Fermi National Accelerator Laboratory

CU-TP-581
FERMILAB-Pub-92/363-A
astro-ph/9302002
submitted to *Physical Review Letters*

INFLATION AS THE UNIQUE SOLUTION TO THE HORIZON AND FLATNESS PROBLEMS

Yue Hu,[1] Michael S. Turner,[2,3] and Erick J. Weinberg[1]

[1]*Department of Physics, Columbia University, New York, NY 10027*

[2]*Departments of Physics and of Astronomy & Astrophysics*
Enrico Fermi Institute, The University of Chicago, Chicago, IL 60637-1433

[3]*NASA/Fermilab Astrophysics Center, Fermi National Accelerator Laboratory, Batavia, IL 60510-0500*

ABSTRACT: We define inflation to be a cosmic scenario characterized by a period of superluminal expansion and massive entropy production. We show, subject to minimal assumptions, that the resolution of the horizon and flatness problems associated with the standard cosmology necessarily involves both features. This rules out a class of adiabatic solutions in which the Planck mass varies by many orders of magnitude, and points to inflation as the unique dynamical solution to horizon and flatness problems.

As successful as the standard cosmology is, it has a number of well known shortcomings: the horizon and flatness problems [1, 2] and, in the context of grand unified theories (GUTs), the monopole problem [3]. While these shortcomings do not involve logical inconsistencies, they indicate that in the standard cosmology the present state of the Universe depends strongly upon the initial state—a feature that many consider undesirable. (It is possible of course, that these cosmological conundrums do not need a dynamical explanation; e.g., Penrose has suggested that there may be a law of physics that governs the initial state of the Universe [4].)

Guth's inflationary Universe scenario provides an elegant solution to these problems involving the microphysics of the very early Universe ($t \lesssim 10^{-34}$ sec and $T \gtrsim 10^{14}$ GeV) [2]. While many implementations of Guth's original idea now exist, all involve two key elements: a period of superluminal expansion (driven by false-vacuum energy) and massive entropy production (conversion of the false-vacuum energy to radiation) [5]. We shall take the presence of these two elements as our definition of inflation. In this *Letter* we show, subject to minimal assumptions, that both elements are required for the solution of the horizon and flatness problems. This rules out the possibility of adiabatic solutions based on a time-varying Planck mass [6, 7, 8], and indicates that inflation is the unique dynamical solution to these problems.

Let us begin with the horizon problem. One of the many ways of stating the horizon problem is that the comoving Hubble radius at early times, H^{-1}/R, was much, much smaller than the present comoving Hubble radius, H_0^{-1}/R_0; here $R(t)$ is the cosmic-scale factor and $H \equiv \dot{R}/R$ is the expansion rate. This precludes causal physics operating at early times from explaining the smoothness of our current Hubble volume [9]. To solve the horizon problem one must arrange to have

$$(R_1H_1)^{-1} \gtrsim (R_0H_0)^{-1}; \tag{1}$$

where subscript 1 refers to some early epoch t_1 and subscript 0 refers to the present epoch [10]. Since $HR = \dot{R}$, this implies that $\dot{R}$ must increase from time t_1 until the present: this can only occur if $\ddot{R}$ is greater than zero at some time between t_1 and t_0, which is the definition of superluminal expansion.

The flatness problem and its resolution are intimately related to the horizon problem. The flatness problem involves the observation that Ω, the ratio of the energy density of the Universe to the critical density, is close to unity today in spite of the fact that $|\Omega - 1|$ grows as a power of the scale factor. Since the curvature radius is proportional to $|\Omega - 1|^{-1/2}$,

$$R_{\rm curv} = R(t)|k|^{-1/2} = \frac{H^{-1}}{|\Omega - 1|^{1/2}}, \tag{2}$$

this implies that the Universe becomes relatively "less flat" with time, provided $k \neq 0$. (In alternative theories of gravity, we take the 3-curvature to be $6k/R(t)^2$ and use Eq. (2) as the definition of Ω.) Suppose the horizon problem is solved, i.e., $(R_1H_1)^{-1} = \beta(R_0H_0)^{-1}$ where $\beta \geq 1$; it follows that

$$|\Omega_0 - 1| = |\Omega_1 - 1|/\beta^2. \tag{3}$$

That is, the deviation of Ω from unity is reduced by the square of the factor by which the horizon problem is solved, insuring that Ω is close to unity today and that the Universe is relatively flat.

