COSMOLOGY WITH DYNAMICAL EXTRA DIMENSIONS

Joel K. Erickson

A Dissertation

Presented to the Faculty

of Princeton University

in Candidacy for the Degree

of Doctor of Philosophy

Recommended for Acceptance

by the Department of

Physics

September 2005

Abstract. Nearly every attempt to unify the fundamental forces incorporates the idea of compact extra dimensions. The notion was introduced by Kaluza and Klein in the 1920s and is an essential part of contemporary string theory and M-theory. In most treatments the extra dimensions are static. We consider the consequences of extra dimensions with time-varying radii. The radii are modeled by light scalar fields. These may have unusual properties which produce observable effects, such as non-canonical kinetic energies, couplings to matter and radiation, and non-minimal coupling to gravity.

Extra dimensions may be responsible for dark energy in the late universe. The simplest model of dark energy is characterized by its equation of state. We show that constraints placed on realistic models by the universality of free fall, variation of fundamental constants and metric tests of gravity are often stricter than bounds on the equation of state. Testing the equivalence principle may be an effective way of distinguishing some quintessence models from a cosmological constant.

In certain dark energy models the speed of sound is much less than the speed of light. We calculate how this affects the cosmic microwave background and show that the speed of sound may be measurable, provided dark energy is sufficiently dense at decoupling. This is another possible signature of quintessence.

Dynamical extra dimensions may have consequences for the early universe. In the cyclic model, the universe is described in terms of a series of contractions and expansions of an extra dimension. The big bang is preceded by a big crunch and quantum fluctuations of the scalar field produce structure in universe. We consider how the fluctuations evolve and build over many cycles and show that there are no observable instabilities or adverse effects.

In the cyclic model extra dimensions act as both dark energy and as an agent to cause contraction and a big crunch. Previous theorems suggested that contraction necessarily leads to chaotic behavior and unacceptable inhomogeneity. We show that homogeneous contraction is possible if the pressure-to-density ratio of the scalar field is sufficiently large.

For my father

Contents

Acknowledgments

Forsan et haec olim meminisse iuvabit.

— Virgil, *Aeneid*, 1.203

Above all, I thank my advisor Paul Steinhardt, for his ideas, guidance, time and good taste. I also extend thanks to my collaborators: Christian Armendáriz-Picón, Rob Caldwell, Steven Gratton, Slava Mukhanov, Neil Turok and Dan Wesley. I am indebted to the other members of Paul's group for valuable discussions over the years. My appreciation of the physics department, in particular the gravity group – faculty, staff, postdocs and students – cannot be overstated. There could not have been a better place to stumble upon cosmology. Thank you all.

Good friends make graduate school – and life in Princeton – gratifying. For that, I thank my friends. I am particularly grateful to my colleagues Justin Khoury and Jeff McMahon, for discussions, gossip, commiseration, beer and trash-talking. I credit Carrie with keeping me off the cliff, with sympathy and frequent bribes of food, in the final weeks of writing.

I was partly supported by the National Science and Engineering Research Council of Canada.

CHAPTER 1

Introduction

Cosmology and particle physics are both described, at present, by models in remarkable agreement with the data. The simple ΛCDM cosmological model assumes that the universe was once in a hot, roughly homogeneous and isotropic initial state, with a spectrum of nearly scale invariant perturbations. In addition to known fundamental physics, the model adds a cosmological constant (Λ) and cold dark matter (CDM) [69]. These additions give good agreement with all the data, notably the cosmic microwave background [167], large scale structure [175] and nucleosynthesis [144]. Likewise, the standard model of particle physics, despite its inelegance, is stubbornly in agreement with data from accelerators. High energy theory – supersymmetry, string theory and M-theory in particular – is a source for a rich phenomenology beyond this model [89, 90, 152, 153, 192].

The intersection of these two pictures – of particle theory and cosmology – is murky. The fundamental explanation for dark matter, baryogenesis, dark energy and the initial conditions of the hot big bang [121] is unknown. In part this is because gravity is a crude tool for particle physics. Conversely, nothing is known about what new gravitational predictions particle theory will ultimately make. One robust feature is that string theory and M-theory are consistent only with a number of extra dimensions.

The idea of extra dimensions has a long history in physics. An extra dimension was first suggested by Nordström in 1914 [142] as a way to unify gravity and electromagnetism in his scalar theory of relativistic gravity. His work was quickly forgotten. In 1919, however, Kaluza [109] rediscovered this idea in general relativity,[1] by showing that the vector potential

[1]For a review of Kaluza-Klein theory, see Overduin and Wesson [145]. Duff [71] discusses Kaluza-Klein compactification in string theory. A historical review with many relevant reprints, including modern papers, is Appelquist [7].

of electromagnetism arises naturally in a five dimensional setting. Since the visible universe is unmistakably four dimensional, the extra dimension must be difficult to probe. Kaluza ensured this by imposing translation invariance on the extra dimension. He called this the "cylinder condition", and regarded it as mainly a formal innovation. (The "cylinder" refers to translation invariance, not to a compact topology.)

Oskar Klein [118, 119] – and later Einstein and Bergmann [77] – were the first to ascribe any physical significance to the extra dimension. Klein thought of it as a circle. He noticed that momentum was quantized in the extra dimension and that the higher Fourier modes of the scalar – the "tower" of Kaluza-Klein states – satisfy Klein-Gordon equations. Thus, in 1926 he suggested that quantum mechanics and the quantization of charge could arise from a compact extra dimension of size roughly 10^{-32} meters, the GUT (grand unified theory) scale.

Kaluza-Klein phenomenology has been extended to non-Abelian gauge groups. Witten [195] has demonstrated that eleven dimensions is the minimum number necessary to obtain the standard model group in a Kaluza-Klein theory compactified to four dimensions. Despite the auspicious coincidence that the largest supergravity is formulated in eleven dimensions, interest in the Kaluza-Klein program has petered out. Problems with chiral fermions and constructing theories with flat backgrounds seems to have made Kaluza-Klein theory an inelegant way to derive Yang-Mills theory from gravity.

However, extra dimensions are an essential element of string theory [89, 90, 152, 153] for quite different reasons. While Kaluza-Klein theory attempts to unify gravity and gauge theory by adding extra compact dimensions, string theory already naturally incorporates both gravity and gauge theory. However, string theory cannot be consistently formulated in four dimensions: the extra dimensions are required for consistency. Thus, the problem in string theory is to eliminate the extra dimensions while preserving the appealing aspects of the theory, such as low energy supersymmetry. An extensive framework in algebraic geometry has been developed for this purpose. It has been shown that compactification on

six dimensional Calabi-Yau manifolds [89, 91] allows for low-energy supersymmetry. These manifolds typically have a large number of internal degrees of freedom, called moduli, which correspond to massless scalar fields.

Another *a priori* possibility for reducing a higher dimensional theory – like string theory – to lower dimensions is to confine the visible universe to a membrane with three spatial dimensions, a three-brane. This idea, first proposed by Dirac [68], is appealing because various kinds of branes are part of the string theory menagerie.[2] By itself, however, the idea is problematic, because gravity is not confined to the brane and so radical departures from the inverse square law are predicted. Randall and Sundrum [155] have proposed a geometry in which gravity is well confined to four dimensions and departures from general relativity are small. However, it is not known if this idea is compatible with string theory [92], so trying to construct phenomenologically viable models in string theory usually involves considering both branes and small compact dimensions.

In heterotic M-theory [95, 96, 196] the four dimensional universe comes from compactified eleven dimensional M-theory [126]. This is an important example because it is the best known candidate for embedding the standard model of particle physics in string theory and M-theory. Recent results have been encouraging: it has been shown that the theory has vacua with standard model gauge groups and three families of quarks and leptons [30, 31, 70].

An important feature of heterotic M-theory is the S^1/Z_2 orbifold: one dimension is compactified on an interval, with distinguished endpoints. The endpoints form spatial boundaries, called fixed planes. The fixed planes contain E_8 gauge theories and can have branes trapped at them. The size of the intervening space is modeled – in the low-energy description – by a scalar field, called the radion.

Thus, virtually every attempt to unify the fundamental forces of physics incorporates the idea of extra dimensions. This thesis considers the idea that compact extra dimensions are

[2]For a discussion of the role of branes in string theory, see [72, 152, 153]. For reviews of brane cosmology, see [32, 75, 122].

dynamical. Rather than being fixed, their sizes are varying in time with the radius modeled by a scalar field, the radion. Remarkably, the early efforts in Kaluza-Klein theory simply ignored this scalar mode. Its physical significance was recognized first by Jordan and later by Einstein and Bergmann [77, 104]. Fierz pointed out that time-varying dimensions could cause large observable effects [81]. This thesis is motivated by the idea that, if stringent observational constraints are satisfied, light scalar fields can play a profound role in the evolution of the universe today, as the origin of dark energy [38, 157] and a solution of the cosmic coincidence problem [10, 173, 198]; and in the early universe, by establishing the initial conditions for the hot big bang in the inflationary [121, 124], ekpyrotic [114, 115] and pre-big bang scenarios [86, 87, 184].

A theory in which a dynamical dimension plays a crucial role is the cyclic scenario of Steinhardt and Turok [170–172]. In the low energy effective action description of heterotic M-theory, the radion is massless. There does not appear to be a reason, such as a fundamental symmetry, that the radion must be massless. It is generally thought that the radion will acquire a mass from non-perturbative effects such as supersymmetry breaking. The non-perturbative potential vanishes as the fixed planes approach each other. The ekpyrotic scenario proposes that, before this happens, the potential has a steep negative exponential segment. The cyclic scenario proposes that, in addition, the potential has a flat segment which corresponds to dark energy domination. In these models, the radius of the S^1/Z_2 orbifold is perpetually changing in a cyclic fashion, going from large values to zero and back again. In this scenario, the collision between the two orbifold planes generated the radiation that triggered the hot big bang, so time-varying compact dimensions are an essential aspect of the model.

This thesis discusses the observational and theoretical ramifications of dynamical extra dimensions. If the radion is evolving today, then it may be the field responsible for dark energy. In chapters 3 and 4 we discuss some observational constraints and possibilities. Later,

we discuss theoretical challenges in such models, such as the chaotic behavior of general relativity (chapter 5) and constructing cyclic models (chapter 6).

Although we use the idea of compactification as a useful geometrical motivation, many of the ideas in this thesis are applicable in the more general context of light scalar fields in the universe.

1. Overview

Chapter 2 reviews some background concepts and notation. Section 1 discusses the cosmological evolution of a scalar field dominated universe. In section 1, we mention some results in cosmological perturbation theory. We show how perturbations are generated from the quantum fluctuations of a scalar field in the inflationary and ekpyrotic scenarios, and discuss how perturbations in the density of matter and radiation evolve as they reenter the horizon. Section 3 derives the Kaluza-Klein and S^1/Z_2 orbifold compactification, and briefly describes the Randall-Sundrum and Hořava-Witten models from the literature.

Chapters 3 and 4 discuss the observational consequences of the idea that a time-varying extra dimension might be the source of dark energy. One of the critical problems of modern cosmology is determining if dark energy is a cosmological constant, with an inexplicably small value, or has a dynamical origin. Therefore, in these chapters we focus on observational signatures that might arise from extra dimensions: dark energy with a non-canonical kinetic energy, or with couplings to matter and radiation, or with Brans-Dicke couplings.

Chapter 3 investigates the speed of sound of quintessence. This is the speed with which perturbations propagate. In the usual Klein-Gordon scalar field models, the speed of sound is equal to the speed of light. However a scalar field ϕ may have an action which is of higher than quadratic order in derivatives: it may contain terms like $(\partial\phi)^4$. These models have a speed of sound different from unity. We show that, if the dark energy density is at least a few percent of the critical density at the surface of last scattering, then the cosmic microwave background anisotropy is sensitive to the sound speed of dark energy at decoupling. Near future observations of the cosmic microwave background, such as those to be performed by

the Planck mission, should be able to distinguish dark energy with sound speed near zero to canonical models with sound speed equal to the speed of light. This is a feature of the k-essence models [10, 11] and potentially of other models with higher derivative terms in their actions.

Chapter 4 discusses the constraints placed by the equivalence principle on models with dynamical extra dimensions. We quickly review the tremendous progress made in testing the equivalence principle – in particular, tests of the universality of free fall, constraints on variations of the fundamental constants, and precision tests of post-Newtonian gravity – and point out that many theories, such as the Randall-Sundrum models and heterotic M-theory, predict violations of the equivalence principle if the radion is unstabilized and evolving on cosmological time-scales. Improved tests of the equivalence principal can, for some forms of dynamical dark energy, be a much better way of distinguishing it from a cosmological constant than only measuring the equation of state.

In chapters 5 and 6, we turn to the effect of dynamical extra dimensions in the early universe. Chapter 5 is concerned with the the behavior of contracting universes in cosmology. The pre-big bang, ekpyrotic and cyclic models all envision the universe as contracting before the big bang. Until recently it was thought that such universes undergo chaotic mixmaster oscillations due to curvature and anisotropy. These oscillations would destroy the observed homogeneity of the universe. We show that the chaos can be avoided if the universe is dominated by a fluid with a sufficiently large equation of state. This is the case in the ekpyrotic and cyclic models, in which the steep negative potential generates an equation of state $w \gg 1$. In this case, the contraction is locally described by the Friedmann equation for a homogeneous and isotropic universe: this is a "no-hair theorem" for contracting universes, analogous to the inflationary no-hair theorem [121].

For every combination of matter fields there is a critical scalar field equation of state $w_{\rm crit}$ which, for $w > w_{\rm crit}$ ensures a stable, isotropic contraction. In a universe with just a single scalar field, $w_{\rm crit} = 1$, but if there is a more general combination of p-forms with couplings

to the scalar field, $w_{\rm crit}$ can be larger. However, for each combination of p-forms, there is always an equation of state sufficiently large that oscillations are suppressed. We show that Z_2 orbifold compactification also contributes to suppressing chaotic behavior. In particular, chaos is avoided in contracting heterotic M-theory models if $w > 1$ at the crunch.

Chapter 6 discusses the cyclic model of the universe. The model contains an expanding, dark energy phase which transitions to a contracting ekpyrotic phase. The dark energy phase can be thought of as a sort of very low-energy inflation. Both phases, then, produce a nearly-scale invariant spectrum of perturbations. We analyze the spectrum of density fluctuations of a simple class cyclic model potentials. Contrary to the intuition from inflation, the amplitude of modes is not fixed as they cross the horizon; rather modes that exit the horizon in the expanding phase continue to grow in the contracting phase. The dark energy modes are amplified by a huge factor in the ekpyrotic phase and ultimately have the same amplitude as the ekpyrotic modes: the two nearly scale invariant parts of the spectrum are smoothly joined by a small feature.

Next, we consider the effect of these fluctuations on the structure of the cyclic model over many cycles. We tabulate the evolution of the scale factor and the horizon H^{-1} in each phase of the cyclic model. The cyclic model, despite a brief contracting phase, expands by many e-folds over each cycle. On scales far outside the horizon, modes receive an huge amplification each cycle. We show that this does not lead to any observable effects, such as a diverging physical curvature. On metaphysical scales, however, the cyclic model can be thought of as different Hubble patches cycling asynchronously, continually expanding and fragmenting into causally disconnected patches.

Finally, we show that a long period of dark energy domination in the cyclic model is not, as was previously thought, necessary to prevent the fluctuations produced in one cycle of the model from contaminating fluctuations in future cycles. The attractor behavior of the contracting, ekpyrotic phase, combined with the natural suppression of scalar field perturbations inside the horizon, is sufficient to prevent any adverse effects.

2. Conventions

Except where explicitly stated, we use the following conventions. We use reduced Planck units throughout, with with $8\pi G = c = \hbar = 1$, where G is Newton's constant, c the speed of light and $\hbar$ the reduced Planck constant. The metric has a $(- + \cdots +)$ signature, with a negative time eigenvalue and positive spatial eigenvalues. Greek indices starting with μ are used for space-time indices: $\mu, \nu, \cdots = 0, 1, 2, 3$. Lowercase Roman indices starting with i for purely spatial indices: $i, j, \cdots = 1, 2, 3$. We will have occasion to use other kinds of indices in several chapters, and they will be defined as they are used. All repeated tensorial indices are summed, unless otherwise specified. The Riemann tensor for a metric $g_{\mu\nu}$ is given by

$$R^{\mu}{}_{\nu\rho\sigma} = \partial_\rho \Gamma^{\mu}{}_{\nu\sigma} - \partial_\sigma \Gamma^{\mu}{}_{\nu\rho} + \Gamma^{\mu}{}_{\xi\rho}\Gamma^{\xi}{}_{\nu\sigma} - \Gamma^{\mu}{}_{\xi\sigma}\Gamma^{\xi}{}_{\nu\rho}, \tag{1.1}$$

in terms of the connection coefficients

$$\Gamma^{\xi}{}_{\mu\nu} = \frac{1}{2} g^{\xi\sigma}[\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}]. \tag{1.2}$$

The Ricci tensor is given by $R_{\mu\nu} = R^{\sigma}{}_{\mu\sigma\nu}$ and the Einstein tensor by $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R$. The scalar curvature is $R = \operatorname{tr} R_{\mu\nu} = g^{\mu\nu} R_{\mu\nu}$. The Einstein-Hilbert action S is

$$S = \frac{1}{2} \int d^4x \sqrt{-g} R, \tag{1.3}$$

where g is the determinant of the metric. We refer to any metric in which the scalar curvature appears in this form as in *Einstein frame*. From this action, the Einstein equations can be obtained,

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = T_{\mu\nu}, \tag{1.4}$$

where $T_{\mu\nu}$ is the stress-energy tensor which comes from the variation of the non-gravitational part of the Lagrangian density $\mathcal{L}_{\mathrm{NG}}$,

$$T_{\mu\nu} = \frac{2}{\sqrt{-g}} \frac{\delta(\sqrt{-g}\mathcal{L}_{\mathrm{NG}})}{\delta g^{\mu\nu}}, \tag{1.5}$$

where δ represents a functional variation. A *canonically normalized* scalar field ϕ has action

$$S = \int d^4x \left[-\tfrac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi) \right], \tag{1.6}$$

where V is the potential. The Friedmann-Lemaître-Robertson-Walker metric is:

$$ds^2 = -dt^2 + a(t)^2 \sum_{i=1}^{3} (dx^i)^2 = a(\tau)^2 \left[-d\tau^2 + \sum_{i=1}^{3} (dx^i)^2 \right], \tag{1.7}$$

where a is the scale factor and the x^i are comoving coordinates; t is proper time, and τ is conformal time. They are related by

$$\tau(t) = \int^t \frac{d\hat{t}}{a(\hat{t})}. \tag{1.8}$$

Dots are always reserved for proper time derivatives, whereas primes are often used for conformal time derivatives: $\dot{a}$ and a', respectively. The Hubble parameter is given by $H = \dot{a}/a = a'/a^2$. We will occasionally use the "dimensionless Hubble parameter" $\mathcal{H} = aH = \dot{a} = a'/a$. The redshift z of something occurring at time t is given by

$$1 + z = \frac{a_0}{a(t)}, \tag{1.9}$$

where a_0 indicates the scale factor today. The present value of the Hubble parameter is likewise written H_0. A wavenumber k is always comoving, so the associated physical wavenumber is k/a. Finally, for a perfect fluid with energy density ρ and pressure p, the *equation of state* w is defined by $p = w\rho$.

CHAPTER 2

Scalar fields and compactification

In this chapter, we introduce the fundamental tools that will be used throughout this dissertation: the scalar field in homogeneous cosmology, cosmological perturbation theory and some basic results in compactification. This chapter focuses on the elements that are essential for the issues considered in the thesis.[1]

1. Cosmological solutions with scalar fields

The most general spatially flat metric compatible with the cosmological principle – that is, a homogeneous and isotropic spacetime – is the Friedmann-Lemaître-Robertson-Walker (FRW) metric. Written in four dimensions in terms of proper time, it is

$$ds^2 = -dt^2 + a(t)^2(dx_1^2 + dx_2^2 + dx_3^2). \tag{2.1}$$

The x_i are comoving coordinates with no absolute physical meaning, since any dilatation of the x_i can be absorbed by a rescaling of a. The *Hubble parameter* $H = \dot{a}/a$ measures the rate of expansion of the universe. The *Hubble length* H^{-1}, measures the distance between two points on a comoving surface whose relative separation is increasing (or decreasing, in a contracting universe) at the speed of light and measures the horizon, or the largest scale at which causal interactions occur. However, H^{-1} can also be interpreted as the *Hubble time* which measures the time taken for one e-fold of expansion (or contraction). In an expanding radiation or matter dominated universe, it is roughly the elapsed time since the big bang.

1.1. Perfect fluids. The Einstein equation for this metric is the acceleration equation,

$$6\frac{\ddot{a}}{a} = -(\rho + 3p) = -(1 + 3w)\rho, \tag{2.2}$$

[1]For a more complete discussion, see [69, 121, 124, 146, 148].

where $\rho(t)$ is the matter energy density, $p(t)$ is its pressure and $w = p/\rho$ is the equation of state. The expansion or contraction accelerates if $w < -1/3$.

A perfect fluid has has stress-energy tensor

$$T_{\mu\nu} = (\rho + p) u_\mu u_\nu + p\, g_{\mu\nu}, \tag{2.3}$$

where $g_{\mu\nu}$ is the background metric and u_μ is the unit ($u^\mu u_\mu = -1$) four-velocity. The only condition compatible with homogeneity and isotropy is a comoving fluid, $u_\mu = (1, 0, 0, 0)$, so

$$T^\mu{}_\nu = \begin{bmatrix} -\rho & & & \\ & p & & \\ & & p & \\ & & & p \end{bmatrix}. \tag{2.4}$$

A fluid in an evolving background obeys the conservation equation

$$\dot{\rho} = -3H(\rho + p) = -3H(1 + w)\rho. \tag{2.5}$$

The Friedmann equation, which is a constraint on (2.2), is

$$3H(t)^2 = \rho(t). \tag{2.6}$$

This is the equation that determines the rate of expansion of a flat universe in terms of its energy density. The *critical density* is simply $3H^2$: it is the density required for the universe to be flat. The solution of (2.2), (2.5) and (2.6) for general constant w is

$$a \propto \begin{cases} |t|^{\frac{2}{3(1+w)}} & w \neq -1 \\ e^{Ht} & w = -1 \end{cases}, \tag{2.7}$$

$$\rho \propto \begin{cases} |t|^{-3(1+w)} & w \neq -1 \\ \text{(constant)} & w = -1 \end{cases}. \tag{2.8}$$

In the case $w = -1$, a cosmological constant, the Hubble parameter H is a constant. Note that these solutions are valid for a contracting ($t < 0$ or $H < 0$) solutions as well as expanding ($t > 0$ or $H > 0$) solutions.

In a universe with multiple non-interacting fluids, these results still hold, where now w is an average quantity and ρ and p measure the total energy density and pressure. In this case,

$$\rho = \sum_X \rho_X, \tag{2.9}$$

$$p = \sum_X p_X, \tag{2.10}$$

where the sum is over the various fluids. We may define densities relative to the critical density

$$\Omega_X = \rho_X / \rho, \tag{2.11}$$

and the *effective* equation of state

$$w = \frac{\sum_X p_X}{\sum_X \rho_X} = \sum_X w_X \Omega_X, \tag{2.12}$$

where $w_X = p_X/\rho_X$. In this case, (2.6) and (2.2) hold with the quantities ρ, p and w defined as in (2.9), (2.10) and (2.12), respectively. The continuity equation is now

$$\dot{\rho}_X = -3H(\rho_X + p_X) = -3H(1 + w_X)\rho_X. \tag{2.13}$$

1.2. Scalar fields. Now consider a scalar field with Lagrangian density

$$\mathcal{L}_\phi = -\frac{1}{2}(\partial\phi)^2 - V(\phi). \tag{2.14}$$

The stress-energy tensor is given by

$$T_{\mu\nu} = \frac{2}{\sqrt{-g}} \frac{\delta(\sqrt{-g}\mathcal{L}_\phi)}{\delta g^{\mu\nu}} = 2\frac{\delta \mathcal{L}_\phi}{\delta g^{\mu\nu}} + g_{\mu\nu}\mathcal{L}_\phi \tag{2.15}$$

so that

$$T_{\mu\nu} = -\partial_\mu\phi\,\partial_\nu\phi + g_{\mu\nu}\left(-\tfrac{1}{2}(\partial\phi)^2 - V(\phi)\right), \tag{2.16}$$

so assuming that $\phi = \phi(t)$,

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi), \tag{2.17}$$

$$p_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi), \tag{2.18}$$

$$w_\phi = \frac{\frac{1}{2}\dot{\phi}^2 - V(\phi)}{\frac{1}{2}\dot{\phi}^2 + V(\phi)}, \tag{2.19}$$

where ρ_ϕ, p_ϕ and w_ϕ are the energy density, pressure and equation of state of the scalar field, respectively. The equation of state generally varies in time. The conservation equation (2.5) is the equation of motion

$$\ddot{\phi} + 3H\dot{\phi} = -V_{,\phi}, \tag{2.20}$$

where the subscript "$_{,\phi}$" denotes a ϕ derivative. This equation can be rewritten in conformal time

$$\phi'' + 2\mathcal{H}\phi' = -a^2 V_{,\phi}. \tag{2.21}$$

The acceleration equation (2.2) becomes

$$3\frac{\ddot{a}}{a} = -\dot{\phi}^2 + V(\phi). \tag{2.22}$$

The free scalar field, with $V = 0$, is the simplest case. It has

$$a = a_0\left|\frac{t}{t_0}\right|^{1/3}, \qquad \phi = \sqrt{\frac{2}{3}}\log\left|\frac{t}{t_0}\right|, \tag{2.23}$$

where a_0 and t_0 are constants of integration. The free scalar field has $w = 1$. Another particularly simple case is the exponential potential,

$$V = V_0 e^{b\phi}, \tag{2.24}$$

where b and V_0 are constants, with $V_0 > 0$. A solution of the equations of motion (2.20), (2.6) and (2.22) is

$$a = a_0 \left| \frac{t}{t_0} \right|^{2/b^2}, \qquad \phi = \frac{2}{b} \log|t/t_0|. \tag{2.25}$$

The constant $t_0 = V_0^{-1/2}(4/b^4 - 2/b^2)^{1/2}$ is unimportant. This is known as a *scaling solution* since a has a simple power-law behavior. It is valid for any $|b| \leq \sqrt{6}$ and has equation of state $b^2 = 3(1+w)$ (so $-1 \leq w \leq 1$). If $b > \sqrt{6}$, the potential energy decreases more quickly than the kinetic energy, and $w \to 1$ from below as $t \to \infty$.

Likewise, we can consider a negative potential

$$V = -V_0 e^{-c\phi}, \tag{2.26}$$

where V_0 and c are positive constants, which has scaling solution

$$a = a_0 \left| \frac{t}{t_0} \right|^{2/c^2}, \qquad \phi = \frac{2}{c} \log|t/t_0|. \tag{2.27}$$

Again, t_0 is an unimportant constant. The solution is valid for any $c \geq \sqrt{6}$ and gives equation of state $c^2 = 3(1+w)$, so $w \geq 1$. If $c < \sqrt{6}$, the kinetic energy blue-shifts more quickly than the potential, so $w \to 1^+$ from above as $t \to 0^-$.

These two solutions, a positive exponential with a small constant $b \leq \sqrt{6}$ and a negative exponential with a large constant $c \geq \sqrt{6}$ are particularly simple solutions of the Einstein equations in the presence of a scalar field. They have constant equation of state, and the expression for the scale factor and equation of state take simple forms.

In an expanding universe, the positive potential solution (2.25) is a dynamical attractor: any set of initial conditions will approach it. This is true even in the presence of matter if $b \ll 1$, and is known as the inflationary "no-hair" theorem [121]. It is one of the central results of inflationary cosmology, and explains why the inflationary model solves the problems of homogeneity, isotropy and flatness. The condition $b \ll 1$ is equivalent to the slow-roll condition of inflation. In this case the model generates a spectrum of nearly scale-invariant quantum fluctuations.

The situation for the contracting, negative potential solution (2.27) is completely analogous. The solutions is a dynamical attractor in a contracting universe, and if $c \gg \sqrt{6}$ this is true even in the presence of matter. This is proved in chapter 5, and establishes a "no-hair" theorem for contracting universes. Moreover, $c \gg \sqrt{6}$ is equivalent to the fast-roll condition of the ekpyrotic model [112, 113], so in this case the model generates a spectrum of nearly-scale invariant fluctuations. This remarkable duality is discussed in [25, 80]. In [88] it is shown that these solutions are representative of the the only two kinds of stable, cosmological solutions – $w \approx -1$ and expanding, and $w \gg 1$ and contracting – which generate the observed nearly scale-invariant spectrum of fluctuations. The connection between the two will be discussed further in chapter 6.

An alternative way to understand the properties of these solutions is to rewrite the above equations in terms of $y = \log a$. The conservation equation (2.5) becomes, simply,

$$\frac{d\rho}{dy} = -3(\rho + p) \qquad \text{or} \qquad \frac{d\log\rho}{dy} = -3(1+w). \tag{2.28}$$

The scalar field case is more complicated. The equations (2.6), (2.2) and (2.20) can be rewritten

$$3y'^2 = \frac{1}{2}y'^2\Big(\frac{d\phi}{dy}\Big)^2 + e^{2y}V, \tag{2.29}$$

$$3y'' = -y'^2\Big(\frac{d\phi}{dy}\Big)^2 + e^{2y}V(\phi), \tag{2.30}$$

$$y'^2\frac{d^2\phi}{dy^2} + y''\frac{d\phi}{dy} + 2y'^2\frac{d\phi}{dy} = -e^{2y}V_{,\phi}. \tag{2.31}$$

The primes here are conformal time derivatives: $y' = a\dot{y}$. We can eliminate two variables from these three equation. Eliminating y' and y'', we obtain

$$2\frac{d^2\phi}{dy^2} = -\Big(\frac{d\phi}{dy} - \frac{V_{,\phi}}{V}\Big)\Big(6 - \Big(\frac{d\phi}{dy}\Big)^2\Big). \tag{2.32}$$

If V is exponential,

$$V(\phi) = V_0 e^{c\phi}, \tag{2.33}$$

where V_0 and c are constants, then we can rewrite this as

$$\frac{d\psi}{dy} = 3(\psi - c/\sqrt{6})(\psi - 1)(\psi + 1), \tag{2.34}$$

where $d\phi/dy = \sqrt{6}\psi$. From (2.19) and (2.29) the equation of state is given by

$$w_\phi = \frac{1}{3}\Big(\frac{d\phi}{dy}\Big)^2 - 1 = 2\psi^2 - 1, \tag{2.35}$$

so we can see that potentials with $|c| = \sqrt{3(1+w_\phi)}$ have an attractor with equations of state w_ϕ. The sign of V_0 does not enter into (2.34), but it determines the initial conditions from the equation of state (2.35). If V_0 is positive then $|\psi| < 1$, whereas if V_0 is negative then $|\psi| > 1$. For $|c| \leq \sqrt{6}$ and V_0 negative, the attractor solution has $w_\phi = 1$. For $|c| \geq \sqrt{6}$ and V_0 positive, the attractor solution has $w_\phi = 1$.

2. Perturbation theory

Cosmological perturbation theory seeks to understand the generation and evolution of linear perturbations in a background cosmology. The comoving scale of the horizon is measured by $(aH)^{-1}$. Since perturbations have fixed comoving scale, in an epoch in which $|aH|$ is decreasing, such as today, perturbations are entering the horizon. This is the source of the horizon problem resolved by the inflationary [124] and ekpyrotic [112] cosmologies: perturbations on the horizon today have never been in causal contact in the standard hot big bang cosmology. This is resolved by the inflationary or ekpyrotic scenarios, in which the universe has a period in which $|aH|^{-1}$ is decreasing, so perturbations are exiting the horizon before the start of the hot big bang. In between, there is a period generically called *reheating*, in which the perturbations are converted into long-wavelength adiabatic fluctuations in the cosmological fluids.

The most general perturbation is

$$g_{\mu\nu} = g^{(0)}_{\mu\nu} + g^{(1)}_{\mu\nu}, \tag{2.36}$$

where $g^{(0)}$ is the background FRW metric. The perturbation $g^{(1)}$ generally contains ten independent functions. Fortunately, it is possible to reduce this. A fundamental result in cosmological perturbation theory is the scalar-vector-tensor (SVT) decomposition [124], which states that the perturbation equations are separable. Perturbations to the FRW metric can be decomposed into scalar fields, spatial divergence-free vector fields, and traceless, symmetric and divergence-free tensor fields. Each is completely decoupled from the others at linear order. The vector perturbations vanish in inflation and the ekpyrotic scenario: there is no coupling between the fluctuations of a scalar field and a divergence-free vector. We focus exclusively on scalar perturbations to the FRW background. (Ekpyrotic tensor perturbations are treated in [24]. They are negligible at the scales we consider.)

The most general scalar fluctuation of metric (2.1) are

$$ds^2 = -(1+2\Psi)dt^2 - 2(\partial_i B)dt\,dx^i + a^2(t)\left[(1-2\Phi)\delta_{ij} + 2(\partial_i\partial_j - \tfrac{1}{3}\delta_{ij}\nabla^2)E\right]dx^i\,dx^j \quad (2.37)$$

It is possible to choose a gauge in which $B = E = 0$. This is a unique gauge choice, called longitudinal gauge or conformal Newtonian gauge. It can be defined with either proper or conformal time used as the time variable:

$$ds^2 = -\big(1+2\Psi(x,t)\big)dt^2 + a^2(t)\big(1-2\Phi(x,t)\big)\delta_{ij}dx^i\,dx^j \quad (2.38)$$

$$= a^2(t)\left[-\big(1+2\Psi(x,t)\big)d\tau^2 + \big(1-2\Phi(x,t)\big)\delta_{ij}dx^i\,dx^j\right]. \quad (2.39)$$

The functions Φ and Ψ are called the gravitational potentials. On small scales, they correspond to the Newtonian potential of Newtonian gravity. Specifically, Φ satisfies the equation

$$\nabla^2\Phi = \frac{1}{2}\delta\rho, \quad (2.40)$$

where ρ is the density perturbation.

We do not need to consider anisotropic stress (that is, a traceless part of the spatial stress-energy tensor T_{ij}) and so the Einstein equations set $\Phi = \Psi$ [69, 124]. The perturbations to

the stress-energy tensor are, for a comoving fluid,

$$\delta T^0{}_0 = -\delta\rho, \tag{2.41}$$

$$\delta T^0{}_i = (\rho + p)v_i, \tag{2.42}$$

$$\delta T^i{}_j = \delta p \delta^i{}_j, \tag{2.43}$$

where $\delta\rho$ and δp are the energy density and pressure perturbations, and v_i is the velocity perturbation of the fluid. By the SVT decomposition, the relevant part of the velocity perturbation is determined by a scalar V, so that $v_i = \partial_i V$. These quantities are all gauge dependent. The sound speed c_s^2 of a fluid is defined by

$$c_s^2 = \frac{\delta p}{\delta\rho}, \tag{2.44}$$

and measures the speed at which perturbations travel. For a canonical scalar field such as (1.6), c_s^2 is unity. The effect of the sound speed on perturbations will be discussed in chapter 3.

2.1. Generation of perturbations. The primordial perturbations of both the inflationary and ekpyrotic cosmologies arise as perturbations of a scalar field evolving in a cosmological background. Writing the field as $\phi(t) + \delta\phi(x, t)$, where $\phi(t)$ solves the background equation (2.20), it is possible to obtain the first order correction to the scalar-field stress-energy tensor (2.16). It is

$$\delta T^0{}_0 = -\dot\phi\,\delta\dot\phi - V_{,\phi}\delta\phi, \tag{2.45}$$

$$\delta T^i{}_j = \delta_{ij}(\dot\phi\,\delta\dot\phi - V_{,\phi}\delta\phi) \tag{2.46}$$

The spatial gradients are all second order in $\delta\phi$ (since $\partial_i\phi(t) = 0$) and do not contribute. If we write $\delta\phi$ as a sum of plane waves,

$$\delta\phi(x, t) = \sum_k e^{ik_i x^i}\delta\phi_k(t) \tag{2.47}$$

where k_i is the *comoving* wavenumber, then the conservation of stress-energy gives

$$\delta\ddot{\phi} + 3H\delta\dot{\phi} + \left(\frac{k}{a}\right)^2 \delta\phi + V_{,\phi\phi}\delta\phi = 0, \tag{2.48}$$

or

$$\delta\phi'' + 2\mathcal{H}\delta\phi' + (k^2 + a^2 V_{,\phi\phi})\delta\phi = 0. \tag{2.49}$$

in conformal time. These equations do not include gravitational back-reaction. The beauty of our gauge choice is that, including back-reaction, the evolution equations for $\delta\phi$ and Φ reduce to a single second order differential equation in the Newtonian potential [36]:

$$\ddot{\Phi} + \left(H - \frac{2\ddot{\phi}}{\dot{\phi}}\right)\dot{\Phi} + 2\left(\dot{H} - \frac{H\ddot{\phi}}{\dot{\phi}}\right)\Phi - \frac{\nabla^2\Phi}{a^2} = 0. \tag{2.50}$$

The relation between Φ and $\delta\phi$ is

$$\dot{\phi}\,\delta\phi/2 = \dot{\Phi} + H\Phi. \tag{2.51}$$

To eliminate the first derivative damping term, we change variables and write the equation in terms of conformal time and the variable[2]

$$u = \Phi/\dot{\phi} = a\Phi/\phi', \tag{2.52}$$

to obtain

$$u_k'' = -\left(k^2 - U_{\rm pot}\right)u_k, \tag{2.53}$$

where

$$U_{\rm pot} = z(1/z)'' \qquad \text{and} \qquad z = a\phi'/\mathcal{H}. \tag{2.54}$$

This is the equation that we will use to compute perturbations. There are two regimes, depending on which term dominates the left hand side of (2.53). Generically

$$U_{\rm pot} \sim (aH)^2, \tag{2.55}$$

[2]This variable, introduced by Mukhanov is sometimes called v [124, 139].

so U_{pot} is crudely related to the horizon. When a mode has $k^2 \gg U_{\text{pot}}$, it is inside the horizon and (2.53) is the equation for an oscillator. When $k^2 \ll U_{\text{pot}}$ the mode is well outside the horizon. The approximation (2.55) can, however, be off by several orders of magnitude (as in inflation or the ekpyrotic model) and can briefly fail altogether at transitional epochs. This will be discussed in detail in chapter 6. When we write that a mode is outside or inside the horizon, then, we refer to the precise relation between k^2 and U_{pot} in the evolution equation for the mode.

We assume that the modes with $k^2 \gg (aH)^2$ are in their Minkowski vacuum state [20] given by

$$\delta\phi_k \sim e^{-ik\tau}/(a\sqrt{2k}), \tag{2.56}$$

up to a random phase. Using the relation $a\phi'\delta\phi/2 = (\phi' u)'$ from (2.51) and neglecting ϕ'' (which is small in an accelerating universe) this gives

$$u_k \sim e^{-ik\tau}/(2k)^{3/2}, \tag{2.57}$$

again up to a phase. Since $(aH)^2$ is increasing in this phase, these modes are moving outside the horizon. For the positive exponential potential considered (2.24) above, (2.25) and (2.54) give $U_{\text{pot}} \approx b^2/2\tau^{-2}$ (for small b). Thus (2.53) reduces to the Bessel equation [2]

$$u_k'' = -\left(k^2 - (b^2/2)\tau^{-2}\right)u_k. \tag{2.58}$$

For the negative exponential potential (2.26), the situation is very similar: (2.27) and (2.54) give

$$u_k'' = -\left(k^2 - (c^2/2)\tau^{-2}\right)u_k. \tag{2.59}$$

These equations are solved in terms of Hankel functions in [24, 112]. The result is that

$$\langle u_k^2 \rangle \propto k^{-3-2b^2/(2-b^2)} \tag{2.60}$$

in the expanding, positive potential case (2.24) and

$$\langle u_k^2 \rangle \propto k^{-3-4/(c^2-2)} \tag{2.61}$$

in the contracting case. The angle brackets indicate an ensemble average, as the random phase in (2.56) implies that the plane waves form a Gaussian random field. A *scale invariant* or Harrison-Zel'dovich spectrum has

$$\langle \Phi^2 \rangle \sim k^{-3} \qquad \text{or} \qquad \langle u^2 \rangle \sim k^{-3}. \tag{2.62}$$

The *scalar spectral index* n_s measures the deviation for scale invariance

$$n_s - 1 = \frac{d \ln \langle \Phi^2 \rangle}{d \ln k} + 3. \tag{2.63}$$

The spectra (2.60) and (2.61) are *nearly* scale invariant because $n_s - 1$ is small for small b or large c, respectively. Large scale structure date [167, 175] has confirmed that the observed spectrum of the universe is nearly scale invariant.

