density of baryons – expressed as a fraction of the current epoch critical density – of $\Omega_{\rm B} \simeq 0.05$
  - i.e. considerably less than the dynamical estimates of the current total matter density

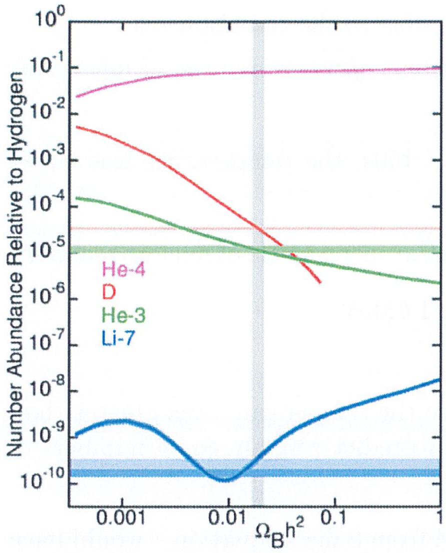


Figure 3: As well as being a triumphant success in explaining the hitherto unexplained fact that stars are apparently born with a mix of hydrogen and helium with the latter composing about 25% of the density, the abundances of the other light elements depends on the ratio of baryons (neutrons plus protons) to photons. The curves in the plot on the left shows the computed abundance of the light elements as a function of the density of baryons – the quantity plotted is $\Omega_{\rm B} h^2$ where $h \equiv H_0/100{\rm km/s/Mpc}$; since the density of radiation is accurately known, the relative density of the baryons scales as the square of the then poorly known expansion rate. The horizontal lines show observed abundances. Evidently, consistency with observation requires $\Omega_{\rm B} h^2 \simeq 0.02$, or, since we now know that $h \simeq 0.7$ a baryonic density parameter $\Omega_{\rm B} \simeq 0.04$. Remarkably, all of the measurements are consistent with this value.

  - but nicely consistent with $\Omega_{\rm m} \simeq 0.3$ and the baryon to dark-matter ratio observed in clusters

This was a remarkable development (well described in Steven Weinberg's book "*The First Three Minutes*"). Combining 20th century nuclear physics with the bold assumption of the applicability of the homogeneous FLRW models and coming up with empirically verifiable – and correct – results showed that the universe is explicable over at least 17 orders of magnitude in time (and therefore something like 34 orders of magnitude in density). This gave cosmologists enormous confidence, and the hubris to continue to the much more speculative physics of inflation that purports to explain what happened at still earlier epochs.

2 Problems with FRW Cosmology

To summarize the story so far:

- The hot-big bang model gained traction with the discovery of the cosmic microwave background by Penzias and Wilson in 1965
  - though its existence had been previously postulated by Gamov and Zel'dovich and others
- And it's 'prediction' of the correct abundances of light elements in the oldest stars in the 70's was an enormous success
  - this gave cosmologists confidence that the universe was understandable using 20th century physics at least back to a fraction of a second, if not earlier

But there were a few 'flies in the ointment' as several cosmologists of the time – and, in particular Bob Dicke – pointed out.

2.1 The fine-tuning problem

- the expansion rate is $H_0 \simeq 70{\rm km/s/Mpc}$
- this gives the critical density ρ_c
  - the density the universe would have to have to be just marginally bound (i.e. $\Omega_m = 1$)
- dividing by the observed *luminosity density* of the universe gives $\rho_c/\mathcal{L} \sim 1000 M_\odot/L_\odot$
  - the mass to light ratio of the universe if it has critical density
- but the mass to light ratio of the Coma cluster is $M/L \simeq 300 M_\odot/L_\odot$
- equivalently, if the M/L of Coma is representative of the universe, $\Omega_m \simeq 0.3$

- this is a bit of an embarrassment, for the following reason:
  - In the Friedmann energy equation – in the form
    * $\dot{a}^2 = (8\pi/3)G\rho a^2 - kc^2$