We now return to Eq. (1); since at present the Universe is matter dominated with $\rho_{\rm matter} \sim 10^4 T_0^4$, we have $H_0 \sim 10^2 T_0^2/m_{\rm Pl}$. At some earlier time t_1 when the Universe was radiation dominated, $H_1 \sim T_1^2/m_{\rm Pl}$; if radiation was not the dominant contribution to H_1, then H_1 was greater than this. In any case, the horizon/flatness-problem-solving condition Eq. (1) implies

$$\frac{m_{\rm Pl}}{T_1 S_1^{1/3}} \gtrsim 10^{-2} \frac{m_{\rm Pl}}{T_0 S_0^{1/3}}; \qquad (4)$$

where $S \propto (RT)^3$ is the entropy per comoving volume. (We will treat g_*, the number of effectively massless degrees of freedom, as a constant. Closer analysis shows that the variation of g_* does not affect our conclusions unless the initial value of g_* was at least 10^{180}.) Equation (4) makes clear that the horizon and flatness problems can be solved by large-scale entropy production; specifically, $S_0/S_1 \gtrsim (10^{-2}T_1/T_0)^3$. This is the strategy underlying the inflationary solution, in which a period of superluminal (usually exponential) growth of the scale factor is followed by a reheating period characterized by large entropy production.

We now proceed to show that entropy production is in fact necessary. In the standard cosmology, where $m_{\rm Pl}$ is constant, this is easy to see. If the expansion were adiabatic, then $S_1 = S_0$ and Eq. (4) would require $T_1 < 10^2 T_0$, and thus, by adiabaticity, $R_1 > 10^{-2}R_0$. Since the early superluminal epoch t_1 was certainly well before the time of recombination ($R_{\rm rec} \sim 10^{-3}R_0$), this means that R must have decreased at some point. However, this cannot be, since any Friedmann-Robertson-Walker model that is now expanding must always have been expanding, provided only that we make the reasonable assumption that the energy density has always been positive.

It might seem that a variable Planck mass could drastically alter this: A larger Planck mass at early times would imply a weaker effective gravitational constant. This would slow the expansion, leading to an older Universe and a larger horizon [6, 7, 8]. Specifically, the horizon-problem-solving condition would be satisfied without entropy production if it could be arranged that

$$m_{\rm Pl}(t_1) \gtrsim (10^{-2}\, T_1/T_0)\, m_{\rm Pl,0} \sim 10^{30} T_1; \tag{5}$$

where $m_{\rm Pl,0}$ denotes the current value of the Planck mass. $m_{\rm Pl,0} \equiv m_{\rm Pl}(t_0) = G_N^{-1/2} = 1.22 \times 10^{19}\,{\rm GeV}$, and $m_{\rm Pl}(t) = G_N(t)^{-1/2}$ denotes its value as a function of time. (Equation (5) can also be derived from entropy considerations. The current Hubble volume contains an entropy $S_0 \sim H_0^{-3} T_0^3 \sim 10^{88}$ while the entropy within the horizon at early times is $S_t \sim H^{-3} T^3 \sim [m_{\rm Pl}(t)/T]^3 \sim 10^{57} [m_{\rm Pl}(t)/m_{\rm Pl,0}]^3 ({\rm GeV}/T)^3$. To ensure that $S_{t_1} \gtrsim S_0$, we must require that $m_{\rm Pl}(t_1) \gtrsim 10^{10} (T_1/\,{\rm GeV}) m_{\rm Pl,0}$.)

We now show that it is not possible to decrease the Planck mass rapidly enough and thus that an adiabatic solution to the horizon/flatness problems is not possible. Our strategy is to focus on the quantity $T/m_{\rm Pl}$. To reproduce the successful predictions of primordial nucleosynthesis, the Planck mass must have reached its present value by a temperature of order 10 MeV. From this and Eq. (5) it follows that at some early time t_1 the value of $T/m_{\rm Pl}$ must have been smaller than its value at nucleosynthesis by a factor of 10^8 in order to solve the horizon/flatness problems. We now demonstrate that such an increase in $T/m_{\rm Pl}$ from t_1 to $t_{\rm BBN}$ is incompatible with the assumption of adiabatic expansion.