2.2. Reentering the horizon. After reheating, the primordial curvature perturbations considered in the last section are converted into adiabatic perturbations of the fluids that make up the present-day cosmology: baryons, radiation, dark matter and, potentially, quintessence. The adiabatic condition for a (perfect) fluid ρ_X is

$$\frac{1}{1+w_X} \frac{\delta \rho_X}{\rho_X} = \frac{1}{1+w} \frac{\delta \rho}{\rho}, \tag{2.64}$$

where w is the equation of state of the universe defined by (2.12). This condition ensures that the fractional entropy perturbation $\delta s_X / s_X$ is the same for all species.

It is easiest to understand the evolution of these modes as they come back inside the horizon in terms of their *density contrast*. This is the fractional perturbation in the background density. For the fluid X, the density contrast δ_X is

$$\delta_X = \frac{\delta \rho_X}{\rho_X}. \tag{2.65}$$

The total density contrast δ is defined as

$$\delta = \frac{\sum_X \delta\rho_X}{\sum_X \rho_X} = \sum_X \Omega_X \delta_X. \tag{2.66}$$

Its k-th Fourier mode is related to the Newtonian potential by

$$\delta_k = -\frac{2}{3}\Big(\frac{k}{aH}\Big)^2 \Phi_k, \tag{2.67}$$

from (2.40). The perturbation equations in such a multi-fluid universe have been derived by Padmanabhan (see [146] equations (4.100))

$$\delta\dot{\rho}_X = -3(\rho_X + p_X)\delta H_X - 3H\delta\rho_X - 3H(\delta p_X - \theta_X \delta p), \tag{2.68}$$

$$\delta\dot{H}_X = -2H\delta H_X - \frac{1}{6}\delta\rho - \frac{1}{3}\frac{\nabla^2\delta p_X}{\rho_X + p_X} + \frac{\dot{p}_X}{\rho_X + p_X}\Big[\sum_Y \theta_Y \delta H_Y - \delta H_X\Big], \tag{2.69}$$

where δH_X is the "perturbation" in the local Hubble parameter due to density perturbation $\delta\rho_X$. It is proportional to the velocity divergence $3\delta H = a^{-1}\partial_i v^i = \nabla^2 V$. Also,

$$\theta_X = \frac{\rho_X + p_X}{\rho + p} = \Omega_X \frac{1 + w_X}{1 + w}. \tag{2.70}$$

The first equation (2.68) is the perturbed continuity equation (2.13) while (2.69) comes from the Raychaudhuri (or Euler) equation for the local expansion of space. In both equations, the last term on the right hand side vanishes in the case of a single fluid, or for a dominant fluid in a multi-fluid universe. The gauge invariant expressions for the pressure perturbation in terms of density and δH are

$$\delta p_X = v_X^2 \delta\rho_X + \Big(\frac{aH}{k}\Big)^2 \big[(1 + w_X)(v_X^2 - w_X) + \tfrac{1}{3}H^{-1}\dot{w}_X\big]\rho_X \frac{\delta H_X}{H}, \tag{2.71}$$

$$= v_X^2 \delta\rho_X + (1 + w_X)\Big(v_X^2 - \frac{\dot{p}_X}{\dot{\rho}_X}\Big)\Big(\frac{aH}{k}\Big)^2 \rho_X \frac{\delta H_X}{H}, \tag{2.72}$$

where v_X^2 is the sound speed of fluid X, and, from the continuity equation (2.13),

$$\frac{\dot{p}_X}{\dot{\rho}_X} = w_X - \frac{1}{3}\frac{H^{-1}\dot{w}_X}{1 + w_X}. \tag{2.73}$$

In the case of a perfect fluid, $v_X^2 = \dot\rho_X/\dot p_X$ and so (2.72) reduces to $\delta p_X = v_X^2\,\delta\rho_X$. Using these expressions, (2.68) and (2.69) may be rewritten in Fourier space

$$\begin{aligned}\frac{d\delta_X}{d\log a} &= -3(1+w_X)\Delta_X + 3w_X\delta_X - 3\Big[v_X^2\delta_X - \frac{1+w_X}{1+w}\sum_Y v_Y^2\Omega_Y\delta_Y\Big] \\ &\quad - 3(1+w_X)\Big(\frac{aH}{k}\Big)^2\Big[\Big(v_X^2 - \frac{\dot p_X}{\dot\rho_X}\Big)\Delta_X - \sum_Y \frac{1+w_Y}{1+w}\Big(v_Y^2 - \frac{\dot p_Y}{\dot\rho_Y}\Big)\Omega_Y\Delta_Y\Big],\end{aligned} \tag{2.74}$$

$$\begin{aligned}\frac{d\Delta_X}{d\log a} &= -\frac{1}{2}(1-3w)\Delta_X - \frac{1}{2}\sum_Y \Omega_Y\delta_Y + \frac{1}{3}\Big(\frac{k}{aH}\Big)^2\frac{v_X^2}{1+w_X}\delta_X \\ &\quad - 3\frac{\dot p_X}{\dot\rho_X}\Big[\sum_Y \frac{1+w_Y}{1+w}\Omega_Y\Delta_Y - \Delta_X\Big] + \frac{1}{3}\Big[v_X^2 - \frac{\dot p_X}{\dot\rho_X}\Big]\Delta_X,\end{aligned} \tag{2.75}$$

where Δ_X is the contrast

$$\Delta_X = \frac{\delta H_X}{H}, \tag{2.76}$$

chosen so that the equations can be written entirely in terms of dimensionless ratios. The penultimate terms on the right hand sides of (2.74) and (2.75) vanish for perfect fluids. The last terms of these equations vanish for a single fluid, or for the dominant ($\Omega_X = 1$) fluid in a multi-component universe. The dominant fluid evolves independently of the other components in the universe, as would be expected. The initial conditions (for perfect fluids) are adiabatic, from (2.64) and have

$$\frac{1}{1+w_X}\delta_X = \frac{1}{1+w}\delta \qquad \text{and} \qquad \Delta_X = -\tfrac{1}{4}\delta. \tag{2.77}$$

In a universe with only one fluid, (2.74) and (2.75) reduce to

$$\begin{aligned}\delta'' + \Big(\tfrac{1}{2} - \tfrac{9}{2}w - \frac{w'}{1+w}\Big)\delta' + \Big(\tfrac{9}{2}w^2 - 3w - \tfrac{3}{2} - \frac{3w'}{1+w} + \Big(\frac{k}{aH}\Big)^2 c_s^2\Big)\delta \\ = \Big(c_s^2 - \frac{\dot p}{\dot\rho}\Big)\big(\tfrac{1}{3}\delta' - w\delta\big),\end{aligned} \tag{2.78}$$

where in this equation primes denote derivatives by $\log a$. The equation allows for time-varying sound speed c_s^2 and equation of state w. The two solutions of this equation are, for a perfect fluid (so the right hand side vanishes) with w constant and far super-horizon modes

$(k^2 \ll (aH)^2)$

$$\delta \sim a^{1+3w} \qquad \text{(growing)}, \tag{2.79}$$

$$\delta \sim a^{-3(1-w)/2} \qquad \text{(decaying)}. \tag{2.80}$$

The density contrast for fluids with $w < -1/3$ does not grow on large scales. The exponents obtain a small correction – which can be derived from (2.78) and (2.73) – for a general fluid with $c_s^2 \neq w$. For example, the dominant mode for a $c_s^2 = 0$ fluid does not grow for $w < -0.35$ and for a $c_s^2 = 1$ fluid for $w < -0.43$.

These equations describe the evolution of structure in the universe until the densities become non-linear, or the effects of anisotropic pressure, such as diffusion or free-streaming become important. This model of matter and radiation is too simplistic for the CMB calculations of chapter 3, as it does not account for perturbations to the photon distribution. For this, the full Boltzmann equations are needed [69, 100, 124].

3. Compactification

We now turn to compactification. While modern high energy theory uses elaborate technology – algebraic geometry, orbifolds and Calabi-Yau compactifications – to extract standard model physics from higher dimensional theories, it is instructive to review the simplest case, the Kaluza-Klein compactification.

3.1. Kaluza-Klein. We derive the Kaluza-Klein action, starting with the five dimensional spacetime. We write the metric G_{MN} (we let uppercase Roman indices run from $1, \ldots, 5$) in the block form

$$G = \begin{pmatrix} g_{\mu\nu} + e^{2\psi} A_\mu A_\nu & e^{2\psi} A_\nu \\ e^{2\psi} A_\mu & e^{2\psi} \end{pmatrix}, \tag{2.81}$$

where A_μ is a four-vector and ψ is a scalar field, the *radion*, which determines the physical radius of the extra dimension. The matrix is written in this form so that its determinant

comes out as simply

$$\det G = e^{2\psi} \det g. \tag{2.82}$$

We assume that the x^4 direction is periodic, with period one, so $x^4 \equiv x^4 + n$, for any integer n. Since x^4 is then topologically a circle, this is known as S^1 compactification. Moreover, we take h, ψ and A_μ to be *independent* of x^4. This corresponds to keeping only the massless Kaluza-Klein modes, which dominate when the extra dimension is very small: the masses of the higher modes go as the reciprocal of the Kaluza-Klein radius.

In this case, the five-dimensional scalar curvature $R_5[G]$ may be rewritten in terms of the four dimensional scalar curvature derived from the metric g, $R_4[g]$, as

$$R_5[G] = R_4[g] - 2(\partial\psi)^2 - 2\Box_g \psi - \frac{1}{4} e^{2\psi} F_{\mu\nu} F^{\mu\nu}, \tag{2.83}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the field strength associated with the vector potential A_μ, $\Box$ is the four-dimensional covariant d'Alembertian, $\Box_g = \nabla^\mu \nabla_\mu = g^{\mu\nu} \nabla_\mu \nabla_\nu$ [152]. The Einstein-Hilbert action reduces to

$$S_{\rm KK} = \frac{1}{2} \int d^5x \sqrt{-G} R_5[G] = \frac{1}{2} \int d^4x\, e^{\psi} \sqrt{-g} \big(R_4[g] - 2(\partial\psi)^2 - 2\Box_g \psi - \tfrac{1}{4} e^{2\psi} F^2 \big). \tag{2.84}$$

The action can be put in Einstein frame by a conformal transformation [20]. If $g_{\mu\nu} = e^{2C\psi} h_{\mu\nu}$ with C a constant, then

$$R_4[g] = e^{-2C\psi} \big(R_4[h] - 6C^2 (\partial\psi)^2 - 6C \Box_h \psi \big), \tag{2.85}$$

where the covariant d'Alembertian is now evaluated with the h metric. Note that

$$\Box_g = e^{-2C\psi} \Box_h + 2C e^{-2C\psi} (\nabla_\mu \psi) h^{\mu\nu} \nabla_\nu, \tag{2.86}$$

so

$$S_{\rm KK} = \int d^4x\sqrt{-h}\left(\tfrac{1}{2}R_4[h] - \tfrac{3}{4}(\partial\psi)^2 + \tfrac{1}{8}e^{3\psi}F^2\right) \tag{2.87}$$

$$= \frac{1}{2}\int d^4x\sqrt{-h}\left(R_4[h] - (\partial\phi)^2 + \tfrac{1}{4}e^{\sqrt{6}\phi}F^2\right), \tag{2.88}$$

where $C = -1/2$ and we have integrated by parts to discard the d'Alembertian and corrected the scalar-field normalization, $\psi = \sqrt{2/3}\phi$.

The original five-dimensional metric (2.81) can now be written

$$G = \begin{pmatrix} e^{-\sqrt{2/3}\phi}h_{\mu\nu} + e^{\sqrt{8/3}\phi}A_\mu A_\nu & e^{\sqrt{8/3}\phi}A_\nu \\ e^{\sqrt{8/3}\phi}A_\mu & e^{\sqrt{8/3}\phi} \end{pmatrix}. \tag{2.89}$$

If we include some matter fields Φ_i in the five-dimensional theory,

$$S = \int d^5x\sqrt{-G}\left(\tfrac{1}{2}R + \mathcal{L}_5[\Phi_i; G_{MN}]\right), \tag{2.90}$$

where $\mathcal{L}_5$ is the matter Lagrangian, then the four-dimensional theory will have a matter sector which includes complicated couplings between the gravitational sector fields ϕ and A_μ. In chapters 4 and 5, we will discuss the relation between these couplings and chaos and the equivalence principle.

3.2. Orbifolding. A compactification that is in some respects simpler is the S^1/Z_2 orbifold. Orbifolds are manifolds which are identified under the action of a discrete group. Orbifolds have distinguished points, or hypersurfaces, which are the fixed points of the group action. We consider the S^1/Z_2 orbifold, which is the simplest example, an interval with two boundary points. It is also very important for high energy physics, as the boundary planes which arise when a higher dimensional theory is identified under the S^1/Z_2 action are an essential part of eleven-dimensional $E_8 \times E_8$ heterotic M theory [95, 96] and are crucial components of the ekpyrotic and cyclic cosmologies [112, 170].

To go from five to four dimensions, we identify the fifth dimension x^4 on a circle, with period two, $x^4 \equiv x^4 + 2n$ for integers n. The orbifold identifies $x^4 \equiv -x^4$, which has two inequivalent fixed points, at $x^4 = 0$ and $x^4 = 1$.

The map $x^4 \to -x^4$, however, takes $G_{\mu 5} \to -G_{\mu 5}$ or equivalently $A_\mu \to -A_\mu$. Thus, the orbifolding projects out the constant (*i.e.* massless) Kaluza-Klein vector mode: $A_\mu = 0$. The component G_{55} survives, however, as $G_{55} \to +G_{55}$. The metric (2.89) reduces to

$$G = \begin{pmatrix} e^{-\sqrt{2/3}\phi} h_{\mu\nu} & 0 \\ 0 & e^{\sqrt{8/3}\phi} \end{pmatrix}, \tag{2.91}$$

and the action (2.88) to the comparatively simple form

$$S_{S^1/Z_2} = \frac{1}{2} \int d^4x \sqrt{-h} \big(R_4[h] - (\partial\phi)^2 \big). \tag{2.92}$$

A theory with matter fields $\Phi_{i,0}$ and $\Phi_{i,1}$, and matter Lagrangians $\mathcal{L}_0$ and $\mathcal{L}_1$ on the orbifold planes at $x^4 = 0$ and 1, respectively, has action

$$\begin{aligned} S = \frac{1}{2} \int d^5x \sqrt{-G} R_5[G] + \int \delta(x^4) d^4x \sqrt{-\det G_{\mu\nu}}\, \mathcal{L}_0[\Phi_{i,0}; G_{\mu\nu}] \\ + \int \delta(x^4 - 1) d^4x \sqrt{-\det G_{\mu\nu}}\, \mathcal{L}_1[\Phi_{i,1}; G_{\mu\nu}] \end{aligned} \tag{2.93}$$

$$\begin{aligned} = \int d^5x \sqrt{-h} \Big(\tfrac{1}{2} R_4[h] - \tfrac{1}{2} (\partial\phi)^2 + e^{-\sqrt{8/3}\phi} \mathcal{L}_0[\Phi_{i,0}; e^{-\sqrt{2/3}\phi} h_{\mu\nu}] \\ + e^{-\sqrt{8/3}\phi} \mathcal{L}_1[\Phi_{i,1}; e^{-\sqrt{2/3}\phi} h_{\mu\nu}] \Big), \end{aligned} \tag{2.94}$$

where δ is the Dirac delta function. In the second equality, the compactified form of the action is given. The coupling between the gravitational sector and the matter sector, in this case, is much simpler than in the Kaluza-Klein sector: the compactified matter sector simply couples to a different metric $e^{-\sqrt{2/3}\phi} h_{\mu\nu}$. In chapter 4 we discuss the relevance of these couplings to tests of general relativity. In chapter 5 we mention how the missing vector modes can turn a chaotic theory into a stable one.

Neglecting the orbifold plane Lagrangians, the action (2.94) is the action for a free scalar field coupled to gravity. Thus, in a homogeneous FRW background, the equations of motion are solved by (2.23). At a collision of the fixed planes, when $\phi \to -\infty$, the scale factor on the planes, $e^{-\sqrt{2/3}\phi}a^2$, tends to a constant. This means that when the orbifold planes collide, the scale factor remains finite, as do all physical quantities on the fixed planes (*e.g.* temperature and density). This confirms that ignoring the radiation and tension on the branes is reasonable at the collision. Their contribution to the equations of motion is finite, while the kinetic energy $\frac{1}{2}\dot{\phi}^2$ diverges.

3.3. Randall-Sundrum scenario. The calculations of the last section were idealized: the metric in the extra dimension was treated as homogeneous. This is obviously wrong, because the orbifold planes are localized and we would expect curvature in the extra dimension. Nonetheless, it is a good approximation in the limit of small brane separations. In general, however, it is necessary to consider non-factorizable geometries, in which the four dimensional metric does not depend on the extra dimension through a conformal factor, but rather has a more complicated dependence. However, it is possible to work out some of the dynamics by considering models homogeneous in the spatial directions x^1, x^2 and x^3. Such a geometry has metric:

$$ds^2 = -n(t,y)^2 dt^2 + a(t,y)^2\left((dx^1)^2 + (dx^2)^2 + (dx^3)^2\right) + r(t,y)^2 dy^2. \tag{2.95}$$

A static solution to the Einstein equations, valid on the interval $y \in [-1, 1]$, is

$$n(y) = a(y) = e^{-kb_0|y|} \qquad \text{and} \qquad r = r_0, \tag{2.96}$$

where r_0 and k are positive constants. This solution has a bulk negative cosmological constant Λ, a positive tension brane, with tension T, at $y = 0$ and a negative tension brane, with tension $-T$, at $y = 1$. These are given by

$$T = 6k, \qquad \text{and} \qquad \Lambda = -6k^2. \tag{2.97}$$

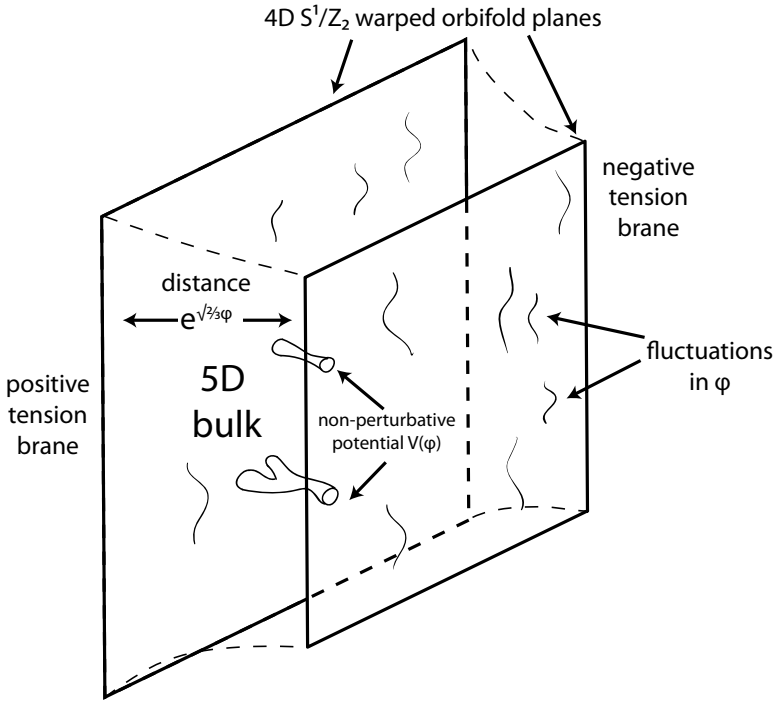

FIGURE 1. A cartoon of the ingredients of a typical five-dimensional S^1/Z_2 orbifold scenario, such as heterotic M-theory, the Randall-Sundrum scenario, or the cyclic scenario. Three-branes, one with positive tension and one with negative tension, are trapped at orbifold planes of a five dimensional bulk (perhaps with a cosmological constant). One of these branes, the visible brane, contains the matter and radiation observed today. The distance between the branes at any point is given in the four-dimensional effective theory by the value of the radion ϕ, whose potential $V(\phi)$, is determined by (unknown) non-perturbative effects. Because of the warped geometry, the positive tension brane has a larger four dimensional scale factor, a_+ than the negative tension brane, a_-.

The quantity $e^{-kr_0|y|}$ is called the warp factor, and measures a scale difference between the two branes: the metric on the positive tension brane is $\eta_{\mu\nu}$, the Minkowski metric, whereas the negative tension brane has the effective metric $e^{-2kr_0}\eta_{\mu\nu}$. The tensions are set by the jump in the derivative of the warp factor at the branes. This solution was discovered by Randall and Sundrum [156] who used it to solve the electroweak scale hierarchy problem: scales on the negative tension brane are exponentially suppressed by the warp factor, so kr_0 can have modest values.

Csáki *et al.* have explicitly derived the Einstein equations for (2.95). They arise from the four-dimensional effective action

$$S_{\rm RS} = -\frac{6}{k}\int dt\, a^3\Big((1-\Omega^2)\frac{\dot a^2}{a^2} + 2kr\Omega^2\frac{\dot a}{a}\frac{\dot r}{r} - k^2r^2\Omega^2\frac{\dot r^2}{r^2} - V(r)\Big) + \int d^4x\sqrt{-g}\,\mathcal{L}_+[g_{\mu\nu}] + \int d^4x\sqrt{-g}\,\Omega^4\mathcal{L}_-[\Omega^2 g_{\mu\nu}]. \quad (2.98)$$

where

$$n(t,y) = e^{-kr(t)y}, \quad (2.99)$$

$$a(t,y) = a(t)e^{-kr(t)y}, \quad (2.100)$$

$$r(t,y) = r(t), \quad (2.101)$$

$$\Omega(t) = e^{-kr(t)}, \quad (2.102)$$

V is the potential from the interaction of orbifold planes, and $\mathcal{L}_+$ and $\mathcal{L}_-$ (which couples to the metric scaled by Ω^2) are the Lagrangian densities of the positive and negative tension branes, respectively. We can change variables in (2.98) to write the metric in a more symmetric form, in terms of the scale factors on the two branes

$$a_+(t) = a(t) \qquad \text{and} \qquad a_-(t) = a(t)\Omega(t), \quad (2.103)$$

which, introducing the lapse function N by the change of variables $dt = Na\,d\tau$, gives the very simple form

$$S_{\rm eff} = \frac{1}{2k}\int d\tau\, N^{-1}6(\dot a_-^2 - \dot a_+^2), \quad (2.104)$$

where we have omitted the potential and the brane actions [110, 112]. The negative tension brane appears with a positive kinetic energy, whereas the positive tension brane has a kinetic energy with negative sign.

When the orbifold planes collide, the warp factor $\Omega \to 1$ and so the solution reduces to the S^1/Z_2 solution in which the two scale factors are equal. In the action (2.104), the

collision corresponds to the line $a_- = a_+$. Moreover, in M-theory, the non-perturbative potential vanishes as the orbifold planes collide, and so it does not affect the geometry of the collision discussed in 3.2. This is discussed in detail in [110].

The static scenario was considered by Randall and Sundrum, whereas the ekpyrotic [112] and cyclic [170, 171] models consider the branes to be dynamical. This is also the case we consider in this thesis. In figure 1 the general situation considered in much of this thesis is illustrated: an S^1/Z_2 orbifold with a warped background and a non-perturbative potential.

3.4. Hořava-Witten scenario. A compactified theory that comes directly from high energy theory is heterotic M theory [95, 96]. This theory, compactified from eleven to four dimensions, seems to be the best way to extract phenomenologically viable models from string theory [196]. Gauge theory and quantum gravity can be unified in this framework. Since it is derived from heterotic M-theory, is a quantum mechanical theory of gravity. Vacua of the theory have been found with standard model gauge groups and three families of quarks and leptons [30, 31, 70].

The four-dimensional, low energy effective action of Hořava-Witten theory, and its supergravity description [192], has been calculated by Lukas *et al.* [126]. Because the couplings will play a role in our discussion in chapter 4, factors of Newton's constant are retained in this section. We record the effective action in the full formalism of [126], although it is principally the interactions of the radion that will be important to us in later chapters.

The action is derived from the full eleven dimensional theory by compactification on the manifold $\mathcal{M}_4 \times \mathrm{CY}_3 \times S^1/Z_2$, where $\mathcal{M}$ is an arbitrary four-manifold and CY_3 is a Calabi-Yau three-fold (*i.e.* it has six real dimensions). The effective action is complicated, because it contains contributions from the decomposition of the M theory three form (with a four form field strength) along various directions, two E_8 gauge supermultiplets on the orbifold planes and moduli from the compactification (the radion and Calabi-Yau moduli).

The action contains a gravity sector, including the radion c and Calabi-Yau volume modulus a, the GUT sector (which arises from E_8 gauge fields in eleven dimensions) and the

universal hypermultiplet. The action is (from [126] equations (81), (114) and (121))

$$S = \frac{\pi\rho V}{\kappa^2}\int_{\mathcal{M}_4}\sqrt{-g}d^4x\Big[R - 18\partial_\mu a\partial^\mu a - \tfrac{3}{2}\partial_\mu\hat{c}\partial^\mu\hat{c} - 3\big(e^{-\hat{c}} + \tfrac{1}{3}\xi\alpha_0 e^{-6a}\big)\partial_\mu C\partial^\mu\bar{C}$$
$$- \tfrac{3}{8}e^{-2\hat{c}}\big(C^pC^q\partial_\mu\bar{C}^p\partial^\mu\bar{C}^q + \bar{C}^p\bar{C}^q\partial_\mu C^p\partial^\mu C^q - 2|C|^2\partial_\mu C\partial^\mu\bar{C}\big)$$
$$- \frac{3k^2}{4}\big(e^{-2\hat{c}-6a} - \tfrac{1}{3}\xi\alpha_0 e^{-\hat{c}-12a}\big)\big(|d_{pqr}C^pC^q|^2 + \tfrac{1}{8}(\bar{C}T^iC)^2\big)\Big]$$
$$- \frac{V}{8\pi\kappa^2}\Big(\frac{\kappa}{4\pi}\Big)^{2/3}\int_{\mathcal{M}_4}\sqrt{-g}\,d^4x\big((e^{6a} + \xi\alpha_0 e^{\hat{c}})\,\mathrm{tr}(F^{(1)})^2$$
$$+ (e^{6a} - \xi\alpha_0 e^{\hat{c}})\,\mathrm{tr}(F^{(2)})^2\big), \quad (2.105)$$

where we have set two moduli that arise from the bulk three form (σ and χ) to zero. Now g is the four-dimensional metric, R is the associated scalar curvature, the separation of the branes is given by $\pi\rho e^c$, the volume of the Calabi-Yau manifold is given by Ve^{6a} and the modulus $\hat{c} = c + 2a$. The C are scalar fields transforming in the **27** of E_6, $F^{(1)}$ and $F^{(2)}$ are the gauge fields strengths on the fixed planes. The constants are κ^2, the eleven-dimensional gravitational coupling, $k = 4\sqrt{2\rho}\pi(4\pi/\kappa)^{1/3}$, $\xi = \sqrt{2}\pi\rho/16$, the T^i ($i = 1,\ldots,78$) are E_6 generators in the fundamental representation of **27**, d_{pqr} is the tensor that projects out the singlet of $\mathbf{27}^3$. Finally, α_0 is the distortion of the Calabi-Yau three-fold X

$$\alpha_0 = -\frac{1}{\sqrt{2}\pi V}\Big(\frac{\kappa}{4\pi}\Big)^{2/3}\int_X \omega\wedge\mathrm{tr}\,R^{(X)}\wedge\mathrm{tr}\,R^{(X)}, \quad (2.106)$$

where ω is the Kähler form and the Ricci tensor $R^{(X)}$ is calculated with respect to X.

Newton's constant and the GUT coupling for this theory are

$$G = \frac{\kappa^2}{16\pi^2 V\rho} \qquad \text{and} \qquad \alpha_{\mathrm{GUT}} = \frac{(4\pi\kappa^2)^{2/3}}{2V}. \quad (2.107)$$

Phenomenology suggests, roughly, $V^{1/6}/\kappa^{2/9} \sim 2$ and $\pi\rho/\kappa^{2/9} \sim 8$ [13]. The GUT mass, in a reasonably isotropic Calabi-Yau manifold is given by $M_{\mathrm{GUT}} = V^{-1/6} \sim 10^{16}\,\mathrm{GeV}$. This suggests $\kappa^{2/9} \sim 5\times10^{-16}\,\mathrm{GeV}$. Witten [196] has pointed out that in this case, the effective action (2.105) is an expansion in the parameter $\varepsilon = \kappa^{2/3}\rho/V^{2/3}$. This expansion breaks

down for $\varepsilon \sim 1$. In this case, the Calabi-Yau becomes so distorted along the orbifold interval that its volume becomes negative. Including higher terms in $\kappa^{2/3}$ ameliorates the situation somewhat [52] but it is not known what effect higher order terms in the scalar curvature (such as R^4 terms) will have [6].

An important feature of this solution is that it must have a non-vanishing bulk three form flux. This prevents the theory from having the anti-de Sitter bulk of the Randall-Sundrum theory (2.98) and gives, to the order considered by Lukas *et al.*, a *linear* ($\Omega(y) \propto 1 + \alpha|y|$), rather than exponential warping ($\Omega(y) \propto e^{C|y|}$). It [52] it is pointed out that higher order corrections give a quadratic warp factor. Including higher order scalar curvature terms, such as R^4, also does not give an exponential warp factor [6]. Therefore, it unfortunately does not seem easy to achieve the Randall-Sundrum solution to the weak hierarchy problem in this model [92], although it does explain the smaller hierarchy between the GUT and Planck scales in (2.107). The form of the warp factor will be significant for our discussion of the equivalence principle in chapter 4.

Considerable additional technology has been developed for compactification, particularly sophisticated techniques in algebraic geometry [90, 91]. However these effective actions provide the broad features we will need for applications to cosmology.

CHAPTER 3

The speed of sound of dark energy

Recent evidence suggests that most of the critical density of the universe is made up of a component with large, negative pressure. Determining the nature of this dark energy component is one of the central problems of modern cosmology. It is not known if the origin of dark energy is a cosmological constant (such as a field theory vacuum density) or a dynamical component, such as quintessence [198]. Both approaches are beset by difficult theoretical problems. The cosmological constant, although the simplest theory, is fine-tuned by roughly 120 orders of magnitude compared to the Planck density expected from field theory. Barring an anthropic explanation [190] it is not known how such a small constant could be obtained from fundamental physics. Quintessence has the advantage that the scale of dark energy can be determined dynamically by "tracker" solutions, but it is unclear how the theory would be protected from quantum mechanical corrections which would give the field a large mass and couplings to other fields.

Quintessence is modeled by a scalar field slowly rolling down a flat potential [38, 157, 198]. The scalar field may be regarded as real, or simply as a device for modeling more general cosmic fluids with negative pressure. These models can be distinguished from a cosmological constant, at least in principle, by the equation of state $w = p/\rho$. A cosmological constant has $w = -1$: it has constant energy density. A scalar field generally has w different from unity and time-varying. Measurements of supernovae [120, 150, 158, 159], large-scale structure [175] and the cosmic microwave background [167] all constrain the equation of state.

These observations of dark energy are all indirect, however. The redshift-distance relation obtained by studying type Ia supernova can, in principle, measure the expansion history of the universe very precisely. However, this provides only an indirect constraint on the equation of state. The relationship between redshift z and luminosity distance d_L (the ratio of luminosity to flux for a distant object) is

$$d_L = \frac{1+z}{H_0}(1+\Omega_{\rm m}/\Omega_{\rm Q})^{1/2}\int_1^{1+z}\frac{dx}{x^{3/2}}\Big[\frac{\Omega_{\rm m}}{\Omega_{\rm Q}}+\exp\Big(3\int_1^x\frac{dy}{y}w_{\rm Q}(y)\Big)\Big]^{-1/2}, \qquad (3.1)$$

where $w_{\rm Q}(z)$ is the equation-of-state of dark energy at redshift z, and $\Omega_{\rm m}$ and $\Omega_{\rm Q}$ are the fractional densities (2.11) of matter and quintessence today. This multiple-integral relationship means that the redshift-luminosity relation is principally sensitive to an average, effective equation of state $\bar{w}$, and is quite insensitive to the fine details of its time evolution [131, 132]. If the equation of state of dark energy is different from -1, there is no reason to expect it to be constant. Observations of supernovae, then, reveal little detailed information about the underlying quintessence. This situation may not improve dramatically with better observations, because even large collections of supernovae are still restricted by underlying systematic errors. It is possible that current methods will not be able to constrain w to better than 10% accuracy [182].

Therefore it is important to develop complementary techniques to confront the dark energy problem. If dark energy in our universe is dynamical, then it is possible that there are other observable effects that would distinguish it from a cosmological constant. In this chapter, the effect of another parameter, the sound-speed of dark energy – *i.e.* the speed with which dark energy perturbations propagate – is investigated. In the next chapter, the role of tests of the equivalence principle in constraining dark energy is investigated.

In standard models of quintessence [38, 84, 149, 157, 198] the scalar field ϕ has a canonical kinetic term, $X = -\frac{1}{2}(\partial_\mu\phi)^2$, as in (1.6). A second category consists of models in which the kinetic energy is not canonical and could be a general function of X. While these have not often been considered in the literature, they are well-defined Lorentz invariant field theories

with (nonlinear) second-order equations of motion. They are significant, because they are corrections that are expected to occur in brane and string models. A well-known example of such an action is the scalar Born-Infeld action [21],

$$S = -\int d^4x\sqrt{-g}f(\phi)\sqrt{1-\alpha X}, \tag{3.2}$$

where α is a constant and f some function of ϕ. This action is used in string theory as a higher order generalization of p-forms actions (in this case, a 0-form or scalar field). It is also the action describing the tension of a three dimensional brane displaced in an extra dimension X^4 by an amount proportional to $X^4 = \phi$.

A prominent class of such models are k-essence models, which are designed to address the issue of why cosmic acceleration has begun only recently [10, 11]. The equation of state in k-essence models is positive, and mirrors the background equation of state, until the onset of matter-domination triggers a change to negative pressure. A key difference between standard quintessence and k-essence models is the time-evolution of the equation of state. The equation of state for a k-essence component approaches -1 soon after the onset of matter-domination and then increases towards a less negative value in the present epoch as the component begins to dominate the energy density. In standard quintessence tracker models, the equation of state is generically monotonically decreasing and approaching -1 today. Quintessence relies on a particular form of the potential for its attractor behavior [173], whereas these models rely on dynamical attractor behavior that comes from the non-canonical kinetic energy density. This feature was discussed in detail in [11].

In this chapter, we focus on a second physical property – the speed of sound – which also distinguishes standard quintessence from k-essence and, more generally, from other cosmic fluids described by a non-canonical kinetic energy density. The sound speed greatly influences the fluctuations of the quintessence fluid and can also, in principle, have an effect on the cosmic microwave background (CMB) and matter power spectrum. We investigate how the variable speed of sound influences the fluctuations of the CMB compared to the case

of standard quintessence where $c_s^2 = 1$. In general, the effect is small, but we show that it is detectable in cases like k-essence models in which the speed of sound is nearly zero during most of the period between last scattering and the present epoch. This behavior produces the greatest difference from standard quintessence [98, 99].

After describing how the speed of sound arises in cosmic fluids and scalar field models, we compare models with exactly the same equation of state as a function of redshift, $w(z)$, but different sound speed. We find that models with near-zero sound speed today (such as k-essence models) are distinguishable from models with $c_s^2 = 1$ based on measurements of the CMB power spectrum, provided the quintessence energy density is greater than a few percent of the critical density at last scattering. The density requirement, which is satisfied by typical k-essence models, for example, is needed so that the sound speed has a measurable effect on the acoustic oscillation peaks of the CMB which are sensitive to conditions at the last scattering surface. Similar results can be obtained for more general forms of dark energy [43, 98, 99]. We then consider whether the effect can be mimicked by varying other cosmic parameters or by introducing a time-varying equation-of-state. To perform the studies, we introduce a spline technique that is useful in exploring models with time-varying w. Our conclusion is that the sound speed effect is distinguishable from all other standard parameter effects. Hence, the CMB can provide a useful constraint on the sound speed of dark energy.

This chapter is largely based on work in [79]. Subsequently, DeDeo *et al.* [64] showed that related effects due to the sound speed of quintessence may occur in the matter power spectrum. Bean and Doré [15] tested the model against WMAP data but found that it does not yet clearly discriminate sound speed. Hannestad [93] did a joint analysis with large scale structure and supernovae and likewise found that current data do not differentiate the sound speed. Both Hannestad and Bean and Doré have pointed out that the ability to constrain the sound speed is diminished if w is close to -1.

1. The sound speed of a scalar field

The perturbed equation of motion for a scalar field against a background $\phi(t)$ is

$$-\Box\delta\phi = -V_{,\phi\phi}\big(\phi(t)\big)\delta\phi, \tag{3.3}$$

omitting the back-reaction on the metric. Since the covariant d'Alembertian is applied to ϕ it always propagates with the speed of light: $c_s^2 = 1$. A perfect fluid, however, $w = c_s^2$ if w is a constant, or, generally (2.73),

$$\dot{w} = 3H(1+w)(w - c_s^2). \tag{3.4}$$

This comes from $c_s^2 = \dot{p}/\dot{\rho}$. For a perfect fluid, the behavior of perturbations is closely tied to the equation of state, which is a property of the background evolution.

Let us now consider a scalar field with a more general Lagrangian density $\mathcal{L}_\phi(X, \phi)$ and action

$$S_\phi = \int d^4x\sqrt{-g}\,\mathcal{L}_\phi(X, \phi), \tag{3.5}$$

where[1] $X = -\frac{1}{2}(\partial\phi)^2$. In order for the equations of motion to make sense, $\mathcal{L}_\phi$ must be a differentiable function of X. The canonical scalar field has $\mathcal{L}_\phi = X - V$. The stress-energy tensor is

$$T_{\mu\nu} = \mathcal{L}_{\phi,X}\partial_\mu\phi\partial_\nu\phi - \mathcal{L}_\phi g_{\mu\nu}, \tag{3.6}$$

so comparing with the perfect fluid stress energy tensor (2.3) the field has pressure, energy density and equation of state

$$p_\phi = \mathcal{L}_\phi, \tag{3.7}$$

$$\rho_\phi = 2X\mathcal{L}_{\phi,X} - \mathcal{L}_\phi, \tag{3.8}$$

$$w_\phi = \frac{\mathcal{L}_\phi}{2X\mathcal{L}_{\phi,X} - \mathcal{L}_\phi}, \tag{3.9}$$

[1]Our definition of X differs from [8–11, 79] because we use the opposite sign convention for the metric.

respectively. The general equation of motion is complicated. However, if we assume a homogeneous background $\phi = \phi(t)$ and $H = H(t)$, it simplifies considerably:

$$\ddot{\phi}(\mathcal{L}_{\phi,X} + 2X\mathcal{L}_{\phi,XX}) + 3H\dot{\phi}\mathcal{L}_{\phi,X} + 2X\mathcal{L}_{\phi,X\phi} = \mathcal{L}_{\phi,\phi}. \tag{3.10}$$

The perturbed equations of motion about this background are

$$-(\mathcal{L}_{\phi,X} + 2X\mathcal{L}_{\phi,XX})\delta\ddot{\phi} + \mathcal{L}_{\phi,X}a^{-2}\nabla^2\delta\phi + \cdots = 0, \tag{3.11}$$

where a is the scale factor and the dots represent terms of lower order in the derivatives. Thus, the speed of sound of this field is

$$c_s^2 = \frac{\mathcal{L}_{\phi,X}}{\mathcal{L}_{\phi,X} + 2X\mathcal{L}_{\phi,XX}}, \tag{3.12}$$

which is equal to unity only when $X = 0$ or $\mathcal{L}_{\phi,XX} = 0$, as for a canonical scalar field.

The Born-Infeld action (3.2) reduces, after a field redefinition, to the canonical action in the slowly varying limit $|\alpha X| \ll 1$. However, when αX is larger, it begins to differ appreciably. The model has equation of state $-1+\alpha X$ and sound speed $1-\alpha X$. Abramo *et al.* [1] have used CMB data to try to see if the a Born-Infeld action is preferred over the standard Klein-Gordon scalar field. They determined that without the specific attractor behavior of *k*-essence, the dark energy density at decoupling is insufficient to distinguish the models in a statistically significant way.