    – this means that the first two terms differ by about a factor 3 at present
  - but the second scales inversely with $a(t)$, since $\rho \propto 1/a^3$
  - and since the last term is constant, $\dot{a}^2$ must be proportional to $1/a$ also
  - but in this model $a(t)$ tends to zero at early times
  - this means that at very early times $a \ll a_0$, the first and second terms in the Friedmann were both extremely large compared to their present values, *but were not exactly equal*
  - the physics of the early universe somehow conspired to make the kinetic and potential terms almost exactly equal but with a tiny difference such that at the present epoch the difference is just becoming appreciable appreciable
  - this seems unnatural – and this is called the *'fine-tuning'* problem
  - it is also sometimes called the *'flatness problem'*

2.2 The horizon problem

- We have already seen that the comoving distance that light can travel is $\chi(z) = (c/a_0)\int dz/H(z)$ and remains finite as $z \to \infty$
  - so there is evidently a *horizon*; sources of infinite redshift are at finite distance
  - moreover, the comoving distance to the horizon was smaller in the past
    * consider a particular galaxy that we can observe today (which has a finite redshift and a distance that is less than the horizon distance)
    * in the past, it was receding from us faster, and was at a higher redshift
    * and at a finite time in the past, its redshift would have become infinite
    * and before that time it was outside of our horizon
- in discussing this, a key concept is the *'comoving horizon scale'*
  - the *physical distance* light can travel in one *'expansion time'* $t \sim 1/H$ is $r \sim ct = c/H$
  - at early times, and independent of the exact value of Ω_m, the size of the universe scales as
    * $a \propto t^{2/3}$ in the matter dominated era and
    * $a \propto t^{1/2}$ in the radiation era
- so the comoving horizon size is:
- $ct/a \propto \begin{matrix} t^{1/3} \\ t^{1/2} \end{matrix} \to 0 \quad \text{as} \quad t \to 0$
- so the distance over which light, and therefore any causal influence, can propagate shrinks to zero faster than the size of the universe itself; the big bang was apparently acausal!
- this is most starkly embarrassing in the face of the observation that the cosmic microwave background has a temperature that, aside from a 'dipolar modulation' that is believed to be caused by our motion, has the same temperature in all directions to about one part in 10^5
- yet in the FRW models the regions that last scattered the radiation we see from opposite hemispheres have never been in causal contact (see figure 5)
- the root cause of both the fine-tuning and horizon problem is that, in FRW models, the universe is *decelerating*
- this means that the expansion energy is decreasing with time

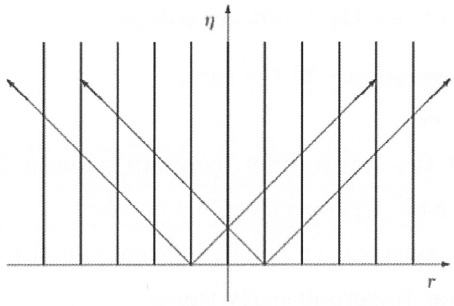


Figure 28.4: The causal structure of FRW models is most clearly shown if we plot particle world-lines in $\eta - r$ coordinates (i.e. conformal time *vs* comoving spatial coordinate). The vertical lines represent comoving observers, while the diagonal lines are light rays. Causal influences are constrained to lie within the light-cones. Since conformal time $\eta \to 0$ at the big-bang, this means that the the initial singularity (the line $\eta = 0$) is acausal, in the sense that comoving observers are initially out of causal contact with each other. If, however, the Universe *accelerates* at early times, the initial singularity is pushed back to $\eta = -\infty$ in this plot, and the big-bang is then causal.

Figure 4: Causal structure of the matter dominated FRW model. The spatial coordinate r in this plot is what we have denoted by χ (or x in the Newtonian analysis).