In describing a generic theory with a variable Planck mass we shall represent the Planck mass squared by a Brans-Dicke type field $\Phi = {m_{\rm Pl}}^2$. We write the action in the form

$$S = \int d^4x \sqrt{-g} \left[-\frac{\Phi}{16\pi} \mathcal{R} + \frac{\omega(\Phi)}{16\pi\Phi} \partial_\mu \Phi \partial^\mu \Phi - V(\Phi) + \mathcal{L}_{\rm matter} \right]. \tag{6}$$

The unusual form of the Φ kinetic energy term is not essential: it can be put in the standard form by transforming to a field $\psi(\Phi)$ obeying $(d\psi/d\Phi)^2 =$

$\omega/(8\pi\Phi)$. For constant ω and vanishing $V(\Phi)$ this reduces to the Brans-Dicke theory. We will assume that both the matter energy density ρ and $V(\Phi)$ are non-negative (a negative potential would lead to a negative cosmological constant). It is also reasonable to require that $\omega(\Phi)$ be positive to avoid the instabilities (and quantum mechanical inconsistencies) associated with negative kinetic and gradient energy terms; actually, only the weaker condition $\omega \geq -3/2$ is needed for our purposes here, and this is what we shall assume. We do not consider the possibility of terms of second or higher order in the curvature; for the case of second-order terms, the theory can be reformulated as Einstein gravity with an additional field [11] and an analysis similar to ours can be applied.

This action leads to a Friedmann equation of the form

$$H^2 = \frac{8\pi(\rho + V)}{3\Phi} - H\left(\frac{\dot{\Phi}}{\Phi}\right) + \frac{\omega}{6}\left(\frac{\dot{\Phi}}{\Phi}\right)^2 - \frac{k}{R^2}. \qquad (7)$$

It is convenient to rewrite this as

$$\left(H + \frac{1}{2}\frac{\dot{\Phi}}{\Phi}\right)^2 = \frac{8\pi(\rho + V)}{3\Phi} + \frac{1}{6}\left(\omega + \frac{3}{2}\right)\left(\frac{\dot{\Phi}}{\Phi}\right)^2 - \frac{k}{R^2}. \qquad (8)$$

The quantity appearing on the lefthand side of this equation is

$$H + \frac{1}{2}\frac{\dot{\Phi}}{\Phi} = \frac{d}{dt}\ln(Rm_{\rm Pl}) = -\frac{d}{dt}\ln(T/m_{\rm Pl}); \qquad (9)$$

where the second equality follows from the assumption of adiabaticity. Hence, during epochs where the righthand side of Eq. (8) is nonzero, the quantity $T/m_{\rm Pl}$ must evolve monotonically. For $k \leq 0$, the righthand side can never be negative, and so the variation of $T/m_{\rm Pl}$ is always monotonic. Since $T/m_{\rm Pl}$ has certainly been decreasing since the time of nucleosynthesis, it must always have been doing so; we thus have our result for an open or flat Universe.

The proof for a closed Universe requires a bit more work. With $k > 0$, the righthand side of Eq. (8) has no definite sign, and so the Universe might have alternated between eras of increasing and decreasing $T/m_{\rm Pl}$. Consider the possibility that the current era of decreasing $T/m_{\rm Pl}$ did not extend back to the beginning, but rather began at some time t^*, after the early era of rapid Planck mass variation and before nucleosynthesis. From the vanishing of the righthand side of Eq. (8) at $t = t^*$, we obtain

$$\frac{k}{R^2(t^*)} \geq \frac{8\pi}{3}\frac{\rho_{\rm rad}(t^*)}{m_{\rm Pl}{}^2(t^*)}; \tag{10}$$

These quantities can be related to the corresponding quantities at the time of nucleosynthesis. By adiabaticity $\rho_{\rm rad}(t^*) = \rho_{\rm rad}(t_{\rm BBN})R^4(t_{\rm BBN})/R^4(t^*)$, while the assumption that $T/m_{\rm Pl}$ has been decreasing since $t = t^*$ implies $\rho_{\rm rad}(t^*)/m_{\rm Pl}{}^4(t^*) > \rho_{\rm rad}(t_{\rm BBN})/m_{\rm Pl}{}^4(t_{\rm BBN})$, since $\rho_{\rm rad} \propto T^4$. Substituting into Eq. (10), we obtain

$$\frac{k}{R^2(t_{\rm BBN})} \geq \frac{8\pi}{3}\frac{\rho_{\rm rad}(t_{\rm BBN})}{m_{\rm Pl}{}^2(t_{\rm BBN})}. \tag{11}$$

This last inequality is false, since at the time of nucleosynthesis the curvature term in the Friedmann equation was in fact much smaller than the radiation energy density. Hence, the assumption of the existence of a time t^* must be abandoned, and we have proven our result.

Our proof for a closed Universe illustrates how difficult it is to have $d(T/m_{\rm Pl})/dt$ change sign. Even if one were to relax the assumptions $\omega \geq -3/2$ and the positivity of other contributions to the energy density, the righthand side of Eq. (8) must vanish and the negative contribution must thereafter decrease more rapidly than R^{-4} to guarantee a radiation-dominated Universe at nucleosynthesis. To see how challenging this is, it is instructive to use a conformal transformation to rewrite our action with a constant Planck mass, but time-varying particle masses. In the conformal frame, changing the sign of $d(T/m_{\rm Pl})/dt$ is equivalent to constructing a cosmological model which bounces, i.e., in which $\dot{R}$ changes sign.