A potential concern is that the sound speed (3.12) can be greater than unity. Indeed, (3.2) has $c_s^2 > 1$ for $\alpha < 0$ in a homogeneous ($X > 0$) background. If $\mathcal{L}_\phi$ is not everywhere a convex function of X, then there will be intervals on which $c_s^2 > 1$. Upon closer examination, one can see that this condition is physically allowed: this means that perturbations of the background scalar field can travel faster than light as measured in the *preferred* frame in which the background field has vanishing spatial gradient. For a time dependent background field, this frame is uniquely defined and means the perturbation theory is not Lorentz invariant. There is no violation of causality. The underlying theory is manifestly

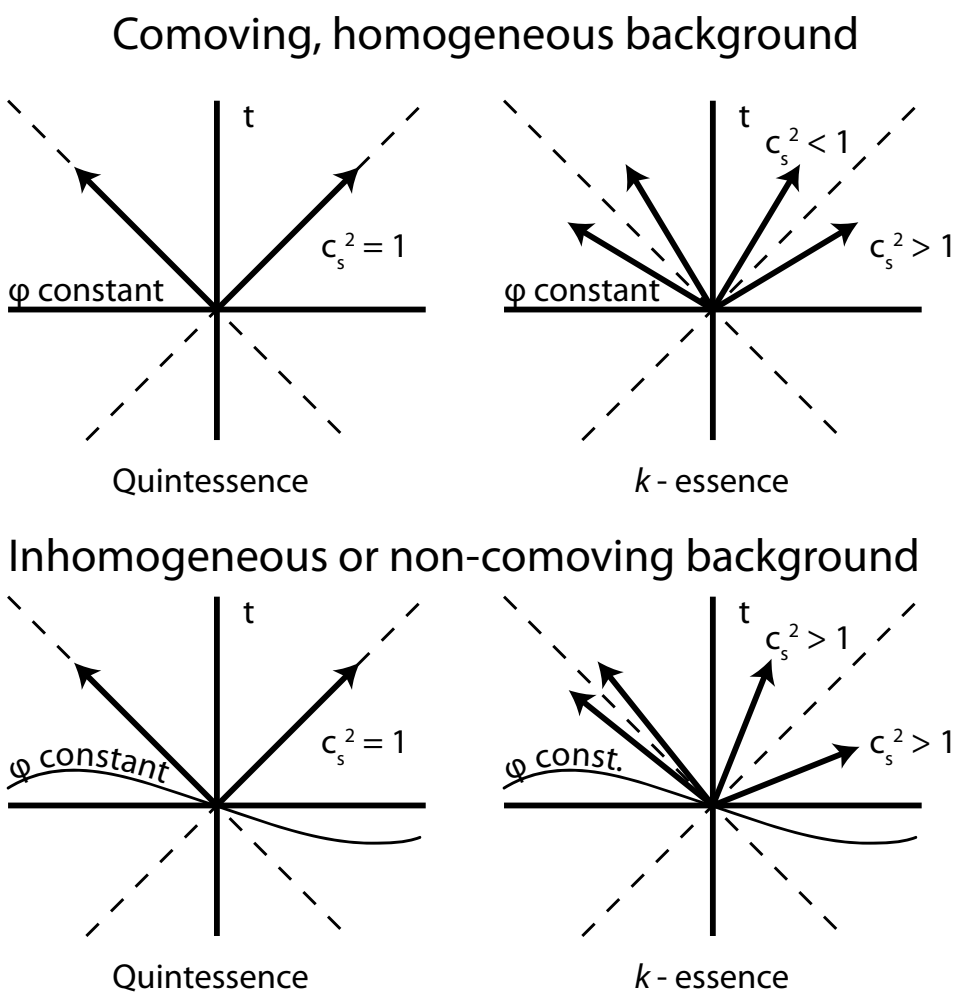

FIGURE 1. The propagation of perturbations for a scalar field with canonical kinetic energy (left) and non-canonical kinetic energy (right). In a homogeneous background (top) perturbations of both fields propagate isotropically. The perturbations of the canonical field propagate on the light cone, but the non-canonical field has sound speed generically different from the speed of light (depending on the Lagrangian, either faster or slower than light). In an inhomogeneous background (bottom) – or a homogeneous, time-dependent background seen from an accelerated frame – the perturbations of the canonical field still propagate on the light cone. The perturbations of the non-canonical field propagate anisotropically, with the speed increased in one direction and decreased in the other. Since perturbations in the $c_s^2 > 1$ version do not propagate along arbitrary spacelike trajectories, there are no violations of causality from closed causal curves in this model.

Lorentz invariant and it is not possible to transmit information faster than light along *arbitrary* space-like directions or create closed time-like curves. While (3.10), (3.11) and (3.12) were derived in a preferred frame, in which the background is homogeneous, it is possible to use Lorentz invariance to see that the field perturbations propagate anisotropically in a general background. See figure 1. The condition that $c_s^2 \leq 1$ everywhere is the condition that $\mathcal{L}_\phi$ be a convex function of ϕ: $\mathcal{L}_{\phi,XX} \geq 0$. In this chapter, our results depend only on the $c_s^2 \sim 0$ behavior of k-essence at moderate redshifts, and not on any $c_s^2 > 1$ behavior. For a more thorough discussion of the stability of these models, see [9].

2. Computing the cosmic microwave background

The cosmic microwave background is a 2.7 Kelvin thermal relic of the big bang. It is emitted at decoupling, when ions in the primordial plasma combine with free electrons to form atoms, and the Compton scattering of radiation ceases. This happened at a redshift of roughly $1,000$. The radiation has traveled relatively unimpeded until today, so the observed CMB is an image of the last scattering surface, the spherical surface surrounding us at this large redshift. Minute anisotropies, due to fluctuations in the photon temperature at decoupling, are encoded in the CMB. By decomposing the temperature of the CMB into spherical harmonics, the power spectrum may be studied. The power spectrum comes from a nearly-scale invariant spectrum of primordial perturbations multiplied by a transfer function describing the evolution of perturbations after they reenter the horizon. These are described in section 2 of chapter 2.

The temperature on the sky of the CMB can be written as $T_0(1 + \Theta)$, where T_0 is the background temperature and the fluctuation Θ depends on the direction being observed. But Θ can be decomposed into spherical harmonics Y_{lm},

$$\Theta = \sum_{l=1}^{\infty} \sum_{m=-l}^{l} a_{lm} Y_{lm}, \tag{3.13}$$

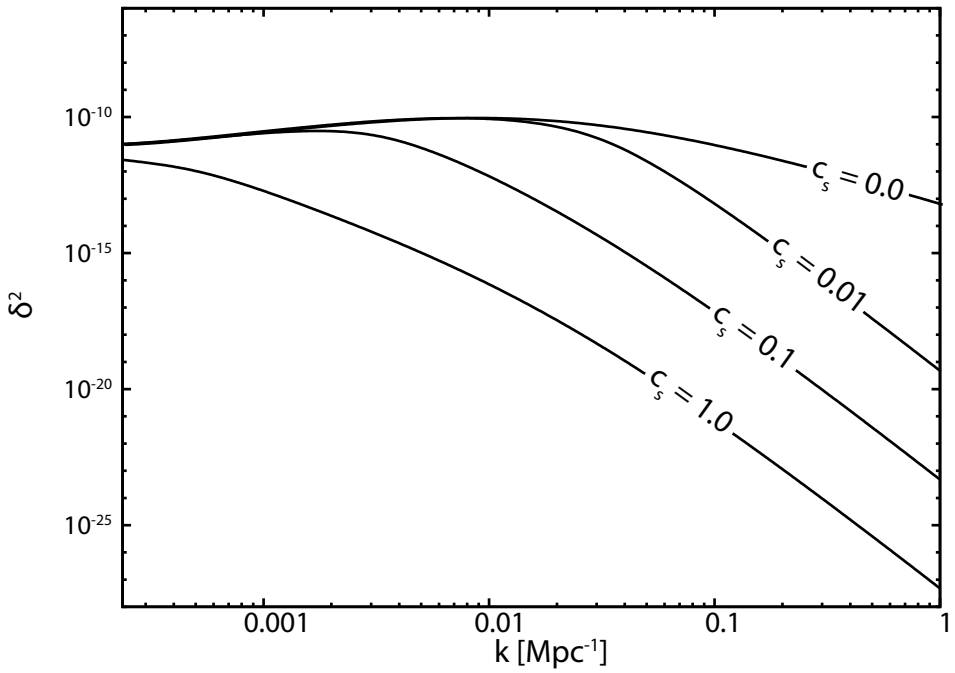

FIGURE 2. The effect of the speed of sound on scalar field perturbations is shown, for a model with $w = -0.75$ and $\Omega_Q = 0.71$ today. The horizon, at 5000 Mpc, corresponds to the left axis. The scale c_s/H_0 acts as a cutoff for the quintessence power spectrum. For larger wavenumbers, the spectrum falls off very rapidly, as $\delta^2 \sim k^{-6.6}$, regardless of the equation of state.

where the coefficients a_{lm} are independent Gaussian random variables, as the initial conditions discussed in section 2.1 are a Gaussian random field. For fixed l, they are identically distributed. If this picture is correct, the only information in the CMB is contained in C_l, the variance of the a_{lm} for fixed l. Since there are only $2l+1$ of the a_{lm} for a given l, there is a fundamental limit, the *cosmic variance*, to how well this variance C_l can be measured from our observations of the CMB, given by

$$\frac{\Delta C_l}{C_l} = \sqrt{\frac{2}{2l+1}}. \tag{3.14}$$

This is used to compute the absolute likelihood, given perfect observations, of being able to detect the speed of sound effects.

The effect of the speed of sound on the CMB perturbation equations comes from the fact that, for $c_s^2 \ll 1$, k-essence will collapse via gravitational instability into cold dark matter

(CDM) potentials, whereas in the $c_s^2 \sim 1$, the growth of density perturbations is strongly suppressed. This can be seen from (2.74). The term

$$\frac{1}{3}\Big(\frac{k}{aH}\Big)^2 \frac{c_s^2}{1+w}\delta, \tag{3.15}$$

where δ is the density contrast, sets the scale of the Jeans instability: clustering on scales smaller than c_s/H is suppressed by pressure effects. For $c_s = 1$, clustering inside the horizon is heavily suppressed, while for $c_s = 0$ there is no scale on which it is suppressed (figure 2).

The power spectrum for a given cosmological model may be computed by integrating the Boltzmann equation for the various cosmic fluids. The CMB power spectra for our models are computed by modifying the standard CMBFAST code [38, 163]. These calculations are performed in synchronous gauge, which is written

$$ds^2 = a^2(\tau)[-d\tau^2 + (\delta_{ij} + h_{ij})dx^i\, dx^j], \tag{3.16}$$

where δ_{ij} is the unperturbed spatial metric, and h_{ij} is the metric perturbation. This choice of gauge is efficient for solving the Boltzmann equations, but has the disadvantage that the gauge slicing (constant τ surfaces) is not uniquely defined: the surface $\tau = 0$ is arbitrary. It is usually chosen so that the threading (the lines of constant x) follow the geodesics of a particular particle species. In the case of CMBFAST, these are the dark matter particles. The Newtonian interpretation of synchronous gauge is also not as straightforward as in the longitudinal gauge of section 2. See [100, 130] for a detailed comparison of the two gauges, and formulae for transforming variables between one and the other.

The modifications of CMBFAST are straightforward. We use h to represent the trace of the spatial metric perturbation h_{ij}. The effect we are examining is due to the perturbations

to the k-essence stress-energy in the synchronous gauge for a mode with wavenumber k

$$\delta\rho = -2\rho\frac{\delta\phi}{\phi} - (\rho + p)\frac{\delta y}{y}c_s^{-2} \tag{3.17}$$

$$\delta p = -2p\frac{\delta\phi}{\phi} - (\rho + p)\frac{\delta y}{y} \tag{3.18}$$

$$\theta = \frac{1}{\sqrt{2}}k^2 y\delta\phi \tag{3.19}$$

where $y \equiv 1/\sqrt{X}$ and the θ variable of CMBFAST is the divergence of the fluid velocity, so $k^i T^0{}_i = \theta(\rho + p)$ (see (2.42)). It is related to the variable δH of section 2.2 of chapter 2 by $\theta = 3a\delta H$. The density contrast, $\delta \equiv \delta\rho/\rho$, obeys the equation

$$\delta' = -(1 + w)\left(\theta + \tfrac{1}{2}h'\right) - 3\mathcal{H}\left(\frac{\delta p}{\delta\rho} - w\right)\delta, \tag{3.20}$$

where the derivative is with respect to conformal time and $\mathcal{H} \equiv a'/a$. The quantity $\delta p/\delta\rho$ appears, and is generally different from c_s^2 (see chapter 2, section 2.2). In fact it can be expressed in gauge invariant form by

$$\delta p = c_s^2\delta\rho + \frac{\theta\rho}{k^2}\left[3\mathcal{H}(1 + w)(c_s^2 - w) + w'\right], \tag{3.21}$$

so they agree only on small scales (or for perfect fluids, for which the second term vanishes). This leads to a simplified evolution equation for the velocity gradient

$$\theta' = (3c_s^2 - 1)\mathcal{H}\theta + c_s^2 k^2\delta/(1 + w). \tag{3.22}$$

Kinetic quintessence is distinguished from regular quintessence, for which $c_s^2 = 1$ in equations (3.21) and (3.22). Linearized perturbations in k-essence can propagate non-relativistically, with $c_s^2 \ll 1$. We can see in equation (3.22) that a small sound speed will cause the velocity gradient to decay; with the conventional gauge choice that $\theta_{\rm CDM} = 0$, the inhomogeneities in the k-essence will describe a fluid which is comoving with the cold dark matter. From equation (3.21), we see that the second term on the right hand side will be negligible even on scales approaching the horizon. The overall effect is that the pressure

fluctuations δp are too weak to prevent k-essence collapse via gravitational instability into the CDM gravitational potentials.

The CMBFAST code takes $w(a)$ and $c_s(a)$ as inputs, so it is possible to manually adjust these functions to have any values (including, of course, $c_s = 1$). Once we have computed the CMB anisotropy for two models, they can be compared by computing their likelihood difference, the probability that they could be confused due to the cosmic variance in local measurements of the CMB. Given models A and B, the negative log-likelihood, $-\log L$, is derived in [101] from (3.14):

$$-\log L = \sum_l (l+\tfrac{1}{2})\Big(1 - \frac{C_{lA}}{C_{lB}} + \log\frac{C_{lA}}{C_{lB}}\Big). \tag{3.23}$$

The condition $-\log L > 6$ corresponds to distinguishability at the 3σ level. The relative normalization of the spectra is chosen so as to minimize the likelihood difference, including l up to 1500.

3. Measuring the speed of sound of quintessence

In the case of k-essence the Lagrangian generically has the form $\mathcal{L}_\phi = \tilde{p}(X)/\phi^2$, where $\tilde{p}(X)$ is some function with $\tilde{p}_{,XX} \neq 0$. We have chosen a specific form for $\tilde{p}$ for our fiducial model. The conditions $\tilde{p}$ must satisfy are discussed in [11]. The equation of state and sound speed for this model are shown in figures 3 and 4 respectively; they are expressed as functions of the scale factor a ($a = 1$ today) by integrating the equations of motion. To completely specify the model, we fix the cosmological parameters today to reasonable values: $\Omega_{\rm b} = 0.05$, $\Omega_{\rm CDM} = 0.3$, $\Omega_{\rm Q} = 0.65$ and $h_0 = 0.5$ (where h_0 is the Hubble parameter in units of $100\,{\rm km\,sec^{-1}\,Mpc^{-1}}$). The energy density as a fraction of the critical density, $\Omega_{k\text{-essence}}$, is shown in figure 5. While $c_s^2 > 1$ for large redshifts, this is not an important feature of the model: Ω_Q is reasonably small whenever $c_s^2 > 1$, so the value of c_s at these times has negligible effect on the CMB. We have verified this by rerunning the calculation after artificially truncating the speed of sound at $c_s \leq 1$ and comparing the CMB power spectra.

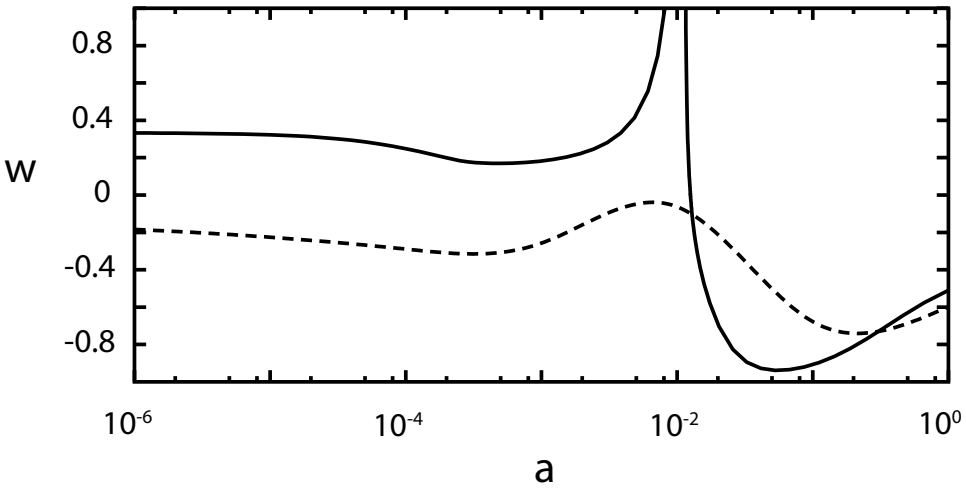

FIGURE 3. The equation of state for the fiducial model (solid line) and the closest spline fit (dashed line) quintessence model. The spline fit is chosen so that the CMB spectrum matches the fiducial model to within the cosmic variance limit. However, the spline itself is smoother than the actual $w(a)$. As discussed in the text, there is a large degeneracy in the spline parameters, so the spline need not mimic the fiducial equation of state very closely to obtain a good match to the CMB spectrum.

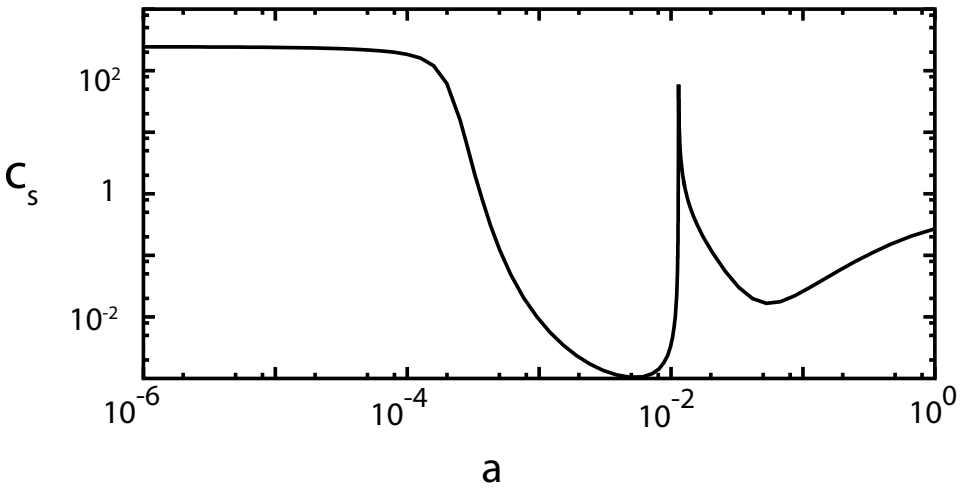

FIGURE 4. The speed of sound as a function of the scale factor a ($a_{\rm today} = 1$). Note that at $c_s^2 \ll 1$ near the last scattering surface and at the present epoch.

Moreover, with slightly different parameters, we can obtain a model in which $c_s < 1$ which at early times and which has the same behavior at late times.) The most important feature of

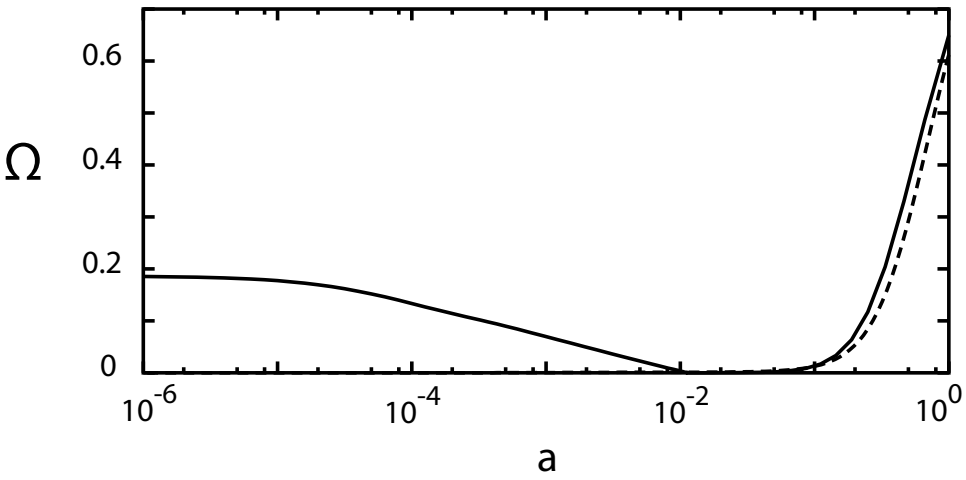

FIGURE 5. $\Omega_{k\text{-essence}}$ and Ω_Q as a function of z for the k-essence (solid line) and best-fit spline models (dashed line), respectively. Note that Ω_Q falls off to zero at large redshifts, whereas $\Omega_{k\text{-essence}}$ approaches a finite value; w approaches 1/3 for this model at large redshifts.

the fiducial model is that $c_s^2 \ll 1$ whenever k-essence contributes significantly to the energy density of the universe.

The power spectrum for our fiducial model is given by the solid line in figure 6. We can see the effect of the unusual speed of sound on this model by computing a new spectrum, which has the same equation of state and cosmological parameters, but whose speed of sound has been set equal to 1 for all a. This corresponds to a quintessence field – with canonical kinetic energy – rolling down a potential. The power spectrum for the model is shown in figure 6. The two models have log likelihood difference $-\log L = 127$: they are easily distinguishable. The speed of sound has a significant effect on the CMB anisotropy. Figure 7 compares the dark matter and dark energy contributions for quintessence and k-essence models. The small sound speed results in distinctive oscillations in the case of k-essence.

Can the effect of the sound speed be distinguished from that of other cosmological parameters? There is already a large degeneracy [101] in these parameters, so it would not have been too surprising if allowing a variable speed of sound merely expanded the pre-existing degeneracy. This problem is addressed by considering, as above, quintessence models which have $c_s^2 = 1$, but allowing the values of Ω_b, Ω_{CDM}, Ω_Q (quintessence) and h_0 to vary (subject

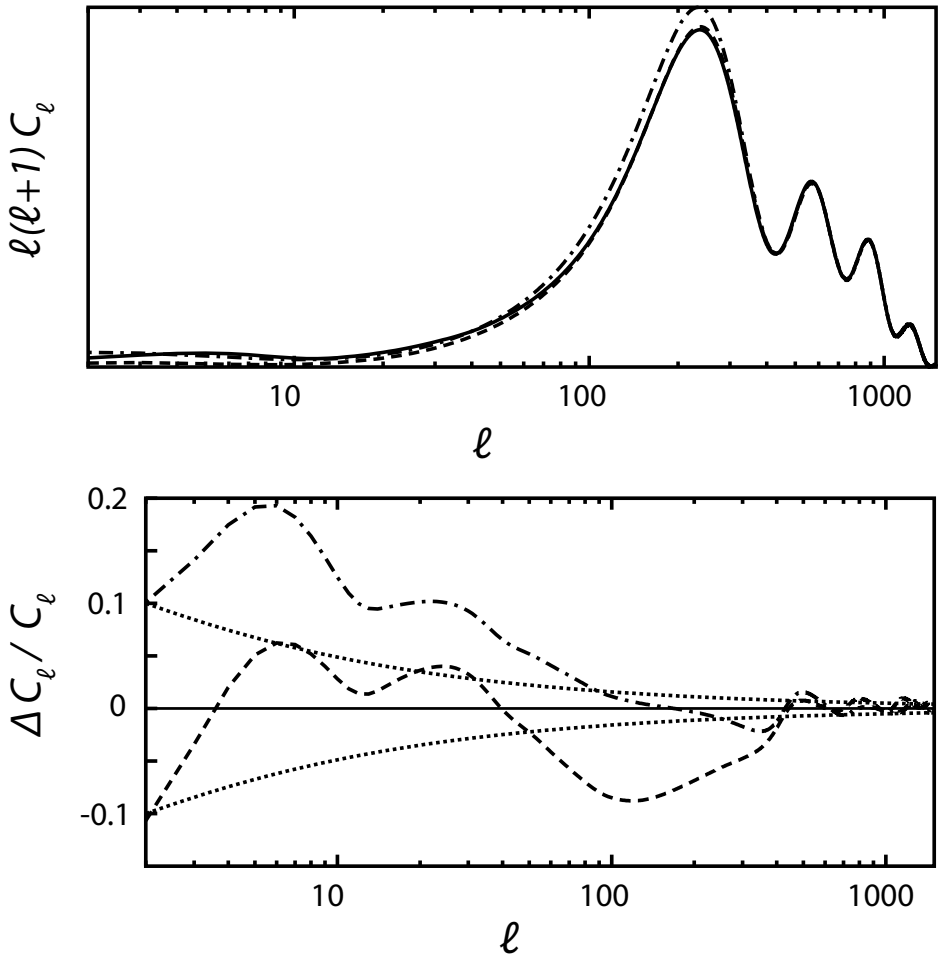

FIGURE 6. The CMB power spectrum (with no particular normalization) for k-essence (solid line), the model with $c_s = 1$ (dot-dashed line, $-\log L = 127$) and the best-fit spline model (dashed, $-\log L = 28$). Models that are distinguishable can still fit quite closely. The lower diagram shows $\Delta C_l / C_l$ (relative to the fiducial model) and the cosmic variance envelope.

to the flatness condition $\Omega_{\rm b} + \Omega_{\rm CDM} + \Omega_{\rm Q} = 1$). For the comparison models, the equation of state, w, is taken to be constant as a function of the scale factor a, but is allowed to vary from model to model. Minimizing the log-likelihood over these parameters, the best fit gives $-\log L = 32$, with parameters $\Omega_b = 0.05$, $\Omega_{\rm CDM} = 0.35$, $\Omega_Q = 0.60$, $h_0 = 0.48$ and $w = -0.78$. This fit was found using well known minimization schemes [154]. It seems likely that $-\log L = 32$ is the best that can be done for this class of models, as the result is quite

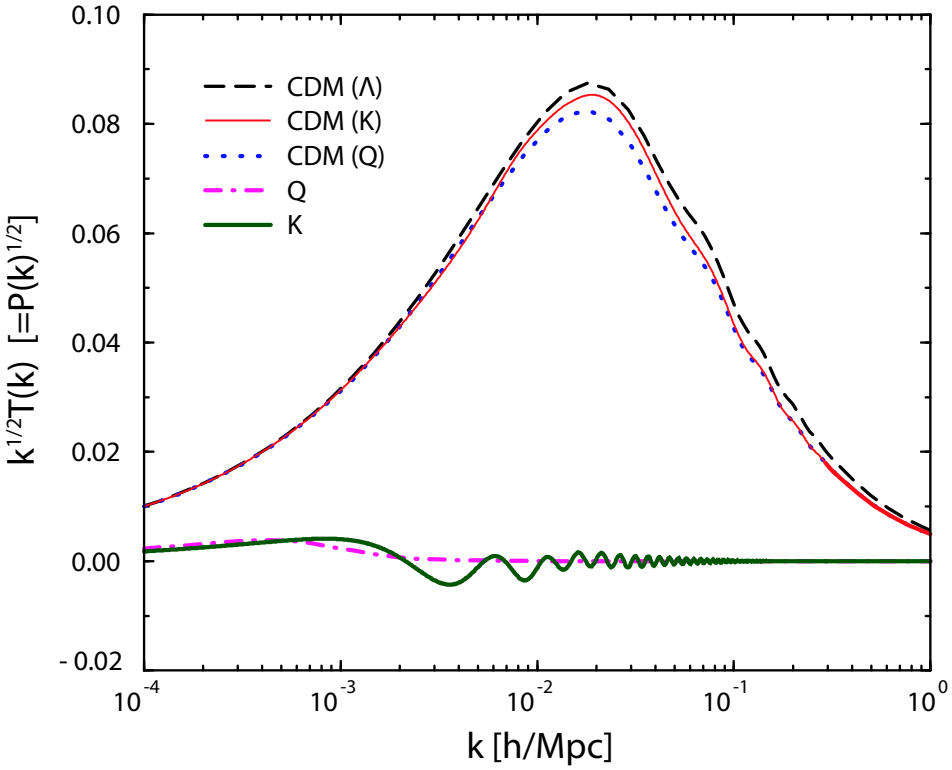

FIGURE 7. Comparison of the power spectra for dark matter (CDM) and dark energy (Q or K) for models with dark energy in the form of cosmological constant (Λ), quintessence (Q) and k-essence (K). All models have $\Omega_{\rm CDM} = 0.3$, $\Omega_{\rm b}h^2 = 0.02$, and $h = 0.65$. For models with cosmological constant, there is no perturbed dark energy component. Note the distinctive oscillations of the k-essence component associated with having $c_s \ll 1$. These are discussed further by DeDeo *et al.* [64].

insensitive to the parameter values with which the minimization is started or to which minimization algorithm (conjugate gradients or a simplex annealing method) is used. Hence, the speed of sound *is* distinguishable.

Thus far, our fiducial model has been compared with two kinds of $c_s = 1$ models: one with all other parameters, including $w(a)$, identical and one with constant w, with the parameters (including w) adjusted so as to minimize the likelihood difference. In both cases the fiducial model was easily distinguishable. We next test the possibility that some other form of quintessence, with general $w(a)$ and canonical kinetic energy, can reproduce the CMB anisotropy of the time-varying c_s model. For this purpose, we consider models with

an equation of state given by a cubic spline [154]. That is, we introduce six new parameters into our model: the values of w at $a = 10^{-4}$, 10^{-3}, 10^{-2}, 10^{-1} and w at the two extremes of a, which lie at $a = 10^{-14}$ and 3.8. (Introducing more spline points has a negligible effect on our results.) The equation of state at other values of a is then given by a piecewise cubic (in $\log a$) function whose coefficients are chosen so that it passes through the points selected above and has a continuous second derivative. We still allow $\Omega_{\rm b}$, $\Omega_{\rm CDM}$, $\Omega_{\rm Q}$ and h_0 to vary (again enforcing flatness), and now allow, for completeness, the spectral tilt n_s to vary as well. The model therefore has a total of ten free parameters. The minimum log-likelihood difference found was $-\log L = 28$, which is not a significant improvement. This model has $\Omega_{\rm b} = 0.05$, $\Omega_{\rm CDM} = 0.34$, $\Omega_{\rm Q} = 0.61$, $h_0 = 0.47$, $n_s = 1.02$. The equation of state is shown in figure 3 (dashed line), Ω_Q as a function of z in figure 5 (dashed line) the CMB power spectrum is shown in figure 6 (dotted line).

The spline technique always produces a very smooth looking equation of state, compared to the rapidly varying equation of state (particularly near $a = 0.01$) given by the actual k-essence model. To see that this does not affect the analysis – that the spline equation of state has sufficient freedom to closely mimic that for the fiducial model – we compare two models to the fiducial model, one with $c_s = 1$ and one with an identical speed of sound to the k-essence model. The equations of state of these models are allowed to vary using the spline technique, but the cosmological parameters are fixed to be the same as for the fiducial model. The minimum $-\log L$ for the $c_s = 1$ model is 70, whereas the log-likelihood difference for the model with the fiducial sound speed is much less than one. Thus, the spline technique appears to do a very good job of modeling the relevant details of the equation of state. Since the equation of state is constrained so indirectly by cosmological observables such as the redshift-luminosity relation (3.1), it is not surprising that such a parameterization is good at exploring the available equations of state.

The cubic spline is a useful method for studying varying a general equation of state $w(a)$ with a finite number of fitting quantities. The same technique can be extended to include

general time-varying c_s. In this way, near-future observations of the CMB may be used to constrain general models of quintessence without introducing strong priors into its nature. A disadvantage of using the spline, however, is that it has such a large parameter degeneracy that it does not provide a clear way of sorting out which features of the equation of state are best constrained by a particular set of observations.

We have seen from our example that it is possible to robustly (much better than 3σ) distinguish models with $c_s = 1$ dark energy (*e.g.* scalar fields with canonical kinetic energy density) from models with $c_s^2 \ll 1$, such as k-essence. Our studies show that distinction depends on $c_s \ll 1$ and Ω_Q being at least a few percent at the last scattering surface, so that the scalar field fluctuations are significant enough that they affect the acoustic peaks, which can be precisely measured, as well as the large-angular scale anisotropy. If Ω_Q is too small at the last scattering surface, then the speed of sound only affects the large angular scale anisotropy and is difficult to distinguish because of the large cosmic variance at those scales. Our results are summarized in 1.

A positive detection of this sound speed effect would be encouraging, as it would provide definitive evidence that dark energy has a dynamical origin. Bean and Doré have looked for this effect, and have found the very tentative 1σ result that WMAP data suggests $c_s^2 < 0.04$ at the surface of last scattering. Hannestad [93] extended this analysis to large scale structure and supernovae and did not find any constraint on c_s^2. However, with improved data – in particular, better measurements of the high-l power spectrum – it is likely that more stringent constraints can be placed on c_s^2. Given the known degeneracies in the cosmological parameters, it is important to continue look for effects that make a clear distinction between dynamical dark energy and a cosmological constant.

Model	Ω_b	Ω_{CDM}	Ω_Q	h_0	n_s	w	c_s	$-\log L$
Fiducial k-essence	0.050	0.300	0.650	50.0	1.00	Fig. 3 (k-essence)	Fig. 4	
Imposed $c_s^2 = 1$	0.050	0.300	0.650	50.0	1.00	Fig. 3 (k-essence)	1	127.7
Best-fit $c_s^2 = 1$	0.056	0.412	0.532	45.8	1.00	Fig. 3 (k-essence)	1	24.2
Spline w and $c_s^2 = 1$	0.052	0.336	0.612	47.6	1.02	Fig. 3 (spline)	1	28.1
Constant w and $c_s^2 = 1$	0.050	0.351	0.599	48.4	1.00	-0.778	1	32.4
ΛCDM	0.041	0.290	0.668	53.2	1.00	-1	N/A	32.7

TABLE 1. The fiducial k-essence model (first line) is compared against several $c_s^2 = 1$ models. Simply setting $c_s^2 = 1$ (second line), holding the other parameters fixed, gives a much worse fit than using minimization methods to get a best fit (third line). The best $c_s^2 = 1$ model has the same equation of state as the fiducial model, but different parameters. The others, fit with a spline (fourth line), a constant equation of state (fifth line) and a cosmological constant (last line) are somewhat worse.

CHAPTER 4

Dark energy and the principle of equivalence

Although most attempts to distinguish whether dark energy is a cosmological constant or quintessence rely on measuring the equation of state, we saw in the last chapter that there is at least one other characteristic – the speed of sound – that can also be used as a test. However, dark energy was still modeled as a scalar field with no coupling to visible matter and minimal gravitational coupling. There is little reason to think this should be the case. In the compactified models considered in section 3 of chapter 2, the radion had couplings to matter and gravity. More generally, quantum corrections to scalar fields can generate couplings throughout the matter sector.

String theory is expected to contain a large number of scalar fields – moduli and the dilaton – which are light. It is usually argued that these fields are stabilized by non-perturbative effects, such as symmetry breaking. Since there are strong experimental limits on long-range scalar interactions in the universe, this is probably the case for most of these fields. There are sound theoretical reasons, however, for supposing that scalar fields might play a crucial role in cosmology, by determining the initial conditions for the hot big bang in the inflationary [121, 124], ekpyrotic [112, 113] and cyclic [170–172] scenarios, and as a dynamical solution of the cosmic coincidence problem [10, 173, 198]. Damour [53, 54] and Brax *et al.* [34, 35] have also pointed out that studying equivalence principle violating scalar fields is well motivated by brane worlds and other string models.

Our approach is to assume that dark energy is driven by a light quintessence field with interactions that violate the equivalence principle. This may be a more realistic model of dynamical dark energy than the idealized, minimally coupled model. We explore the compactifications discussed in section 3 of chapter 2 and describe the extent to which they agree

with the equivalence principle. We consider several models including, briefly, the minimally coupled scalar field [157], Brans-Dicke theory [28], a more general coupled scalar field, the exponentially warped Randall-Sundrum geometry [155, 156], four-dimensional heterotic M-theory [126], the cosmic chameleon model [116, 117] and the runaway dilaton model [85, 185]. These will be described later in the chapter.

The early experimental foundation of general relativity rested on two measurements: the gravitational deflection of light by the sun and the precession of the perihelion of Mercury. In 1959 a program of precision experimental tests of relativity was initiated, largely at the impetus of Dicke [66]. These can be thought of in the framework of the equivalence principle [193]. This included precision tests of the universality of free fall [26, 160, 174], bounds on the variation of the fundamental constants [3, 56, 166] and metric tests of gravity [19, 165, 194]. These experiments are a powerful incentive to understand the connection between dark energy and the equivalence principle. By contrast, the equation of state of dark energy was first detected in 1998 [150, 158] and, as was mentioned in chapter 3, there may be limits on how well it can be measured. We argue that experimental tests of gravity are a better tool for constraining many models of dynamical dark energy than the equation of state.

The equivalence principle can be formulated in different ways. In section 1, we discuss the different forms of the principle and the present state of experimental tests. In section 2, we discuss universal models, such as Brans-Dicke theory and the S^1/Z_2 compactification, which modify gravity in such a way that non-gravitational physics is preserved. In section 3 we discuss more general models, such as heterotic M-theory, in which the fundamental constants may vary and the universality of free-fall is violated. Section 4 discusses our results, which are summarized in table 1.

Although we do not discuss it here, Carroll [40] has proposed a model in which an approximate global symmetry (translation invariance in the scalar field) might suppress some scalar field couplings enough that they are consistent with observations. Other couplings (to the derivatives of the scalar field) would be present at first order. This could lead to an

unusual effect in which the polarization vector of radiation from cosmological sources is rotated.

1. The equivalence principles

The equivalence principle asserts that all particles, regardless of their composition or structure, fall along the same set of trajectories. That is, there exists a set of universal, inertial frames in general relativity. More generally, the principle can be taken to imply the Copernican principle that local experiments should always have an outcome that is independent of their velocity or position in space-time. This principle has been a powerful tool for testing theories of gravity as part of a program for precision tests of general relativity first laid out by Dicke in 1959 [66, 67].[1] There are three forms of the equivalence principle in frequent use.

1.1. The weak equivalence principle. The weak equivalence principle, or universality of free fall, asserts that *uncharged test particles follow identical trajectories, regardless of their mass, composition or structure.* The universality of free fall is a local statement, so that tidal forces do not play a role. This basic principle has been known since Galileo and Newton as the equivalence of inertial and gravitational mass, but in general relativity it is a consequence of the geometrical nature of the theory: it is true for all metric theories of gravity, in which matter is minimally coupled to gravity by the factor of $\sqrt{-g}$ in the action. In such theories, matter all follows the geodesics of the same metric.

The weak equivalence principle is tested by measuring the differential acceleration of two masses in a gravitational field. Departures from the universality of free fall are quantified by a parameter η, which is the fractional difference in acceleration between two masses:

$$\eta = 2\frac{|a_1 - a_2|}{|a_1 + a_2|}. \tag{4.1}$$

[1] See Will [193] for a review of experimental tests of gravity and Brans [29] for an interesting historical discussion.

It is usually thought that a violation of the weak equivalence principle would show up as the effective mass associated with different forms of energy – say leptons, QCD binding energy, or electromagnetic binding energy – falling at different rates. In practice, the weak equivalence principle is tested by Eötvös torsion-balance experiments [78] testing the differential acceleration of pairs of different metals which have different fractional contributions from the different kinds of energy. The pairs are chosen to optimize the difference in their energy compositions [53]. The most stringent limits [26, 160] are set by looking for differential acceleration towards the sun. Braginsky and Panov [26] have set a limit of $\eta < 10^{-12}$ (for a platinum and aluminum pair). The Eöt-Wash group [174] have set limits at different ranges and for different kinds of sources. In particular, they have set limits for differential acceleration towards the Earth ($\eta < 3\times 10^{-12}$) and dark matter at the galactic center $\eta^{\rm DM} < 2\times 10^{-3}$.

A useful way to think of violating the weak equivalence principle is that different sorts of matter couple to different, but conformally related metrics. If particle X couples to the metric $f(\phi)g_{\mu\nu}$, then the scalar field ϕ will generate a force between two X particles that is of order $(f'/f)^2$ times the gravitational force. A theory of gravity is *universal* if all sources of energy couple to the geodesics of the same metric, and which therefore satisfy the weak equivalence principle. Apart from general relativity, another universal theory of gravity is Brans-Dicke [28] theory,

$$S_{\rm BD} = \int d^4x\sqrt{-g}\big(\phi R - \omega\phi^{-1}(\partial\phi)^2 + \mathcal{L}_{\rm NG}[\Psi_i; g_{\mu\nu}]\big), \tag{4.2}$$

where $\mathcal{L}_{\rm matter}$ is the non-gravitational (matter) Lagrangian and the Ψ_i are the matter fields. This theory is manifestly universal, because, as in general relativity, the matter sector is minimally coupled to the metric through the factor of $\sqrt{-g}$, and does not couple to the other gravitational variable ϕ. After a change of variables and conformal transformation (2.85), the Brans-Dicke action can be written

$$S_{\rm BD} = \int d^4x\sqrt{-h}\big(\tfrac{1}{2}R_h - \tfrac{1}{2}(\partial\psi)^2 + e^{4\beta\psi}\mathcal{L}_{\rm NG}[\Psi_i; e^{2\beta\psi}h_{\mu\nu}]\big), \tag{4.3}$$

where $\beta = (6 + 4\omega)^{-1/2}$. The theory is also universal. The powers of the metric h and $e^{\beta\psi}$ match up exactly, so that all matter couples to the same metric $g_{\mu\nu} = e^{2\beta\psi} h_{\mu\nu}$ and there is no differential acceleration.