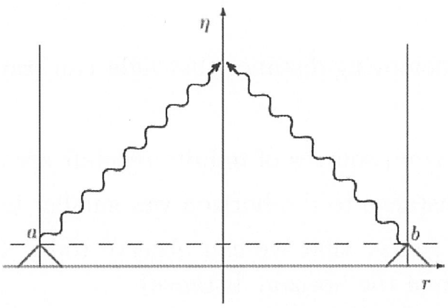


Figure 28.5: The wiggly diagonal lines in this conformal space-time plot show the world lines of two photons which we now see arriving from opposite directions, and which were last scattered at points a, b. The horizontal dashed line indicates the moment of recombination of the plasma. Also shown are the past light cones of the events a, b. The regions which can have causally influenced the photons prior to their departure are very small compared to the separation of the scattering events.

Figure 5: CMB photon paths in conformal coordinates.

- so the relative difference between this and the potential energy must be increasing

- and having non-vanishing radiation density in the radiation era only exacerbates the problem
  - so it is still the case that the distance light can travel in an expansion time ct shrinks to zero faster than the size of the universe
- while if, for example, we had had a power law expansion $a \propto t^\gamma$ with $\gamma > 1$
  - so $\dot{a} \propto t^{\gamma-1}$ would be increasing with time
- then the comoving horizon size ct/a would not shrink to zero at $t \to 0$.

2.3 The age and monopole problems

Another problem is called the age problem. If, as certainly seems natural, and, as we will see is predicted by inflation, the universe is spatially flat and dominated by matter, then the deceleration at late times results in the age of the universe being uncomfortably low as compared to the estimated age of the oldest stars.

Another different type of problem is known as the *monopole problem.* These problems arise if one has fields in the expanding universe which are in thermal equilibrium – and therefore spatially incoherent – at early times, but have potentials such that, as the universe cools, the field falls into one or other minima and 'defects' are formed. Possibilities are cosmic strings, domain walls and monopoles, which would have had separation, at the time of their formation, on the order of the horizon size. Strings are relatively benign, but walls or monopoles, if formed in this way, would come to dominate the density of the universe, in conflict with observation. These problems could be avoided if the fields were, at early times, spatially ordered. But with different parts of the universe out of contact with one another that seems hard to achieve.

2.4 The singularity theorems

So the hot big bang model provided a good working model for the evolution of the universe, but its origin seems to be required to be 'put in by hand'. Many cosmologists were unhappy with this, but it was hard to see how the problem could be avoided. While what happened at $t \to 0$ was, of course, a matter of speculation, Hawking and Penrose, building on Raychaudhuri's work, obtained their *singularity theorems* that showed that, for any sensible equation of state, an initial singularity was unavoidable.

3 Inflation

The flatness and fine-tuning problems can be avoided if the universe was dominated by a scalar field – the 'inflaton' – in the early universe.

The age problem is avoided if the universe undergoes 'late-time inflation'. Either driven by another scalar field or by Einstein's cosmological constant.

3.1 Inflation in the early universe

- prior to Maxwell there were 3 forces: electricity, magnetism and gravity, but Maxwell unified E & M in the middle of the 19th century, so then there were 2
- with the discovery of radioactivity and nuclear physics there were 4 known forces in the universe
  - the strong and weak nuclear forces, electromagnetism and gravity.
- but the weak and electro-magnetic interactions are unified by the 'electro-weak' theory invented in the middle of the 20th century
  - central to which is spontaneous symmetry breaking with the Higgs field – a scalar field – settling into its present value at a temperature on the order of a GeV and resulting in the asymmetry between the mass-less photon and the other 'gauge-bosons' ($W^{\pm}$ and Z)
- and there are indications that the strong nuclear force may be unified with the other non-gravity forces at a temperature of about 10^{15} GeV
- Alan Guth – who was working on grand unified theories of particle physics at the time – describes hearing Bob Dicke giving a talk on cosmology in which he described the horizon and flatness/fine-tuning problems. He realised that a scalar field of the kind that was considered as a way to bring the strong force into a unified theory could, potentially, solve these problems. He called this theory 'inflation'.

