

Cosmic bounces and cyclic universes

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2011 Class. Quantum Grav. 28 204004

(<http://iopscience.iop.org/0264-9381/28/20/204004>)

[View the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 194.94.224.254

The article was downloaded on 05/01/2012 at 11:34

Please note that [terms and conditions apply](#).

Cosmic bounces and cyclic universes

Jean-Luc Lehners

Max-Planck-Institute for Gravitational Physics (Albert-Einstein-Institute),
D-14476 Potsdam/Golm, Germany

E-mail: jlehners@aei.mpg.de

Received 29 March 2011, in final form 27 May 2011

Published 3 October 2011

Online at stacks.iop.org/CQG/28/204004

Abstract

Cosmological models involving a bounce from a contracting to an expanding universe can address the standard cosmological puzzles and generate ‘primordial’ density perturbations without the need for inflation. Some such models, in particular the ekpyrotic and cyclic models that we focus on, fit rather naturally into string theory. We discuss a number of topics related to these models: the reasoning that leads to the ekpyrotic phase, the predictions for upcoming observations, the differences between singular and non-singular models of the bounce as well as the predictive and explanatory power offered by these models.

PACS numbers: 98.80.-k, 98.80.Bp, 95.36.+x, 11.25.-w

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Our universe is special in that it evolved from an *a priori* highly unlikely state, in the sense of Boltzmann: the very early universe was exceptionally flat, homogeneous and isotropic (hence it had very low entropy), with approximately scale-invariant density perturbations of magnitude $Q \sim 10^{-5}$ added on top. Now it may be that string theory, once better understood, leads to a unique initial state with these characteristics, but there is currently little evidence pointing in that direction. Therefore, we would like to find a *dynamical* explanation for these ‘initial’ conditions.

The most popular attempt at formulating such a model is inflationary cosmology [1–3]. In recent years, many attempts have been made at obtaining inflationary dynamics in string theory (see for example the contributions by McAllister and Burgess, Quevedo and Mazumdar to this focus issue). All models to date require rather elaborate constructions (with multiple branes of diverse dimensionalities, orientifold planes and warped geometries induced by p -form fluxes all carefully balanced) together with their own set of initial conditions. Thus, currently, these models simply report the question of initial conditions to an earlier epoch, perhaps alleviating it, but not solving it. What these models do achieve is that they provide a mechanism with a

microphysical basis for generating the primordial density perturbations, and this is their main merit. All models to date still have to be fine-tuned in order for the characteristics of the generated perturbations, such as their amplitude and spectrum, to agree with observations (see also the contribution by Mulryne and Ward). This unfortunately reduces the explanatory power of inflationary models. A recourse that is sometimes advocated is to invoke eternal inflation [4] (which creates an infinite number of universes of all possible configurations—hence it will also create universes with the elaborate characteristics necessary for ordinary inflation to occur) together with anthropic arguments [5] (which ‘justify’ disregarding all universes with characteristics substantially different from ours). Such an approach may however turn out to be fallacious: eternal inflation leads to infinities, which must be regulated. However, all results turn out to depend on the regulator, i.e. eternal inflation does not make regulator-independent predictions (this is the so-called measure problem of inflationary cosmology, see the contributions by Freivogel [77] and Kleban [78] to this focus issue). Hence, one may suspect that eternal inflation is not a consistent theory¹, and we will adopt this point of view in this paper. In this respect, it certainly seems worthwhile to examine alternative dynamical models for the early universe. In this paper, we will discuss models that are based on a cosmic bounce, i.e. models in which the current expanding phase of the universe was preceded by a contracting phase. We will see that such models, in particular the *ekpyrotic* [8] and *cyclic* [9] models of the universe, may provide a framework that can explain the ‘initial’ conditions of the universe dynamically, without recourse to anthropic arguments. Moreover, these models make rather distinctive predictions for upcoming cosmological observations, which we will discuss.

The original ekpyrotic model was inspired by string theory [8], in particular by the Hořava–Witten theory [10, 11]. We start by reviewing a simplified version of the string theoretic setup in question (section 2), and build up the logic that leads to a cyclic universe from there. In brief, the logic goes as follows: the big bang may be seen as a collision of branes in higher dimensions, but where the scale factor on the branes does not shrink to zero size, only the distance between the branes goes to zero. Hence, the big bang may have been a much milder, and potentially tractable, event. If so, the universe would have been in a contracting phase previously. However, a contracting universe is unstable to the buildup of anisotropies and the onset of chaotic mixmaster behavior, unless there is a matter component in the universe with an ultra-stiff equation of state. This matter component is ‘ekpyrotic’ matter (section 3), and it can be modeled as a scalar field with a steep and negative potential (this scalar may be the scalar parameterizing the distance between the branes; its potential would then describe an attractive inter-brane force). Surprisingly, this ekpyrotic phase not only manages to flatten and isotropize the universe, but it also amplifies quantum fluctuations which may obtain a scale-invariant spectrum via the entropic mechanism reviewed below (section 4). As the potential is steep, the scalar is strongly self-coupled, and the resulting perturbations are distinguished by large non-Gaussianities of the local form. However, as for inflation, for the ekpyrotic phase to be successful, the universe must again have had the right initial conditions, so it looks like we might not have made any progress at all. This is where the idea of a cyclic universe comes in (section 6): it turns out that by joining the ekpyrotic phase to the future of the current dark energy phase, it becomes manifest that the current dark energy phase sets up precisely the right conditions for a subsequent ekpyrotic phase to be successful. This then leads to the idea of a cyclic universe, in which the most recent ekpyrotic phase (during which our perturbations were generated) was also preceded by a dark energy phase, etc. Toward the end of this paper, we will present different models for the all-important bounce phase (section 5) and discuss the reasons why such a cyclic framework may be powerfully predictive (section 6).

¹ Albrecht [6] and Banks [7] have suggested that holography may prevent eternal inflation from occurring.

2. Colliding branes and the big bang

Consider pure gravity in five dimensions. Motivated by the Hořava–Witten theory [10, 11], we want to look for a solution where one spatial dimension consists of a line segment. A line segment can also be described as the quotient, or *orbifold*, of a circle by a Z_2 reflection symmetry across a diagonal of the circle. Under the action of the Z_2 reflection symmetry, the two endpoints of the line segment are fixed points—from the five-dimensional point of view these define the location of two orbifold planes, which are $(3+1)$ -dimensional boundaries of spacetime. We can write the metric as

$$ds_5^2 = e^{-\sqrt{2/3}\phi} ds_4^2 + e^{2\sqrt{2/3}\phi} dy^2, \quad (1)$$

where $-y_0 < y < y_0$ denotes the orbifold coordinate and ϕ is the radion field parameterizing the size of the orbifold. The ϕ -dependent prefactor in front of the four-dimensional metric ds_4^2 ensures that after dimensional reduction we are left with the canonically normalized Lagrangian²

$$\mathcal{L} = \sqrt{-g} \left[\frac{1}{2} R - \frac{1}{2} (\partial\phi)^2 \right]. \quad (2)$$

Thus, the equations of motion become

$$3H^2 = \frac{1}{2}\dot{\phi}^2 = -\dot{H} \quad \ddot{\phi} + 3H\dot{\phi} = 0. \quad (3)$$

They imply that $a \propto e^{\phi/\sqrt{6}}$ and are solved by

$$a = a_0(-t)^{1/3} \quad \phi = \sqrt{\frac{2}{3}} \ln(-t) + \phi_0, \quad (4)$$

for some integration constants a_0 , ϕ_0 . This means that we can now plug solution (4) into the original five-dimensional metric to find (up to obvious re-scalings)

$$ds_5^2 = (-t)^{-2/3} \left[-dt^2 + (-t)^{2/3} dx_3^2 \right] + (-t)^{4/3} dy^2 \quad (5)$$

$$= -dT^2 + T^2 dy^2 + dx_3^2, \quad (6)$$

where we have changed the time coordinate to $T \propto (-t)^{2/3}$. This model spacetime, sometimes called the compactified Milne spacetime, describes two (empty) orbifold planes approaching each other, colliding and receding away from each other again. Under the further change of coordinates $u = T \cosh y$, $v = T \sinh y$, the metric becomes simply that of a five-dimensional Minkowski space

$$ds_5^2 = -du^2 + dv^2 + dx_3^2, \quad (7)$$

which shows that this spacetime is flat, except at the moment of collision ($T = 0$), see also figure 1. From the four-dimensional point of view, the moment of collision corresponds to $a(t) = 0$, i.e. it corresponds to the big bang. However, as seen from the full five-dimensional theory, the singularity corresponds to only the orbifold dimension shrinking to zero size, and not the other spatial dimensions. Indeed, as can be seen from (1), the brane scale factors are really given by $e^{-\phi/\sqrt{6}}a$, and they are constant and non-zero at the collision. The density of matter (which is stuck on the branes and subdominant regarding the dynamics of the collision) as well as the temperature remain finite at the collision.

This simple example shows that, in a higher dimensional context, the big bang singularity may be much milder than in four-dimensional general relativity, and may be tractable. We will discuss models of cosmic bounces (both singular, like this one, and non-singular) in more detail in section 5. For now, let us just go with the idea that the big bang may not have been

² We use reduced Planck units $\hbar = c = 1$ and $8\pi G = M_{Pl}^{-2} = 1$.