It is useful to write η as a sum of contributions from different types of matter. If we write the masses m_1 and m_2 as sums over different kinds of energy $m_{1,2} = \sum_A m_{1,2}^A$ then we may write [63, 193]

$$\eta = \Big| \sum_A \eta_A \Big(\frac{m_1^A}{m_1} - \frac{m_2^A}{m_2} \Big) \Big|, \tag{4.4}$$

where η_A quantifies the magnitude of the equivalence principle violation for a particular type of matter.[2] If the matter has Einstein-frame Lagrangian density $\mathcal{L}_A$ which couples to a (canonically normalized) scalar field ψ, then it is given by [54]

$$\eta_A = \Big(\frac{\partial \log \mathcal{L}_A}{\partial \psi} \Big)^2. \tag{4.5}$$

This is a measure of the strength of the equivalence principle violation of the field ψ. It is squared because the coupling comes from two bodies, the test mass and the source of the gravitational field.

Consider, for example, platinum and aluminum, the two metals used in the Braginsky and Panov experiment. If we consider these metals to be composed entirely of fermion (quarks and leptons) rest mass and gauge field energy, the two factors in (4.4) are very roughly 3×10^{-4} (this is the fractional difference of the contribution of lepton and quark masses to the total mass of the atoms). If these masses are independent of ψ but the gauge fields have an equivalence principle violating interaction,

$$\mathcal{L}_{\rm GUT} = -\frac{1}{16\pi\alpha_{\rm GUT}} f(\psi) F^2, \tag{4.6}$$

[2]This framework neglects other ways in which the equivalence principle could be violated, such as accelerations that depend on the aggregate size of the test mass. The composition dependent effects are the important ones for the theory under consideration.

then the measured $\eta^{\rm GUT}$ for attraction to any object made of nucleons is $(f'/f)^2$.[3] In our example, since $\eta < 10^{-12}$ and the fractional mass difference in (4.4) is of order 10^{-4} the limit on f is roughly $(f'/f) < 10^{-4.5}$.

1.2. The Einstein equivalence principle. The Einstein equivalence principle assumes that the weak equivalence principle holds, and further asserts that *the outcome of any (inertially moving) experiment should be independent of its velocity or position in space-time.* This applies only to closed, non-gravitational experiments that are sufficiently local that tidal forces may be neglected.

Schiff [161, 193] has conjectured that any complete, self-consistent theory of gravity that satisfies the weak equivalence principle also satisfies the more general Einstein equivalence principle. Nonetheless, the two principles are tested in very different ways. The weak equivalence principle is a local statement which is tested in laboratory measurements, while the Einstein equivalence principle is a global statement and is tested by gravitational redshift experiments, tests of Lorentz invariance and variation of the fundamental "constants." We consider only Lorentz invariant theories. Gravitational redshift experiments test for local variation of the fundamental constants in a gravitational potential. We are considering cases in which the scalar field ϕ is unstabilized. In these theories, the strongest observational constraints come from variation of the fundamental constants on cosmological scales.[4]

In practice, it is dimensionless numbers, such as the gauge coupling constants and the ratios of masses of fundamental constants, whose variation are constrained. The best limits come from the natural Oklo fission reactor [166], which was active two billion years ago. Remnant isotopic abundances are extremely sensitive to the fundamental constants. The strongest limits are on the variation of the fine-structure constant, so we consider it exclusively. Damour and Dyson [56] found the limit $|\Delta\alpha/\alpha| < 10^{-7}$ which is, in terms of the Hubble parameter, $|H_0^{-1}\dot{\alpha}/\alpha| < 8 \times 10^{-7}$.

[3]Since the gauge theory action is conformally invariant, there is no way that a term like (4.6) could appear in a universal theory like (4.3).
[4]For a review of tests of the constancy of the constants see Uzan [183].

Webb *et al.* [140, 188, 189] have reported a 4σ detection of a time-varying fine-structure constant, corresponding to $H_0^{-1}\dot{\alpha}/\alpha \approx 9\times 10^{-6}$, using the absorption lines of distant ($z\sim 2$) quasars observed by Keck/HIRES, but this is still controversial, with results from VLT/UVES [42, 123, 168] suggesting $|H_0^{-1}\dot{\alpha}/\alpha| < 3\times 10^{-6}$ at 3σ. It also seems difficult to reconcile the quasar variation with the Oklo results in any simple field-theory model [55, 169] as together they suggest spatial variation but no time variation.

1.3. The strong equivalence principle. The Einstein equivalence principle does not apply to the gravitation of gravitational self-energy (*e.g.* self-gravitating bodies, like stars and black holes). This is considered part of the strong equivalence principle, which the Brans-Dicke action (4.3) violates.

The strong equivalence principle generalizes the above principles to include gravitational phenomena. It assumes the weak equivalence principle holds, even for self-gravitating bodies, such as stars, black holes and Cavendish experiments. That is, *any uncharged mass, placed at the same initial event in space-time with the same initial velocity, will follow an identical trajectory.* The bodies must be small enough that tidal forces may be neglected. Further, it asserts that *the outcome of any local experiment, gravitational or non-gravitational, is independent of its velocity and its position in space time.*

The strong equivalence principle is the most restrictive principle, and includes the Einstein equivalence principle (which includes the weak equivalence principle). General relativity is the only theory known to be in complete agreement with this form of the principle. Weak-field, nonrelativistic deviations from the principle are generally described in the parameterized post-Newtonian framework [193]. However, we consider only theories that reduce, in the gravitational sector, to Brans-Dicke theory (4.3) with a potential. Thus, we consider principally limits on the cosmological evolution of Newton's constant G and on the Brans-Dicke parameter ω. We consider only positive ω. In the limit $\omega\to+\infty$, Brans-Dicke theory approaches general relativity.[5]

[5]The general relation between the Brans-Dicke parameter ω and the parameterized post-Newtonian parameter γ is $\gamma = \frac{1+\omega}{2+\omega}$, or $\omega = (1-\gamma)^{-1}$ for γ near the general relativity limit $\gamma = 1$.

The strong equivalence principle is amenable to both local and cosmological tests. The cosmological constraints come from observational limits on the time-variation of Newton's constant G (or, equivalently, coherent variation of all the masses) and are obtained from planetary science, stellar physics, pulsar timing, laser rangefinding and nucleosynthesis [183]. A tight, relatively model independent limit constrains G to have evolved by no more than 40% since nucleosynthesis (or $H_0^{-1}|\dot{G}/G| < 10^{-2}$) [3]. This is obtained by studying how the expansion rate of the universe at nucleosynthesis affects the primordial helium-4 abundance.

Many local tests of general relativity constrain the Brans-Dicke parameter. The classic test is the Nordtvedt effect [143], a perturbation on the orbits of astrophysical bodies that comes from their gravitational self-energies falling at different rates. This is measured with precision lunar rangefinding, and is a sort of lunar Eötvös experiment, measuring the polarization of the Moon's orbit due to differential acceleration of the Moon and the Earth towards the Sun. The most recent limit is $\omega > 1,100$ [194]. Better limits are placed by very long baseline interferometry measurements of the deflection by the sun of signals from distant radio sources [165], which give $\omega \geq 2,500$, and by the Cassini time-delay experiments [19] which gives the strongest limit $\omega > 40,000$.

2. Universal models

In this section, we discuss the restricted class of models that satisfy the weak and Einstein equivalence principles: they are metric theories of gravity. Damour [53, 54] has pointed out that in the context of string theory and supergravity, it is unnatural to consider only universal couplings. Nonetheless, as we have seen in section 3, the couplings of simple brane models are universal, at least to leading order.

Since our motivation is to study dark energy, we study models with the scaling potential (2.24) of section 1

$$V(\phi) = V_0 e^{b\phi}. \tag{4.7}$$

We choose this potential because it gives a constant equation of state in the limit of a scalar field dominated universe. Here, however, we must consider the effect of couplings in the non-gravitational sector.

2.1. Brans-Dicke. In this section, we investigate the relationship between the compactified, universal theories we are considering and Brans-Dicke theory in more detail. Let $\mathcal{L}_{\rm NG}$ be the non-gravitational Lagrangian and the Ψ_i the non-gravitational fields. The two actions,

$$S_1 = \int d^4x\sqrt{-g}\{\tfrac{1}{2}R[g] - \tfrac{1}{2}(\partial\psi)^2 - V(\psi) + e^{4\beta\psi}\mathcal{L}_{\rm NG}[\Psi_i; e^{2\beta\psi}g_{\mu\nu}]\}, \tag{4.8}$$

in Einstein frame and

$$S_2 = \int d^4x\sqrt{-h}\{\phi R - \omega\phi^{-1}(\partial\phi)^2 + 2\phi\lambda(\phi) + \mathcal{L}_{\rm NG}[\Psi_i; h_{\mu\nu}]\}, \tag{4.9}$$

in Brans-Dicke frame, are equivalent up to a conformal transformation and change of variable $\phi = e^{-2\beta\psi}$.[6] The constant

$$\omega = \frac{1-6\beta^2}{4\beta^2}, \tag{4.10}$$

and

$$2\lambda(\phi) = -\psi V(\psi). \tag{4.11}$$

The effective gravitational constant in this theory is [193]

$$G = \phi^{-1}\Big(\frac{4+2\omega}{3+2\omega}\Big), \tag{4.12}$$

so for small ω the scalar field acts as an additional degree of freedom in the inverse square-law, and increases the gravitational force. In order for this theory to be compatible with fifth-force constraints (tests of the inverse-square law) and tests of the strong equivalence

[6]This theory, which is Brans-Dicke theory with a potential, was first studied by Bergmann [18] and Wagoner [186]. They considered the more general case where ω is a function of ϕ.

principle, the mass V'' must be either large, $m \gg (1\,\mathrm{mm})^{-1}$ [97], or the Brans-Dicke parameter must be large. Since we are interested in a rolling scalar field, the theory must satisfy the observational constraints on ω.

A scaling potential (4.7) gives a power-law $\lambda = -V_0\phi^n$, where

$$n = 1 - b/2c. \tag{4.13}$$

The equations of motion are simplest in Brans-Dicke frame [193]:

$$3H^2 = \phi^{-1}\left(\rho + \tfrac{1}{2}\omega\phi^{-1}\dot\phi^2 - \phi\lambda - 3H\dot\phi\right), \tag{4.14}$$

$$\ddot\phi + 3H\dot\phi = \frac{1}{3+2\omega}\left(T + 2(n-1)V_0\phi^{n+1}\right) = \frac{2\beta}{1+3\beta^2}\left(\beta T + bV_0\phi^{2-b/2\beta}\right), \tag{4.15}$$

where $\rho = T_{00}$ is the density associated with $\mathcal{L}_{\mathrm{NG}}$ and T is the trace of corresponding stress-energy tensor (in Brans-Dicke frame). Since the effective gravitational constant (4.12) goes as ϕ^{-1}, these equations can be integrated to give the time-evolution of the gravitational constant. If $b < 0$, the potential and the Brans-Dicke couplings conspire to enhance the rate of change of the gravitational constant. The cosmological evolution of the ϕ field is constrained by the nucleosynthesis limits on ΔG. However, these limits are much weaker (roughly $|b| < 2$ and $\beta < 1/10$) than the the equation of state constraints $|b| \lesssim 1$ (*i.e.* $w < -0.7$ from supernovae [159, 175]) and the Cassini [19] limit $\beta < 1/400$, from (4.10). Together, these observations suggest that ΔG could have been no more than 2×10^{-3} since nucleosynthesis.

There is a relation between the Brans-Dicke parameter ω for dark energy, the equation of state of dark energy w and the cosmological rate of change of the gravitational constant $H_0^{-1}\dot G/G$. Since ω is known to be large, we treat the field ψ in (4.9) as a minimally coupled scalar field plus small corrections in β. Then, using the Friedmann equation (2.6), the equation of state of a scalar field (2.19) can be rewritten

$$w_\psi + 1 = \frac{1}{3}\Omega_{\mathrm{Q}}^{-1}(H_0^{-1}\dot\psi)^2, \tag{4.16}$$

where Ω_Q is the fractional energy density in the ψ-field relative to the critical density. This generalizes (2.35) to a multi-fluid universe. However, since Newton's constant $G \propto \phi^{-1} = e^{2\beta\psi}$,

$$\frac{\dot{G}}{G} = 2\beta\dot{\psi} = \omega^{-1/2}\dot{\psi}, \tag{4.17}$$

so

$$w_\psi + 1 = \frac{\omega}{3\Omega_Q}\Big(\frac{d\log G}{d\log a}\Big)^2, \tag{4.18}$$

where a is the scale factor. Measurements of w and ω therefore give $H_0^{-1}\dot{G}/G < 4\times 10^{-3}$. Since most of the variation of the dark energy field ψ has occurred in the last e-fold, any cosmological variation of the gravitational constant due to a dark energy fluid would be at the 4×10^{-3} level or less. This is two orders of magnitude more stringent than the present observational limit [3]. Thus, precision solar-system tests of general relativity are much better for constraining this theory.

2.2. Randall-Sundrum. The S^1/Z_2 orbifold (2.94) has $\beta = 1/\sqrt{6}$ (so $\omega = 0$) and manifestly violates the strong equivalence principle. The warped, non-factorizable compactifications considered by Randall and Sundrum [156] behave very differently. Consider their four-dimensional effective action [51], which we rewrite in terms of covariant quantities by integrating (2.98) by parts

$$S_{\rm RS} = \frac{1}{k}\int d^4x\sqrt{-g}\Big((1-\Omega^2)R - \frac{3}{2k}(\partial\Omega)^2 - 2V(r)\Big) + \int d^4x\sqrt{-g}\,\mathcal{L}_+[g_{\mu\nu}] + \int d^4x\sqrt{-g}\,\Omega^4\mathcal{L}_-[\Omega^2 g_{\mu\nu}]. \tag{4.19}$$

This is not a simple Brans-Dicke theory. Nonetheless, it is possible to derive an effective coupling. A change of variables gives the Brans-Dicke parameter ω_+ on the positive tension brane

$$\omega_+(\Omega) = \frac{3(1-\Omega^2)}{8\Omega^2 k}. \tag{4.20}$$

Since $\Omega^2 \approx 10^{-15}$ in the scenario, the result is $\omega_+ \approx 10^{14}$ so that gravity on the brane is well approximated by general relativity. This is true for the same reason as in the second Randall-Sundrum model [155], in which five dimensional gravity reduces to four dimensional gravity on a brane in a warped anti de Sitter background. For the negative tension brane, the coupling to the metric is much more important than the coupling to the scalar curvature, which is heavily suppressed. The effective β is given by

$$\beta_- = (3/2k)^{1/2} \frac{\partial \log \Omega}{\partial \Omega} = \left(\frac{3}{2k\Omega^2}\right)^{1/2}, \tag{4.21}$$

so $\omega_- \approx -3/2$ for any $\Omega \ll 1$. Gravity on the negative tension brane violates the strong equivalence principle, and additional effects (such as radion stabilization) are necessary to restore Einstein gravity on the brane.

3. General models

In this section we consider more general models which contain non-universal couplings of the form

$$\frac{1}{16\pi\alpha} e^{-\lambda\psi} F^2, \tag{4.22}$$

where F is a Yang-Mills field strength. The gauge couplings vary as $\alpha \sim e^{\lambda\psi}$, or

$$\frac{\dot\alpha}{\alpha} = \lambda\dot\psi, \qquad \frac{\Delta\alpha}{\alpha} \approx \lambda\Delta\psi, \tag{4.23}$$

where Δ represents the change since some epoch, such as big bang nucleosynthesis. Since we are considering quintessence, which has undergone most of its variation in the past Hubble time, we have $H_0^{-1}\dot\psi \approx \Delta\psi$. The interaction (4.22) violates the Einstein equivalence principle – the coupling α varies – and, as we saw in section 1.1, the universality of free fall. Thus, just as we saw for Brans-Dicke theory, they are amenable to both local and cosmological tests: limits on the variation of the fine-structure constant and Eövös-experiment limits on the universality of free fall. The parameter $\eta = C_1\lambda^2$, where C_1 is the fractional mass difference in (4.4). The value for C_1 depends on the particular test masses and the model of how

the other coupling constants vary. Values can be computed from the semi-empirical mass formula and are typically are of order 10^{-2}–10^{-4} [53, 63, 193]. Thus, in order for (4.22) to be consistent with observations, $\lambda \lesssim 10^{-4}$ is required. This is much less than the natural value for the gravitational coupling, $\lambda \sim 1$.

It is possible to write down the analog of (4.18) for these interactions, which relates the equation of state w to the cosmological variation of the fine-structure constant $\alpha_{\rm EM}$ and the universality of free fall parameter η (4.1). From (4.23) and (4.16),

$$w_\psi + 1 = \frac{1}{3}\frac{C}{\Omega_{\rm Q}\eta}\Big(\frac{d\log\alpha_{\rm EM}}{d\log a}\Big)^2. \tag{4.24}$$

If $w_\psi + 1 \approx 0.1$ and we substitute the Oklo limit in 4.24, gives $\eta/C \sim 10^{-13}$, much less than the current Braginsky and Panov limit $\eta < 10^{-12}$. The strongest limit on these variations now come from Oklo. However, forthcoming satellite tests of the universality of free fall [141, 179, 197] promises to reduce this limit to $\eta < 10^{-18}$. Thus, dramatically improved limits on such a model can be obtained by testing the equivalence principle.

Of course the coupling (4.22) also contains violations of the strong equivalence principle. This generally imposes a weaker constraint. If the couplings are not universal, not all the mass of a typical object is coupled to ψ. The Brans-Dicke parameter implied by (4.22) is then

$$\omega = (4C_2^2\lambda^2)^{-1} \gg 10^8 \tag{4.25}$$

from (4.10), where $C_2 < 1$ is the fraction of a typical test mass that couples to ψ.

3.1. Hořava-Witten. The four-dimensional low-energy effective action of heterotic M theory, derived by Lukas *et al* [126], couples to the radion c nearly universally, as it derives from a simple S^1/Z_2 compactification. The linear warping of the background corrects this: for example, the vector bosons on the brane couple to the radion, at the second order in the eleven-dimensional gravitational constant. The volume modulus a in the effective action for heterotic M theory (2.105) must be stabilized in any realistic model: variation in the

Calabi-Yau volume affects the gauge couplings at the first order in the eleven dimensional gravitational coupling $\kappa^{2/3}$. However, the radion is universal at this order.

Thus, setting the volume modulus $a = 1$ conformally rescaling the metric $g_{\mu\nu} = e^{\hat{c}} h_{\mu\nu}$ in the Lukas-Ovrut-Waldram action (2.105) yields

$$S = \frac{\pi\rho V}{\kappa^2}\int_{\mathcal{M}^4}\sqrt{-h}d^4x\Big[e^c R - 0\times e^c(\partial c)^2 - 3\big(1+\tfrac{1}{3}\xi\alpha_0 e^c\big)\partial_\mu C\partial^\mu\bar{C}$$
$$-\tfrac{3}{8}e^{-c}C\bar{C}\bar{C} - \frac{3k^2}{4}\big(1-\tfrac{1}{3}\xi\alpha_0 e^c\big)C\bar{C}\bar{C}$$
$$-\frac{V}{8\pi\kappa^2}\Big(\frac{\kappa}{4\pi}\Big)^{2/3}\int_{\mathcal{M}^4}\sqrt{-g}\,d^4x\big((1+\xi\alpha_0 e^c)\,\mathrm{tr}(F^{(1)})^2 + (1-\xi\alpha_0 e^c)\,\mathrm{tr}(F^{(2)})^2\big), \quad (4.26)$$

V is the Calabi-Yau volume, $L = \pi\rho e^c$ is the S^1/Z_2 length (ρ is a constant) and κ^2 is the eleven-dimensional gravitational coupling. The $F^{(i)}$ are the Yang-Mills field strengths on the two branes and the C are scalar fields. We have written terms quadratic in the C schematically. Now, the variable C^2 is of order $\kappa^{2/3}$ as are the parameters α_0 and the parameter k^{-2} (see section 3.4), so to first order in $\kappa^{2/3}$, the theory is universal. The Brans-Dicke parameter, however, is $\omega = 0$. So in this case, as in the case of the negative tension brane in the Randall-Sundrum scenario, gravity on the brane is far from general relativity.

An intriguing attribute of (4.26) are the so-called threshold corrections to the gauge couplings. These corrections come from the linear warping of the Calabi-Yau manifold in the background solution. The corrections are proportional to the length of the orbifold interval $L = 2^{7/2}\xi e^c$ times the Calabi-Yau distortion α_0. These corrections violate the universality of free fall and would cause cosmological variation of the GUT couplings. The value of $\xi\alpha_0$ is

$$\xi\alpha_0 \sim \frac{L}{2^4\pi V^{2/3}}\Big(\frac{\kappa}{4\pi}\Big)^{2/3}\theta = 4\times 10^{-3}\frac{L\kappa^{2/3}}{V^{2/3}}\theta = 4\times 10^{-3}\theta\varepsilon, \quad (4.27)$$

where $\varepsilon = L\kappa^{2/3}/\pi V^{2/3}$ is the expansion parameter of the low-energy effective action and we have rewritten the integral in (2.106) as $V^{1/3}\theta$, where the dimensionless variable θ is

$$\theta = -V^{-1/3}\int_X \omega\wedge\mathrm{tr}\,R^{(X)}\wedge\mathrm{tr}\,R^{(X)}, \quad (4.28)$$

an integral over the Calabi-Yau manifold X (the notation is explained below (2.106)). The magnitude of the differential acceleration (4.1) predicted in the Braginsky and Panov experiment [26], using the Einstein-frame action (2.105),

$$\eta = \tfrac{1}{3}(\xi\alpha_0)^2 \times 3 \times 10^{-4} = 10^{-9}\theta^2\varepsilon^2. \tag{4.29}$$

The factor of $1/3$ comes from the normalization of c and the factor of 3×10^{-4} is the difference, as in (4.4), for aluminum and platinum, between the contribution of the fermion masses and gauge field binding energy (*e.g.* electrostatic and QCD). The calculation of η (4.29) assumes that the masses of fermions scale with the radion in a similar (same order of magnitude, or zero). Since the scalar fields C do, it seems reasonable to assume that their super-partners will have similar couplings [192]. The parameter ε is expected to be of order unity [13, 127, 196] as is the integral θ. Thus, for moderately small values of $\theta \sim 10^{-1}$, the theory predicts equivalence principle violations quite near present observations, with a significant constraint already placed by the Bragninsky and Panov [26] limit $\eta < 10^{-12}$.

Limits on cosmological variation of the gauge coupling constants constrain the cosmological evolution of the radion. The Oklo limit $\dot{\alpha}/\alpha < 10^{-7}$ at redshift 2 [56] implies $H_0^{-1}\dot{\psi} < 10^{-3}$. In the theory without a potential, c couples to the background radiation density. Because of the suppression of the coupling by $\xi\alpha_0$ and the low radiation density today, this only suggests that the fine-structure constant will have varied by roughly 10^{-11}, which is four orders of magnitude smaller than the Oklo constraint and is inconsistent with the variation claimed by Webb *et al.* [140, 188, 189].

3.2. Runaway dilaton. The runaway dilaton scenario of [85, 185] is a string-inspired model that naturally satisfies the equivalence principle constraints discussed above. In this model, the string theory dilaton Φ is moving towards infinity. This is the strong coupling limit: the non-perturbative potential and couplings associated with Φ are generally a power series in Φ. However, in [185] Veneziano argued that the couplings and potential may, in fact, approach finite limits as $\Phi \to \infty$. This could be a natural explanation for our failure

to observe equivalence principle violations associated with the dilaton. The cosmological consequences are discussed in Damour *et al.* [61, 62], where it was shown that deviations from the universality of free fall near the present experimental threshold are possible. It predicts that general relativity should hold to high accuracy, as η and ω are related by

$$|\eta| \approx 2.6 \times 10^{-5} \omega^{-1}. \tag{4.30}$$

For a "dark energy" dilaton with a potential of the form (4.7), the runaway dilaton model predicts:

$$\left|\frac{d\log\alpha}{d\log a}\right| \approx 6.2 \times 10^{-7} \sqrt{10^{12}\eta}. \tag{4.31}$$

These constraints suggest that ω should be at least two orders of magnitude larger than current bounds, but η and $\dot{\alpha}$ may be well within the reach of future experiments.

(A related scenario is the least coupling scenario of Damour and Polyakov [63], in which the dilaton is attracted to a finite value.)

3.3. Chameleon model. Khoury and Weltman have proposed a chameleon model for dark energy [116, 117] in which equivalence principle constraints can be avoided for terrestrial tests of the universality of free fall. It contains a "thin-shell" effect: while the universe contains a light scalar field, in dense regions its couplings drive it to a fixed value and give it a large effective mass. In the universe today, this model requires $w = -1$ to high accuracy and that variations in the fine-structure constant should be exponentially suppressed [33]. However, because the thin-shell effect applies only in high density regions, future space-based tests of the equivalence principle [4, 141, 179, 197] could detect large – order unity – deviations from the universality of free fall and measure a vastly different gravitational constant.

4. Discussion

If the potential has a scaling form, as (4.7), then the equation of state is fixed by the constant equation-of-state attractor solution. Another possibility is that the equation of state is

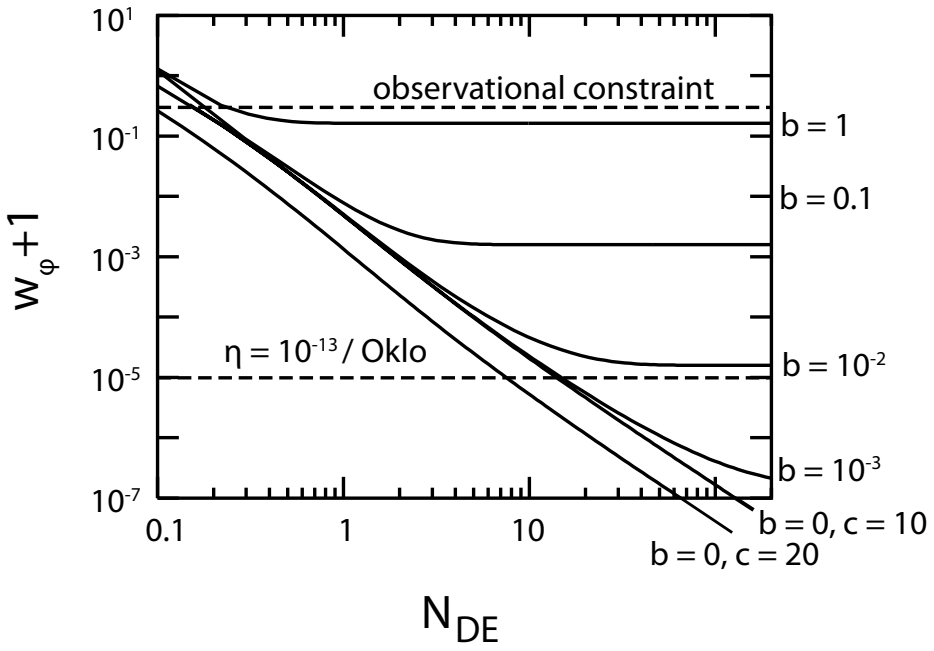

FIGURE 1. $w+1$ as a function of the number of e-folds of expansion remaining for the potential (4.33). The observational constraint $w \lesssim 0.7$ is indicated by a dashed line. It is satisfied by all the models with $\gtrsim 0.5$ e-fold remaining. The other dashed line is the upper bound on $w_\phi + 1$ from (4.24) if a violation of the universality of free fall is found at the $\eta = 10^{-13}$ level. This could be the case in the M-theory model (4.29).

changing. We mentioned in chapter 3 that quintessence trackers have a decreasing equation of state in the late universe and *k*-essence attractors have an increasing equation of state in the late universe. It is difficult to construct models in which which w crosses the $w = -1$ barrier [37, 39]. Given that w is presently measured to be near -1, future observations are probably more sensitive to increasing w, than to w approaching -1 from above. One possibility is that the dark energy potential will eventually become negative and the universe will begin contracting. This is discussed in chapters 5 and 6.

In [108, 187], the possibility of using the equation of state to constrain the end of dark energy domination and the reversal to contraction was discussed. They used a linear potential

$$V(\phi) = V_0(1 + \alpha\phi), \tag{4.32}$$

where α and V_0 are positive constants and found that the 2σ constraint that dark energy domination end more than two e-folds (24 billion years) from today. We consider the slightly different potential

$$V(\phi) = V_0(e^{b\phi} - e^{-c\phi}), \tag{4.33}$$

where V_0, b and c are positive constants. This potential behaves like the constant equation of state scaling potential (4.7) at large values of ϕ, but has a cutoff at $\phi = 0$ whose severity is determined by the constant c. The potential (4.33), while slightly more complicated than (4.32) is relevant to the cyclic model discussed in chapter 6. We define the number of e-folds $N_{\rm DE}$ of dark energy domination remaining as the number of e-folds of expansion before the Hubble parameter vanishes and the universe begins contracting. In figure 1 we plot $N_{\rm DE}$ against the equation of state measured today for a variety of parameters b and c (these are implicit functions of the value of ϕ today in the attractor solution of (4.33). The equation of state constraints [175] impose the limit $N_{DE} \gtrsim 0.5$, which is not particularly strong. These are weaker than the limits on (4.32) because of the sharp cutoff.

If the strength of equivalence principle violations was known either from a theoretical model (as in the heterotic M-theory case of section 3.1) or from a measured violation of the universality of free fall then a much stronger constraint could be placed on w. For example, if a measurement showed that the universality of free fall was violated at *e.g.* the $\eta = 10^{-13}$ level, then the Oklo constraint on the time-variation of α and (4.24) imply that $w+1 \lesssim 10^{-5}$. Figure 1 shows that this places a strong constraint on the potential (4.33) and the number of e-folds of dark energy domination $b \lesssim 10^{-2}$ and $N_{\rm DE} \gtrsim 20$. A similar idea holds for the relation between w, $\dot{G}$ and ω (4.18), but since the constraints on $\dot{G}$ are comparatively weak, this could not realistically place a constraint on w.

Our results are summarized in table 1, which compares the models of dark energy we have discussed against observational constraints, in the top row. Each model can be compared with observations, to see what the best observational strategy for constraining the model is.

	$w+1$	$\lvert\Delta G_{\rm BBN}/G\rvert$	$H_0^{-1}\lvert\dot\alpha_{\rm EM}\rvert$	$\lvert\eta\rvert$	$\lvert\omega\rvert$	tightest constraints
Observations	$\lesssim 0.3$	$\lesssim 0.4$	$\lesssim 10^{-7}$	$\lesssim 10^{-12}$	$\gtrsim 40,000$	
Scalar fields						
Minimally interacting	$\lesssim 0.3$	—	—	—	—	w
Brans-Dicke (§2.1)	$\lesssim 0.3$	$\lesssim 2\times 10^{-3}$	—	—	$\gtrsim 40,000$	w, ω
General (§3)	$\lesssim 0.3$	$\lesssim 10^{-4}$	$\lesssim 10^{-7}$	$\lesssim 10^{-12}$	$\gtrsim 10^{8}$	w, η, $\dot\alpha$
Compactifications						
S^1/Z_2 (Chapter 2 §3.2)	$\lesssim 0.3$	large	—	—	0	ruled out by ω
Randall-Sundrum[a] (§2.2),	$\lesssim 0.3$	large	—	—	$-3/2$	ruled out by ω
positive tension brane	$\lesssim 0.3$	—	—	—	—	dark matter EP tests?
String inspired						
Heterotic M-theory (§3.1)	$\lesssim 10^{-6}$	large	$\lesssim 10^{-7}$	$\approx 10^{-12}$	0	η, $\dot\alpha$ (ruled out by ω)
Runaway dilaton (§3.2)	$\lesssim 0.3$	$\lesssim 10^{-4}$	$\lesssim 10^{-7}$	$\lesssim 10^{-13}$	$\gtrsim 10^{8}$	w, η, $\dot\alpha$
Cosmic chameleon (§3.3)	—	—	—	$\lesssim 10^{-12}$ [b]	$\gtrsim 10^{12}$	η (space-based)

TABLE 1. Constraints on light scalar fields in various models of dark energy. The constraints come from the equation of state w and various equivalence principle parameters described in the text. The first row indicates the constraints imposed by observations. The models, listed below, are described in the text. Columns for which the effect vanishes, or is many orders of magnitude beyond the present state of observations, are marked with a dash, —. The last column indicates the best observational opportunities for constraining a given model.

[a] Negative tension brane.

[b] Equivalence principle constraint for ground based experiments. Future space based experiments could give large violations of the equivalence principle in this scenario.

Each models must have a potential which is sufficiently flat to be consistent with equation of state constraints. Other parameters of must be small enough to ensure that they satisfy observational constraints. This is natural for some of these models, but in the case of Brans-Dicke theory and the scalar field with general interactions, the couplings must be tuned to very small values for these simple models to be consistent with observations. The table is discussed line-by-line below.

Minimally interacting scalar: The ideal minimally interacting scalar field [38, 149, 157, 198], such as that discussed in section 1 of chapter 2 is a simple toy model of quintessence. It is defined by one parameter, the equation of state. In chapter 3, we saw that an additional parameter which characterizes its perturbations, the speed of sound, could be measured.

Brans-Dicke theory: The Cassini constraint [19] requires that the Brans-Dicke parameter $\omega > 40,000$ [28].

Scalar field with general couplings: The model with general couplings, discussed in section 3, generally violates tests of post-Newtonian gravity and the universality of free fall if the couplings have their natural, gravitational strength (*i.e.* are of order one in Planck units). Since these are local tests, this is true even with a very flat potential. In special cases, such as the chameleon model, these constraints be averted, but it is generally necessary to have the coupling in (4.22) satisfy $\lambda \lesssim 10^{-4}$. Although this is not well motivated, if we nonetheless assume this is the case, we obtain the figures in the table.

S^1/Z_2 orbifold: The S^1/Z_2 orbifold of section 3.2 of chapter 2 violates tests of the strong equivalence principle.

Randall-Sundrum: Even though they interact with the same metric, the conformal coupling to the radion causes matter on each of the two branes in the Randall-Sundrum model [156] to have very different gravitational behavior. Like the S^1/Z_2 orbifold, the negative tension brane strongly violates the equivalence principle, and

must be stabilized if the model is to succeed. The positive tension brane, however, has a large value of ω and gravity behaves very much like general relativity on this brane [155].

Cosmic chameleon: The best tests of the cosmic chameleon model [33, 116, 117] are tests of the equivalence principle in space, which could measure unexpectedly large deviations. The chameleon field has a quite flat potential, so effects due to w and the time-variation of the fundamental constants are small.

Runaway dilaton: The predictions of the runaway dilaton model [61, 62, 85, 185] are the same as the scalar field with general suppressed couplings. The model simply provides a natural mechanism by which a light scalar field might have small, but still observable, couplings to matter.

Heterotic M-theory: Although the heterotic M-theory scenario we have discussed violates the strong equivalence principle to a degree that is inconsistent with present bounds, it is conceivable that other compactifications would look like general relativity at this order. If, as is likely, the warped background geometry persists, then the predicted effect on the universality of free fall is likely to remain. In this case, the main constraints would come from time-variation of the fine structure constant, as in the table.

In sum, light scalar fields are important for cosmology and abundant in string theory. They are predicted to have equivalence principle violating interactions. Violations of the equivalence principle has been very well constrained by an array of precise experiments. However, there are natural models in which these interactions are suppressed to be consistent with observations, so there is a reason to expect that future experiments may detect such a violation. The best way to differentiate quintessence from a cosmological constant is to adopt a multi-pronged approach, in which a program of tests of the equivalence principle – metric tests of gravity, Eötvös experiments and constraints on the variation of the fine structure constant – continues along with measurements of the equation of state.

CHAPTER 5

Chaos and contracting universes

A common feature of the pre-big bang [86, 87, 184], ekpyrotic [114, 115] and cyclic [170–172] models of brane cosmology is that they match a contracting universe onto an expanding hot big bang universe. The behavior of bouncing universes is actively being investigated in string theory and cosmology[1]. In section 3.2 of chapter 2, we described the collision of two S^1/Z_2 orbifold planes, described by the action (2.94). In this solution, the Einstein-frame scale factor vanishes as the fixed planes collide, while the brane scale factors remain finite.

Understanding the behavior of the universe as it contracts is a key issue in these scenarios. In the 1970s, Belinsky, Khalatnikov and Lifshitz (BKL) [16, 17] showed that chaotic oscillations generically occur in the approach to a crunch. This has would destroy the observed homogeneity and isotropy of the universe in the cyclic and ekpyrotic models, and could have dramatic consequences in the pre-big bang scenario. In this chapter, we show how chaos can be suppressed by the S^1/Z_2 orbifold and by scalar field in the cyclic model, which has a large pressure-to-density ratio in the contracting phase.

Previous studies [16, 17, 57–60, 65, 125], have focused on models in which the universe contains matter and radiation, or, more generally, an energy component whose equation of state is $w \leq 1$ (where $w \equiv p/\rho$ is defined as the ratio of the pressure p to the energy density ρ). If $w < 1$, a contracting homogeneous and isotropic solution is unstable to small perturbations in the anisotropy and spatial curvature. As the overall volume shrinks, the anisotropy causes the universe to expand along one axis and contract along the others, a state that can be approximated by the anisotropic Kasner solution. The spatial curvature causes the axes and rates of contraction to undergo sudden jumps from one Kasner-like

[1] See [14, 49, 133, 151, 177, 178, 180].

solution to another, an effect known as "mixmaster" [136, 138] behavior. If the curvature is not spatially uniform, then the chaotic behavior in different regions is not synchronized and the universe becomes highly inhomogeneous at the big crunch. Hence, mixmaster behavior could potentially wreak havoc in cosmological models with a big crunch/big bang transition, making them inconsistent with the observed large scale homogeneity of the universe.

In this chapter, we show that the behavior of the universe as it approaches the big crunch is very different if there is an energy component with $w > 1$. The chaotic behavior is suppressed and the universe contracts homogeneously and isotropically as it approaches the singularity. The reason is that the anisotropy and curvature terms in the Einstein equations grow rapidly and become dominant if $w < 1$, but they remain negligible compared to the energy density if $w > 1$. In the latter case, the Einstein equations converge to the Friedmann equations with purely time-dependent terms, a condition sometimes referred to as "ultralocality." The effect can be viewed as a generalization of the "cosmic no-hair theorem" invoked in a rapidly inflating universe. Here we demonstrate analogous behavior in a slowly contracting universe with $w > 1$. This chapter is adapted from [80]. A related result of Dunsby *et al.* [44, 45, 73] shows that models with $0 < w < 1$ but with ρ^2 terms in the stress-energy tensor are also driven towards isotropy. This is related to our result, because $\rho^2 \sim a^{-6(1+w)}$, so the contributions from a ρ^2 term in the stress energy tensor scale in the same way as a component with equation of state $1 + 2w$. The convergence rate of the $w \geq 1$ no-hair behavior has since been studied by Coley and Lim [46].

The cosmic no hair theorem for a contracting universe containing a perfect fluid with $w \geq 1$ is discussed in section 1. A common example of a perfect fluid is a scalar field ϕ with a potential $V(\phi)$. In section 2, we consider the interaction of the scalar ϕ with a p-form field F_{p+1} through an exponential coupling,

$$e^{\lambda\phi} F_{p+1}^2, \tag{5.1}$$

where λ is a constant. We consider this case because scalar fields with exponential couplings to p-form fields are common in Kaluza-Klein, supergravity and superstring models. For the case $w = 1$, it is known [16, 60] that the contraction is not chaotic if λ lies within a bounded interval. Here we show that, for *any* λ and p, there is a critical value $w_{\rm crit}(\lambda, p)$ for which the chaotic behavior is suppressed if $w > w_{\rm crit}(\lambda, p)$.

Our results are of particular importance for the ekpyrotic and cyclic cosmological models, which have a big crunch/big bang transition with a contraction phase dominated by a scalar field with $w \geq 1$ [115]. These results are relevant because they suggests that the universe can remain homogeneous and isotropic on large scales. Once the evolution becomes ultralocal, the whole universe is following the same homogeneous and isotropic evolution all the way to the big crunch. These consequences are discussed further in chapter 6.

In section 3, we explore how time-variation of w affects our conclusions, and in particular how w approaching $w_{\rm crit}$ from above may suppress chaotic behavior. In section 4 we discuss some specific models. In particular, we show how orbifolding can remove p-forms that might induce chaotic behavior and discuss the special case of heterotic M-theory, which, to leading order in the eleven dimensional gravitational coupling κ, is on boundary between chaotic and smooth behavior.