- A scalar field that is sufficiently smooth – i.e. not varying much in space – has energy density and pressure given by
  - $$\boxed{\begin{aligned} \rho c^2 &= \tfrac{1}{2}\dot{\phi}^2 + V(\phi) \\ P &= \tfrac{1}{2}\dot{\phi}^2 - V(\phi) \end{aligned}}$$
- where $V(\phi)$ is the *potential* for the field
  - for a simple massive scalar field, this is $V(\phi) = (mc^2/\hbar)^2\phi^2/2$
  - another alternative – which would give a self-interacting field – would be $V(\phi) = \lambda\phi^4$ with λ a constant (see figure 7)
- if the field happens to be also sufficiently constant in *time* and has a a sufficiently large value that $V \gg \dot{\phi}^2$ this material has, it turns out, the interesting property of being in a state of great tension.

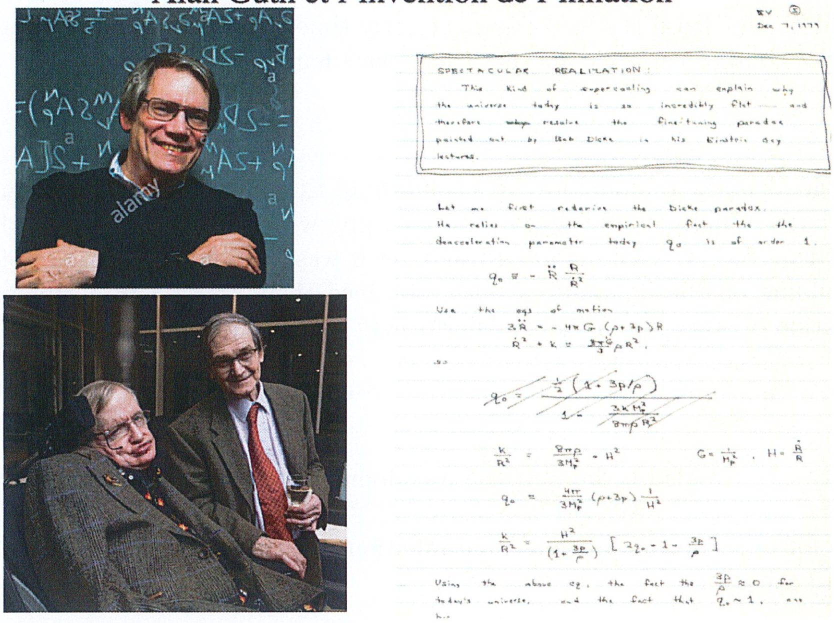


Figure 6: Alan Guth, upper left, was listening to a talk in 1979 in which Bob Dicke explained the triumphs – and problems – of cosmology. His notebook on the right shows his realisation that, by invoking a scalar field analogous to the Higgs field during a phase of 'grand unification' he might be able to avoid the constraints imposed by the singularity theorems of Hawking and Penrose (bottom left). As always 'success has many parents' and others were thinking along similar lines. One might also consider Fred Hoyle's 'steady state theory' with its continual creation of matter by the 'C-field' to be a precursor.

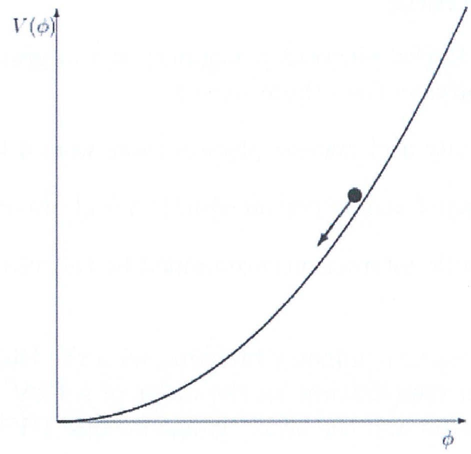


Figure 29.2: In the *chaotic inflation* scenario, the potential function $V(\phi)$ is assumed to be a monotonically increasing function with $V(0) = 0$. The potential could simply be a mass term $V \propto \phi^2$ or perhaps $V \propto \phi^4$.