The conformal transformation is accomplished by defining a new metric

$$\tilde{g}_{\mu\nu} = \frac{\Phi}{{m_{\rm Pl,0}}^2} g_{\mu\nu} \tag{12}$$

where $m_{\rm Pl,0} \equiv m_{\rm Pl}(t_0)$ is the present value of the Planck mass. When expressed in terms of this metric and the corresponding Ricci scalar $\tilde{\mathcal{R}}$, the action of Eq. (6) becomes (after an integration by parts)

$$\tilde{S} = \int d^4x \sqrt{-\tilde{g}} \left[-\frac{{m_{\rm Pl,0}}^2}{16\pi} \tilde{\mathcal{R}} \quad + \quad \left(\omega + \frac{3}{2}\right) \frac{{m_{\rm Pl,0}}^2}{16\pi\Phi^2} \partial_\mu \Phi \partial^\mu \Phi \right.$$
$$\left. - \quad \frac{{m_{\rm Pl,0}}^4}{\Phi^2} V(\Phi) + \frac{{m_{\rm Pl,0}}^4}{\Phi^2} \tilde{\mathcal{L}}_{\rm matter} \right] \tag{13}$$

where indices are raised and lowered with the new metric and, because of the metric factors which it contains, $\tilde{\mathcal{L}}_{\rm matter}$ has further Φ-dependence not displayed explicitly. If rewritten in terms of the field $\chi = \ln \Phi$, this would be the usual dilaton action.

Since the gravitational part of the action has the usual form, the energy momentum-tensor of the Φ field can be immediately read off. For a spatially homogeneous Φ field, the energy density and pressure are

$$\rho_\Phi = \left(\omega + \frac{3}{2}\right) \frac{{m_{\rm Pl,0}}^2}{16\pi\Phi^2} \left(\frac{d\Phi}{d\tilde{t}}\right)^2 + \frac{{m_{\rm Pl,0}}^4}{\Phi^2} V(\Phi); \tag{14}$$

$$p_\Phi = \left(\omega + \frac{3}{2}\right) \frac{{m_{\rm Pl,0}}^2}{16\pi\Phi^2} \left(\frac{d\Phi}{d\tilde{t}}\right)^2 - \frac{{m_{\rm Pl,0}}^4}{\Phi^2} V(\Phi). \tag{15}$$

Furthermore, the time and scale factor for the transformed metric are related to those for the original metric by $d\tilde{t}/dt = \sqrt{\Phi/{m_{\rm Pl,0}}^2}$ and $\tilde{R} = \sqrt{\Phi/{m_{\rm Pl,0}}^2}\, R$.

so the Hubble parameter for the transformed metric is

$$\tilde{H} \equiv \frac{1}{\tilde{R}} \frac{d\tilde{R}}{d\tilde{t}} = \frac{m_{\rm Pl,0}}{\sqrt{\Phi}} \left(H + \frac{1}{2} \frac{\dot{\Phi}}{\Phi} \right). \tag{16}$$

We now recognize Eq. (8) as the Friedmann equation, in standard form, in the new frame. Moreover, $\tilde{H}$ is, up to to a positive numerical factor, equal to $-d\ln(T/m_{\rm Pl})/dt$. If $d(T/m_{\rm Pl})/dt$ is to change sign, the expansion rate in the conformal frame must vanish and change sign—$\tilde{R}$ must undergo a bounce.

In retrospect, it is actually quite natural that superluminal expansion should be followed by entropy production. Recall that the energy density of a fluid with equation of state $p = \gamma\rho$ evolves as $\rho \propto R^{-3(1+\gamma)}$ and that if it dominates the energy density of the Universe the scale factor evolves as $t^{2/3(1+\gamma)}$. Superluminal expansion requires that the effective equation of state have $\gamma < -1/3$, which implies that ρ decreases more slowly than R^{-2}. During the superluminal phase, the energy density of the "fluid" that drives inflation increases as R^2 (or faster) relative to the radiation energy density and as R (or faster) relative to the matter density. Since we can be confident that the latter phases of the evolution of the Universe involve a radiation-dominated phase (from a temperature at least as high as 10 MeV until a temperature of about 10 eV) and a matter-dominated phase (from a temperature of about 10 eV until the present or close to it), this fluid must eventually "decay" into radiation, thereby increasing the entropy by a very large amount.