1. A "cosmic no-hair theorem" for contracting universes

The cornerstone of the inflationary paradigm is an argument known as the "cosmic no-hair theorem", according to which a universe containing a perfect fluid component with $w < -1/3$ will rapidly approach flatness, homogeneity and isotropy at late times, for a wide range of initial data (namely those for which the space curvature, inhomogeneity and anisotropy are not very large) [121]. In the Friedmann equation, the energy density for a component with equation of state w is proportional to $1/a^x$, where the exponent $x = 3(1 + w)$. The anisotropy term is proportional to a^{-6} and the spatial curvature term is proportional to a^{-2}. As the universe expands, the contribution with the smallest values of x redshifts away more slowly than components with larger values of x and so come to dominate the Friedmann

equation and the components with the smallest value of x overall ultimately dominate. If the energy component with the smallest value of w has $w < -1/3$, then $x < 2$ and this component dominates. For a wide range of initial data, convergence to a homogeneous and isotropic expanding universe is assured.

Below, we will present an analogous "cosmic no-hair theorem" for contracting universes. In a contracting universe, the component with the largest value of x will dominate the Friedmann equation. Starting from an inhomogeneous and anisotropic initial state, we will show that the existence of a perfect fluid with $w > 1$ (or $x > 6$) will suppress chaotic behavior, and enable a smooth and isotropic contraction to the big crunch. We will find that curvature plays a more complicated role compared to the case of expansion. Hence, we first obtain a cosmic no-hair theorem for the case of zero spatial curvature and then generalize to the case of arbitrary spatial curvature. We intentionally take a pedagogical approach that encompasses known results for $w \leq 1$ to make our discussion self-contained. Our analysis assumes the initial inhomogeneity is small; it is possible that the universe evolves towards other attractors for sufficiently large deviations from homogeneity.

All of our computations are performed in synchronous gauge,

$$ds^2 = -dt^2 + h_{ab}(t, \mathbf{x})\, dx^a\, dx^b, \tag{5.2}$$

where we use our freedom to choose a spatial slicing to ensure that the big crunch occurs everywhere at $t = 0$ ($\det h_{ab} \to 0$ as $t \to 0$). For a perfect, comoving fluid with equation of state $p = w\rho$, the Einstein equations are [125]:

$$\frac{\partial}{\partial t}\kappa_j{}^j + \kappa_j{}^k \kappa_k{}^j = -\left(\frac{1+3w}{2}\right)\rho, \tag{5.3}$$

$$\frac{\partial}{\partial x^a}\kappa^j{}_j - \frac{\partial}{\partial x^j}\kappa^j{}_a = 0, \tag{5.4}$$

$$P_a{}^b + \frac{1}{\sqrt{h}}\frac{\partial}{\partial t}\left(\sqrt{h}\kappa_a{}^b\right) = +\left(\frac{1-w}{2}\right)\rho, \tag{5.5}$$

where ${P_a}^b$ is the Ricci tensor on spacelike surfaces, and κ_{ab} is defined by

$$\kappa_{ab} = \frac{1}{2}\frac{\partial}{\partial t} h_{ab}, \tag{5.6}$$

$${\kappa_a}^b = \kappa_{aj} h^{jb}. \tag{5.7}$$

Near the big crunch, the dynamics of the metric (5.2) are *ultralocal* [5, 17, 59, 60]. That is, the evolution of adjacent spatial points decouples because spatial gradients increase more slowly than other terms in the equations of motion. Therefore, analyzing the dynamics of this metric near the singularity and at fixed spatial coordinate $\mathbf{x}_0$ is equivalent to analyzing the much simpler system

$$ds^2 = -dt^2 + \sum_{ij} e^{2\beta_{ij}(t;\mathbf{x}_0)} \sigma^{(i)}(\mathbf{y};\mathbf{x_0}) \sigma^{(j)}(\mathbf{y};\mathbf{x}_0), \tag{5.8}$$

where the $\sigma^{(i)}$ are $\mathbf{y}$-dependent one-forms that are linearly independent at each point and form a *homogeneous* (but possibly curved) space such as Bianchi type IX [138]. The β_{ij}, which do not depend on $\mathbf{y}$, describe the (generally anisotropic) contraction of this space. Both the $\sigma^{(i)}$ and the β_{ij} depend on the parameter $\mathbf{x}_0$, the spatial point being studied. The dynamics of the inhomogeneous universe at a fixed spatial point can be approximated, near $t = 0$, by the dynamics of a homogeneous (but curved and anisotropic) universe. Differences in curvature and anisotropy between different $\mathbf{x}_0$ are encoded in the different $\sigma^{(i)}$ and β_{ij} associated with these points.

In each Kasner-like epoch, we may perform a rotation so that β is diagonal. Furthermore, we may separate out the trace of β and write it as the "volume scale-factor" $a(t)$, in analogy to the isotropic Friedman-Robertson-Walker universe, to obtain the metric

$$ds^2 = -dt^2 + a^2(t) \sum_i e^{2\beta_i(t)} (\sigma^{(i)})^2, \tag{5.9}$$

$$\beta_1(t) + \beta_2(t) + \beta_3(t) = 0, \tag{5.10}$$

where the dependence of $a(t)$, the β_i and the $\sigma^{(i)}$ on $\mathbf{x}_0$ has been suppressed. The combination $a\, e^{\beta_i}$ can be thought of as the effective scale factor along the i^{th} direction, and the functions β_i then describe the contraction or expansion of each direction relative to the overall volume contraction. We may use our freedom to rescale the σ to ensure that at some time t_0, $a(t_0) = 1$, $\beta_i(t_0) = 0$ and $\det(\sigma^{(1)}, \sigma^{(2)}, \sigma^{(3)}) = 1$. Quantities with a subscript zero (such as ρ_0) refer to their values at this fixed time.

The Einstein equations (5.3)–(5.5) close with the equation of energy conservation (2.5) for the fluid,

$$\frac{d\log\rho}{d\log a} = -3(1+w). \tag{5.11}$$

For constant w, this equation has the familiar solution (2.8),

$$\rho(a) = \rho_0 a^{-3(1+w)}. \tag{5.12}$$

While we could have included several perfect fluids, with different equations of state w_i, the fluid with the largest equation of state will always dominate near the crunch, so it is sufficient to consider only one energy component. We have taken this fluid to be comoving, because small perturbations of a comoving background are suppressed in a $w > 0$, contracting universe. In particular, the $T^0{}_i$ terms that would appear on the right hand side of (5.4) grow only as $t^{-2/(1+w)}$, which is slower than the t^{-2} rate at which the diagonal terms grow [2]

1.1. The curvature-free case. We first examine the case of Ricci flat spatial 3-surfaces, for which $P_a{}^b = 0$. In this case, we write $\sigma^{(i)} = dx^i$. Then, the Einstein equations (5.3)-(5.5)

[2]In fact, in a homogeneous universe, $T^0{}_i$ scales as a^{-3}. In this case, the non-comoving component of a scalar field or perfect fluid never grows to dominate in a cosmology with $w > 0$. For a universe with a scalar field and negative exponential potential, such as (5.40), it is possible to show that inhomogeneous perturbations to the scaling background may be neglected near the crunch in a contracting universe.

reduce to

$$3\left(\frac{\dot{a}}{a}\right)^2 - \frac{1}{2}(\dot{\beta}_1^2 + \dot{\beta}_2^2 + \dot{\beta}_3^2) = \rho, \tag{5.13}$$

$$\ddot{\beta}_i + 3\frac{\dot{a}}{a}\dot{\beta}_i = 0, \tag{5.14}$$

where a dot indicates a derivative with respect to the proper time t. Integration of (5.14) gives,

$$\dot{\beta}_i = c_i a^{-3}, \tag{5.15}$$

while the constraint (5.10) implies,

$$c_1 + c_2 + c_3 = 0. \tag{5.16}$$

Combining these results, equation (5.13) becomes a Friedmann equation,

$$3\left(\frac{\dot{a}}{a}\right)^2 = \rho(a) + \frac{\sigma^2}{a^6} = \frac{\rho_0}{a^{3(1+w)}} + \frac{\sigma^2}{a^6}, \tag{5.17}$$

where we define

$$\sigma^2 = \frac{1}{2}(c_1^2 + c_2^2 + c_3^2). \tag{5.18}$$

An anisotropic universe has $\dot{\beta}_i \neq 0$, *i.e.* $c_i \neq 0$. The constant σ^2 parameterizes the anisotropic contribution to the Friedmann equation in (5.17). The anisotropy evolves as $1/a^6$ or $x = 6$. We define the fractional energy densities Ω_ρ and Ω_σ as

$$\Omega_\rho = \frac{\rho(a)}{\rho(a) + \sigma^2/a^6}, \tag{5.19}$$

$$\Omega_\sigma = \frac{\sigma^2/a^6}{\rho(a) + \sigma^2/a^6}. \tag{5.20}$$

These quantities represent the contribution of the perfect fluid and anisotropy to the critical density for closure of the universe. Since we are neglecting curvature, $\Omega_\rho + \Omega_\sigma = 1$.

The solution for the β_i as a function of the scale factor a is,

$$\beta_i(a) = c_i\sqrt{3}\int_a^1 \frac{da'}{a'}\left(\rho(a')a'^6 + \sigma^2\right)^{-1/2}. \tag{5.21}$$

The limits of integration have been chosen to ensure $\beta_i(1) = 0$. For the remainder of the chapter, we will assume a universe contracting towards $a \to 0$ as t approaches zero from below. Let us now examine the behavior of these solutions for various w.

$w < 1$: When $w < 1$, the $\rho(a)$ part of the integral (5.21) is negligible as $a \to 0$, and so the solution converges to the vacuum ($\rho = 0$) Kasner universe during contraction,

$$a(t) = \left(\frac{t}{t_0}\right)^{1/3} \tag{5.22}$$

$$\beta_i(t) = \frac{c_j}{\sigma\sqrt{3}}\ln\left(\frac{t}{t_0}\right). \tag{5.23}$$

The Kasner universe is parameterized by three *Kasner exponents* p_i,

$$p_i = \frac{1}{3} + \frac{c_j}{\sigma\sqrt{3}}. \tag{5.24}$$

The scale factors in (5.9) are powers of t:

$$a e^{\beta_i} = |t/t_0|^{p_j}, \tag{5.25}$$

and the relations (5.16) and (5.18) become

$$p_1 + p_2 + p_3 = 1 \tag{5.26}$$

$$p_1^2 + p_2^2 + p_3^2 = 1, \tag{5.27}$$

known as the *Kasner conditions*. These describe the intersection of a plane, the *Kasner plane*, and a unit sphere, the *Kasner sphere*, as illustrated in Fig. 1. We will denote the intersection, which represents the allowed values of the p_i, as the *Kasner circle*. The outermost circle in Fig. 1 corresponds to the limit where $w < 1$, as the energy density scales away and only a vacuum, anisotropic universe remains.

There are three degenerate solutions where exactly one of the p_i is one, and the other exponents are zero (the solid black circles in Fig. 1). At all other points on the (dashed) Kasner circle exactly one of the p_i is negative. Thus, although the geometric mean of the three scale factors $a(t) = |t|$ is contracting, a single scale factor corresponding to the negative Kasner exponent is undergoing expansion to infinity.

For the curvature-free case, the universe becomes increasingly anisotropic near the big crunch if $w < 1$. In particular, the isotropic solution, $p_1 = p_2 = p_3 = 1/3$, is inconsistent with the Kasner conditions (5.26) and (5.27).

$w = 1$: Inspection of (5.12) reveals that, when $w = 1$, the matter density and the anisotropy terms in the Friedmann equation (5.17) scale with the same power of a, so Ω_ρ and Ω_σ remain fixed. The solutions are

$$a(t) = \left(\frac{t}{t_0}\right)^{1/3} \tag{5.28}$$

$$\beta_i(t) = \frac{c_j}{\sqrt{3(\sigma^2 + \rho_0)}} \ln\left(\frac{t}{t_0}\right). \tag{5.29}$$

This solution is very similar to the $\rho = 0$ case, and indeed we may define the Kasner exponents,

$$p_i = \frac{1}{3} + \frac{c_j}{\sigma\sqrt{3}}\left(1 + \frac{\rho_0}{\sigma^2}\right)^{-1/2}. \tag{5.30}$$

The Kasner conditions are different. If we define

$$q^2 \equiv \frac{2}{3}\frac{\rho_0}{\sigma^2 + \rho_0} = \frac{2}{3}(1 - \Omega_\sigma) \tag{5.31}$$

then the Kasner conditions are

$$p_1 + p_2 + p_3 = 1, \tag{5.32}$$

$$p_1^2 + p_2^2 + p_3^2 = 1 - q^2 = \tfrac{1}{3} + \tfrac{2}{3}\Omega_\sigma. \tag{5.33}$$

The first condition is unchanged from (5.27) but the right hand side of the second condition has been modified. Increasing Ω_σ corresponds to increasing the radius of the Kasner sphere.

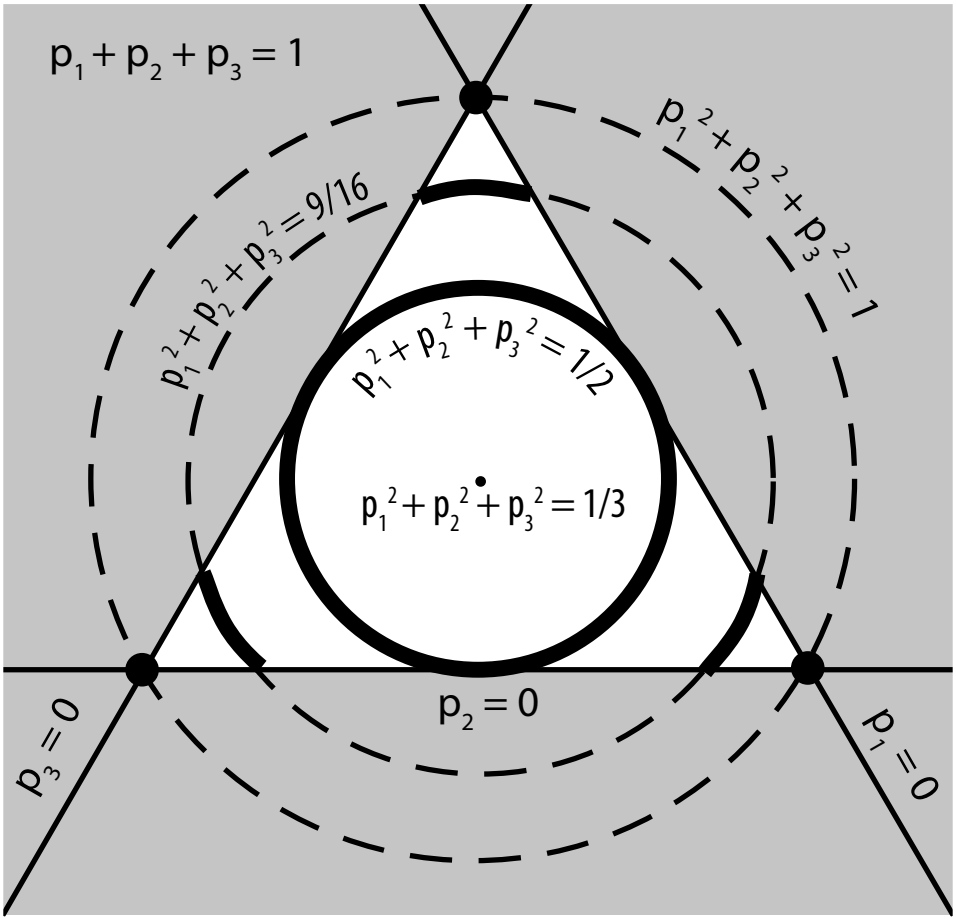

FIGURE 1. The Kasner plane $p_1+p_2+p_3 = 1$ and its intersections (the Kasner circles) with various spheres $p_1^2+p_2^2+p_3^2 = 1-q^2$ where $q^2 = \frac{2}{3}(1-\Omega_\sigma)$; see (5.31). The vacuum solution corresponds to $\Omega_\sigma = 1$ (the outermost circle). The inner circles are relevant to the case where $w = 1$ and $\Omega_\sigma < 1$. In the white regions, the Kasner exponents are all positive (corresponding to contraction); in gray regions, one exponent is negative (expanding). If the spatial curvature is non-zero, points along the circles in the white region (thick parts of circles) are stable but points in the gray regions (dashed parts of circles) are unstable, jumping to new values after a short period of contraction. If a model (*i.e.* a circle) has an open set of stable points (the three innermost circles but not the outermost circle), the contracting phase does not exhibit chaotic mixmaster behavior.

The $w = 1$ model allows us to explore the behavior of the contracting universe as a function of Ω_σ. The perfectly isotropic case corresponds to $\Omega_\sigma = 0$, which is the usual flat Friedmann-Robertson-Walker solution (innermost circle, in the limit where the circle has

shrunk to a point, in Fig. 1). Unlike the vacuum Kasner case, all of the Kasner exponents are positive (*i.e.* lie within the white region of Fig. 1) provided that $\Omega_\sigma < 1/4$ (within the larger, solid circle inscribed in the triangle). For this range, none of the scale factors is increasing during the contraction, although they are decreasing at different rates. When $\Omega_\sigma > 1/4$ (third largest circle), then some points on the Kasner circle have a negative Kasner exponent (dashed part of circle) and other points may have all positive Kasner exponents (solid, thick parts of circle).

Thus, ignoring the curvature, the $w = 1$ case with non-zero Ω_σ contracts smoothly but anisotropically to the crunch. In the special case where $\Omega_\sigma = 0$, the contraction is isotropic.

$w > 1$: For $w > 1$, the energy density dominates ($\Omega_\rho \to 1$) as $a \to 0$, and the metric approaches the approximate form

$$a(t) = \left(\frac{t}{t_0}\right)^{2/3(1+w)}, \tag{5.34}$$

$$\beta_i(t) = c_j \frac{2}{\sqrt{3\rho_0}} \frac{1}{w-1}\left[\left(\frac{t}{t_0}\right)^{\frac{w-1}{w+1}} - 1\right], \tag{5.35}$$

where we have chosen the constants of integration so $\beta_i = 0$ at $t = t_0$. The crucial feature is that the time-varying part of the β_i is proportional to t^α where α is *positive* if $w > 1$. This means that the β_i approach a constant and the universe becomes isotropic at the crunch[3]

This simple result is a "no-hair theorem" for universes without spatial curvature: When $w > 1$, an initially anisotropic universe becomes isotropic ($\Omega_\sigma \to 0$) near the big crunch. The $w > 1$ case is stable under anisotropic perturbations. For $w < 1$, the universe becomes increasingly anisotropic in the sense that $\Omega_\sigma \to 1$ as $a \to 0$. For $w = 1$, Ω_σ remains fixed as $a \to 0$. Evolution is smooth (no mixmaster behavior) in all cases, and is well-approximated as a Kasner metric with constant coefficients for sufficiently small a.

[3]Conversely, when $w < 1$, α is negative and the β_i grow rapidly. Thus even if the energy density ρ is dominant initially, the anisotropy grows and eventually dominates near the crunch, a result consistent with our earlier analysis.

1.2. Curvature and chaos. Complex behavior can arise when there is non-zero spatial curvature in a contracting universe. This may seem surprising at first, since the spatial curvature for a homogeneous and isotropic universe grows as $1/a^2$, which increases more slowly than either the anisotropy or the energy density of a component with $w > -1/3$. However, we have seen above that the contracting phase for $w \leq 1$ is anisotropic. We will show below that this can produce rapidly growing curvature perturbations and chaotic behavior. On the other hand, we will see that chaotic behavior is suppressed if $w > 1$ and the contraction approaches isotropy as $a \to 0$.

We now allow the $\sigma^{(i)}$ to have an $\mathbf{x}$-dependence and consider a curved manifold. The spatial Ricci tensor for the metric (5.9) has the form [125]

$$P_a{}^b = \frac{1}{a^2} \sum_{ijk} S_a{}^b{}_{ijk}(\sigma) e^{2(\beta_i - \beta_j - \beta_k)}. \tag{5.36}$$

The functions $S_a{}^b{}_{ijk}$ depend only on the σ^i and their space derivatives, and are independent of time.

The expression (5.36) reveals a crucial connection between the behavior of anisotropy and curvature near the big crunch. In the isotropic limit, $\beta_i = 0$ and (5.36) reduces to the homogeneous and isotropic $1/a^2$ scaling discussed above. However, the terms in (5.36) are essentially ratios of scale factors. Thus, if the anisotropy is growing as $a \to 0$, some terms – involving ratios of expanding and contracting scale factors – will grow, and the corresponding curvature components will scale faster than $1/a^2$. For $w < 1$ the anisotropy dominates near the crunch, and, as we will discuss below, this causes the curvature to grow and induce chaos. By contrast, in the $w > 1$ model, the anisotropy vanishes at the crunch, and the curvature scales as the usual $1/a^2$, which may be neglected.

$w < 1$: In this case, we begin by assuming that the behavior near the crunch is described by the vacuum Kasner solution, with Kasner conditions (5.26) and (5.27). Using the Kasner solution, it is readily seen that the Einstein equation (5.3) contains a leading order term with time dependence t^{-2}.

The second Einstein equation (5.4) is a consistency check for our assumption of ultralocality. For an appropriate choice of the $\sigma^{(i)}$ – a basis for one of the Bianchi universes – this equation vanishes identically and the metric (5.9) solves the Einstein equations.

The third Einstein equation (5.5) indicates that the simple Kasner solutions must break down near the big crunch. If we order the Kasner exponents as $p_1 \leq p_2 \leq p_3$, then the most divergent term in the third Einstein equation comes from terms in the spatial curvature (5.36), with leading time dependence,

$$t^{-2(1-2p_1)}. \tag{5.37}$$

The leading term is more divergent than t^{-2}, since the Kasner conditions (5.26) and (5.27) imply p_1 is always negative. Therefore, our smoothly contracting solutions are not stable to perturbations in the spatial 3-curvature. A small amount of curvature will grow and come to dominate the dynamics before the big crunch.

The behavior of the universe in this regime has been extensively studied and is known to be chaotic [16, 17, 57–59, 65]. The spatial curvature terms cause the Kasner exponents p_i and the principal directions σ^i to become time-dependent during contraction.

More precisely, the exponents and principal directions are nearly constant for stretches of Kasner-like contraction, during which the curvature is negligible. These Kasner-like epochs are punctuated by short intervals when the curvature momentarily dominates. The exponents and principal directions suddenly jump to new values, and then a new stretch of Kasner-like contraction begins during which the curvature terms are again negligible. The universe undergoes an infinite number of such jumps before the big crunch. The chaotic, non-integrable evolution is equivalent to that of a billiard ball [59], which experiences free motion interrupted by collisions with walls. Models with this oscillatory behavior are called *chaotic*.

This presents a problem for cosmological models, as one expects curvature perturbations in any realistic universe will cause the local value of the curvature to vary from point to point.

If each spatial point evolves independently and chaotically, the evolution of nearby points diverges very quickly as contraction continues, and the universe rapidly becomes highly inhomogeneous as $a \to 0$. If $w < 1$ throughout the contracting phase, it seems unlikely that the observed homogeneous universe could emerge from this state after the bounce to an expanding phase [147].

$w = 1$: The chaotic behavior is mitigated in the $w = 1$ case. Recalling our discussion of the curvature-free scenario, it is clear that there are regions of non-zero measure on the Kasner circle for which all of the p_i are positive. We will refer to these points as *stable*. All choices of p_i when $\Omega_\sigma < 1/4$ are stable. If the universe begins at a stable point, the curvature term remains negligible as $a \to 0$ and the contraction is smoothly Kasner-like.

However, when $\Omega_\sigma > 1/4$, some choices of the p_i will have one $p_i < 0$. If the universe begins at one of these points, the curvature term will grow and become dominant, causing the values of p_i and the principal axes σ_i to change. We refer to these points as *unstable*. A more complete analysis [16] reveals that, after a finite number of jumps, the universe hits a point in the open set of stable p_i. From this point onwards, the universe contracts smoothly and without any further jumps.

We call these models *non-chaotic*, since the universe is guaranteed to arrive at a stable point as $a \to 0$. Non-chaotic models (Kasner circles) may contain both stable and unstable points, but they will always oscillate only a finite number of times before arriving in the set of stable points, after which the behavior is integrable.

$w > 1$: For $w > 1$, curvature does not affect the contraction. The key is the time-dependence of the β_i in (5.34) and (5.35), which approach zero as a positive power of t as $t \to 0$. Consequently, the exponential factors e^{β_i} in the metric approach constants. The leading order time-behavior of ${P_a}^b$ is simply that of a homogeneous and isotropic universe,

$$P \sim \frac{1}{a^2} \sim |t|^{-\frac{4}{3(1+w)}}. \tag{5.38}$$

This is always less divergent that t^{-2} for $w > 1$. Thus, even in the presence of initial anisotropy and curvature, the solution for $w > 1$ converges to the isotropic solution represented by the central point on the Kasner sphere in Fig. 1.

We can generalize our cosmic no-hair theorem (described at the end of section 1.1) to include models with spatial curvature. The Einstein equations for a contracting universe with anisotropy and inhomogeneous spatial curvature converge to the Friedmann equation for a homogeneous, flat and isotropic universe if it contains energy with $w > 1$, and that for a homogeneous, flat but anisotropic universe if $w = 1$. The $w < 1$ case becomes highly inhomogeneous and the no-hair theorem is inapplicable.

2. Coupling to p-forms

In section 1, we assumed that the evolution of the universe was dominated by an energy component with fixed equation of state evolving independently of other matter in the universe. The component could have been a scalar field or a perfect fluid. We found chaotic behavior for $w < 1$ in the presence of curvature but non-chaotic behavior for $w \geq 1$.

In this section, we want to consider how the behavior for $w \geq 1$ can change if the fluid is imperfect or couples to other components. In many theories, including Kaluza-Klein, supergravity and superstring models, the relevant energy consists of a scalar field that is coupled to p-forms. Consequently, we will focus on this important example, as others have in the past [16, 57–60].

To determine the effect of the coupling to p-forms on chaotic behavior, our approach is similar to our analysis for spatial curvature, where we assume an initial state in which the spatial curvature is negligible and then check that it remains small. Here we assume that the p-form field strength is initially negligible and ask how its contribution evolves relative to the energy density with equation of state w. Our action is

$$S = \int d^4x\sqrt{-g}\left(\tfrac{1}{2}R - \tfrac{1}{2}(\partial\phi)^2 - V(\phi) - \tfrac{1}{2(p+1)!}e^{\lambda\phi}F_{\mu_1\cdots\mu_{p+1}}{}^2\right), \tag{5.39}$$

where g is the metric, R is the scalar curvature, V is a potential for the scalar field ϕ, p is the rank of the p-form, F is the associated field strength tensor and λ is the coupling constant. The potential $V(\phi)$ is chosen to give fixed equation of state $w \geq 1$ in the absence a p-form coupling, as in (2.24):

$$V(\phi) = -V_0 e^{-\sqrt{3(1+w)}\phi}, \tag{5.40}$$

where V_0 is a positive constant. Throughout this chapter, we assume without loss of generality that $\phi \to -\infty$ as $a \to 0$.

For a given equation of state w and p-form rank, the behavior of the system as $t \to 0$ depends on the coupling λ. We can extend the terminology introduced earlier to describe the properties for a given λ. We classify the p-form coupling parameter λ as *supercritical* if the p-form terms grow relative to the scalar field energy density. We call these models supercritical, as opposed to chaotic, because if $w > 1$ it is not known whether chaos occurs or whether the p-forms merely play a non-negligible role in integrable dynamics. In the special case $w = 1$, chaos is known to occur, and we call these models *chaotic* [57–60]. Values of λ for which the contracting solution with negligible p-forms is stable are called *non-chaotic* (some authors use *subcritical*). These two cases are analogous to those introduced in section 1. If λ is on the boundary between supercritical and non-chaotic, we call λ *critical*. The behavior of critical models may be novel, and will be discussed at the end of this section.

We are assuming that initially the spatial curvature, the anisotropy and the p-form terms are small, and then we check if these conditions are maintained as the universe contracts. Since we are considering models where $w \geq 1$, the model is non-chaotic if the p-forms are negligible. The universe may be approximated initially by the homogeneous isotropic Friedmann-Robertson-Walker form in (5.9) with $\beta_i \approx 0$ and $\sigma^{(i)} = dx^i$. If $w > 1$ and the p-form terms are negligible, $\Omega_\sigma \to 0$ as the crunch approaches. For $w = 1$, Ω_σ remains small but finite. If the isotropic case is unstable, then adding anisotropy cannot restore stability; just as in section 1.2, the isotropic scale factors are the most stable.

It can be shown that the p-form terms involving the spatial gradients of F grow slower than the leading homogeneous time-derivative terms, another example of the ultralocal behavior discussed previously. Hence, we neglect all spatial derivatives of the field strength.

The components of F with purely spatial indices, $F_{i_1\cdots i_{p+1}}$ are called *magnetic* and the components with one time index, $F^{0i_1\cdots i_p}$, are called *electric*, in analogy with the Maxwell action. We will use the labels E and B to indicate their respective contributions. F has a vanishing exterior derivative $dF = 0$. In coordinate notation, neglecting the spatial derivatives of F, this corresponds to

$$\partial_{[0}F_{i_1\cdots i_{p+1}]} = 0, \tag{5.41}$$

where the brackets $[\cdots]$ indicate antisymmetrization. Thus, the magnetic components are constant,

$$F_{i_1\cdots i_{p+1}} = (\text{constant}) \tag{5.42}$$

The equation of motion for F is

$$\nabla_\mu(e^{\lambda\phi}F^{\mu\mu_2\cdots\mu_{p+1}}) = \partial_\mu(e^{\lambda\phi}F^{\mu\mu_2\cdots\mu_{p+1}}) + \Gamma^\mu{}_{\mu\sigma}e^{\lambda\phi}F^{\sigma\mu_2\cdots\mu_{p+1}} = 0. \tag{5.43}$$

Only one set of Christoffel symbols appears due to the antisymmetry of F. Since $\Gamma^\mu{}_{\mu 0} = \frac{\partial}{\partial t}\log\sqrt{-g}$ and $\Gamma^\mu{}_{\mu i} = 0$, we can integrate to find,

$$F^{0i_1\cdots i_p} = \frac{e^{-\lambda\phi}}{\sqrt{-g}} \times (\text{constant}). \tag{5.44}$$

The p-form part of the stress-energy tensor is

$$T_{\mu\nu} = \frac{e^{\lambda\phi}}{(p+1)!}\big((p+1)F_{\mu\mu_2\cdots\mu_{p+1}}F_\nu{}^{\mu_2\cdots\mu_{p+1}} - \tfrac{1}{2}g_{\mu\nu}F^2\big). \tag{5.45}$$

Decomposing (5.45) into electric and magnetic components, and including factors of the metric, we compute the energy density for the p-forms $\rho_p = -T_0{}^0$, which is,

$$\rho_p = \frac{e^{\lambda\phi}}{(p+1)!}\Big(\tfrac{p+1}{2}F^{0i_1\cdots i_p}F^{0j_1\cdots j_p}g_{i_1j_1}\cdots g_{i_pj_p} + \tfrac{1}{2}F_{i_1\cdots i_{p+1}}F_{j_1\cdots j_{p+1}}g^{i_1j_1}\cdots g^{i_{p+1}j_{p+1}}\Big) \quad (5.46)$$

$$= \frac{e^{-\lambda\phi}}{a^{2(3-p)}}\alpha_E^2 + \frac{e^{\lambda\phi}}{a^{2(p+1)}}\alpha_B^2, \quad (5.47)$$

where the positive constants α_E^2 and α_B^2 represent the magnitude of the electric and magnetic energy, respectively. We can now define a new set of fractional energy densities,

$$\Omega_\phi = \rho^{-1}\big(\dot\phi^2/2 + V(\phi)\big), \quad (5.48)$$

$$\Omega_E = \rho^{-1}\frac{e^{-\lambda\phi}}{a^{2(3-p)}}\alpha_E^2, \quad (5.49)$$

$$\Omega_B = \rho^{-1}\frac{e^{\lambda\phi}}{a^{2(p+1)}}\alpha_B^2, \quad (5.50)$$

$$\rho = \tfrac{1}{2}\dot\phi^2 + V(\phi) + \frac{e^{-\lambda\phi}}{a^{2(3-p)}}\alpha_E^2 + \frac{e^{\lambda\phi}}{a^{2(p+1)}}\alpha_B^2. \quad (5.51)$$

where Ω_ϕ is the energy density in the scalar field and Ω_E and Ω_M are the energy densities in electric and magnetic modes. We are assuming that the anisotropy is negligible, so $\Omega_\sigma \approx 0$.

The scaling solution for a ϕ-dominated universe with equation of state w is (2.25),

$$\phi = q\ln|t|, \qquad q = \sqrt{\frac{4}{3(1+w)}}, \quad (5.52)$$

and $a = |t/t_0|^{2/3(1+w)}$. Substituting in (5.47), two terms in ρ_p may be written as

$$\rho_p = \alpha_E^2|t|^{p_E} + \alpha_B^2|t|^{p_B}, \quad (5.53)$$

where p_E and p_B are called the electric and magnetic exponents, respectively. They are,

$$p_E = -\frac{4(3-p)}{3(1+w)} - \lambda\sqrt{\frac{4}{3(1+w)}}, \quad (5.54)$$

$$p_B = -\frac{4(p+1)}{3(1+w)} + \lambda\sqrt{\frac{4}{3(1+w)}}. \quad (5.55)$$

Note that these expressions are invariant under a duality transformation, which takes $p \to 2-p$, interchanges the electric and magnetic modes, and takes $\phi \to -\phi$.

In the Friedmann equation, the scalar field energy density scales as t^{-2}. Consequently, $\Omega_\phi \to 1$ and $\Omega_{E,B} \to 0$ as the universe contracts if both p_E and p_B are both greater than -2. In this case, the p-form contribution is negligible and λ is non-chaotic. Alternatively, if either p_E or p_B is less than -2, the respective p-form terms become large and alter the dynamics.

For $w = 1$, the non-chaotic values of λ are

$$\begin{cases} -\sqrt{8/3} < \lambda < 0 & p = 0 \\ -\sqrt{2/3} < \lambda < \sqrt{2/3} & p = 1 \\ 0 < \lambda < \sqrt{8/3} & p = 2 \end{cases} \tag{5.56}$$

Increasing w causes the interval of non-chaotic couplings to grow, as shown in Fig. 2. In particular, for any p and λ, there exists a critical value $w_{\rm crit}(\lambda, p)$ such that, for $w > w_{\rm crit}(\lambda, p)$ the p-form terms remains negligible. For any set of p-forms and couplings there exists a $\bar{w}_{\rm crit}$, the maximum of 1 (the critical equation of state for curvature) and the $w_{\rm crit}(\lambda, p)$ for each p and λ. Then the contraction is non-chaotic if $w > \bar{w}_{\rm crit}$.

The behavior can be understood in terms of an *effective* equation of state for the action (5.39), using the conservation equation

$$\dot{\rho} = -3\frac{\dot{a}}{a}(1 + w_{\rm eff})\rho, \tag{5.57}$$

where ρ is given by (5.51). Using (5.48)–(5.51), the equation of motion for ϕ and the Friedmann equation, we find

$$w_{\rm eff} = w_\phi \Omega_\phi + \tfrac{3-2p}{3}\Omega_E + \tfrac{2p-1}{3}\Omega_B, \tag{5.58}$$

where

$$w_\phi = \frac{\dot{\phi}^2/2 - V(\phi)}{\dot{\phi}^2/2 + V(\phi)} \tag{5.59}$$

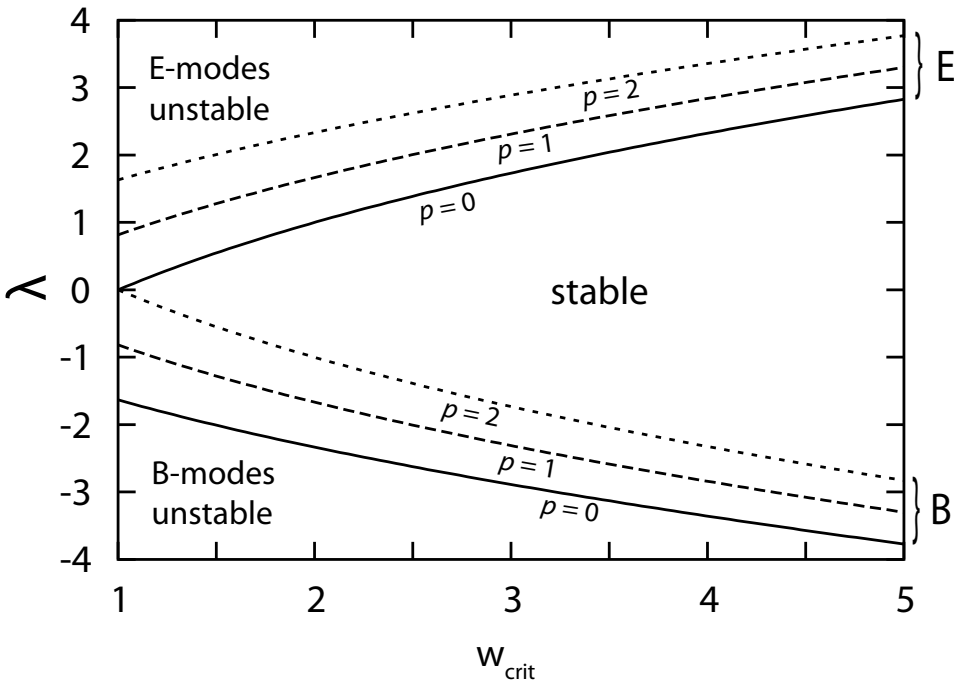

FIGURE 2. The four dimensional electric and magnetic couplings λ as a function of the critical equation of state for $p = 0, 1, 2$. The upper and lower three curves represent the critical electric and magnetic exponents, respectively. A form with given p and λ is stable in a universe with equation of state w if the point (w, λ) lies between the two curves for the given p.

is the equation of state for the decoupled scalar field and $\Omega_\phi + \Omega_E + \Omega_B = 1$. The expression (5.58) is exact, valid for all values of the Ω_i assuming the background is homogeneous, flat and isotropic. For the electric and magnetic contributions, we can introduce $w_E = \frac{3-2p}{3}$ and $w_B = \frac{2p-1}{3}$, respectively. The w_{eff} is just the Ω-weighted average of w_ϕ, w_E and w_B.

All the λ dependence of w_{eff} is contained in the time evolution of the Ω_i; w_ϕ, w_E and w_B do not depend on λ. Both w_E and w_B are always less than or equal to unity, and at least one is strictly less.

If the p-form coupling λ is non-chaotic, the behavior is simple. The quantities Ω_E and Ω_B rapidly approach zero as Ω_ϕ approaches one, and the universe is dominated by the scalar field, with the equation of state w_ϕ. This is the non-chaotic case, discussed in section 2.

Alternatively, if the p-form coupling is supercritical, Ω_E and Ω_B grow. The averaging of the E and B component ensures $w_{\text{eff}} < w_\phi$. If $w_\phi = 1$ then $w_{\text{eff}} < 1$. In this case, the anisotropy grows and chaotic oscillations occur. It is not known if this happens in the $w_\phi > 1$ case. If in addition, the p-form coupling is critical (so $w_{\text{crit}} = 1$), it turns out that the model is equivalent to an infinite-dimensional hyperbolic Toda system. There are an infinite number of jumps from one Kasner-like solution to the next, but the system may be formally integrable [59, 60]. It is not clear what the physical ramifications of this behavior are.

3. Time-varying equation of state

In a realistic cosmological model, the equation of state will not be constant, but will depend on the scale factor and approach some limiting value $w \to \bar{w}$ as $a \to 0$. If $\bar{w} \neq w_{\text{crit}}$, none of the above analysis changes substantially. The model is supercritical if $\bar{w} < w_{\text{crit}}$ or non-chaotic if $\bar{w} > w_{\text{crit}}$. The critical case, $\bar{w} = w_{\text{crit}}$, is more subtle, and the time dependence of $\bar{w}$ can be significant. In this section, we assume $w_{\text{crit}} = 1$, as this is the most important case, and analyze what happens when $w_\phi \to 1$ at the crunch. We can expand w_ϕ as

$$w_\phi(a) = 1 + \gamma(a), \tag{5.60}$$

where γ is a small function of the scale factor such that $\gamma \to 0$ as $a \to 0$.

If there is no p-form with critical coupling, then using (5.11) and (5.21), it can be shown that if $\gamma(a) \log a$ approaches a constant as $a \to 0$, then the behavior is essentially the same as the $w = 1$ case, *i.e.* non-chaotic. The radius of the Kasner circle in figure 1 shrinks, if $w \to 1^+$, or expands, if $w \to 1^-$. If $\gamma \to 0$ so slowly that $\gamma(a) \log a$ diverges as $a \to 0$, then the anisotropy is eliminated if γ approaches zero from above or the chaos is restored if γ approaches from below.