Figure 7: Potential for the scalar field in the 'chaotic' inflation scenario.

- it has an equation of state $P = -\rho c^2$

- it follows immediately from the continuity equation $\dot{\rho} = -3H(\rho + P/c^2)$ that a universe dominated by a scalar field could expand with no decrease in its energy density since $\dot{\rho} = 0$
  - in expanding, the scalar field would do *negative* PdV work
  - i.e. it would *gain* energy
  - and it would do so at just the rate required to keep ρ – and therefore also the expansion rate H – constant in time
  - the inflationary universe is, as Guth says, *'the ultimate free-lunch'*
- the tension in the scalar field in equation is somewhat analogous to that generated in creating a magnetic field by pulling a fridge magnet off a fridge
  - the energy density is proportional to the field squared times the volume the field and the stress along the field lines is of the same magnitude as the energy density

- the electro-magnetic field, however, is a *vector* field – and, as a consequence, the 'stress-tensor' is anisotropic, and the magnetic field has tension along the field lines and positive stress – i.e. positive pressure – in the other two directions
- for a scalar field the stress is isotropic and there is, it turns out, tension in all three directions
- a weakness of this analogy is that in the case of the magnet there is a spatial gradient of the EM potential $\vec{A}$, which is analogous to the field ϕ. But in inflation there is no gradient of ϕ.

3.2 What came before the hot-big bang'?

- The inflationary universe is the answer to the question *"what came before the hot-big bang"?*
  - the 'inflationary scenario' postulates that there was some expanding region of space with roughly uniform and non-time varying scalar field
  - this would rapidly come to dominate the universe
  - during the 'inflationary phase' the acceleration equation is $\ddot{a}/a = -(4\pi/3)G(\rho + 3P/c^2) = (8\pi/3)G\rho$
  - so the universe *accelerates* with $\ddot{a} > 0$ rather than decelerates as is the case for matter or radiation dominated universes as had always been assumed previously
  - and, since ρ is constant, the scale factor obeys $a(t) \propto \exp(Ht)$ with $H =$ constant
- in this scenario, the region of space we can currently observe was, at the onset of inflation, of sub-horizon size
  - the expansion velocity across it was less than c
- but as its size increased (exponentially) with $H = \dot{a}/a$ constant the velocity also increased until, at some point, the velocity reached c and the now observable universe expanded beyond the horizon size
- after some further e-foldings, with smaller (in comoving scale) regions exiting the horizon size, the inflationary phase is assumed to have ended by 'reheating' – with the energy density of the scalar field converted to a thermal plasma – and transitioning to a radiation dominated and then finally a matter dominated era
- this allows the possibility that the very precise uniformity we see on large scales was established by physical processes as the currently observable region – opposite sides of which were, in the big-bang model, always out of causal contact – were, during the inflationary phase, in causal contact
- thus solving the horizon problem (see figure 8)

- the other thing that inflation solves is the fine-tuning problem: if the universe is accelerating (i.e. has $\dot{a} > 0$) then, unlike in the matter or radiation eras where the kinetic ($\dot{a}^2$) and potential energy terms in the Friedmann equation were decreasing, they were (exponentially) increasing with time. Thus, if there were any initial imbalance it would tend to get erased with time.
  - thus an inflationary phase naturally 'prepares' the universe in a state where potential and kinetic energy are nearly perfectly in balance
  - this explains the reluctance on the part of some cosmologists to be reluctant to accept the evidence that $\Omega_{\rm m}$ was close to but still substantially different from unity
  - but the observations stubbornly refused to cooperate

3.3 Fluctuogenesis in the inflationary era

- In the left part of figure 8 we see that the region of space containing a galaxy say (dashed line) can be thought of as appearing when it was only a 'Planck length' $\sim 10^{-32}$m in size.
- We can consider the line labelled $r_{\rm pl}$ to be a boundary of our current ignorance.