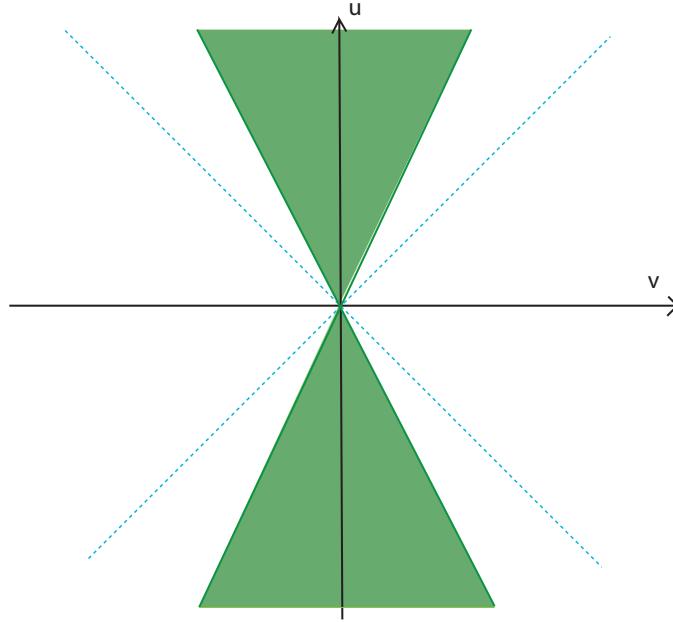


Figure 1. The compactified Milne universe (the shaded region) is a useful model for a big bang caused by the collision of two boundary branes. This figure illustrates the embedding in five-dimensional Minkowski space, the three ordinary spatial dimensions being suppressed and with the dashed lines delineating the light cone at the moment of the collision.

the beginning of the universe, and that there may have been a previous phase during which the universe contracted. This immediately leads to a potential problem, namely the wild growth of anisotropies, combined with a chaotic mixmaster big crunch, that is expected to occur in a contracting universe. In such a case, we evidently would not expect a flat and isotropic universe to re-emerge from the crunch. One way to avoid this would be for the universe to be exceptionally flat at the start of the contracting phase, but this would amount once more to imposing special initial conditions—the very thing we want to explain. But there is another possibility, described in [12], which involves the addition of a new matter component. We discuss it next.

3. The ekpyrotic phase

For concreteness, consider a Friedmann–Robertson–Walker (FRW) metric

$$ds^2 = -dt^2 + a(t)^2 \left(\frac{dr^2}{1 - \kappa r^2} + r^2 d\Omega_2^2 \right), \quad (8)$$

where $a(t)$ denotes the scale factor of the universe and $\kappa = -1, 0, 1$ for an open, flat or closed universe respectively. If the universe is filled with a number of fluids interacting only via gravity and with energy densities ρ_i and constant equations of state w_i , then the equations of continuity

$$\dot{\rho}_i + 3\frac{\dot{a}}{a}(\rho_i + p_i) = 0 \quad (9)$$

(where dots denote derivatives with respect to time t) imply that they will evolve according to

$$\rho_i \propto a^{-3(1+w_i)}. \quad (10)$$

The Einstein equations lead to the Friedmann equation, which involves the Hubble parameter $H \equiv \dot{a}/a$:

$$H^2 = \frac{1}{3} \left(\frac{-3\kappa}{a^2} + \frac{\rho_{m,0}}{a^3} + \frac{\rho_{r,0}}{a^4} + \frac{\rho_{a,0}}{a^6} + \dots + \frac{\rho_{\phi,0}}{a^{3(1+w_\phi)}} \right). \quad (11)$$

The $\rho_{i,0}$ s are constants giving the energy densities at scale factor $a = 1$ of the various constituents of the universe: non-relativistic matter (subscript m), radiation (r) and the energy density associated with anisotropies in the curvature of the universe (a). There is also a term due to the mean curvature of space, and we assume the presence of a scalar field ϕ .

In a contracting universe, as $a \rightarrow 0$, the term scaling with the largest negative power of a will eventually come to dominate. And this will be the anisotropy term $\propto a^{-6}$, unless the scalar field has an equation of state $w_\phi > 1$. In that case, the scalar field will come to dominate and suppress the anisotropies, as well as the mean curvature, matter and radiation. Such a phase is called an *ekpyrotic* phase. Since the equation of state of a scalar field is given by

$$w_\phi = \frac{p}{\rho} = \frac{\frac{1}{2}\dot{\phi}^2 - V(\phi)}{\frac{1}{2}\dot{\phi}^2 + V(\phi)}, \quad (12)$$

we can see that $w_\phi > 1$ can be achieved by having a negative potential $V(\phi)$. A simple, but relevant example is the negative exponential

$$V(\phi) = -V_0 e^{-c\phi}, \quad (13)$$

where V_0 and c are constants. As mentioned previously, in string theory such a scalar could be the radion, i.e. the scalar field describing the distance between the two boundary branes, and in that case a negative exponential potential can be generated by instanton effects, see e.g. [13, 14]. Once the scalar comes to dominate the dynamics, the equations of motion reduce to

$$3H^2 = \frac{1}{2}\dot{\phi}^2 - V_0 e^{-c\phi} \quad \ddot{\phi} + 3H\dot{\phi} - cV_0 e^{-c\phi} = 0, \quad (14)$$

and they are solved by

$$a = (-t)^{1/\epsilon}, \quad \phi = \sqrt{\frac{2}{\epsilon}} \ln(-\sqrt{\epsilon V_0} t), \quad \epsilon = \frac{c^2}{2}. \quad (15)$$

The equation of state is

$$w_\phi = \frac{2\epsilon}{3} - 1 = \frac{c^2}{3} - 1, \quad (16)$$

so we need $c > \sqrt{6}$ for ekpyrosis. In realistic models (realistic in the sense of agreeing with observations), ϵ typically turns out to be of $\mathcal{O}(100)$. Note that the same parameter, in inflationary models, is the ‘slow-roll’ parameter and is typically of $\mathcal{O}(1/100)$. Here, ϵ is rather a ‘fast-roll’ parameter. Note that when ϵ is large, the potential is steep and the universe is contracting very slowly.

4. Predictions for observations

As described above, the ekpyrotic phase suppresses the mean curvature and anisotropies, and so, if this phase lasts long enough (typically a fraction of a second is already enough, see e.g. [15]), the flatness problem is solved. Incidentally, the horizon problem is automatically solved too, as, during the contracting phase, there is ample time for the different regions

we observe in the cosmic microwave background (CMB) to have been in causal contact. However, the ekpyrotic phase not only resolves the standard cosmological puzzles, it also generates density perturbations by amplifying quantum fluctuations [16], much like inflation, but with some crucial differences in the details. Here, we simply summarize the results—the detailed derivations were recently reviewed in [17].

During the ekpyrotic phase, the universe is contracting very slowly, but the horizon $1/H \propto t \rightarrow 0$ is shrinking fast. Hence, similarly to what happens during inflation, quantum fluctuations with progressively smaller wavelengths exit the horizon and turn into classical density perturbations. Now it turns out that if there is a single scalar field, with a steep and negative potential, driving the ekpyrotic phase, then the resulting curvature fluctuations have a blue spectrum (with spectral index $n_s \approx 3$) [16, 18–20]³. On large scales, these perturbations are completely subdominant (although they could make a re-appearance during the bounce phase, as discussed further on) and hence these cannot give rise to the temperature fluctuations we observe in the CMB. But if there is more than one ekpyrotic scalar present, scale-invariant curvature perturbations can be generated via the *entropic mechanism* [23–25]. In fact, from the point of view of string theory, it is very natural to expect more than one scalar. For example, in the colliding brane setup described above, there was one scalar field describing the distance between the two orbifold planes along a fifth dimension. However, in the full string/M-theory setting, we would expect there to be six additional internal dimensions, for example, in the shape of a Calabi–Yau manifold [26]. The volume of the Calabi–Yau manifold is then described by an additional scalar, and many more scalars are required to describe all the shape deformations of the internal space. Since the inclusion of one additional scalar already entails the needed qualitative differences to the single-field case, we will focus on having two scalars for simplicity, with action

$$S = \int \sqrt{-g} \left[R - \frac{1}{2}(\partial\phi_1)^2 - \frac{1}{2}(\partial\phi_2)^2 - V(\phi_1, \phi_2) \right]. \quad (17)$$

Let us assume also that both fields get an ekpyrotic-type potential:

$$V(\phi_1, \phi_2) = -V_1 e^{-c_1\phi_1} - V_2 e^{-c_2\phi_2}. \quad (18)$$

Then, it is much more natural to discuss the dynamics in terms of the new variables σ and s pointing transverse and perpendicular to the field velocity respectively [27, 28]; they are defined, up to unimportant additive constants which we will fix below, via

$$\sigma \equiv \frac{\dot{\phi}_1\phi_1 + \dot{\phi}_2\phi_2}{\dot{\sigma}}, \quad s \equiv \frac{\dot{\phi}_1\phi_2 - \dot{\phi}_2\phi_1}{\dot{\sigma}}, \quad (19)$$

with $\dot{\sigma} \equiv (\dot{\phi}_1^2 + \dot{\phi}_2^2)^{1/2}$. For later use, we also define the angle θ of the trajectory in field space, via [29]

$$\cos \theta = \frac{\dot{\phi}_1}{\dot{\sigma}}, \quad \sin \theta = \frac{\dot{\phi}_2}{\dot{\sigma}}. \quad (20)$$