Alternatively, if the model has a p-form with critical coupling, the Kasner contraction will be stable if the p-form contribution to the equations of motion remain subdominant, or, equivalently, if the ratio of the p-form terms to the other terms vanishes in the $a \to 0$ limit.

For magnetic modes with critical coupling $\lambda_{\text{crit}} < 0$, we find:

$$\log \frac{\Omega_M}{\Omega_\phi} = \log \frac{e^{\lambda_{\text{crit}}\phi}/a^{2(p+1)}}{\dot{\phi}^2/2 + V(\phi)} \sim -C \int_a^{a_0} \frac{da'}{a'} \gamma(a'), \tag{5.61}$$

where C is a positive constant and by $\sim$ we mean up to terms finite in the $a \to 0$ limit. The behavior is identical if the electric modes have critical coupling. If $\gamma \to 0$ very slowly, for example

$$\gamma(a) \sim |\log a|^{-1} \tag{5.62}$$

so that the integral diverges as $a \to 0$, then the ratio goes to zero and Ω_M becomes negligible in the $a \to 0$ limit. This ensures that the term is small, and never grows to influence the dynamics.

Let us investigate what conditions on the potential will give us a γ of this form. If we combine the Friedmann equation and equation of motion for ϕ, we obtain (2.34)

$$\frac{d\psi}{d\log a} = 3\Big(\psi - \frac{V_{,\phi}}{\sqrt{6}V}\Big)(\psi - 1)(\psi + 1), \tag{5.63}$$

where , ϕ denotes a derivative by ϕ and

$$\sqrt{6}\psi = \frac{d\phi}{d\log a}. \tag{5.64}$$

The equation of state (5.59) can be expressed in terms of ψ,

$$w_\phi = 1 + \gamma = 2\psi^2 - 1. \tag{5.65}$$

We can obtain $w_\phi \to 1^+$ as $a \to 0$ for any negative potential which is bounded (for large negative values of ϕ) by $-Ce^{-\sqrt{6}\phi}$, where C is a positive constant (see (5.40)). The kinetic energy increases more rapidly than the potential energy in these cases, and so w_ϕ approaches unity at the crunch. In particular, the potential need not be bounded below. In general, any potential which can be expressed in the form

$$V(\phi) = 2\,W'(\phi)^2 - 3\,W(\phi)^2 \tag{5.66}$$

satisfies positive energy [22]. Hertog *et al.* [94] have shown that the potential

$$-V_0 e^{-c\phi}, \tag{5.67}$$

where V_0 and c are positive constants, can be expressed in this form provided $c < \sqrt{6}$, and so satisfies positive energy. For $c \geq \sqrt{6}$, solutions exist with total ADM energy that is unbounded below.

For the potential (5.67), $V_{,\phi}/V = c$. In the case $c < \sqrt{6}$, we find

$$\gamma \propto a^{\gamma}, \tag{5.68}$$

where γ is a positive constant. Consequently, $\gamma \log a \to 0$ as $a \to 0$ and the p-form with critical coupling is not suppressed. However, when $c = \sqrt{6}$, the solution to the equations of motion show that $\gamma \log a$ approaches a constant, so the p-form can be suppressed when positive energy is violated.

The potential

$$V(\phi) = -V_0 e^{-\sqrt{6}\phi} |\phi|^n, \tag{5.69}$$

(or more generally, an exponential times any finite order polynomial) satisfies positive energy (*i.e.* can be expressed in the form (5.66)) for $n \leq -1$. Solving the equation of motion (5.63) for large ϕ, we find that for $n \leq -1$ the p-form with critical coupling is not suppressed. Surprisingly, for $n > -1$ the ratio (5.61) goes to zero, and the solution is stable. For the broad class of potentials (5.67) and (5.69), the parameters for which they satisfy positive energy turn out to be exactly those which do not suppress the p-form. It is an open question whether any potential can be constructed which will suppress the $w_{\text{crit}} = 1$ p-form and satisfy positive energy.

4. Extra dimensions and orbifolds

In models in which gravity is fundamentally higher dimensional, the detailed global structure of the extra dimensions can suppress or enhance chaos in the four dimensional

theory. We consider two simple compactifications of five dimensional gravity, on S^1 and S^1/Z_2. In the first, the chaotic nature of pure five dimensional gravity descends to the four dimensional theory. In the second, the chaotic behavior is suppressed. Models of quantum gravity also generally include additional matter fields in the extra dimensions. As an example, we discuss the compactification of heterotic M-theory to four dimensions and find that its behavior during gravitational contraction is on the borderline between smooth and chaotic.

Consider a five-dimensional, flat universe without matter fields. We know from the study of general Kasner universes [65] that it will exhibit chaotic behavior. Now compactify one dimension on S^1. We know that the four-dimensional effective theory describes Einstein gravity coupled to a free scalar field. The scalar field describes the volume of the S^1 – it is a simple example of a moduli field. As all of our preceding arguments regarding gravitational contraction are local in nature, we expect that the resulting system should be chaotic as well. However, a free scalar field has equation of state $w = 1$. According to our analysis in section 1, one might think that the behavior should be non-chaotic. What has happened to the chaos?

The resolution comes from the fact that we have neglected the Kaluza-Klein one-form. The dimensionally reduced action (2.88),

$$S = \int d^4x\sqrt{-g}\left(\tfrac{1}{2}R - \tfrac{1}{2}(\partial\phi)^2 - \tfrac{1}{4}e^{\sqrt{6}\phi}F^2\right), \tag{5.70}$$

describes a vector field coupled to a free scalar and to gravity. The coupling $\lambda = \sqrt{6}$ is outside the stable range for a one-form in four dimensions. Therefore, the four dimensional theory is chaotic, as we would have guessed, but we have to include the interactions with p-forms to see that this is so.

If, instead of compactifying the fifth dimension on S^1, we compactify on the orbifold S^1/Z_2 we obtain the action (2.94), which has the Kaluza-Klein one-form projected out: it

sets $A_\mu = 0$ in (5.70). The absence of this vector field in the effective action thus implies that the four dimensional theory is no longer chaotic.

While orbifolding suppresses some gauge fields and p-forms that would cause chaotic behavior, in some models there are additional p-forms in the bulk. These p-forms, after dimensional reduction, may themselves lead to chaotic behavior. An illustrative example is heterotic M-theory, which includes a three-form field. To zeroth order in the eleven dimensional gravitational coupling κ, the low-energy four dimensional effective action is from (2.105) and [126]:

$$S^{(0)} = \frac{\pi\rho V}{\kappa^2}\int dx^4\sqrt{-g}\big(R - (\partial a)^2 - (\partial c)^2 - e^{-\sqrt{8/3}c}(\partial\chi)^2 - e^{-\sqrt{8}a}(\partial\sigma)^2\big), \tag{5.71}$$

where we have rescaled the fields in the action so the kinetic energies are canonically normalized. The scalar field c is the radion, which governs the brane separation. The Calabi-Yau volume modulus a and scalar field σ (which comes from the eleven dimensional three-form) do not couple to c, and so can be ignored. However, the three-form modulus χ couples to c and the exponent is critical $\lambda = -\sqrt{8/3}$. Hence, the theory does not lead to stable Kasner contraction. Including the first order ($\kappa^{2/3}$) correction to the action does not change the result. As we discussed in section 3.1 of chapter 4, the Yang-Mills gauge fields have a coupling to the radion proportional to $1 \pm \xi\alpha_0 e^c$. Damour *et al.* [59, 60] have shown that Yang-Mills one-forms have the same asymptotic behavior as the Abelian case considered here. Since the radion $c \to -\infty$, the effective $\lambda = 0$ and according to (5.56), the gauge theories on the fixed planes do not cause chaos.

As heterotic M-theory is critical, it is quite conceivable that higher order corrections will lead to a different behavior during cosmological contraction. There are a number of kinds of corrections to (5.71) that could push the theory away from criticality and render it either chaotic or non-chaotic; but, it is not yet known which behavior occurs.

Wesley *et al.* [191] consider the consequences of compactifications on more complex manifolds, and show that the chaotic properties are determined by the cohomology and Killing vectors of the compact manifold.

The new results presented in this chapter build on over three decades of preceding research on the behavior of cosmological models contracting to a big crunch. The classic work focused on cases where the equation of state of the dominant energy component is $w \leq 1$ and w is constant. The essential results in this case are:

- For a perfect fluid with $w < 1$, the contraction is smooth and anisotropic in the absence of curvature and chaotic mixmaster if there is non-zero curvature.
- For a perfect fluid with $w = 1$, the contraction is smooth and *anisotropic* in the absence of curvature. With curvature, the contraction is *anisotropic* also, although, depending on the initial anisotropy, the contraction may undergo a finite number of jumps from one Kasner-like behavior to another.
- For a free scalar field coupled to p-forms with coupling $e^{\lambda\phi}$, the contraction is *chaotic mixmaster* if the coupling λ is outside a finite interval of non-chaotic λ. The *mixmaster case is non-integrable* and the *critical case may be integrable*.

In this chapter, we have extended this work to include cases where $w > 1$, a situation that arises naturally in some recent models with a big crunch/big bang transition, such as the cyclic and ekpyrotic models. We have added the following results:

- For perfect fluid with $w > 1$, the contraction is smooth and converges to *isotropic* at the crunch. The Einstein equations converge to ultralocal, homogeneous and isotropic Friedmann equations.
- For a scalar field coupled to p-forms, there exists a w_{crit} such that the contraction is smooth and *isotropic* for $w > w_{\text{crit}}$.
- If w is time-varying and approaches one from above sufficiently slowly the contraction is *smooth* and *non-chaotic*, even in the presence of a p-form with critical equation of state $w_{\text{crit}} = 1$.

- In models with an extra dimension, compactification generically produces a scalar field and p-forms. Z_2 orbifolding forces some p-forms to zero and, thereby, suppresses their contributions to chaos.

In this chapter we have studied how chaotic mixmaster behavior may be suppressed in models involving a big crunch/big bang transition. In particular, the ekpyrotic and cyclic models already include some of the required ingredients including a scalar field with $w > 1$ and Z_2 orbifolding.

CHAPTER 6

Ekpyrotic and cyclic applications

Nearly scale invariant cosmological perturbations are an important ingredient of modern cosmology. Observations of the cosmic microwave background [167] and large scale structure [175] suggest that the spectrum of initial perturbations must be scale invariant to about one part in ten. It has recently been shown that there are only two stable mechanisms for generating nearly scale invariant perturbations, depending on the equation of state of the universe w: an expanding universe with $w \approx -1$ and a contracting universe with $w \gg 1$ [88]. These are called the inflationary [124] and ekpyrotic [110, 112, 113] mechanisms, respectively.

In this chapter, we examine what happens to perturbations during a transition between these two regimes. A contracting universe cannot begin expanding without undergoing a bounce, which generically involves either a singularity or a violation of the null energy condition [110]. However, it is quite simple to construct models that move from an expanding phase to a contracting phase: a scalar field with a potential that crosses zero is all that is required.

This is particularly relevant to the cyclic scenario [170–172] which contains a scalar field, the radion, which models both the accelerating, dark energy epoch and the contracting ekpyrotic phase. We study the full perturbation spectrum of this model, including the transition from expansion to contraction, and point out that the modes generated in the dark energy epoch continue to grow, even when they are far outside the horizon, and have roughly the same amplitude as the modes produced in the ekpyrotic phase. This is surprising because the dark energy and ekpyrotic phases in the cyclic model are generated at vastly different

energy scales: the dark energy scale is 10^{-3} eV whereas the ekpyrotic scale must be many GeV.

The growth of dark energy modes when far outside the horizon seems to contradict the usual intuition from inflation, in which modes do not grow outside the horizon. We point out that this is due to a choice of the variable ζ, the comoving curvature perturbation, which projects out the growing mode we are interested in. The Newtonian potential Φ does grow, and whether this growth is physically important depends on the details of the theory under consideration.

We show that the modes produced in the dark energy phase are not observationally relevant in the model. The model requires that the visible perturbations are produced in the ekpyrotic phase, leaving the perturbations produced in the dark energy epoch far outside the horizon. It is possible to tune the ekpyrotic (but not the cyclic) model so that dark energy modes, or even the transition feature, are part of the observable spectrum, although there is no reason to prefer such a model.

We also discuss the behavior of the cyclic model over longer intervals. Over the course of each cycle, the scale factor is multiplied by some exponentially large factor e^N. A large band of dark energy and ekpyrotic perturbations are produced and exit the horizon. Some fraction of these modes reenter during the kinetic energy and radiation dominated expansion.

We show that dark energy domination is not necessary for the global consistency of the cyclic model. The original treatment of cyclic model [171] pointed out that sixty e-folds of dark energy domination ensures that there was less than one particle per Hubble volume at the transition to the ekpyrotic phase. This is sufficient to ensure that fluctuations in one cycle do not interfere with the quantum generation of perturbations in the next. However, we show that the classical suppression of scalar field fluctuations inside the horizon ensures that the model is consistent without any dark energy domination.

We discuss the global structure of the model, and show that once perturbations exit the horizon for the last time, they are continually amplified and grow to ever larger scales.

This means that superhorizon perturbations quickly become nonlinear, and that once two patches fall out of causal contact they become asynchronous over the course of one cycle. This is a purely metaphysical issue, however, as physical observable, such as the curvature perturbation inside a Hubble patch, do not grow.

In section 1 we briefly review the cyclic model and the ekpyrotic mechanism for generating nearly scale invariant perturbations. In section 2 we discuss the transition from dark energy domination to the ekpyrotic phase, including a qualitative discussion and numerical results. In section 3, we discuss constraints on the cyclic model parameters and show that the feature generated at the transition must be far outside the horizon. In section 4 we discuss the global features of the cyclic model.

1. The ekpyrotic and cyclic models

The ekpyrotic scenario [110, 112] is a cosmological scenario that reproduces many of the successes of inflation in a very different context. The crucial difference is that the primordial density perturbations are produced in a *contracting* phase by a scalar field with a steep potential, rather than an expanding phase with a flat potential. In terms of the equation of state, the ekpyrotic universe is contracting with $w \gg 1$, whereas the inflationary universe is expanding with $w \approx -1$. In chapter 5, we saw that these regimes are both attractors, to which no-hair theorems apply. Gratton *et al.* [88] have shown, moreover, that these are the only two stable regimes which produce nearly scale-invariant perturbations.

In order to match the ekpyrotic phase onto the hot big bang, it is necessary for the empty, homogeneous contracting universe to bounce: that is, the big crunch needs to match onto a big bang. The ekpyrotic model is essentially higher-dimensional, and implements the bounce by assuming that the bounce comes from colliding branes in a universe with additional dimensions. Because of the interest in constructing phenomenologically viable string models from Horăva-Witten theory [95, 96, 126], a particularly interesting way to implement the theory is the collision of two S^1/Z_2 orbifold fixed planes. The scenario, in its simplest form, consists of a five-dimensional bulk spacetime bounded by four-dimensional

S^1/Z_2 orbifold planes. Branes trapped at each fixed plane contain the matter and radiation of the universe. The five-dimensional geometry is similar to the Randall-Sundrum model [156], in which one brane has positive tension and the other has negative tension, with a cosmological constant in the bulk. Near the brane collision, effects due to a warped bulk go away. Moreover, non-perturbative string effects should vanish, as the string coupling $g_s \to 0$ as the separation between the branes vanishes ($\phi \to -\infty$). Thus, the geometry simplifies to a five dimensional Milne universe:

$$ds^2 = -dt^2 + dx_1^2 + dx_2^2 + dx_3^2 + |vt|^2 dy^2, \tag{6.1}$$

where $y \in [0, 1]$ is the orbifold direction and v is the velocity with which the branes collide. The bounce is a singular event. In five dimensions, a dimension instantaneously vanishes when the branes are coincident. In four dimensions, it is a big crunch: the scale factor vanishes instantaneously. Results about the bounce differ widely depending on what prescription for regulating the singularity is used [14, 27, 41, 74, 76, 82, 106, 107, 111, 113, 128, 133, 134, 151, 180]. Regulating in four dimensions is problematic, because a nonsingular bounce invariably violates the null energy condition [110] whereas it is difficult to know what matching rule to apply at a singular bounce. We assume the results of Tolley *et al.* [177, 178] (see also [180]) which assume that a five-dimensional Milne-like bounce occurs in M-theory and use unitarity to match incoming and outgoing modes across the singularity. Note that in the S^1/Z_2 model (2.94), the temperature on the orbifold planes remains finite at the bounce, even though the bulk scale factor goes to zero: this is because the orbifold planes couple to the metric $e^{-\sqrt{2/3}\phi} h_{\mu\nu}$, and the bulk scale factor and radion conspire to keep the brane scale factor from contracting to zero.

One brane, the visible brane, has the standard model on it. It is assumed that fields on the other, hidden brane, do not interact with the visible brane, other than gravitationally. Away from the bounce, the ekpyrotic model may be described by an effective four dimensional

theory. The separation of the branes is given by a canonically normalized scalar field ϕ, and goes as as $L \sim e^{\sqrt{2/3}\phi}$ (just as in the S^1/Z_2 compactifications of chapter 2 section 3).

The branes have an interaction that is modeled in four dimensions by a potential ϕ. This is the force that draws them together and causes the collision. It is put in by hand, but it should ultimately be derivable from M-theory, perhaps from the exchange of M2-branes between M5-branes, the exchange of strings between D-branes, effects from interacting moduli, or other dynamics in the bulk. The most important feature of the potential is that it needs a steep, negative segment in order for the ekpyrotic mechanism to operate (see section 2.1 of chapter 2). The simplest such potential is just a negative exponential

$$V(\phi) = -V_0 e^{-c\phi}, \qquad \text{for} \qquad \phi_{\text{end}} < \phi < 0, \tag{6.2}$$

where c and V_0 are positive constants, with $c \gg 1$. Nearly scale-invariant perturbations are produced by a scalar field with rolling down this potential. The universe is very slowly contracting, with $w \gg 1$. The magnitude of the Hubble parameter is rapidly increasing, so $(aH)^2$ is growing rapidly and modes are exiting the horizon. Some authors have argued that the perturbations produced in the Newtonian potential do not produce physical curvature perturbations [49, 102, 129] but in the Tolley *et al.* prescription, this is not the case, and the Newtonian potential is converted into fluctuations in the density of radiation at the bounce.

The cyclic model is a more ambitious version of the basic ekpyrotic scenario in which the orbifold planes collide periodically throughout history. The solution is an attractor, and thus explains the initial condition problem in cosmology. The potential for the radion, $V(\phi)$, illustrated in figure 1, has three basic regions (the point $\phi = 0$ is fixed arbitrarily):

$\phi > 0$: A region where $V(\phi)$ is flat and has a small, positive value. This is the "dark energy" region, which accounts for the dark energy density observed today. Other forms of this potential, such as a metastable minimum, may also be possible.

$0 > \phi > \phi_{\text{end}}$: A very steep region in which $V(\phi)$ is negative. This is called the *ekpyrotic phase*. The potential must satisfy the "fast-roll" conditions, that $\frac{1}{2}(V'/V)^2 \gg 1$

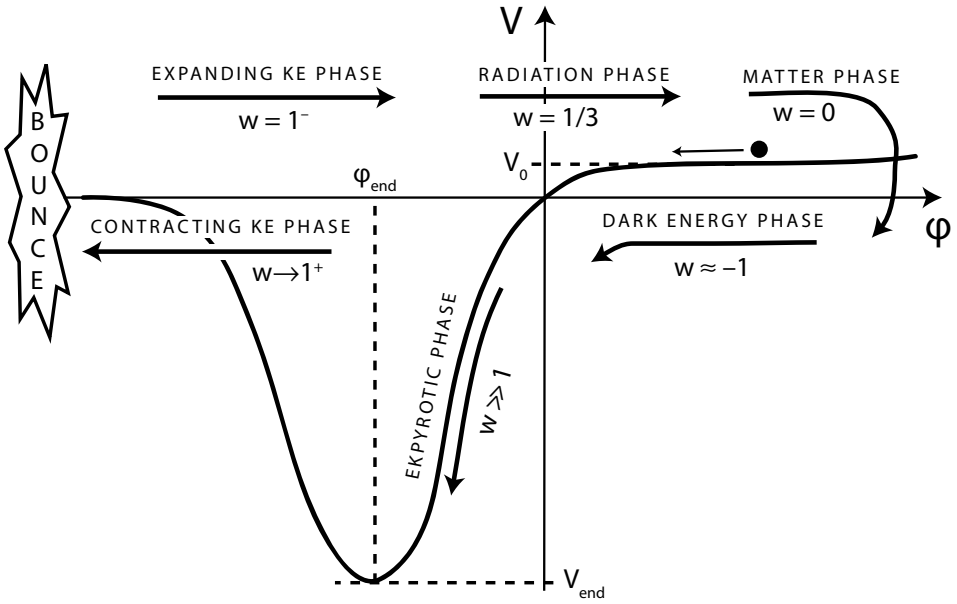

FIGURE 1. A schematic plot of the scalar field potential in the cyclic model. The motion of the scalar field (illustrated by a ball rolling down from the plateau) during the course of a cycle is illustrated. The different phases are marked along with the equation of state w. The arrows show the direction of motion of the scalar field in that phase. To the right represents a growing separation between the orbifold planes; to the left represents a decreasing separation.

and $|VV''/(V')^2 - 1| \approx 0$. The value of the potential when these conditions cease to be satisfied is $V(\phi_{\rm end}) = -V_{\rm end}$.

$\phi_{\rm end} > \phi$: String theory seems to require that $V(\phi) \to 0$ as $\phi \to -\infty$, as non-perturbative effects shut off at small separations. However, consistency of the model only requires only the weaker condition, that $V(\phi)e^{\sqrt{6}\phi} \to 0$ as $\phi \to -\infty$. (This implies $w \to 1^+$ as $\phi \to -\infty$.)

The constraints on the potential are discussed in detail in [115]. In our analysis, we assume a particularly simple form for $V(\phi)$:

$$V(\phi) = V_0(e^{b\phi} - e^{-c\phi})F(\phi), \tag{6.3}$$

where $c \gg 1$ and $0 \geq b \ll 1$. The function F cuts off the potential at $\phi_{\rm end}$: $F(\phi) \to 1$ for $\phi \gg \phi_{\rm end}$ and $F(\phi) \to 0$ rapidly for $\phi \ll \phi_{\rm end}$. The particular form of F is not important for our analysis, as its shape will only correct the amplitude of the overall spectrum by a factor of order unity. A potential exception is when analyzing the conditions for chaos, as discussed in chapter 5.

A cycle goes as follows. Starting today, shortly after the start of dark energy domination, the universe remains dark energy dominated for some number of e-foldings, labeled $N_{\rm DE}$. In this epoch, the universe has $w \approx -1$.

After the period of dark energy domination, the potential crosses the point $\phi = 0$ and $V = 0$ and turns negative. In the Friedmann equation,

$$3H^2 = \rho + \tfrac{1}{2}\dot{\phi}^2 + V(\phi), \tag{6.4}$$

with $V < 0$, matter will continually redshifts away, as V decreases, until $H = 0$. At this time, $\dot{\phi}$ begins to grow rapidly until the $w \gg 1$ ekpyrotic attractor – in which $\frac{1}{2}\dot{\phi}^2 \approx -V(\phi)$ – is very quickly reached.

In both the dark energy and ekpyrotic regimes, $|aH|^2$ is increasing and so modes are exiting the horizon. This occurs for quite different reasons in the two regimes. In the dark energy epoch, H is nearly constant and a is increasing exponentially, whereas in the ekpyrotic epoch, a is nearly constant and the magnitude of H is rapidly increasing.

When $\phi \approx \phi_{\rm end}$, the ekpyrotic phase ends, and the kinetic energy $\dot{\phi}^2/2$ dominates over the negative potential. This phase has equation of state $w = 1$. In a short time, $\phi \to -\infty$ and the four dimensional scale factor $a \to 0$. A small amount of radiation is produced on the branes at the big crunch. For our purposes, this is parameterized by $T_{\rm RH}$ the "reheat

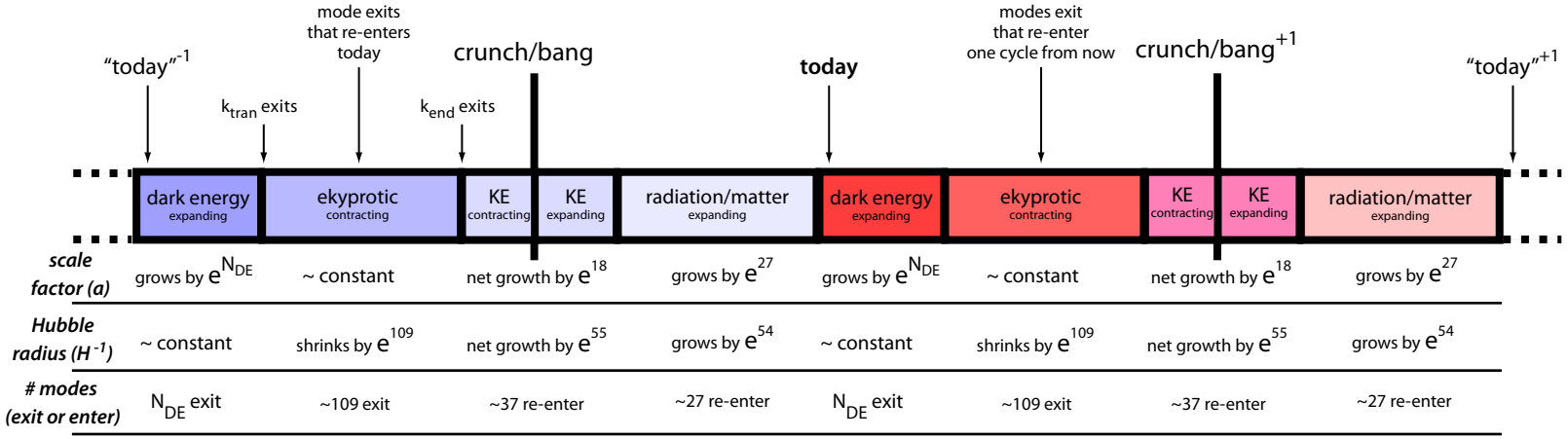

FIGURE 2. Time-line for two complete cycles in the cyclic model, with time running from left to right. Typical values of the reheat temperature $T_{\rm RH} = 1\,{\rm GeV}$ and potential minimum $V_{\rm end} = 10^{12}\,{\rm GeV}$ were used.

temperature." This is the temperature at which the densities of kinetic energy and radiation are equal. In order to cycle correctly [115, 171] the velocity of the branes coming out of the bounce must be greater than that going in. This is possible only because one of the branes has negative tension.

The kinetic phase is nearly symmetric, with the contraction going into the bounce canceled by the expansion coming out of it. However, since the scalar field has more kinetic energy coming out of the bounce, the universe remains kinetic energy dominated as the scalar field crosses the ekpyrotic part of the potential and reaches the dark energy plateau. This asymmetry causes a net expansion during this epoch.

Radiation domination begins at the reheat temperature, which must be greater than the temperature of big bang nucleosynthesis in order for the successes of the regular hot big bang model to be reproduced in the cyclic model. As the universe expands, matter domination eventually begins and structures form. Eventually, dark energy domination begins. This marks the completion of one cycle.

At some point, the field rolling up the plateau turns around and begins rolling back down Whether this happens during radiation, matter or dark energy domination is immaterial. The complete time-line is reproduced in figure 2, with some notes about the relevant scales. These will be analyzed in section 3. First, we focus on the transition from expansion to contraction.

2. The transition from expansion to contraction

In this section, we study the production of perturbations in universes that move from an expanding, dark energy epoch to a contracting, ekpyrotic epoch. We use the potential

$$V(\phi) = V_0(e^{b\phi} - e^{-c\phi}), \tag{6.5}$$

where V_0, b and c are positive constants. The solution for $\phi > 0$ is the dark energy attractor, whereas for $\phi < 0$ it is the ekpyrotic attractor. Thus, we require $0 \leq b \ll 1$ and $c \gg 1$. It is

possible to analytically describe the evolution of modes in these attractor regimes in terms of Bessel functions. The dark energy phase is really very low energy inflation, and can be treated using the same techniques. However, we are also interested in the transition regime, where a numerical solution is necessary. First, however, we discuss qualitative expectations for the spectrum.

We follow the perturbations in the Mukhanov variable [139] $u = \Phi/\dot{\phi}$, where Φ is the Newtonian potential. The evolution equation is given in Fourier space by (2.53)

$$u_k'' = -\left(k^2 - U_{\rm pot}\right) u_k, \tag{6.6}$$

where $U_{\rm pot}$ is (2.54)

$$U_{\rm pot} = z(1/z)'' \qquad \text{and} \qquad z = a\phi'/\mathcal{H}, \tag{6.7}$$

and we use primes to denote conformal time derivatives throughout this chapter. We assume modes start in their Minkowski ground state (2.57), so that $|u_k|^2 = (2k)^{-3}$. This assumption should be valid for modes at sufficiently small scales in the dark energy epoch.

2.1. Qualitative expectations. When $k^2 \gg U_{\rm pot}$ in (6.6), the modes are well within the horizon and oscillate, $u_k \sim e^{\pm ik\tau}$. However, for long-wavelength modes, with $k^2 \ll U_{\rm pot}$, the explicit solutions to (6.6) are

$$u_1(\tau) = z(\tau)^{-1}, \tag{6.8}$$

$$u_2(\tau) = \frac{1}{z(\tau)} \int^{\tau} d\tau' \, z(\tau')^2. \tag{6.9}$$

In general, one of these modes will be dominant. For dark energy domination, where z is roughly constant, it is the second (6.9) whereas in the ekpyrotic phase, the magnitude of H is growing rapidly and it is (6.8). When modes are outside the horizon, they all evolve in concert according to these solutions. The limits of integration on u_2 should be defined to

make the integral finite. In the dark energy dominated epoch, the growing mode is

$$u_{2,\mathrm{DE}}(\tau) = \frac{1}{z(\tau)} \int_{-\infty}^{\tau} d\tau' \, z(\tau')^2 . \tag{6.10}$$

The ekpyrotic decaying mode is

$$u_{2,\mathrm{ek}}(\tau) = -\frac{1}{z(\tau)} \int_{\tau}^{0} d\tau' \, z(\tau')^2 , \tag{6.11}$$

where τ is negative and $\tau \to 0$ corresponds to the crunch. The modes (6.10) and (6.11) have finite, equal limits as $H \to 0$ (*i.e.* $z \to \pm\infty$). However, their derivatives do not match, which demonstrates that the dark energy growing mode does not entirely match onto the ekpyrotic decaying mode, but rather must match on to a combination of the growing and decaying modes.

The usual intuition from inflationary perturbation theory – that the amplitude of modes is conserved outside the horizon – breaks down. Instead, the amplitude continue to grow, while the shape of the spectrum is preserved, and modes produced at radically different energy scales, the dark energy and inflationary scales, smoothly match. The conflict arises because the inflationary result is derived using the variable ζ, the canonical conjugate to Φ:

$$\zeta = \frac{2}{3a^2(1+w)} \left(\frac{\Phi}{a'/a^3} \right)' = \frac{2}{3a^2(1+w)} (zu)' . \tag{6.12}$$

The $1/z$ mode is thus projected out by ζ. In an expanding universe, this is the decaying mode of Φ, while the growing mode is preserved. However, in a contracting universe, the two modes are interchanged, and the $1/z$ mode is the growing mode. Thus, the variable ζ, while it is conserved outside the horizon, is incapable of seeing the growth of perturbations in the contracting phase, because it is orthogonal to the growing mode.

As we alluded in chapter 2, U_{pot} is only loosely connected to the horizon, whose comoving scale is given by aH. In fact, U_{pot} can differ from $(aH)^2$ by several orders of magnitude: in the expanding phase the scale is much larger than the horizon, $U_{\mathrm{pot}}/(aH)^2 \sim b^2$, whereas

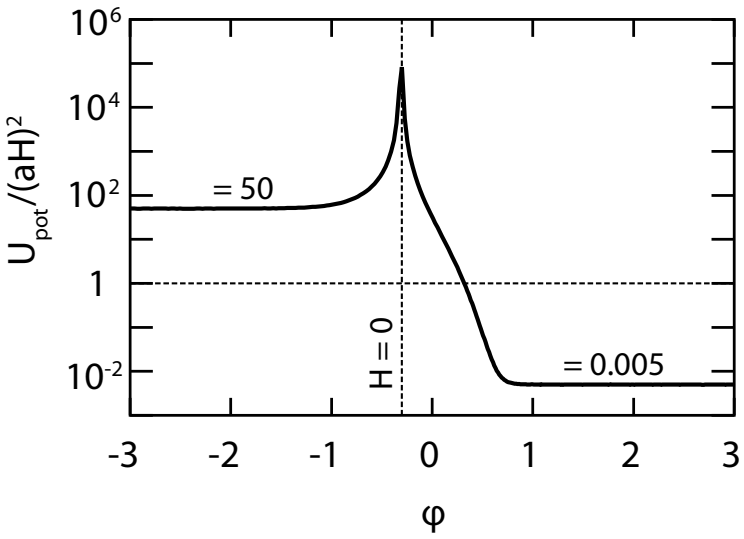

FIGURE 3. Plot of $U_{\rm pot}/(aH)^2$ against ϕ for the potential $V_0(e^{0.1\phi} - e^{-10\phi})$. In the attractor regimes, $U_{\rm pot}$ takes values which differ from the horizon scale aH by two orders of magnitude in each direction. The value of ϕ for which the Hubble parameter vanishes, $\phi = -0.31$, is marked.

in the contracting phase it is much smaller, $U_{\rm pot}/(aH)^2 \sim c^2$ (see figure 3). Thus, in this section we use the terms "outside the horizon" and "inside the horizon" loosely to refer to modes that satisfy the conditions $k^2 \ll U_{\rm pot}$ and $k^2 \gg U_{\rm pot}$, respectively.

We have seen that $U_{\rm pot}$ increases without limit during contraction, and so all modes eventually leave the horizon. If $U_{\rm pot}$ increases monotonically, we can associate a critical time τ_k with each mode, defined implicitly through $k^2 = U_{\rm pot}(\tau_k)$, marking the mode exiting the horizon. If this transition occurs while the universe is in the dark energy epoch, we call the mode a dark energy mode one. If it occurs well into the contracting phase, we call the mode ekpyrotic. In between we have transition modes. The dark energy modes necessarily have the longest wavelengths; the ekpyrotic ones the shortest. If $U_{\rm pot}$ does not increase monotonically, or even goes negative, these labels are not so well defined. This does not pose any practical problem interpreting the spectrum, as any decrease in $U_{\rm pot}$ is typically brief and quite severe, and thus will only affect the shape of the spectrum for modes that left

the horizon very recently. Moreover, $U_{\rm pot}$ increases without limit as $\phi \to -\infty$, so each mode has a time at which it last moves outside the horizon.

Well after τ_k, any mode behaves like the $k = 0$ growing mode (6.8); it "freezes out". Hence the spectrum of frozen-out dark energy modes at a given time preserves its shape as time passes, even through the transition from dark energy domination to decelerating expansion to contraction. This is useful since it allows us to apply some results from inflationary theory to predict the shape of this part of the spectrum. Likewise, we can use ekpyrotic calculations to predict the shape of the spectrum of ekpyrotic modes. The amplitude of modes, however, is not conserved outside the horizon in the ekpyrotic phase; rather, modes grow as $1/z(\tau) = \mathcal{H}/a\phi'$. The scalar spectral index is defined as

$$n_s - 1 = \frac{d\ln|u|^2}{d\ln k} + 3. \tag{6.13}$$

Analytic calculations (see section 2.1 of chapter 2) give the spectral index for our potential (6.5) in the two scaling regimes. For the dark energy modes, $n_s - 1$ is $-2b^2/(2 - b^2)$, whereas it is $-4/(c^2 - 2)$ for the ekpyrotic modes. Both slopes are the same if we set $c^2 = 4/b^2$. (An interesting duality between the scalar spectrum of expanding and contracting phases is discussed in [25].) These qualitative arguments demonstrate that we would expect two nearly scale-invariant parts of the spectrum to be joined by a transition feature. We now analyze the shape of this feature.

2.2. Numerical results. We rewrite $U_{\rm pot}$ in terms of the potential and proper time, to obtain

$$U_{\rm pot} = a^2\Big(\tfrac{1}{2}\dot\phi^2 + 2V + \frac{8HV_{,\phi}}{\dot\phi} + \frac{2V_{,\phi}^2}{\dot\phi^2} + V_{,\phi\phi}\Big). \tag{6.14}$$

We use this expression when evaluating $U_{\rm pot}$ numerically. By rescaling $V(\phi)$ and τ by dimensionless variables such that $\tau^2 V(\phi)$ is fixed, the form of (6.6) is preserved and we can set V_0 to 1 in (6.5). In figure 4, we show parametric plots of $U_{\rm pot}/a^2$ against ϕ for the attractor solutions with potentials of the form (6.5) with $c = 10$. We have plotted $U_{\rm pot}$ against ϕ rather than t or τ in order to ease comparison between the models.

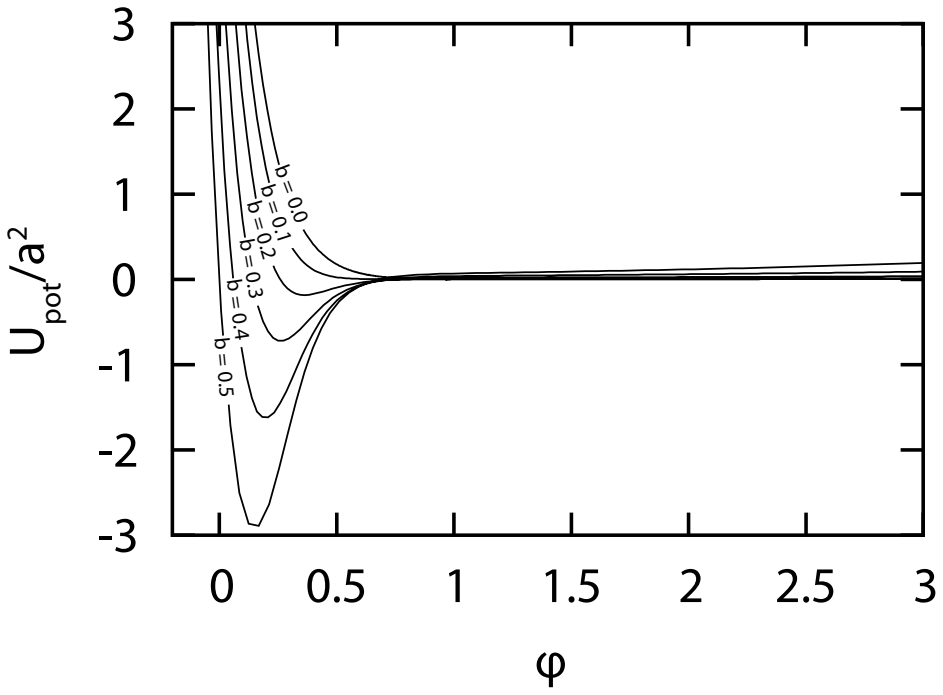

FIGURE 4. A plot of $U_{\rm pot}/a^2$ against ϕ for the attractor solutions with potentials of the form $V_0(e^{b\phi} - e^{-10\phi})$, with b ranging from 0.0 to 0.5 in steps of 0.1. In the solutions with $b \gtrsim 0.2$, $U_{\rm pot}$ takes negative values.

With $c = 10$, as in the figure, $U_{\rm pot}$ stops behaving monotonically if $b \gtrsim 0.2$. This is because the $V_{,\phi\phi}$ term in $U_{\rm pot}$ (6.14) can dominate over the other terms at the transition. This term can even make $U_{\rm pot}$ go negative. In this case, just as $(aH)^2$ instantaneously vanishes at the transition, our notion of the horizon, defined by $U_{\rm pot}$, temporarily breaks down. This has no effect on the modes that left the horizon many e-folds before the transition, as the terms in the definition of $U_{\rm pot}$ are roughly of order $a^2 V_0$ and the $k^2 \ll a^2 V_0$ term in the evolution of u_k (6.6) is still negligible. However, the feature can affect the overall amplitude of the spectrum and the shape of a moderate band of wavenumbers. When $U_{\rm pot}$ is negative (6.6) is an equation for a time-dependent harmonic oscillator. The frequency is changing rapidly, not adiabatically: $(\phi')^2 \sim a^2 V_0 \sim k_{\rm tran}^2$, where $k_{\rm tran}$ is the transition wavenumber. Since the feature in figure 4 spans a range of $\Delta\phi \sim 0.5$, the transition modes will not oscillate more than once at the transition. For such a short switch the amplitude of u_k may simply begin to decrease.