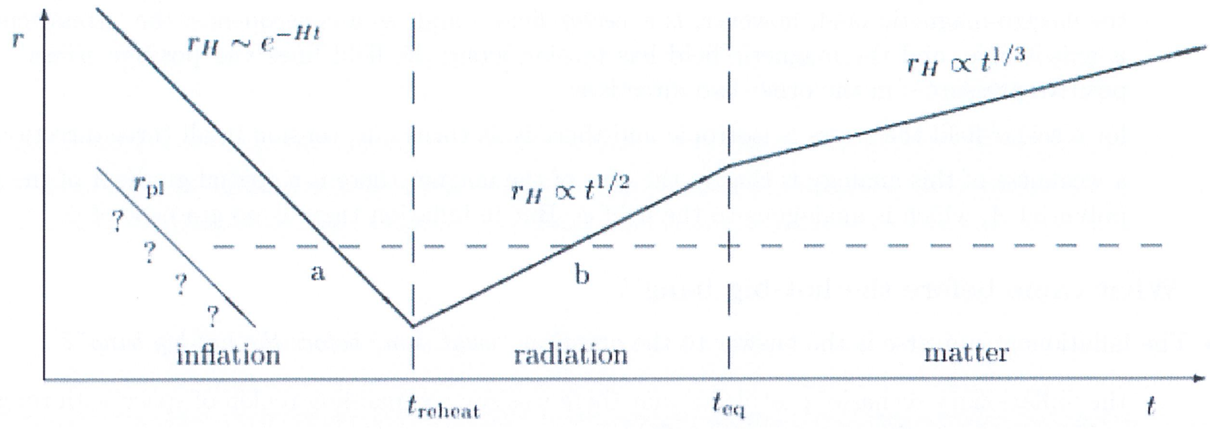


Figure 29.1: The evolution of the comoving horizon scale (heavy solid line) in a Universe which passes through three phases; an inflationary stage followed by a radiation dominated and then a matter dominated era. The ordinate is logarithmic and the abscissa is linear in the inflationary era and logarithmic thereafter. The diagonal line labeled $r_{\rm pl}$ indicates the Planck length. The horizontal dashed line indicates the comoving separation of a pair of observers. This length scale appears at the Planck scale at some time during the inflationary era. We have no adequate description of physics below and to the left of the Planck scale line. The observers then accelerate away from one another. If they were to exchange signals they would perceive an increasing redshift. The separation between the observers reaches the horizon scale at the point labeled 'a'. At that time their relative velocity reaches the speed of light and their relative redshift becomes infinite. Subsequently the observers are unable to exchange signals. At the reheating epoch the Universe starts to decelerate, and at point 'b' the recession velocity falls below the speed of light. The observers then re-appear on each other's horizon; they can exchange signals which are received with steadily decreasing redshift. The separation chosen here is such that it re-enters the horizon during the radiation dominated era. Larger separations enter the horizon at later times. The current horizon scale is $ct_0 \sim c/H_0 \simeq 3000$Mpc. Since the comoving horizon scale is proportional to $t^{1/3}$ in the matter era, the horizon scale at $t_{\rm eq}$ is smaller than the current horizon by a factor ~ 100, or about ~ 30Mpc, the scale of super-clusters. The horizontal dashed line might represent the size of a region encompassing the matter now comprising a galaxy say.

Figure 8: The comoving horizon size vs time in the inflationary scenario.