In terms of these new variables, the potential can be re-expressed as

$$V_{ek} = -V_0 e^{\sqrt{2}\epsilon\sigma} \left[1 + \epsilon s^2 + \frac{\kappa_3}{3!} \epsilon^{3/2} s^3 + \frac{\kappa_4}{4!} \epsilon^2 s^4 + \dots \right], \quad (21)$$

where for exact exponentials of the form (18), one has $\kappa_3 = 2\sqrt{2}(c_1^2 - c_2^2)/|c_1 c_2|$ and $\kappa_4 = 4(c_1^6 + c_2^6)/(c_1^2 c_2^2 (c_1^2 + c_2^2))$. However, in the absence of a microphysical derivation of the

³ Right at the onset of the ekpyrotic phase, a small range of scale-invariant curvature perturbations are produced by the *adiabatic mechanism* [21, 22]. However, it has so far not been possible to incorporate this mechanism into a full ekpyrotic or cyclic model of the universe.

potential, we will simply take $\kappa_3, \kappa_4 \sim \mathcal{O}(1)$ and express all results in terms of κ_3, κ_4 . The ekpyrotic scaling solution becomes

$$a(t) = (-t)^{1/\epsilon} \quad \sigma = -\sqrt{\frac{2}{\epsilon}} \ln(-\sqrt{\epsilon V_0} t) \quad s = 0, \quad (22)$$

with the angle θ being constant. Thus, we can see that, up to a sign determined by convention, the field σ now plays the role of the single scalar we discussed earlier, and hence this is the field that picks up a blue spectrum of curvature perturbations. The transverse field s , on the other hand, feels a different potential. Because of the overall minus sign in (21), we can see that the potential for s is actually *unstable* [25, 30], and the motion during the ekpyrotic phase is along a ridge in the potential. This has profound consequences, which we will discuss in section 6. In the present context, the instability implies that the s field fluctuations δs , which are entropy (or isocurvature) fluctuations, get amplified. More detailed considerations show that they obtain an amplitude Q_s and a spectral index n_s given by [25]

$$Q_s \approx (\epsilon V_{\min})^{1/2} \quad (23)$$

$$n_s - 1 = \frac{2}{\epsilon} - \frac{\epsilon_{,N}}{\epsilon^2}, \quad (24)$$

where V_{\min} denotes the value of the ekpyrotic potential at the end of the ekpyrotic phase. The fast-roll parameter ϵ is now allowed to vary slowly and $dN \equiv d \ln a$. Hence, for ϵ of $\mathcal{O}(10^2)$ (or more), the spectrum is close to scale-invariant, with the first term on the right-hand side tending to make the spectrum blue and the second term tending to make it red⁴. A simple estimate, based on [31], of the natural range of n_s gives the range $0.97 < n_s < 1.02$ in good agreement with current data.

In this way, the ekpyrotic phase generates nearly scale-invariant entropy perturbations. But what we are really interested in are the curvature perturbations ζ , which are the perturbations that give rise to the temperature fluctuations in the CMB. In that respect it is useful to examine the evolution equation for the (linear) curvature perturbations on large scales, given by [29]

$$\dot{\zeta} = -\frac{2H}{\dot{\sigma}} \dot{\theta} \delta s. \quad (25)$$

As soon as the background trajectory bends ($\dot{\theta} \neq 0$), the entropy perturbations source the curvature perturbations, implying that on large scales the curvature perturbations thus created obtain the same spectrum as the entropy fluctuations (there is no wavelength dependence in (25)), with an amplitude Q_ζ of [17]

$$Q_\zeta^2 \approx \frac{\epsilon V_{\min}}{10^3}. \quad (26)$$

We will discuss the amplitude in much more detail in section 6. In the most concrete models of the entropic mechanism studied in the literature (reviewed in section 5), the indispensable bending of the trajectory occurs automatically. These models show that the bending occurs after the ekpyrotic phase has come to an end, in the phase leading up to the bounce, when the ekpyrotic potential has become unimportant and the dynamics is dominated by the kinetic energy of the σ field⁵.

The curvature perturbations created in this way are indistinguishable from those created by a phase of inflation, for example. However, at higher orders, the ekpyrotic perturbations are

⁴ Note that during the ekpyrotic phase, N is decreasing, and so must ϵ in order for the ekpyrotic phase to come to an end eventually.

⁵ It was also previously envisaged that the trajectory could bend due to the instability of the potential [27, 32]. This case is now disfavored, as it leads to predictions for non-Gaussianities in disagreement with observations at the 4σ level, see [17, 33].

characterized by sizeable and specific non-Gaussian signatures [17, 33–35]. There is a simple reason for this: the potential is necessarily steep during ekpyrosis, so that the scalar fields are necessarily significantly self-coupled, leading to non-Gaussian statistics for the perturbations. It turns out that these non-Gaussian corrections are of the so-called local form. If the full, nonlinear curvature perturbations ζ are re-written as a linear, Gaussian piece ζ_g , plus correction terms,

$$\zeta = \zeta_g + \frac{3}{5} f_{\text{NL}} \zeta_g^2 + \frac{9}{25} g_{\text{NL}} \zeta_g^3, \quad (27)$$

then the predictions for ekpyrotic curvature perturbations generated via the entropic mechanism are [36–39]

$$f_{\text{NL}} = \frac{3}{2} \kappa_3 \sqrt{\epsilon} + 5 \quad (28)$$

$$g_{\text{NL}} = \left(\frac{5}{3} \kappa_4 + \frac{5}{4} \kappa_3^2 - 40 \right) \epsilon. \quad (29)$$

With κ_3, κ_4 expected to be of $\mathcal{O}(1)$, and ϵ typically of $\mathcal{O}(10^2)$ (ϵ needs to be above about 50 in order for the power spectrum to be in agreement with observations), we would thus expect the parameter f_{NL} to be of $\mathcal{O}(\pm 10)$, with the sign being typically determined by the sign of κ_3 , and g_{NL} of $\mathcal{O}(-1000)$ and *negative* in sign. The current observational bounds are that $f_{\text{NL}} = 38 \pm 21$, where the errors are quoted at one standard deviation [40], while currently no strong constraints exist on g_{NL} . It will therefore be extremely interesting to see the improved data likely to be gathered by the PLANCK satellite and ongoing and future large scale surveys, especially in light of the fact that the ekpyrotic predictions are significantly different than those of simple inflationary models.

Before proceeding to discuss the bounce in more detail, we should also discuss the predictions for primordial gravitational waves [8, 41, 42]. In fact, a quick heuristic argument is sufficient to get at the main result: at the linear level, the generation of gravitational waves depends entirely on the behavior of the scale factor. And during the ekpyrotic phase, the scale factor shrinks very slowly, so that, to a first approximation, the background is Minkowski space. This immediately implies that no significant large-scale gravitational waves are generated during ekpyrosis (otherwise gravitational waves would also be produced all around us all the time), in stark contrast to large-field inflationary models. It turns out that, at second order in perturbation theory, small-amplitude gravitational waves are sourced by the scalar density perturbations, and these are the dominant gravitational wave signal expected from an ekpyrotic phase [43], though they are unlikely to be observable in the near future. Hence, a detection of large-amplitude primordial gravitational waves would rather conclusively rule out the models described here.

These predictions are interesting, because they are quite specific: reasonably large f_{NL} , reasonably large and negative g_{NL} plus an absence of primordial gravitational waves. It is clear that inflationary models can also be constructed with exactly these predictions; hence, even if future observations remain in agreement with the predictions, no clear-cut conclusions can be drawn. However, inflationary models with exactly these predictions would have to be rather contrived and fine-tuned, and would have to be adjusted *a posteriori*, which would render them disfavored in the author's opinion (and according to standard results in probability theory).

5. Cosmic bounces: singular or non-singular?

In cosmological models involving a reversal from contraction to expansion, the bounce phase is the most crucial ingredient. This is the moment that gets identified with the ‘big bang’,

and it is the time when quantum gravity effects are likely to play a dominant role. Since we currently do not possess an experimentally supported theory of quantum gravity, all current models are necessarily rather speculative. Nevertheless, using current candidate theories, in particular string theory, we can construct and analyze various models for a bounce, and explore their consequences.