To compute the power spectrum of u, it is sufficient to know its modulus r. We solve for r directly. This is helpful, as it removes the rapidly changing phase from u. Set

$$u_k = r_k e^{i\theta_k}/(2k)^{3/2}, \tag{6.15}$$

where we have absorbed the initial condition (2.57) into r, so that $r = 1$ at early times. In these units, the spectral index is $n_s - 1 = d\ln r^2/d\ln k$, and scale invariance corresponds to r independent of k. Substituting into (6.6) and integrating the imaginary part gives $r^2\theta' = -k$. Substituting this into the real part leaves

$$r'' = k^2\left(\frac{1}{r^3} - r\right) + U_{\rm pot} r. \tag{6.16}$$

Thus, r behaves like a ball rolling in a time-dependent potential, resting at $r = 1$ in the far past where $U_{\rm pot}$ is negligible. As $U_{\rm pot}$ changes slowly from zero, but remains much less than k^2, r stays nestled at the slowly-moving minimum of its time-dependent potential. In integrating (6.16), we start suitably early that $|U_{\rm pot}| \ll k^2$, where taking $r = 1$ and $\dot{r} = 0$ suffices.

We integrate from a time when $U_{\rm pot} \ll k^2$ to a time when $U_{\rm pot} \gg k^2$, for a large range of k^2. To do this efficiently, we tabulate the solutions to the background equations of motion. Since both conformal and proper time vary over many orders of magnitude, ϕ is treated as the independent variable. We use (6.16) to integrate r over the range of ϕ between the initial and final time, and to construct a table of r versus k. This technique is quite efficient, and can be used to quickly compute the spectrum for a wide range of potentials. The background evolution needs only be solved once for each potential, and the r for each k need only be integrated over the relevant range of ϕ, when $U_{\rm pot} \sim k^2$ (otherwise r is either constant or behaves identically to the $k = 0$ mode). These techniques work quite generally, so the code could also be used to compute the spectra associated with various inflationary potentials in models without a contracting phase.

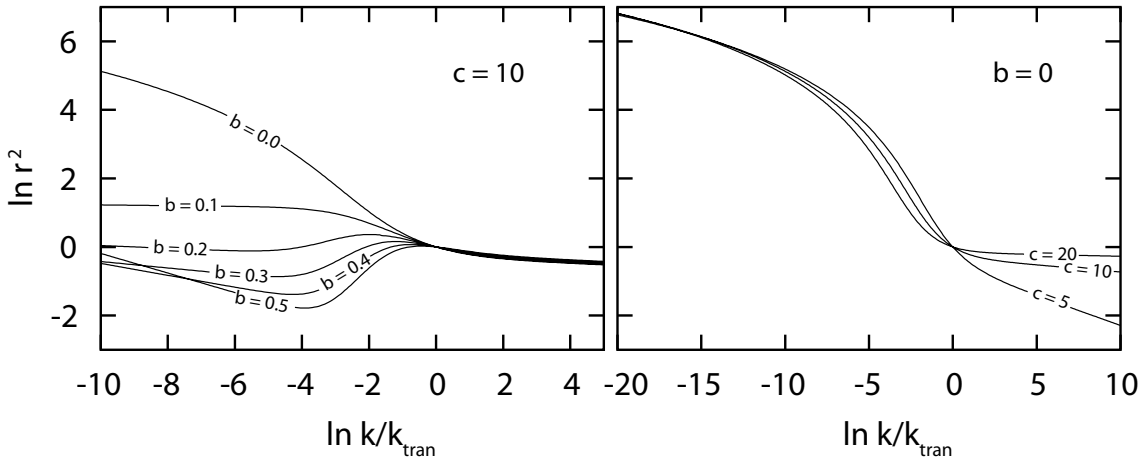

FIGURE 5. Plots of $\ln r_k^2$ (equivalently, $\ln|u_k|^2 - 3\ln k$) against $\ln k$ for potentials of the form $V_0(e^{b\phi} - e^{-c\phi})$. The background evolves according to the attractor. For definiteness, we have chosen the normalization so that the modes with $k^2 = U_{\rm pot}$ when $H = 0$ coincide for each of the solutions. The plots on the left have fixed $c = 10$ and those on the right have fixed $b = 0$. The slope gives the spectral index: a horizontal line corresponds to a scale-invariant spectrum.

We plot $\ln r^2$ against $\ln k$ in figure 2.2 for the attractor solutions with potentials of the form (6.5). The observed scale and amplitude of the fluctuations depend on other parameters, discussed in section 4. Here, we rescale the curves to put the mode with $k^2 = U_{\rm pot}(H = 0)$ at the origin. At large k all curves tend to their expected ekpyrotic limit, and similarly at small k those corresponding to $b > 0$ go to their power-law inflation limit. There is a feature in the power spectrum for those modes which leave the horizon near $H = 0$, at the transition between expansion and contraction in all models. This also corresponds to the modes that exit the horizon in the transition between the dark energy and ekpyrotic epochs. The models with larger b have a more abrupt change in V and $U_{\rm pot}$, and hence a more pronounced feature. Solutions for which $U_{\rm pot}$ goes negative have spectra which are not monotonically decreasing: as discussed above, the modes that leave the horizon near $H = 0$ have time to begin an oscillation.

In the cyclic scenario perturbations are produced in the expanding phase, when the universe is dark-energy dominated. This phase can be thought of as a phase of ultra-low energy inflation: the height of the potential is the cosmological constant, 10^{-120} in Planck units. The perturbations, however, are amplified throughout the ekpyrotic phase, when the associated energy scale is much larger. At the bounce, while the dark energy perturbations are on much larger physical scales, their magnitude is comparable to those produced during the ekpyrotic phase.

The spectral index in this model is of the sort marginally preferred by recent surveys of large scale structure [103, 167] (but see also [162]). However, we will see in the next section that these modes are far outside the horizon in the cyclic model, so that the feature is not, in general, visible.

3. Scales and the cyclic model

In order to understand how the cyclic model behaves over multiple cycles, we record some scales in terms of the fundamental parameters governing the model. These are the value of the potential at the end of the ekpyrotic phase $V_{\rm end}$ (defined to be positive) and the reheat temperature $T_{\rm RH}$, which determines how much radiation is produced at the bounce and sets the scale at which scalar field kinetic energy density and radiation density are equal. The collision velocity of the branes, $v_{\rm coll}$, is a five dimensional parameter which is determined by $V_{\rm end}$ and the condition from observations that the spectrum of perturbations have amplitude 10^{-5}. The constraints on these parameters were considered in detail in [115] and are summarized in figure 6.

The cyclic model has four phases: the dark energy phase, the ekpyrotic phase, the kinetic energy dominated phase and the radiation and matter dominated phase. We consider the change of the scale factor a and the comoving horizon aH in each of those phases. Our notation for the subscripts and results are summarized in table 1 and in the time-line figure 2. Note, in particular that we use the subscript 1 to denote quantities evaluated today, and the subscript 0 to denote quantities evaluated "today, one cycle ago."

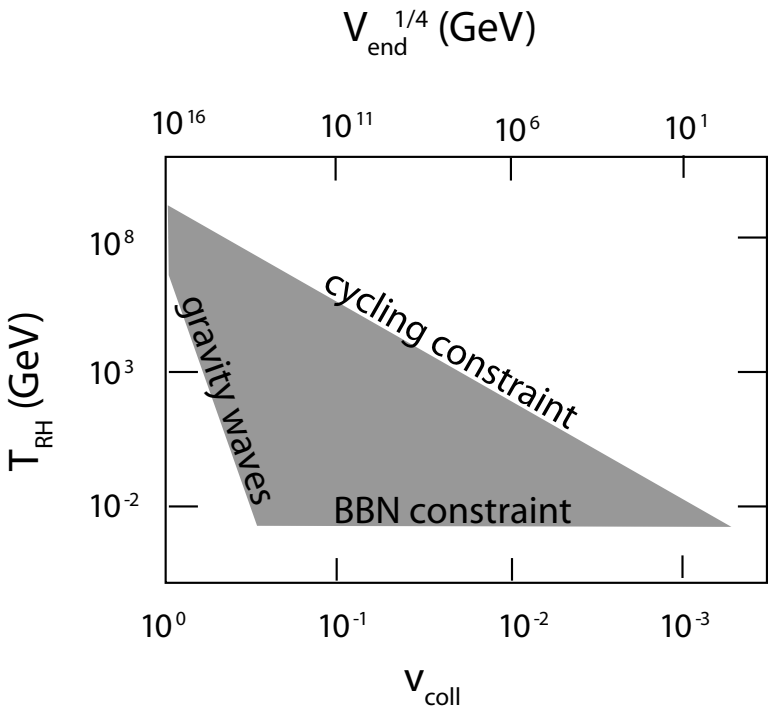

FIGURE 6. Constraints on two cyclic model parameters: the reheat temperature $T_{\rm RH}$ and $V_{\rm end}$, the value of the potential at the end of the ekpyrotic phase. The latter parameter can be written in terms of the collision velocity $v_{\rm coll}$ of the two orbifold planes. The shaded region shows the allowed parameters for fast-roll parameter $\varepsilon = 10^{-2}$ and warping scale 3×10^4 Planck units. The BBN constraint comes from the condition that the model is radiation dominated at big bang nucleosynthesis temperatures. The cycling constraint comes from the constraint that the scalar field reach the dark energy plateau before reheating. The gravity wave constraint comes from the requirement that no more than 10% of the critical density be made up by gravitational waves at BBN. Adapted with permission from [115].

We assume that there are $N_{\rm DE}$ e-folds of expansion in the dark energy phase. Since the Hubble parameter H is approximately constant, the ratio of scale factors before and after $a_{\rm DE}/a_0 \approx e^{N_{\rm DE}}$ and likewise for the horizon $(aH)_{\rm DE}/(aH)_0 \approx e^{N_{\rm DE}}$.

The ekpyrotic phase has equation of state $w \gg 1$. From (2.27) and (2.26), we see that

$$\frac{a_{\rm end}}{a_{\rm DE}} = \left(\frac{V_0}{V_{\rm end}}\right)^{1/c^2}, \tag{6.17}$$

	scale factor		expansion	horizon	example value	
epoch	initial	final	$a_\mathrm{f}/a_\mathrm{i}$	$(aH)_\mathrm{f}/(aH)_\mathrm{i}$	$\ln\frac{a_\mathrm{f}}{a_\mathrm{i}}$	$\ln\frac{(aH)_\mathrm{f}}{(aH)_\mathrm{i}}$
dark energy	a_0	a_tran	N_DE	N_DE	N_DE	N_DE
ekpyrotic	a_tran	a_end	$\left(\frac{V_\mathrm{end}}{V_0}\right)^{-1/c^2}$	$\left(\frac{V_\mathrm{end}}{V_0}\right)^{1/2}$	≈ 1	111
kinetic energy	a_end	a_RH	$\left(\frac{V_\mathrm{end}^{1/4}}{T_\mathrm{RH}}\right)^{2/3}$	$\left(\frac{V_\mathrm{end}^{1/4}}{T_\mathrm{RH}}\right)^{-4/3}$	18	-37
matter/radiation	a_RH	a_1	$\frac{T_\mathrm{RH}}{V_0^{1/4}}$	$\left(\frac{T_\mathrm{RH}}{V_0^{1/4}}\right)^{-1}$	27	-27

TABLE 1. The scale factor notation and approximate evolution of the scale factor and horizon in the cyclic model in terms of the reheat temperature T_RH, the potential at the end of the ekpyrotic phase V_end and the scale of dark energy V_0. The illustrative values at the right of the table take $T_\mathrm{RH} = 1\,\mathrm{GeV}$, $V_\mathrm{end} = 10^{12}\,\mathrm{GeV}$ and $V_0 = 10^{-3}\,\mathrm{eV}$.

which is a very small contraction, typically 1–2 e-folds. Moreover,

$$e^{N_\mathrm{ek}} = \frac{(aH)_\mathrm{end}}{(aH)_\mathrm{DE}} \approx \frac{H_\mathrm{end}}{H_\mathrm{DE}} = \left(\frac{V_\mathrm{end}}{V_0}\right)^{1/2} = 1.6\times 10^{23}\left(\frac{V_\mathrm{end}^{1/4}}{\mathrm{GeV}}\right)^2, \tag{6.18}$$

which is very large, so an exponentially large number, N_ek of modes are generated in the ekpyrotic phase.

The $w = 1$ kinetic energy dominated phase is symmetric between contracting and expanding phases for $\phi < \phi_\mathrm{end}$. However, we have seen that the ekpyrotic contraction is negligible and does not compensate the expansion that occurs when ϕ is in the interval before reheating, $\phi_\mathrm{end} < \phi < \phi_\mathrm{RH}$. In this interval, the energy density of the universe goes from approximately V_end to $\pi^2 g\, T_\mathrm{RH}^4/30$, where g is the number of relativistic species at reheating. Since $\rho \propto a^{-6}$:

$$\frac{a_\mathrm{RH}}{a_\mathrm{end}} = \left(\frac{V_\mathrm{end}^{1/4}}{T_\mathrm{RH}}\right)^{2/3}, \tag{6.19}$$

and

$$\frac{(aH)_\mathrm{RH}}{(aH)_\mathrm{end}} = \left(\frac{T_\mathrm{RH}}{V_\mathrm{end}^{1/4}}\right)^{4/3}, \tag{6.20}$$

where we have omitted the constants, which make a difference of only a few e-folds.

Finally the results for the radiation and matter dominated phase can be extracted from the ratio of the CMB temperature and the reheat temperature:

$$\frac{a_1}{a_{\rm RH}} = \frac{T_{\rm RH}}{T_0}\Big(\frac{T_{\rm eq}}{T_{\rm dec}}\Big)^{1/2} = 1.4\times 10^{13}\frac{T_{\rm RH}}{\rm GeV}, \tag{6.21}$$

where the small square root term is from the slower rate of cooling $T \propto a^{-1/2}$ between matter-radiation equality and decoupling. The corresponding result for the comoving horizon is

$$\frac{(aH)_1}{(aH)_{\rm RH}} = \Big(\frac{a_{\rm RH}}{a_1}\Big)\Big(\frac{a_1}{a_{\rm eq}}\Big)^{1/2} = 4.2\times 10^{-12}\Big(\frac{T_{\rm RH}}{\rm GeV}\Big)^{-1}, \tag{6.22}$$

where again we have corrected for the different rate of expansion in matter domination.

We have derived a complete history of the scale factor and Hubble constant over a cycle. The Hubble constant returns to the same value after each cycle, $H_1 = H_0$. The aggregate expansion in the model over the course of the cycle is then

$$N_{\rm tot} = \log\frac{(aH)_1}{(aH)_0} = \log\frac{a_1}{a_0} \approx 30 + N_{\rm DE} + \frac{2}{3}\log\frac{V_{\rm end}^{1/4}}{\rm GeV} + \frac{1}{3}\log\frac{T_{\rm RH}}{\rm GeV} \tag{6.23}$$

This is also the number of e-folds outside the horizon today the mode that was on the horizon in the last cycle is. Since the only contraction is due to the ekpyrotic phase, this net expansion is an irrevocable feature of the cyclic model. This is important, because it dilutes entropy generated each cycle. For the parameters used in table 1, $N_{\rm tot} \approx 48 + N_{\rm DE}$. Therefore if $N_{\rm DE} = 0$, then the horizon next cycle has physical scale, today, of roughly 100 km.

The transition ($H = 0$) mode produced in the last cycle has scale

$$N_{\rm tran} = -\log\frac{k_{\rm tran}}{(aH)_1} = \log\frac{(aH)_1}{(aH)_{\rm DE}} = 27 + \frac{1}{3}\log\frac{T_{\rm RH}}{\rm GeV} + \frac{2}{3}\log\frac{V_{\rm end}^{1/4}}{\rm GeV}, \tag{6.24}$$

which is much larger than the horizon today for any reasonable parameters. The number of visible e-folds of ekpyrotic modes is

$$N_{\rm vis} = \log\frac{k_{\rm end}}{(aH)_1} = -\log\frac{(aH)_1}{(aH)_{\rm end}} = 26 + \frac{4}{3}\log\frac{V_{\rm end}^{1/4}}{\rm GeV} - \frac{1}{3}\log\frac{T_{\rm RH}}{\rm GeV}, \tag{6.25}$$

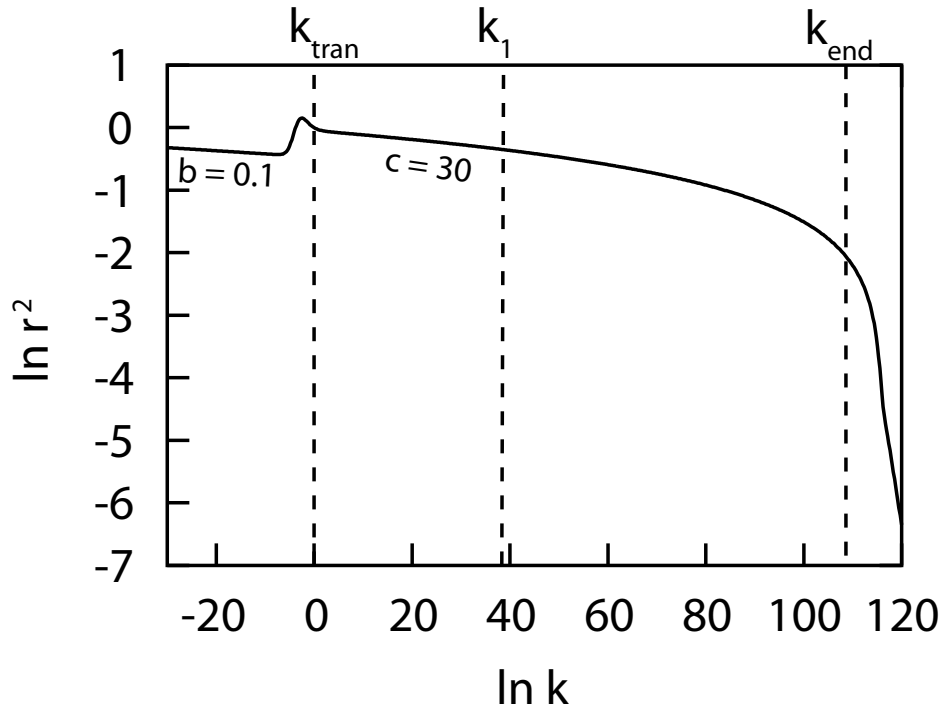

FIGURE 7. A typical spectrum of primordial perturbations for the cyclic model, with roughly 110 e-folds of ekpyrotic modes. The potential has $V_{\rm end} \approx 10^{12}\,$GeV. The mode that left the horizon when the universe moved from expansion to contraction ($H = 0$) has wavenumber labeled by $k_{\rm tran}$, where $k_{\rm tran}^2 = U_{\rm pot}(H = 0)$. The wavenumber that is on the horizon today is labeled k_1 and is calculated with $T_{\rm RH} = 1\,$GeV. All modes with $k > k_1$ are inside the horizon, including the last ekpyrotic mode to exit the horizon, when $V = -V_{\rm end}$, $k_{\rm end}$. The steep falloff for $k > k_{\rm end}$ is from modes that exit in the kinetic energy dominated phase.

This is a positive number for any allowed parameters. Thus, the cyclic model robustly predicts that the visible modes should all be ekpyrotic. The smallest scale ekpyrotic mode visible today (*i.e.* produced in the last ekpyrotic phase) has a physical wavelength of roughly 10 cm today, for the parameters in table 1.

Figure 3 shows all the modes, including those that exit the horizon in the kinetic phase. Modes continue to exit in this phase, although they are not scale-invariant. The kinetic modes go as $|u|^2 \sim k^{-4}$ and so have deeply red spectral index $n_s - 1 = -1$. In the kinetic

Mode	$k/(aH)_1$	physical scale
Horizon last cycle	$e^{-N_{\rm tot}}$	10^{24} Mpc
Largest ekpyrotic mode	$e^{-N_{\rm tran}}$	10^{24} Mpc
Horizon today	1	5000 Mpc
Largest ekpyrotic, next cycle	$e^{N_{\rm tot}-N_{\rm tran}}$	5000 Mpc
Horizon next cycle	$e^{N_{\rm tot}}$	100 km
Smallest ekpyrotic mode	$e^{N_{\rm vis}}$	10 cm
Planck mode	$e^{N_{\rm vis}+N_{\rm kin}}$	1 Å
Smallest ekpyrotic, next cycle	$e^{N_{\rm tot}+N_{\rm vis}}$	10^{-22} m

TABLE 2. Physical scales in the cyclic model evaluated today. They are computed with $V_{\rm end} = 10^{12}\,\text{GeV}$, $T_{\rm RH} = 1\,\text{GeV}$ and $N_{\rm DE} \approx 0$.

phase, $aH \sim \rho^{1/3}$, so a band of

$$N_{\rm kin} = 60 - \frac{4}{3}\log\frac{V_{\rm end}^{1/4}}{\text{GeV}} \tag{6.26}$$

modes are produced before $\rho \sim 1$ (in Planck units). Whether these modes are physical depends on the details of the bounce, and in particular what effect it has on short wavelength modes. They have scales of at most a few centimeters, which is too small to produce structure in the universe. All the physical scales are summarized in table 2.

The largest visible ekpyrotic mode – the mode on the horizon today – left the horizon in the ekpyrotic phase when $V = -V_{\rm end}e^{-2N_{\rm vis}}$. Equations (6.24) and (6.25) indicate that the mode on the horizon today was produced very roughly halfway through the ekpyrotic epoch $V(\phi) \sim -\sqrt{V_0 V_{\rm end}}$. If $V_{\rm end}^{1/4} > T_{\rm RH}$, as in most of the allowed parameter region (figure 6), then there are more visible ekpyrotic modes than ekpyrotic modes outside the horizon today. This is illustrated in figure 3.

In the ekpyrotic scenario with the potential (6.3) V_0 need not be the density of dark energy today, since it is assumed that the potential turns off after the big crunch, and dark energy has some other origin. In this case, the relation (6.24) becomes considerably more

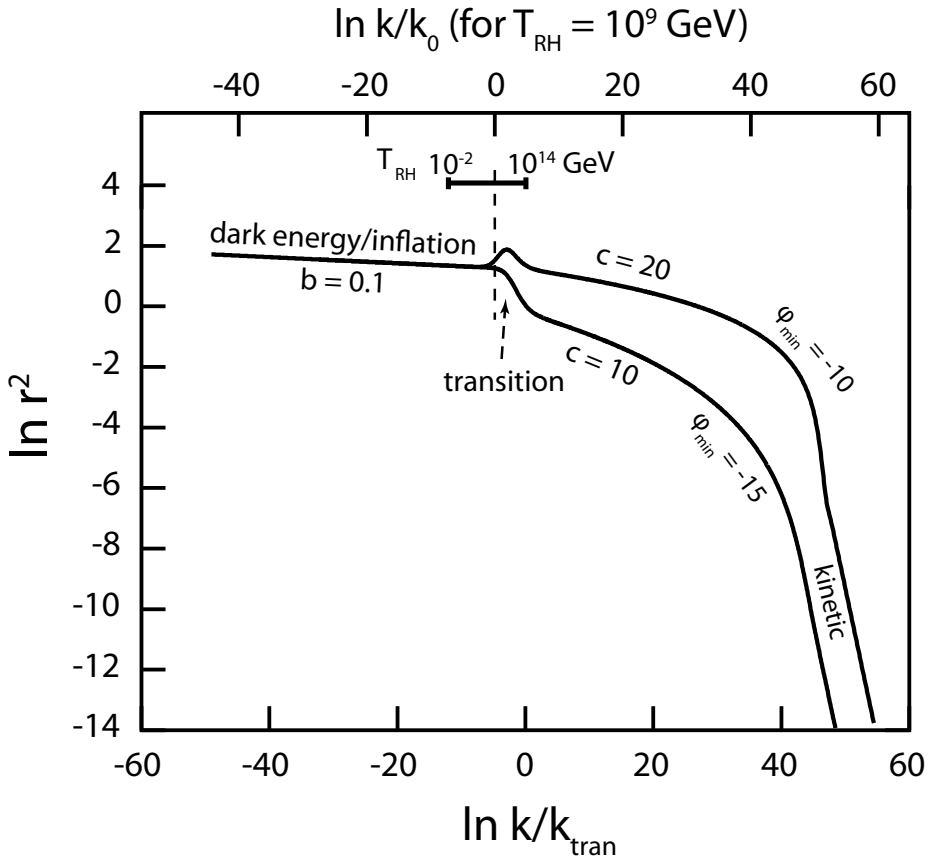

FIGURE 8. The perturbation spectrum for two potentials with cut-offs. There is a flat, inflationary spectrum for small k; a transition feature; a flat, ekpyrotic spectrum for intermediate k; and a steep falloff at large k for modes that leave the horizon in the kinetic-dominated phase. Two models are shown, both with $b = 0.1$. The model with the "bump" feature has $c = 20$, $V_0 = 10^2\,\mathrm{GeV}^4$, $V_{\mathrm{end}} = 10^{45}\,\mathrm{GeV}^4$, $\phi_{\mathrm{end}} = -10$ and $N_{\mathrm{ek}} = 50$; the model with the "step" feature has $c = 10$, $V_0 = 10^{-3}\,\mathrm{GeV}^4$, $V_{\mathrm{end}} = 10^{29}\,\mathrm{GeV}^4$, $\phi_{\mathrm{end}} = -15$ and $N_{\mathrm{ek}} = 37$. The two models are normalized so the dark energy modes have equal amplitude, and $k = k_{\mathrm{tran}}$ corresponds to the mode that leaves the horizon when $H = 0$. The scale of a mode of these particular models relative to the horizon today is shown on the top axis. This is for a reheat temperature of $10^9\,\mathrm{GeV}$; how the axis would shift by varying the reheat temperature within the allowed range of 10^{-2}–$10^{14}\,\mathrm{GeV}$ is indicated with a horizontal bar.

flexible, and it is possible to have the transition mode inside the horizon provided the condition

$$-\log\frac{k_{\rm tran}}{(aH)_1} = -26 + \frac{1}{3}\log\frac{T_{\rm RH}}{\rm GeV} + \frac{2}{3}\log\frac{V_{\rm end}^{1/4}}{\rm GeV} - 2\log\frac{V_0^{1/4}}{\rm GeV} < 0 \tag{6.27}$$

is satisfied. Adjusting V_0 is equivalent to changing the scale of the transition from dark energy epoch to the ekpyrotic epoch, which moves the transition feature to higher k. Figure 3 illustrates an example of such a mixed model. There is no cycling constraint (the upper line in figure 6) for this model. However, in order to avoid forming GUT monopoles at reheating, we require $T_{\rm RH} \lesssim 10^{16}\,{\rm GeV}$. While this model has a natural feature in the spectrum, which may be marginally suggested by observations, such a model would need to be fine tuned for the feature to be visible at the correct wavenumber.

In the following chapter, we consider the effect of the ekpyrotic amplification has on the global structure of the model.

4. The global structure of the cyclic model

We now turn to the global structure of the cyclic model. In the previous section, we considered the scales of various modes. Here, we study how the model behaves over multiple cycles: how the scales relate, and how the amplitude of the modes behave. First, we will show that there is no constraint on the number of e-folds of dark energy domination required by the cyclic model. The suppression of dark energy fluctuations inside the horizon and the attractor behavior of the contracting phase – discussed, respectively, in chapters 3 and 5 – prevent fluctuations from one cycle from contaminating the next.

Next, we discuss the global structure of the model, the structure on scales much greater than the horizon. In particular, as modes pass outside the horizon for the last time, their amplitudes diverge. We show that this has no physical ramifications for the model, and merely suggests that a fixed gauge choice over such huge distances makes no sense over multiple cycles.

In this section, we again use the Newtonian potential, which is related to the variable u by $u = \Phi/\dot{\phi}$. The superhorizon solutions (6.8) and (6.9) then become

$$\Phi_1(t) = \frac{H}{a}, \tag{6.28}$$

$$\Phi_2(t) = \frac{H}{a}\int_0^t dt'\, a(t')\big(w(t') + 1\big), \tag{6.29}$$

where w is the equation of state. Again, the integral mode (6.29) is the dominant mode in the dark energy epoch. It is constant (for the scaling solution (2.25)), whereas the mode (6.28) decays rapidly. This is consistent with the inflationary expectation that Φ is conserved outside the horizon.

The ekpyrotic growing mode, (6.28), grows as $1/|t|$ where $t \to 0^-$, from (2.27), whereas (6.29) is constant. As with the u variables, the dark energy growing mode matches onto a linear combination of the ekpyrotic growing and decaying modes at the reversal from expansion to contraction.

4.1. The dark energy epoch. In [171] it was pointed out that sixty e-folds of dark energy domination were sufficient to reduce the visible universe to less than one particle per Hubble volume, and so would prevent the possibility of physical perturbations in our universe from disturbing the quantum mechanical generation of fluctuations for the next cycle. This is a very conservative limit on the number of e-folds required, since we are concerned about the horizon $(aH)_1$ which has a much smaller scale than the horizon a cycle previously $(aH)_0$.

Moreover, in chapter 5 we saw that the $w \gg 1$ ekpyrotic solution was an attractor. It satisfies a no-hair theorem that says that, since matter and radiation blueshift much slower than the scalar field in the ekpyrotic phase, they make a negligible contribution to the dynamics well into this phase. In voids, the dark energy density is already higher than in dense regions. However, after the reversal to contraction, the horizon will shrink and the increasing scalar field energy will rapidly suppress matter effects in even highly non-linear regions,

such as galaxies. Therefore, it is only scalar field perturbations – not matter perturbations – that contribute deep in the ekpyrotic phase, as they are the perturbations which grow.

If dark energy fluctuations have a sound speed $c_s^2 > 0$ then they are suppressed by a factor $\langle \delta_\phi^2 \rangle \propto k^{-6.6}$ inside the horizon (see chapter 3, in particular, figure 2). This is robust, and holds for any $w < -1/3$ and $c_s^2 > 0$. Reducing the sound speed merely changes the scale, relative to the horizon, at which this rapid falloff begins. Since the Newtonian potential is related to the density contrast by the Poisson equation (2.40) and (2.67), this gives $\langle \Phi_\phi^2 \rangle \sim k^{-10.6}$ today, where the subscript indicates we are only considering contributions from the scalar field perturbation $\delta\phi$. Since the mode on the horizon today has amplitude 2×10^{-5},

$$\langle \Phi_\phi^2 \rangle \approx 4 \times 10^{-10} \Big(\frac{k}{(aH)_0} \Big)^{-10.6}, \tag{6.30}$$

where $(aH)_0$ is the horizon scale. This is the amplitude today. This mode will receive some amplification in the ekpyrotic phase. The ekpyrotic amplification depends on when it exits the horizon. From (6.28), a mode is amplified by a factor of $k/k_{\rm end}$, where $k_{\rm end}$, which receives no amplification, is the last mode to leave at the end of the ekpyrotic phase. Thus (6.30) becomes

$$\langle \Phi^2 \rangle_{\phi=\phi_{\rm end}} \approx 4 \times 10^{-10} \frac{V_{\rm end}}{V_0} \Big(\frac{k}{(aH)_0} \Big)^{-8.6}, \tag{6.31}$$

using (6.18). The matter perturbations are negligible at the end of the ekpyrotic epoch. This spectrum is deeply red, so the largest mode will be that which is on the horizon today, which has $k = (aH)_1$. Thus, we require

$$4 \times 10^{-10} \frac{V_{\rm end}}{V_0} \Big(\frac{(aH)_1}{(aH)_0} \Big)^{-8.6} \approx 10^{-75} \Big(\frac{V_{\rm end}^{1/4}}{\rm GeV} \Big)^{-1.7} \Big(\frac{T_{\rm RH}}{\rm GeV} \Big)^{-2.9} \ll 4 \times 10^{-10}, \tag{6.32}$$

from (6.23) with $N_{\rm DE} = 0$. The right hand side of the inequality comes from the contribution of the ekpyrotic modes visible today. This neglects the contribution of the amplitude factor at the bounce, which changes the constraint on the right hand side. However, since the amplitude of perturbations is reduced in the bounce studied by Tolley *et al.* [115, 178], this will only increase the right hand side of (6.32) and weaken the constraint. However,

the inequality is easily satisfied for all the parameters allowed in figure 6, and thus there is no circumstance in which dark energy domination is necessary to separate the ekpyrotic perturbations from one cycle from the perturbations of the next cycle.

4.2. The horizon. Each cycle involves an exponentially large range of modes. We analyze what happens in terms of the comoving wavenumber k. Recalling the notation of section 3, over the course of a cycle, the horizon shifts to larger values of k by a large number $N_{\rm tot}$ of e-folds (6.23). In a cycle, a band of $N_{\rm DE} + N_{\rm ek}$ e-folds of nearly scale-invariant modes are generated by the inflationary and ekpyrotic mechanisms. These modes pass outside the horizon. However, $N_{\rm tot}$ e-folds reenter, of which $N_{\rm vis}$ are nearly scale-invariant ekpyrotic modes. The remaining modes are part of the red spectrum of kinetic modes.

From (6.23) and (6.18) it can be seen that the total number of e-folds of nearly scale-invariant fluctuations satisfies an inequality:

$$N_{\rm ek} + N_{\rm DE} \lesssim 2N_{\rm tot}. \tag{6.33}$$

If $N_{\rm DE}$ is comparable to $N_{\rm ek}$, this inequality is easily satisfied and a mode will exit the horizon at most twice. Some modes will exit the horizon first in the ekpyrotic phase, reenter the radiation dominated universe and exit again in the dark energy phase, after which it will never again reenter the horizon. Other modes, which exit near the transition from dark energy domination to radiation domination will never be visible in the radiation dominated epoch, and will merely exit the horizon once. This is illustrated in figure 9.

The situation is different if the universe has a very small number of e-folds of dark energy domination, $N_{\rm DE} = 0$. In this case,

$$N_{\rm ek} - 2N_{\rm tot} \approx \frac{2}{3}\log\frac{V_{\rm end}^{1/4}}{T_{\rm RH}}, \tag{6.34}$$

and some modes will exit the horizon *three* times, in the first cycle as small scale modes produced at the end of the ekpyrotic phase, when $V \approx -V_{\rm end}$e; in the second cycle as modes produced roughly halfway down the ekpyrotic potential, visible today near but just inside

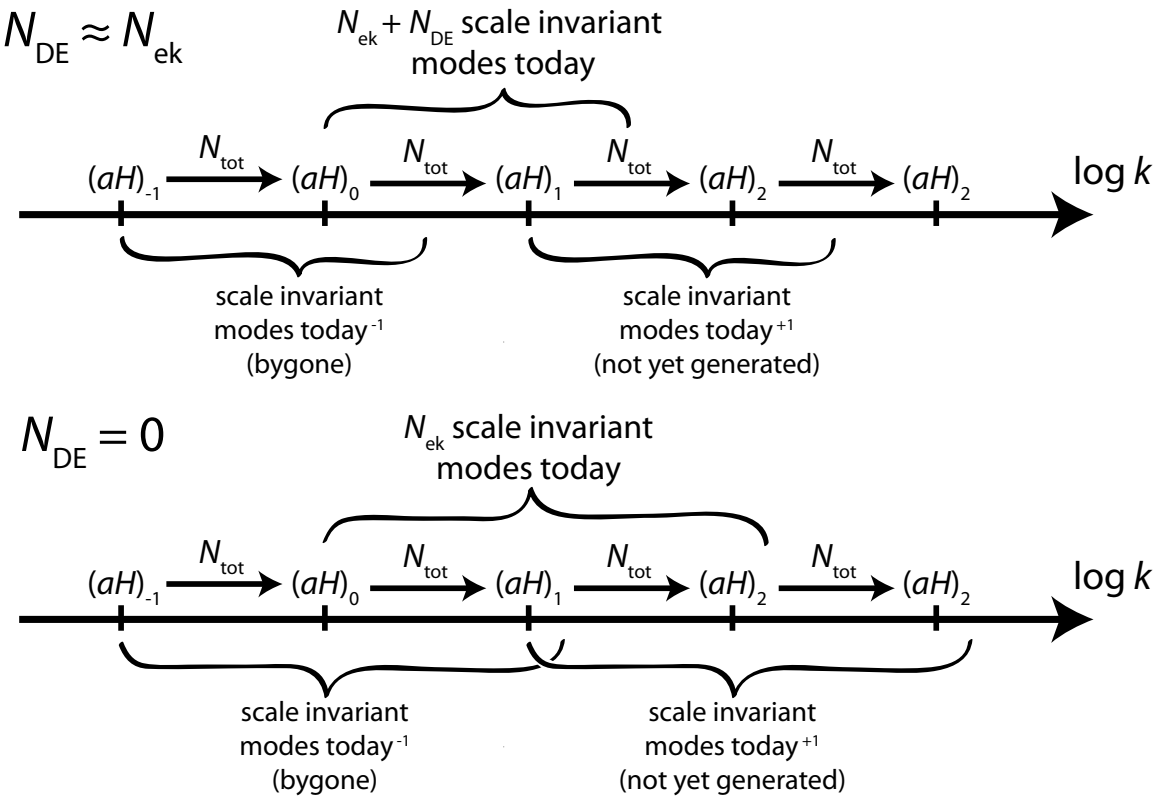

FIGURE 9. How the comoving wavenumbers evolve over a cycle. In each cycle, the comoving horizon shifts to the left by $N_{\rm tot}$ e-folds, because the scale factor increases and H returns to the same value. Over each cycle, a band of $N_{\rm ek} + N_{\rm DE}$ e-folds of nearly scale-invariant perturbations is also produced. If $N_{\rm DE}$ is small, the same comoving wavenumber will leave the horizon up to three times. However, if $N_{\rm DE}$ is large, this will only happen twice before the mode is permanently on superhorizon scales.

the horizon; and finally, as modes which exit near the top of the potential and are never again reenter. We saw that these scalar field fluctuations are heavily suppressed in the dark energy phase, and so the modes of one cycle are not affected by the previous cycle: each time they leave the horizon, except perhaps the last, they form new, independent Gaussian random perturbations for the big bang of the next cycle. Moreover, (6.34) is small compared to the total number of modes generated. Thus most modes exit the horizon only twice (and reenter only once). This situation is also illustrated in figure 9.

4.3. Beyond the horizon. As modes move beyond the horizon for the last time, they continue expanding to ever larger scales, as their physical wavelength grows by $e^{N_{\rm tot}}$ each cycle. They evolve in concert, according to the exact solutions for the superhorizon behavior (6.28) and (6.29). While the dominant mode in the expanding phase, the integral mode (6.28), is roughly constant, the modes receive a contribution from the dominant mode in the contracting phase (6.29). This mode grows by the large factor $\sqrt{V_{\rm end}/V_0}$ each cycle, which is between 10^{25} and 10^{55} (from figure 6). When cyclic modes reenter the horizon, they do not receive a net amplification from cycle to cycle, as pressure effects rapidly smooth out inhomogeneities. However, when modes become acausal, they receive a huge amplification each cycle. Since the modes have amplitude $\Phi \sim 10^{-5}$, they will be non-linear only one cycle after last exiting.

Such a huge amplification factor is necessary because the scale of quantum perturbations is set by the Hubble constant. The amplitude of scalar field fluctuations, before they exit the horizon, is of order the Hubble parameter, or roughly $\delta\phi \sim H \sim \sqrt{V}$. Likewise $\Phi \sim \sqrt{V}$. In the inflationary mechanism, the amplitude is therefore set by the potential. In the ekpyrotic mechanism, however, the potential begins at the tiny dark energy scale, 10^{-120} in Planck units. Since the spectrum is nearly scale-invariant, this means that the dark energy modes must be amplified by 10^{55} once the exit the horizon.

This can be interpreted as the cyclic model becoming very inhomogeneous on superhorizon scales. If two disjoint Hubble patches reach the ekpyrotic epoch at slightly different times, the time delay will be amplified by a huge factor. However, with superhorizon modes there is no causal mechanism to suppress the growth of inhomogeneities in the model so causally disconnected Hubble patches become more and more asynchronous. Thus, on metaphysical scales, the cyclic model can be interpreted as a series of uncorrelated patches, expanding and fragmenting into yet more asynchronous Hubble patches.

It is clear that these large superhorizon modes are irrelevant for observations. Physically, these modes do not affect the conditions inside our horizon because the physical curvature

goes as

$$\left(\frac{k}{a}\right)^2 \Phi \propto e^{-2M(N_{\rm tot}+N_{\rm ek})}, \tag{6.35}$$

where M is the number of cycles. Since $2N_{\rm tot} \gtrsim N_{\rm ek}$, the contribution of a mode to the physical curvature perturbation diminishes. Thus, this is a purely metaphysical issue.

Conclusion

This year, 2005, is the centenary of Einstein's *annus mirabilis*. Much as in 1905, we have an enormously successful physical theory – the standard model of particle physics – that is beginning to show some cracks. Although the difficulty of reconciling gravity with the standard model is notorious, we cannot conduct experiments to investigate this directly. It is auspicious, then, that there are new cosmological observations – dark matter and dark energy – that indicate that our understanding of gravity, or of particle physics, is incomplete. Moreover, there are now two competing, testable solutions of the initial condition problem of the hot big bang, the inflationary and ekpyrotic scenarios. Future experimental tests of gravity, cosmic microwave background polarization measurements and particle physics at the Large Hadron Collider promise to improve this situation.