- All other sized would have appeared at the Planck scale and then expanded exponentially to leave the horizon in the same manner
- so it was immediately recognised that this could potentially allow the generation of fluctuations with the so-called 'Harrison-Zel'dovich' scale invariant spectrum which, if of the correct amplitude ($\delta\rho/\rho \sim 10^{-5}$ at horizon crossing), could form the seeds of structure.
- rapid activity followed with the publication of several papers in 1982 showing that a desirable result could indeed be obtained by appealing to zero-point fluctuations of the inflaton field

3.4 Late Time Inflation and the 'Concordance Model'

- an inflationary phase provides a highly appealing scenario for how the universe got started
  - it would naturally prepare the universe in a state which is spatially flat (i.e. expansion and potential energy in almost perfect balance)
  - and it turns out that the minimal level of quantum fluctuations – zero-point fluctuations – in the field would give rise to something close to a 'scale-invariant' spectrum of density fluctuations (fluctuations in density at horizon-crossing being independent of the scale) which is what is needed for structure formation

- but this leaves us with a puzzle
  - why do we observe $\Omega_m = 0.3$ and not $\Omega_m = 1$?
- one way out of this dilemma would be to question the extrapolation from the mass-to-light ratio of clusters of galaxies to the universe and to assume that the formation of galaxies and stars in clusters of galaxies was *biased* and more efficient, by a factor of a few, relative to the average for the universe
- if so, this could allow $\Omega_m = 1$. In favour of this we might note that:
  - after all, it is observed that the *types* of galaxies are quite different in clusters – dominated by early type galaxies – as compared to the 'field galaxies' which are predominantly spirals
  - it might not then seem hugely implausible that the overall efficiency of creation of starlight was biased
  - another clue is that clusters of galaxies have very strong clustering strength
    * this can be explained quite naturally as a consequence of these objects being identified with high peaks of the initial mass fluctuation field
    * so it is natural to assume that bright galaxies also have an enhanced clustering – and this is indeed that case
- but while this had some appeal, it also had some serious problems
  - one is that it is not hard to understand the bias in types of galaxies in clusters as being caused by star formation in galaxies being extinguished when the galaxies fell in and had gas stripped by ram pressure from the hot intra-cluster medium
  - but this would cause a bias of the mass-to-light in the opposite direction from what is needed to reconcile the data with $\Omega_m = 1$
  - another is that if our universe has $\Omega_m = 1$ then the age of the universe is only $2/3H_0$ which, with $H_0 = 70$km/s/Mpc, is uncomfortably short compared with the ages of the oldest stars
- another way out might be to assume that our universe has $\Omega_m = 0.3$ (and $\Omega_k = 0.7$) – despite inflation – and is actually spatially hyperbolically curved
  - but this is very hard to reconcile with observations of the microwave background anisotropy (see figure 9)
    * the microwave-sky shows ripples on a scale of about a degree that are very well explained as being the imprint of sound-waves in the primordial plasma
    * but this only matches what is observed if the universe has Ω_k – which is just a measure of the mis-match between $\dot{a}^2$ and $(8\pi/3)G\rho a^2$ – very close to zero

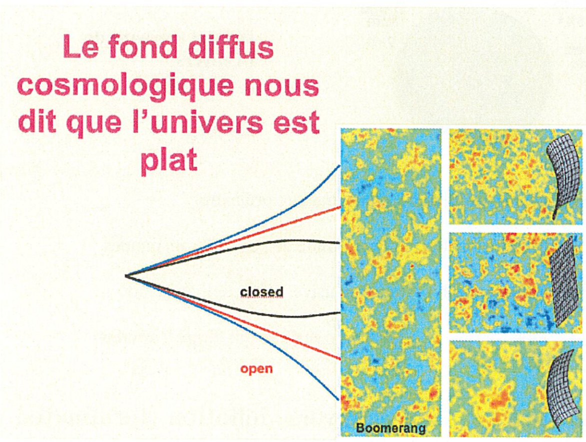
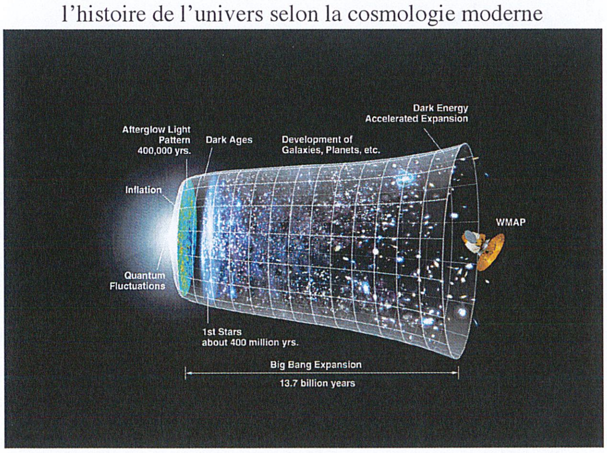