During the contracting phase, the Hubble rate is negative, while it is positive during the subsequent expanding phase of the universe. Hence, the bounce must allow the Hubble rate H to increase. If we describe the dominant energy component in the universe as a perfect fluid with energy density ρ and pressure p , then the Einstein equations imply that

$$\dot{H} = -\frac{1}{2}(\rho + p). \quad (30)$$

Obtaining the required increase in the Hubble rate can then happen in only two ways: either $\rho + p < 0$, i.e. the null energy condition (NEC) is violated, or the bounce must be classically singular with the Hubble rate instantaneously jumping from negative to positive values. Of course, especially in the second case, it becomes crucially important that the quantum corrections to the classical description be understood. We will discuss models of both types below. Before doing so, it may be useful to point out that the case of a singular bounce is not a limit of a classically non-singular bounce. In the non-singular case, the scale factor of the universe must slow down its contraction, come to a halt and then increase again. This is typically a very fine-tuned process. By contrast, in the singular case, there is typically no slowing down, as will be evident in the example described below. Hence, we should not expect these two types of models to lead to equivalent physical results.

5.1. Classically singular bounces

The most interesting example of a classically singular bounce is the collision of the orbifold branes in the Hořava–Witten theory/heterotic M-theory, a simplified version of which has already been described in section 2. Here, we will elaborate on this process, and list the reasons for why this example is particularly encouraging.

Hořava–Witten theory describes the strong coupling limit of the ten-dimensional $E_8 \times E_8$ heterotic string theory [10, 11]. The value of the string coupling is determined by a field (the ϕ field of section 2) which gets identified with the size of a new spatial dimension. As the coupling constant grows, this 11th dimension grows too, and in this newly opened space the theory becomes that of 11-dimensional supergravity. The endpoints of the new 11th dimension consist of two orbifold fixed planes, which are endowed with an E_8 gauge theory each. If in addition six of the spatial dimensions are compactified on a Calabi–Yau manifold, then the theory, now usually called *heterotic M-theory*, contains $(3+1)$ -dimensional boundary branes with realistic particle physics on them [26, 44, 45]. Restricting our attention to gravity and the Calabi–Yau volume modulus V_{CY} only (these are the only fields that are relevant close to the bounce), the action is

$$S_5 = \frac{1}{2\kappa_5^2} \int_{5d} \sqrt{-g} \left[R - \frac{1}{2} V_{\text{CY}}^{-2} \partial_m V_{\text{CY}} \partial^m V_{\text{CY}} - 6\alpha^2 V_{\text{CY}}^{-2} \right] + \frac{1}{2\kappa_5^2} \left\{ -12\alpha \int_{4d, y=+1} \sqrt{-g} V_{\text{CY}}^{-1} + 12\alpha \int_{4d, y=-1} \sqrt{-g} V_{\text{CY}}^{-1} \right\}, \quad (31)$$

where α denotes the boundary brane tension. The static vacuum solution to this theory is

$$\begin{aligned} ds^2 &= h^{2/5}(y)[A^2(-d\tau^2 + d\vec{x}^2) + B^2 dy^2], \\ V_{\text{CY}} &= Bh^{6/5}(y), \\ h(y) &= 5\alpha y + C, \quad -1 \leq y \leq +1, \end{aligned} \quad (32)$$

where A , B and C are integration constants. The y coordinate spans the orbifold with fixed points at $y = \pm 1$ (we have re-scaled y compared to section 2), a negative-tension brane being located at $y = -1$ and a positive-tension brane at $y = +1$. As we will justify below, we are interested in the situation in which the branes are moving slowly. In order to describe that case, we can simply promote the moduli A , B , C to functions of (conformal) time τ . Then, after the field redefinitions

$$a^2 \equiv A^2 B I_{\frac{3}{5}}, \quad (33)$$

$$e^{\phi_1/\sqrt{2}} \equiv B(I_{\frac{3}{5}})^{3/4}, \quad (34)$$

$$\phi_2 \equiv -\frac{\sqrt{6}}{20} \int dC(I_{\frac{3}{5}})^{-1} [9(I_{-\frac{2}{5}})^2 + 16I_{-\frac{7}{5}}I_{\frac{3}{5}}]^{1/2}, \quad (35)$$

$$I_n = \int_{-1}^1 dy h^n = \frac{1}{5\alpha(n+1)} [(C+5\alpha)^{(n+1)} - (C-5\alpha)^{(n+1)}], \quad (36)$$

the action simplifies considerably to [46]

$$S_{\text{mod}} = \int_{4d} \left[-3a'^2 + \frac{1}{2}a^2(\phi_1'^2 + \phi_2'^2) \right], \quad (37)$$

where $' \equiv \frac{d}{d\tau}$. This is simply gravity with scale factor a minimally coupled to two scalars, just like the theory we used to describe the entropic mechanism in section 4, except that we have written it out here on a FRW background and in terms of conformal time. Higher dimensional quantities are then determined in terms of the fields a , ϕ_1 , ϕ_2 , for example the distance d between the branes and the values of the brane scale factors a_{\pm} at the locations of the boundaries at $y = \pm 1$:

$$d = \frac{1}{3(2\alpha)^{1/4}} e^{\phi_1/\sqrt{2}-\sqrt{3/2}\phi_2} [(1 + e^{2\sqrt{2/3}\phi_2})^{3/2} - |1 - e^{2\sqrt{2/3}\phi_2}|^{3/2}], \quad (38)$$

$$a_{\pm} = (2\alpha)^{1/8} a e^{-\phi_1/2\sqrt{2}} \begin{cases} (\cosh \sqrt{2/3}\phi_2)^{1/4} \\ (-\sinh \sqrt{2/3}\phi_2)^{1/4}. \end{cases} \quad (39)$$

The axis $\phi_2 = 0$ is special: positive values of ϕ_2 would lead to unphysical imaginary values for the brane scale factor a_- , and are thus forbidden. One can show that in the presence of matter on the boundary branes, there arises an effective repulsive potential along the ϕ_2 axis, which causes the field space trajectory to bend in its vicinity [46, 47]. This is of course just what is required for the conversion of entropy into curvature perturbations described in section 4, and the example here provides a microphysical basis for the entropic mechanism.

The effective theory (37) contains two scalar fields, unlike the toy model of section 2. However, by requiring that at the brane collision the brane scale factors do not shrink to zero or blow up to infinity, we are forced to look at solutions where close to the collision $\phi_1 - \sqrt{3}\phi_2$ approaches a constant. This effectively reduces the number of scalars to 1, when looking at the background evolution (for the perturbations, it remains of course important that the theory really contains two scalars). The theory (37) is then solved by

$$a = |2y_0\tau|^{1/2}, \quad e^{\phi_1} = |2y_0\tau|^{3/2\sqrt{2}}, \quad e^{\phi_2} = |4\alpha y_0\tau|^{\sqrt{3}/2\sqrt{2}}, \quad (40)$$

where $2y_0$ has the interpretation of being the velocity of the branes at the collision, as (38) implies that $d \approx 2y_0\tau$ for small τ . Note that y_0 is the only parameter of the solution.

Close to the collision, the brane tension α becomes irrelevant and these solutions approach the compactified Milne solution described in section 2, with the speed at collision being given

by $2y_0$. Their most important feature is that in the approach to the collision the spacetime that they describe becomes increasingly flat (of course, at the moment of collision there is still a δ -function curvature singularity). Hence, higher derivative corrections induced by quantum loops become more and more negligible too [48]! This is similar in spirit to a conjecture due to Penrose, which states that quantum corrections are small when the Weyl curvature is small (as is the case here too) [49]. So, even though this solution breaks supersymmetry⁶, the quantum corrections can be expected to be mild. This hope is given further support by the fact that the distance between the orbifold branes also determines the string coupling constant, as seen from ten dimensions. As the branes approach each other, this coupling decreases, and it is zero at the moment of collision. Thus, during the bounce, we expect interactions to be significantly suppressed. This can be made more quantitative by studying certain aspects of the brane collision semi-classically. In particular, membranes that wind around the compact dimension are expected to become light as the size of the 11th dimension shrinks. Thus, such states must be taken into account close to the collision. Turok *et al* have shown that the equations of motion for winding M2-branes remain regular at all times, and that they propagate smoothly through the collision [51]—see also [52–54]. Instanton techniques then enable one to estimate the production of such light membranes at the collision. The result is that the density of produced membranes is finite and their gravitational backreaction is small, provided the collision speed y_0 remains non-relativistic. This last finding has interesting consequences regarding the amplitude of the cosmological perturbations that are produced in these models, as will be discussed in section 6.

Eventually, one would like to be able to treat the collision of the orbifold planes of heterotic M-theory fully quantum mechanically. Unfortunately, this seems to require that a full non-perturbative definition of M-theory be found first. This drawback notwithstanding, it is interesting that all indications found so far imply that this particular example of a singular bounce may well be viable.

5.2. Classically non-singular bounces

As we saw at the beginning of this section, in order for the universe to bounce in a non-singular way, the dominant energy component in the universe must violate the NEC. From a model-building point of view, this is difficult to achieve, as violations of the NEC are commonly associated with various instabilities in the theory (most notably the appearance of ghosts, i.e. fields with negative kinetic energy). Moreover, since the energy density scales as $a^{-3(1+w)}$ for matter with equation of state w , and since violations of the NEC correspond to $w < -1$, we can see immediately that in a contracting universe such matter becomes quickly less and less relevant. In fact, as discussed in section 3, at the end of the ekpyrotic phase all other types of matter have become vastly subdominant to the ekpyrotic scalar field. Hence, the only option seems to be that the ekpyrotic scalar itself must turn into the required NEC-violating matter [32]. At that point, our options are currently rather limited: there exist only two types of scalar field theories in which the NEC can be violated in a stable and controlled way: ghost condensates [55] and Galileons [56–58]⁷. We will mainly focus on the most studied case, that of a ghost condensate bounce, which has been used in ‘new ekpyrotic’ models [32], see also [35, 60]. It works as follows: suppose that, as the kinetic energy for the scalar field ϕ becomes very small, its kinetic term becomes non-canonical. In fact, if we denote $-\frac{1}{2m^4}(\partial\phi)^2 \equiv X$, so

⁶ Even the exactly flat compactified Milne solution breaks supersymmetry, because the fermions are not invariant under the boost identifications that need to be performed in order for one dimension to be a line segment [50].