Finally, suppose for a moment that it were somehow possible for $T/m_{\rm Pl}$ to increase as required to make an adiabatic solution viable. We would argue that the flatness problem is still not solved in a truly satisfactory way. The flatness problem involves the size of the curvature radius relative to the Hubble radius at the initial epoch: In the standard cosmology, the apparently "natural" initial condition that $R_{\rm curv}(t_i) = R(t_i)|k|^{-1/2}$ be comparable to the Hubble radius H_i^{-1} leads to a Universe that quickly recollapses for $k > 0$, or goes into free expansion, $R(t) \propto t$ for $k < 0$, neither of which is consistent with our Universe. The survival of the Universe to at least the ripe old age of 10 Gyr before becoming curvature dominated requires the "unnatural" initial condition $R_{\rm curv}(t_i) \gtrsim (10^{30}T_i/m_{\rm Pl})H_i^{-1}$. In a variable-Planck-mass model the curvature radius is initially (i.e., at t_1) comparable to the Hubble radius; however, the price of achieving this is an initial temperature that is thirty orders of magnitude smaller than the initial value of the Planck mass, i.e. $T_1 \sim 10^{-30}m_{\rm Pl}(t_1)$. Thus, an unnatural initial choice of curvature radius has simply been traded for an unnatural choice of initial temperature. In fact, had we chosen the initial epoch for standard cosmology by this criteria, $T_i \sim 10^{-30}m_{\rm Pl}$, there would be no horizon/flatness problem since the initial epoch would essentially be today.

In sum, we have shown that the search for a dynamical resolution of the horizon and flatness problems associated with the standard cosmology naturally leads to both superluminal expansion and massive entropy production, the two generic features of inflation. This precludes adiabatic solutions that attempt to solve the horizon/flatness by a large variation in the Planck mass [12]. Further, subject to our minimal assumptions, it implies that inflation is the unique dynamical solution to horizon and flatness problems.

We thank J. Levin and K. Freese for extensive discussions. We understand that they have independently arrived at some of the conclusions of this paper. We also thank J. Harvey for helpful conversations. This work was supported in part by the DOE (at Chicago, Columbia, and Fermilab) and by NASA through NAGW-2381 (at Fermilab).

References

[1] R.H. Dicke and P.J.E. Peebles, in *General Relativity: An Einstein Centenary Survey*, edited by S.W. Hawking and W. Israel (Cambridge Univ. Press, Cambridge, 1979), p. 504.

[2] A.H. Guth, *Phys. Rev. D* **23**, 347 (1981).

[3] J. Preskill, *Ann. Rev. Nucl. Part. Sci.* **34**, 461 (1984).

[4] R. Penrose, in *General Relativity: An Einstein Centenary Survey*, edited by S.W. Hawking and W. Israel (Cambridge Univ. Press, Cambridge, 1979), p. 581.

[5] For a textbook discussion of inflation see e.g., E.W. Kolb and M.S. Turner, *The Early Universe* (Addison-Wesley, Redwood City, CA, 1990), Ch. 8.

[6] A. Zee, *Phys. Rev. Lett.* **44**, 703 (1980).

[7] See e.g., M.V. Fischetti, J.B. Hartle, and B.-L. Hu, *Phys. Rev. D* **20**, 1757 (1979); P. Anderson, *ibid* **28**, 271 (1983).

[8] K. Freese and J.J. Levin, Univ. of Mich. preprint UM AC 92-16 (1992); J.J. Levin and K. Freese, Univ. of Mich. preprint UM AC 92-15 (1992).

[9] In any Robertson-Walker cosmology without superluminal expansion the Hubble radius H^{-1} is, up to a numerical factor, equal to the horizon distance, $d_H(t) = R(t)\int_0^t dt'/R(t')$.

[10] It may seem paradoxical to discuss the horizon problem in the context of a Robertson-Walker model, which, by assumption, is homogeneous. The underlying idea is to view an inhomogeneous cosmological model as being comprised of locally smooth patches of size comparable to the local Hubble radius. Once superluminal expansion sets in, the use of a Robertson-Walker model to analyze the subsequent behaviour of the patch is well justified. See e.g., A.A. Starobinskii, *JETP Lett.* **37**, 66 (1983); L. Jensen and J. Stein-Schabes, *Phys. Rev. D* **35**, 1146 (1987).

[11] A.A. Starobinskii, *Phys. Lett. B* **91**, 99 (1980).

[12] Of course it is still possible that the Planck mass has varied during the evolution of the Universe; in fact, this occurs in an interesting model of inflation, D. La and P.J. Steinhardt, *Phys. Rev. Lett.* **62**, 376 (1989).