Figure 9: Early universe inflation most naturally predicts a very nearly spatially flat universe. It is possible, however, to solve the age problem, at least, by assuming a universe with only (dark plus baryonic matter) and fit the observations of the amount of matter in clusters etc. One would still have the fine-tuning problem. But this would be very hard to reconcile with CMB observations, as the ripples would appear at too small an angular scale. This observation provided strong evidence that something else like dark energy or a cosmological constant was needed. When the observation of the Hubble diagram from SN1a appeared in 1999, it provided the clincher.

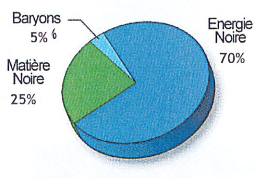
- the solution that cosmologists have come to accept is that our universe is undergoing *late-time inflation* and has just very recently entered into a phase of accelerated expansion

- while this possibility – which was known to be allowed if the simplest form of GR is augmented to include Einstein's *cosmological constant* – was often argued about, the convincing evidence, first presented in 1999, is that the Hubble diagram for type 1a supernovae (white dwarfs in binary systems being slowly fed by their partners until they reach the Chandrasehkar limit and explode) is incompatible with a decelerating universe
- there is still some debate as to whether the data really require that the universal expansion be *accelerating*
  - the problem is that the acceleration has only started rather recently
  - and the data at low redshift are arguably influenced by peculiar motions and need to be treated with caution
- but together with the spatial flatness from the CMB, the SN1a Hubble diagram argues strongly for another constituent to the universe – 'dark energy' (DE) – that has negative pressure and has just come to dominate the total energy density quite recently
  - this might be another scalar field – analogous to the 'inflaton' – and this goes by the name of 'quintessence'
  - or it might be Einstein's 'cosmological constant' Λ
- the cosmological model incorporating this is known as the 'concordance model' or ΛCDM
- in this picture there are 4 important phases in the history of the universe
  - inflation in the early universe – perhaps associated with grand unification
  - the radiation era
  - the matter dominated era
  - late-time inflation, with the universe once again accelerating
- this model seems, so far, to be compatible with many observations
- but it requires, in addition to the known constitutents (the baryons, leptons and photons etc.) not only dark matter but also dark energy and, in the early universe, the inflaton.
- Many cosmologists worry that the apparent existence of these unexplained components – which are forced on us if we believe in Einstein's gravity – are an indication that this theory is wrong, and 'modified gravity' theory is a hot area of research.

l'histoire de l'univers selon la cosmologie moderne



Le contenu de l'Univers



Densité moyenne de l'Univers
$\rho_c \sim 10^{-29}$ **g/cm³**

$\Omega = \rho/\rho_c$

→Baryons= protons, neutrons, matière ordinaire

→ Matière sombre: exotique, faite de particules inconnues

→ Energie sombre: force répulsive, accélère l'expansion

→ Plus l'inflaton — un champ scalaire qui dirige l'inflation

Figure 10: The current model for the universe has 4 significant phases: Early universe inflation (terminated by 'reheating'); the radiation era; the matter era and late-time inflation. With seeds for structure from quantum fluctuations of the inflaton field it is remarkably successful in explaining many observations. But to do so, it requires invoking two mysterious dark components.