⁷ In fact, it has recently been discovered in [59] that both types of theories are closely related: they are virtually identical when one considers time dependence alone, the main differences residing in spatial gradient terms.

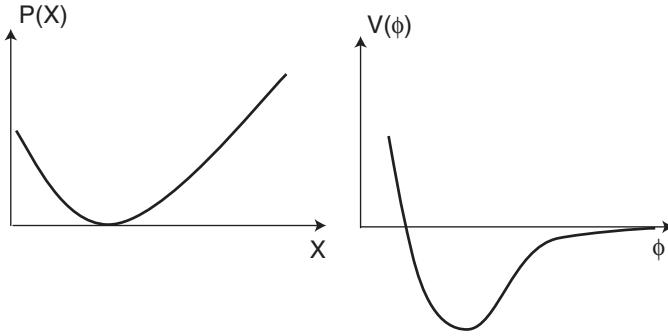


Figure 2. The kinetic and potential functions for the ghost condensate bounce, as used in new ekpyrotic models.

that the ordinary kinetic term in the Lagrangian is simply X , then we imagine that, for small X , the Lagrangian (which we denote by $\mathcal{L} = P(X)$) is of the form

$$\mathcal{L}_{\text{small } X} = P(X)_{\text{small } X} = M^4(-X + X^2), \quad (41)$$

i.e. it has a local minimum. Here, m and M are two mass scales, which we will discuss in more detail below. The full kinetic function is sketched in figure 2.

When the scalar reaches the minimum, the vacuum shifts, in the sense that the scalar develops a time-dependent vacuum expectation value (vev) that is growing linearly with time, $\langle \phi \rangle = m^2 t$. The Lagrangian, perturbed to quadratic order around this time-dependent vev, is proportional to

$$\mathcal{L} \propto (2XP_{,XX} + P_{,X})(\dot{\phi})^2 - P_{,X}(\nabla \delta \phi)^2. \quad (42)$$

Amazingly, the perturbations in the vicinity of the minimum ($P_{,X} \approx 0$) are not ghost-like, despite the $-X$ term in the Lagrangian. The energy density and pressure are given by

$$\rho = 2P_{,X}X - P, \quad p = P, \quad (43)$$

so that $\rho + p = 2P_{,X}X$. Thus, at the minimum, the ghost condensate is right on the verge of violating the NEC, its equation of state being that of a cosmological constant $p = -\rho$. But if the field is pushed to even smaller X values, where $P_{,X} < 0$, then it violates the NEC, *without the appearance of ghosts*, and can cause the universe to revert smoothly from contraction to expansion.

For such a scenario to work, a number of conditions must be met: first, there needs to be a steeply rising potential to slow the scalar down after the end of the ekpyrotic phase, in order for the ghost condensate phase to start—see figure 2 for a sketch of the shape of the potential that is required. In fact, as described in [32], the potential must rise roughly as $V \propto -\frac{1}{m^2}\phi$ for consistency.

Secondly, even though there are no ghosts, there are still gradient instabilities when the NEC is violated, as is apparent from the last term in (42). However, as shown in [32, 57], these are harmless as long as the bounce is fast (on the order of one e-fold), as there is then no time for the instability to grow. But having a fast bounce requires the parameters of the model, in particular the mass scales m and M , to be rather finely tuned. It turns out that, for an ekpyrotic phase with the fast-roll parameter ϵ , lasting for N_{ek} e-folds and ending at a potential value of V_{\min} , the mass scales must satisfy the hierarchy

$$m^4 e^{2N_{ek}} \ll |V_{\min}| \ll \epsilon M^4. \quad (44)$$

In other words, the effective field theory cut-off scale M must be at or above the scale of the ekpyrotic potential, with the scale m of the ghost condensate vev many orders of magnitude below. Such a tuning does not seem unreasonable.

There is one further requirement that the ghost condensate model must satisfy, and it stems from a potential problem that has recently been discovered by Xue and Steinhardt, and which actually threatens to invalidate all non-singular models of a bounce [61]. The problem arises from an unlikely source, namely an apparently doubly subdominant cosmological perturbation mode. Let us temporarily neglect the entropy perturbations that are thought to source the primordial density fluctuations in the present model. Then, as discussed at the beginning of section 4, the dominant perturbations have a blue spectrum and ultimately provide a negligible contribution to the observed temperature fluctuations in the CMB. In slightly more detail, the comoving curvature perturbation ζ obeys the equation of motion (in Fourier space and conformal time)

$$\zeta_k'' + 2\frac{z'}{z}\zeta_k' + k^2\zeta_k = 0, \quad (45)$$

where $z \equiv a\sqrt{2\epsilon/c_s^2}$. c_s denotes the speed of sound of the fluctuations, and can be read off from (42):

$$c_s^2 = \frac{P_{,X}}{2XP_{,XX} + P_{,X}}. \quad (46)$$

On large scales, i.e. for small k , the equation of motion is solved by

$$\zeta_k = \frac{c_1}{\sqrt{k}} + c_2\sqrt{k}\int \frac{d\tau}{z^2}, \quad (47)$$

where c_1 and c_2 are constants of $\mathcal{O}(1)$, as can be seen by matching to the Minkowski vacuum in the far past. The first mode is the one with the blue spectrum $n_s = 3$, and which can be neglected. The second mode has an even bluer spectrum, $n_s = 5$, and is typically not even mentioned. But it is instructive to re-write the integral in (47) as

$$\frac{\zeta_k^{\text{int}}}{c_2\sqrt{k}} \equiv \int \frac{d\tau}{z^2} = \int \frac{dt c_s^2}{2a^3\epsilon} = \int \frac{dt c_s^2}{3a^3(1+w)}. \quad (48)$$

During the bounce, we must have $w < -1$, and so it looks like the integral might blow up in the approach to the bounce phase, which would signal a breakdown of perturbation theory. However, the speed of sound also goes to zero as w reaches -1 , and so the integral is finite. It is nevertheless dominated by the contribution close to the bounce phase, and can be evaluated using (43), (46) and the approximation of neglecting the friction term in the equation for X , $\dot{X} \approx -V_{,\phi}\sqrt{X}$:

$$\frac{\zeta_k^{\text{int}}}{c_2\sqrt{k}} = \int \frac{2XP_{,X} - P}{2a^3X(2XP_{,XX} + P_{,X})} \frac{dX}{(-V_{,\phi})\sqrt{X}} \approx \frac{-1}{V_{,\phi}\sqrt{X_{\text{bounce}}}}. \quad (49)$$

Since the speed of the scalar at the bounce $\sqrt{X_{\text{bounce}}}$ is necessarily very small, we can see that in the approach to the bounce, this initially doubly subdominant term gets massively amplified, and threatens to dominate over all other perturbation modes. As shown in [61], the ratio of this mode to the curvature perturbation ζ_s produced by the entropic mechanism is

$$\frac{\zeta_k^{\text{int}}}{\zeta_s} \approx \frac{-1}{V_{,\phi}} e^{N_{ek} - 2N_k}, \quad (50)$$

where N_k denotes the number of e-folds of ekpyrosis remaining after mode k of ζ^{int} has exited the horizon. For modes of observational interest $N_k \approx (20 - 30)$, while $N_{ek} > 60$, so for

many models of non-singular bounces this effect will be problematic. However, as we saw above, in the present model the bounce must be fast in order to avoid gradient instabilities, and this means that the potential must be steep, $|V_{,\phi}| \approx 1/m^2 \gg 1$. In fact, the consistency relation (44) implies that for the model discussed here, $m^2 e^{N_{ek}}$ is always small, and so the doubly blue mode remains subdominant. Hence, the condition that the bounce be *fast* avoids the onset of all known potential instabilities.

From a model-building point of view, it is very satisfying to see that a consistent non-singular bounce model can be constructed⁸. What is not so clear yet is whether such a model could also arise from string theory. There exist some general arguments to the contrary [63, 64], but history has shown that no-go arguments have a tendency to be overcome. In a recent effort to clarify this question, the ghost condensate model has been supersymmetrized [65], which may help in elucidating the connection to string theory. It would also be interesting to see if non-singular bounce models can be constructed using the closely related Galileon theories. These theories do not suffer from gradient instabilities, and so might lead to more general bounce models. The relation of these models to string theory is equally unsettled at present. They are thought not to arise in string theory by the arguments of [63], but on the other hand, they can be derived, very suggestively, from the dynamics of branes [66]. They have also been supersymmetrized recently [59], and so these questions may find a clarification soon.