Perhaps the most exciting development of recent years has been the 1998 discovery of dark energy [150, 158]. Nothing is presently known about dark energy except that it is very homogeneous, has an equation of state near -1 and unaccountably makes up roughly seven tenths of the critical density of the universe. A potential explanation is Einstein's cosmological constant. The field theory values predicted for this constant are around 120 orders of magnitude too large. It has recently been proposed that string theory contains a very large "discretuum" of metastable vacua corresponding to different choices of form fluxes on a Calabi-Yau compactification [105]. If the discretuum is large enough (*i.e.* there are more than 10^{120} vacua), as seems likely, then some of the vacua have a vacuum density of the right order to account for dark energy. Anthropic arguments [23, 190] then suggest that we live in such a universe (figure 1).

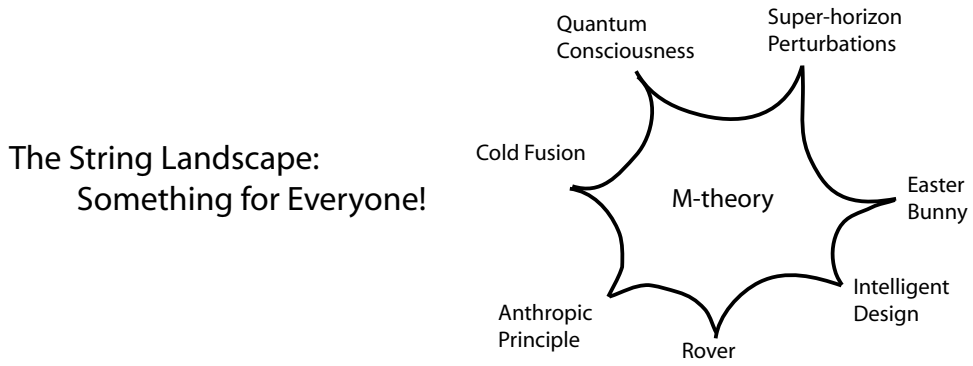

FIGURE 1. The multitude of vacua in M-theory provide sufficient freedom to explain all known physics as well as many exciting extensions.

Such an explanation is an abstract possibility. Apart from the metaphysical issues raised by the anthropic principle, however, it is probably premature to adopt such a position. If dark energy is dynamical, then it could be a rich source of new physical observations which help us understand physics beyond the standard model. In this thesis, we have explored some of the possibilities afforded by dynamical extra dimensions.

Jordan, Einstein and Bergmann [77, 104] were the first to take the Kaluza-Klein scalar – the radius of the extra dimension – seriously: previous authors fixed the size of the extra dimension. Fierz, in 1956, pointed out that a time-varying extra dimension would lead to unacceptably large violations of the equivalence principle. This led to the emphasis on "metric theories" of gravity, such as Brans-Dicke theory [28] in which matter is minimally coupled to the gravitational degrees of freedom, by the determinant of the metric $\sqrt{-g}$.

The modern perspective has changed. Tests of the equivalence principle have improved to the point where post-Newtonian tests of gravity [19, 165, 194] are of comparable precision to tests of the universality of free-fall [26, 160, 174] and constraints on the variation of the fundamental constants [56, 166, 183]. Surprisingly, these measurements do not rule out observable equivalence principle violations in the universe. In string theory, there must be corrections to general relativity at some level. In chapter 4, we described a number of realistic scenarios in which violations can occur with magnitudes near current observational limits. These ideas will be tested by future experiments, including improved metric tests of gravity

from lunar and potentially martian ranging [181], satellite tests of the universality of free-fall [141, 179, 197], quasar tests of the variation of the fine-structure constant [42, 123, 140, 168], sub-millimeter fifth-force searches [97] and space-based Cavendish experiments [4]. Each of these experiments provides a potential window into physics beyond the standard model.

An intriguing possibility to consider is if dark *matter* violates the equivalence principle [83]. In brane models, such as the S^1/Z_2 orbifold, it is possible to have visible matter on one brane and dark matter on the other. In this scenario, they would chiefly interact gravitationally. In the discussion of the Randall-Sundrum model in section 2.2 of chapter 4, we pointed out that gravity on the two branes of the model behaves very differently if the radion is not stabilized. Gravity on the positive tension brane approaches general relativity in a highly warped background, whereas gravity on the negative tension brane approaches Brans-Dicke theory with parameter $\omega = -3/2$. Instead of viewing the Randall-Sundrum geometry as a solution to the hierarchy problem [156] in which matter is on a negative tension brane, if visible matter is on the positive tension brane, then we do not see violations of the equivalence principle because the warp factor causes gravity on the to behave very much like general relativity, as in the second Randall-Sundrum scenario [155]. In this scenario, dark matter on the negative tension brane will generally have large equivalence principle violations.

In the early universe, the cosmic microwave background is another window on new fundamental physics. Chapter 3 demonstrated that under certain conditions cosmic microwave background measurements can discriminate the presence of dark energy perturbations. The B-mode polarization of the cosmic microwave background allows primordial gravitational waves to be measured [50, 164]. These observations allow inflation to be distinguished from the cyclic and ekpyrotic scenarios discussed in chapter 6 [25]. Moreover, if inflation is the correct theory, they determine its energy scale.

Many fundamental questions in string cosmology have yet to be resolved. Chapters 5 and 6 touched on some of these. In chapter 6 we discussed the global structure of the cyclic

model, and showed that dark energy domination is not necessary for consistency. On vast, superhorizon scales the model was shown to become ever more asynchronous, as slight differences between adjacent Hubble patches were exponentially amplified each cycle. Since two asynchronous patches are never again in causal contact, this does not create any instability inside the horizon. An important issue for the cyclic model is the origin of the steep, non-perturbative potential that gives rise to the ekpyrotic phase. Little is known about the interaction of branes in string [12, 135] or M-theory [180].

Chapter 5 showed that the ekpyrotic attractor is instrumental in the cyclic model, not only for generating a spectrum of scale-invariant perturbations, but also for ensuring that the homogeneity and isotropy of the universe are preserved as the crunch is approached. When the cyclic model potential vanishes near the brane collision, the large suppression of curvature and anisotropy in the ekpyrotic phase may be sufficient to suppress chaos until the Planck epoch. This seems particularly likely given the intriguing fact that heterotic M-theory is only critically chaotic for $w = 1$, so chaotic contributions to curvature and anisotropy grow much more slowly than they do in supercritical theories. This idea is discussed in detail by Wesley *et al.* [191].

What happens when the Planck epoch is reached? Toy models of quantum gravity – such as the homogeneous minisuperspace approximations – have been used to scrutinize the quantum mechanical behavior of the big crunch, starting with the work of Misner [137]. However, string theory in time dependent backgrounds is still quite poorly understood. Although some progress has been made, there is still much work to be done understanding gravitational singularities and collisions of orbifold planes in string theory [47, 48, 176–178, 180]. The resolution of these problems will be important for understanding how advances in particle theory relate to cosmology.

Bibliography

[1] L. R. Abramo, F. Finelli and T. S. Pereira, Constraining Born-Infeld models of dark energy with CMB anisotropies, *Phys. Rev.* **D70**, 063517 (2004), [astro-ph/0405041].

[2] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables* (Dover, New York, 1964).

[3] F. S. Accetta, L. M. Krauss and P. Romanelli, New limits on the variability of G from big bang nucleosynthesis, *Phys. Lett.* **B248**, 146 (1990).

[4] A. D. Alekseev *et al.*, SEE project for testing gravity in space: current status and new estimates, *Grav. Cosmol.* **5**, 67 (1999), [gr-qc/0002088].

[5] L. Andersson and A. D. Rendall, Quiescent cosmological singularities, *Commun. Math. Phys.* **218**, 479 (2001), [gr-qc/0001047].

[6] L. Anguelova and D. Vaman, R^4 corrections to heterotic M-theory, hep-th/0506191.

[7] T. Appelquist, A. Chodos and P. G. O. Freund (eds.), *Modern Kaluza-Klein theories* (Addison-Wesley, Reading, Massachusetts, 1987).

[8] C. Armendáriz-Picón, T. Damour and V. Mukhanov, k-inflation, *Phys. Lett.* **B458**, 209 (1999), [hep-th/9904075].

[9] C. Armendáriz-Picón and E. A. Lim, Haloes of k-Essence, astro-ph/0505207.

[10] C. Armendáriz-Picón, V. Mukhanov and P. J. Steinhardt, A dynamical solution to the problem of a small cosmological constant and late-time cosmic acceleration, *Phys. Rev. Lett.* **85**, 4438 (2000), [astro-ph/0004134].

[11] C. Armendáriz-Picón, V. Mukhanov and P. J. Steinhardt, Essentials of k-essence, *Phys. Rev.* **D63**, 103510 (2001), [astro-ph/0006373].

[12] C. Bachas, D-brane dynamics, *Phys. Lett.* **B374**, 37 (1996), [hep-th/9511043].

[13] T. Banks and M. Dine, Couplings and scales in strongly coupled heterotic string theory, *Nucl. Phys.* **B479**, 173 (1996), [hep-th/9605136].

[14] T. J. Battefeld, S. P. Patil and R. Brandenberger, Perturbations in a bouncing brane model, *Phys. Rev.* **D70**, 066006 (2004), [hep-th/0401010].

[15] R. Bean and O. Doré, Probing dark energy perturbations: the dark energy equation of state and speed of sound as measured by WMAP, *Phys. Rev.* **D69**, 083503 (2004), [astro-ph/0307100].

[16] V. A. Belinsky and I. M. Khalatnikov, Effects of scalar and vector fields on the nature of the cosmological singularity, *Sov. Phys. JETP* **36**, 591 (1973).

[17] V. A. Belinsky, I. M. Khalatnikov and E. M. Lifshitz, Oscillatory approach to a singular point in the relativistic cosmology, *Adv. Phys.* **19**, 525 (1970).

[18] P. G. Bergmann, Comments on the scalar-tensor theory, *Int. J. Theor. Phys.* **1**, 25 (1968).

[19] B. Bertotti, L. Iess and P. Tortora, A test of general relativity using radio links with the Cassini spacecraft, *Nature* **425**, 374 (2003).

[20] N. D. Birrell and P. C. W. Davies, *Quantum fields in curved space* (Cambridge University Press, Cambridge, UK, 1982).

[21] M. Born and L. Infeld, Foundations of the new field theory, *Proc. Roy. Soc. Lond.* **A144**, 425 (1934).

[22] W. Boucher, Positive energy without supersymmetry, *Nucl. Phys.* **B242**, 282 (1984).

[23] R. Bousso and J. Polchinski, Quantization of four-form fluxes and dynamical neutralization of the cosmological constant, *JHEP* **06**, 006 (2000), [hep-th/0004134].

[24] L. A. Boyle, P. J. Steinhardt and N. Turok, The cosmic gravitational wave background in a cyclic universe, *Phys. Rev.* **D69**, 127302 (2004), [hep-th/0307170].

[25] L. A. Boyle, P. J. Steinhardt and N. Turok, A new duality relating density perturbations in expanding and contracting Friedmann cosmologies, *Phys. Rev.* **D70**, 023504 (2004), [hep-th/0403026].

[26] V. B. Braginsky and V. I. Panov, Verification of the equivalence of inertial and gravitational mass, *Sov. Phys. JETP* **34**, 463 (1972).

[27] R. Brandenberger and F. Finelli, On the spectrum of fluctuations in an effective field theory of the ekpyrotic universe, *JHEP* **11**, 056 (2001), [hep-th/0109004].

[28] C. Brans and R. H. Dicke, Mach's principle and a relativistic theory of gravitation, *Phys. Rev.* **124**, 925 (1961).

[29] C. H. Brans, The roots of scalar-tensor theory: an approximate history, gr-qc/0506063.

[30] V. Braun, Y.-H. He, B. A. Ovrut and T. Pantev, A heterotic standard model, *Phys. Lett.* **B618**, 252 (2005), [hep-th/0501070].

[31] V. Braun, Y.-H. He, B. A. Ovrut and T. Pantev, A standard model from the $E(8) \times E(8)$ heterotic superstring, *JHEP* **06**, 039 (2005), [hep-th/0502155].

[32] P. Brax and C. van de Bruck, Cosmology and brane worlds: A review, *Class. Quant. Grav.* **20**, R201 (2003), [hep-th/0303095].

[33] P. Brax, C. van de Bruck, A.-C. Davis, J. Khoury and A. Weltman, Detecting dark energy in orbit: the cosmological chameleon, *Phys. Rev.* **D70**, 123518 (2004), [astro-ph/0408415].

[34] P. Brax, C. van de Bruck, A. C. Davis and C. S. Rhodes, Cosmological evolution of brane world moduli, *Phys. Rev.* **D67**, 023512 (2003), [hep-th/0209158].

[35] P. Brax, C. van de Bruck, A. C. Davis and C. S. Rhodes, Varying constants in brane world scenarios, *Astrophys. Space Sci.* **283**, 627 (2003), [hep-ph/0210057].

[36] M. Bucher, A. S. Goldhaber and N. Turok, An open universe from inflation, *Phys. Rev.* **D52**, 3314 (1995), [hep-ph/9411206].

[37] R. R. Caldwell, A phantom menace?, *Phys. Lett.* **B545**, 23 (2002), [astro-ph/9908168].

[38] R. R. Caldwell, R. Dave and P. J. Steinhardt, Cosmological imprint of an energy component with general equation-of-state, *Phys. Rev. Lett.* **80**, 1582 (1998), [astro-ph/9708069].

[39] R. R. Caldwell and M. Doran, Dark-energy evolution across the cosmological-constant boundary, astro-ph/0501104.

[40] S. M. Carroll, Quintessence and the rest of the world, *Phys. Rev. Lett.* **81**, 3067 (1998), [astro-ph/9806099].

[41] C. Cartier, R. Durrer and E. J. Copeland, Cosmological perturbations and the transition from contraction to expansion, *Phys. Rev.* **D67**, 103517 (2003), [hep-th/0301198].

[42] H. Chand, R. Srianand, P. Petitjean and B. Aracil, Probing the cosmological variation of the fine-structure constant: Results based on VLT-UVES sample, *Astron. Astrophys.* **417**, 853 (2004), [astro-ph/0401094].

[43] T. Chiba, N. Sugiyama and T. Nakamura, Observational tests of x-matter models, *Mon. Not. Roy. Astron. Soc.* **301**, 72 (1998), [astro-ph/9806332].

[44] A. Coley, No chaos in brane-world cosmology, *Class. Quant. Grav.* **19**, L45 (2002), [hep-th/0110117].

[45] A. A. Coley, Dynamics of brane-world cosmological models, *Phys. Rev.* **D66**, 023512 (2002), [hep-th/0110049].

[46] A. A. Coley and W. C. Lim, Asymptotic analysis of spatially inhomogeneous stiff and ultra-stiff cosmologies, gr-qc/0506097.

[47] L. Cornalba and M. S. Costa, Time-dependent orbifolds and string cosmology, *Fortsch. Phys.* **52**, 145 (2004), [hep-th/0310099].

[48] L. Cornalba, M. S. Costa and C. Kounnas, A resolution of the cosmological singularity with orientifolds, *Nucl. Phys.* **B637**, 378 (2002), [hep-th/0204261].

[49] P. Creminelli, A. Nicolis and M. Zaldarriaga, Perturbations in bouncing cosmologies: dynamical attractor vs scale invariance, *Phys. Rev.* **D71**, 063505 (2005), [hep-th/0411270].

[50] R. Crittenden, J. R. Bond, R. L. Davis, G. Efstathiou and P. J. Steinhardt, The imprint of gravitational waves on the cosmic microwave background, *Phys. Rev. Lett.* **71**, 324 (1993), [astro-ph/9303014].

[51] C. Csáki, M. Graesser, L. Randall and J. Terning, Cosmology of brane models with radion stabilization, *Phys. Rev.* **D62**, 045015 (2000), [hep-ph/9911406].

[52] G. Curio and A. Krause, Four-flux and warped heterotic M-theory compactifications, *Nucl. Phys.* **B602**, 172 (2001), [hep-th/0012152].

[53] T. Damour, Testing the equivalence principle: Why and how?, *Class. Quant. Grav.* **13**, A33 (1996), [gr-qc/9606080].

[54] T. Damour, Questioning the equivalence principle, gr-qc/0109063.

[55] T. Damour, String theory, cosmology and varying constants, *Astrophys. Space Sci.* **283**, 445 (2003), [gr-qc/0210059].

[56] T. Damour and F. Dyson, The Oklo bound on the time variation of the fine-structure constant revisited, *Nucl. Phys.* **B480**, 37 (1996), [hep-ph/9606486].

[57] T. Damour and M. Henneaux, Chaos in superstring cosmology, *Gen. Rel. Grav.* **32**, 2339 (2000).

[58] T. Damour and M. Henneaux, $E(10)$, $BE(10)$ and arithmetical chaos in superstring cosmology, *Phys. Rev. Lett.* **86**, 4749 (2001), [hep-th/0012172].

[59] T. Damour, M. Henneaux and H. Nicolai, Cosmological billiards, *Class. Quant. Grav.* **20**, R145 (2003), [hep-th/0212256].

[60] T. Damour, M. Henneaux, A. D. Rendall and M. Weaver, Kasner-like behaviour for subcritical Einstein-matter systems, *Annales Henri Poincare* **3**, 1049 (2002), [gr-qc/0202069].

[61] T. Damour, F. Piazza and G. Veneziano, Runaway dilaton and equivalence principle violations, *Phys. Rev. Lett.* **89**, 081601 (2002), [gr-qc/0204094].

[62] T. Damour, F. Piazza and G. Veneziano, Violations of the equivalence principle in a dilaton- runaway scenario, *Phys. Rev.* **D66**, 046007 (2002), [hep-th/0205111].

[63] T. Damour and A. M. Polyakov, The string dilaton and a least coupling principle, *Nucl. Phys.* **B423**, 532 (1994), [hep-th/9401069].

[64] S. DeDeo, R. R. Caldwell and P. J. Steinhardt, Effects of the sound speed of quintessence on the microwave background and large scale structure, *Phys. Rev.* **D67**, 103509 (2003), [astro-ph/0301284].

[65] J. Demaret, M. Henneaux and P. Spindel, Nonoscillatory behavior in vacuum Kaluza-Klein cosmologies, *Phys. Lett.* **B164**, 27 (1985).

[66] R. H. Dicke, New research on old gravitation, *Science* **129**, 3349 (1959).

[67] R. H. Dicke, Mach's principle and equivalence, in *Evidence for gravitational theories: proceedings of the international school of physics "Enrico Fermi"*, edited by C. Möller (Academic Press, New York, 1962).

[68] P. A. M. Dirac, An Extensible model of the electron, *Proc. Roy. Soc. Lond.* **A268**, 57 (1962).

[69] S. Dodelson, *Modern cosmology* (Academic Press, San Diego, 2003).

[70] R. Donagi, Y.-H. He, B. A. Ovrut and R. Reinbacher, The spectra of heterotic standard model vacua, *JHEP* **06**, 070 (2005), [hep-th/0411156].

[71] M. J. Duff, Kaluza-Klein theory in perspective, hep-th/9410046.

[72] M. J. Duff, Supermembranes, hep-th/9611203.

[73] P. Dunsby, N. Goheer, M. Bruni and A. Coley, Is the brane-world born isotropic?, hep-th/0312174.

[74] R. Durrer, Clarifying perturbations in the ekpyrotic universe: a web note, hep-th/0112026.

[75] R. Durrer, Braneworlds, hep-th/0507006.

[76] R. Durrer and F. Vernizzi, Adiabatic perturbations in pre big bang models: matching conditions and scale invariance, *Phys. Rev.* **D66**, 083503 (2002), [hep-ph/0203275].

[77] A. Einstein and P. Bergmann, On a generalization of Kaluza's theory of electricity, *Annals Math.* **39**, 683 (1938).

[78] R. v. Eötvös, D. Pekar and E. Fekete, Beitrage zum Gesetze der Proportionalität von Trägheit und Gravität (Contributions to the law of proportionality of inertia and gravity), *Annalen Phys.* **68**, 11 (1922).

[79] J. K. Erickson, R. R. Caldwell, P. J. Steinhardt, C. Armendáriz-Picón and V. Mukhanov, Measuring the speed of sound of quintessence, *Phys. Rev. Lett.* **88**, 121301 (2002), [astro-ph/0112438].

[80] J. K. Erickson, D. H. Wesley, P. J. Steinhardt and N. Turok, Kasner and mixmaster behavior in universes with equation of state $w \geq 1$, *Phys. Rev.* **D69**, 063514 (2004), [hep-th/0312009].

[81] M. Fierz, Uber die physikalische Deutung der erweiterten Gravitationstheorie P. Jordans, *Helv. Phys. Acta* **29**, 128 (1956).

[82] F. Finelli and R. Brandenberger, On the generation of a scale-invariant spectrum of adiabatic fluctuations in cosmological models with a contracting phase, *Phys. Rev.* **D65**, 103522 (2002), [hep-th/0112249].

[83] J. A. Friedman and B.-A. Gradwohl, Dark matter and the equivalence principle, *Phys. Rev. Lett.* **67**, 2926 (1991).

[84] J. A. Frieman, C. T. Hill, A. Stebbins and I. Waga, Cosmology with ultralight pseudo Nambu-Goldstone bosons, *Phys. Rev. Lett.* **75**, 2077 (1995), [astro-ph/9505060].

[85] M. Gasperini, F. Piazza and G. Veneziano, Quintessence as a run-away dilaton, *Phys. Rev.* **D65**, 023508 (2002), [gr-qc/0108016].

[86] M. Gasperini and G. Veneziano, Pre-big bang in string cosmology, *Astropart. Phys.* **1**, 317 (1993), [hep-th/9211021].

[87] M. Gasperini and G. Veneziano, The pre-big bang scenario in string cosmology, *Phys. Rept.* **373**, 1 (2003), [hep-th/0207130].

[88] S. Gratton, J. Khoury, P. J. Steinhardt and N. Turok, Conditions for generating scale-invariant density perturbations, *Phys. Rev.* **D69**, 103505 (2004), [astro-ph/0301395].

[89] M. B. Green, J. H. Schwarz and E. Witten, *Superstring theory, volume one: introduction* (Cambridge University Press, Cambridge, UK, 1987).

[90] M. B. Green, J. H. Schwarz and E. Witten, *Superstring theory, volume two: loop amplitudes, anomalies and phenomenology* (Cambridge University Press, Cambridge, UK, 1987).

[91] B. R. Greene, String theory on Calabi-Yau manifolds, hep-th/9702155.

[92] B. R. Greene, K. Schalm and G. Shiu, Warped compactifications in M and F theory, *Nucl. Phys.* **B584**, 480 (2000), [hep-th/0004103].

[93] S. Hannestad, Constraints on the sound speed of dark energy, *Phys. Rev.* **D71**, 103519 (2005), [astro-ph/0504017].

[94] T. Hertog, G. T. Horowitz and K. Maeda, Negative energy density in Calabi-Yau compactifications, *JHEP* **05**, 060 (2003), [hep-th/0304199].

[95] P. Hořava and E. Witten, Eleven-dimensional supergravity on a manifold with boundary, *Nucl. Phys.* **B475**, 94 (1996), [hep-th/9603142].

[96] P. Hořava and E. Witten, Heterotic and type I string dynamics from eleven dimensions, *Nucl. Phys.* **B460**, 506 (1996), [hep-th/9510209].

[97] C. D. Hoyle *et al.* (Eöt-Wash), Sub-millimeter tests of the gravitational inverse-square law: A search for 'large' extra dimensions, *Phys. Rev. Lett.* **86**, 1418 (2001), [hep-ph/0011014].

[98] W. Hu, Structure formation with generalized dark matter, *Astrophys. J.* **506**, 485 (1998), [astro-ph/9801234].

[99] W. Hu, D. J. Eisenstein, M. Tegmark and M. J. White, Observationally determining the properties of dark matter, *Phys. Rev.* **D59**, 023512 (1999), [astro-ph/9806362].

[100] W. T. Hu, *Wandering in the background: a cosmic microwave background explorer*, PhD thesis, University of California, Berkeley (1995), [astro-ph/9508126].

[101] G. Huey, L.-M. Wang, R. Dave, R. R. Caldwell and P. J. Steinhardt, Resolving the cosmological missing energy problem, *Phys. Rev.* **D59**, 063005 (1999), [astro-ph/9804285].

[102] J.-c. Hwang, Cosmological structure problem in the ekpyrotic scenario, *Phys. Rev.* **D65**, 063514 (2002), [astro-ph/0109045].

[103] M. Ishak, C. M. Hirata, P. McDonald and U. Seljak, Weak Lensing and CMB: Parameter forecasts including a running spectral index, *Phys. Rev.* **D69**, 083514 (2004), [astro-ph/0308446].

[104] P. Jordan, Erweiterung der projektiven Relativitatstheorie (Extension of projective relativity), *Ann. Phys. (Leipzig)* **1**, 219 (1947).

[105] S. Kachru, R. Kallosh, A. Linde and S. P. Trivedi, De Sitter vacua in string theory, *Phys. Rev.* **D68**, 046005 (2003), [hep-th/0301240].

[106] R. Kallosh, L. Kofman and A. D. Linde, Pyrotechnic universe, *Phys. Rev.* **D64**, 123523 (2001), [hep-th/0104073].

[107] R. Kallosh, L. Kofman, A. D. Linde and A. A. Tseytlin, BPS branes in cosmology, *Phys. Rev.* **D64**, 123524 (2001), [hep-th/0106241].

[108] R. Kallosh, J. Kratochvil, A. Linde, E. V. Linder and M. Shmakova, Observational bounds on cosmic doomsday, *JCAP* **0310**, 015 (2003), [astro-ph/0307185].

[109] T. Kaluza, On the problem of unity in physics, *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* **1921**, 966 (1921).

[110] J. Khoury, B. A. Ovrut, N. Seiberg, P. J. Steinhardt and N. Turok, From big crunch to big bang, *Phys. Rev.* **D65**, 086007 (2002), [hep-th/0108187].

[111] J. Khoury, B. A. Ovrut, P. J. Steinhardt and N. Turok, A brief comment on "the pyrotechnic universe", hep-th/0105212.

[112] J. Khoury, B. A. Ovrut, P. J. Steinhardt and N. Turok, The ekpyrotic universe: colliding branes and the origin of the hot big bang, *Phys. Rev.* **D64**, 123522 (2001), [hep-th/0103239].

[113] J. Khoury, B. A. Ovrut, P. J. Steinhardt and N. Turok, Density perturbations in the ekpyrotic scenario, *Phys. Rev.* **D66**, 046005 (2002), [hep-th/0109050].

[114] J. Khoury, P. J. Steinhardt and N. Turok, Great expectations: inflation versus cyclic predictions for spectral tilt, *Phys. Rev. Lett.* **91**, 161301 (2003), [astro-ph/0302012].

[115] J. Khoury, P. J. Steinhardt and N. Turok, Designing cyclic universe models, *Phys. Rev. Lett.* **92**, 031302 (2004), [hep-th/0307132].

[116] J. Khoury and A. Weltman, Chameleon cosmology, *Phys. Rev.* **D69**, 044026 (2004), [astro-ph/0309411].

[117] J. Khoury and A. Weltman, Chameleon fields: awaiting surprises for tests of gravity in space, *Phys. Rev. Lett.* **93**, 171104 (2004), [astro-ph/0309300].

[118] O. Klein, The atomicity of electricity as a quantum theory law, *Nature* **118**, 516 (1926).

[119] O. Klein, Quantum theory and five-dimensional theory of relativity, *Z. Phys.* **37**, 895 (1926).

[120] R. A. Knop *et al.* (The Supernova Cosmology Project), New constraints on Ω_M, Ω_Λ, and w from an independent set of eleven high-redshift supernovae observed with HST, *Astrophys. J.* **598**, 102 (2003), [astro-ph/0309368].

[121] E. W. Kolb and M. S. Turner, *The early universe* (Addison-Wesley, Redwood City, USA, 1990).

[122] D. Langlois, Brane cosmology: An introduction, *Prog. Theor. Phys. Suppl.* **148**, 181 (2003), [hep-th/0209261].

[123] S. A. Levshakov, M. Centurion, P. Molaro and S. D'Odorico, VLT/UVES constraints on the cosmological variability of the fine-structure constant, astro-ph/0408188.

[124] A. R. Liddle and D. H. Lyth, Cosmological inflation and large-scale structure.

[125] E. M. Lifshitz and I. M. Khalatnikov, Investigations in relativistic cosmology, *Adv. Phys.* **12**, 185 (1963).

[126] A. Lukas, B. A. Ovrut and D. Waldram, On the four-dimensional effective action of strongly coupled heterotic string theory, *Nucl. Phys.* **B532**, 43 (1998), [hep-th/9710208].

[127] A. Lukas, B. A. Ovrut and D. Waldram, Non-standard embedding and five-branes in heterotic M theory, *Phys. Rev.* **D59**, 106005 (1999), [hep-th/9808101].

[128] D. H. Lyth, The failure of cosmological perturbation theory in the new ekpyrotic scenario, *Phys. Lett.* **B526**, 173 (2002), [hep-ph/0110007].

[129] D. H. Lyth, The primordial curvature perturbation in the ekpyrotic universe, *Phys. Lett.* **B524**, 1 (2002), [hep-ph/0106153].

[130] C.-P. Ma and E. Bertschinger, Cosmological perturbation theory in the synchronous and conformal Newtonian gauges, *Astrophys. J.* **455**, 7 (1995), [astro-ph/9506072].

[131] I. Maor, R. Brustein, J. McMahon and P. J. Steinhardt, Measuring the equation-of-state of the universe: pitfalls and prospects, *Phys. Rev.* **D65**, 123003 (2002), [astro-ph/0112526].

[132] I. Maor, R. Brustein and P. J. Steinhardt, Limitations in using luminosity distance to determine the equation-of-state of the universe, *Phys. Rev. Lett.* **86**, 6 (2001), [astro-ph/0007297].

[133] J. Martin, P. Peter, N. Pinto Neto and D. J. Schwarz, Passing through the bounce in the ekpyrotic models, *Phys. Rev.* **D65**, 123513 (2002), [hep-th/0112128].

[134] J. Martin, P. Peter, N. Pinto-Neto and D. J. Schwarz, Comment on 'density perturbations in the ekpyrotic scenario', *Phys. Rev.* **D67**, 028301 (2003), [hep-th/0204222].

[135] L. McAllister and I. Mitra, Relativistic D-brane scattering is extremely inelastic, *JHEP* **02**, 019 (2005), [hep-th/0408085].

[136] C. W. Misner, Mixmaster universe, *Phys. Rev. Lett.* **22**, 1071 (1969).

[137] C. W. Misner, Quantum cosmology I, *Phys. Rev.* **186**, 1319 (1969).

[138] C. W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation* (W. H. Freeman and Co., New York, 1973).

[139] V. F. Mukhanov, H. A. Feldman and R. H. Brandenberger, Theory of cosmological perturbations. Part 1. Classical perturbations. Part 2. Quantum theory of perturbations. Part 3. Extensions, *Phys. Rept.* **215**, 203 (1992).

[140] M. T. Murphy, J. K. Webb and V. V. Flambaum, Further evidence for a variable fine-structure constant from Keck/HIRES QSO absorption spectra, *Mon. Not. Roy. Astron. Soc.* **345**, 609 (2003), [astro-ph/0306483].

[141] A. M. Nobili, D. Bramanti, E. Polacco, G. Catastini, A. Anselmi, S. Portigliotti, A. Lenti and A. Severi, The "Galileo Galilei" (GG) project: testing the equivalence principle in space and on earth, *Advances in Space Research* **25**, 1231 (2000).

[142] G. Nordström, Über die Möglichkeit, das elektromagnetishe Feld und das Gravitationsfeld zu vereinigen (On the possibility of a unification of the electromagnetic and gravitational fields), *Physik. Zeitschr.* **15**, 504 (1914).

[143] K. L. Nordtvedt, Testing relativity with laser ranging to the moon, *Phys. Rev.* **170**, 1186 (1968).

[144] K. A. Olive, G. Steigman and T. P. Walker, Primordial nucleosynthesis: theory and observations, *Phys. Rept.* **333**, 389 (2000), [astro-ph/9905320].

[145] J. M. Overduin and P. S. Wesson, Kaluza-Klein gravity, *Phys. Rept.* **283**, 303 (1997), [gr-qc/9805018].

[146] T. Padmanabhan, *Structure formation in the universe* (Cambridge University Press, Cambridge, UK, 1993).

[147] P. J. E. Peebles, Light out of darkness vs order out of chaos, *Comments on Astrophysics and Space Physics* **4**, 53 (1972).

[148] P. J. E. Peebles, *Principles of physical cosmology* (Princeton University Press, Princeton, NJ, 1993).

[149] P. J. E. Peebles and B. Ratra, Cosmology with a time variable cosmological 'constant', *Astrophys. J.* **325**, L17 (1988).

[150] S. Perlmutter *et al.* (Supernova Cosmology Project), Measurements of Ω and Λ from 42 high-redshift supernovae, *Astrophys. J.* **517**, 565 (1999), [astro-ph/9812133].

[151] P. Peter and N. Pinto-Neto, Primordial perturbations in a non singular bouncing universe model, *Phys. Rev.* **D66**, 063509 (2002), [hep-th/0203013].

[152] J. Polchinski, *String theory, volume one: an introduction to the bosonic string* (Cambridge University Press, Cambridge, UK, 1998).

[153] J. Polchinski, *String theory, volume two: superstring theory and beyond* (Cambridge University Press, Cambridge, UK, 1998).

[154] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in C: the art of scientific computing* (Cambridge University Press, 1995), second edition.

[155] L. Randall and R. Sundrum, An alternative to compactification, *Phys. Rev. Lett.* **83**, 4690 (1999), [hep-th/9906064].

[156] L. Randall and R. Sundrum, A large mass hierarchy from a small extra dimension, *Phys. Rev. Lett.* **83**, 3370 (1999), [hep-ph/9905221].

[157] B. Ratra and P. J. E. Peebles, Cosmological consequences of a rolling homogeneous scalar field, *Phys. Rev.* **D37**, 3406 (1988).

[158] A. G. Riess *et al.* (Supernova Search Team), Observational evidence from supernovae for an accelerating universe and a cosmological constant, *Astron. J.* **116**, 1009 (1998), [astro-ph/9805201].

[159] A. G. Riess *et al.* (Supernova Search Team), Type Ia supernova discoveries at $z > 1$ from the Hubble space telescope: evidence for past deceleration and constraints on dark energy evolution, *Astrophys. J.* **607**, 665 (2004), [astro-ph/0402512].

[160] P. G. Roll, R. Krotkov and R. H. Dicke, The Equivalence of inertial and passive gravitational mass, *Ann. Phys.* **26**, 442 (1964).

[161] L. I. Schiff, On experimental tests of the general theory of relativity, *Am. J. Phys.* **28**, 340 (1960).

[162] U. Seljak, P. McDonald and A. Makarov, Cosmological constraints from the CMB and Ly-alpha forest revisited, *Mon. Not. Roy. Astron. Soc.* **342**, L79 (2003), [astro-ph/0302571].

[163] U. Seljak and M. Zaldarriaga, A line of sight approach to cosmic microwave background anisotropies, *Astrophys. J.* **469**, 437 (1996), [astro-ph/9603033].

[164] U. Seljak and M. Zaldarriaga, Signature of gravity waves in polarization of the microwave background, *Phys. Rev. Lett.* **78**, 2054 (1997), [astro-ph/9609169].

[165] S. S. Shapiro, J. L. Davis, D. E. Lebach and J. S. Gregory, Measurement of the solar gravitational deflection of radio waves using geodetic very-long-baseline interferometry data, 1979–1999, *Phys. Rev. Lett.* **92**(12), 121101 (2004).

[166] A. I. Shlyakhter, Direct test of the constancy of fundamental nuclear constants, *Nature* **264**, 340 (1976).

[167] D. N. Spergel *et al.* (WMAP), First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: determination of cosmological parameters, *Astrophys. J. Suppl.* **148**, 175 (2003), [astro-ph/0302209].

[168] R. Srianand, H. Chand, P. Petitjean and B. Aracil, Limits on the time variation of the electromagnetic fine-structure constant in the low energy limit from absorption lines in the spectra of distant quasars, *Phys. Rev. Lett.* **92**, 121302 (2004), [astro-ph/0402177].

[169] C. L. Steinhardt, Constraints on field theoretical models for variation of the fine structure constant, *Phys. Rev.* **D71**, 043509 (2005), [hep-ph/0308253].

[170] P. J. Steinhardt and N. Turok, A cyclic model of the universe, hep-th/0111030.

[171] P. J. Steinhardt and N. Turok, Cosmic evolution in a cyclic universe, *Phys. Rev.* **D65**, 126003 (2002), [hep-th/0111098].

[172] P. J. Steinhardt and N. Turok, A cyclic model of the universe, *Science* **296**, 1436 (2002).

[173] P. J. Steinhardt, L.-M. Wang and I. Zlatev, Cosmological tracking solutions, *Phys. Rev.* **D59**, 123504 (1999), [astro-ph/9812313].

[174] Y. Su *et al.* (Eöt-Wash), New tests of the universality of free fall, *Phys. Rev.* **D50**, 3614 (1994).

[175] M. Tegmark *et al.* (SDSS), Cosmological parameters from SDSS and WMAP, *Phys. Rev.* **D69**, 103501 (2004), [astro-ph/0310723].

[176] A. J. Tolley, String propagation through a big crunch / big bang transition, hep-th/0505158.

[177] A. J. Tolley and N. Turok, Quantum fields in a big crunch / big bang spacetime, *Phys. Rev.* **D66**, 106005 (2002), [hep-th/0204091].

[178] A. J. Tolley, N. Turok and P. J. Steinhardt, Cosmological perturbations in a big crunch / big bang space-time, *Phys. Rev.* **D69**, 106005 (2004), [hep-th/0306109].

[179] P. Touboul, B. Foulon, L. Lafargue and G. Metris, The MICROSCOPE mission, *Acta Astronautica* **50**, 433 (2002).

[180] N. Turok, M. Perry and P. J. Steinhardt, M theory model of a big crunch / big bang transition, *Phys. Rev.* **D70**, 106004 (2004), [hep-th/0408083].

[181] S. G. Turyshev *et al.*, Laser ranging to the moon, mars and beyond, gr-qc/0411082.

[182] A. Upadhye, M. Ishak and P. J. Steinhardt, Dynamical dark energy: current constraints and forecasts, astro-ph/0411803.

[183] J.-P. Uzan, The fundamental constants and their variation: observational status and theoretical motivations, *Rev. Mod. Phys.* **75**, 403 (2003), [hep-ph/0205340].

[184] G. Veneziano, Scale factor duality for classical and quantum strings, *Phys. Lett.* **B265**, 287 (1991).

[185] G. Veneziano, Large N bounds on, and compositeness limit of, gauge and gravitational interactions, *JHEP* **06**, 051 (2002), [hep-th/0110129].

[186] R. V. Wagoner, Scalar-tensor theory and gravitational waves, *Phys. Rev.* **D1**, 3209 (1970).

[187] Y. Wang, J. M. Kratochvil, A. Linde and M. Shmakova, Current observational constraints on cosmic doomsday, *JCAP* **0412**, 006 (2004), [astro-ph/0409264].

[188] J. K. Webb, V. V. Flambaum, C. W. Churchill, M. J. Drinkwater and J. D. Barrow, Evidence for time variation of the fine structure constant, *Phys. Rev. Lett.* **82**, 884 (1999), [astro-ph/9803165].

[189] J. K. Webb *et al.*, Further evidence for cosmological evolution of the fine structure constant, *Phys. Rev. Lett.* **87**, 091301 (2001), [astro-ph/0012539].

[190] S. Weinberg, Anthropic bound on the cosmological constant, *Phys. Rev. Lett.* **59**, 2607 (1987).

[191] D. H. Wesley, P. J. Steinhardt and N. Turok, Controlling chaos through compactification in cosmological models with a collapsing phase, hep-th/0502108.

[192] J. Wess and J. Bagger, *Supersymmetry and supergravity* (Princeton University Press, Princeton NJ, 1992).

[193] C. M. Will, *Theory and experiment in gravitational physics* (Cambridge University Press, Cambridge, UK, 1993).

[194] J. G. Williams, S. G. Turyshev and D. H. Boggs, Progress in lunar laser ranging tests of relativistic gravity, *Phys. Rev. Lett.* **93**, 261101 (2004), [gr-qc/0411113].

[195] E. Witten, Search for a realistic Kaluza-Klein theory, *Nucl. Phys.* **B186**, 412 (1981).

[196] E. Witten, Strong coupling expansion of Calabi-Yau compactification, *Nucl. Phys.* **B471**, 135 (1996), [hep-th/9602070].

[197] P. Worden, R. Torii, J. C. Mester and C. W. F. Everitt, The STEP payload and experiment, *Advances in Space Research* **25**, 1205 (2000).

[198] I. Zlatev, L.-M. Wang and P. J. Steinhardt, Quintessence, cosmic coincidence, and the cosmological constant, *Phys. Rev. Lett.* **82**, 896 (1999), [astro-ph/9807002].