5.3. Nonperturbative bounces

The AdS/CFT correspondence is thought to provide a non-perturbative definition of a theory of quantum gravity. This makes one wonder whether, in this framework, one can finally tackle the question of cosmological singularities in a full quantum mechanical way, and study bounces non-perturbatively. In this vein, Craps and Evnin [79] have attempted to construct models of a crunch of anti-de Sitter space [67–69], which are described in detail in his contribution to this issue. Hence, we will limit ourselves to just a few remarks here to point out the main differences with the colliding brane bounce discussed above: the AdS crunch corresponds to a high curvature crunch, and thus can be expected to lead to overproduction of particles at the crunch, with subsequent rapid re-collapse. Numerical studies in the context of a five-dimensional model confirm this expectation [70], but there is the hope that a better behavior of the coupling constant in the dual conformal field theory for the four-dimensional case might evade this problem. Also, in the models studied so far, there is no ekpyrotic matter present, and thus there is no matter component that can halt the rapid growth of anisotropies. Thus, the models in question may describe the quantum version of chaotic mixmaster collapse. This in itself will be very interesting, but at the same time it will also be of interest to see if the framework can be extended to low-curvature crunches more like that of compactified Milne space.

5.4. Reheating

In all models of a cosmic bounce, the ordinary radiation and matter content of the universe is thought to arise at the bounce itself. For singular bounces, one expects the brane collision to be slightly inelastic, so that a small part of the collision energy will be converted into radiation and matter [9]. Note that in that case the scalar fields do not disappear by decaying into radiation and matter, as is typically assumed in inflation. Rather, the scalar fields become

⁸ See the review [62] for other attempts at constructing bounce models.

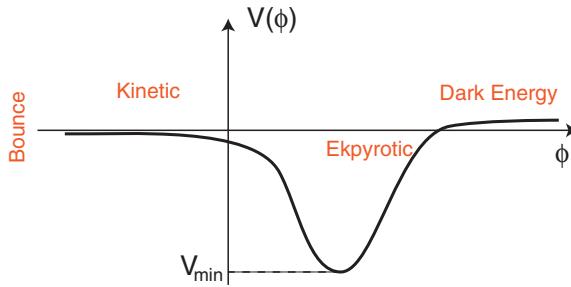


Figure 3. The effective potential in the cyclic universe, with the different phases of evolution.

cosmologically relevant again at later times, when they act as quintessence, and their non-decay is essential in the construction of a cyclic model of the universe, as detailed in the next section. For non-singular bounces on the other hand, the reheating mechanism most discussed so far involves the decay of the field governing the bounce, e.g. the ghost condensate field, into radiation and matter just after the bounce [32]. Such a mechanism is most suitable for single-bounce models of the universe, but appears more difficult to incorporate into a cyclic model.

It seems fair to say that the topic of reheating has not been studied in sufficient detail and generality in the context of cosmic bounces; this is certainly an area that deserves further research.

6. Cosmic cycles, the phoenix and predictivity

By having a contracting phase before the current expanding phase, have we not simply pushed the question of initial conditions back to an earlier time, in much the same way as the theory of inflation does, though with the added complication of a bounce? Then all we have gained is a new mechanism for generating cosmological perturbations (though we already had one with inflation), but no answers to any other basic cosmological questions. However, imagine that we link up alternating contracting and expanding phases into a cyclic universe, could it then be that one cycle prepares the initial conditions for the next one? Or, assuming that there was some original big bang-like event, but which did not produce a universe like ours—could subsequent cycles evolve more and more into universes like ours? If so, such an attractor framework might be highly predictive, much more so than a theory which produces all types of universes, but with no clear preference for one like ours⁹. The cyclic model which we will discuss now is an attempt to build precisely such a model.

The cyclic universe, first proposed by Steinhardt and Turok [9] posits that the current dark energy phase that we find ourselves in will be followed in the distant future by a contracting ekpyrotic phase, which will lead to a bounce and a new phase of radiation and matter domination, followed once more by dark energy, and so forth. As an effective theory, this can be described by a scalar field moving in the potential shown in figure 3. The figure shows the potential only in one direction, and one should keep in mind that other, transverse, directions can also be relevant: in particular, during the ekpyrotic phase, there is a transverse unstable potential, as described in section 4, which will play an important role here.

⁹ This kind of argument can be made precise, see for example [71].

The cyclic model is formulated as an effective theory, without specifying a precise microphysical origin, although the most precise formulation to date is based on the colliding brane picture stemming from heterotic M-theory¹⁰. It is useful to briefly go over the different phases of the model, and to follow the evolution of the scale factor and the Hubble parameter [72]. Let us start with the current dark energy phase, when the scalar field takes its largest values. During this phase, the Hubble parameter is roughly constant and the universe grows by N_{de} e-folds as the scalar field is very slowly rolling to smaller field values. Eventually, the potential becomes negative, the field speeds up, the universe starts to contract again and a new ekpyrotic phase takes place. From section 3 we know that the scale factor now remains roughly constant, while the Hubble parameter grows by $e^{N_{ek}} = V_{\min}^{1/2}/V_0^{1/2}$, where V_0 denotes the current dark energy density. The ekpyrotic phase comes to an end when the scalar reaches the bottom of the potential at V_{\min} , and is followed by the bounce phase. If we assume the colliding brane model of the bounce, the dynamics will be dominated by the kinetic energy of the scalar, with the scalar shooting off to $-\infty$ and bouncing back after the brane collision. By the end of the bounce/kinetic phase, the universe has grown by the modest amount $V_{\min}^{1/6}/T_r^{2/3}$, where T_r denotes the temperature to which the universe reheats at the brane collision. Meanwhile, the Hubble parameter has shrunk by a factor $T_r^2/V_{\min}^{1/2}$. After the bounce, the scalar field flies back over the potential well, slows down due to Hubble friction and comes to a halt on the plateau at large field values. Right before the dark energy phase, while the energy density of the universe is still dominated by radiation and matter, the universe grows by a factor of about T_r/T_0 , with the Hubble parameter shrinking by T_0^2/T_r^2 , where $T_0 = V_0^{1/4}$ is the current temperature of the universe. Thus, putting everything together, the Hubble parameter returns to its starting value after one cycle (as do all such locally measurable quantities), while the scale factor grows by the large amount of

$$\left(\frac{V_{\min}^{1/4}}{T_r}\right)^{2/3} \left(\frac{T_r}{T_0}\right) e^{N_{de}}, \quad (51)$$

over the course of one cycle. This net growth is essential to the viability of the cyclic universe. For one, it solves a puzzle about the second law of thermodynamics, which asks how a cyclic universe is possible, given that the entropy must increase all the time. The resolution here is that, although the total entropy does increase continually, the entropy *density* gets diluted to the same small value again after each cycle.

We should now go back to the ekpyrotic phase, and address the instability of the potential that is necessary for the generation of perturbations. Cyclic universe and instability are two concepts that do not seem to go very well together *a priori*. However, this instability may actually lie at the heart of the high level of predictivity of the model. As described in section 4, during the ekpyrotic phase the background trajectory of interest is the one that rolls unstably along a ridge in the potential, with the potential falling off steeply both left and right. For slightly different initial conditions, the ekpyrotic phase will be of short duration: the trajectory will veer off to the side, ekpyrosis and the associated flattening of the universe will end prematurely and the corresponding region of the universe will collapse in a chaotic mixmaster crunch. At the end of such a crunch, we will conjecture that a gravitationally collapsed region remains, in the form of a black hole presumably, so that this region will stop growing. Only a region with very special field values at the start of the ekpyrotic phase will make it all the way down the ridge through the bounce and on to the next expanding phase. However, that region, unlike all the other collapsed regions, will grow by the enormous factor (51) until the start

¹⁰ It is at present an open question whether a consistent cyclic model of the universe can also be constructed using a non-singular bounce of the ghost condensate type, for example.

of the next ekpyrotic phase. Moreover, for $N_{de} > 60$, it turns out that this newly expanded universe is bigger than its originator universe one cycle before [73]. Thus, as long as the dark energy phase lasts sufficiently long (sufficiently long meaning at least 600 billion years), the cyclic universe grows from cycle to cycle and *selects* regions of the universe with just the right initial conditions to produce a large, flat and habitable space during the next cycle. In this context, it is wrong to say that the ekpyrotic phase requires special initial conditions. Rather, this *phoenix universe* [73, 74], reminiscent in spirit of the phoenix universe discussed by Lemaitre [75] and Dicke, produces a vast range of ‘initial’ conditions before the onset of the ekpyrotic phase¹¹. Almost the entire universe will then collapse irredeemably, except for some small regions, which are necessarily exceptionally flat, homogeneous, isotropic but with small quantum-generated density perturbations, growing into the kind of universe we see around us. And should the universe that is produced in such a way happen to be small (compared to the presently observed Hubble volume, say), then this is hardly an obstacle, as each subsequent cycle will make it bigger and bigger. Unlike eternal inflation, the phoenix universe renders the vast majority of space habitable. Note also that it is not necessary for the cycles to have occurred indefinitely into the past. A beginning of the universe, via a tunneling event for example, is entirely conceivable in this framework, as long as such an event produces one region with the right conditions for the ekpyrotic phase to proceed successfully. Whether the universe had a beginning or not is unfortunately difficult to establish experimentally in this framework. For better or worse, the instability implies a near total isolation from even the immediately preceding cycle, with only the quantum-generated fluctuations surviving the bounce. Any small excess energy present at the onset of the last ekpyrotic phase in the region that is supposed to evolve into our observable universe would automatically cause the corresponding field trajectory to fall off the potential ridge and prevent that region from making it through the bounce. On the positive side, this means that there is no paradox in us not having found messages across the sky from far-advanced civilizations of the last cycle(s).

As speculative as some of these discussions may sound, the cyclic universe described above leads to concrete pre- and postdictions, which can be confronted by experiment. The most important predictions, regarding non-Gaussianities in the CMB and gravitational waves, were described in section 4. But there is also an important postdiction concerning the amplitude Q of the primordial density perturbations [76]. In other models of the early universe, notably inflationary models, this quantity is always adjusted by hand to fit the observed value of $Q_{\text{obs}} \sim 10^{-5}$, even though the observed value is typically not favored from a theoretical point of view. By contrast, the cyclic model (with a bounce modeled by colliding branes) *dynamically selects* regions of the universe with the observed value of Q ! This is entirely analogous in spirit to the selection of flat regions of the universe discussed in the previous paragraph. Since this is a significant result, it is worth deriving it in a little bit more detail. The ekpyrotic phase evolves according to the scaling solution (22), or, using conformal time,

$$a = (-\tau)^{1/\epsilon}, \quad \frac{d\phi}{d\tau} = \sqrt{2}/(\sqrt{\epsilon}\tau), \quad V = -\epsilon\mathcal{H}^2, \quad (52)$$

with the conformal Hubble rate $\mathcal{H} \equiv (\frac{da}{d\tau})/a$. The subsequent kinetic phase is characterized by

$$a = (-\tau)^{1/2}, \quad \frac{d\phi}{d\tau} = \sqrt{3}/(\sqrt{2}\tau). \quad (53)$$

¹¹ Note that, given the assumed shape of the potential, this range of ‘initial’ conditions does not however include conditions which would allow eternal inflation to occur. If this were the case, we would of course face the measurement problem once again.

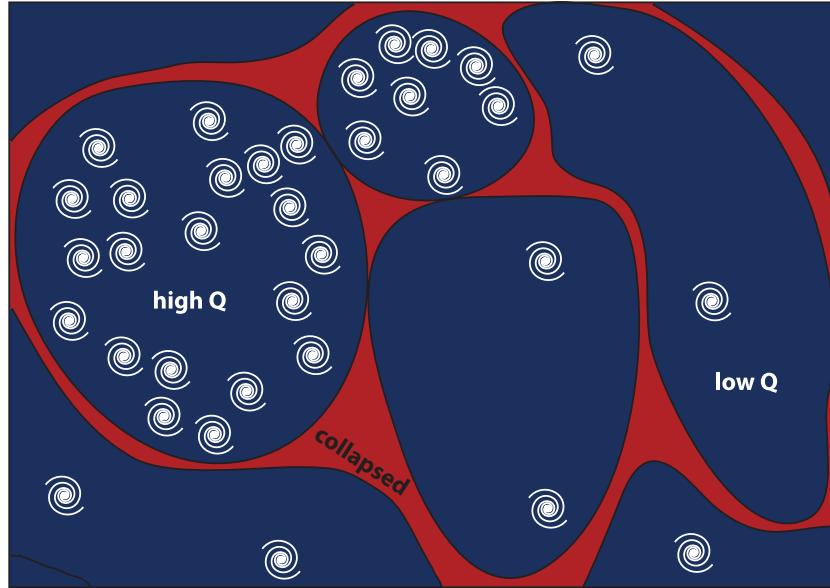


Figure 4. The global structure of the cyclic universe.

Matching the conformal Hubble rates in the two phases at the time τ_k that the kinetic phase begins implies

$$-2\tau_k = \sqrt{\epsilon/V_{\min}}. \quad (54)$$

Recalling also from the discussion below equation (40) that the inter-brane distance is given by

$$d = 2y_0|\tau_k|, \quad (55)$$

we get the relation $V_{\min} = \epsilon y_0^2/d^2$. Important additional information is that the extension of solution (40) to the ekpyrotic phase implies that the inter-brane distance is approximately constant during that phase, with the value d given above (see [46, 76]). Hence, the amplitude Q_ζ of the curvature perturbations generated by the entropic mechanism, given above in (26), can be re-expressed very simply as

$$Q_\zeta \approx \frac{\epsilon y_0}{10^{3/2} d} \approx \frac{y_0}{d}, \quad (56)$$

where we have taken $\epsilon \sim \mathcal{O}(10^2)$ both because this is consistent with observations of the spectral index, and because, from a theoretical point of view, the lowest allowed values for ϵ correspond to the least amount of fine-tuning in the potential. For d , we can assume the value $d \approx 10^{3.5}$ suggested by the Hořava–Witten theory, and which implies the observed value of Newton's constant. The crux of the argument comes from considering y_0 , the brane collision velocity. As discussed in section 5, it cannot be arbitrarily high, as this would lead to an overproduction of matter (wrapped M2-branes) at the brane collision, with a subsequent rapid collapse of the corresponding region. In other words, if y_0 is relativistic in some region, then that region undergoes gravitational collapse at the bounce, and becomes irrelevant regarding the large-scale evolution of the cyclic universe in much the same way as those regions with the

wrong initial conditions for ekpyrosis described above. Thus, the vast majority of the universe must have $y_0 \lesssim 0.1$, or

$$Q_\zeta \lesssim 10^{-4.5}. \quad (57)$$

Regions in which the brane collision velocity is higher lead to a larger Q_ζ and thus to more structure, making it highly likely to find oneself in a region of the universe with an observed value of Q as close as possible to the upper bound (57), see also figure 4. It is certainly intriguing that the observed value of Q agrees so closely with this postdicted value!

To conclude, it seems worthwhile to dwell for a little longer on the main theme discussed in this last section, namely the contrast between the instability of the ekpyrotic phase and the net expansion over the course of each cycle. The discarding of large regions of the universe at the end of each cycle provides a type of dissipation, which is just what is needed to have attractor behavior of the model. Hence, perhaps somewhat counter-intuitively, it is precisely the interplay between *fragility* and *amplification* that lends the model its high predictive power. If it was easy to make all sorts of large, reasonably flat universes, with all sorts of density perturbations, all allowing for the formation of gravitationally bound structures, it would be very difficult to understand the broad features of our universe. We would have to conclude that our part of the universe is as it is because of sheer coincidence. Before accepting such a conclusion though, we should test those models in which the specialness of our universe is not due to mere historical accident, but due to mathematical and physical necessity.

Acknowledgments

I would like to thank Thorsten Battefeld, Justin Khoury, Jens Niemeyer, Burt Ovrut, Paul Steinhardt, Neil Turok and Bingkan Xue for illuminating discussions on the topics discussed in this paper. I also gratefully acknowledge the European Research Council's support in the form of a Starting Grant.

References

- [1] Guth A H 1981 *Phys. Rev. D* **23** 347
- [2] Linde A D 1982 *Phys. Lett. B* **108** 389
- [3] Albrecht A and Steinhardt P J 1982 *Phys. Rev. Lett.* **48** 1220
- [4] Guth A H 2007 *J. Phys. A: Math. Theor.* **40** 6811 (arXiv:[hep-th/0702178](#))
- [5] Tegmark M and Rees M J 1998 *Astrophys. J.* **499** 526 (arXiv:[astro-ph/9709058](#))
- [6] Albrecht A 2009 *J. Phys.: Conf. Ser.* **174** 012006 (arXiv:[0906.1047](#))
- [7] Banks T 2009 *J. Phys. A: Math. Theor.* **42** 304002 (arXiv:[0809.3951](#))
- [8] Khoury J, Ovrut B A, Steinhardt P J and Turok N 2001 *Phys. Rev. D* **64** 123522 (arXiv:[hep-th/0103239](#))
- [9] Steinhardt P J and Turok N 2002 *Phys. Rev. D* **65** 126003 (arXiv:[hep-th/0111098](#))
- [10] Horava P and Witten E 1996 *Nucl. Phys. B* **460** 506 (arXiv:[hep-th/9510209](#))
- [11] Horava P and Witten E 1996 *Nucl. Phys. B* **475** 94 (arXiv:[hep-th/9603142](#))
- [12] Erickson J K, Wesley D H, Steinhardt P J and Turok N 2004 *Phys. Rev. D* **69** 063514 (arXiv:[hep-th/0312009](#))
- [13] Moore G W, Peradze G and Saulina N 2001 *Nucl. Phys. B* **607** 117 (arXiv:[hep-th/0012104](#))
- [14] Lima E, Ovrut B A, Park J and Reinbacher R 2001 *Nucl. Phys. B* **614** 117 (arXiv:[hep-th/0101049](#))
- [15] Lehners J-L 2008 *Phys. Rep.* **465** 223 (arXiv:[0806.1245](#))
- [16] Khoury J, Ovrut B A, Steinhardt P J and Turok N 2002 *Phys. Rev. D* **66** 046005 (arXiv:[hep-th/0109050](#))
- [17] Lehners J-L 2010 *Adv. Astron.* **2010** 903907 (arXiv:[1001.3125](#))
- [18] Lyth D H 2002 *Phys. Lett. B* **524** 1 (arXiv:[hep-ph/0106153](#))
- [19] Tsujikawa S, Brandenberger R and Finelli F 2002 *Phys. Rev. D* **66** 083513 (arXiv:[hep-th/0207228](#))
- [20] Creminelli P, Nicolis A and Zaldarriaga M 2005 *Phys. Rev. D* **71** 063505 (arXiv:[hep-th/0411270](#))

- [21] Khoury J and Steinhardt P J 2010 *Phys. Rev. Lett.* **104** 091301 (arXiv:0910.2230)
- [22] Khoury J and Steinhardt P J 2011 arXiv:1101.3548
- [23] Finelli F 2002 *Phys. Lett. B* **545** 1 (arXiv:hep-th/0206112)
- [24] Notari A and Riotto A 2002 *Nucl. Phys. B* **644** 371 (arXiv:hep-th/0205019)
- [25] Lehners J-L, McFadden P, Turok N and Steinhardt P J 2007 *Phys. Rev. D* **76** 103501 (arXiv:hep-th/0702153)
- [26] Lukas A, Ovrut B A, Stelle K S and Waldram D 1999 *Phys. Rev. D* **59** 086001 (arXiv:hep-th/9803235)
- [27] Koyama K and Wands D 2007 *J. Cosmol. Astropart. Phys.* **JCAP04**(2007)008 (arXiv:hep-th/0703040)
- [28] Koyama K, Mizuno S and Wands D 2007 *Class. Quantum Grav.* **24** 3919 (arXiv:0704.1152)
- [29] Gordon C, Wands D, Bassett B A and Maartens R 2001 *Phys. Rev. D* **63** 023506 (arXiv:astro-ph/0009131)
- [30] Tolley A J and Wesley D H 2007 *J. Cosmol. Astropart. Phys.* **JCAP05**(2007)006 (arXiv:hep-th/0703101)
- [31] Khoury J, Steinhardt P J and Turok N 2003 *Phys. Rev. Lett.* **91** 161301 (arXiv:astro-ph/0302012)
- [32] Buchbinder E I, Khoury J and Ovrut B A 2007 *Phys. Rev. D* **76** 123503 (arXiv:hep-th/0702154)
- [33] Koyama K, Mizuno S, Vernizzi F and Wands D 2007 *J. Cosmol. Astropart. Phys.* **JCAP11**(2007)024 (arXiv:0708.4321)
- [34] Buchbinder E I, Khoury J and Ovrut B A 2008 *Phys. Rev. Lett.* **100** 171302 (arXiv:0710.5172)
- [35] Creminelli P and Senatore L 2007 *J. Cosmol. Astropart. Phys.* **JCAP11**(2007)010 (arXiv:hep-th/0702165)
- [36] Lehners J-L and Steinhardt P J 2008 *Phys. Rev. D* **77** 063533 (arXiv:0712.3779)
- [37] Lehners J-L and Steinhardt P J 2008 *Phys. Rev. D* **78** 023506 (arXiv:0804.1293)
- [38] Lehners J-L and Renaux-Petel S 2009 *Phys. Rev. D* **80** 063503 (arXiv:0906.0530)
- [39] Lehners J-L and Steinhardt P J 2009 *Phys. Rev. D* **80** 103520 (arXiv:0909.2558)
- [40] Smith K M, Senatore L and Zaldarriaga M 2009 *J. Cosmol. Astropart. Phys.* **JCAP09**(2009)006 (arXiv:0901.2572)
- [41] Boyle L A, Steinhardt P J and Turok N 2004 *Phys. Rev. D* **69** 127302 (arXiv:hep-th/0307170)
- [42] Boyle L A, Steinhardt P J and Turok N 2004 *Phys. Rev. D* **70** 023504 (arXiv:hep-th/0403026)
- [43] Baumann D, Steinhardt P J, Takahashi K and Ichiki K 2007 *Phys. Rev. D* **76** 084019 (arXiv:hep-th/0703290)
- [44] Lukas A, Ovrut B A, Stelle K S and Waldram D 1999 *Nucl. Phys. B* **552** 246 (arXiv:hep-th/9806051)
- [45] Braun V, He Y-H, Ovrut B A and Pantev T 2005 *J. High Energy Phys.* **JHEP06**(2005)039 (arXiv:hep-th/0502155)
- [46] Lehners J-L, McFadden P and Turok N 2007 *Phys. Rev. D* **76** 023501 (arXiv:hep-th/0612026)
- [47] Lehners J-L and Turok N 2008 *Phys. Rev. D* **77** 023516 (arXiv:0708.0743)
- [48] Lehners J-L, McFadden P and Turok N 2007 *Phys. Rev. D* **75** 103510 (arXiv:hep-th/0611259)
- [49] Penrose R 2009 *J. Phys.: Conf. Ser.* **174** 012001
- [50] Khoury J, Ovrut B A, Seiberg N, Steinhardt P J and Turok N 2002 *Phys. Rev. D* **65** 086007 (arXiv:hep-th/0108187)
- [51] Turok N, Perry M and Steinhardt P J 2004 *Phys. Rev. D* **70** 106004 (arXiv:hep-th/0408083)
- [52] Niz G and Turok N 2007 *Phys. Rev. D* **75** 026001 (arXiv:hep-th/0601007)
- [53] Niz G and Turok N 2007 *Phys. Rev. D* **75** 126004 (arXiv:0704.1727)
- [54] Copeland E J, Niz G and Turok N 2010 *Phys. Rev. D* **81** 126006 (arXiv:1001.5291)
- [55] Arkani-Hamed N, Cheng H-C, Luty M A and Mukohyama S 2004 *J. High Energy Phys.* **JHEP05**(2004)074 (arXiv:hep-th/0312099)
- [56] Nicolis A, Rattazzi R and Trincherini E 2009 *Phys. Rev. D* **79** 064036 (arXiv:0811.2197)
- [57] Nicolis A, Rattazzi R and Trincherini E 2010 *J. High Energy Phys.* **JHEP05**(2010)095 (arXiv:0912.4258)
- [58] Creminelli P, Nicolis A and Trincherini E 2010 *J. Cosmol. Astropart. Phys.* **JCAP11**(2010)021 (arXiv:1007.0027)
- [59] Khoury J, Lehners J-L and Ovrut B A 2011 arXiv:1103.0003
- [60] Lin C, Brandenberger R H and Levasseur L P 2010 arXiv:1007.2654
- [61] Xue B and Steinhardt P J 2010 *Phys. Rev. Lett.* **105** 261301 (arXiv:1007.2875)
- [62] Novello M and Bergliaffa S 2008 *Phys. Rep.* **463** 127 (arXiv:0802.1634)
- [63] Adams A, Arkani-Hamed N, Dubovsky S, Nicolis A and Rattazzi R 2006 *J. High Energy Phys.* **JHEP10**(2006)014 (arXiv:hep-th/0602178)
- [64] Kallosh R, Kang J U, Linde A D and Mukhanov V 2008 *J. Cosmol. Astropart. Phys.* **JCAP04**(2008)018 (arXiv:0712.2040)
- [65] Khoury J, Lehners J-L and Ovrut B 2010 arXiv:1012.3748
- [66] de Rham C and Tolley A J 2010 *J. Cosmol. Astropart. Phys.* **JCAP05**(2010)015 (arXiv:1003.5917)
- [67] Turok N, Craps B and Hertog T 2007 arXiv:0711.1824
- [68] Craps B, Hertog T and Turok N 2007 arXiv:0712.4180
- [69] Craps B, Hertog T and Turok N 2009 *Phys. Rev. D* **80** 086007 (arXiv:0905.0709)
- [70] Battarra L and Hertog T 2010 *J. High Energy Phys.* **JHEP12**(2010)017 (arXiv:1009.0992)

-
- [71] Neal R M 2006 arXiv:[math/0608592](#)
 - [72] Erickson J K, Gratton S, Steinhardt P J and Turok N 2007 *Phys. Rev. D* **75** 123507 (arXiv:[hep-th/0607164](#))
 - [73] Lehners J-L and Steinhardt P J 2009 *Phys. Rev. D* **79** 063503 (arXiv:[0812.3388](#))
 - [74] Lehners J-L, Steinhardt P J and Turok N 2009 *Int. J. Mod. Phys. D* **18** 2231 (arXiv:[0910.0834](#))
 - [75] Lemaitre G 1933 *Ann. Soc. Sci. Brux. I A* **53** 51
 - [76] Lehners J-L and Steinhardt P J 2011 *Phys. Rev. Lett.* **106** 081301 (arXiv:[1008.4567](#))
 - [77] Freivogel B 2011 *Class. Quantum Grav.* **28** 204007
 - [78] Kleban M 2011 *Class. Quantum Grav.* **28** 204008
 - [79] Craps B and Evnin O 2011 *Class. Quantum Grav.* **28** 204006